Convergent LMI relaxations for nonconvex quadratic programs

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

Abstract

We consider the general nonconvex quadratic programming problem and provide a series of convex positive semidefinite programs (or LMI relaxations) whose sequence of optimal values is monotone and converges to the optimal value of the original problem. It improves and includes as a special case the well-known Shor's LMI formulation. Often, the optimal value is obtained at some particular early relaxation as shown on some nontrivial test problems from Floudas and Pardalos [9].

1 Introduction

In this paper we consider the general nonconvex quadratic programming problem

$$\mathbb{P} \mapsto p^* := \min_{x\in\mathbb{R}^n} \{g_0(x) \,|\, g_i(x) \geq 0,\, i = 1,2,\dots,m\}. \quad (1)$$

where the $g_i(x)$ are all quadratic polynomials, $i = 0,\dots m$, and where neither the criterion nor the constraint set are assumed to be convex. We also allow equality constraints since they can be written with two opposite inequalities. In fact, any optimization problem with polynomials can be put in this form in an augmented space (see Shor [13], [14] and Ferrier [3]). This problem is very difficult in general, for several NP-hard problems can be put in this form. However, there is a well-known relaxation of this problem, known as Shor's relaxation [13], which replaces $\mathbb{P}$ by a convex positive semidefinite (psd) program (or LMI relaxation). This relaxation has been proved very useful for computing estimates of global minima, even in combinatorial optimization. For instance, for the well-known MAX-CUT problem (a special case of (1)), very good approximations of a global minimum have been obtained in Goemans and Williamson [4] using Shor's relaxation, followed by a randomized rounding procedure.

In this paper, we propose a family $\{\mathbb{Q}_i\}$ of convex LMI relaxations (or psd programs) that contains Shor's relaxation as its first member $\mathbb{Q}_1$ and with an associated *increasing* sequence of lower bounds $\{\inf \mathbb{Q}_i\}$.

When the constraint set is compact, the nondecreasing sequence of these lower bounds converges to the global optimal value p^* in (1). In fact, in many cases, the global optimal value is reached exactly in a few steps. The approach is based on the theory of moments and recent results on the representation of polynomials that are strictly positive on a compact semi algebraic set. For results on the theory of moments and the representation of positive polynomials, the reader is referred to Curto and Fialkow [2], Berg [1], Schmüdgen [11], Putinar [10], Jacobi [5], Jacobi and Prestel [6]. It turns out that this theory is a natural and appropriate tool for global optimization since $g_0(x) - p^*$ is precisely a (non strictly) positive polynomial, a feature that distinguishes p^* from the other (local) minima. Moreover, the LMI relaxations are well-suited since both primal and dual psd programs perfectly match both sides of the same theory (moments and positive polynomials). Indeed, when the optimal value p^* is obtained at some relaxation, say $\mathbb{Q}_i$, the primal psd program provides a global minimizer whereas the dual psd program $\mathbb{Q}_i^*$ provides the coefficients of polynomials in the decomposition of $g_0(x) - p^*$ into a weighted sum of squares.

Of course, there is a price to pay for these refined relaxations. The number of variables in the LMI relaxation $\mathbb{Q}_i$ is the dimension of the vector space of real-valued polynomials in n variables of degree $2i$, that is, $O(n^{2i})$.

However, the $\mathbb{Q}_2$ relaxation is already interesting. On a sample of 50 randomly generated MAX-CUT problems in $\mathbb{R}^{10}$, the $\mathbb{Q}_2$ relaxation always provided the optimal solution, in contrast to only a few cases for the $\mathbb{Q}_1$ relaxation. The results obtained on a sample of nontrivial test problems taken from Floudas and Pardalos [9] are also promising.

2 Notation, Definitions and Preliminaries

For all $i = 0,1,\dots,m$, let A_i be a real-valued symmetric (n,n)-matrix, $c_i \in \mathbb{R}^n$, and consider the general nonconvex quadratic programming problem (1) with $g_i(x) = x'A_ix + c_i'x + b_i$, $i = 1,\dots m$ and $g_0(x) = x'A_0x + c_0'x$. Let $\mathbb{K}$ be the feasible set of $\mathbb{P}$, that is, $\mathbb{K} = \{g_i(x) \geq 0,\, i = 1,\dots m\}$.

Given any two real-valued symmetric matrices A,B let $\langle A,B\rangle$ denote the usual scalar product trace(AB) and let $A \succeq B$ (resp. $A \succ B$) stand for $A-B$ psd (resp. $A-B$ positive definite). Let

$$1, x_1, x_2, \dots x_n, x_1^2, x_1x_2 \dots, x_1^r, \dots, x_n^r, \quad (2)$$

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

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

where

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}, \quad \text{with } \sum_{i=1}^n \alpha_i = k,$$

is a monomial of degree k with coefficient p_α. Denote by $p = \{p_\alpha\} \in \mathbb{R}^{s(r)}$ the coefficients of the polynomial $p(x)$ in the basis (2).

Hence, the respective vectors of coefficients of the polynomials $g_i(x)$, $i = 0, 1, \ldots m$, are denoted $\{(g_i)_\alpha\} =: g_i \in \mathbb{R}^{s(2)}$, $i = 0, 1, \ldots m$.

Given a $s(2r)$-sequence $(1, y_1, \ldots,)$, let $M_r(y)$ be the **moment** matrix of dimension $s(r)$, with rows and columns labelled by (2). For instance, for illustration purposes, and for clarity of exposition, consider the 2-dimensional case. The moment matrix $M_r(y)$ is the block matrix $\{M_{i,j}(y)\}_{0 \le i,j \le r}$ defined by

$$M_{i,j}(y) = \begin{bmatrix} y_{i+j,0} & y_{i+j-1,1} & \cdots & y_{i,j} \\ y_{i+j-1,1} & y_{i+j-2,2} & \cdots & y_{i-1,j+1} \\ \cdots & \cdots & \cdots & \cdots \\ y_{j,i} & y_{i+j-1,1} & \cdots & y_{0,i+j} \end{bmatrix}. \tag{3}$$

To fix ideas, when $n = 2$ and $r = 2$ one obtains

$$M_2(y) = \begin{bmatrix} 1 & | & y_{10} & y_{0,1} & | & y_{20} & y_{11} & y_{0,2} \\ - & - & - & - & - & - & - & - \\ y_{10} & | & y_{20} & y_{11} & | & y_{30} & y_{21} & y_{12} \\ y_{01} & | & y_{11} & y_{02} & | & y_{21} & y_{12} & y_{03} \\ - & - & - & - & - & - & - & - \\ y_{20} & | & y_{30} & y_{21} & | & y_{40} & y_{31} & y_{22} \\ y_{11} & | & y_{21} & y_{12} & | & y_{3,1} & y_{22} & y_{13} \\ y_{02} & | & y_{12} & y_{03} & | & y_{22} & y_{13} & y_{04} \end{bmatrix}$$

For the 3-dimensional case, $M_r(y)$ is defined via blocks $\{M_{i,j,k}(y)\}$, $0 \le i, j, l \le r$ in a similar fashion, and so on.

If the entry (i, j) of the matrix $M_m(y)$ is y_β, let $\beta(ij)$ denotes the subscript β of y_β. Next, for a polynomial $\theta(x) : \mathbb{R}^n \to \mathbb{R}$, let $M_r(\theta y)$ be the matrix defined by

$$M_r(\theta y)(i, j) = \sum_\alpha \theta_\alpha y_{\{\beta(i,j)+\alpha\}}. \tag{4}$$

For instance, with

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

$M_1(\theta y)$ reads

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

$M_r(y)$ define a bilinear form $\langle .,. \rangle_y$ on the space A_r of r-degree real-valued polynomials by

$$\langle q(x), v(x) \rangle_y := \langle q, M_r(y) v \rangle, \quad q(x), v(x) \in A_r,$$

and if y is a sequence of moments of some measure μ_y, then

$$\langle q, M_r(y) q \rangle = \int q(x)^2 \mu_y(dx) \ge 0, \tag{5}$$

so that $M_r(y) \succeq 0$, and similarly,

$$\langle q, M_r(\theta y) q \rangle = \int \theta(x) q(x)^2 \mu_y(dx), \tag{6}$$

so that $M_r(\theta y) \succeq 0$ whenever μ_y has its support contained in $\{\theta(x) \ge 0\}$. The theory of *moments* identifies those sequences y with $M_r(y) \succeq 0$, which are moment-sequences. For details and recent results, the interested reader is referred to Berg [1], Curto and Fialkow [2], Simon [12], Schmüdgen [11] and the many references therein.

3 A family of convex LMI relaxations

Consider the problem $\mathbb{P}$ with feasible set $\mathbb{K} = \{g_i(x) \ge 0, i = 1, \ldots m\}$ and criterion $g_0(x)$. When needed below, the vectors $g_i \in \mathbb{R}^{s(2)}$ are extended to vectors of $\mathbb{R}^{s(2i)}$ by completing with zeros, i.e., the quadratic polynomials polynomials $g_i(x)$, $i = 0, \ldots m$, are considered as $2i$-degree polynomials with null coefficients for terms of degree larger than 2. As we minimize g_0, we assume that its constant term is zero, that is, $g_0(0) = 0$.

For $i = 1, 2, \ldots$, consider the following family of convex LMI problems

$$\mathbb{Q}_i \left\{ \begin{array}{lll} \min\limits_y \sum\limits_\alpha (g_0)_\alpha y_\alpha & & \\ \qquad M_i(y) & \succeq & 0 \\ \qquad M_{i-1}(g_k y) & \succeq & 0, \quad k = 1, \ldots m. \end{array} \right. \tag{7}$$

with respective dual problems

$$\mathbb{Q}_i^* \left\{ \begin{array}{l} \max\limits_{X, Z_k \succeq 0} -X(1,1) - \sum\limits_{k=1}^m g_k(0) Z_k(1,1) \\ \langle X, B_\alpha \rangle + \sum\limits_{k=1}^m \langle Z_k, C_\alpha^k \rangle = (g_0)_\alpha, \ \forall \alpha \ne 0, \end{array} \right. \tag{8}$$

where we have written

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

(with $y_0 = 1$) for appropriate real-valued symmetric matrices B_α, C_α^k, $k = 1, \ldots m$.

At this stage, it is worthy to write down the LMI program $\mathbb{Q}_1$, that is, when $i = 1$.

$$\mathbb{Q}_1 \left\{ \begin{array}{lll} \min\limits_y \sum\limits_\alpha (g_0)_\alpha y_\alpha & & \\ \qquad M_1(y) & \succeq & 0 \\ \quad \sum\limits_\alpha (g_k)_\alpha y_\alpha & \ge & 0 \quad k = 1, \ldots, m \end{array} \right. \tag{9}$$

with dual

$$
\mathbb{Q}_1^* \begin{cases} \max_{\lambda \geq 0, X \succeq 0} -X(1,1) - \sum_{k=1}^{m} \lambda_k g_k(0) \\ \langle X, B_\alpha \rangle + \sum_{k=1}^{m} \lambda_k (g_k)_\alpha \,=\, (g_0)_\alpha, \, \forall \alpha \neq 0. \end{cases} \tag{10}
$$

Observe that, if we write

$$
M_1(y) \,=\, \begin{bmatrix} 1 & y' \\ y & Y \end{bmatrix},
$$

and remembering that $g_k(x) = x'A_k x + c_k' x + b_k$, $k = 0, 1, \ldots m$, then $\mathbb{Q}_1$ has the equivalent form

$$
\mathbb{Q}_1 \begin{cases} \min_{Y,y} \langle A_0, Y \rangle + c_0' y & & \\ \quad \langle A_k, Y \rangle + c_k' y & \geq & -b_k, k = 1, \ldots m \\ \quad \begin{bmatrix} 1 & y' \\ y & Y \end{bmatrix} & \succeq & 0 \end{cases} \tag{11}
$$

which is the dual of the well-known Shor's relaxation for quadratic programs (see e.g. [15]). Therefore, Shor's relaxation is the first LMI program $\mathbb{Q}_1$ in the hierarchy of relaxations $\{\mathbb{Q}_i\}$.

We need the following intermediate result:

Proposition 3.1 *The family of LMI problems $\{\mathbb{Q}_i\}$ satisfies*

$$
\inf \mathbb{Q}_i \,\leq\, \inf \mathbb{Q}_{i+1} \,\leq\, \inf \mathbb{P}, \qquad i = 1, 2, \ldots. \tag{12}
$$

Proof: That $\inf \mathbb{Q}_i \leq \inf \mathbb{P}$ is immediate because to every admissible point x of $\mathbb{P}$, corresponds a solution

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

admissible for $\mathbb{Q}_i$, and thus $\inf \mathbb{Q}_i \leq \inf \mathbb{P}$ for every $i = 1, \ldots$.

Next, consider an LMI problem $\mathbb{Q}_i$, with $i > 1$, and let $y = \{y_\alpha\}$ a feasible sequence for $\mathbb{Q}_i$, that is, $(1, y)$ is a vector of dimension $s(2i)$, the dimension of the vector space of real-valued polynomials $p(x) : \mathbb{R}^n \to \mathbb{R}$, of degree $2i$. If we write $y = (y_1, y_2)$ with $(1, y_1)$ of dimension $s(2(i-1))$, we have

$$
M_i(y) \,=\, \begin{bmatrix} M_{i-1}(y_1) & M \\ M' & N \end{bmatrix},
$$

and

$$
M_{i-1}(g_k y) \,=\, \begin{bmatrix} M_{i-2}(g_k y_1) & V \\ V' & NS \end{bmatrix}, k = 1, \ldots m
$$

for appropriate matrices M, N, V, S. Therefore,

$$
M_i(y) \succeq 0 \Rightarrow M_{i-1}(y_1) \succeq 0,
$$

and

$$
M_{i-1}(g_k y) \succeq 0 \Rightarrow M_{i-2}(g_k y_1) \succeq 0, \, k = 1, \ldots m,
$$

so that y_1 is admissible for $\mathbb{Q}_{i-1}$. Moreover, the value of y in $\mathbb{Q}_i$ is the same as the value of y_1 in $\mathbb{Q}_{i-1}$, and the result follows. ∎

Proposition 3.1 ensures that better and better lower bounds on $\mathbb{P}$ can be obtained by solving the relaxations $\mathbb{Q}_i$, $i = 1, \ldots$. The next result shows that in fact, whenever $\mathbb{K}$ is compact, one may approach as closely as desired, the optimal value $p^* = \inf \mathbb{P}$. We will use the fact that under a certain condition on the feasible set $\mathbb{K}$, every polynomial $p(x)$, strictly positive on $\mathbb{K}$, has the following representation:

$$
p(x) \,=\, \sum_{j=1}^{r_0} q_j(x)^2 + \sum_{k=1}^{m} g_k(x) \left[\sum_{j=1}^{r_k} q_{kj}(x)^2 \right] \tag{13}
$$

for a finite family of polynomials $\{q_j(x)\}$, $j = 1, \ldots r_0$, and $\{q_{kj}(x)\}$, $j = 1, \ldots r_k$, $k = 1, \ldots m$. In fact, a necessary and sufficient condition for the representation (13) to exist is that there exists a polynomial $u(x)$ of the form (13) such that $\{u(x) \geq 0\}$ is compact (see Putinar [10] and Jacobi [5]).

For instance, the representation (13) holds whenever $\{g_k(x) \geq 0\}$ is compact for some k, or when all the $g_k(x)$ are linear and $\mathbb{K}$ is compact. In particular, it holds for every 0-1 program. Indeed, write the integral constraints as $x_i^2 - x_i \geq 0$ and $x_i - x_i^2 \geq 0$, for all $i = 1, \ldots n$. Then, consider the polynomial $u(x) := \sum_{i=1}^{n} (x_i - x_i^2)$. Its level set $\{u(x) \geq 0\}$ is compact. Moreover, if one knows that a global minimizer is contained in some ball $||x||^2 \leq M$, for some M large enough, then one may add the redundant constraint $g_{m+1}(x) := M - \sum_{i=1}^{n} x_i^2 \geq 0$, and the set $\mathbb{K}$ (defined as previously with the latter additional constraint $g_{m+1}(x) \geq 0$) has the required property. And, we have:

Theorem 3.2 *Assume that there is some polynomial $u(x) : \mathbb{R}^n \to \mathbb{R}$ of the form (13) and with $\{u(x) \geq 0\}$ compact. Then, as $i \to \infty$,*

$$
\inf \mathbb{Q}_i \uparrow \min \mathbb{P}. \tag{14}
$$

If $\mathbb{K}$ has a nonempty interior, then as $i \to \infty$,

$$
\max \mathbb{Q}_i^* \,=\, \inf \mathbb{Q}_i \uparrow \min \mathbb{P}. \tag{15}
$$

The equality $\max \mathbb{Q}_i^* = \min \mathbb{P}$ *occurs for all $i \geq i_0$ (for some index i_0) if and only if $g_0(x) - p^*$ is of the form (13).*

Proof: Let $\varepsilon > 0$ be fixed arbitrary. Then, the polynomial $x \mapsto G_0(x) := g_0(x) - p^* + \varepsilon$ is strictly positive on K. From the assumption on $\mathbb{K}$, it follows that $G_0(x)$ has the representation

$$
G_0(x) \,=\, \sum_{j=1}^{r_0} q_j(x)^2 + \sum_{k=1}^{m} g_k(x) \left[\sum_{j=1}^{r_k} q_{kj}(x)^2 \right] \tag{16}
$$

for some polynomials $q_j(x)$, $j = 1, \ldots r_0$, and $q_{kj}(x)$, $j = 1, \ldots r_k$, $k = 1, \ldots m$ (see e.g. Putinar [10], Jacobi [5]). Now, let $i_1(\varepsilon)$ (that for notational convenience we simply write i_1) be the maximum degree of $q_j(x)$, $j = 1, \ldots r_0$, and

$i_2(\varepsilon)$ (noted i_2) be the maximum degree of the polynomials $\{q_{kj}(x)\}$, so that the polynomials $\{g_k(x)q_{kj}(x)^2\}$ have maximum degree $2i_2+2$ (as the $g_k(x)$'s are quadratic polynomials). Let $i := \max[i_1, i_2+1]$. Let $q_j \in \mathbb{R}^{s(i)}$ and $q_{kj} \in \mathbb{R}^{s(i-1)}$ be the respective vectors of coefficients of the polynomials $\{q_j(x)\}$ and $\{q_{kj}(x)\}$, and write

$$X := \sum_{j=1}^{r_0} q_j q_j', \;\; Z_k := \sum_{j=1}^{r_k} q_{kj} q_{kj}', \; k = 1, \ldots m. \qquad (17)$$

Observe that with

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

then,

$$g_k(x)q_{kj}(x)^2 = \langle q_{kj}, M_{i-1}(g_k y^x) q_{kj}\rangle = \langle q_{kj}q_{kj}', M_{i-1}(g_k y^x)\rangle,$$

so that

$$\begin{aligned} g_k(x) \sum_{j=1}^{r_k} q_{kj}(x)^2 &= \langle Z_k, M_{i-1}(g_k y^x)\rangle \qquad (18) \\ &= \sum_\alpha x^\alpha \langle Z_k, C_\alpha^k\rangle, \end{aligned}$$

and similarly,

$$\sum_{j=1}^{r_0} q_j(x)^2 = \sum_\alpha x^\alpha \langle X, B_\alpha\rangle, \qquad (19)$$

with B_α, C_α^k as in $\mathbb{Q}_i^*$.

Therefore, from the representation (16), it follows that

$$\sum_\alpha x^\alpha \left[\langle X, B_\alpha\rangle + \sum_{k=1}^m \langle Z_k, C_\alpha^k\rangle\right] = \sum_\alpha x^\alpha (g_0)_\alpha - p^* + \varepsilon. \quad (20)$$

Identifying terms of same power yields

$$\langle X, B_\alpha\rangle + \sum_{k=1}^m \langle Z_k, C_\alpha^k\rangle = (g_0)_\alpha \;\; \forall \alpha \neq 0,$$

and, for the constant term,

$$X(1,1) + \sum_{k=1}^m Z_k(1,1) g_k(0) = -p^* + \varepsilon,$$

which proves that $(X, Z_1 \ldots Z_m)$ is admissible for $\mathbb{Q}_i^*$ with value $p^* - \varepsilon$. Therefore, $p^* - \varepsilon \leq \inf \mathbb{Q}_i \leq p^*$, and from the monotonicity and the fact that ε was arbitrary, $\inf \mathbb{Q}_i \uparrow p^* = \min \mathbb{P}$.

The statement (15) follows from a standard result in convexity. Indeed, we have just proved that $\mathbb{Q}_i^*$ has a feasible solution $(X, Z_1, \ldots Z_m)$. Moreover, if $\mathbb{K}$ has a nonempty interior, let μ be a probability measure with support $\mathbb{K}$ and with a uniform distribution. Denote by y^μ the vector of all the (well-defined) moments of μ. From (5)-(6) and the fact that μ has a uniform distribution on $\mathbb{K}$, it follows easily that $M_i(y^\mu) \succ 0$ and $M_{i-1}(g_k y^\mu) \succ 0$ for all $k = 1, \ldots m$, so that y^μ is a strictly admissible solution. Therefore, there is no duality gap between $\mathbb{Q}_i^*$ and $\mathbb{Q}_i$, and (15) follows.

Finally, consider the *if part* of the final statement. From the representation (13) of $g_0(x) - p^*$ and using the same above arguments, the corresponding matrices $X, Z_1, \ldots Z_m$ form an admissible solution for $\mathbb{Q}_i^*$ with value $p^* = \min \mathbb{P}$, hence proving $\max \mathbb{Q}_i^* = \min \mathbb{P}$.

Conversely, the *only if part* of the final statement also follows easily from the above proof. Let $(X, Z_1, \ldots Z_m)$ be an optimal solution of $\mathbb{Q}_i^*$ with $\max \mathbb{Q}_i^* = p^* = \min \mathbb{P}$. As $X, Z_1, \ldots Z_m \succeq 0$, write them as in (17). The admissiblity of $(X, Z_1, \ldots Z_m)$ implies (20) with p^* in lieu of $p^* - \varepsilon$, which in turn, using (18)-(19) yields the desired result. ∎

The representation (13) has a nice interpretation as a **global optimality condition** à la Karush-Kuhn-Tucker. Indeed,

Proposition 3.3 *Assume that $g_0(x) - p^*$ has the representation (13). Then, at a global minimizer x^*, we have*

$$\nabla g_0(x^*) = \sum_{k=1}^m \nabla g_k(x^*) \left[\sum_{j=1}^{r_k} q_{kj}(x^*)^2\right] \qquad (21)$$

and

$$g_k(x^*) \left[\sum_{j=1}^{r_k} q_{kj}(x^*)^2\right] = 0, \;\; k = 1, \ldots m. \qquad (22)$$

Proof: The proof is immediate. As $g_0(x^*) - p^* = 0$, the representation (13) leads

$$\sum_{j=1}^{r_0} q_j(x^*)^2 = 0; \; \sum_{k=1}^m g_k(x^*) \left[\sum_{j=1}^{r_k} q_{kj}(x^*)^2\right] = 0,$$

so that (22) follows. Differentiating in (13) and using the above relationship also yields (21). ∎

Hence, the theory of representation of polynomials, positive on the feasible set $\mathbb{K}$, can be viewed as a global optimality condition. The polynomials (sums of squares) $\sum_j q_{kj}(x)^2$ in the representation (13) of $g_0(x) - p^*$ (when it holds) are nothing less than *Lagrange Karush-Kuhn-Tucker* "multipliers". In contrast to the usual Karush-Kuhn-Tucker necessary optimality condition with "scalar multipliers" λ_k, a non-active constraint $g_k(x)$ at a global minimizer x^* may have an nontrivial associated "polynomial" multiplier (when this constraint plays a role to eliminate some better (non feasible) solutions. However, this polynomial multiplier vanishes at x^*, as in the usual Karush-Kuhn-Tucker conditions where $\lambda_k = 0$ at a non-active constraint $g_k(x^*) > 0$.

4 Examples

For illustration purposes we will consider the MAX-CUT problem and some nontrivial nonconvex quadratic test problems from Floudas and Pardalos [9].

4.1 The max-cut problem

Roughly speaking, the MAX-CUT problem is a special case of (1) with

- $g_0(x) = x'A_0x$ and $\mathrm{diag}(A_0) = 0$.
- $g_i(x) = x_i^2 - 1$ and the constraint is an equality constraint.

As for 0-1 programs, the feasible set $\mathbb{K}$ satisfies the condition required in Theorem 3.2. Indeed, write the integral constraint $x_i^2 = 1$ as $x_i^2 - 1 \geq 0$ and $1 - x_i^2 \geq 0$ for all $i = 1, \ldots n$, and consider the polynomial $u(x) := \sum_{i=1}^n (1 - x_i^2)$. Its level set $\{u(x) \geq 0\}$ is compact. Shor's relaxation (equivalently, $\mathbb{Q}_1$), is the convex LMI problem:

$$\mathbb{Q}_1 \left\{ \begin{array}{l} \min \langle Y, A_0 \rangle \\ \\ \begin{bmatrix} 1 & | & y' \\ - & - & - \\ y & | & Y \end{bmatrix} \succeq 0; \ \ \mathrm{diag}(Y) = e \end{array} \right. \tag{23}$$

where e is a vector of ones. To visualize the two relaxations $\mathbb{Q}_1$ and $\mathbb{Q}_2$, take an example in $\mathbb{R}^3$. In that case, $\mathbb{Q}_1$ reads

$$\mathbb{Q}_1 \left\{ \begin{array}{l} \min q_{110}y_{110} + q_{101}y_{101} + q_{011}y_{011} \\ \\ \begin{bmatrix} 1 & y_{100} & y_{010} & y_{001} \\ y_{100} & 1 & y_{110} & y_{101} \\ y_{010} & y_{110} & 1 & y_{011} \\ y_{001} & y_{101} & y_{011} & 1 \end{bmatrix} \succeq 0, \end{array} \right.$$

whereas $\mathbb{Q}_2$ reads

$$\mathbb{Q}_2 \left\{ \begin{array}{l} \min q_{110}y_{110} + q_{101}y_{101} + q_{011}y_{011} \\ \\ \begin{bmatrix} M_1(y) & | & B \\ - & & - \\ B' & | & C \end{bmatrix} \succeq 0, \end{array} \right.$$

with

$$B = \begin{bmatrix} 1 & y_{110} & y_{101} & 1 & y_{011} & 1 \\ y_{100} & y_{010} & y_{001} & y_{100} & y_{111} & y_{100} \\ y_{010} & y_{100} & y_{111} & y_{110} & y_{001} & y_{010} \\ y_{001} & y_{111} & y_{100} & y_{001} & y_{010} & y_{001} \end{bmatrix}$$

and

$$C = \begin{bmatrix} 1 & y_{110} & y_{101} & 1 & y_{011} & 1 \\ y_{110} & 1 & y_{011} & y_{110} & y_{101} & y_{110} \\ y_{101} & y_{011} & 1 & y_{101} & y_{110} & y_{101} \\ 1 & y_{110} & y_{101} & 1 & y_{011} & 1 \\ y_{011} & y_{101} & y_{110} & y_{011} & 1 & y_{011} \\ 1 & y_{110} & y_{101} & 1 & y_{011} & 1 \end{bmatrix}.$$

where we have used the fact that the equality $x_{100}^2 = 1$ translates into $y_{200} = 1$ and therefore, a term like y_{400} is replaced by 1, a term like y_{210} is replaced by y_{010}, y_{211} by y_{011}, etc. so that only the variables y_α with $\alpha_i \leq 1$ and $\sum_i \alpha_i \leq 4$ are present.

Hence, the $\mathbb{Q}_2$ relaxation of a MAX-CUT problem of dimension n is an LMI problem with $\binom{n+1}{4} + \binom{n+1}{2}$ variables and one LMI constraint of dimension $(n+1)(n+2)/2$ in comparison with $\binom{n+1}{2}$ variables and one LMI constraint of dimension $n+1$ for the $\mathbb{Q}_1$ relaxation. Of course, as i increases, then so does the computational burden, a price to pay to obtain better and better lower bounds.

For a sample of 50 randomly generated problems in $\mathbb{R}^{10}$ with positive and negative weights, the $\mathbb{Q}_2$ relaxation provided the optimal solution in all cases whereas the $\mathbb{Q}_1$ relaxation in less than 20% of the cases only.

In both relaxations $\mathbb{Q}_1$ and $\mathbb{Q}_2$, Slater's interior point condition fails, and we have to mention that in using the MATLAB LMI toolbox, the running time for $\mathbb{Q}_2$ was surprisingly very high compared to $\mathbb{Q}_1$, despite the fact that $\mathbb{Q}_2$ contains relatively few variables (15 for $\mathbb{R}^4$ and 30 for $\mathbb{R}^5$, 375 fro $\mathbb{R}^{10}$).

4.2 Nonconvex quadratic problems

The nonconvex quadratic test problems below are from Floudas and Pardalos [9].

4.2.1 : Problem 2.2 in Floudas and Pardalos [9].

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

with $Q := I$ and $c = [-10.5, -7.5, -3.5, -2.5, -1.5]$. The optimal value -213 is obtained at the $\mathbb{Q}_2$ relaxation.

4.2.2 : Problem 2.6 in Floudas and Pardalos [9].

$$\left\{ \begin{array}{l} \min_{x} f(x) := c^T x - 0.5x^T Qx \\ Ax \leq b \\ 0 \leq x_i \leq 1, i = 1, \ldots, 10 \\ y_5 \leq 2 \end{array} \right.$$

with A being the matrix

$$\begin{bmatrix} -2 & -6 & -1 & 0 & -3 & -3 & -2 & -6 & -2 & -2 \\ 6 & -5 & 8 & -3 & 0 & 1 & 3 & 8 & 9 & -3 \\ -5 & 6 & 5 & 3 & 8 & -8 & 9 & 2 & 0 & -9 \\ 9 & 5 & 0 & -9 & 1 & -8 & 3 & -9 & -9 & -3 \\ -8 & 7 & -4 & -5 & -9 & 1 & -7 & -1 & 3 & -2 \end{bmatrix}$$

$c = [48, 42, 48, 45, 44, 41, 47, 42, 45, 46]$, $Q = 100I$ and $b = [-4, 22, -6, -23, -12]$. The optimal value -39 is obtained at the $\mathbb{Q}_2$ relaxation.

4.2.3 : Problem 2.9 in Floudas and Pardalos [9].

$$\left\{ \begin{array}{l} \max_{x} f(x) := \sum_{i=1}^{9} x_i x_{i+1} + \sum_{i=1}^{8} x_i x_{i+2} \\ +x_1x_7 + x_1x_9 + x_1x_{10} + x_2x_{10} + x_4x_7 \\ \sum_{i=1}^{10} x_i = 1 \\ x_i \geq 0, i = 1, \ldots 10. \end{array} \right.$$

The optimal value 0.375 is obtained at the $\mathbb{Q}_2$ relaxation.

4.2.4 : Problem 3.3 in Floudas and Pardalos [9].

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

The optimal value -310 is obtained at the $\mathbb{Q}_2$ relaxation.

4.2.5 : Problem 3.4 in Floudas and Pardalos [9].

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

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 $\mathbb{Q}_4$ relaxation whereas $\inf \mathbb{Q}_3 = -4.0685$.

5 Conclusion

We have proposed a sequence of LMI relaxations $\{\mathbb{Q}_i\}$ and an associated sequence of nondecreasing lower bounds that converges to the global minimum p^*, and in many cases, p^* is obtained at a particular relaxation. It has been shown that the primal and dual psd programs perfectly match both sides of the same theory (moments and positive polynomials). Moreover, the representation of polynomials, positive on the feasible set $\mathbb{K}$, is interpreted as a Karush-Kuhn-Tucker global optimality condition with “polynomial” Lagrange multipliers instead of scalar multipliers. However, although efficient LMI software packages are now available, high order relaxations require a lot of variables. It is hoped that for a large class of problems, low order relaxations like $\mathbb{Q}_2$, $\mathbb{Q}_3$ or $\mathbb{Q}_4$ will provide the optimal value, or at least, a good lower bound that could be exploited in other optimization methods.

References

[1] C. Berg, *The multidimensional moment problem and semi-groups*, in: Moment in Mathematics (1980), pp. 110–124, ed. H.J. Landau, Amer. Math. Soc.

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

[3] C. Ferrier, *Hilbert's 17th problem and best dual bounds in quadratic minimization*, Cybern. Syst. Anal. **34** (1998), pp. 696–709.

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

[5] T. Jacobi, *A representation theorem for certain partially ordered commutative rings*, Mathematische Zeitschrift, to appear.

[6] T. Jacobi and A. Prestel, *On special representations of strictly positive polynomials*, Technical-report, Konstanz University, January 2000.

[7] J.B. Lasserre, *The global minimization of a polynomial is an easy convex problem*, LAAS Technical report #99485, Toulouse, December 1999.

[8] J.B. Lasserre, *Global minimization with polynomials and the problem of moments*, LAAS Technical report #00025, Toulouse, January 2000, submitted.

[9] C. Floudas and P.M. Pardalos, *A Collection of Test Problems for Constrained Global Optimization Algorithms*, Springer-Verlag, Berlin, 1990.

[10] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Ind. Univ. Math. J. **42** (1993), pp. 969–984.

[11] K. Schmüdgen, *The K-moment problem for compact semi-algebraic sets*, Math. Ann. **289** (1991), pp. 203-206.

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

[13] N.Z. Shor, *Quadratic optimization problems*, Tekhnicheskaya Kibernetika **1** (1987), pp. 128-139.

[14] N.Z. Shor, *Nondifferentiable Optimization and Polynomial Problems*, Kluwer Academic Publishers, Dordrecht, 1998.

[15] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Review **38** (1996), pp. 49-95.