

A SUM OF SQUARES APPROXIMATION OF NONNEGATIVE POLYNOMIALS

JEAN B. LASSERRE*

Abstract. We show that every real nonnegative polynomial f can be approximated as closely as desired (in the l_1 -norm of its coefficient vector) by a sequence of polynomials $\{f_\epsilon\}$ that are sums of squares. The novelty is that each f_ϵ has a simple and explicit form in terms of f and ϵ .

Key words. Real algebraic geometry; positive polynomials; sum of squares; semidefinite programming.

AMS subject classifications. 12E05, 12Y05, 90C22

1. Introduction. The study of relationships between *nonnegative* and *sums of squares* (s.o.s.) polynomials, initiated by Hilbert, is of real practical importance in view of numerous potential applications, notably in polynomial programming. Indeed, checking whether a given polynomial is nonnegative is a NP-hard problem whereas checking whether it is s.o.s. reduces to solving a (convex) Semidefinite Programming (SDP) problem for which efficient algorithms are now available. (For instance, it is known that up to an a priori fixed precision, an SDP is solvable in time polynomial in the input size of the problem.)

For instance, recent results in real algebraic geometry, most notably by Schmüdgen [16], Putinar [13], Jacobi and Prestel [5], have provided s.o.s. representations of polynomials, positive on a compact semialgebraic set; the interested reader is referred to Prestel and Delzell [12] and Scheiderer [15] for a nice account of such results. This in turn has permitted to develop efficient SDP-relaxations in polynomial optimization (see e.g. Lasserre [6, 7, 8], Parrilo [10, 11], Schweighofer [17], and the many references therein).

So, back to a comparison between nonnegative and s.o.s. polynomials, on the negative side, Blekherman [4] has shown that if a degree > 2 is *fixed* (and for a large fixed number of variables), then the cone of nonnegative polynomials is much *larger* than that of s.o.s. However, on the positive side, a denseness result [2] states that the cone of s.o.s. polynomials is *dense* in the space of polynomials that are nonnegative on $[-1, 1]^n$ (for the l_1 -norm $\|f\|_1 = \sum_\alpha |f_\alpha|$ whenever f is written $\sum_\alpha f_\alpha x^\alpha$ in the usual canonical basis); see e.g. Berg et al. [2, Theorem 9.1, p. 273]).

Contribution. We show that *every nonnegative* polynomial f is almost a s.o.s., namely we show that f can be approximated by a sequence of s.o.s. polynomials $\{f_\epsilon\}_\epsilon$, in the specific form

$$(1.1) \quad f_\epsilon = f + \epsilon \sum_{k=0}^{r(f,\epsilon)} \sum_{j=1}^n \frac{x_j^{2k}}{k!},$$

for some $r(f, \epsilon) \in \mathbb{N}$, so that $\|f - f_\epsilon\|_1 \rightarrow 0$ as $\epsilon \downarrow 0$. (Notice that in (1.1), one may replace $r(f, \epsilon)$ with any $r \geq r(f, \epsilon)$ and still get the same result.)

This result is in the spirit of the previous denseness result. However we here provide in (1.1) an *explicit* converging approximation with a very specific (and simple)

*LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse Cédex 4, France. Tel: 33 5 61336415; Email: lasserre@laas.fr.

form; namely it suffices to slightly perturbate f by adding a small coefficient $\epsilon > 0$ to each square monomial x_i^{2k} for all $i = 1, \dots, n$ and all $k = 0, 1, \dots, r$, with r sufficiently large. To prove this result we combine

- (generalized) **Carleman's** sufficient condition (due to Nussbaum [9]) for a moment sequence $\mathbf{y} = \{y_\alpha\}$ to have a unique *representing measure* μ (i.e., such that $y_\alpha = \int x^\alpha d\mu$ for all $\alpha \in \mathbb{N}^n$), and

- a **duality** result from convex optimization.

As a consequence, we may thus define a procedure to approximate the global minimum of a polynomial f , at least when there is a global minimizer x^* that satisfies $\|x^*\|_\infty \leq M$ for some known M . It consists in solving a sequence of SDP-relaxations which are simpler and easier to solve than those defined in Lasserre [6]; see §3.

Finally, we also consider the case where f is a *convex* polynomial, nonnegative on a convex semi-algebraic set \mathbb{K} defined by (concave polynomial) inequalities $g_j \geq 0$. We show that the approximation f_ϵ of f , defined in (1.1), has a certificate of positivity on \mathbb{K} (or a representation) similar to Putinar's s.o.s. representation [13], but in which the s.o.s. polynomial coefficients of the g_j 's now become simple nonnegative *scalars*, the Lagrange multipliers of a related convex optimization problem.

2. Notation and definitions. For a real symmetric matrix A , the notation $A \succeq 0$ (resp. $A \succ 0$) stands for A positive semidefinite (resp. positive definite). The sup-norm $\sup_j |x_j|$ of a vector $x \in \mathbb{R}^n$, is denoted by $\|x\|_\infty$. Let $\mathbb{R}[x_1, \dots, x_n]$ be the ring of real polynomials, and let

$$(2.1) \quad v_r(x) := (1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_1x_n, x_2^2, x_2x_3, \dots, x_n^r)$$

be the canonical basis for the \mathbb{R} -vector space \mathcal{A}_r of real polynomials of degree at most r , and let $s(r)$ be its dimension. Similarly, $v_\infty(x)$ denotes the canonical basis of $\mathbb{R}[x_1, \dots, x_n]$ as a \mathbb{R} -vector space, denoted \mathcal{A} . So a vector in \mathcal{A} has always *finitely* many non-zero entries.

Therefore, a polynomial $p \in \mathcal{A}_r$ is written

$$x \mapsto p(x) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha = \langle \mathbf{p}, v_r(x) \rangle, \quad x \in \mathbb{R}^n,$$

(where $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$) for some vector $\mathbf{p} = \{p_\alpha\} \in \mathbb{R}^{s(r)}$, the vector of coefficients of p in the basis (2.1).

Extending \mathbf{p} with zeros, we can also consider \mathbf{p} as a vector indexed in the basis $v_\infty(x)$ (i.e. $\mathbf{p} \in \mathcal{A}$). If we equip \mathcal{A} with the usual scalar product $\langle \cdot, \cdot \rangle$ of vectors, then for every $p \in \mathcal{A}$,

$$p(x) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha = \langle \mathbf{p}, v_\infty(x) \rangle, \quad x \in \mathbb{R}^n.$$

Given a sequence $\mathbf{y} = \{y_\alpha\}$ indexed in the basis $v_\infty(x)$, let $L_{\mathbf{y}} : \mathcal{A} \rightarrow \mathbb{R}$ be the linear functional

$$p \mapsto L_{\mathbf{y}}(p) := \sum_{\alpha \in \mathbb{N}^n} p_\alpha y_\alpha = \langle \mathbf{p}, \mathbf{y} \rangle.$$

Given a sequence $\mathbf{y} = \{y_\alpha\}$ indexed in the basis $v_\infty(x)$, the *moment* matrix $M_r(\mathbf{y}) \in \mathbb{R}^{s(r) \times s(r)}$ with rows and columns indexed in the basis $v_r(x)$ in (2.1), satisfies

$$[M_r(\mathbf{y})(1, j) = y_\alpha \text{ and } M_r(\mathbf{y})(i, 1) = y_\beta] \Rightarrow M_r(\mathbf{y})(i, j) = y_{\alpha+\beta}.$$

For instance, with $n = 2$,

$$M_2(\mathbf{y}) = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix}.$$

A sequence $\mathbf{y} = \{y_\alpha\}$ has a *representing* measure $\mu_{\mathbf{y}}$ if

$$(2.2) \quad y_\alpha = \int_{\mathbb{R}^n} x^\alpha \mu_{\mathbf{y}}(dx), \quad \forall \alpha \in \mathbb{N}^n.$$

In this case one also says that \mathbf{y} is a *moment sequence*. In addition, if $\mu_{\mathbf{y}}$ is unique then \mathbf{y} is said to be a *determinate* moment sequence.

The matrix $M_r(\mathbf{y})$ defines a bilinear form $\langle \cdot, \cdot \rangle_{\mathbf{y}}$ on \mathcal{A}_r , by

$$\langle q, p \rangle_{\mathbf{y}} := \langle \mathbf{q}, M_r(\mathbf{y})\mathbf{p} \rangle = L_{\mathbf{y}}(qp), \quad q, p \in \mathcal{A}_r,$$

and if \mathbf{y} has a *representing* measure $\mu_{\mathbf{y}}$ then

$$(2.3) \quad \langle \mathbf{q}, M_r(\mathbf{y})\mathbf{q} \rangle = \int_{\mathbb{R}^n} q(x)^2 \mu_{\mathbf{y}}(dx) \geq 0,$$

so that $M_r(\mathbf{y}) \succeq 0$.

Next, given a sequence $\mathbf{y} = \{y_\alpha\}$ indexed in the basis $v_\infty(x)$, let $y_{2k}^{(i)} := L_{\mathbf{y}}(x_i^{2k})$ for every $i = 1, \dots, n$ and every $k \in \mathbb{N}$. That is, $y_{2k}^{(i)}$ denotes the element in the sequence \mathbf{y} , corresponding to the monomial x_i^{2k} .

Of course not every sequence $\mathbf{y} = \{y_\alpha\}$ has a representing measure $\mu_{\mathbf{y}}$ as in (2.2). However, there exists a *sufficient* condition to ensure that it is the case. The following result stated in Berg [3, Theorem 5, p. 117] is from Nussbaum [9], and is re-stated here, with our notation.

THEOREM 2.1. *Let $\mathbf{y} = \{y_\alpha\}$ be an infinite sequence such that $M_r(\mathbf{y}) \succeq 0$ for all $r = 0, 1, \dots$. If*

$$(2.4) \quad \sum_{k=1}^{\infty} (y_{2k}^{(i)})^{-1/2k} = \infty, \quad i = 1, \dots, n,$$

then \mathbf{y} is a determinate moment sequence.

The condition (2.4) in Theorem 2.1 is called *Carleman's condition* as it extends to the multivariate case the original Carleman's sufficient condition given for the univariate case.

3. Preliminaries. Let B_M be the closed ball

$$(3.1) \quad B_M = \{x \in \mathbb{R}^n \mid \|x\|_\infty \leq M\}.$$

PROPOSITION 3.1. *Let $f \in \mathbb{R}[x_1, \dots, x_n]$ be such that $-\infty < f^* := \inf_x f(x)$. Then, for every $\epsilon > 0$ there is some $M_\epsilon \in \mathbb{N}$ such that*

$$f_M^* := \inf_{x \in B_M} f(x) < f^* + \epsilon, \quad \forall M \geq M_\epsilon.$$

Equivalently, $f_M^* \downarrow f^*$ as $M \rightarrow \infty$.

Proof. Suppose it is false. That is, there is some $\epsilon_0 > 0$ and an infinite sequence $\{M_k\} \subset \mathbb{N}$, with $M_k \rightarrow \infty$, such that $f_{M_k}^* \geq f^* + \epsilon_0$ for all k . But let $x_0 \in \mathbb{R}^n$ be such that $f(x_0) < f^* + \epsilon_0$. With any $M_k \geq \|x_0\|_\infty$, one obtains the contradiction $f^* + \epsilon_0 \leq f_{M_k}^* \leq f(x_0) < f^* + \epsilon_0$. \square

To prove our main result (Theorem 4.1 below), we first introduce the following related optimization problems.

$$(3.2) \quad \mathbb{P} : \quad f^* := \inf_{x \in \mathbb{R}^n} f(x),$$

and for $0 < M \in \mathbb{N}$,

$$(3.3) \quad \mathcal{P}_M : \quad \inf_{\mu \in \mathcal{P}(\mathbb{R}^n)} \left\{ \int f d\mu \mid \int \sum_{i=1}^n e^{x_i^2} d\mu \leq ne^{M^2} \right\},$$

where $\mathcal{P}(\mathbb{R}^n)$ is the space of probability measures on \mathbb{R}^n . The respective optimal values of \mathbb{P} and \mathcal{P}_M are denoted $\inf \mathbb{P} = f^*$ and $\inf \mathcal{P}_M$, or $\min \mathbb{P}$ and $\min \mathcal{P}_M$ if the infimum is attained.

PROPOSITION 3.2. *Let $f \in \mathbb{R}[x_1, \dots, x_n]$ be such that $-\infty < f^* := \inf_x f(x)$, and consider the two optimization problems \mathbb{P} and \mathcal{P}_M defined in (3.2) and (3.3) respectively. Then, $\inf \mathcal{P}_M \downarrow f^*$ as $M \rightarrow \infty$. If f has a global minimizer $x^* \in \mathbb{R}^n$, then $\min \mathcal{P}_M = f^*$ whenever $M \geq \|x^*\|_\infty$.*

Proof. Let $\mu \in \mathcal{P}(\mathbb{R}^n)$ be admissible for \mathcal{P}_M . As $f \geq f^*$ on \mathbb{R}^n then it follows immediately that $\int f d\mu \geq f^*$, and so, $\inf \mathcal{P}_M \geq f^*$ for all M .

As B_M is closed and bounded, it is compact and so, with f_M^* as in Proposition 3.1, there is some $\hat{x} \in B_M$ such that $f(\hat{x}) = f_M^*$. In addition let $\mu \in \mathcal{P}(\mathbb{R}^n)$ be the Dirac probability measure at the point \hat{x} . As $\|\hat{x}\|_\infty \leq M$,

$$\int \sum_{i=1}^n e^{x_i^2} d\mu = \sum_{i=1}^n e^{(\hat{x}_i)^2} \leq ne^{M^2},$$

so that μ is an admissible solution of \mathcal{P}_M with value $\int f d\mu = f(\hat{x}) = f_M^*$, which proves that $\inf \mathcal{P}_M \leq f_M^*$. This latter fact, combined with Proposition 3.1 and with $f^* \leq \inf \mathcal{P}_M$, implies $\inf \mathcal{P}_M \downarrow f^*$ as $M \rightarrow \infty$, the desired result. The final statement is immediate by taking as feasible solution for \mathcal{P}_M , the Dirac probability measure at the point $x^* \in B_M$ (with $M \geq \|x^*\|_\infty$). As its value is now f^* , it is also optimal, and so, \mathcal{P}_M is solvable with optimal value $\min \mathcal{P}_M = f^*$. \square

Proposition 3.2 provides a rationale for introducing the following Semidefinite Programming (SDP) problems. Let $2r_f$ be the degree of f and for every $r_f \leq r \in \mathbb{N}$, consider the SDP problem

$$(3.4) \quad \mathbb{Q}_r \begin{cases} \min_{\mathbf{y}} L_{\mathbf{y}}(f) (= \sum_{\alpha} f_{\alpha} y_{\alpha}) \\ \text{s.t. } M_r(\mathbf{y}) \succeq 0 \\ \sum_{k=0}^r \sum_{i=1}^n \frac{y_{2k}^{(i)}}{k!} \leq ne^{M^2}, \\ y_0 = 1, \end{cases}$$

and its associated *dual* SDP problem

$$(3.5) \quad \mathbb{Q}_r^* \begin{cases} \max_{\lambda \geq 0, \gamma, q} & \gamma - ne^{M^2} \lambda \\ \text{s.t.} & f - \gamma \\ & q \end{cases} = q - \lambda \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!}$$

s.o.s. of degree $\leq 2r$,

with respective optimal values $\inf \mathbb{Q}_r$ and $\sup \mathbb{Q}_r^*$ (or $\min \mathbb{Q}_r$ and $\max \mathbb{Q}_r^*$ if the optimum is attained, in which case the problems are said to be solvable). For more details on SDP theory, the interested reader is referred to the survey paper [18].

The SDP problem \mathbb{Q}_r is a relaxation of \mathcal{P}_M , and we next show that in fact

- \mathbb{Q}_r is solvable for all $r \geq r_0$,
- its optimal value $\min \mathbb{Q}_r \rightarrow \inf \mathcal{P}_M$ as $r \rightarrow \infty$, and
- \mathbb{Q}_r^* is also solvable with same optimal value as \mathbb{Q}_r , for every $r \geq r_f$.

This latter fact will be crucial to prove our main result in the next section. Let l_∞ (resp. l_1) be the Banach space of bounded (resp. summable) infinite sequences with the sup-norm (resp. the l_1 -norm).

THEOREM 3.3. *Let $f \in \mathbb{R}[x_1, \dots, x_n]$ be of degree $2r_f$, with global minimum $f^* > -\infty$, and let $M > 0$ be fixed. Then :*

- (i) *For every $r \geq r_f$, \mathbb{Q}_r is solvable, and $\min \mathbb{Q}_r \uparrow \inf \mathcal{P}_M$ as $r \rightarrow \infty$.*
- (ii) *Let $\mathbf{y}^{(r)} = \{y_\alpha^{(r)}\}$ be an optimal solution of \mathbb{Q}_r and complete $\mathbf{y}^{(r)}$ with zeros to make it an element of l_∞ . Every (pointwise) accumulation point \mathbf{y}^* of the sequence $\{\mathbf{y}^{(r)}\}_{r \in \mathbb{N}}$ is a determinate moment sequence, that is,*

$$(3.6) \quad y_\alpha^* = \int_{\mathbb{R}^n} x^\alpha d\mu^*, \quad \alpha \in \mathbb{N}^n,$$

for a unique probability measure μ^* , and μ^* is an optimal solution of \mathcal{P}_M .

- (iii) *For every $r \geq r_f$, $\max \mathbb{Q}_r^* = \min \mathbb{Q}_r$.*

For a proof see §5.1.

So, one can approximate the optimal value f^* of \mathcal{P} as closely as desired, by solving SDP-relaxations $\{\mathbb{Q}_r\}$ for sufficiently large values of r and M . Indeed, $f^* \leq \inf \mathcal{P}_M \leq f_M^*$, with f_M^* as in Proposition 3.1. Therefore, let $\epsilon > 0$ be fixed, arbitrary. By Proposition 3.2, we have $f^* \leq \inf \mathcal{P}_M \leq f^* + \epsilon$ provided that M is sufficiently large. Next, by Theorem 3.3(i), one has $\inf \mathbb{Q}_r \geq \inf \mathcal{P}_M - \epsilon$ provided that r is sufficiently large, in which case, we finally have $f^* - \epsilon \leq \inf \mathbb{Q}_r \leq f^* + \epsilon$.

For instance, if the infimum f^* is attained and one knows an upper bound M on $\|x^*\|_\infty$ for some global minimizer x^* , then the sequence of SDP-relaxations \mathbb{Q}_r in (3.4) with M being fixed, will suffice. Notice that this SDP-relaxations \mathbb{Q}_r is simpler than the one defined in Lasserre [6]. Both have the same variables $\mathbf{y} \in \mathbb{R}^{s(2r)}$, but the former has *one* SDP constraint $M_r(\mathbf{y}) \succeq 0$ and one scalar inequality (as one substitutes y_0 with 1) whereas the latter has the same SDP constraint $M_r(\mathbf{y}) \succeq 0$ and one additional SDP constraint $M_{r-1}(\theta\mathbf{y}) \succeq 0$ for the localizing matrix associated with the polynomial $x \mapsto \theta(x) = M^2 - \|x\|^2$. This results in a significant simplification.

4. Sum of squares approximation. Let \mathcal{A} be equipped with the norm

$$f \mapsto \|f\|_1 := \sum_{\alpha \in \mathbb{N}^n} |f_\alpha|, \quad f \in \mathcal{A}.$$

THEOREM 4.1. *Let $f \in \mathbb{R}[x_1, \dots, x_n]$ be nonnegative with global minimum f^* , that is,*

$$0 \leq f^* \leq f(x), \quad x \in \mathbb{R}^n.$$

(i) *There is some $r_0 \in \mathbb{N}, \lambda_0 \geq 0$ such that, for all $r \geq r_0$ and $\lambda \geq \lambda_0$,*

$$(4.1) \quad f + \lambda \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} \quad \text{is a sum of squares.}$$

(ii) *For every $\epsilon > 0$, there is some $r(f, \epsilon) \in \mathbb{N}$ such that,*

$$(4.2) \quad f_\epsilon := f + \epsilon \sum_{k=0}^{r(f, \epsilon)} \sum_{j=1}^n \frac{x_j^{2k}}{k!} \quad \text{is a sum of squares.}$$

Hence, $\|f - f_\epsilon\|_1 \rightarrow 0$ as $\epsilon \downarrow 0$.

For a detailed proof, the reader is referred to §5.2.

Remark 4.2. Notice that whenever $r \geq r(f, \epsilon)$ (with $r(f, \epsilon)$ as in Theorem 4.1(ii)), the polynomial

$$f_{\epsilon r} := f + \epsilon \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!}$$

is also a sum of squares, as we add up squares to f_ϵ in (4.2).

In some specific examples, one may even obtain adhoc perturbations f_ϵ of f , simpler than the one in (4.2), and with same properties. This is illustrated in the following nice example, kindly provided by Bruce Reznick.

Example 1: Consider the Motzkin polynomial $(x, y) \mapsto f(x, y) = 1 + x^2y^2(x^2 + y^2 - 3)$, which is nonnegative, but *not* a sum of squares. Then, for all $n \geq 3$, the polynomial

$$f_n := f + 2^{4-2n}x^{2n}$$

is a sum of squares, and $\|f - f_n\|_1 \rightarrow 0$ as $n \rightarrow \infty$. To prove this, write

$$f(x, y) = (xy^2 + x^3/2 - 3x/2)^2 + p(x),$$

with $p(x) = 1 - (x^3/2 - 3x/2)^2 = (1 - x^2)^2(1 - x^2/4)$. Next, the univariate polynomial $x \mapsto q(x) := p(x) + 2^{4-2n}x^{2n}$ is nonnegative on \mathbb{R} , hence a sum of squares. Indeed, if $x^2 \leq 4$ then $p \geq 0$ and so $q \geq 0$. If $x^2 > 4$ then $|p(x)| \leq (x^2)^2x^2/4 = x^6/4$. From

$$q(x) \geq 2^{4-2n}x^{2n} - |p(x)| \geq \frac{x^6}{4}((x^2/4)^{n-3} - 1),$$

and the fact that $n \geq 3, x^2 > 4$, we deduce that $q(x) \geq 0$.

Theorem 4.1(ii) is a *denseness* result in the spirit of Theorem 9.1 in Berg *et al.* [2, p. 273] which states that the cone of s.o.s. polynomials is dense (also for the norm

$\|f\|_1$) in the cone of polynomials that are nonnegative on $[-1, 1]^n$. (However, notice that Theorem 4.1(ii) provides an *explicit* converging sequence $\{f_\epsilon\}$ with a simple and very specific form.) One may thus wonder whether the specific s.o.s. approximation f_ϵ in (4.2) is also valid for polynomials f that are only nonnegative on $[-1, 1]^n$, and not necessarily on the whole \mathbb{R}^n . The answer is *no*. To see this, let $x_0 \in \mathbb{R}^n$ be such that $f(x_0) < 0$, and let $M > \|x_0\|_\infty$. Observe that with f_ϵ as in (4.2), one has $f_\epsilon(x) < f(x) + \epsilon \sum_{i=1}^n e^{x_i^2}$, for all $x \in \mathbb{R}^n$, no matter the value of the parameter $r(f, \epsilon)$. Therefore, $f_\epsilon(x_0) < f(x_0) + \epsilon n e^{M^2}$. Hence, for $\epsilon < |f(x_0)|e^{-M^2}/n$, we have $f_\epsilon(x_0) < 0$. On the other hand, other adhoc perturbations of the same flavour, may work. Consider the following example, again kindly provided by B. Reznick.

Example 2: Let f be the univariate polynomial $x \mapsto f(x) := 1 - x^2$, nonnegative on $[-1, 1]$. The following adhoc perturbation $x \mapsto f_n(x) := f(x) + c_n x^{2n}$ is s.o.s. whenever $c_n \geq (n-1)^{n-1}/n^n$, and so, one may choose c_n so as to also obtain $\|f_n - f\|_1 \rightarrow 0$, as $n \rightarrow \infty$. In this case, one has to be very careful in the choice of the coefficient c_n . It cannot be too small because the degree of f_n is fixed apriori ($2n$), whereas for a nonnegative polynomial f , the parameter ϵ can be fixed arbitrarily, and independently of f . On the other hand, $r(f, \epsilon)$ is not known.

We next consider the case of a convex polynomial, nonnegative on a convex semi-algebraic set. Given $\{g_j\}_{j=1}^m \subset \mathbb{R}[x_1, \dots, x_n]$, let $\mathbb{K} \subset \mathbb{R}^n$ be the semi-algebraic set

$$(4.3) \quad \mathbb{K} := \{x \in \mathbb{R}^n \mid g_j(x) \geq 0, \quad j = 1, \dots, m\}.$$

COROLLARY 4.3. *Let \mathbb{K} be as in (4.3), where all the g_j 's are concave, and assume that Slater's condition holds, i.e., there exists $x_0 \in \mathbb{K}$ such that $g_j(x_0) > 0$ for all $j = 1, \dots, m$.*

Let $f \in \mathbb{R}[x_1, \dots, x_n]$ be convex, nonnegative on \mathbb{K} , and with a minimizer on \mathbb{K} , that is, $f(x^) \leq f(x)$ for all $x \in \mathbb{K}$, for some $x^* \in \mathbb{K}$. Then there exists a nonnegative vector $\lambda \in \mathbb{R}^m$ such that for every $\epsilon > 0$, there is some $r_\epsilon = r(f, \lambda, g_1, \dots, g_m, \epsilon) \in \mathbb{N}$ for which*

$$(4.4) \quad f + \epsilon \sum_{k=0}^{r_\epsilon} \sum_{i=1}^n \frac{x_i^{2k}}{k!} = f_0 + \sum_{j=1}^m \lambda_j g_j,$$

with $f_0 \in \mathbb{R}[x_1, \dots, x_n]$ being a sum of squares. (Therefore, the degree of f_0 is less than $\max[2r_\epsilon, \deg f, \deg g_1, \dots, \deg g_m]$.)

Proof. Consider the convex optimization problem $f^* := \min\{f(x) \mid x \in \mathbb{K}\}$. As f is convex, \mathbb{K} is a convex set and Slater's condition holds, the Karush-Kuhn-Tucker optimality condition holds. That is, there exists a *nonnegative* vector $\lambda \in \mathbb{R}^m$ of Lagrange-KKT multipliers, such that

$$\nabla f(x^*) = \sum_{j=1}^m \lambda_j \nabla g_j(x^*); \quad \lambda_j g_j(x^*) = 0, \quad j = 1, \dots, m.$$

(See e.g. Rockafellar [14].) In other words, x^* is also a (global) minimizer of the convex Lagrangian $L := f - \sum_{j=1}^m \lambda_j g_j$. Then $f^* = f(x^*) = L(x^*)$ is the (global) minimum of f over \mathbb{K} , as well as the global minimum of L over \mathbb{R}^n , i.e.,

$$(4.5) \quad f - \sum_{j=1}^m \lambda_j g_j - f^* \geq 0, \quad x \in \mathbb{R}^n.$$

As $f \geq 0$ on \mathbb{K} , $f^* \geq 0$, and so $L \geq 0$ on \mathbb{R}^n . Then (4.4) follows from Theorem 4.1(ii), applied to the polynomial L . \square

When \mathbb{K} is *compact* (and so, f has necessarily a minimizer $x^* \in \mathbb{K}$), one may compare Corollary 4.3 with Putinar's representation [13] of polynomials, positive on \mathbb{K} . When f is nonnegative on \mathbb{K} (compact), and with

$$(4.6) \quad f_\epsilon := f + \epsilon \sum_{k=0}^{r_\epsilon} \sum_{i=1}^n \frac{x_i^{2k}}{k!},$$

one may rewrite (4.4) as

$$(4.7) \quad f_\epsilon = f_0 + \sum_{j=1}^m \lambda_j g_j,$$

which is indeed a *certificate* of positivity of f_ϵ on \mathbb{K} . In fact, as $f_\epsilon > 0$ on \mathbb{K} , (4.7) can be seen as a special form of Putinar's s.o.s. representation, namely

$$(4.8) \quad f_\epsilon = q_0 + \sum_{j=1}^m q_j g_j, \quad \text{with } q_0, \dots, q_m \text{ s.o.s.}$$

(which holds under an additional assumption on the g_j 's). So, in the convex compact case, and under Slater's condition, Corollary 4.3 states that if $f \geq 0$ on \mathbb{K} , then its approximation f_ϵ in (4.6), has the simplified Putinar representation (4.7), in which the s.o.s. coefficients $\{q_j\}$ of the g_j 's in (4.8), become now simple nonnegative *scalars* in (4.7), namely, the Lagrange-KKT multipliers $\{\lambda_j\}$.

5. Proofs.

5.1. Proof of Theorem 3.3. We will prove (i) and (ii) together. We first prove that \mathbb{Q}_r is solvable. This is because the feasible set (which is closed) is compact. Indeed, the constraint

$$\sum_{k=0}^r \sum_{i=1}^n y_{2k}^{(i)} / k! \leq ne^{M^2}$$

implies that every diagonal element $y_{2k}^{(i)}$ of $M_r(\mathbf{y})$ is bounded by $\tau_r := nr!e^{M^2}$. By Lemma 6.2, this in turn implies that its diagonal elements (i.e., $y_{2\alpha}$, with $|\alpha| \leq r$) are *all* bounded by τ_r .

This latter fact and again $M_r(\mathbf{y}) \succeq 0$, also imply that in fact *every* element of $M_r(\mathbf{y})$ is bounded by τ_r , that is, $|y_\alpha| \leq \tau_r$ for all $|\alpha| \leq 2r$. Indeed, for a symmetric matrix $A \succeq 0$, every non diagonal element A_{ij} satisfies $A_{ij}^2 \leq A_{ii}A_{jj}$ so that $|A_{ij}| \leq \max_i A_{ii}$.

Therefore the set of feasible solutions of \mathbb{Q}_r is a closed bounded subset of $\mathbb{R}^{s(2r)}$, hence compact. As $L_{\mathbf{y}}(f)$ is linear in \mathbf{y} , the infimum is attained at some feasible point. Thus, for all $r \geq r_f$, \mathbb{Q}_r is solvable with optimal value $\min \mathbb{Q}_r \leq \inf \mathcal{P}_M$. The latter inequality is because the moment sequence \mathbf{y} associated with an arbitrary feasible solution μ of \mathcal{P}_M , is obviously feasible for \mathbb{Q}_r , and with value $L_{\mathbf{y}}(f) = \int f d\mu$.

Next, as the sequence $\{\min \mathbb{Q}_r\}_r$ is obviously monotone non decreasing, one has $\min \mathbb{Q}_r \uparrow \rho^* \leq \inf \mathcal{P}_M$, as $r \rightarrow \infty$. We have seen that every entry of $M_r(\mathbf{y})$ is bounded

by τ_r , and this bound holds for all $r \geq r_f$. Moreover, $M_r(\mathbf{y})$ is also a (north-west corner) submatrix of $M_s(\mathbf{y})$ for every $s > r$. Indeed, whenever $s > r$, one may write

$$M_s(\mathbf{y}) = \left[\begin{array}{c|c} M_r(\mathbf{y}) & B \\ \hline & \\ B' & C \end{array} \right]$$

for some appropriate matrices B and C . Therefore, for the same reasons, any feasible solution \mathbf{y} of \mathbb{Q}_s satisfies $|y_\alpha| \leq \tau_r$, for all $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq 2r$. Therefore, for every $s \in \mathbb{N}$, and every feasible solution \mathbf{y} of \mathbb{Q}_s , we have

$$|y_\alpha| \leq \tau_r, \quad \forall \alpha \in \mathbb{N}^n, \quad 2r - 1 \leq |\alpha| \leq 2r, \quad r = 1, \dots, s.$$

Thus, given $\mathbf{y} = \{y_\alpha\}$, denote by $\hat{\mathbf{y}} = \{\hat{y}_\alpha\}$ the new sequence obtained from \mathbf{y} by the scaling

$$\hat{y}_\alpha := y_\alpha / \tau_r \quad \forall \alpha \in \mathbb{N}^n, \quad 2r - 1 \leq |\alpha| \leq 2r, \quad r = 1, 2, \dots$$

So let $\mathbf{y}^{(r)} = \{y_\alpha^{(r)}\}$ be an optimal solution of \mathbb{Q}_r and complete $\mathbf{y}^{(r)}$ with zeros to make it an element of l_∞ . Hence, all the elements $\hat{y}_\alpha^{(r)}$ are in the unit ball B_1 of l_∞ , defined by

$$B_1 = \{\mathbf{y} = \{y_\alpha\} \in l_\infty \mid \|\mathbf{y}\|_\infty \leq 1\}.$$

By the Banach-Alaoglu Theorem, this ball is sequentially compact in the $\sigma(l_\infty, l_1)$ (weak*) topology of l_∞ (see e.g. Ash [1]). In other words, there exists an element $\hat{\mathbf{y}}^* \in B_1$ and a subsequence $\{r_k\} \subset \mathbb{N}$, such that $\hat{\mathbf{y}}^{(r_k)} \rightarrow \hat{\mathbf{y}}^*$ for the weak* topology of l_∞ , that is, for all $u \in l_1$,

$$(5.1) \quad \langle \hat{\mathbf{y}}^{(r_k)}, u \rangle \rightarrow \langle \hat{\mathbf{y}}^*, u \rangle, \quad \text{as } k \rightarrow \infty.$$

In particular, *pointwise* convergence holds, that is, for all $\alpha \in \mathbb{N}^n$,

$$\hat{y}_\alpha^{(r_k)} \rightarrow \hat{y}_\alpha^*, \quad \text{as } k \rightarrow \infty,$$

and so, defining \mathbf{y}^* from $\hat{\mathbf{y}}^*$ by

$$\mathbf{y}_\alpha^* = \tau_r \hat{y}_\alpha^*, \quad \forall \alpha \in \mathbb{N}^n, \quad 2r - 1 \leq |\alpha| \leq 2r, \quad r = 1, 2, \dots$$

one also obtains the pointwise convergence

$$(5.2) \quad \text{for all } \alpha \in \mathbb{N}^n, \quad y_\alpha^{(r_k)} \rightarrow y_\alpha^*, \quad \text{as } k \rightarrow \infty.$$

We next prove that \mathbf{y}^* is the moment sequence of an optimal solution μ^* of problem \mathcal{P}_M . From the pointwise convergence (5.2), we immediately get $M_r(\mathbf{y}^*) \succeq 0$ for all $r \geq r_f$, because $M_r(\mathbf{y})$ belongs to the cone of positive semidefinite matrices of size $s(r)$, which is closed. Next, and again by pointwise convergence, for every $s \in \mathbb{N}$,

$$\sum_{j=0}^s \sum_{i=1}^n (y^*)_{2j}^{(i)} / j! = \lim_{k \rightarrow \infty} \sum_{j=0}^s \sum_{i=1}^n (y^{(r_k)})_{2j}^{(i)} / j! \leq ne^{M^2},$$

and so, by the Monotone Convergence Theorem

$$(5.3) \quad \sum_{j=0}^{\infty} \sum_{i=1}^n (y^*)_{2j}^{(i)} / j! = \lim_{s \rightarrow \infty} \sum_{j=0}^s \sum_{i=1}^n (y^*)_{2j}^{(i)} / j! \leq ne^{M^2}.$$

But (5.3) implies that \mathbf{y}^* satisfies Carleman's condition (2.4). Indeed, from (5.3), for all $i = 1, \dots, n$, we have $(y^*)_{2k}^{(i)} < \rho k!$ for all $k \in \mathbb{N}$, and so, as $k! \leq k^k = \sqrt{k}^{2k}$,

$$[(y^*)_{2k}^{(i)}]^{-1/2k} > (\rho)^{-1/2k} / \sqrt{k},$$

which in turn implies

$$\sum_{k=0}^{\infty} [(y^*)_{2k}^{(i)}]^{-1/2k} > \sum_{k=0}^{\infty} \frac{\rho^{-1/2k}}{\sqrt{k}} = +\infty.$$

Hence, by Theorem 2.1, \mathbf{y}^* is a determinate moment sequence, that is, there exists a unique measure μ^* on \mathbb{R}^n , such that

$$y_{\alpha}^* = \int_{\mathbb{R}^n} x^{\alpha} d\mu^*, \quad \alpha \in \mathbb{N}^n.$$

By (5.3),

$$\int \sum_{i=1}^n e^{x_i^2} d\mu^* = \sum_{j=0}^{\infty} \sum_{i=1}^n (y^*)_{2j}^{(i)} / j! \leq ne^{M^2},$$

which proves that μ^* is admissible for \mathcal{P}_M .

But then, again by the pointwise convergence (5.2) of $\mathbf{y}^{(r_k)}$ to \mathbf{y}^* , we get $L_{\mathbf{y}^{(r_k)}}(f) \rightarrow L_{\mathbf{y}^*}(f) = \int f d\mu^*$ as $k \rightarrow \infty$, which, in view of $L_{\mathbf{y}^{(r_k)}}(f) \leq \inf \mathcal{P}_M$ for all k , implies

$$\int f d\mu^* = L_{\mathbf{y}^*}(f) \leq \inf \mathcal{P}_M.$$

But this proves that μ^* is an optimal solution of \mathcal{P}_M because μ^* is admissible for \mathcal{P}_M with value $\int f d\mu^* \leq \inf \mathcal{P}_M$. As the converging subsequence $\{r_k\}$ was arbitrary, it is true for every limit point. Hence, we have proved (i) and (ii).

(iii) Let \mathbf{y} be the moment sequence associated with the probability measure μ on the ball $B_{M/2} \subset \mathbb{R}^n$

$$B_{M/2} = \{x \in \mathbb{R}^n \mid \|x\|_{\infty} \leq M/2\},$$

with uniform distribution. That is,

$$\mu(B) = M^{-n} \int_{B \cap B_{M/2}} dx, \quad B \in \mathcal{B},$$

where \mathcal{B} is the sigma-algebra of Borel subsets of \mathbb{R}^n .

As μ has a continuous density $f_{\mu} > 0$ on $B_{M/2}$, it follows easily that $M_r(\mathbf{y}) > 0$ for all $r \geq r_f$. In addition,

$$\sum_{k=0}^r \sum_{i=1}^n y_{2k}^{(i)} / k! < \int \sum_{i=1}^n e^{x_i^2} d\mu < ne^{M^2},$$

so that \mathbf{y} is a strictly admissible solution for \mathcal{Q}_r . Hence, the SDP problem \mathcal{Q}_r satisfies Slater's condition, and so, there is no duality gap between \mathcal{Q}_r and \mathcal{Q}_r^* , and \mathcal{Q}_r^* is solvable if $\inf \mathcal{Q}_r$ is finite; see e.g. Vandenberghe and Boyd [18]. Thus, \mathcal{Q}_r^* is solvable because we proved that \mathcal{Q}_r is solvable. In other words, $\sup \mathcal{Q}_r^* = \max \mathcal{Q}_r^* = \min \mathcal{Q}_r$, the desired result.

5.2. Proof of Theorem 4.1. It suffices to prove (i) and (ii) for the case $f^* > 0$. Indeed, if $f^* = 0$ take $\epsilon > 0$ arbitrary, fixed. Then $f + n\epsilon \geq f_\epsilon^* = f^* + n\epsilon > 0$ and so, suppose that (4.1) holds for $f + n\epsilon$ (for some r_0, λ_0). In particular, pick $\lambda \geq \lambda_0 + \epsilon$, so that

$$f + n\epsilon + (\lambda - \epsilon) \sum_{k=0}^{r_\lambda} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = q_\lambda,$$

(with q_λ s.o.s.), for $r_\lambda \geq r_0$. Equivalently,

$$f + \lambda \sum_{k=0}^{r_\lambda} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = q_\lambda + \epsilon \sum_{k=1}^{r_\lambda} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = \hat{q}_\lambda,$$

where \hat{q}_λ is a s.o.s. Hence (4.1) also holds for f (with $\lambda_0 + \epsilon$ in lieu of λ_0).

Similarly, for (4.2). As $f^* = 0$, $f + n\epsilon > 0$ and so, suppose that (4.2) holds for $f + n\epsilon$. In particular,

$$f + n\epsilon + \epsilon \sum_{k=0}^{r_\epsilon} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = q_\epsilon,$$

(with q_ϵ s.o.s.), for some r_ϵ . Equivalently,

$$f + 2\epsilon \sum_{k=0}^{r_\epsilon} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = q_\epsilon + \epsilon \sum_{k=1}^{r_\epsilon} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = \hat{q}_\epsilon,$$

where \hat{q}_ϵ is a s.o.s. Hence (4.2) also holds for f . Therefore, we will assume that $f^* > 0$.

(i) As $f^* > 0$, let M_0 be such that $f^* > 1/M_0$, and fix $M > M_0$. Consider the SDP problem \mathbb{Q}_r^* defined in (3.5), associated with M . By Proposition 3.2, $f^* \leq \inf \mathcal{P}_M$. By Theorem 3.3, $\max \mathbb{Q}_r^* = \min \mathbb{Q}_r \uparrow \inf \mathcal{P}_M \geq f^*$. Therefore, there exists some $r_M \geq r_f$ such that $\max \mathbb{Q}_{r_M}^* \geq f^* - 1/M > 0$. That is, if $(q_M, \lambda_M, \gamma_M)$ is an optimal solution of $\mathbb{Q}_{r_M}^*$, then $\gamma_M - n\lambda_M e^{M^2} \geq f^* - 1/M > 0$. In addition,

$$f - \gamma_M = q_M - \lambda_M \sum_{k=0}^{r_M} \sum_{j=1}^n \frac{x_j^{2k}}{k!},$$

that we rewrite

$$(5.4) \quad f - (\gamma_M - n\lambda_M e^{M^2}) = q_M + \lambda_M \left(n e^{M^2} - \sum_{k=0}^{r_M} \sum_{j=1}^n \frac{x_j^{2k}}{k!} \right).$$

Equivalently,

$$f + \lambda_M \sum_{k=0}^{r_M} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = q_M + n\lambda_M e^{M^2} + (\gamma_M - n\lambda_M e^{M^2}).$$

Define \hat{q}_M to be the s.o.s. polynomial

$$\hat{q}_M := q_M + n\lambda_M e^{M^2} + (\gamma_M - n\lambda_M e^{M^2}),$$

so that we obtain

$$(5.5) \quad f + \lambda_M \sum_{k=0}^{r_M} \sum_{j=1}^n \frac{x_j^{2k}}{k!} = \hat{q}_M,$$

the desired result.

If we now take $r > r_M$ and $\lambda \geq \lambda_M$ we also have

$$\begin{aligned} f + \lambda \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} &= f + \lambda_M \sum_{k=0}^{r_M} \sum_{j=1}^n \frac{x_j^{2k}}{k!} \\ &\quad + \lambda_M \sum_{k=r_M+1}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} + (\lambda - \lambda_M) \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} \\ &= \hat{q}_M + \lambda_M \sum_{k=r_M+1}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} + (\lambda - \lambda_M) \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} \\ &= \hat{\hat{q}}_M, \end{aligned}$$

that is,

$$(5.6) \quad f + \lambda \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!} = \hat{\hat{q}}_M,$$

where $\hat{\hat{q}}_M$ is a s.o.s. polynomial, the desired result.

(ii) Let M be as in (i) above. Evaluating (5.4) at $x = 0$, and writing $f(0) = f(0) - f^* + f^*$, yields

$$f(0) - f^* + f^* - (\gamma_M - n\lambda_M e^{M^2}) = q_M(0) + n\lambda_M (e^{M^2} - 1),$$

and as $1/M \geq f^* - (\gamma_M - n\lambda_M e^{M^2})$,

$$\lambda_M \leq \frac{1/M + f(0) - f^*}{n(e^{M^2} - 1)}.$$

Now, letting $M \rightarrow \infty$, yields $\lambda_M \rightarrow 0$.

Now, let $\epsilon > 0$ be fixed, arbitrary. There is some $M > M_0$ such that $\lambda_M \leq \epsilon$ in (5.5). Therefore, (4.2) is just (5.6) with $\lambda := \epsilon > \lambda_M$ and $r = r_\epsilon \geq r_M$. Finally, from this, we immediately have

$$\|f - f_\epsilon\|_1 \leq \epsilon \sum_{i=1}^n \sum_{k=0}^{\infty} \frac{1}{k!} = \epsilon n e \rightarrow 0, \quad \text{as } \epsilon \downarrow 0.$$

6. Appendix. In this section we derive two auxiliary results that are helpful in the proofs of Theorem 3.3 and Theorem 4.1 in §5.

LEMMA 6.1. *Let $n = 2$ and let \mathbf{y} be a sequence indexed in the basis (2.1), and such that $M_r(\mathbf{y}) \succeq 0$. Then all the diagonal entries of $M_r(\mathbf{y})$ are bounded by $\tau_r := \max_{k=1, \dots, r} \max[y_{2k,0}, y_{0,2k}]$.*

Proof. It suffices to prove that all the entries $y_{2\alpha,2\beta}$ with $\alpha + \beta = r$ are bounded by $s_r := \max[y_{2r,0}, y_{0,2r}]$, and repeat the argument for entries $y_{2\alpha,2\beta}$ with $\alpha + \beta =$

$r-1, r-2$, etc ... Then, take $\tau_r := \max_{k=1, \dots, r} s_k$. So, consider the odd case $r = 2p+1$, and the even case $r = 2p$.

- The odd case $r = 2p+1$. Let $\Gamma := \{(2\alpha, 2\beta) \mid \alpha + \beta = r, \alpha, \beta \neq 0\}$, and notice that

$$\Gamma = \{(2r-2k, 2k) \mid k = 1, \dots, r-1\} = \Gamma_1 \cup \Gamma_2$$

with

$$\Gamma_1 := \{(r, 0) + (r-2k, 2k), \mid k = 1, \dots, p\},$$

and

$$\Gamma_2 := \{(0, r) + (2j, r-2j), \mid j = 1, \dots, p\}.$$

Therefore, consider the two rows (and columns) corresponding to the indices $(r, 0)$ and $(r-2k, 2k)$, or $(0, r)$ and $(2j, r-2j)$. In view of $M_r(\mathbf{y}) \succeq 0$, one has

$$(6.1) \quad \begin{cases} y_{2r,0} \times y_{2r-4k,4k} & \geq (y_{2r-2k,2k})^2, & k = 1, \dots, p, \\ y_{0,2r} \times y_{4j,2r-4j} & \geq (y_{2j,2r-2j})^2, & j = 1, \dots, p. \end{cases}$$

Thus, let $s := \max\{y_{2\alpha,2\beta} \mid \alpha + \beta = r, \alpha\beta \neq 0\}$, so that either $s = y_{2r-2k^*,2k^*}$ for some $1 \leq k^* \leq p$, or $s = y_{2j^*,2r-2j^*}$ for some $1 \leq j^* \leq p$. But then, in view of (6.1), and with $s_r := \max[y_{2r,0}, y_{0,2r}]$,

$$s_r \times s \geq y_{2r,0} \times y_{2r-4k^*,4k^*} \geq (y_{2r-2k^*,2k^*})^2 = s^2,$$

or,

$$s_r \times s \geq y_{0,2r} \times y_{4j^*,2r-4j^*} \geq (y_{2j^*,2r-2j^*})^2 = s^2,$$

so that $s \leq s_r$, the desired result.

- The even case $r = 2p$. Again, the set $\Gamma := \{(2\alpha, 2\beta) \mid \alpha + \beta = r, \alpha\beta \neq 0\}$ can be written $\Gamma = \Gamma_1 \cup \Gamma_2$, with

$$\Gamma_1 := \{(r, 0) + (r-2k, 2k), \mid k = 1, \dots, p\},$$

and

$$\Gamma_2 := \{(0, r) + (2j, r-2j), \mid j = 1, \dots, p\}.$$

The only difference with the odd case is that $\Gamma_1 \cap \Gamma_2 = (2p, 2p) \neq \emptyset$. But the rest of the proof is the same as in the odd case. \square

LEMMA 6.2. *Let $r \in \mathbb{N}$ be fixed, and let \mathbf{y} be a sequence such that the associated moment matrix $M_r(\mathbf{y})$ is positive semidefinite, i.e., $M_r(\mathbf{y}) \succeq 0$. Assume that there is some $\tau_r \in \mathbb{R}$ such that the diagonal elements $\{y_{2k}^{(i)}\}$ satisfy $y_{2k}^{(i)} \leq \tau_r$, for all $k = 1, \dots, r$, and all $i = 1, \dots, n$.*

Then, the diagonal elements of $M_r(\mathbf{y})$ are all bounded by τ_r (i.e., $y_{2\alpha} \leq \tau_r$ for all $\alpha \in \mathbb{N}^n$, with $|\alpha| \leq r$).

Proof. The proof is by induction on the the number n of variables. By our assumption it is true for $n = 1$, and by Lemma 6.1, it is true for $n = 2$. Thus, suppose it is true for $k = 1, 2, \dots, n-1$ variables and consider the case of n variables (with $n \geq 3$).

By our induction hypothesis, it is true for all elements $y_{2\alpha}$ where at least one index, say α_i , is zero ($\alpha_i = 0$). Indeed, the submatrix $A_r^{(i)}(\mathbf{y})$ of $M_r(\mathbf{y})$, obtained from $M_r(\mathbf{y})$ by deleting all rows and columns corresponding to indices $\alpha \in \mathbb{N}^n$ in the basis (2.1), with $\alpha_i > 0$, is a moment matrix of order r , with $n - 1$ variables $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$. Hence, by a permutation of rows and columns, we can write

$$M_r(\mathbf{y}) = \left[\begin{array}{c|c} A_r^{(i)}(\mathbf{y}) & B \\ \hline B' & C \end{array} \right],$$

for some appropriate matrices B and C . In particular, all elements $y_{2\alpha}$ with $\alpha_i = 0$, are diagonal elements of $A_r^{(i)}(\mathbf{y})$. In addition, its diagonal elements $y_{2k}^{(j)}$, $j \neq i$, are all bounded by τ_r . And of course, $A_r^{(i)}(\mathbf{y}) \succeq 0$. Therefore, by our induction hypothesis, all its diagonal elements are bounded by τ_r . As i was arbitrary, we conclude that all elements $y_{2\alpha}$ with at least one index being zero, are all bounded by τ_r .

We next prove it is true for an arbitrary element $y_{2\alpha}$ with $|\alpha| \leq r$ and $\alpha > 0$, i.e., $\alpha_j \geq 1$ for all $j = 1, \dots, n$. With no loss of generality, we assume that $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$.

Consider the two elements $y_{2\alpha_1, 0, \beta}$ and $y_{0, 2\alpha_2, \gamma}$, with $\beta, \gamma \in \mathbb{N}^{n-2}$ such that:

$$|\beta| = |\alpha| - 2\alpha_1; \quad |\gamma| = |\alpha| - 2\alpha_2,$$

and

$$(2\alpha_1, 0, \beta) + (0, 2\alpha_2, \gamma) = (2\alpha_1, 2\alpha_2, \beta + \gamma) = 2\alpha.$$

So, for instance, take $\beta = (\beta_3, \beta_4, \dots, \beta_n)$, $\gamma = (\gamma_3, \gamma_4, \dots, \gamma_n)$, defined by

$$\beta := (\alpha_3 + \alpha_2 - \alpha_1, \alpha_4, \dots, \alpha_n), \quad \gamma := (\alpha_3 + \alpha_1 - \alpha_2, \alpha_4, \dots, \alpha_n).$$

By construction, we have $4\alpha_1 + 2|\beta| = 4\alpha_2 + 2|\gamma| = 2|\alpha| \leq 2r$, so that both $y_{4\alpha_1, 0, 2\beta}$ and $y_{0, 4\alpha_2, 2\gamma}$ are diagonal elements of $M_r(\mathbf{y})$ with at least one entry equal to 0. Hence, by the induction hypothesis,

$$y_{4\alpha_1, 0, 2\beta} \leq \tau_r, \quad y_{0, 4\alpha_2, 2\gamma} \leq \tau_r.$$

Next, consider the two rows and columns indexed by $(2\alpha_1, 0, \beta)$ and $(0, 2\alpha_2, \gamma)$. The constraint $M_r(\mathbf{y}) \succeq 0$ clearly implies

$$\tau_r^2 \geq y_{4\alpha_1, 0, 2\beta} \times y_{0, 4\alpha_2, 2\gamma} \geq (y_{2\alpha_1, 2\alpha_2, \beta + \gamma})^2 = y_{2\alpha}^2.$$

Hence, $y_{2\alpha} \leq \tau_r$, the desired result. \square

Acknowledgement: The authors wishes to thank anonymous referees for their valuable comments and suggestions, as well as Bruce Reznick who kindly provided the two examples in §4.

REFERENCES

- [1] R. ASH, *Real Analysis and Probability*, Academic Press, San Diego, 1972.
- [2] C. BERG, J.P.R. CHRISTENSEN, P. RESSEL, *Positive definite functions on abelian semigroups*, Math. Ann. 223 (1976), pp. 253–274.

- [3] C. BERG, *The multidimensional problem and semigroups*, in: *Moments in Mathematics*, AMS short course, San Antonio, Texas, 1987, Proc. Symp. Appl. Math. 37 (1987), pp. 110–124.
- [4] G. BLEKHERMAN, *There are significantly more nonnegative polynomials than sums of squares*, Department of Mathematics, University of Michigan, Ann Arbor, USA, 2004.
- [5] T. JACOBI, A. PRESTEL, *Distinguished representations of strictly positive polynomials*, J. Reine. Angew. Math. 532 (2001), pp. 223–235.
- [6] J.B. LASSERRE. *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. 11 (2001), pp. 796–817.
- [7] J.B. LASSERRE. *Polynomials nonnegative on a grid and discrete optimization*, Trans. Amer. Math. Soc. 354 (2002), pp. 631–649.
- [8] J.B. LASSERRE. *Semidefinite programming vs. LP relaxations for polynomial programming*, Math. Oper. Res. 27 (2002), pp. 347–360.
- [9] A.E. NUSSBAUM, *Quasi-analytic vectors*, Ark. Mat. 6 (1966), pp. 179–191.
- [10] P.A. PARRILO. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, PhD thesis, California Institute of Technology, Pasadena, CA, 2000.
- [11] P.A. PARRILO. *Semidefinite programming relaxations for semialgebraic problems*, Math. Progr. Ser. B 96 (2003), pp. 293–320.
- [12] A. PRESTEL, C.N. DELZELL, *Positive Polynomials*, Springer, Berlin, 2001.
- [13] M. PUTINAR. *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. 42 (1993), pp. 969–984.
- [14] R.T. ROCKAFFELAR, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
- [15] C. SCHEIDERER, *Positivity and sums of squares: A guide to some recent results*, Department of Mathematics, University of Duisburg, Germany.
- [16] K. SCHMÜDGEN, *The K -moment problem for compact semi-algebraic sets*, Math. Ann. 289 (1991), pp. 203–206.
- [17] M. SCHWEIGHOFER, *Optimization of polynomials on compact semialgebraic sets*, SIAM J. Optim. to appear.
- [18] L. VANDENBERGHE AND S. BOYD, *Semidefinite programming*, SIAM Review 38 (1996), pp. 49–95.