SIAM J. OPTIM.
Vol. 11, No. 3, pp. 796–817

GLOBAL OPTIMIZATION WITH POLYNOMIALS AND THE PROBLEM OF MOMENTS*

JEAN B. LASSERRE†

Abstract. We consider the problem of finding the unconstrained global minimum of a real-valued polynomial $p(x) : \mathbb{R}^n \to \mathbb{R}$, as well as the global minimum of $p(x)$, in a compact set K defined by polynomial inequalities. It is shown that this problem reduces to solving an (often finite) sequence of convex linear matrix inequality (LMI) problems. A notion of Karush–Kuhn–Tucker polynomials is introduced in a global optimality condition. Some illustrative examples are provided.

Key words. global optimization, theory of moments and positive polynomials, semidefinite programming

AMS subject classifications. 90C22, 90C25

PII. S1052623400366802

1. Introduction.

Given a real-valued polynomial $p(x) : \mathbb{R}^n \to \mathbb{R}$, we are interested in solving the problem

$$\mathbb{P} \mapsto p^* := \min_{x \in \mathbb{R}^n} p(x), \tag{1.1}$$

that is, finding the *global minimum* p^* of $p(x)$ and, if possible, a global minimizer x^*. We are also interested in solving

$$\mathbb{P}_K \mapsto p_K^* := \min_{x \in K} p(x), \tag{1.2}$$

where K is a (not necessarily convex) compact set defined by polynomial inequalities $g_i(x) \geq 0$, $i = 1, \ldots, r$, which includes many applications of interest and standard problems like quadratic, linear, and 0-1 programming as particular cases.

In the one-dimensional case, that is, when $n = 1$, Shor [17] first showed that (1.1) reduces to a convex problem. Next, Nesterov [13], invoking a well-known representation of nonnegative polynomials as a sum of squares of polynomials, provided a self-concordant barrier for the cone K_{2n} of nonnegative univariate polynomials so that efficient interior point algorithms are available to compute a global minimum.

However, the multivariate case radically differs from the one-dimensional case, for not every nonnegative polynomial can be written as a sum of squares of polynomials. Even more, as mentioned in Nesterov [13], the global unconstrained minimization of a 4-degree polynomial is an NP-hard problem. Via successive changes of variables, Shor [18] (see also Ferrier [5]) proposed to transform (1.2) into a quadratic, quadratically constrained optimization problem and then solve a standard convex linear matrix inequality (LMI) relaxation to obtain good lower bounds. By adding redundant quadratic constraints one may improve the lower bound and sometimes obtain the optimal value.

In this paper, we will show that the global unconstrained minimization (1.1) of a polynomial can be approximated as closely as desired (and often can be obtained exactly) by solving a finite sequence of convex LMI optimization problems of the

*Received by the editors January 28, 2000; accepted for publication (in revised form) August 28, 2000; published electronically January 19, 2001.
http://www.siam.org/journals/siopt/11-3/36680.html

†LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse Cédex 4, France (lasserre@laas.fr).

same flavor as in the one-dimensional case. A similar conclusion also holds for the constrained optimization problem $\mathbb{P}_K$ in (1.2), when K is a compact set, not necessarily convex, defined by polynomial inequalities. The difference between nonnegative and strictly positive polynomials is the reason why, in some cases, only an asymptotic result is possible. Indeed, for the latter, several representations in terms of weighted sums of squares are always possible, whereas few results are known for the former. However, from a numerical point of view, the distinction is irrelevant. In the constrained case, the nonnegative squared polynomials in the representation of the polynomial $p(x)-p_K^*$ can be interpreted as generalized Karush–Kuhn–Tucker multipliers whose value at a global minimizer are precisely the original Karush–Kuhn–Tucker scalar multipliers. This representation of nonnegative polynomials thus provides a natural optimality condition for global optimality.

When the optimal value is obtained at a particular LMI relaxation, the constrained *global* optimization problem thus has a natural "primal" LMI formulation, whose optimal solution provides a global minimizer, whereas an optimal solution of the dual LMI problem provides the Karush–Kuhn–Tucker polynomial multipliers in a representation of $p(x) - p_K^*$. Hence, the primal and dual LMI formulations perfectly match both sides of the same theory (moments and positive polynomials).

This approach is also valid for handling combinatorial problems, e.g., 0-1 programming problems, since the integrality constraint $x_i \in \{0,1\}$ can be written $x_i^2 - x_i \geq 0$ and $x_i - x_i^2 \geq 0$. An elementary illustrative example is provided. We finally consider the general convex quadratic, quadratically constrained problem and provide a natural exact LMI formulation for both primal and dual problems (the Shor relaxation and its dual). The standard linear programming problem also appears as a particular case.

In [13], for the univariate case, the idea was to characterize the nonnegative polynomial $p(x) - p^*$ as a sum of squares. However, we will adopt a dual point of view. Namely, we replace $\mathbb{P}$ and $\mathbb{P}_K$ with the equivalent problems

$$\mathcal{P} \mapsto p^* := \min_{\mu \in \mathcal{P}(\mathbb{R}^n)} \int p(x)\, \mu(dx) \tag{1.3}$$

and

$$\mathcal{P}_K \mapsto p^* := \min_{\mu \in \mathcal{P}(K)} \int p(x)\, \mu(dx), \tag{1.4}$$

respectively, where $\mathcal{P}(\mathbb{R}^n)$ (respectively, $\mathcal{P}(K)$) is the space of finite Borel signed measures on $\mathbb{R}^n$ (respectively, on K). That $\mathcal{P}$ is equivalent to $\mathbb{P}$ is trivial. Indeed, as $p(x) \geq p^*$, then $\int p d\mu \geq p^*$ and thus $\inf \mathcal{P} \geq p^*$. Conversely, if x^* is a global minimizer of $\mathbb{P}$, then the probability measure $\mu^* := \delta_{x^*}$ (the Dirac at x^*) is admissible for $\mathcal{P}$. We then observe that if p is a polynomial of degree, say m, the criterion to minimize is a linear criterion $a'y$ on the finite collection of moments $\{y_\alpha\}$, up to order m, of the probability measure μ. We can then in turn replace $\mathcal{P}$ (respectively, $\mathcal{P}_K$) with an optimization problem on the y_α variables with the constraint that the y_α's must be moments of some probability measure μ. The theory of *moments* provides adequate conditions on the y_α variables. It has been known for a long time that the theory of moments is strongly related to—and in fact, in duality with—the theory of nonnegative polynomials and Hilbert's 17th problem on the representation of nonnegative polynomials. For the historical development and recent results on the theory of moments, the interested reader is referred to Berg [1], Curto and Fialkow [2], [3], Jacobi [8], Putinar [15], Putinar and Vasilescu [14], Simon [19], Schmüdgen [16], and references therein.

The paper is organized as follows. We introduce the notation and some preliminary results in section 2. The unconstrained case is treated in section 3 and the constrained case (1.2) in section 4. Some elementary as well as nontrivial examples are presented for illustration. In the last section we show that when $p(x) - p^*_K$ is a weighted sum of squares, then the squared polynomials can be interpreted as generalized Karush–Kuhn–Tucker multipliers. The convex quadratic case is also investigated.

2. Notation and preliminary results. Let

$$1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1x_n, x_2x_3, \ldots, x_n^2, \ldots, x_1^m, \ldots, x_n^m \tag{2.1}$$

be a basis for the m-degree real-valued polynomials $p(x) : \mathbb{R}^n \to \mathbb{R}$, and let $s(2m)$ be its dimension. We adopt the following standard notation. If $p(x) : \mathbb{R}^n \to \mathbb{R}$ is an m-degree polynomial, write

$$p(x) = \sum_\alpha p_\alpha x^\alpha \quad \text{with } x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \quad \text{and } \sum_i \alpha_i \le m, \tag{2.2}$$

where $p = \{p_\alpha\} \in \mathbb{R}^{s(m)}$ is the coefficient vector of $p(x)$ in the basis (2.1). When needed, a polynomial of degree m can be considered as a polynomial of higher degree, say r, with coefficient vector $p \in \mathbb{R}^{s(r)}$, where the coefficients of monomials of degree higher than m are set to zero.

Given an $s(2m)$-vector $y := \{y_\alpha\}$ with first element $y_{0,\ldots,0} = 1$, let $M_m(y)$ be the *moment* matrix of dimension $s(m)$, with rows and columns labeled by (2.1). For instance, for illustration and clarity of exposition, consider the two-dimensional case. The moment matrix $M_m(y)$ is the block matrix $\{M_{i,j}(y)\}_{0\le i,j\le 2m}$ defined by

$$M_{i,j}(y) = \begin{bmatrix} y_{i+j,0} & y_{i+j-1,1} & \cdots & y_{i,j} \\ y_{i+j-1,1} & y_{i+j-2,2} & \cdots & y_{i-1,j+1} \\ \cdots & \cdots & \cdots & \cdots \\ y_{j,i} & y_{i+j-1,1} & \cdots & y_{0,i+j} \end{bmatrix}, \tag{2.3}$$

where $y_{i,j}$ represents the $(i+j)$-order moment $\int x^i y^j \, \mu(d(x,y))$ for some probability measure μ. To fix ideas, when $n = 2$ and $m = 2$, one obtains

$$M_2(y) = \left[\begin{array}{c|cc|ccc} 1 & y_{1,0} & y_{0,1} & y_{2,0} & y_{1,1} & y_{0,2} \\ \hline y_{1,0} & y_{2,0} & y_{1,1} & y_{3,0} & y_{2,1} & y_{1,2} \\ y_{0,1} & y_{1,1} & y_{0,2} & y_{2,1} & y_{1,2} & y_{0,3} \\ \hline y_{2,0} & y_{3,0} & y_{2,1} & y_{4,0} & y_{3,1} & y_{2,2} \\ y_{1,1} & y_{2,1} & y_{1,2} & y_{3,1} & y_{2,2} & y_{1,3} \\ y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4} \end{array}\right].$$

For the three-dimensional case, $M_m(y)$ is defined via blocks $\{M_{i,j,k}(y)\}$, $0 \le i, j, l \le 2m$ in a similar fashion, and so on.

Let $y = \{y_\alpha\}$ (with $y_{0,\ldots,0} = 1$) be the vector of moments up to order $2m$ of some probability measure μ_y. Let $\mathcal{A}_m$ be the vector space of real-valued polynomials $q(x) : \mathbb{R}^n \to \mathbb{R}$ of degree at most m. Identifying $q(x)$ with its vector $q \in \mathbb{R}^{s(m)}$ of coefficients in the basis (2.1), one may then define a bilinear form $\langle ., \rangle_y : \mathcal{A}_m \times \mathcal{A}_m \to \mathbb{R}$ by

$$\langle q(x), p(x)\rangle_y \;=\; \langle q, M_m(y)p\rangle \;=\; \sum_\alpha (qp)_\alpha y_\alpha \;=\; \int q(x)p(x)\,\mu_y(dx). \tag{2.4}$$

This bilinear form also defines a positive semidefinite form on $\mathcal{A}_m$ since

$$\langle q(x), q(x)\rangle_y \;=\; \sum_\alpha (q^2)_\alpha y_\alpha \;=\; \int q(x)^2\,\mu_y(dx) \;\geq\; 0 \tag{2.5}$$

for all polynomials $q(x) \in \mathcal{A}_m$. The theory of moments identifies those sequences y with $M_m(y) \succeq 0$ that correspond to moments of some probability measure μ_y on $\mathbb{R}^n$.

We first briefly outline the idea developed in the next section: With K an arbitrary (Borel) subset of $\mathbb{R}^n$, one first reduces $\mathbb{P}_K$ to the equivalent convex optimization problem $\mathcal{P}_K$,

$$\mathcal{P}_K \mapsto \min_{\mu\in\mathcal{P}(K)} \int p(x)\,d\mu, \tag{2.6}$$

on the space of Borel probability measures μ with support contained in K. Indeed, we have the following.

PROPOSITION 2.1. *The problems $\mathcal{P}_K$ and $\mathbb{P}_K$ are equivalent, that is,*

(a) $\inf \mathbb{P}_K \;=\; \inf \mathcal{P}_K$.

(b) *if x^* is a global minimizer of $\mathbb{P}_K$, then $\mu^* := \delta_{x^*}$ is a global minimizer of $\mathcal{P}_K$.*

(c) *assuming $\mathbb{P}_K$ has a global minimizer, then, for every optimal solution μ^* of $\mathcal{P}_K$, $p(x) = \min \mathbb{P}_K$, μ^*-almost everywhere (μ^*-a.e.).*

(d) *if x^* is the unique global minimizer of $\mathbb{P}_K$, then $\mu^* := \delta_{x^*}$ is the unique global minimizer of $\mathcal{P}_K$.*

Proof. (a) As for every $x \in K$, $p(x) = \int p\,d\delta_x$, it follows that $\inf \mathcal{P}_K \leq \inf \mathbb{P}_K$ (including the case $-\infty$). Conversely, assume that $p^* := \inf \mathbb{P}_K > -\infty$. As $p(x) \geq p^*$ for all $x \in K$, it follows that $\int p d\mu \geq p^*$ for every probability measure μ with support contained in K.

(b) This proof is trivial.

(c) From (b), $\mathbb{P}_K$ has at least one optimal solution. For an arbitrary optimal solution μ^*, we have $\int p d\mu^* = p^*$ with $p^* = \min \mathbb{P}_K$. Assume that there is a Borel set $B \subset K$ such that $\mu^*(B) > 0$ and $p(x) \neq p^*$ on B, that is, $p(x) > p^*$ on B. Then,

$$\int p\,d\mu^* \;=\; \int_B p\,d\mu^* + \int_{K-B} p\,d\mu^* \;>\; p^*,$$

in contradiction with $\int p d\mu^* = p^*$.

(d) This proof follows from (c). $\square$

Observe that since $p(x)$ is a polynomial of degree, say m, the criterion $\int p d\mu$ involves only the moments of μ, up to order m and, in addition, is *linear* in the moment variables. Therefore, one next replaces μ with the finite sequence $y = \{y_\alpha\}$ of all its moments, up to order m, that is,

$$y_\alpha := \int x^\alpha\,d\mu, \qquad \sum_{i=1}^n \alpha_i = k,\; k = 0, 1, \ldots, m,$$

and one works with the finite sequence y of the moments of μ, up to order m, instead of μ itself. Of course, not every sequence y has a *representing measure* μ; that is, given an arbitrary finite sequence y, there might not be any probability measure μ, all of whose moments up to order m coincide with the y_α scalars. In the one-dimensional case, characterizing those sequences y that have a representing measure on X (respectively, on $[0,\infty)$ and $[a,b]$) is called the (truncated) *Hamburger* (respectively, *Stieltjes* and *Hausdorff*) *moment problem* (see Curto and Fialkow [3] or Simon [19] and references therein). The various necessary and sufficient conditions for the existence of a representing measure μ_y all invoke the positive semidefiniteness of the related (Hankel) moment matrix

$$H_m(y) := \begin{bmatrix} y_0 & y_1 & y_2 & \cdot & y_m \\ y_1 & y_2 & \cdot & \cdot & y_{m+1} \\ \cdot\cdot & \cdot & \cdot & \cdot & \cdot \\ y_m & y_{m+1} & \cdot & y_{2m-1} & y_{2m} \end{bmatrix} \tag{2.7}$$

(see, for instance, the various conditions related to the truncated Hamburger, Stieltjes, and Hausdorff moment problems in Curto and Fialkow [3]). For trigonometric polynomials, Toeplitz matrices are the analogues of the Hankel matrices.

As mentioned earlier, this theory of moments is in duality with the theory of nonnegative polynomials and Hilbert's 17th problem on the representation of nonnegative polynomials as sum of squares (always possible in the one-dimensional case). However, the multivariate case radically differs from the univariate case, for not every nonnegative polynomial can be written as a sum of squares. Also, in contrast to the univariate case, with $M_m(y)$ the moment matrix previously introduced (in lieu of the Hankel matrix (2.7)), there are vectors y for which $M_m(y) \succ 0$ but with *no* representing measure μ_y.

3. Unconstrained global optimization. Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be a real-valued polynomial of degree $2m$ with coefficient vector $p \in \mathbb{R}^{s(2m)}$. Since we wish to minimize $p(x)$, we may and will assume that the constant term vanishes, that is, $p_0 = 0$. Let us introduce the following convex LMI optimization problem (or positive semidefinite (psd) program):

$$\mathbb{Q} \mapsto \begin{cases} \inf_y \sum_\alpha p_\alpha y_\alpha, \\ \qquad M_m(y) \succeq 0, \end{cases} \tag{3.1}$$

or equivalently,

$$\mathbb{Q} \mapsto \begin{cases} \inf_y \sum_\alpha p_\alpha y_\alpha, \\ \qquad \sum_{\alpha\neq 0} y_\alpha B_\alpha \succeq -B_0, \end{cases} \tag{3.2}$$

where the matrices B_0 and B_α are easily understood from the definition of $M_m(y)$. The dual problem $\mathbb{Q}^*$ of $\mathbb{Q}$ is the convex LMI problem defined by

$$\mathbb{Q}^* \mapsto \begin{cases} \sup_X \langle X, -B_0\rangle \,(= -X(1,1)), \\ \qquad \langle X, B_\alpha\rangle = p_\alpha, \qquad \alpha \neq 0, \\ \qquad X \succeq 0, \end{cases} \tag{3.3}$$

where X is a real-valued symmetric matrix and $\langle A, B\rangle$ stands for the usual Frobenius inner product trace(AB) for real-valued symmetric matrices. The reader is referred to Vandenberghe and Boyd [20] for a survey on semidefinite programming.

We first have the following result.

PROPOSITION 3.1. *Assume that $\mathbb{Q}^*$ has a feasible solution. Then $\mathbb{Q}^*$ is solvable and there is no duality gap, that is,*

$$\inf \mathbb{Q} \;=\; \max \mathbb{Q}^*. \tag{3.4}$$

Proof. The result follows from the duality theory of convex programming if we can prove that there is a feasible solution y of $\mathbb{Q}$ with $M_m(y) \succ 0$. Let μ be a probability measure on $\mathbb{R}^n$ with a *strictly positive* density f with respect to the Lebesgue measure and with all its moments finite; that is, μ is such that

$$y_\alpha \;:=\; \int x^\alpha \, d\mu \;<\; \infty$$

for every combination $\alpha_1 + \alpha_2 + \alpha_n = r$, $r = 1, 2, \ldots$. Then the matrix $M_m(y)$, with y as above, is such that $M_m(y) \succ 0$. Indeed, for every polynomial $q(x) : \mathbb{R}^m \to \mathbb{R}$, we have

$$\begin{aligned} \langle q(x), q(x)\rangle_y \;=\; \langle q, M_m(y)q\rangle &= \int q^2(x)\,\mu(dx) \;\; \text{(by (2.5))} \\ &= \int q(x)^2 f(x)\,dx \\ &> 0 \;\; \text{whenever } q \neq 0 \text{ (as } f > 0). \end{aligned}$$

Therefore, y is feasible for $\mathbb{Q}$ and $M_m(y) \succ 0$, the desired result. ∎

Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be a real-valued polynomial with $p_0 := p(0) = 0$. The first result of this paper is as follows.

THEOREM 3.2. *Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be a $2m$-degree polynomial as in* (2.2) *with global minimum $p^* = \min \mathbb{P}$.*

(i) *If the nonnegative polynomial $p(x) - p^*$ is a sum of squares of polynomials, then $\mathbb{P}$ is equivalent to the convex LMI problem $\mathbb{Q}$ defined in* (3.1). *More precisely,* $\min \mathbb{Q} = \min \mathbb{P}$ *and, if x^* is a global minimizer of $\mathbb{P}$, then the vector*

$$y^* \;:=\; (x_1^*, \ldots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \ldots, (x_1^*)^{2m}, \ldots, (x_n^*)^{2m}) \tag{3.5}$$

is a minimizer of $\mathbb{Q}$.

(ii) *Conversely, if $\mathbb{Q}^*$ has a feasible solution, then* $\min \mathbb{P} = \min \mathbb{Q}$ *only if $p(x) - p^*$ is a sum of squares.*

Proof. (i) Let $p(x) - p^*$ be a sum of squares of polynomials, that is,

$$p(x) - p^* \;=\; \sum_{i=1}^{r} q_i(x)^2, \qquad x \in \mathbb{R}^n, \tag{3.6}$$

for some polynomials $q_i(x) : \mathbb{R}^n \to \mathbb{R}$, with coefficient vector $q_i \in \mathbb{R}^{s(m)}$, $i = 1, 2, \ldots, r$. Equivalently,

$$p(x) - p^* \;=\; \langle X, M_m(y)\rangle, \qquad x \in \mathbb{R}^n, \tag{3.7}$$

with $X = \sum_1^r q_i q_i'$ and $y = (x_1, \ldots, (x_1)^{2m}, \ldots, (x_n)^{2m})$. But from (3.7) and

$$p(x) - p^* \;=\; -p^* + \sum_\alpha p_\alpha x^\alpha,$$

it follows that

$$X(1,1) \;=\; -p^* \quad \text{and } \langle X, B_\alpha \rangle \;=\; p_\alpha \quad \text{for all } \alpha \neq 0$$

so that (as $X \succeq 0$) X is feasible for $\mathbb{Q}^*$ with value $-X(1,1) = p^*$. Next, observe that y^* in (3.5) is feasible for $\mathbb{Q}$ with value p^* so that $\min \mathbb{Q} = \max \mathbb{Q}^*$ and y^* and X are optimal solutions of $\mathbb{Q}$ and $\mathbb{Q}^*$, respectively.

(ii) Assume that $\mathbb{Q}^*$ has a feasible solution and $\min \mathbb{P} = \min \mathbb{Q}$. Then, from Proposition 3.1, $\mathbb{Q}^*$ is solvable and there is no duality gap, that is, $\max \mathbb{Q}^* = \inf \mathbb{Q} = \min \mathbb{Q}$. Let X^* be an optimal solution of $\mathbb{Q}^*$, guaranteed to exist. Write $X^* = \sum_{i=1}^r \lambda_i q_i q_i'$ with the q_i's being the eigenvectors of X^* corresponding to the positive eigenvalues λ_i, $i = 1, \ldots, r$.

As $\lambda^* := \max \mathbb{Q}^* = \min \mathbb{Q}$, and $\min \mathbb{Q} = \min \mathbb{P}$, let y^* as in (3.5) be an optimal solution of $\mathbb{Q}$. From the optimality of both X^* and y^*, we must have

$$\langle X^*, M_m(y^*) \rangle \;=\; 0.$$

Equivalently,

$$0 \;=\; \sum_{i=1}^r \lambda_i \langle q_i, M_m(y^*) q_i \rangle \;=\; \sum_{i=1}^r \lambda_i q_i(x^*)^2.$$

For an arbitrary $x \in \mathbb{R}^n$, let

$$y \;:=\; (x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_1^{2m}, \ldots, x_n^{2m})$$

so that, as we did for x^*,

$$\langle X^*, M_m(y) \rangle \;=\; \sum_{i=1}^r \lambda_i q_i(x)^2.$$

On the other hand,

$$\begin{aligned} \langle X^*, M_m(y) \rangle &= \lambda^* + \sum_{\alpha \neq 0} y_\alpha \langle X^*, B_\alpha \rangle \\ &= \lambda^* + \sum_{\alpha \neq 0} p_\alpha y_\alpha \;=\; \lambda^* + p(x). \end{aligned}$$

Therefore, as X^* is optimal, $-X^*(1,1) = -\lambda^* = p^*$, and we obtain

$$\sum_{i=1}^r \lambda_i q_i(x)^2 \;=\; p(x) - p^*,$$

the desired result. $\square$

From the proof of Theorem 3.2, it is obvious that if $\min \mathbb{Q} = \min \mathbb{P}$, then x^* is a root of each polynomial $q_i(x)$, where $X^* = \sum_{i=1}^r q_i q_i'$ at an optimal solution X^* of

$\mathbb{Q}^*$. When $p(x) - p^*$ is a sum of squares, solving the dual LMI problem $\mathbb{Q}^*$ provides the q_i polynomials of such a decomposition. As a corollary, we obtain the following.

Corollary 3.3. *Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be a real-valued polynomial of degree $2m$. Assume that $\mathbb{Q}^*$ has a feasible solution. Then,*

$$p(x) - p^* = \sum_{i=1}^{r} q_i(x)^2 - [\min \mathbb{P} - \inf \mathbb{Q}] \tag{3.8}$$

for some real-valued polynomials $q_i(x) : \mathbb{R}^n \to \mathbb{R}$ of degree at most m, $i = 1, 2, \ldots, r$.

The proof is the same as the proof of Theorem 3.2(ii), except now we may not have $\min \mathbb{Q} = \min \mathbb{P}$, but instead $\inf \mathbb{Q} \leq \min \mathbb{P}$. Hence, $\inf \mathbb{Q}$ always provides a lower bound on p^*.

Corollary 3.3 states that one may always write $p(x) - p^*$ as a sum of squares of polynomials up to some constant whenever $\mathbb{Q}^*$ has a feasible solution.

One may ask whether a nonnegative polynomial can be "approached" by polynomials that are the sum of squares. The answer is yes (see Remark 3.6 below).

Example 1. Consider the polynomial $p(x) : \mathbb{R}^2 \to \mathbb{R}$

$$(x_1, x_2) \mapsto (x_1^2 + 1)^2 + (x_2^2 + 1)^2 + (x_1 + x_2 + 1)^2.$$

It is not obvious a priori that with x^* a global minimizer, $p(x) - p^*$ is a sum of squares. Solving $\mathbb{Q}$ yields a minimum value of -0.4926, and from the solution y, one may check that

$$y = (x_1^*, x_2^*, (x_1^*)^2, x_1^* x_2^*, (x_2^*)^2, \ldots, (x_1^*)^4, \ldots, (x_2^*)^4),$$

with $x_1^* = x_2^* = -0.2428$, is a good approximation of a global minimizer of $\mathbb{P}$ since the gradient vector

$$\frac{\partial p(x_1^*,{}_2^*)}{\partial x_1} = \frac{\partial p(x_1^*, x_2^*)}{\partial x_2} = 4 * 10^{-9}.$$

Solving $\mathbb{Q}^*$ yields

$$X^* \approx \begin{bmatrix} 0.4926 & 1.0000 & 1.0000 & -0.0196 & -0.0316 & -0.0668 \\ 1.0000 & 3.0392 & 1.0316 & 0 & -0.0276 & -0.1666 \\ 1.0000 & 1.0316 & 3.1335 & 0.0276 & 0.1666 & 0 \\ -0.0196 & 0 & 0.0276 & 1.0000 & 0 & -0.5539 \\ -0.0316 & -0.0276 & 0.1666 & 0 & 1.1078 & 0 \\ -0.0668 & -0.1666 & 0 & -0.5539 & 0 & 1.0000 \end{bmatrix}$$

with eigenvalues

$$[1.0899,\ 1.5414,\ 2.0885,\ 0.4410,\ 0.0000,\ 4.6123]$$

and corresponding eigenvectors

$$\begin{bmatrix} 0.0579 & 0.0144 & 0.0163 & 0.0675 & 0.9414 & -0.3246 \\ -0.0972 & 0.1118 & 0.6999 & -0.0759 & -0.2286 & -0.6559 \\ 0.1010 & -0.0657 & -0.6861 & 0.0114 & -0.2286 & -0.6800 \\ 0.0224 & -0.7105 & 0.0503 & -0.6993 & 0.0555 & -0.0092 \\ -0.9882 & -0.0334 & -0.1368 & -0.0028 & 0.0555 & -0.0242 \\ -0.0006 & 0.6907 & -0.1337 & -0.7075 & 0.0555 & 0.0377 \end{bmatrix}.$$

Example 2. Consider the polynomial $p(x):\mathbb{R}^2\to\mathbb{R}$

$$(x_1,x_2)\;\mapsto\;(x_1^2+1)^2+(x_2^2+1)^2-2(x_1+x_2+1)^2.$$

Solving $\mathbb{Q}$ yields an optimal value $p^*\approx -11.4581$ and an optimal solution

$$(x_1^*,x_2^*)\;=\;(1.3247,1.3247)$$

with corresponding gradient

$$\frac{\partial p(x_1^*,x_2^*)}{\partial x_1}\;=\;\frac{\partial p(x_1^*,x_2^*)}{\partial x_2}\;=\;-8.7*10^{-6}.$$

Solving $\mathbb{Q}^*$ yields an optimal solution

$$X^*\;\approx\;\begin{bmatrix} 11.4581 & -2.0000 & -2.0000 & -1.0582 & -1.1539 & -1.2977\\ -2.0000 & 2.1164 & -0.8461 & 0 & -0.0868 & 0.2676\\ -2.0000 & -0.8461 & 2.5953 & 0.0868 & -0.2676 & 0\\ -1.0582 & 0 & 0.0868 & 1.0000 & 0 & -0.4625\\ -1.1539 & -0.0868 & -0.2676 & 0 & 0.9250 & 0\\ -1.2977 & 0.2676 & 0 & -0.4625 & 0 & 1.0000 \end{bmatrix}.$$

The eigenvalues of X^* are

$$[1.2719,\;1.4719,\;0.5593,\;0.0000,\;3.2582,\;12.5336]$$

with corresponding eigenvectors (in columns below)

$$\begin{bmatrix} 0.0854 & -0.0552 & -0.0615 & 0.2697 & 0.0177 & -0.9554\\ 0.5477 & -0.1658 & -0.3615 & 0.3573 & -0.6204 & 0.1712\\ 0.3274 & -0.2171 & -0.2965 & 0.3573 & 0.7740 & 0.1760\\ 0.2384 & 0.6906 & 0.4831 & 0.4733 & 0.0403 & 0.0847\\ -0.6736 & 0.2490 & -0.4967 & 0.4733 & -0.0744 & 0.0896\\ -0.2740 & -0.6191 & 0.5454 & 0.4733 & -0.0919 & 0.1081 \end{bmatrix}.$$

Hence,

$$\begin{aligned}
&p(x_1,x_2)-p^*\\
\approx&1.2719(0.0854+0.5477x_1+0.3274x_2+0.2384x_1^2-0.6736x_1x_2-0.2740x_2^2)^2\\
+&1.4719(-0.0552-0.1658x_1-0.2171x_2+0.6906x_1^2+0.2490x_1x_2-0.6191x_2^2)^2\\
+&0.5593(-0.0615-0.3615x_1-0.2965x_2+0.4831x_1^2-0.4967x_1x_2+0.5454x_2^2)^2\\
+&3.2582(0.0177-0.6204x_1+0.7740x_2+0.0403x_1^2-0.0744x_1x_2-0.0919x_2^2)^2\\
+&12.5336(-0.9554+0.1712x_1+0.1760x_2+0.0847x_1^2+0.0896x_1x_2+0.1081x_2^2)^2.
\end{aligned}$$

General case. We now provide a result valid in the general case, that is, when $p(x)-p^*$ is not necessarily a sum of squares.

We first need to introduce some notation: Let $q(x):\mathbb{R}^n\to\mathbb{R}$ be a real-valued polynomial of degree w with coefficient vector $q\in\mathbb{R}^{s(w)}$.

If the entry (i,j) of the matrix $M_m(y)$ is y_β, let $\beta(ij)$ denote the subscript β of y_β. Let $M_m(qy)$ be the matrix defined by

$$M_m(qy)(i,j)\;=\;\sum_\alpha q_\alpha y_{\{\beta(i,j)+\alpha\}}. \tag{3.9}$$

For instance, with

$$M_1(y) = \begin{bmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix} \quad \text{and } x \mapsto q(x) = a - x_1^2 - x_2^2,$$

we obtain

$$M_1(qy) = \begin{bmatrix} a - y_{20} - y_{02}, & ay_{10} - y_{30} - y_{12}, & ay_{01} - y_{21} - y_{03} \\ ay_{10} - y_{30} - y_{12}, & ay_{20} - y_{40} - y_{22}, & ay_{11} - y_{31} - y_{13} \\ ay_{01} - y_{21} - y_{03}, & ay_{11} - y_{31} - y_{13}, & ay_{02} - y_{22} - y_{04} \end{bmatrix}.$$

Let $\{y_\alpha\}$ (with $y_0 = 1$) be an $s(2m)$-vector of moments up to order $2m$ of some probability measure μ_y on $\mathbb{R}^n$. Then, for every polynomial $v(x) : \mathbb{R}^n \to \mathbb{R}$, of degree at most m, with coefficient vector $v \in \mathbb{R}^{s(m)}$,

$$(3.10) \qquad \langle v, M_m(qy)v\rangle = \int q(x)v(x)^2 \, \mu_y(dx).$$

Therefore, with $K_q := \{x \in \mathbb{R}^n \,|\, q(x) \geq 0\}$, if μ_y has its support contained in K_q, then it follows from (3.10) that $M_m(qy) \succeq 0$.

Suppose that we know in advance that a global minimizer x^* of $p(x)$ has norm less than a for some $a > 0$, that is, $p(x^*) = p^* = \min \mathbb{P}$ and $\|x^*\| \leq a$. Then, with $x \mapsto \theta(x) = a^2 - \|x\|^2$, we have $p(x) - p^* \geq 0$ on $K_a := \{\theta(x) \geq 0\}$.

We will use the fact that every polynomial $p(x)$, strictly positive on K_a, can be written

$$p(x) = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2,$$

for some polynomials $q_i(x), t_j(x)$, $i = 1, \ldots, r_1$, $j = 1, \ldots, r_2$ (see, e.g., Berg [1, p. 119]). For every $N \geq m$, let $\mathbb{Q}_a^N$ be the convex LMI problem

$$(3.11) \qquad \mathbb{Q}_a^N \begin{cases} \inf\limits_y \displaystyle\sum_\alpha p_\alpha y_\alpha, & \\ \qquad M_N(y) \succeq 0, & \\ \quad M_{N-1}(\theta y) \succeq 0. & \end{cases}$$

Writing $M_{N-1}(\theta y) = \sum_\alpha y_\alpha C_\alpha$, for appropriate matrices $\{C_\alpha\}$, the dual of $\mathbb{Q}_a^N$ is the convex LMI problem

$$(3.12) \qquad (\mathbb{Q}_a^N)^* \begin{cases} \sup\limits_{X,Z \succeq 0} -X(1,1) - a^2 Z(1,1), & \\ \qquad \langle X, B_\alpha\rangle + \langle Z, C_\alpha\rangle = p_\alpha, \ \alpha \neq 0. & \end{cases}$$

Now we have the following theorem.

THEOREM 3.4. *Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be a $2m$-degree polynomial as in* (2.2) *with global minimum $p^* = \min \mathbb{P}$ and such that $\|x^*\| \leq a$ for some $a > 0$ at some global minimizer x^*. Then*

(a) *as $N \to \infty$, one has*

$$(3.13) \qquad \inf \mathbb{Q}_a^N \uparrow p^*.$$

Moreover, for N sufficiently large, there is no duality gap between $\mathbb{Q}_a^N$ and its dual $(\mathbb{Q}_a^N)^*$, and $(\mathbb{Q}_a^N)^*$ is solvable.

(b) $\min \mathbb{Q}_a^N = p^*$ *if and only if*

$$p(x) - p^* \;=\; \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2 \tag{3.14}$$

for some polynomials $q_i(x)$, $i = 1, \dots, r_1$, of degree at most N, and some polynomials $t_j(x)$, $j = 1, \dots, r_2$, of degree at most $N-1$. In this case, the vector

$$y^* \;:=\; (x_1^*, \dots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \dots, (x_1^*)^{2N}, \dots, (x_n^*)^{2N}) \tag{3.15}$$

is a minimizer of $\mathbb{Q}_a^N$. In addition, $\max(\mathbb{Q}_a^N)^* = \min \mathbb{Q}_a^N$ and for every optimal solution (X^*, Z^*) of $(\mathbb{Q}_a^N)^*$,

$$p(x) - p^* \;=\; \sum_{i=1}^{r_1} \lambda_i q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} \gamma_j t_j(x)^2, \tag{3.16}$$

where the vectors of coefficients of the polynomials $q_i(x)$,$t_j(x)$ are the eigenvectors of X^* and Z^* with respective eigenvalues λ_i, γ_j.

Proof. (a) From $x^* \in K_a$, and with

$$y^* \;:=\; (x_1^*, \dots, (x_1^*)^{2N}, \dots, (x_n^*)^{2N}),$$

it follows that $M_N(y^*), M_{N-1}(\theta y^*) \succeq 0$ so that y^* is admissible for $\mathbb{Q}_a^N$ and thus $\inf \mathbb{Q}_a^N \le p^*$.

Now, fix $\epsilon > 0$ arbitrary. Then, $p(x) - (p^* - \epsilon) > 0$ and, therefore, there is some N_0 such that

$$p(x) - p^* + \epsilon \;=\; \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2$$

for some polynomials $q_i(x)$, $i = 1, \dots, r_1$, of degree at most N_0, and some polynomials $t_j(x)$, $j = 1, \dots, r_2$, of degree at most $N_0 - 1$ (see Berg [1, p. 119]).

Let $q_i \in \mathbb{R}^{s(N_0)}, t_j \in \mathbb{R}^{s(N_0-1)}$ be the vector of coefficients of the polynomials $q_i(x), t_j(x)$, respectively, and let

$$X \;:=\; \sum_{i=1}^{r_1} q_i q_i', \;\; Z \;:=\; \sum_{j=1}^{r_2} t_j t_j'$$

so that $X, Z \succeq 0$. It is immediate to check that (X, Z) is admissible for $(\mathbb{Q}_a^{N_0})^*$ with value $-X(1,1) - a^2 Z(1,1) = (p^* - \epsilon)$. From weak duality it follows that $\inf \mathbb{Q}_a^{N_0} \ge -(X(1,1) + a^2 Z(1,1)) = p^* - \epsilon$, and the desired result follows from

$$p^* - \epsilon \;\le\; \inf \mathbb{Q}_a^{N_0} \;\le\; p^*.$$

We next prove that there is no duality gap between $\mathbb{Q}_a^N$ and its dual $(\mathbb{Q}_a^N)^*$ as soon as $N \ge N_0$. Indeed, let μ be a probability measure with uniform distribution in K_a. Let $y_\mu = \{y_\alpha\}$ with

$$y_\alpha \;:=\; \int x^\alpha \mu(dx)$$

for all combinations $(\alpha_1, \ldots, \alpha_n) = r$, $r = 1, \ldots, N$. All the y_α's are well defined since μ has its support contained in the compact set K_a. From (2.4),

$$\langle q, M_N(y_\mu)q\rangle \;=\; \int q(x)^2\mu(dx) \;>\; 0 \text{ whenever } 0 \neq q \in \mathbb{R}^{s(N)},$$

and from (3.10),

$$\langle q, M_{N-1}(\theta y_\mu)q\rangle \;=\; \int \theta(x)q(x)^2\mu(dx) \;>\; 0 \text{ whenever } 0 \neq q \in \mathbb{R}^{s(N-1)}.$$

It follows that $M_N(y_\mu), M_{N-1}(\theta y_\mu) \succ 0$; that is, y_μ is (strictly) admissible for $\mathbb{Q}_a^N$ and, as $(\mathbb{Q}_a^N)^*$ has an admissible solution, from a standard result in convex optimization, there is no duality gap between $(\mathbb{Q}_a^N)^*$ and $\mathbb{Q}_a^N$. In addition, $(\mathbb{Q}_a^N)^*$ is solvable, that is, $\sup(\mathbb{Q}_a^N)^* = \max(\mathbb{Q}_a^N)^*$.

That $\inf \mathbb{Q}_a^N \uparrow p^*$ follows from the fact that, obviously, $M_N(y) \succeq 0$ implies $M_{N'}(y) \succeq 0$ for every $N \geq N'$ (since $M_{N'}(y)$ is a submatrix of $M_N(y)$) and similarly for $M_{N-1}(\theta y)$. Therefore, for every solution y of $\mathbb{Q}_a^N$, the adequate truncated vector y' is admissible for $\mathbb{Q}_a^{N'}$, whenever $N' \leq N$, with the same value. Hence, $\inf \mathbb{Q}_a^N \geq \inf \mathbb{Q}_a^{N'}$ whenever $N \geq N'$.

(b) *Only if part.* That y^* in (3.15) is a minimizer of $\mathbb{Q}_a^N$ is obvious. From (a) we know that there is no duality gap between $\mathbb{Q}_a^N$ and $(\mathbb{Q}_a^N)^*$ for N sufficiently large, and $(\mathbb{Q}_a^N)^*$ is solvable. Therefore, for N sufficiently large, let (X^*, Z^*) be an optimal solution of $(\mathbb{Q}_a^N)^*$, guaranteed to exist.

As $X^* \succeq 0, Z^* \succeq 0$, write

$$X^* \;=\; \sum_{i=1}^{r_1} \lambda_i q_i q_i'; \; Z^* \;=\; \sum_{j=1}^{r_2} \gamma_j t_j t_j',$$

where the q_i's (respectively, the t_j's) are the eigenvectors of X^* (respectively, Z^*), with eigenvalues λ_i (respectively, γ_j). With

$$y = (x_1, \ldots, x_n, \ldots, (x_1)^{2N}, \ldots, (x_n)^{2N}),$$

we have

$$\begin{aligned}\langle X^*, M_N(y)\rangle + \langle Z^*, M_{N-1}(\theta y)\rangle &= X^*(1,1) + a^2 Z^*(1,1) \\ &\quad + \sum_{\alpha \neq 0} y_\alpha[\langle X^*, B_\alpha\rangle + \langle Z^*, C_\alpha\rangle] \\ &= X^*(1,1) + a^2 Z^*(1,1) + p(x) \\ &= p(x) - p^*,\end{aligned}$$

where the last equality follows from

$$\min \mathbb{Q}_a^N \;=\; p^* \;=\; \max(\mathbb{Q}_a^N)^* \;=\; -X^*(1,1) - a^2 Z^*(1,1).$$

On the other hand,

$$\langle X^*, M_N(y)\rangle \;=\; \sum_{i=1}^{r_1} \lambda_i \langle q_i, M_N(y) q_i\rangle \;=\; \sum_{i=1}^{r_1} \lambda_i q_i(x)^2$$

and

$$\langle Z^*, M_{N-1}(\theta y)\rangle \;=\; \sum_{j=1}^{r_2} \gamma_j \langle t_j, M_{N-1}(\theta y) t_j\rangle \;=\; \theta(x) \sum_{j=1}^{r_2} \gamma_j t_j(x)^2.$$

Therefore,

$$p(x) - p^* \;=\; \sum_{i=1}^{r_1} \lambda_i q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} \gamma_j t_j(x)^2,$$

the desired result.

If part. If (3.14) holds, then one proves as in (a) (but with $\epsilon = 0$) that $\sup(\mathbb{Q}_a^N)^* \geq p^*$ so that, in fact, $\max(\mathbb{Q}_a^N)^* = p^* = \min \mathbb{Q}_a^N$ for N sufficiently large. $\square$

Thus, one may approach the global optimal value p^* as closely as desired by solving a finite number of convex LMI problems $\mathbb{Q}_a^N$, and if $p(x) - p^*$ (which is only nonnegative and not strictly positive) can be written as a weighted sum of squares, one obtains the exact optimal value by solving a finite number of problems $\mathbb{Q}_a^N$. However, from a computational point of view, the remark is irrelevant, especially if one solves $\mathbb{Q}_a^N$ with an interior point method.

REMARK 3.5. *Theorem* 3.4 *also applies for the global minimization of $p(x)$ on K_a if K_a does not contain any global minimizer of $p(x)$ on $\mathbb{R}^n$. It suffices to replace p^* with $\delta^* := \min_{x \in K_a} p(x)$.*

Example 3. Consider the polynomial $p(x) : \mathbb{R}^2 \to \mathbb{R}$,

$$x \;\mapsto\; p(x) \;:=\; x_1^2 x_2^2 (x_1^2 + x_2^2 - 1).$$

$1 + p(x)$ is positive but is not a sum of squares (see Berg [1]). A global minimizer of $p(x)$ is $x_1^2 = x_2^2 = 1/3$ with optimal value $p^* = -1/27$.

Solving the LMI problem $\mathbb{Q}$ in Theorem 3.2 yields an approximated optimal value of $-33.157352 < p^*$. With K_1 (the unit ball), solving $\mathbb{Q}_1^3$, one obtains *exactly* the global minimum p^* and a global minimizer x^*. In fact, as $p(x)$ contains only even powers of x_1 and x_2, y^* is the convex combination of $0.5y_1^* + 0.5y_2^*$ with y_1^*, y_2^* being the sequences of moments corresponding to the Dirac measures at $x_1^* = -\sqrt{1/3}$ and at $x_2^* = \sqrt{1/3}$, respectively.

This shows that in some cases one will obtain the exact global optimal value with few trials. In the present example, $p(x)$ is of degree 6 and we do not need to increase the degree to get the weighted sum of squares (3.14) when it exists; that is, $q_i(x)^2, t_j(x)^2$ in (3.14) are of degree at most 6, as $p(x)$.

REMARK 3.6. *One may ask whether a nonnegative polynomial can be "approached" by polynomials that are sums of squares. An answer is given in Berg* [1]. *Indeed, let $\mathcal{A}$ be the space of real-valued polynomials $p(x) : \mathbb{R}^n \to \mathbb{R}$ equipped with the norm $\|p(x)\|_{\mathcal{A}} = \|p\|$ with p the (finite-dimensional) vector of the coefficients of $p(x)$ (for instance, in the (extended) basis* (2.1)). *Then, the cone Σ of polynomials that are sums of squares is dense (for the norm $\|.\|_{\mathcal{A}}$) in the set of polynomials that are nonnegative on $[-1, 1]^n$.*

For instance, as we know that $p(x) - p^*$ is positive in $[-1, 1]^2$, for the polynomial $p(x)$ in the above example we may try to solve the LMI problem $\mathbb{Q}$ with $M_4(y) \succeq 0$ instead of $M_3(y) \succeq 0$ and perturbate $p(x)$ by adding the terms $0.01(x_1^8 + x_2^8)$ whose effect in $[-1, 1]^2$ is negligible. Solving $\mathbb{Q}$ for $\tilde{p}(x) = p(x) + 0.01(x_1^8 + x_2^8)$ yields the optimal value $\tilde{p}^* = -0.036792$ to compare with -0.037037 and a global minimizer $(\tilde{x}_1^*)^2 = (\tilde{x}_2^*)^2 = 0.3319$ to compare with $(x_1^*)^2 = (x_2^*)^2 = 1/3$. In this case, $\tilde{p}(x) - \tilde{p}^*$ is a sum of squares. However, the smaller perturbation $0.001(x_1^8 + x_2^8)$ does not work.

4. Constrained case. We now consider the constrained case, that is,

$$\mathbb{P}_K \mapsto p_K^* := \min_{x\in K} p(x), \tag{4.1}$$

where

- $p(x) : \mathbb{R}^n \to \mathbb{R}$ is a real-valued polynomial of degree at most m.
- K is a *compact* set defined by polynomials inequalities $g_i(x) \geq 0$ with $g_i(x) : \mathbb{R}^n \to \mathbb{R}$ being a real-valued polynomial of degree at most w_i, $i = 1, 2, \ldots, r$.

Concerning the semi-algebraic compact set K, we make the following assumption.

ASSUMPTION 4.1. *The set K is compact and there exists a real-valued polynomial $u(x) : \mathbb{R}^n \to \mathbb{R}$ such that $\{u(x) \geq 0\}$ is compact, and*

$$u(x) = u_0(x) + \sum_{k=1}^{r} g_i(x)u_i(x) \quad \textit{for all } x \in \mathbb{R}^n, \tag{4.2}$$

where the polynomials $u_i(x)$ are all sums of squares, $i = 0, \ldots, r$.

Assumption 4.1 is satisfied in many cases, for instance, if there is one polynomial $g_i(x)$ such that $\{g_i(x) \geq 0\}$ is compact (take $u_k(x) \equiv 0$ except $u_i(x) \equiv 1$ in (4.2)). It is also satisfied if all the p_i's are linear (see Jacobi and Prestel [9]) and for 0-1 programs, that is, when K includes the inequalities $x_i^2 \geq x_i$ and $x_i \geq x_i^2$ for all i. Therefore, one way to ensure that Assumption 4.1 holds is to add to the definition of K the extra constraint $g_{r+1}(x) = a^2 - \|x\|^2 \geq 0$ for some a sufficiently large.

It is important to emphasize that we do *not* assume that K is convex (it may even be disconnected). We will use the fact that whenever Assumption 4.1 holds, every polynomial $p(x)$, strictly positive on K, can be written

$$p(x) = q(x) + \sum_{k=1}^{r} g_k(x)t_k(x) \quad \text{for all } x \in \mathbb{R}^n \tag{4.3}$$

for some polynomials $q(x)$, $t_k(x)$, $k = 1, \ldots, r$, that are all sums of squares (see, e.g., Lemma 4.1 in Putinar [15] and also Jacobi [8]). In fact, Assumption 4.1 is an *if and only if* condition for (4.3) to hold. Of course, one does not know in advance the degrees of these polynomials.

As we did for $\theta(x)$ in the previous section, for every $i = 1, \ldots, r$, let $M_m(g_i y)$ be the matrices defined as in (3.9), with $g_i(x)$ in lieu of $\theta(x)$. Therefore, if y is an $s(2m)$ moment vector for some probability measure μ on $\mathbb{R}^n$, then for every $i = 1, 2, \ldots, r$, and every polynomial $q(x)$ of degree at most m,

$$\langle q(x), q(x)\rangle_{g_i y} := \langle q, M_m(g_i y)q\rangle = \int g_i(x)q(x)^2\,\mu(dx) \tag{4.4}$$

so that, if μ has its support contained in K, then $M_m(g_i y) \succeq 0$ for all $i = 1, 2, \ldots, r$.

Let $\tilde{w}_i := \lceil w_i/2 \rceil$ be the smallest integer larger than $w_i/2$, and with $N \geq \lceil m/2 \rceil$ and $N \geq \max_i \tilde{w}_i$, consider the convex LMI problem

$$\mathbb{Q}_K^N \left\{ \begin{array}{rcl} \inf_y \sum_\alpha p_\alpha y_\alpha, & & \\ M_N(y) & \succeq & 0, \\ M_{N-\tilde{w}_i}(g_i y) & \succeq & 0, \quad i = 1, \ldots, r. \end{array} \right. \tag{4.5}$$

Writing $M_{N-\tilde{w}_i}(g_i y) = \sum_\alpha C_{i\alpha} y_\alpha$, for appropriate symmetric matrices $\{C_{i\alpha}\}$, the dual of $\mathbb{Q}_K^N$ is the convex LMI problem

$$
(4.6) \qquad (\mathbb{Q}_K^N)^* \begin{cases} \sup_{X, Z_i} \; -X(1,1) - \sum_{i=1}^r g_i(0) Z_i(1,1), & \\ \quad \langle X, B_\alpha\rangle + \sum_{i=1}^r \langle Z_i, C_{i\alpha}\rangle \;=\; p_\alpha, \; \alpha \neq 0, & \\ \quad X, Z_i \;\succeq\; 0, \;\; i = 1, \ldots, r. & \end{cases}
$$

Now we have the following theorem.

THEOREM 4.2. *Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be an m-degree polynomial and K be the compact set $\{g_i(x) \geq 0, \; i = 1, \ldots, r\}$. Let Assumption 4.1 hold, and let $p_K^* := \min_{x \in K} p(x)$. Then*

(a) *as $N \to \infty$, one has*

$$
(4.7) \qquad \inf \mathbb{Q}_K^N \uparrow p_K^*.
$$

Moreover, for N sufficiently large, there is no duality gap between $\mathbb{Q}_K^N$ and its dual $(\mathbb{Q}_K^N)^*$ if K has a nonempty interior.

(b) *if $p(x) - p_K^*$ has the representation (4.3), that is,*

$$
(4.8) \qquad p(x) - p_K^* \;=\; q(x) + \sum_{i=1}^r g_i(x) t_i(x)
$$

for some polynomial $q(x)$ of degree at most $2N$, and some polynomials $t_i(x)$ of degree at most $2N - w_i$, $i = 1, \ldots, r$, all sums of squares, then $\min \mathbb{Q}_K^N = p_K^* = \max \, (\mathbb{Q}_K^N)^*$ *and the vector*

$$
(4.9) \qquad y^* \;:=\; (x_1^*, \ldots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \ldots, (x_1^*)^{2N}, \ldots, (x_n^*)^{2N})
$$

is a global minimizer of $\mathbb{Q}_K^N$. In addition, for every optimal solution $(X^*, Z_1^*, \ldots, Z_r^*)$ of $(\mathbb{Q}_K^N)^*$,

$$
(4.10) \qquad p(x) - p_K^* \;=\; \sum_{i=1}^{r_0} \lambda_i q_i(x)^2 + \sum_{i=1}^r g_i(x) \sum_{j=1}^{r_i} \gamma_{ij} t_{ij}(x)^2,
$$

where the vectors of coefficients of the polynomials $q_i(x)$,$t_{ij}(x)$ are the eigenvectors of X^* and Z_i^* with respective eigenvalues λ_i, γ_{ij}.

Proof. The proof is similar to that of Theorem 3.4. For (a) it is immediate that $\inf \mathbb{Q}_K^N \leq p_K^*$ since the sequence of moments y^* constructed from a global minimizer x^* is admissible with value p_K^*. Also, as in Theorem 3.4, the sequence $\{\inf \mathbb{Q}_K^N\}$ is easily seen to be monotone nondecreasing in N. Moreover,

(i) given $\epsilon > 0$ arbitrary, the polynomial $p(x) - p_K^* + \epsilon$ is strictly positive on K and thus can be written as in (4.3) for some polynomial $q(x)$ of degree at most $2N$ and some polynomials $t_i(x)$, $i = 1, \ldots, r$, of degree at most $2N - w_i$, that are all sums of squares. As in Theorem 3.4, writing $q(x) = \sum_i q_i(x)^2$ and $t_i(x) = \sum_j t_{ij}(x)^2$, from the vector of coefficients q_i of $q_i(x)$ (and t_{ij} of $t_{ij}(x)$), one may construct matrices $X := \sum_i q_i q_i' \succeq 0$ and $Z_i := \sum_j t_{ij} t_{ij}' \succeq 0$, $i = 1, \ldots, r$, that are admissible for $(\mathbb{Q}_K^N)^*$, with value $-X(1,1) - \sum_i g_i(0) Z_i(1,1) = p_K^* - \epsilon$. Indeed, with

$$
y = (x_1, \ldots, (x_1)^{2N}, \ldots, (x_n)^{2N}),
$$

we obtain

$$\langle X, M_N(y)\rangle + \sum_{i=1}^{r} \langle Z_i, M_{N-\tilde{w}_i}(g_i y)\rangle \;=\; p(x) - p_K^* + \epsilon$$

so that, as x was arbitrary,

$$\langle X, B_\alpha\rangle + \sum_{i=1}^{r} \langle Z_i, C_{i\alpha}\rangle \;=\; p_\alpha \;\text{ for all } \alpha \neq 0,$$

and $X(1,1) + \sum_i g_i(0) Z_i(1,1) = -(p_K^* - \epsilon)$. Hence, $p_K^* - \epsilon \leq \sup(\mathbb{Q}_K^N)^* \leq \inf \mathbb{Q}_K^N \leq p_K^*$. As ϵ was arbitrary, (4.7) follows.

(ii) That there is no duality gap between $\mathbb{Q}_K^N$ and its dual $(\mathbb{Q}_K^N)^*$ follows from the fact that $\mathbb{Q}_K^N$ admits a strictly admissible solution. It suffices to consider a probability measure μ with uniform distribution on K. The vector y_μ of its moments up to order $2N$ is such that $M_N(y) \succ 0$ and $M_{N-\tilde{w}_i}(g_i y) \succ 0$. Therefore, as $(\mathbb{Q}_K^N)^*$ has a feasible solution, by a standard result in convexity, $\sup(\mathbb{Q}_K^N)^* = \max(\mathbb{Q}_K^N)^* = \inf \mathbb{Q}_K^N$.

The proof of (b) is also similar. If $p(x) - p_K^*$ has the representation (4.3), then from the polynomials $q(x)$ and $\{t_i(x)\}$, of degree at most $2N$ and $2N - w_i$, respectively, one may construct matrices $X, Z_i \succeq 0$, $i = 1, \ldots, r$, as in (a), such that $(X, Z_1, \ldots, Z_r)$ is an admissible solution for $(\mathbb{Q}_K^N)^*$, with value $-X(1,1) - \sum_i g_i(0) Z_i(1,1) = p_K^*$. From $p_K^* \leq \sup(\mathbb{Q}_K^N)^* \leq \inf \mathbb{Q}_K^N \leq p_K^*$, it follows immediately that $\max(\mathbb{Q}_K^N)^* = p_K^* = \min \mathbb{Q}_K^N$ and $(X, Z_1, \ldots, Z_k)$ is an optimal solution of $(\mathbb{Q}_K^N)^*$. The last statement is obtained in a similar fashion. □

One may also prove that if K has a nonempty interior, then (4.8) is also necessary for $\min Q_K^N = p_K^*$ to hold.

When K is compact and Assumption 4.1 does not hold, there is still a representation of polynomials, strictly positive on K (see Corollary 3 in Schmüdgen [16]). But, instead of being "linear" as in (4.3), there are product terms of the form $g_{i_1}(x) g_{i_2}(x) \ldots g_{i_l}(x)$ times a sum of squares of polynomials, with $i_1, \ldots, i_l \in \{1, \ldots, r\}$. It then suffices to include the corresponding constraints $M_m(g_{i_1} \ldots g_{i_l} y) \succeq 0$ in the LMI problem Q_K^N. However, the number of LMI constraints in $\mathbb{Q}_K^N$ grows exponentially with the number of constraints.

Example 4. Let $p(x) : \mathbb{R}^2 \to \mathbb{R}$ be the polynomial $x \mapsto p(x) := -a_1 x_1^2 - a_2 x_2^2$ and K be the compact set

$$K \;:=\; \{x \in \mathbb{R}^2 \,|\, x_1 + x_2 \leq b_1; a x_1 + y \leq b_2;\; x_1, x_2 \geq 0\}.$$

Whenever $a_i > 0$, $p(x)$ is concave so that we have a concave minimization problem and thus, some vertex of K is a global minimizer.

We have solved $\mathbb{Q}_K^2$ for several values of $a_i > 0, b_i$, $i = 1, 2$, and $a < 0$, each time providing a global minimizer exactly, so that

$$p(x) - p_K^* \;=\; q(x) + (b_1 - x_1 - x_2) t_1(x) + (b_2 - a x_1 - x_2) t_2(x) + x_1 t_3(x) + x_2 t_4(x)$$

for some 4-degree polynomial $q(x)$ and 2-degree polynomials $t_i(x)$, all sums of squares, $i = 1, \ldots, 4$.

Example 5. Let $p(x) : \mathbb{R}^2 \to \mathbb{R}$ be the concave polynomial $x \mapsto p(x) := -(x_1 - 1)^2 - (x_1 - x_2)^2 - (x_2 - 3)^2$ and

$$K := \{(x_1, x_2) \in \mathbb{R}^2 \,|\, 1 - (x_1 - 1)^2 \geq 0;\; 1 - (x_1 - x_2)^2 \geq 0;\; 1 - (x_2 - 3)^2 \geq 0\}.$$

The point $(1,2)$ is a global minimizer with optimal value -2. Solving $\mathbb{Q}_K^1$, that is, with $N=1$ and $\tilde{p}(x) = p(x) + 10$ (since we eliminate the constant term -10), yields an optimal value of 7 instead of the desired value 8. On the other hand, solving $\mathbb{Q}_K^2$ yields an optimal value 8.00017 and an approximate global minimizer $(1.0043, 2.0006)$ (the error 0.00017 is likely due to the use of an interior point method in the LMI toolbox of MATLAB). Hence, with polynomials of degree 4 instead of 2, one obtains a good approximation of the correct value. Observe that there exist $\lambda_i = 1 \geq 0$ such that

$$p(x) + 3 \;=\; 0 + \sum_{i=1}^{r} \lambda_i g_i(x),$$

but $p(x) - p_K^*$ $(= p(x) + 2)$ cannot be written that way.

Therefore, for the general nonconvex and quadratically constrained quadratic problem, $\mathbb{Q}_K^1$ may sometimes provide directly the exact global minimum, but in general a lower bound only (if $(\mathbb{Q}_K^1)^*$ has a feasible solution).

Solving some test problems. We have also solved the following test problems proposed in Floudas and Pardalos [6].

Problem 2.2 in [6].

$$\left\{\begin{array}{l} \min\limits_{x,y} p(x,y) \;:=\; c^T x - 0.5 x^T Q x + d^T y; \\ 6x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 \;\leq\; 6.5; \\ 10x_1 + 10x_3 + y \;\leq\; 20; \\ 0 \;\leq\; y;\; 0 \;\leq\; x_i \;\leq\; 1,\; i = 1, \ldots, 5 \end{array}\right.$$

with $Q := I$ and $c = [-10.5, -7.5, -3.5, -2.5, -1.5]$. The optimal value -213 is obtained at the $\mathbb{Q}_K^2$ relaxation.

Problem 2.6 in [6].

$$\left\{\begin{array}{l} \min\limits_{x} p(x) \;:=\; c^T x - 0.5 x^T Q x; \\ Ax \;\leq\; b; \\ 0 \;\leq\; x_i \leq 1,\; i = 1, \ldots, 10 \end{array}\right.$$

with A being the matrix

$$\left[\begin{array}{cccccccccc} -2 & -6 & -1 & 0 & -3 & -3 & -2 & -6 & -2 & -2 \\ 6 & -5 & 8 & -3 & 0 & 1 & 3 & 8 & 9 & -3 \\ -5 & 6 & 5 & 3 & 8 & -8 & 9 & 2 & 0 & -9 \\ 9 & 5 & 0 & -9 & 1 & -8 & 3 & -9 & -9 & -3 \\ -8 & 7 & -4 & -5 & -9 & 1 & -7 & -1 & 3 & -2 \end{array}\right],$$

$c = [48, 42, 48, 45, 44, 41, 47, 42, 45, 46]$, $b = [-4, 22, -6, -23, -12]$, and $Q = 100I$. The optimal value -39 is obtained at the $\mathbb{Q}_K^2$ relaxation.

Problem 2.9 in [6].

$$\left\{\begin{array}{l} \max\limits_{x} p(x) \;:=\; \sum_{i=1}^{9} x_i x_{i+1} + \sum_{i=1}^{8} x_i x_{i+2} + x_1 x_7 + x_1 x_9 + x_1 x_{10} + x_2 x_{10} + x_4 x_7; \\ \sum_{i=1}^{10} x_i \;=\; 1;\; x_i \;\geq\; 0,\; i = 1, \ldots, 10. \end{array}\right.$$

The optimal value 0.375 is obtained at the $\mathbb{Q}_K^2$ relaxation.

Problem 3.3 in [6].

$$\begin{cases} \min_x p(x) := -25(x_1-2)^2 - (x_2-2)^2 \\ \qquad\qquad -(x_3-1)^2 - (x_4-4)^2 - (x_5-1)^2 - (x_6-4)^2; \\ (x_3-3)^2 + x_4 \geq 4;\ (x_5-3)^2 + x_6 \geq 4; \\ x_1 - 3x_2 \leq 2;\ -x_1 + x_2 \leq 2; \\ x_1 + x_2 \leq 6;\ x_1 + x_2 \geq 2; \\ 1 \leq x_3 \leq 5;\ 0 \leq x_4 \leq 6; \\ 1 \leq x_5 \leq 5;\ 0 \leq x_6 \leq 10; \\ x_1, x_2, \geq 0. \end{cases}$$

The optimal value -310 is obtained at the $\mathbb{Q}_K^2$ relaxation.

Problem 3.4 in [6].

$$\begin{cases} \min_x p(x) := -2x_1 + x_2 - x_3; \\ x_1 + x_2 + x_3 \leq 4; \\ x_1 \leq 2;\ x_3 \leq 3;\ 3x_2 + x_3 \leq 6; \\ x_i \geq 0,\ i = 1,2,3; \\ x^T B^T B x - 2r^T B x + \|r\|^2 - 0.25\|b - v\|^2 \geq 0 \end{cases}$$

with $r = [1.5, -0.5, -5]$ and

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -2 & 1 & -1 \end{bmatrix};\ b = \begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix};\ v = \begin{bmatrix} 0 \\ -1 \\ -6 \end{bmatrix}.$$

The optimal value -4 is obtained at the $\mathbb{Q}_K^4$ relaxation, whereas $\inf \mathbb{Q}_K^3 = -4.0685$.

Problem 4.6 in [6].

$$\begin{cases} \min_x p(x) := -x_1 - x_2; \\ x_2 \leq 2x_1^4 - 8x_1^3 + 8x_1^2 + 2; \\ x_2 \leq 4x_1^4 - 32x_1^3 + 88x_1^2 - 96x_1 + 36; \\ 0 \leq x_1 \leq 3;\ 0 \leq x_2 \leq 4. \end{cases}$$

The feasible set K is almost disconnected. The $\mathbb{Q}_K^4$ relaxation provides the optimal value -5.5079, the best value known so far, and therefore proves its global optimality.

Problem 4.7 in [6].

$$\begin{cases} \min_x p(x) := -12x_1 - 7x_2 + x_2^2; \\ -2x_1^4 + 2 - x_2 = 0; \\ 0 \leq x_1 \leq 2;\ 0 \leq x_2 \leq 3. \end{cases}$$

The $\mathbb{Q}_K^5$ relaxation provides the optimal value -16.73889, the best known solution so far, and therefore proves its global optimality.

0-1 programming. It is also worth mentioning that constrained and unconstrained 0-1 programming problems can also be treated by solving convex LMI problems $\mathbb{Q}_K^N$ since the integral constraints $x_i \in \{0,1\}$ can be written $x_i^2 \geq x_i; x_i^2 \leq x_i$ for all $i = 1, \dots, n$. Therefore, the set

$$K_1 := \{x_i - x_i^2 \geq 0;\ x_i^2 - x_i \geq 0;\ i = 1, 2, \dots, n\} \tag{4.11}$$

(and its intersection with other additional polynomial constraints) is compact and Assumption 4.1 holds. However, there is no strictly admissible solution (or interior point). For illustration we have solved the elementary problem

$$\min\{-ax_1 - bx_2 \,|\, r - x_1 - cx_2 \geq 0; x_1, x_2 \geq 0;\ x_1, x_2 \in \{0,1\}\},$$

replacing the integrality constraints with $(x_1, x_2) \in K_1$, and K_1 as in (4.11).

Solving $\mathbb{Q}_K^2$ with $0 < a < b$, and several random values of c, yields the global optimal value in all cases and a global minimizer at one of the integral points $(0,1)$, $(1,0)$, and $(1,1)$ of K.

Also, the first experimental results on a sample of randomly generated MAX-CUT problems in $\mathbb{R}^n$ (that is, maximizing a quadratic form with no squared terms under the integrality constraints $x_i^2 = 1$ for all i) are encouraging. Indeed, the optimal value was obtained at the $\mathbb{Q}_K^2$ relaxation in all cases (see Lasserre [12]) for $n = 5$ and even $n = 10$.

5. Karush–Kuhn–Tucker global optimality conditions. In this section we still consider the problem $\mathbb{P}_K$ with a compact set K defined by polynomials inequalities $g_i(x) \geq 0$, $i = 1, \dots, r$.

PROPOSITION 5.1. *Let $p_K^* := \min \mathbb{P}_K$ and assume that $x^* \in K$ is a global minimizer. If $p(x) - p_K^*$ can be written*

$$p(x) - p_K^* = \sum_{i=1}^{r_0} q_i(x)^2 + \sum_{k=1}^{r} g_k(x) \sum_{j=1}^{r_k} t_{kj}(x)^2, \qquad x \in \mathbb{R}^n, \tag{5.1}$$

for some polynomials $q_i(x)$, $t_{kj}(x)$, $i = 1, \dots, r_0$, $k = 1, \dots, r$, $j = 1, \dots r_k$, then

$$0 = g_k(x^*) \left[\sum_{j=1}^{r_k} t_{kj}(x^*)^2\right], \qquad k = 1, \dots, r. \tag{5.2}$$

$$\nabla p(x^*) = \sum_{k=1}^{r} \nabla g_k(x^*) \left[\sum_{j=1}^{r_k} t_{kj}(x^*)^2\right]. \tag{5.3}$$

Moreover, if there exist associated Lagrange Karush–Kuhn–Tucker multipliers $\lambda^* \in (\mathbb{R}^r)^+$ and if the gradients $\nabla g_k(x^*)$ are linearly independent, then

$$\sum_{j=1}^{r_k} t_{kj}(x^*)^2 = \lambda_k^*, \qquad k = 1, \dots, r. \tag{5.4}$$

Proof. As x^* is a global minimizer of $\mathbb{P}_K$, it follows from $p(x^*) - p_K^* = 0$ and (5.1) that

$$0 = \sum_{i=1}^{r_0} q_i(x^*)^2 + \sum_{k=1}^{r} g_k(x^*) \sum_{j=1}^{r_k} t_{kj}(x^*)^2$$

so that

$$0 = q_i(x^*), \quad i = 1, \dots, r_0, \quad \text{and } 0 = g_k(x^*) \sum_{j=1}^{r_k} t_{kj}(x^*)^2, \quad k = 1, \dots, r.$$

Moreover, from (5.1) and in view of the above,

$$\begin{aligned}\nabla p(x^*) &= \sum_{k=1}^{r} \nabla g_k(x^*) \sum_{j=1}^{r_k} t_{kj}(x^*)^2 \\ &= \sum_{k=1}^{r} \lambda_k^* \nabla g_k(x^*)\end{aligned}$$

so that (5.4) follows from the linear independence of the $\nabla g_k(x^*)$. ∎

Hence, the representation (5.1) can be viewed as a global optimality condition of the Karush–Kuhn–Tucker type, where the multipliers are now nonnegative polynomials instead of nonnegative constants. In general, and in contrast to the usual (local) Karush–Kuhn–Tucker optimality conditions, the polynomial multiplier associated to a constraint $g_k(x) \geq 0$, nonactive at x^*, is *not* identically null, but vanishes at x^*.

If $p(x) - p_K^*$ cannot be written as (5.1), we still have that $p(x) - p_K^* + \epsilon$ can be written as (5.1) for every $\epsilon > 0$. Of course, the degrees of $q_i(x)$ and $t_{kj}(x)$ in (5.1) depend on ϵ, but we have

$$\lim_{\epsilon \to 0} \sum_{i=1}^{r_0(\epsilon)} q_i(x^*)^2 = 0 \quad \text{and} \quad \lim_{\epsilon \to 0} \sum_{j=1}^{r_k(\epsilon)} t_{kj}(x^*)^2 = 0$$

for every k such that $g_k(x^*) > 0$.

Convex quadratic programming. In the case where $p(x)$ is a convex quadratic polynomial and $g_k(x)$ are concave quadratic (or linear) polynomials, then, at a Karush–Kuhn–Tucker point (x^*, λ^*), and with the Lagrangian $L(x, \lambda^*) := p(x) - \sum_{k=1}^{r} \lambda_k^* g_k(x)$, we have

$$\begin{aligned}p(x) - p_K^* &= L(x, \lambda^*) - L(x^*, \lambda^*) + \sum_{k=1}^{r} \lambda_k^* g_k(x) \\ &= \frac{1}{2} \langle x - x^*, \nabla_{xx}^2 L(x^*, \lambda^*)(x - x^*) \rangle + \sum_{k=1}^{r} \lambda_k^* g_k(x) \\ &= \sum_{i=1}^{n} \alpha_i (\langle q_i, x - x^* \rangle)^2 + \sum_{k=1}^{r} \lambda_k^* g_k(x),\end{aligned}$$

where the q_i's are the eigenvectors of the psd form $\nabla_{xx}^2 L/2$ with respective eigenvalues α_i, $i = 1, \dots, n$.

In this case, $p(x) - p_K^*$ can be written as (5.1) with $r_k = 1$ and $t_k(x) \equiv \sqrt{\lambda_k^*}$, and

$$q_i(x) = \sqrt{\alpha_i} \langle q_i, x - x^* \rangle, \qquad i = 1, \dots, n.$$

That is, the polynomial $\sum_j t_{kj}(x)^2$ is just the constant λ_k^*. Therefore, we have the following theorem.

THEOREM 5.2. *Let $p(x) : \mathbb{R}^n \to \mathbb{R}$ be a convex quadratic polynomial and $K := \{g_i(x) \geq 0\}$ be a compact convex set defined by concave quadratic polynomials $g_i(x)$,*

$i = 1, \ldots, r$. Let x^* be a local (hence global) minimum of $\mathbb{P}_K$ with associated Karush–Kuhn–Tucker multipliers $\lambda^* \in (\mathbb{R}^r)^+$. Then,

$$y^* = (x_1^*, \ldots, x_n^*, (x_1^*)^2, \ldots, (x_n^*)^2)$$

is an optimal solution of the convex LMI problem

$$\mathbb{Q}_K^1 \begin{cases} \min_y \sum_\alpha p_\alpha y_\alpha, \\ \quad \sum_\alpha (g_i)_\alpha y_\alpha \geq -g_i(0), \quad i = 1, \ldots, r, \\ \quad M_1(y) \succeq 0 \end{cases}$$

and λ^* is an optimal solution of the dual LMI problem

$$(\mathbb{Q}_K^1)^* \begin{cases} \max_{X \succeq 0, \lambda \geq 0} - X(1,1) - \sum_{i=1}^r \lambda_i g_i(0), \\ \qquad \langle X, B_\alpha \rangle + \sum_{i=1}^r \lambda_i (g_i)_\alpha = p_\alpha, \quad \alpha \neq 0. \end{cases}$$

Hence $(\mathbb{Q}_K^1)^*$, which is the well-known Shor's relaxation for nonconvex quadratic programs, is also the natural dual problem of the general convex quadratically constrained quadratic program. In fact, Theorem 5.2 is also true in the more general case where $\nabla^2_{xx} L(x^*, \lambda^*) \succeq 0$, which may also happen at a global minimizer of some nonconvex quadratic programs. For instance, the particular nonconvex quadratic problems investigated in [4] reduce to solving the single LMI problem Q_K^1.

The difference between the convex and nonconvex cases is that $\mathbb{Q}_K^1$ provides an exact solution in the convex case, whereas one has to solve an (often finite) sequence of problems $\{Q_K^N\}$ in the nonconvex case.

In the case where $p(x), g_i(x)$ are all linear, then the standard linear programming problem $\min_x \{c'x | Ax \geq b\}$ is just $\mathbb{Q}_K^0$, with $K := \{Ax \geq b\}$.

6. Conclusion. We have shown that the constrained and unconstrained global optimization problem with polynomials has a natural sequence of convex LMI relaxations $\{Q_K^N\}$ whose optimal values converge to the optimal value p_K^*. In some cases, the exact optimal value and a global minimizer are obtained at a particular relaxation. When this happens, every optimal solution of the dual LMI problem provides the Karush–Kuhn–Tucker polynomials in the representation of the polynomial $p(x) - p_K^*$, nonnegative on K, the analogues of the scalar multipliers in the standard Karush–Kuhn–Tucker (local) optimality condition. Identifying classes of problems, for which the dimension of the LMI problem $\mathbb{Q}_K^N$ to solve is known in advance, is a topic of further research.

Acknowledgments. The author wishes to thank Prof. Jim Renegar and an anonymous referee for their fruitful suggestions that helped improve the paper.

REFERENCES

[1] C. Berg, *The multidimensional moment problem and semi-groups*, in Moments in Mathematics, H.J. Landau, ed., AMS, Providence, RI, 1980, pp. 110–124.

[2] R.E. Curto and L.A. Fialkow, *Recursiveness, positivity, and truncated moment problems*, Houston J. Math., 17 (1991), pp. 603–635.

[3] R.E. Curto, *Flat Extensions of Positive Moment Matrices: Recursively Generated Relations*, Mem. Amer. Math. Soc., 136 (1998), no. 648.
[4] P.-D. Tao and L.-T.-H. An, *Lagrangian stability and global optimality in nonconvex quadratic minimization over Euclidian balls and spheres*, J. Convex Anal., 2 (1995), pp. 263–276.
[5] C. Ferrier, *Hilbert's* 17*th problem and best dual bounds in quadratic minimization*, Cybernet. Systems Anal., 34 (1998), pp. 696–709.
[6] C.A. Floudas and P.M. Pardalos, *A Collection of Test Problems for Constrained Global Optimization Algorithms*, Lecture Notes in Comput. Sci. 455, Springer-Verlag, Berlin, 1990.
[7] J.B. Hiriart-Urruty, *Conditions for global optimality*, in Handbook of Global Optimization, Kluwer, Dordrecht, 1995, pp. 1–24.
[8] T. Jacobi, *A representation theorem for certain partially ordered commutative rings*, Math. Z., to appear.
[9] T. Jacobi and A. Prestel, *On Special Representations of Strictly Positive Polynomials*, Tech. report, Konstanz University, Konstanz, Germany, January 2000.
[10] M. Krein and A. Nudel'man, *The Markov Moment Problem and Extremal Problems*, Ideas and Problems of P.L. Čebyšev and A.A. Markov and Their Further Development, Transl. Math. Monographs 50, AMS, Providence, RI, 1977.
[11] I.J. Landau, *Moments in Mathematics*, AMS, Providence, RI, 1987.
[12] J.B. Lasserre, *Optimality Conditions and LMI Relaxations for* 0-1 *Programs*, Tech. report 2000099, LAAS, Toulouse, France, March 2000, submitted.
[13] Y. Nesterov, *Squared functional systems and optimization problems*, in High Performance Optimization, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., Kluwer, Dordrecht, 2000.
[14] M. Putinar and F.-H. Vasilescu, *Solving moment problems by dimensional extension*, Ann. Math., 149 (1999), pp. 1087–1107.
[15] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Ind. Univ. Math. J., 42 (1993), pp. 969–984.
[16] K. Schmüdgen, *The K-moment problem for compact semi-algebraic sets*, Math. Ann., 289 (1991), pp. 203–206.
[17] N.Z. Shor, *Quadratic optimization problems*, Soviet J. Comput. Systems Sci., 25 (1987), pp. 1–11.
[18] N.Z. Shor, *Nondifferentiable Optimization and Polynomial Problems*, Kluwer, Dordrecht, 1998.
[19] B. Simon, *The classical moment problem as a self-adjoint finite difference operator*, Adv. Math., 137 (1998), pp. 82–203.
[20] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.