MATHEMATICS OF OPERATIONS RESEARCH
Vol. 27, No. 2, May 2002, pp. 347–360
*Printed in U.S.A.*

# SEMIDEFINITE PROGRAMMING VS. LP RELAXATIONS FOR POLYNOMIAL PROGRAMMING

JEAN B. LASSERRE

We consider the global minimization of a multivariate polynomial on a semi-algebraic set $\Omega$ defined with polynomial inequalities. We then compare two hierarchies of relaxations, namely, LP relaxations based on products of the original constraints, in the spirit of the RLT procedure of Sherali and Adams (1990), and recent semidefinite programming (SDP) relaxations introduced by the author. The comparison is analyzed in light of recent results in real algebraic geometry on various representations of polynomials, positive on a compact semi-algebraic set.

**1. Introduction.** In recent years, semidefinite programming (SDP) and LP-based relaxations have become more and more popular for obtaining good lower bounds (or even an optimal solution) for global optimization problems with polynomials. For instance, the well-known Shor's (1987) SDP relaxation has provided good lower bounds for combinatorial problems, notably the MAX-CUT problem, for which the Goemans and Williamson (1995) algorithm yields an approximate solution with guaranteed performance. Also, the *lift-and-project* procedure of Lovász and Schrijver (1991) yields a hierarchy of SDP or LP-based relaxations for 0-1 linear programs, with finite convergence (see also Kojima and Tunçel 2000 for extensions).

Other LP-based relaxations have been proposed in the literature, particularly the so-called reformulation linearization technique (RLT) of Sherali and Adams (1990). (See also Sherali and Tuncbilek 1992, 1997.) The basic idea, elements of which appear in Adams and Sherali (1986) and Shor (1987), is to (i) multiply the original constraints by a family of polynomials (usually products of the original constraints), (ii) linearize in an augmented space (lifting) via introduction of additional variables, and (iii) solve the associated resulting LP program. Depending on the degree of the multiplying polynomials, one obtains a hierarchy of LP-based relaxations. For unconstrained (and a certain class of constrained) 0-1 polynomial programs, the sequence of relaxations converges in at most $n$ steps. For more general problems, the authors propose to include additional constraints and use these relaxations in a branch-and-bound algorithm; see, e.g., Sherali and Tuncbilek (1992, 1997) and also Audet et al. 2000. This methodology illustrates the old idea of using *valid inequalities* to help solving nonconvex problems.

More recently, Lasserre (2001a) has introduced a new hierarchy of SDP relaxations for general optimization problems with polynomials. The resulting sequence of optimal values converges asymptotically to the global optimum, and in many cases, the optimal value is obtained at some particular relaxation. For instance, for MAX-CUT problems, the second relaxation provided the global optimum in a sample of 50 randomly generated instances of MAX-CUT in $\mathbb{R}^{10}$ (see Lasserre 2000). We recently showed that for arbitrary nonlinear (constrained) 0-1 programs, these SDP relaxations have, in fact, *finite* convergence (see Lasserre 2001b).

Received February 27, 2001; revised August 6, 2001.
*MSC 2000 subject classification*. Primary: 90C22, 90C25.
*OR/MS subject classification.* Primary: Programming/nonlinear.
*Key words*. Global optimization, real algebraic geometry, semidefinite relaxations, LP relaxations.



0364-765X/02/2702/0347/$05.00
1526-5471 electronic ISSN, © 2002, INFORMS

In this paper, we compare the relative merits of LP relaxations based on linearizing valid inequalities formed with products of the original constraints (in the spirit of Sherali and Tuncbilek's 1992, 1997 RLT procedure) and the abovementioned SDP relaxations of Lasserre (2001a) for polynomial programming. To do this, we will consider the generic problem

$$\mathbb{P} \mapsto p^* := \min_{x \in \mathbb{R}^n} \{g_0(x) | g_i(x) \geq 0,\ i = 1, \ldots, m\}, \tag{1.1}$$

where $g_i : \mathbb{R}^n \to \mathbb{R}$ is a real-valued polynomial for all $i = 0, \ldots, m$.

We will see that there is a common and natural framework to analyze both relaxations, namely, in terms of the problem

$$\mathcal{P} \to \min_{\mu} \left\{ \int g_0(x)\, d\mu(x) | \mu(\mathbb{K}) = 1;\ \mu \geq 0 \right\} \tag{1.2}$$

(where $\mathbb{K}$ is the feasible set in (1.1)) that is easily seen to be equivalent to $\mathbb{P}$ (see Lasserre 2001a) and in terms of recent results in algebraic geometry on various representations of polynomials, positive on a compact semi-algebraic set.

It turns out that in both SDP and LP relaxations, the variables aim at representing the moments of the probability measure $\mu$ in (1.2). The constraints in the primal SDP and LP relaxations are specific *moment conditions* to ensure that the support of $\mu$ is contained in $\mathbb{K}$. In the SDP relaxations, they are stated in terms of positive semidefiniteness of appropriate matrices, whereas the linear inequalities in the LP relaxations are *Hausdorff-type* moment conditions (only necessary in general). Similarly, the respective duals of both relaxations have a simple interpretation in terms of the representation of polynomials, positive on the feasible set $\mathbb{K}$. While the duals of the SDP relaxations aim at representing the polynomial $g_0(x) - p^*$, nonnegative on $\mathbb{K}$, as a sum of $g_i$'s weighted by *sums of squares* of polynomials, we show that the duals of the LP relaxations aim at representing $g_0(x) - p^*$ as a sum of *products* of the $g_i$'s, weighted by nonnegative *scalars*. In the light of recent results in real algebraic geometry by Putinar (1993) and Jacobi and Prestel (2001), the former representation is far more general than the latter.

The univariate case (that is, when $n = 1$ and $\mathbb{K} = [a, b]$) deserves special attention. Indeed, Shor (1987) was the first to show that $\mathbb{P}$ reduces to a convex minimization problem that could be solved via interior-point methods. Later, Nesterov (2000) provided an LMI formulation of the cones of polynomials nonnegative on $\mathbb{R}$, $\mathbb{R}^+$, and on $[a, b]$ that could be used to solve $\mathbb{P}$ via a positive semidefinite program (SDP) (see also Lasserre 1999). Therefore, in the univariate case, a *single* SDP relaxation solves $\mathbb{P}$ whereas in general, only asymptotic convergence can hold for LP relaxations. In particular, for the latter relaxations, we show that the exact optimal value cannot be reached at a particular relaxation whenever there is a global minimizer in the interior of $[a, b]$. In addition, the LP relaxations are "ill conditioned," as they contain larger and larger binomial coefficients. Thus, in the univariate case, the (single) SDP relaxation clearly outperforms basic LP relaxations based on products of the original constraints. This is confirmed in the sample of problems considered in Sherali and Tuncbilek (1997), where even with additional constraints only a lower bound is obtained.

Of course, an attractive feature of LP relaxations is that it permits us to use powerful LP codes to solve large size problems, which is not (yet?) the case for SDP relaxations. However, we will see that the LP relaxations suffer several drawbacks, namely,

(a) The Hausdorff moment conditions are not numerically stable because of the binomial coefficients involved in the constraints, whereas no such coefficient appears in the SDP relaxations.

(b) In contrast to SDP relaxations, the asymptotic convergence of the LP relaxations is *not* guaranteed in general. However, we prove asymptotic convergence in the univariate case as well as in the multivariate case when $\mathbb{K}$ is a convex polytope with nonempty interior. To the best of our knowledge, this is a new result that we prove by invoking a result of Handelman (1988) in algebraic geometry. Incidentally, this result validates and provides a rationale for the old idea of using valid inequality constraints to help in solving nonconvex optimization problems.

(c) Even in the case of a convex polytope $\mathbb{K}$, the LP relaxations cannot be exact in general (for instance, as soon as there is a global minimizer $x^*$ in the interior of the feasible set or if there is some nonoptimal solution that saturates the same constraints as the global minimizer). In contrast, this is not a problem for the SDP relaxations, because the "polynomial multipliers" of the inactive constraints are not required to be identically null but *vanish* at $x^*$ (which is not possible for a scalar coefficient).

Both drawbacks (b) and (c) are illustrated on simple examples. Therefore, it seems that SDP relaxations are in principle superior to LP-based relaxations (this is already known for the SDP and LP lift-and-project procedures of Lovász and Schrijver 1991 for 0-1 linear programs). However, so far, the present status of SDP software packages excludes their utilization for large-size (or even medium-size) problems, whereas LP software packages can handle very large-size problems. Thus, while it seems that SDP relaxations will outperform LP relaxations for small-size problems $\mathbb{P}$, LP relaxations (with eventual additional constraints and associated with a branch-and-bound procedure in the general case, e.g., as in Sherali and Tuncbilek 1992, 1997 or Audet et al. 2000) are so far the only ones implementable for larger-size problems (up to the numerical stability issue).

Of course, there are alternatives to SDP and LP relaxations. For instance, the recent work by Sherali and Fraticelli (2000) tries to combine the power of LP solvers with the strength of SDP relaxations by translating SDP relationships into RLT types of valid inequalities to tighten the LP-based relaxations. In a different spirit, Burer and Monteiro (2001), and Vanderbei and Benson (2000), solve SDP relaxations as ordinary nonlinear programs with suitable nonlinear programming techniques.

We hope that this paper will stimulate further developments of more efficient solving procedures for large-size (or even moderate-size) SDPs and/or adhoc alternative techniques as in the above-mentioned recent works (Burer and Monteiro 2001, Vanderbei and Benson 2000).

**2. Global minimization of a univariate polynomial.** We first consider the univariate case, that is, the global minimization of a univariate polynomial $g_0(x)$: $\mathbb{R} \to \mathbb{R}$, on an interval $[a, b]$. Of course, one way to solve such a problem is to compute the finitely many real zeros of the polynomial $g_0'$ via an appropriate method and compare the values of $g_0$ at those points (as well as the values of $g_0$ at $a$ and $b$). However, here we compare the SDP and LP-based relaxations. Thus, consider the problem

$$\mathbb{P} \to p^* := \min_{x \in [a, b]} g_0(x), \tag{2.1}$$

where

$$g_0(x) := \sum_{k=0}^{s} (g_0)_k x^k,$$

if $s$ is the degree of $g_0$.

**2.1. SDP relaxations.** Given a vector $y \in \mathbb{R}^{2n+1}$, let $M_n(y)$, $B_n(y)$ be the Hankel matrices,

$$M_n(y) = \begin{bmatrix} 1 & y_1 & \cdot & y_n \\ y_1 & y_2 & \cdot & y_{n+1} \\ \cdot & \cdot & \cdot & \cdot \\ y_n & \cdot & \cdot & y_{2n} \end{bmatrix}; \qquad B_n(y) = \begin{bmatrix} y_1 & y_2 & \cdot & y_{n+1} \\ y_2 & y_3 & \cdot & y_{n+2} \\ \cdot & \cdot & \cdot & \cdot \\ y_{n+1} & \cdot & \cdot & y_{2n+1} \end{bmatrix}.$$

For convenience, and with no loss of generality, we may and will assume that the constant term $g_0(0)$ of the polynomial $g_0$ vanishes.

PROPOSITION 2.1. *Let $g_0(x) : \mathbb{R} \to \mathbb{R}$ be a univariate polynomial of odd degree $2n+1$ with $g_0(0) = 0$ and let $[a, b]$ be an interval of the real line. Then*

$$\min_{x \in [a,b]} g_0(x) = p^* = \begin{cases} \min_y \sum_{k=1}^{2n+1} (g_0)_k y_k, \\ s.t. \;\; bM_n(y) \succeq B_n(y) \succeq aM_n(y). \end{cases} \tag{2.2}$$

PROOF. Observe that from the equivalence of $\mathbb{P}$ and (1.1), the criterion $\int g_0(x)\,\mu(dx)$ is a linear form in the first $2n+1$ moments, that is,

$$\int g_0(x)\,\mu(dx) = \sum_{k=1}^{2n+1} (g_0)_k y_k.$$

Next, the sequence $\{1, y_1, \ldots, y_{2n+1}\}$ is a sequence of moments of some probability measure $\mu$ with support in $[a, b]$ if and only if

$$bM_n(y) \succeq B_n(y) \succeq aM_n(y) \tag{2.3}$$

(see, for instance, Curto and Fialkow 1991, Theorem 4.1, Remark 4.2). Therefore, $\mathbb{P}$ is equivalent to

$$\min_y \left\{ \sum_{k=1}^{2n+1} (g_0)_k y_k \;\middle|\; \text{(2.3) holds} \right\},$$

and the result follows. □

In addition, we also have

$$g_0(x) - p^* = (x-a)q_a(x)^2 + (b-x)q_b(x)^2 \tag{2.4}$$

for some polynomials $q_a(x)^2$, $q_b(x)^2$ of degree at most $n$. The coefficients of the polynomials $q_a, q_b$ in (2.4) are precisely optimal solutions of the dual SDP of (2.2) (see, e.g., Lasserre 2001a, b in a more general framework).

In the case where $g_0$ has even degree, then (after rescaling to obtain $a = -1$ and $b = 1$),

$$g_0(x) - p^* = q(x)^2 + (1-x^2)q_1(x)^2 \tag{2.5}$$

for some polynomials $q(x)$, $q_1(x)$ of degree, at most $n/2$.

Nesterov 2000 characterized the cone of polynomials nonnegative on $[a, b]$ and its dual to obtain (in a slightly different form) the constraints of the above SDP (see Nesterov 2000, Theorem 17.13). The case of polynomials with even degree can be treated in a similar manner using now Remark 4.4 in Curto and Fialkow (2000) or Nesterov (2000, Theorem 17.12). Thus $\mathbb{P}$, equivalent to the *single* SDP (2.2), is a hidden convex problem.

**2.2. LP relaxations.** We now consider LP relaxations in the spirit of the RLT procedure of Sherali and Tuncbilek (1992, 1997). To simplify the exposition, and after an affine transformation, one may and will assume that $[a, b] = [0, 1]$. We consider the $\delta$ LP relaxation obtained by linearizing the constraints

$$x^k(1-x)^m \geq 0, \qquad k, m = 0, 1, \ldots, \delta \tag{2.6}$$

in replacing each term $x^i$ with the new variable $y_i$. One then minimizes $\sum_{i=1}^{2n+1}(g_0)_i y_i$, subject to the (linearized) constraints (2.6). Obviously, the abovementioned constraints (2.6) contain the so-called *bound-factor product* constraints obtained from the linearization of the constraints

$$x^k(1-x)^{\delta-k} \geq 0, \qquad k = 0, \ldots, \delta \tag{2.7}$$

in the LP relaxation of the RLT procedure of Sherali and Tuncbilek (1992, 1997). (As proved in Sherali and Tuncbilek (1992), they imply all the bound-factor products of order less than $\delta$.)

The interpretation of these constraints is easy if one realizes that a probability measure $\mu$ has its support contained in $[0, 1]$ if and only if

$$\int_0^1 x^k(1-x)^m\,\mu(dx) \geq 0 \quad \forall\, k, m \geq 0. \tag{2.8}$$

The abovementioned conditions (2.8) are due to Hausdorff (and also Bernstein) (see Feller 1966, Shohat and Tamarkin 1943), and a sequence $\{y_j\}$ is a moment sequence if and only if $y$ satisfies the conditions

$$\sum_{j=0}^{m}(-1)^j\binom{m}{j}y_{k+j} \geq 0, \quad \forall\, k, m = 0, 1, \ldots, \tag{2.9}$$

obtained from (2.8) after "linearization." Thus, for a fixed $\delta$, and after linearization, the constraints (2.6) of the LP relaxation are a *finite* subset of the *infinitely many* necessary and sufficient Hausdorff moment conditions (2.9), and the variable $y_i$ is to be interpreted as the moment $\int x^i\,d\mu$ of some probability measure $\mu$. The conditions (2.9) on the $y_i$'s will ensure that $\mu$ has its support contained in $[0, 1]$. We therefore consider the LP relaxation

$$\mathbb{P}_\delta \to \rho_\delta := \min_y\left\{\sum_i (g_0)_i y_i \mid \text{s.t. } (2.9),\ 0 \leq k+m \leq \delta\right\}. \tag{2.10}$$

We still assume that the constant term $g_0(0) = 0$.

PROPOSITION 2.2. *Consider the LP relaxation* $\mathbb{P}_\delta$ *in* (2.10). *Then, as* $\delta \to \infty$,

$$\rho_\delta := \min \mathbb{P}_\delta \uparrow p^* := \min_{x\in[0,1]} g_0(x). \tag{2.11}$$

PROOF. We obviously have $\rho_\delta \leq p^*$ for all $\delta$. Next, observe that the LP dual of $\mathbb{P}_\delta$ is the linear program

$$\mathbb{P}_\delta^* \to \begin{cases} \max\limits_{c_{km}\geq 0} -\sum\limits_{m=1}^{\delta} c_{0m}, \\ \sum\limits_{k\leq i;\, k+m\geq i} (-1)^{i-k}\binom{m}{i-k}c_{km} = (g_0)_i, \qquad i = 1, \ldots, \delta. \end{cases} \tag{2.12}$$

Let $\epsilon > 0$ be fixed and arbitrary. Then, $g_0(x) - p^* + \epsilon$ is strictly positive on $[0, 1]$ and, therefore, can be written as

$$g_0(x) - p^* + \epsilon = \sum_{0 \le k+m \le \delta(\epsilon)} c_{km} x^k (1-x)^m, \quad \forall x \in \mathbb{R} \tag{2.13}$$

for some integer $\delta(\epsilon)$ and some nonnegative coefficients $\{c_{km}\}$ (see Powers and Reznick 2000). By identifying terms of same power in both sides of (2.13), the $\{c_{km}\}$ must satisfy

$$\sum_{k \le i;\, k+m \ge i} (-1)^{i-k} \binom{m}{i-k} c_{km} = (g_0)_i, \qquad i = 1, \dots, \delta(\epsilon)$$

and for the constant term

$$\sum_{m=1}^{\delta(\epsilon)} c_{0m} = -p^* + \epsilon.$$

Thus, as soon as $\delta \ge \delta(\epsilon)$, $\{c_{km}\}$ is admissible for $\mathbb{P}_\delta^*$ with value $p^* - \epsilon$. As $\epsilon > 0$ was arbitrary, the result follows. $\square$

We have thus proved that the LP relaxations $\mathbb{P}_\delta$ yield lower bounds as close as desired to the optimal value $p^*$ if one lets $\delta \to \infty$. Although for each $\delta$, the LP relaxations $\mathbb{P}_\delta$ are stronger than those in the RLT procedure; the latter also converge because as $\delta \to \infty$, the constraints (2.7) match the constraints (2.9).

In both SDP and LP relaxations, the vector $y$ in (2.2) and (2.10) has the same interpretation as the *moment* vector of a probability measure $\mu$. The constraints (2.3) and (2.9) are different necessary and sufficient conditions for $\mu$ to be supported on $[a, b]$. Similarly, an optimal solution to the dual problem of each relaxation yields a different representation of the polynomial $g_0(x) - p^* + \epsilon$:

- as a sum of $x - a$ and $b - x$, weighted by *sums of squares* for SDP relaxations (see (2.4)), and
- as a sum of products $(x-a)^k(b-x)^m$, weighted by *nonnegative scalars* for LP relaxations (see (2.13) with $a = 0, b = 1$).

REMARK 2.3. (i) In general, the representation (2.13) holds for polynomials $p$, *strictly positive* on $[0, 1]$. Observe that if $g_0(x) - p^*$ has the representation (2.13), $\delta$ may be larger than $\deg(g_0)$. In addition, assume that there is a global minimizer $x^*$ in the interior of $[0, 1]$. Then, $g_0(x) - p^*$ *cannot have* the representation (2.13) with $\epsilon = 0$, for then taking $x^* \in (0, 1)$ yields $c_{km} = 0$ for all $k, m$. Therefore, in such a case, the LP relaxation can provide only a lower bound $\rho_\delta < p^*$, and therefore, only asymptotic convergence $\rho_\delta \uparrow p^*$ holds if one lets $\delta \to \infty$. This is why one may have to consider infinitely many constraints (2.9) in the LP relaxation $\mathbb{P}_\delta$, even if $g_0(x)$ is a low-degree polynomial. In addition, the constraints (2.9) are ill behaved in view of the binomial coefficients $\binom{m}{j}$, whereas no such coefficient appears in the Hankel matrices $M_n(y)$ and $B_n(y)$ in (2.2).

(ii) After a rescaling, we may have instead considered the minimization of a polynomial $g_0$ on $[-1, 1]$ (instead of $[0, 1]$). Consider the case where there is a global minimizer $x^* \in (-1, 1)$. For every $\epsilon > 0$, we have

$$g_0(x) - p^* + \epsilon = \sum_{0 \le k+m \le \delta(\epsilon)} c_{km}(\epsilon)(1+x)^k(1-x)^m \tag{2.14}$$

for some finite $\delta(\epsilon)$. However, then with $x := 0$,

$$g_0(0) - p^* + \epsilon = \sum_{k+m \le \delta(\epsilon)} c_{km}(\epsilon),$$

so that the sequence $\{c_{km}(\epsilon)\}$ is bounded, that is, when extended with zeros, $\{c_{km}(\epsilon)\} \in l_\infty^+$. Therefore, let $\epsilon_i \downarrow 0$. By a standard diagonal argument, consider a (pointwise) converging subsequence $\{c_{km}(\epsilon_{i_n})\} \to \{c^*_{km}\}$ in $l_\infty$. Fix $x \in (-1, 1)$ as arbitrary and consider (2.14). By Fatou's lemma, we must have

$$g_0(x) - p^* = \liminf_{n\to\infty} \sum_{k,m} c_{km}(\epsilon_{k_n})(1+x)^k(1-x)^m \geq \sum_{k,m} c^*_{km}(1+x)^k(1-x)^m.$$

Taking $x = x^* \in (-1, 1)$ in the above inequality clearly implies that $c^*_{km} = 0$. Therefore, the whole sequence $\{c_{km}(\epsilon)\}$ converges to the null sequence in $l_\infty$. Hence, the duals of the LP relaxations that provide the coefficients $\{c_{km}(\epsilon)\}$ in (2.14) will handle solutions with very small values as $\delta$ grows.

This is not true for the SDP relaxations. Indeed, from the representation (2.4), even if a global minimizer $x^*$ is in the interior of $[-1, 1]$, we have

$$g_0(x^*) - p^* = 0 = q(x^*)^2 + (1-x^*)^2 q_1(x^*)^2,$$

with $q(x^*) = q_1(x^*) = 0$; that is, both "polynomials multipliers" $q, q_1$ vanish at $x^*$. For a complete and detailed discussion on various representations of univariate polynomials positive on an interval, the interested reader is referred to the paper by Powers and Reznick (2000).

The examples for the univariate problems considered in Sherali and Tuncbilek (1997) are disappointing and confirm the drawbacks of the LP relaxations with only the subset (2.6) of the Hausdorff moment conditions (2.9). Indeed, the lower bound obtained by the LP relaxations with those constraints (and even additional *convex variable-bounding* constraints) is very far away from the optimal value $p^*$ ($\delta$ is the degree of the polynomial $g_0$ to minimize) (see Sherali and Tuncbilek 1997, Table 1 column $\nu(C - LB)$). Additional constraints are needed to improve the lower bounds. In contrast, a single SDP relaxation with 6 variables and 3 LMI constraints of size $4 \times 4$ solves exactly each problem.

EXAMPLE 2.4. Consider the (trivial) concave minimization problem $\min_{x\in[0,1]} x - x^2$. The version of the SDP relaxation (2.2) for the *even* case and the LP relaxation with $\delta = 2$ both find the global optimum $p^*$ (the global minimizer is at the boundary of [0, 1]). If we instead consider the convex minimization problem $\min_{x\in[0,1]} -x + x^2$, the SDP relaxation is again exact, whereas for the LP relaxations, we obtain

$$\rho_2 = \rho_4 = -1/3; \qquad \rho_6 = -0.3; \qquad \rho_{10} = -0.27; \qquad \rho_{15} = -0.2695,$$

with $\delta = 2, \ldots, 15$. Observe that in the latter problem, the global minimizer $x^* = 1/2$ is in the interior of the feasible set (see Remark 2.3(i)) and the convergence $\rho_\delta \uparrow p^* = -0.25$ is very slow.

**3. Multivariate polynomials.** In this section, we consider the (considerably more difficult) multivariate case,

$$\mathbb{P} \to \min_x \{g_0(x) \mid g_k(x) \geq 0,\ k = 1, \ldots, m\}, \tag{3.1}$$

where $g_k(x)$: $\mathbb{R}^n \to \mathbb{R}$ is a real-valued polynomial for all $k = 0, 1, \ldots, m$. In the generic problem considered in Sherali and Tuncbilek (1997), one assumes that the constraints $x \in [0, 1]^n$ are included in the constraints $\{g_k(x) \geq 0\}$. Equality constraints are also allowed (via the constraints $g_k(x) \geq 0$ and $-g_k(x) \geq 0$).

**3.1. SDP relaxations.** Let

$$1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1x_n, x_2x_3, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r \tag{3.2}$$

be a basis for the real-valued polynomials of degree at most $r$ and let $s(r)$ be its dimension. Therefore, an $r$-degree polynomial $p(x)$: $\mathbb{R}^n \to \mathbb{R}$ is written

$$p(x) = \sum_\alpha p_\alpha x^\alpha, \qquad x \in \mathbb{R}^n,$$

where $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, with $|\alpha| := \sum_{i=1}^n \alpha_i = k$, is a monomial of degree $k$ with coefficient $p_\alpha$. Denote by $p = \{p_\alpha\} \in \mathbb{R}^{s(r)}$ the vector of coefficients of $p(x)$ in the basis (3.2).

Let $\mathbb{K} := \{x \in \mathbb{R}^n \mid g_k(x) \geq 0, k = 1, \ldots, m\}$ be the feasible set of the problem $\mathbb{P}$ in (1.1). The degree of each polynomial $g_k(x)$ is written $2v_k - 1$ if odd and $2v_k$ if even, for all $k = 1, \ldots, m$. Again, with no loss of generality, we will assume that the constant term $g_0(0) = 0$.

For $i \geq \max_k v_k$, consider the following family $\{\mathbb{Q}_i\}$ of convex SDPs, introduced in Lasserre (2001)

$$\mathbb{Q}_i \begin{cases} \min\limits_y \sum\limits_\alpha (g_0)_\alpha y_\alpha, \\ M_i(y) \succeq 0, \\ M_{i-v_k}(g_k y) \succeq 0, \qquad k = 1, \ldots, m, \end{cases} \tag{3.3}$$

with respective dual problems

$$\mathbb{Q}_i^* \begin{cases} \min\limits_{X, Z_k \succeq 0} -X(1,1) - \sum\limits_{k=1}^m g_k(0) Z_k(1,1), \\ \langle X, B_\alpha \rangle + \sum\limits_{k=1}^m \langle Z_k, C_\alpha^k \rangle = (g_0)_\alpha, \quad \forall\, \alpha \neq 0, \end{cases} \tag{3.4}$$

where we have written

$$M_i(y) = \sum_\alpha B_\alpha y_\alpha; \qquad M_{i-v_k}(g_k y) = \sum_\alpha C_\alpha^k y_\alpha, \qquad k = 1, \ldots, m$$

for appropriate real-valued symmetric matrices $B_\alpha$, $C_\alpha^k$, $k = 1, \ldots, m+n$. The matrices $M_i(y)$ and $M_{i-v_k}(g_k y)$ are called *moment* and *localizing* matrices, respectively (for more details, see, e.g., Lasserre 2001a, b and Curto and Fialkow 1991, 2000). To see that $\mathbb{Q}_i$ is a relaxation of $\mathbb{P}$, let $x \in \mathbb{R}^n$ be a feasible solution of $\mathbb{P}$, let $u_i(x)$ be the vector in $\mathbb{R}^{s(i)}$ of the basis (3.2) for $r := i$, and let $y^x := u_{2i}(x)$ (with $y_0^x = 1$) so that $y^x_{\alpha_1 \cdots \alpha_n} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Then, from the definition of the moment and localizing matrices in Lasserre (2001a), and as $g_k(x) \geq 0$ for all $k = 1, \ldots, m$, we have

$$M_i(y^x) = u_i(x) u_i(x)' \succeq 0 \quad \text{and} \quad M_{i-v_k}(g_k y^x) = g_k(x) u_{i-v_k}(x) u_{i-v_k}(x)' \succeq 0$$

for all $k = 1, \ldots, m$; that is, $y^x$ is feasible for $\mathbb{Q}_i$.

It was shown in Lasserre (2001a) that under some conditions on the feasible set $\mathbb{K}$,

$$\inf \mathbb{Q}_i \uparrow p^* \quad \text{as } i \to \infty, \tag{3.5}$$

and $p^* = \min \mathbb{Q}_i$ for some $i$, whenever

$$g_0(x) - p^* = \sum_j q_j(x)^2 + \sum_{k=1}^m g_k(x) \left[ \sum_j q_{kj}(x)^2 \right] \tag{3.6}$$

for some polynomials $\{q_j(x)\}$ of degree at most $i$ and some polynomials $\{q_{kj}(x)\}$ of degree at most $i - v_k$. The above representation (3.6) is guaranteed for polynomials $p(x)$ that are strictly positive on $\mathbb{K}$, provided $\mathbb{K}$ is such that there is some polynomial $u(x)$ that can be written as in (3.6) and such that $\{x \mid u(x) \geq 0\}$ is compact (see, e.g., Putinar 1993). This condition is satisfied in many cases (like 0-1 nonlinear programs, or $\mathbb{K}$ compact with linear constraints). It suffices that $\{x \mid g_k(x) \geq 0\}$ be compact for some index $k \in \{1, \ldots, m\}$, and if not, one way to ensure the above condition on $\mathbb{K}$ is to add the quadratic constraint $\|x\|^2 \leq M$ for some $M$ large enough (see Lasserre 2001a). In the case of constraints $x \in \{0, 1\}^n$, the condition is satisfied (take $u(x) := \sum_i x_i(1 - x_i)$).

As for the univariate case, when (3.6) holds, the vectors of coefficients of the polynomials $\{q_j(x), q_{kj}(x)\}$ are provided by the eigenvectors of optimal solutions $\{X^*, Z_k^*\}$ of the dual problem $\mathbb{Q}_i^*$ (see Lasserre 2001).

**3.2. LP relaxations.** In the multivariate case, we will consider the generic LP relaxations obtained from the linearization of all possible mixed products of the original constraints, that is, constraints of the form

$$g_1(x)^{\alpha_1} g_2(x)^{\alpha_2} \cdots g_m(x)^{\alpha_m} \geq 0, \qquad |\alpha| := \sum_{i=1}^{m} \alpha_i \leq \delta. \tag{3.7}$$

After developing, each monomial term $x^\alpha$ in (3.7) is replaced with a variable $y_\alpha$, so as to obtain a linear inequality in the $y_\alpha$s.

Observe that the constraints (3.7) contain the so-called *bound-factor product constraints* (when one considers only the bound-constraints $0 \leq x_i \leq 1$) as well as the *constraint factor-based restrictions* of the RLT procedure of Sherali and Tuncbilek (1997).

Again, as in the univariate case, and after linearization, the bound-factor product constraints are nothing less than a *finite subset* of the *infinitely many* (multivariate analogs) necessary and sufficient Hausdorff moment conditions on the variables $y_\alpha$ to be moments of a probability measure $\mu$ supported in $[0, 1]^n$ (see, e.g., Shohat and Tamarkin 1943). The additional linear restrictions coming from the mixed products (3.7) are (only) necessary conditions for $\mu$ to be supported in $\mathbb{K}$.

Hence, the $\delta$ LP relaxation is the linear program $\mathbb{P}_\delta$:

$$\mathbb{P}_\delta \to \min_y \{c_\delta' y \mid A_\delta y \geq b_\delta \text{ deduced from (3.7) for every } |\alpha| \leq \delta\}. \tag{3.8}$$

Thus, the constraints (3.7) for all possible values of $\delta$ are only *necessary* conditions for the variables $y_\alpha$s to be the moments of some probability measure $\mu$ with support contained in $\mathbb{K}$. There is a case where those conditions guarantee convergence of the LP relaxations if one allows $\delta \to \infty$ and all possible products (3.7) are considered. If all the polynomials $g_k$ defining the constraint set are linear and define a convex polytope $\mathbb{K}$ with nonempty interior, then by a theorem of Handelman (1988), every polynomial $p$ (strictly) positive on $\mathbb{K}$ has the representation

$$p = \sum_{|\alpha| \leq m} b_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m} \tag{3.9}$$

for some $m$, and real-valued coefficients $b_\alpha \in \mathbb{R}_+$ (see also Powers and Reznick 2000, Theorem 2). Therefore, *in the linear case*, for every $\epsilon > 0$, as $g_0(x) - p^* + \epsilon > 0$ on $\mathbb{K}$, there is some $m(\epsilon)$ such that

$$g_0(x) - p^* + \epsilon = \sum_{|\alpha| \leq m(\epsilon)} b_\alpha g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m}, \qquad x \in \mathbb{R}^n, \tag{3.10}$$

and thus, for sufficiently large $\delta$, the LP relaxation $\mathbb{P}_\delta$, with $\delta := m(\epsilon)$, provides an optimal value $\rho_\delta$ within $\epsilon$ of $p^*$, and an optimal solution of the dual $\mathbb{P}_\delta^*$ provides the coefficients $\{b_\alpha\}$ in (3.10). The proof is similar to the univariate case and is omitted. As soon as some $g_k$ is not linear, there is no longer a guarantee of convergence because the representation (3.9) does not necessarily hold.

Note in passing that Handelman's (1988) result provides a rationale for the use of valid inequalities in nonconvex optimization on polytopes, the valid inequalities being various products of the original constraints.

Again, we may and will assume with no loss of generality that the constant term of $g_0(x)$ is zero, i.e., $g_0(0) = 0$.

THEOREM 3.1. *Consider the problem* $\mathbb{P}$ *in* (3.1) *and the LP relaxation* $\mathbb{P}_\delta$ *in* (3.8) *defined from the constraints* (3.7). *Let* $\rho_\delta$ *be its optimal value*:

(a) *For every* $\delta$, $\rho_\delta \le p^*$ *and*

$$g_0(x) - \rho_\delta = \sum_{|\alpha|\le\delta} b_\alpha(\delta) g_1(x)^{\alpha_1}\cdots g_m(x)^{\alpha_m}, \qquad x \in \mathbb{R}^n \tag{3.11}$$

*for some nonnegative scalars* $\{b_\alpha(\delta)\}$. *Let* $x^*$ *be a global minimizer of* $\mathbb{P}$ *and let* $I(x^*)$ *be the set of active constraints at* $x^*$. *If* $I(x^*) = \varnothing$ (*i.e.*, $x^*$ *is in the interior of* $\mathbb{K}$) *or if there is some feasible, nonoptimal solution* $x \in \mathbb{K}$ *with* $g_i(x) = 0, \forall i \in I(x^*)$, *then* $\rho_\delta < p^*$ *for all* $\delta$, *that is, no relaxation* $\mathbb{P}_\delta$ *can be exact.*

(b) *If all the* $g_i$ *are linear, that is, if* $\mathbb{K}$ *is a convex polytope, then* (3.11) *holds and* $\rho_\delta \uparrow p^*$ *as* $\delta \to \infty$. *If* $I(x^*) = \varnothing$ *for some global minimizer* $x^*$, *then in* (3.11),

$$\sum_\alpha b_\alpha(\delta) \to 0 \quad \textit{as } \delta \to \infty. \tag{3.12}$$

PROOF. (a) The representation (3.11) follows from the definition of the dual of the LP relaxation $\mathbb{P}_\delta$ (and strong duality in linear programming, as soon as the primal has finite value), in the same manner as was done for the univariate case. Next, let $x^*$ be a global minimizer of $\mathbb{P}$, in the interior of $\mathbb{K}$, that is, $g_k(x^*) > 0$ for all $k = 1, \dots, m$. Then, from (3.11), it follows that $p^* - \rho_\delta > 0$, because $g_1(x^*)^{\alpha_1}\cdots g_m(x^*)^{\alpha_m} > 0$ for every $\alpha = (\alpha_1, \dots, \alpha_m)$. More generally, let $I(x^*)$ be the set of active constraints at a global minimizer $x^* \in \mathbb{K}$, that is, $g_i(x^*) = 0$ whenever $i \in I(x^*)$ and $g_i(x^*) > 0$ for $i \notin I(x^*)$. Then, for the LP relaxation $\mathbb{P}_\delta$ to be exact, one needs to have $\rho_\delta = p^*$ in (3.11), and

$$b_\alpha(\delta) > 0 \Rightarrow \alpha_i > 0, \qquad \text{for some } i \in I(x^*);$$

otherwise, if there is some $b_\alpha(\delta) > 0$ with $\alpha_i = 0, \forall i \in I(x^*)$, then from (3.11), $g_0(x^*) - p^* > 0$, which is a contradiction. However then, let $x \in \mathbb{K}$ be any feasible nonoptimal solution with $g_i(x) = 0$, $\forall i \in I(x^*)$. With the same argument, it follows that $g_0(x) - p^* = 0$, which contradicts $x$ nonoptimal. Thus, $\rho_\delta < p^*$ for every $\delta$.

(b) The proof that $\rho_\delta \uparrow p^*$ follows from Handelman's (1988) result. For every $\epsilon > 0$ (3.10) holds for some $m(\epsilon)$, and thus, if we take $\delta \ge m(\epsilon)$ with $m(\epsilon)$ as in (3.10), the optimal value of the dual of the LP relaxation will be $\rho_\delta = p^* - \epsilon$, and the result follows as $\delta \to \infty$.

Finally, to get (3.12), let $x_0$ be such that $g_k(x_0) > 0$ for all $k = 1, \dots, m$. With no loss of generality, after division of each $g_k(x)$ by $g_k(x_0)$, we may have assumed from the beginning that $g_k(x_0) = 1$ for all $k = 1, \dots, m$. With the same arguments as in the univariate case, the sequence of coefficients $\{b_\alpha(\delta)\}$ is bounded, for we have

$$g_0(x_0) - \rho_\delta = \sum_{|\alpha|\le\delta} b_\alpha(\delta) g_1(x_0)^{\alpha_1}\cdots g_m(x_0)^{\alpha_m} = \sum_{|\alpha|\le\delta} b_\alpha(\delta), \quad \forall \delta.$$

Hence, extended with zeros, the sequence $\{b_\alpha(\delta)\}$ is considered as an element of $l_\infty^+$. Therefore, taking a sequence $\delta_k \to \infty$, the corresponding sequence $\{b_\alpha(\delta_k)\} \in l_\infty$ has a (pointwise) converging subsequence $\{b_\alpha(\delta_{k_n})\} \to \{b_\alpha^*\} \in l_\infty$. With $x \in \mathbb{K}$ being fixed and arbitrary, consider (3.11). By Fatou's Lemma, and from $\rho_\delta \uparrow p^*$,

$$g_0(x) - p^* = \liminf_{n\to\infty} \sum_\alpha b_\alpha(\delta_{k_n}) g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m} \geq \sum_\alpha b_\alpha^* g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m},$$

which, taking $x = x^*$, shows that $\{b_\alpha^*\} \equiv \{0\}$ (as $g_i(x^*) > 0$ for all $i$). Hence, the whole sequence $\{b_\alpha(\delta_k)\}$ converges to the null sequence $\{0\}$, and $\sum_\alpha b_\alpha(\delta) \to 0$ as $\delta \to \infty$. □

Hence, again, in both SDP and LP relaxations, the vector $y$ has the same interpretation, namely, the *moment* vector of some probability measure $\mu$. The constraints in (3.3) and (3.8) are different necessary conditions for $y$ to be the moment vector of a probability measure $\mu$ with support contained in the feasible set $\mathbb{K}$. Similarly, the respective duals $\mathbb{Q}_i^*$ and $\mathbb{P}_\delta^*$ aim at representing the polynomial $g_0(x) - p^* + \epsilon$ in two different ways:

- as a sum of the $g_k$s, weighted by *sum of squares*, for SDP relaxations (see (3.6)) and
- as a sum of products $g_1^{\alpha_1} \cdots g_m^{\alpha_m}$ of the constraints, weighted by *nonnegative scalars* (see 3.10).

Thus the main difference between SDP and LP relaxations is their pursuit (or vocation) (via their respective dual relaxations $\mathbb{Q}_i^*$ and $\mathbb{P}_\delta^*$) in representing $g_0 - p^* + \epsilon$ as in (3.6) or (3.10), respectively, with $\epsilon$ as small as possible (and $\epsilon = 0$ if possible).

Recent results of real algebraic geometry tell us that for arbitrary $\epsilon > 0$, the first representation is indeed legitimate in a rather general framework (see Putinar 1993 and Jacobi and Prestel 2001), whereas the second representation is guaranteed only if $\mathbb{K}$ is a convex polytope (by Handelman's 1988 result). For the important special case of 0-1 nonlinear programs, both SDP and LP relaxations exhibit *finite* convergence (as proved in Lasserre 2001b for SDP relaxations and in Sherali and Adams 1990 for LP relaxations).

REMARK 3.2. Theorem 3.1 has important consequences that we summarize here:

(a) In the case of nonlinear constraints, the LP relaxation cannot be exact in general, no matter how large $\delta$ is. Moreover, in general, $\rho_\delta$ will be bounded away from $p^*$, since the representation (3.11) with $\rho_\delta > p^* - \epsilon$, and $\epsilon$ arbitrary small, does *not* hold in general. A notable exception is the case of 0-1 programs (i.e., with the constraints $x_i^2 = x_i$ for all $i = 1, \ldots, n$) and some special 0-1 constrained programs, as demonstrated in Sherali and Adams (1990), where the (adapted) $\delta$ (=$n$) LP relaxation is exact. Observe that in this case, the bound-constraints $x_i \geq 0$, $(1 - x_i) \geq 0$ are such that the set of active constraints $I(x^*)$ at $x^*$ determines a unique point $x^*$, and the last statement of Theorem 3.1(a) does not apply. Moreover, in a recent paper, Laurent (2001) has shown that for 0-1 programs, the SDP relaxation $\mathbb{Q}_i$ is tighter than the corresponding LP relaxation of Sherali and Adams (1990).

(b) In the linear case, that is, when $\mathbb{K}$ is a convex polytope, the asymptotic (and in general, not finite) convergence $\rho_\delta \uparrow p^*$ holds. However, in this case, the primal LP relaxation is ill conditioned in view of the (large) binomial coefficients involved. Therefore, in practice, it is preferable to fix $\delta$ and use this relaxation in a branch-and-bound procedure as, e.g., in Sherali and Tuncbilek (1997) or Audet et al. (2000). Moreover, if there is a global minimizer in the interior of the feasible set $\mathbb{K}$, the dual LP relaxation (also ill conditioned) yields an almost null solution, for all the coefficients $b_\alpha(\delta)$ in the representation (3.11) vanish as $\delta \to \infty$.

We illustrate the preceding result on the following example.

EXAMPLE 3.3. This example is the global minimization of a fourth-degree polynomial on $\mathbb{R}^4$, found in Bartholomew-Biggs (1976) and also considered in Audet et al. (2000).

$$\begin{cases} \min_{x\in\mathbb{R}^4} x_3 + x_1^2 x_4 + x_1 x_2 x_4 + x_1 x_3 x_4 \\ \qquad\qquad x_1^2 + x_2^2 + x_3^2 + x_4^2 = 40, \\ \qquad\qquad x_1 x_2 x_3 x_4 \geq 25, \\ \qquad\qquad x_i \in [1,5], \qquad i = 1, \ldots, 4. \end{cases}$$

A global optimum is

$$x^* = (1, 4.74319, 3.8209, 1.37944)$$

and is found exactly at the first SDP relaxation $\mathbb{Q}_2$ (the first to consider as we have polynomials of degree 4). It is also found in the branch-and-bound procedure of Audet et al. (2000), based on the LP relaxation of Sherali and Adams's (1990) RLT procedure.

There are 11 polynomial constraints, $g_k(x) \geq 0$, $k = 1, \ldots, 11$ (the equality constraint $\|x\|^2 = 40$ being written as two opposite inequalities $g_1(x) := 40 - \|x\|^2 \geq 0$, $g_2(x) := -g_1(x) \geq 0$). We show that $g_0(x) - p^*$ cannot have the representation

$$g_0 - p^* = \sum_\alpha b_\alpha g_1^{\alpha_1} \cdots g_{11}^{\alpha_{11}} \tag{3.13}$$

for some nonnegative scalars $\{b_\alpha\}$. Indeed, for the representation (3.13) to hold, there must be some products with $\alpha_1 = \alpha_2 = 0$; otherwise, every $x \in \mathbb{K}$ would satisfy $g_0(x) - p^* = 0$. Moreover, as the constraints $g_3(x) := x_1 x_2 x_3 x_4 - 25 \geq 0$ and $g_4(x) := x_1 - 1 \geq 0$ are the only ones binding at $x^*$ (except $g_1, g_2$), we should have $\alpha_3 + \alpha_4 > 0$ whenever $\alpha_1 = \alpha_2 = 0$ (otherwise, $g_0(x^*) - p^* = 0 > 0$, which is a contradiction). However then, every feasible point $x$ with $x_1 = 1$ and $x_1 x_2 x_3 x_4 = 25$ would satisfy $g_0(x) - p^* = 0$! (Take, for example, the (nonoptimal) feasible point $x = (1, 5, 3.44949, 1.44949)$.)

The next trivial example shows that the absence of bound constraints may imply that the lower bounds of the LP relaxations can be bounded away from the optimal value $p^*$.

EXAMPLE 3.4. Consider the following trivial one-dimensional example:

$$\begin{cases} \min_{x\in\mathbb{R}} -x, \\ x^2 - 1 = 0, \\ \dfrac{1}{2} - x \geq 0, \end{cases}$$

with global optimum $p^* = 1$ at the optimal solution $x^* = -1$. The feasible set $\mathbb{K}$ is compact, but we cannot have the representation

$$g_0(x) - p^* = -x - 1 = \sum_\alpha b_\alpha (x^2 - 1)^{\alpha_1} \left(\frac{1}{2} - x\right)^{\alpha_2},$$

with $b_\alpha \geq 0$ whenever $\alpha_1 = 0$. Indeed, $g_0(-1) - 1 = 0$ implies $\alpha_1 > 0$ for all $\alpha$. However, on the other hand, this would yield $g_0(1) - 1 = -2 = 0$, which is a contradiction. Thus, every LP relaxation $\mathbb{P}_\delta$ cannot be exact, so that $\rho_\delta < p^*$. However, from the interpretation of an optimal solution of its dual, we have the representation

$$g_0(x) - \rho_\delta = \sum_{|\alpha| \leq \delta} b_\alpha (x^2 - 1)^{\alpha_1} \left(\frac{1}{2} - x\right)^{\alpha_2},$$

with $\rho_\delta < p^*$. Hence, necessarily, there will be terms like $b_\alpha (1/2 - x)^\alpha$; otherwise, in the representation of $g_0(x) - \rho_\delta$, we would have $0 < p^* - \rho_\delta = 0$. However, this implies that the best lower bound $\rho_\delta$ of the LP relaxation is 0 because we must have (writing $\rho_\delta = 1-\epsilon$) for some $\epsilon > 0$:

$$-x-(1-\epsilon) = \sum_{\alpha=0}^{\delta} b_\alpha (1/2-x)^\alpha + \sum_{|\beta|\le\delta} b_\beta (x^2-1)^{\beta_1}\left(\frac{1}{2}-x\right)^{\beta_2},$$

with $\mathscr{B}_\alpha \ge 0$, which yields (for $x=-1$ and $x=1$),

$$\epsilon = \sum_{\alpha=0}^{\delta} b_\alpha (3/2)^\alpha; \qquad -2+\epsilon = \sum_{\alpha=0}^{\delta} b_\alpha (-1/2)^\alpha;$$

summing up yields $2\epsilon = 2+\sum_{\alpha=0}^{\delta} b_\alpha[(3/2)^\alpha + (-1/2)^\alpha] \ge 2$, so that $\epsilon \ge 1$, and thus $\rho_\delta \le 0 < p^* = 1$.

However, if we include the bound constraints $(1+x)\ge 0$ and $(1-x)\ge 0$, as required in the RLT approach (Sherali and Tuncbilek 1997), then we have

$$-x-1 = 2\left(\frac{1}{2}-x\right)(1+x) + 2(x^2-1),$$

and the LP relaxation with $\delta = 2$ will indeed provide the optimal value $-1$. Observe that the global minimizer saturates a bound constraint (cf., Sherali and Tuncbilek 1997, Remark 3.2(b)).

On the other hand, we have the representation

$$-x-1 = (x^2-1)\left(x+\frac{7}{4}\right)^2 + (1-x^2)\left(x+\frac{5}{4}\right)^2 + \left(\frac{1}{2}-x\right)(x+1)^2.$$

Thus, the SDP relaxation $\mathbb{Q}_2$ yields the optimal value $p^*$.

**4. Conclusion.** We have shown that there is a common natural framework for comparing SDP and LP relaxations, namely, the theory of moments and its dual theory of representation of polynomials, positive on a compact semi-algebraic set. Each relaxation aims at representing the polynomial $g_0(x) - p^*$ in a specific manner. For the SDP relaxations, the polynomials $g_k$ of the constraints are weighted by *sums of squares of polynomials*, whereas for the LP relaxations, all possible products of the $g_k$'s are weighted by *nonnegative scalars*. From results in algebraic geometry, it appears that the former representation is far more general than the latter, which even in the case of convergence, has some drawbacks. However, the present status of SDP software packages is not as advanced as that of their LP counterparts, so that high-order SDP relaxations are not yet a viable alternative for large-size or even medium-size problems. We hope this paper will contribute stimulating efforts toward improving SDP-solving procedures.

**Acknowledgments.** The author wishes to thank anonymous referees for several helpful remarks and suggestions to improve the paper.

## References

Adams, W. P., H. D. Sherali. 1986. A tight linearization and an algorithm for zero-one quadratic programming problems. *Management Sci.* **32** 1274–1290.

Audet, C., P. Hansen, B. Jaumard, G. Savard. 2000. A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. *Math. Progamming Ser. A* **87** 131–152.

Bartholomew-Biggs, M. 1976. A numerical comparison between two approaches to nonlinear programming problems. Technical Report #77, Numerical Optimization Center, Hatfield, U.K.

Berg, C. 1987. The multidimensional moment problem and semi-groups. H. J. Landau, ed. *Proc. Sympos. Appl. Math.*, vol. 37. American Mathematics Society, Providence, RI, 110–124.

Burer, S., R. Monteiro. 2000. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Working paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Curto, R. E., L. A. Fialkow. 1991. Recursiveness, positivity, and truncated moment problems. *Houston J. Math.* **17** 603–635.

———, ———. The truncated complex $K$-moment problem. *Trans. Amer. Math. Soc.* **352** 2825–2855.

Feller, W. 1966. *An Introduction to Probability Theory and Its Applications.* John Wiley & Sons, New York.

Goemans, M. X., D. P. Williamson. 1995. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM* **42** 1115–1145.

Handelman, D. 1988. Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific J. Math.* **132** 35–62.

Jacobi, T., A. Prestel. 2001. Distinguished representations of strictly positive polynomials. *J. Reine Angewandle Mathematik* **532** 223–235.

Kojima, M., L. Tunçel. 2000. Cones of matrices and successive convex relaxations of non convex sets. *SIAM J. Optim.* **10** 750–778.

Lasserre, J. B. 2000. Optimality conditions and LMI relaxations for 0-1 programs. LAAS-CNRS Technical Report No. 00098, LAAS, Toulouse, France.

———. 1999. The global optimization of a polynomial is an easy convex problem. LAAS Technical Report No. 99485, LAAS, Toulouse, France.

———. 2001a. Global optimization with polynomials and the problem of moments. *SIAM J. Optim* **11** 796–817.

———. 2001b. An explicit exact SDP relaxation for nonlinear 0-1 programs. K. Aardal, A. M. H. Gerards, eds. *Lecture Notes in Computer Science*. Springer, Berlin, 293–303.

Laurent, M. 2001. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. Technical Report PNA-R0108, Centrum Voor Wiskunde en Informatica, Amsterdam, The Netherlands.

Lovász, L., A. Schrijver. 1991. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.* **1** 166–190.

Nesterov. Y. 2000. Squared functional systems and optimization problems. H. Frenk, K. Roos, S. Zhang, eds. *High Performance Optimization Methods*. Kluwer Academic Publishers, Dordrecht, The Netherlands. 405–439.

Powers, V., B. Reznick. 2000a. Polynomials that are positive on an interval. *Trans. Amer. Math. Soc.* **352** 4677–4692.

———, ———. 2001. A new bound for Pólya's Theorem with applications to polynomials positive on polyhedra. *J. Pure Appl. Algebra* **164** 221–229.

Putinar, M. 1993. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.* **42** 969–984.

Schmüdgen, K. 1991. The $K$-moment problem for compact semi-algebraic sets. *Mathematische Annalen* **289** 203–206.

Sherali, H. D., W. P. Adams. 1990. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.* **3** 411–430.

———, M. P. Fraticelli. 2000. Enhancing RLT relaxations via a new class of semidefinite cuts. Working paper, Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Insitute and State University, Blacksburg, VA.

———, C. H. Tuncbilek. 1992. A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique. *J. Global Optim.* **2** 101–112.

———, ———. 1997. New reformulation linearization/convexification relaxations for univariate and multivariate polynomial programmimg problems. *Oper. Res. Lett.* **21** 1–9.

Shohat, J. A., J. D. Tamarkin. 1943. The problem of moments. *American Mathematical Society Mathematical Surveys*, vol. II. American Mathematical Society, New York.

Simon, B. 1998. The classical moment problem as a self-adjoint finite difference operator. *Adv. Math.* **137** 82–203.

Shor, N. Z. 1987. Quadratic optimization problems. *Tekhnicheskaya Kibernetika* **1** 128–139.

Vandenberghe, L., S. Boyd. Semidefinite programming. *SIAM Rev.* **38** 49–95.

Vanderbei, R. J., H. Y. Benson, 2000. On formulating semidefinite programming problems as smooth convex nonlinear optimization problems. Working paper, Department of Operations Research and Financial Engineering, Princeton University, NJ.

J. B. Lasserre: LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse Cédex 4, France; e-mail: lasserre@laas.fr