Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

SOS approximation of polynomials nonnegative on an algebraic set

Jean B. Lasserre

***Abstract*— Let $V \subset \mathbb{R}^n$ be a real algebraic set described by finitely many polynomials equations $g_j(x) = 0, j \in J$, and let f be a real polynomial, nonnegative on V. We show that for every $\epsilon > 0$, there exist nonnegative scalars $\{\lambda_j\}_{j \in J}$ such that, for all r sufficiently large,**

$$f_{\epsilon r} + \sum_{j \in J} \lambda_j \, g_j^2, \quad \textbf{is a sum of squares,}$$

for some polynomial $f_{\epsilon r}$ with a simple and *explicit* form in terms of f and the parameters $\epsilon > 0, r \in \mathbb{N}$, and such that $\|f - f_{\epsilon r}\|_1 \to 0$ as $\epsilon \to 0$.

This representation is an obvious certificate of nonnegativity of $f_{\epsilon r}$ on V, and valid with *no* assumption on V. In addition, this representation is also useful from a computational point of view, as we can define semidefinite programming relaxations to approximate the global minimum of f on a real algebraic set V, or a basic closed semi-algebraic set $\mathbf{K}$, and again, with *no* assumption on V or $\mathbf{K}$.

I. Introduction

Let $\mathbb{R}[x](= \mathbb{R}[x_1, \ldots, x_n])$, and let $V \subset \mathbb{R}^n$ be the real algebraic set

$$V \;:=\; \{x \in \mathbb{R}^n \,|\quad g_j(x) \,=\, 0, \quad j = 1, \ldots, m\}, \tag{1}$$

for some family of real polynomials $\{g_j\} \subset \mathbb{R}[x]$.

The main motivation of this paper is to provide a characterization of polynomials $f \in \mathbb{R}[x]$, nonnegative on V, in terms of a *certificate* of positivity. In addition, and in view of the many potential applications (notably in control, as described in e.g. Henrion and Lasserre [4]), one would like to obtain a representation that is also *useful* from a computational point of view.

In some particular cases, when V is *compact*, and viewing the equations $g_j(x) = 0$ as two opposite inequations $g_j(x) \geq 0$ and $g_j(x) \leq 0$, one may obtain Schmüdgen's *sum of squares* (s.o.s.) representation [18] for $f + \epsilon$ ($\epsilon > 0$), instead of f. In this equality case, $f + \epsilon$ reads

$$f + \epsilon \;=\; f_0 + \sum_{j=1}^{m} f_j \, g_j, \tag{2}$$

for some polynomials $\{f_j\} \subset \mathbb{R}[x]$, with f_0 a s.o.s. Hence, if f is nonnegative on V, every approximation $f + \epsilon$ of f (with $\epsilon > 0$) has the representation (2). The interested reader is referred to Marshall [11], Prestel and Delzell [13], and Scheiderer [16], [17] for a nice account of such results.

Contribution. We propose the following result: Let $\|f\|_1 = \sum_\alpha |f_\alpha|$ whenever $x \mapsto f(x) = \sum_\alpha f_\alpha x^\alpha$. Let $f \in \mathbb{R}[x]$ be nonnegative on V, as defined in (1), and let $F := \{f_{\epsilon r}\}_{\epsilon, r}$ be the family of polynomials

$$f_{\epsilon r} \;=\; f + \epsilon \sum_{k=0}^{r} \sum_{i=1}^{n} \frac{x_i^{2k}}{k!}, \qquad \epsilon \geq 0, \quad r \in \mathbb{N}. \tag{3}$$

(So, for every $r \in \mathbb{N}$, $\|f - f_{\epsilon r}\|_1 \to 0$ as $\epsilon \downarrow 0$.)

Then, for every $\epsilon > 0$, there exist nonnegative *scalars* $\{\lambda_j\}_{j=1}^m$, such that for all r sufficiently large (say $r \geq r(\epsilon)$),

$$f_{\epsilon r} \;=\; q_\epsilon - \sum_{j=1}^{m} \lambda_j \, g_j^2, \tag{4}$$

for some s.o.s. polynomial $q_\epsilon \in \mathbb{R}[x]$, that is, $f_{\epsilon r} + \sum_{j=1}^m \lambda_j \, g_j^2$ is s.o.s.

Thus, with *no* assumption on the set V, one obtains a representation of $f_{\epsilon r}$ (which is positive on V as $f_{\epsilon r} > f$ for all $\epsilon > 0$) in the simple and explicit form (4), an obvious *certificate* of positivity of $f_{\epsilon r}$ on V. In particular, when $V \equiv \mathbb{R}^n$, one retrieves the result of [9], which states that every nonnegative real polynomial f can be aproximated as closely as desired, by a family of s.o.s. polynomials $\{f_{\epsilon r(\epsilon)}\}_\epsilon$, with $f_{\epsilon r}$ as in (3).

Notice that $f + n\epsilon = f_{\epsilon 0}$. So, on the one hand, the approximation $f_{\epsilon r}$ in (4) is more complicated than $f + \epsilon$ in (2), valid for the compact case with an additional assumption, but on the other hand, the coefficients of the g_j's in (4) are now *scalars* instead of s.o.s., and (4) is valid for an arbitrary algebraic set V.

The case of a semi-algebraic set $\mathbf{K} = \{x \in \mathbb{R}^n | g_j(x) \geq 0, \; j = 1, \ldots, m\}$ reduces to the case of an algebraic set $V \in \mathbb{R}^{n+m}$, by introducing m *slack* variables $\{z_j\}$, and replacing $g_j(x) \geq 0$ with $g_j(x) - z_j^2 = 0$, for all $j = 1, \ldots, m$. Let $f \in \mathbb{R}[x]$ be nonnegative on $\mathbf{K}$. Then, for every $\epsilon > 0$, there exist nonnegative scalars $\{\lambda_j\}_{j=1}^m$ such that, for all sufficiently large r,

$$f + \epsilon \sum_{k=0}^{r} \left[\sum_{i=1}^{n} \frac{x_i^{2k}}{k!} + \sum_{j=1}^{m} \frac{z_j^{2k}}{k!} \right] \;=\; q_\epsilon - \sum_{j=1}^{m} \lambda_j \, (g_j - z_j^2)^2,$$

for some s.o.s. $q_\epsilon \in \mathbb{R}[x, z]$. Equivalently, everywhere on $\mathbf{K}$, the polynomial

$$x \mapsto f(x) + \epsilon \sum_{k=0}^{r} \sum_{i=1}^{n} \frac{x_i^{2k}}{k!} + \epsilon \sum_{k=0}^{r} \sum_{j=1}^{m} \frac{g_j(x)^k}{k!}$$

coincides with the polynomial $x \mapsto q_\epsilon(x_1, \ldots, x_n, \sqrt{g_1(x)}, \ldots, \sqrt{g_m(x)})$, obviously nonnegative.

The representation (4) is also useful for computational purposes. Indeed, for instance, using (4) with fixed ϵ, one

J.B. Lasserre is with LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse Cédex 4, France. lasserre@laas.fr

0-7803-9568-9/05/$20.00 ©2005 IEEE

can approximate the global minimum of f on V, by solving a sequence of semidefinite programming (SDP) problems. The same applies to an arbitrary semi-algebraic set $\mathbf{K} \subset \mathbb{R}^n$, defined by m polynomials inequalities, as explained above. Again, and in contrast to previous SDP-relaxation techniques as in e.g. [6], [7], [8], [12], [19], no compacity assumption on V or $\mathbf{K}$ is required.

In a sense, the family $F = \{f_{\epsilon r}\} \subset \mathbb{R}[x]$ (with $f_{0r} \equiv f$) is a set of *regularizations* of f, because one may approximate f by members of F, and those members *always* have nice representations when f is nonnegative on an algebraic set V (including the case $V \equiv \mathbb{R}^n$), whereas f itself might not have such a nice representation.

Methodology. To prove our main result, we proceed in three main steps.

1. We first define an infinite dimensional linear programming problem on an appropriate space of measures, whose optimal value is the global minimum of f on the set V.

2. We then prove a crucial result, namely that there is *no* duality gap between this linear programming problem and its dual. The approach is similar but different from that taken in [9] when $V \equiv \mathbb{R}^n$. Indeed, the approach in [9] does *not* work when $V \not\equiv \mathbb{R}^n$. Here, we use the important fact that the polynomial θ_r (defined in (11) below) is a *moment function*. And so, if a set of probability measures Π satisfies $\sup_{\mu\in\Pi} \int \theta_r d\mu < \infty$, it is *tight*, and therefore, by Prohorov's theorem, relatively compact. This latter intermediate result is crucial for our purpose.

3. In the final step, we use our recent result [9] which states that if a polynomial $h \in \mathbb{R}[x]$ is nonnegative on $\mathbb{R}^n$, then $h + \epsilon\theta_r$ ($\epsilon > 0$) is a sum of squares, provided that r is sufficiently large.

The paper in organized as follows. After introducing the notation and definitions in §II, some preliminary results are stated in §III, whereas our main result is stated and discussed in §IV. The detailed proofs can be found in Lasserre [10].

II. Notation and definitions

Let $\mathbb{R}_+ \subset \mathbb{R}$ denote the cone of nonnegative real numbers. 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 $\mathbb{R}[x]$ be the ring of real polynomials, and let

$$v_r(x) := (1, x_1, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^r) \qquad (5)$$

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 $\mathbb{R}[x]$ as a $\mathbb{R}$-vector space, denoted $\mathcal{A}$. So a vector in $\mathcal{A}$ has always *finitely* many zeros.

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

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

(where $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$) for some vector $\mathbf{p} = \{p_\alpha\} \in \mathbb{R}^{s(r)}$, the vector of coefficients of p in the basis (5).

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 $\langle .,. \rangle$ of vectors, then for every $p \in \mathcal{A}$,

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

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

$$p \mapsto L_\mathbf{y}(p) := \sum_{\alpha\in\mathbb{N}^n} p_\alpha y_\alpha = \langle \mathbf{p}, \mathbf{y}\rangle. \qquad (6)$$

Given a sequence $\mathbf{y} = \{y_\alpha\}$ indexed in the basis $v_\infty(x)$, define 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 (5), by

$$M_r(y)(\alpha,\beta) = y_{\alpha+\beta}, \quad |\alpha|, |\beta| \leq r. \qquad (7)$$

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

$$y_\alpha = \int_{\mathbb{R}^n} x^\alpha \, d\mu_\mathbf{y}, \qquad \forall\, \alpha \in \mathbb{N}^n. \qquad (8)$$

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 $\langle .,. \rangle_\mathbf{y}$ 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

$$L_\mathbf{y}(q^2) = \langle \mathbf{q}, M_r(\mathbf{y})\mathbf{q}\rangle = \int_{\mathbb{R}^n} q^2 \, d\mu_\mathbf{y} \geq 0, \quad \forall\, q \in \mathcal{A}_r, \qquad (9)$$

so that $M_r(\mathbf{y})$ is positive semidefinite, i.e., $M_r(\mathbf{y}) \succeq 0$.

III. Preliminaries

Let $V \subset \mathbb{R}^n$ be the real algebraic set defined in (1), and let B_M be the closed ball

$$B_M = \{x \in \mathbb{R}^n \,|\, \quad \|x\|_\infty \leq M\}. \qquad (10)$$

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

$$f_M^* := \inf\{f(x) | x \in B_M \cap V\} < f^* + \epsilon, \qquad \forall M \geq M_\epsilon.$$

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

The proof is rather elementary and left to the reader.

For every $r \in \mathbb{N}$, let $\theta_r \in \mathbb{R}[x]$ be the polynomial

$$x \mapsto \theta_r(x) := \sum_{k=0}^{r} \sum_{i=1}^{n} \frac{x_i^{2k}}{k!}, \qquad x \in \mathbb{R}^n, \qquad (11)$$

and notice that $n \leq \theta_r(x) \leq \sum_{i=1}^n \mathrm{e}^{x_i^2} =: \theta_\infty(x)$, for all $x \in \mathbb{R}^n$. Moreover, θ_r satisfies

$$\lim_{M\to\infty} \inf_{x \in B_M^c} \theta_r(x) = +\infty, \qquad (12)$$

where B_M^c denotes the complement of B_M in $\mathbb{R}^n$; this property is very useful because in particular, a set Π of probability measures satisfying $\sup_{\mu\in\Pi} \int \theta_r \, d\mu < \infty$, is *relatively compact* (see e.g. Hernandez-Lerma and Lasserre [5, p. 10]).

Next, with V as in (1), introduce the following optimization problems.

$$\mathbf{P}: \qquad f^* := \inf_{x\in V} f(x), \qquad (13)$$

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

$$\mathcal{P}_M^r : \left\{ \begin{array}{lll} & \inf\limits_{\mu} \int f \, d\mu & \\ \text{s.t.} & \int g_j^2 \, d\mu & \leq 0, \quad j = 1, \ldots, m \\ & \int \theta_r \, d\mu & \leq n\mathrm{e}^{M^2} \\ & \mu \in \mathcal{P}(\mathbb{R}^n), & \end{array} \right. \qquad (14)$$

where $\mathcal{P}(\mathbb{R}^n)$ is the space of probability measures on $\mathbb{R}^n$ (with $\mathcal{B}$ its associated Borel σ-algebra). The respective optimal values of $\mathbf{P}$ and $\mathcal{P}_M^r$ are denoted $\inf \mathbf{P} = f^*$ and $\inf \mathcal{P}_M^r$, or $\min \mathbf{P}$ and $\min \mathcal{P}_M^r$ if the minimum is attained (in which case, the problem is said to be solvable).

Proposition 3.2: Let $f \in \mathbb{R}[x]$, and let $\mathbf{P}$ and $\mathcal{P}_M^r$ be as in (13) and (14) respectively. Assume that $f^* > -\infty$. Then, for every $r \in \mathbb{N} \cup \{\infty\}$, $\inf \mathcal{P}_M^r \downarrow f^*$ as $M \to \infty$. If f has a global minimizer $x^* \in V$, then $\min \mathcal{P}_M^r = f^*$ whenever $M \geq \|x^*\|_\infty$.

Proof: When M is sufficiently large, $B_M \cap V \neq \emptyset$, and so, $\mathcal{P}_M^r$ is consistent, and $\inf \mathcal{P}_M^r < \infty$. Let $\mu \in \mathcal{P}(\mathbb{R}^n)$ be admissible for $\mathcal{P}_M^r$. From $\int g_j^2 \, d\mu \leq 0$ for all $j = 1, \ldots, m$, it follows that $g_j(x)^2 = 0$ for μ-almost all $x \in \mathbb{R}^n$, $j = 1, \ldots, m$, That is, for every $j = 1, \ldots, m$, there exists a set $A_j \in \mathcal{B}$ such that $\mu(A_j^c) = 0$ and $g_j(x) = 0$ for all $x \in A_j$. Take $A = \cap_j A_j \in \mathcal{B}$ so that $\mu(A^c) = 0$, and for all $x \in A$, $g_j(x) = 0$ for all $j = 1, \ldots, m$. Therefore, $A \subset V$, and as $\mu(A^c) = 0$,

$$\int_{\mathbb{R}^n} f \, d\mu = \int_A f \, d\mu \geq f^* \quad \text{because } f \geq f^* \text{ on } A \subset V,$$

which proves $\inf \mathcal{P}_M^r \geq f^*$.

As V is closed and B_M is closed and bounded, the set $B_M \cap V$ is compact and so, with f_M^* as in Proposition 3.1, there is some $\hat{x} \in B_M \cap V$ 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 $\|\hat{x}\|_\infty \leq M$,

$$\int \theta_r \, d\mu = \theta_r(\hat{x}) \leq n\mathrm{e}^{M^2}.$$

Moreover, as $\hat{x} \in V$, $g_j(\hat{x}) = 0$, for all $j = 1, \ldots, m$, and so

$$\int g_j^2 \, d\mu = g_j(\hat{x})^2 = 0, \quad j = 1, \ldots, m,$$

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

Consider now, the following optimization problem $\mathcal{Q}_M^r$, the dual problem of $\mathcal{P}_M^r$, i.e.,

$$\mathcal{Q}_M^r : \quad \begin{array}{lll} \max\limits_{\lambda,\delta,\gamma} & \gamma - n\delta\mathrm{e}^{M^2} & \\ \text{s.t.} & f + \delta\theta_r + \sum_{j=1}^m \lambda_j g_j^2 & \geq \gamma \\ & \gamma \in \mathbb{R}, \delta \in \mathbb{R}_+, \lambda \in \mathbb{R}_+^m, & \end{array} \qquad (15)$$

with optimal value denoted by $\sup \mathcal{Q}_M^r$. Indeed, $\mathcal{Q}_M^r$ is a dual of $\mathcal{P}_M^r$ because *weak duality* holds. To see this, consider any two feasible solutions $\mu \in \mathcal{P}(\mathbb{R}^n)$ and $(\lambda, \delta, \gamma) \in \mathbb{R}_+^m \times \mathbb{R}_+ \times \mathbb{R}$, of $\mathcal{P}_M^r$ and $\mathcal{Q}_M^r$, respectively. Then, integrating both sides of the inequality in $\mathcal{Q}_M^r$ with respect to μ, yields

$$\int f d\mu + \delta \int \theta_r \, d\mu + \sum_{j=1}^m \lambda_j \int g_j^2 \, d\mu \geq \gamma,$$

and so, using that μ is feasible for $\mathcal{P}_M^r$,

$$\int f d\mu \geq \gamma - \delta n\mathrm{e}^{M^2}.$$

Hence, the value of any feasible solution of $\mathcal{Q}_M^r$ is always smaller than the value of any feasible solution of $\mathcal{P}_M^r$, i.e., weak duality holds. In fact, we have the more important and crucial following result.

Theorem 3.3: Let M be large enough so that $B_M \cap V \neq \emptyset$. Let $f \in \mathbb{R}[x]$, and let $r_0 > \max[\deg f, \deg g_j]$. Then, for every $r \geq r_0$, $\mathcal{P}_M^r$ is solvable, and there is no duality gap between $\mathcal{P}_M^r$ and its dual $\mathcal{Q}_M^r$. That is, $\sup \mathcal{Q}_M^r = \min \mathcal{P}_M^r$.

For a proof see Lasserre [10].

We finally end up this section by re-stating a result proved in [9], which, together with Theorem 3.3, will be crucial to prove our main result.

Theorem 3.4 ([9]): Let $f \in \mathbb{R}[x]$ be nonnegative. Then for every $\epsilon > 0$, there is some $r(f, \epsilon) \in \mathbb{N}$ such that,

$$f_{\epsilon r(f,\epsilon)} \; (= f + \epsilon\theta_{r(f,\epsilon)}) \qquad \text{is a sum of squares,} \qquad (16)$$

and so is $f_{\epsilon r}$, for all $r \geq r(f, \epsilon)$.

IV. Main result

Recall that for given $(\epsilon, r) \in \mathbb{R} \times \mathbb{N}$, $f_{\epsilon r} = f + \epsilon\theta_r$, with $\theta_r \in \mathbb{R}[x]$ being the polynomial defined in (11). We now state our main result:

Theorem 4.1: Let $V \subset \mathbb{R}^n$ be as in (1), and let $f \in \mathbb{R}[x]$ be nonnegative on V. Then, for every $\epsilon > 0$, there exists $\overline{r} \in \mathbb{N}$ and nonnegative scalars $\{\lambda_j\}_{j=1}^m$, such that, for all $r \geq \overline{r}$,

$$f_{\epsilon r} \;=\; q - \sum_{j=1}^m \lambda_j\, g_j^2, \tag{17}$$

for some s.o.s. polynomial $q \in \mathbb{R}[x]$. In addition, $\|f - f_{\epsilon r}\|_1 \to 0$, as $\epsilon \downarrow 0$.
For a proof see Lasserre [10].

Remark 4.2: (i) Observe that (17) is an obvious certificate of positivity of $f_{\epsilon r}$ on the algebraic set V, because everywhere on V, $f_{\epsilon r}$ coincides with the s.o.s. polynomial q. Also, as (17) is true for arbitrary $\epsilon > 0$, it follows easily from (17) that f is nonnegative on V. Therefore, when f is nonnegative on V, one obtains with *no* assumption on the algebraic set V, a certificate of positivity for any approximation $f_{\epsilon r}$ of f (with $r \geq \overline{r}$), whereas f itself might not have such a representation. In other words, the (ϵ, r)–perturbation $f_{\epsilon r}$ of f, has a *regularization* effect on f as it permits to derive nice representations.

(ii) From the proof of Theorem 4.1, instead of the representation (17), one may also provide the alternative representation

$$f_{\epsilon r} \;=\; q - \lambda \sum_{j=1}^m g_j^2,$$

for some s.o.s. polynomial q, and some (single) nonnegative scalar λ (instead of m nonnegative scalars in (17)).

A. The case of a semi-algebraic set

We now consider the representation of polynomials, nonnegative on a semi algebraic set $\mathbf{K} \subset \mathbb{R}^n$, defined as,

$$\mathbf{K} \;:=\; \{x \in \mathbb{R}^n \mid \quad g_j(x) \,\geq\, 0, \quad j = 1, \ldots, m\}, \tag{18}$$

for some family $\{g_j\}_{j=1}^m \subset \mathbb{R}[x]$.

One may apply the machinery developed previously for algebraic sets, because the semi-algebraic set $\mathbf{K}$ may be viewed as the projection on $\mathbb{R}^n$, of an algebraic set in $\mathbb{R}^{n+m}$. Indeed, let $V \subset \mathbb{R}^{n+m}$ be the algebraic set defined by

$$V \;=\; \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid g_j(x) - z_j^2 = 0,\; j = 1, \ldots, m\}. \tag{19}$$

Then every $x \in \mathbf{K}$ is associated with the point $(x, \sqrt{g_1(x)}, \ldots, \sqrt{g_m(x)}) \in V$.

Let $\mathbb{R}[z] := \mathbb{R}[z_1, \ldots, z_m]$, and $\mathbb{R}[x, z] := \mathbb{R}[x_1, \ldots x_n, z_1, \ldots, z_m]$, and for every $r \in \mathbb{N}$, let $\varphi_r \in \mathbb{R}[z]$ be the polynomial

$$z \;\mapsto\; \varphi_r(z) \;=\; \sum_{k=0}^r \sum_{j=1}^m \frac{z_j^{2k}}{k!}. \tag{20}$$

We then get :

Corollary 4.3: Let $\mathbf{K}$ be as in (18), and θ_r, φ_r be as in (11) and (20). Let $f \in \mathbb{R}[x]$ be nonnegative on $\mathbf{K}$. Then, for every $\epsilon > 0$, there exist nonnegative scalars $\{\lambda_j\}_{j=1}^m$ such that, for all r sufficiently large,

$$f + \epsilon\theta_r + \epsilon\varphi_r \;=\; q_\epsilon - \sum_{j=1}^m \lambda_j (g_j - z_j^2)^2, \tag{21}$$

for some s.o.s. polynomial $q_\epsilon \in \mathbb{R}[x, z]$.

Equivalently, everywhere on $\mathbf{K}$, the polynomial

$$x \mapsto f(x) + \epsilon \sum_{k=0}^r \sum_{i=1}^n \frac{x_i^{2k}}{k!} \;+\; \epsilon \sum_{k=0}^r \sum_{j=1}^m \frac{g_j(x)^k}{k!} \tag{22}$$

coincides with the nonnegative polynomial $x \mapsto q_\epsilon(x, \sqrt{g_1(x)}, \ldots, \sqrt{g_m(x)})$.

So, as for the case of an algebraic set $V \subset \mathbb{R}^n$, (21) is an obvious certificate of positivity on the semi-algebraic set $\mathbf{K}$, for the polynomial $f_{\epsilon r} \in \mathbb{R}[x, z]$

$$f_{\epsilon r} \;:=\; f + \epsilon\theta_r + \epsilon\varphi_r,$$

and in addition, viewing f as an element of $\mathbb{R}[x, z]$, one has $\|f - f_{\epsilon r}\|_1 \to 0$ as $\epsilon \downarrow 0$. Notice that *no* assumption on $\mathbf{K}$ or on the g_j's that define $\mathbf{K}$, is needed.

B. Computational implications

The results of the previous section can be applied to compute (or at least approximate) the global minimum of f on V. Indeed, with $\epsilon > 0$ fixed, and $2r \geq \max[\deg f, \deg g_j^2]$, consider the convex optimization problem

$$\mathbf{Q}_{\epsilon r} \left\{ \begin{array}{lll} \min\limits_{\mathbf{y}} \; L_{\mathbf{y}}(f_{\epsilon r}), & & \\ M_r(\mathbf{y}) & \succeq & 0 \\ L_{\mathbf{y}}(g_j^2) & \leq & 0, \quad j = 1, \ldots, m \\ y_0 & = & 1, \end{array} \right. \tag{23}$$

where θ_r is as in (11), $L_{\mathbf{y}}$ and $M_r(\mathbf{y})$ are the linear functional and the moment matrix associated with a sequence $\mathbf{y}$ indexed in the basis (5); see (6) and (7) in §II.

$\mathbf{Q}_{\epsilon r}$ is called a semidefinite programming (SDP) problem, and its associated *dual* SDP problem reads

$$\mathbf{Q}^*_{\epsilon r} \left\{ \begin{array}{ll} \max\limits_{\lambda, \gamma, q} & \gamma \\ f_{\epsilon r} - \gamma & = q - \sum\limits_{j=1}^m \lambda_j g_j^2 \\ \lambda \in \mathbb{R}^m, & \lambda \geq 0, \\ q \in \mathbb{R}[x], & q \text{ s.o.s. of degree } \leq 2r. \end{array} \right. \tag{24}$$

The optimal values are denoted $\inf \mathbf{Q}_{\epsilon r}$ and $\sup \mathbf{Q}^*_{\epsilon r}$, respectively (or $\min \mathbf{Q}_{\epsilon r}$, $\max \mathbf{Q}^*_{\epsilon r}$ if the optimum is attained, in which case the problems are said to be solvable). Both problems $\mathbf{Q}_{\epsilon r}$ and its dual $\mathbf{Q}^*_{\epsilon r}$ are nice convex optimization problems that, in principle, can be solved efficiently by standard software packages. For more details on SDP theory, the interested reader is referred to the survey paper [20].

That weak duality holds between $\mathbf{Q}_{\epsilon r}$ and $\mathbf{Q}^*_{\epsilon r}$ is straightforward. Let $\mathbf{y} = \{y_\alpha\}$ and $(\lambda, \gamma, q) \in \mathbb{R}^m_+ \times \mathbb{R} \times \mathbb{R}[x]$ be feasible solutions of $\mathbf{Q}_{\epsilon r}$ and $\mathbf{Q}^*_{\epsilon r}$, respectively. Then, by linearity of $L_{\mathbf{y}}$,

$$\begin{aligned} L_{\mathbf{y}}(f_{\epsilon r}) - \gamma &= L_{\mathbf{y}}(f_{\epsilon r} - \gamma) \\ &= L_{\mathbf{y}}(q - \sum_{j=1}^m \lambda_j g_j^2) \\ &= L_{\mathbf{y}}(q) - \sum_{j=1}^m \lambda_j L_{\mathbf{y}}(g_j^2) \\ &\geq L_{\mathbf{y}}(q) \quad [\text{because } L_{\mathbf{y}}(g_j^2) \leq 0 \text{ for all } j] \\ &\geq 0, \end{aligned}$$

where the latter inequality follows from because q is s.o.s. and $M_r(\mathbf{y}) \succeq 0$; see (9). Therefore, $L_{\mathbf{y}}(f_{\epsilon r}) \geq \gamma$, the desired conclusion. Moreover, $\mathbf{Q}_{\epsilon r}$ is an obvious relaxation of the perturbed problem

$$\mathbf{P}_{\epsilon r}: \quad f^*_{\epsilon r} := \min_x \{f_{\epsilon r} \mid x \in V\}.$$

Indeed, let $x \in V$ and let $\mathbf{y} := v_{2r}(x)$ (see (5)), i.e., $\mathbf{y}$ is the vector of moments (up to order $2r$) of the Dirac measure at $x \in V$. Then, $\mathbf{y}$ is feasible for $\mathbf{Q}_{\epsilon r}$ because $y_0 = 1$, $M_r(\mathbf{y}) \succeq 0$, and $L_{\mathbf{y}}(g_j^2) = g_j(x)^2 = 0$, for all $j = 1, \ldots, m$. Similarly, $L_{\mathbf{y}}(f_{\epsilon r}) = f_{\epsilon r}(x)$. Therefore, $\inf \mathbf{Q}_{\epsilon r} \leq f^*_{\epsilon r}$.

It is not known whether *strong* duality holds between $\mathbf{Q}_{\epsilon r}$ and its dual $\mathbf{Q}^*_{\epsilon r}$.

Theorem 4.4: Let $V \subset \mathbb{R}^n$ be as in (1), and θ_r as in (11). Assume that f has a global minimizer $x^* \in V$ with $f(x^*) = f^*$. Let $\epsilon > 0$ be fixed. Then

$$\begin{aligned} f^* &\leq \sup \mathbf{Q}^*_{\epsilon r} \leq \inf \mathbf{Q}_{\epsilon r} \\ &\leq f^* + \epsilon \theta_r(x^*) \leq f^* + \epsilon \sum_{i=1}^n \mathrm{e}^{(\mathrm{x}_i^*)^2}, \end{aligned} \qquad (25)$$

provided that r is sufficiently large.

Proof: Observe that the polynomial $f - f^*$ is nonnegative on V. Therefore, by Theorem 4.1, for every ϵ there exists $r(\epsilon) \in \mathbb{N}$ and $\lambda(\epsilon) \in \mathbb{R}^m_+$, such that

$$f - f^* + \epsilon \theta_r + \sum_{j=1}^m \lambda_j(\epsilon) g_j^2 = q_\epsilon,$$

for some s.o.s. polynomial $q_\epsilon \in \mathbb{R}[x]$. But this shows that $(\lambda(\epsilon), f^*, q_\epsilon) \in \mathbb{R}^m_+ \times \mathbb{R} \times \mathbb{R}[x]$ is a feasible solution of $\mathbf{Q}^*_{\epsilon r}$ as soon as $r \geq r(\epsilon)$, in which case, $\sup \mathbf{Q}^*_{\epsilon r} \geq f^*$. Moreover, we have seen that $\inf \mathbf{Q}_{\epsilon r} \leq f_{\epsilon r}(x)$ for any feasible solution $x \in V$. In particular, $\inf \mathbf{Q}_{\epsilon r} \leq f^* + \epsilon \theta_r(x^*)$, from which (25) follows. ■

Theorem 4.4 has a nice feature. Suppose that one knows some bound ρ on the norm $\|x^*\|_\infty$ of a global minimizer of f on V. Then, one may fix à priori the error bound η on $|\inf \mathbf{Q}_{\epsilon r} - f^*|$. Indeed, let η be fixed, and fix $\epsilon > 0$ such that $\epsilon \leq \eta(n\mathrm{e}^{\rho^2})^{-1}$. By Theorem 4.4, one has $f^* \leq \inf \mathbf{Q}_{\epsilon r} \leq f^* + \eta$, provided that r is large enough.

The same approach works to approximate the global minimum of a polynomial f on a semi-algebraic set $\mathbf{K}$, as defined in (18). In view of Corollary 4.3, and via a lifting in $\mathbb{R}^{n+m}$, one is reduced to the case of a real algebraic set $V \subset \mathbb{R}^{n+m}$, so that Theorem 4.4 still applies. It is important to emphasize that one requires *no* assumption on $\mathbf{K}$, or on the g_j's that define $\mathbf{K}$. This is to be compared with previous SDP-relaxation techniques developed in e.g. [6], [7], [8], [12], [19], where the set $\mathbf{K}$ is supposed to be compact, and with an additional assumption on the g_j's to ensure that Putinar's representation [14] holds.

V. ACKNOWLEDGMENTS

The author gratefully acknowledges reviewers' comments.

REFERENCES

[1] E.J. Anderson, P. Nash, *Linear Programming in Infinite-Dimensional Spaces*, John Wiley & Sons, Chichester, 1987.
[2] T. Jacobi, A. Prestel, Distinguished representations of strictly positive polynomials, *J. Reine. Angew. Math.* **532**, 2001, pp. 223–235.
[3] D. Henrion, J.B. Lasserre, Solving nonconvex optimization problems, *IEEE Contr. Systems Mag.* **24**, pp. 72–83.
[4] O. Hernández-Lerma, J.B. Lasserre, *Markov Chains and Invariant Probabilities*, Birkhäuser Verlag, Basel, 2003.
[5] J.B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.* **11**, 2001, pp. 796–817.
[6] J.B. Lasserre, Polynomials nonnegative on a grid and discrete optimization, *Trans. Amer. Math. Soc.* **354**, 2002, pp. 631–649.
[7] J.B. Lasserre, Semidefinite programming vs. LP relaxations for polynomial programming, *Math. Oper. Res.* **27**, 2002, pp. 347–360.
[8] J.B. Lasserre, A sum of squares approximation of nonnegative polynomials, Report #04428, LAAS-CNRS, Toulouse, France, 2004. To appear in *SIAM J. Optim.*
[9] J.B. Lasserre, SOS approximations of polynomials, nonnegative on an algebraic set, Report #04480, LAAS-CNRS, Toulouse, France, 2004. To appear in *SIAM J. Optim.*
[10] M. Marshall, Approximating positive polynomials using sums of squares, *Canad. Math. Bull.* **46**, 2003, pp. 400-418.
[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] A.P. Robertson, W. Robertson, Topological Vector Spaces, Cambridge University Press, Cambridge, U.K., 1964.
[15] C. Scheiderer, Positivity and sums of squares: A guide to some recent results, Department of Mathematics, University of Duisburg, Germany.
[16] C. Scheiderer, Sums of squares on real algebraic curves, Department of Mathematics, University of Duisburg, Germany.
[17] K. Schmüdgen, The K-moment problem for compact semi-algebraic sets, *Math. Ann.* **289**, 1991, pp. 203–206.
[18] M. Schweighofer, Optimization of polynomials on compact semialgebraic sets, *SIAM J. Optim* **15**, 2005, 805–825.
[19] L. Vandenberghe, S. Boyd, Semidefinite programming, *SIAM Review* **38**, 1996, pp. 49-95.