



**HAL**  
open science

# A MAX-CUT FORMULATION OF 0/1 PROGRAMS

Jean-Bernard Lasserre

► **To cite this version:**

Jean-Bernard Lasserre. A MAX-CUT FORMULATION OF 0/1 PROGRAMS. 2015. hal-01154698v1

**HAL Id: hal-01154698**

**<https://hal.science/hal-01154698v1>**

Preprint submitted on 22 May 2015 (v1), last revised 22 Dec 2015 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A MAX-CUT FORMULATION OF 0/1 PROGRAMS

JEAN B. LASSERRE

ABSTRACT. We consider the linear or quadratic 0/1 program

$$\mathbf{P} : f^* = \min\{\mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{A} \mathbf{x} = \mathbf{b}; \mathbf{x} \in \{0, 1\}^n\},$$

for some vectors  $\mathbf{c} \in \mathbb{R}^n$ ,  $\mathbf{b} \in \mathbb{Z}^m$ , some matrix  $\mathbf{A} \in \mathbb{Z}^{m \times n}$  and some real symmetric matrix  $\mathbf{F} \in \mathbb{R}^{n \times n}$ . We show that  $\mathbf{P}$  can be formulated as a MAX-CUT problem whose quadratic form criterion is explicit from the data of  $\mathbf{P}$ . In particular, to  $\mathbf{P}$  one may associate a graph  $(G, V)$  whose connectivity is related to the connectivity of the matrix  $\mathbf{F}$  and  $\mathbf{A}^T \mathbf{A}$ , and  $\mathbf{P}$  reduces to finding a maximum (weighted) cut in such a graph. Hence the whole arsenal of approximation techniques for MAX-CUT can be applied. On a sample of 0/1 knapsack problems, we compare the lower bound on  $f^*$  of the associated standard (Shor) SDP-relaxation with the standard linear relaxation where  $\{0, 1\}^n$  is replaced with  $[0, 1]^n$  (resulting in an LP when  $\mathbf{F} = 0$  and a quadratic program when  $\mathbf{F}$  is positive definite). We also compare our lower bound with that of the first SDP-relaxation associated with the copositive formulation of  $\mathbf{P}$ .

## 1. INTRODUCTION

Consider the linear or quadratic 0/1 program  $\mathbf{P}$  defined by:

$$(1.1) \quad \mathbf{P} : f^* = \min_{\mathbf{x}} \{ \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{A} \mathbf{x} = \mathbf{b}; \mathbf{x} \in \{0, 1\}^n \}$$

for some cost vector  $\mathbf{c} \in \mathbb{R}^n$ , some matrix  $\mathbf{A} \in \mathbb{Z}^{m \times n}$ , some vector  $\mathbf{b} \in \mathbb{Z}^m$ , and some real symmetric matrix  $\mathbf{F} \in \mathbb{R}^{n \times n}$ . If  $\mathbf{F} = 0$  then  $\mathbf{P}$  is a 0/1 linear program and a quadratic 0/1 program otherwise. Obtaining good quality lower bounds on  $f^*$  is highly desirable since the efficiency of Branch & Bound algorithms to solve large scale problems  $\mathbf{P}$  heavily depends on the quality of bounds of this form computed at nodes of the search tree.

To obtain lower bounds for 0/1 programs (1.1) one may solve a relaxation of  $\mathbf{P}$  where the integrality constraints  $\mathbf{x} \in \{0, 1\}^n$  are replaced with the box constraints  $\mathbf{x} \in [0, 1]^n$ . If  $\mathbf{F} = 0$  the resulting relaxation is linear whereas if  $\mathbf{F}$  is positive definite it is a (convex) quadratic program. If  $\mathbf{F}$  is not positive semidefinite then one may also solve a convex quadratic program but now with an appropriate convex quadratic underestimator  $\mathbf{x}^T \hat{\mathbf{F}} \mathbf{x}$  of  $\mathbf{x}^T \mathbf{F} \mathbf{x}$  on  $[0, 1]^n$ . An alternative is to consider an equivalent formulation of  $\mathbf{P}$  as a copositive conic program as advocated by Burer [3] and compute

---

*Date:* May 22, 2015.

Research partially supported by a grant from the MONGE program of the *Fédération Mathématique Jacques Hadamard* (FMJH, Paris).

a sequence of lower bounds by solving an appropriate hierarchy of LP- or SDP-relaxations associated with the copositive cone (or its dual). For more details on the latter approach the interested reader is referred to e.g. De Klerk and Pasechnik [4], Dürr [5], Bomze [1], and Bomze and de Klerk [2].

**Contribution.** The purpose of this note is to show that solving  $\mathbf{P}$  is equivalent to minimizing a quadratic form in  $n+1$  variables on the hypercube  $\{-1, 1\}^{n+1}$  (and the quadratic form is explicit from the data of  $\mathbf{P}$ ). In other words  $\mathbf{P}$  can be viewed as an explicit instance of the MAX-CUT problem. Hence the MAX-CUT problem which at first glance seems to be a very specific combinatorial optimization problem, in fact can be considered as a canonical model of linear and quadratic 0/1 programs. In particular, to each linear or quadratic 0/1 program (1.1) one may associate a graph  $(G, V)$  with  $n+1$  nodes and  $(i, j) \in V$  whenever a product  $x_i x_j$  has a nonzero coefficient in some quadratic form built upon the data  $\mathbf{c}, \mathbf{b}, \mathbf{F}$  and  $\mathbf{A}$  of (1.1). (Among other things, the sparsity of  $(G, V)$  is related to the sparsity of the matrices  $\mathbf{F}$  and  $\mathbf{A}^T \mathbf{A}$ .) Then solving (1.1) reduces to finding a maximum (weighted) cut of  $G$ .

Therefore the whole specialized arsenal of approximation techniques for MAX-CUT can be applied. In particular one may obtain a lower bound  $f_1^*$  on  $f^*$  by solving the standard (Shor) SDP-relaxation associated with the resulting MAX-CUT problem while solving higher levels of the associated Lasserre-SOS hierarchy [6, 7] would provide a monotone nondecreasing sequence of improved lower bounds  $f_1^* \leq f_d^* \leq f^*$ ,  $d = 2, \dots$ , but of course at a higher computational cost. Alternatively one may also apply the Handelman hierarchy of LP-relaxations as described and analyzed in Laurent and Sun [10]. For more details on recent developments on computational approaches to MAX-CUT the interested reader is referred to Wigele and Rendl [13]. If  $\mathbf{F} = 0$  (i.e. when  $\mathbf{P}$  is a linear 0/1 program) the lower bound  $f_1^*$  can be better than the standard LP-relaxation which consists in replacing the integrality constraints  $\mathbf{x} \in \{0, 1\}^n$  with the box  $[0, 1]^n$ , as shown on a (limited) sample of 0/1-knapsack-type examples. On such examples  $f_1^*$  also dominates the one obtained from the first relaxation of the copositive formulation (where the dual cone  $\mathcal{C}^*$  of completely positive matrices is replaced with  $\mathcal{S}^+ \cap \mathcal{N} \supset \mathcal{C}^*$ ) in about 55% of cases and the maximum relative difference is bounded by 0.55% in all cases.

In addition one may also obtain performance guarantees *à la Nesterov* [12] in the form

$$f_1^* \leq f^* \leq \frac{2}{\pi} f_1^* + \left(1 - \frac{2}{\pi}\right) h_1^*,$$

(where  $h_1^*$  is the optimal value of a similar SDP but with a max-criterion instead of a min-criterion) or their improvements by Marshall [11].

In fact, and still on the same sample of linear and quadratic 0/1 knapsack examples, one also observes that the resulting lower bound  $f_1^*$  is almost always better than the lower bound obtained by solving the first SDP-relaxation of the Lasserre-SOS hierarchy applied to the initial formulation

(1.1) of the problem (which is also an SDP of same size). This is good news since typically the SOS-hierarchy is known to produce good lower bounds for general polynomial optimization problems (discrete or not) even at the first level of the hierarchy. Even more, the first level SDP-relaxation has the celebrated Goemans & Williamson performance guarantee ( $\approx 87\%$ ) when the matrix  $\mathbf{Q}$  (associated with the quadratic form) has nonnegative entries and a performance guarantee  $\approx 64\%$  when  $\mathbf{Q} \succeq 0$ . (However note that the matrix  $\mathbf{Q}$  associated with our MAX-CUT problem equivalent to the initial 0/1 program (1.1) does not have all its entries nonnegative.) This explains why in the linear 0/1 knapsack examples the lower bound  $f_1^*$  is almost always better than the one obtained with the standard LP-relaxation and why for quadratic 0/1 knapsack problems (1.1),  $f_1^*$  is also likely to be better than the lower bound obtained by relaxing  $\{0, 1\}^n$  to  $[0, 1]^n$ , replacing  $\mathbf{F}$  with a convex quadratic underestimator of  $\mathbf{F}$  on  $[0, 1]^n$ , and solving the resulting convex quadratic program.

Finally, the same methodology also works for general 0/1 optimization problems with feasible set as in (1.1) and polynomial criterion  $f \in \mathbb{R}[\mathbf{x}]$  of degree  $d > 2$ , except that now the problem reduces to minimizing a new criterion  $\tilde{f}(\mathbf{x})$  on the hypercube  $\{-1, 1\}^n$ .

## 2. MAIN RESULT

Denote by  $\mathbb{Z}$  the set of integer numbers and  $\mathbb{N} \subset \mathbb{Z}$  the set of natural numbers. Let  $\mathbf{P}$  be the 0/1 program defined in (1.1) with  $\mathbf{F}^T = \mathbf{F} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{A} \in \mathbb{Z}^{m \times n}$ ,  $\mathbf{c} \in \mathbb{R}^n$  and  $\mathbf{b} \in \mathbb{Z}^m$ . Let  $|\mathbf{c}| := (|c_i|) \in \mathbb{R}_+^n$ . With  $\mathbf{e} \in \mathbb{Z}^n$  being the vector of all ones, notice first that  $\mathbf{P}$  has an equivalent formulation on the hypercube  $\{-1, 1\}^n$ , by the change of variables  $\tilde{\mathbf{x}} := 2\mathbf{x} - \mathbf{e}$ . Indeed,  $\mathbf{A}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$  and  $\mathbf{F}$  now become  $\mathbf{A}/2$ ,  $\mathbf{b} - \mathbf{A}\mathbf{e}/2$ ,  $(\mathbf{c} + \mathbf{e}^T\mathbf{F})/2$  and  $\mathbf{F}/4$  respectively.

Therefore from now on we consider the discrete program:

$$(2.1) \quad \mathbf{P} : \quad f^* = \min_{\mathbf{x} \in \{-1, 1\}^n} \{ \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{A} \mathbf{x} = \mathbf{b} \},$$

on the hypercube  $\{-1, 1\}^n$ , with  $\mathbf{A} \in \mathbb{Z}^{m \times n}$ ,  $\mathbf{c} \in \mathbb{R}^n$ ,  $\mathbf{b} \in \mathbb{Z}^m$ , and  $\mathbf{F}^T = \mathbf{F} \in \mathbb{R}^{n \times n}$ . With  $\mathbf{c}$  and  $\mathbf{F}$ , let us associate the scalars:

$$r_{\mathbf{c}, \mathbf{F}}^1 = \min \{ \mathbf{c}^T \mathbf{x} + \langle \mathbf{X}, \mathbf{F} \rangle : \begin{bmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{X} \end{bmatrix} \succeq 0; X_{ii} = 1, i = 1, \dots, n \}$$

$$r_{\mathbf{c}, \mathbf{F}}^2 = \max \{ \mathbf{c}^T \mathbf{x} + \langle \mathbf{X}, \mathbf{F} \rangle : \begin{bmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{X} \end{bmatrix} \preceq 0; X_{ii} = 1, i = 1, \dots, n \}$$

(with  $\mathbf{X}^T = \mathbf{X}$ ) and let

$$(2.2) \quad \rho(\mathbf{c}, \mathbf{F}) := \max_{i=1,2} |r_{\mathbf{c}, \mathbf{F}}^i|.$$

It is straightforward to verify that

$$\rho(\mathbf{c}, \mathbf{F}) \geq \max \{ |\mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x}| : \mathbf{x} \in \{-1, 1\}^n \},$$

and  $\rho(\mathbf{c}, \mathbf{F}) = |\mathbf{c}|$  if  $\mathbf{F} = 0$ . Moreover each scalar  $r_{\mathbf{c}, \mathbf{F}}^i$  can be computed by solving an SDP which is the Shor relaxation (or first level of the Lasserre-SOS hierarchy [6, 7]) associated with the problems  $\min (\max) \{ \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{x} \in \{-1, 1\}^n \}$ .

### 2.1. A MAX-CUT formulation of $\mathbf{P}$ .

**Lemma 2.1.** *Let  $\mathbf{P}$  be as (2.1) and let  $\rho(\mathbf{c}, \mathbf{F})$  be as in (2.2). Then  $f^*$  is the optimal value of the quadratic minimization problem:*

$$(2.3) \quad \min_{\mathbf{x} \in \{-1, 1\}^n} \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} + (2\rho(\mathbf{c}, \mathbf{F}) + 1) \cdot \|\mathbf{A} \mathbf{x} - \mathbf{b}\|^2.$$

*Proof.* Let  $\Delta := \{ \mathbf{x} \in \{-1, 1\}^n : \mathbf{A} \mathbf{x} = \mathbf{b} \}$  be the feasible set of  $\mathbf{P}$  defined in (1.1), and let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be the function

$$(2.4) \quad x \mapsto f(\mathbf{x}) := \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} + (2\rho(\mathbf{c}, \mathbf{F}) + 1) \cdot \|\mathbf{A} \mathbf{x} - \mathbf{b}\|^2.$$

On  $\{-1, 1\}^n$  one has  $\max \{ \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{x} \in \{-1, 1\}^n \} \leq \rho(\mathbf{c}, \mathbf{F})$ , and

$$\|\mathbf{A} \mathbf{x} - \mathbf{b}\|^2 \geq 1, \quad \forall \mathbf{x} \in \{-1, 1\}^n \setminus \Delta,$$

because  $\mathbf{A} \in \mathbb{Z}^{m \times n}$  and  $\mathbf{b} \in \mathbb{Z}^m$ . Therefore,

$$f(\mathbf{x}) \begin{cases} = \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} & \text{on } \Delta \\ \geq \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} + 2\rho(\mathbf{c}, \mathbf{F}) + 1 > \rho(\mathbf{c}, \mathbf{F}) & \text{on } \{-1, 1\}^n \setminus \Delta. \end{cases}$$

From this and  $\mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} < \rho(\mathbf{c}, \mathbf{F})$  on  $\Delta$ , the result follows.  $\square$

Next, let  $Q : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  be the homogenization of the quadratic polynomial  $f$ , i.e., the quadratic form  $Q(\mathbf{x}, x_0) := x_0^2 f(\frac{\mathbf{x}}{x_0})$ , or in explicitly form:

$$(2.5) \quad Q(\mathbf{x}, x_0) = x_0 \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} + (2\rho(\mathbf{c}, \mathbf{F}) + 1) \cdot \|\mathbf{A} \mathbf{x} - x_0 \mathbf{b}\|^2.$$

Observe that  $Q(\mathbf{x}, 1) = f(\mathbf{x})$ .

**Theorem 2.2.** *Let  $f^* = \min \{ \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{A} \mathbf{x} = \mathbf{b}; \mathbf{x} \in \{-1, 1\}^n \}$  and let  $Q$  be the quadratic form in (2.5). Then*

$$(2.6) \quad f^* = \min_{(\mathbf{x}, x_0) \in \{-1, 1\}^{n+1}} Q(\mathbf{x}, x_0),$$

*that is,  $f^*$  is the optimal value of the MAX-CUT problem associated with the quadratic form  $Q$ .*

*Proof.* Let  $f$  be as in (2.4). By definition of  $Q$ ,

$$(2.7) \quad \min_{\mathbf{x} \in \{-1, 1\}^n} f(\mathbf{x}) = \min_{(\mathbf{x}, x_0) \in \{-1, 1\}^{n+1}} \{ Q(\mathbf{x}, x_0) : x_0 = 1 \}.$$

On the other hand, let  $(\mathbf{x}^*, x_0^*) \in \{-1, 1\}^{n+1}$  be a global minimizer of  $\min \{ Q(\mathbf{x}, x_0) : (\mathbf{x}, x_0) \in \{-1, 1\}^{n+1} \}$ . Then by homogeneity of  $Q$ ,  $(-\mathbf{x}^*, -x_0^*)$  is also a global minimizer and so one may decide arbitrarily to fix  $x_0 = 1$ . That is,

$$\min_{(\mathbf{x}, x_0) \in \{-1, 1\}^{n+1}} Q(\mathbf{x}, x_0) = \min_{(\mathbf{x}, x_0) \in \{-1, 1\}^{n+1}} \{ Q(\mathbf{x}, x_0) : x_0 = 1 \},$$

which combined with (2.7) yields the desired result.  $\square$

Next, write  $Q(\mathbf{x}, x_0) = (\mathbf{x}, x_0)\mathbf{Q}(\mathbf{x}, x_0)^T$  for an appropriate real symmetric matrix  $\mathbf{Q} \in \mathbb{R}^{(n+1) \times (n+1)}$ , and introduce the semidefinite programs

$$(2.8) \quad \min_{\mathbf{X}} \{ \langle \mathbf{Q}, \mathbf{X} \rangle : \mathbf{X} \succeq 0; X_{ii} = 1, \quad i = 1, \dots, n+1 \}$$

with optimal value denoted by  $\min \mathbf{Q}_+$ , and

$$(2.9) \quad \max_{\mathbf{X}} \{ \langle \mathbf{Q}, \mathbf{X} \rangle : \mathbf{X} \succeq 0; X_{ii} = 1, \quad i = 1, \dots, n+1 \}$$

with optimal value denoted by  $\max \mathbf{Q}^+$ .

**Proposition 2.3.** *Let  $\mathbf{P}$  be the problem defined in (2.1) with optimal value  $f^*$ . Then*

$$(2.10) \quad \min \mathbf{Q}_+ \leq f^* \leq \frac{2}{\pi} \min \mathbf{Q}_+ + \left(1 - \frac{2}{\pi}\right) \max \mathbf{Q}^+$$

where  $\mathbf{Q}$  is the real symmetric matrix associated with the quadratic form (2.5) and  $\min \mathbf{Q}_+$  (resp.  $\max \mathbf{Q}^+$ ) is the optimal value of the semidefinite program (2.8) (resp. (2.9)).

*Proof.* The bounds in (2.10) are from Nesterov [12]. In addition, one may also use the bounds provided in Marshall [11] which sometimes improve those in (2.10).  $\square$

The quality of the upper bound in (2.10) depends strongly on the magnitude of the “penalty coefficient”  $2\rho(\mathbf{c}, \mathbf{F}) + 1$  in the definition of the function  $f$  in (2.4). However for a practical use of relaxations what matters most is the quality of the lower bound  $\min \mathbf{Q}_+$  which in principle is very good for MAX-CUT problems (even if  $\mathbf{Q} \not\geq 0$  or  $\mathbf{Q} \not\leq 0$ ). For instance in a Branch & Bound algorithm the lower bound  $\min \mathbf{Q}_+$  has an important impact in the pruning of nodes in the search tree.

**Sparsity.** Hence to each 0/1 program (1.1) one may associate a graph  $(G, V)$  with  $n+1$  nodes and an arc  $(i, j) \in V$  connects the nodes  $i, j \in G$  if and only if the coefficient  $\mathbf{Q}_{ij}$  of the quadratic form  $Q(\mathbf{x}, x_0)$  does not vanish. *Sparsity* properties of  $(G, V)$  are of primary interest, e.g. for computational reasons. From the definition of the matrix  $\mathbf{Q}$ , this sparsity is in turn related to sparsity of the matrix  $\mathbf{F} + (2\rho_{\mathbf{c}, \mathbf{F}} + 1) \cdot \mathbf{A}^T \mathbf{A}$ , hence of sparsity of  $\mathbf{F}$  and  $\mathbf{A}^T \mathbf{A}$ . In particular two nodes  $i, j$  are not connected if  $\mathbf{F}_{ij} = 0$  and  $\mathbf{A}_{ki} \mathbf{A}_{kj} = 0$  for all  $k = 1, \dots, m$ .

**Example 2.4.** To evaluate the quality of the lower bound obtained with the MAX-CUT formulation consider the following simple linear knapsack-type examples:

$$(2.11) \quad \min \{ \mathbf{c}^T \mathbf{x} : \mathbf{a}^T \mathbf{x} = b; \mathbf{x} \in \{-1, 1\}^n \},$$

on  $\{-1, 1\}^n$ , with 4 and 10 variables. For  $n = 4$ ,  $\mathbf{c} = (13, 11, 7, 3)$  and  $\mathbf{a} = (3, 7, 11, 13)$ , while for  $n = 10$ ,  $\mathbf{c} = (37, 31, 29, 23, 19, 17, 13, 11, 7, 3)$ , and  $\mathbf{a} = (3, 7, 11, 13, 17, 19, 23, 29, 31, 37)$ .

The right-hand-side  $b$  is taken into  $[-|\mathbf{a}|, |\mathbf{a}|] \cap \mathbb{Z}$ . Figure 1 displays the difference  $\min \mathbf{Q}_+ - \min \text{LP}$  where the lower bound  $\min \text{LP}$  is obtained by relaxing the integrality constraints  $\mathbf{x} \in \{-1, 1\}^n$  to the box constraint  $\mathbf{x} \in [-1, 1]^n$  and solving the resulting LP. As expected the lower bound  $\min \mathbf{Q}_+$  is much better than  $\min \text{LP}$ . In fact the cases where the LP-bound is slightly better is for right-hand-side  $b$  such that the relaxation provides the optimal value  $f^*$ .

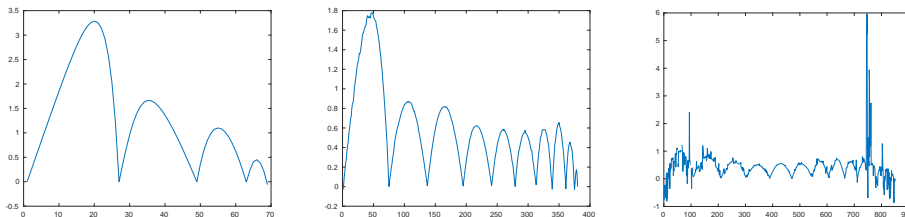


FIGURE 1. Difference  $\min \mathbf{Q}_+ - \min \text{LP}$  with  $n=(4,10,15)$

Moreover Figure 2 displays the difference  $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$  where  $\min \hat{\mathbf{Q}}_+$  is the optimal value of the first SDP-relaxation of the Lasserre-SOS hierarchy applied to the initial formulation (2.11) of the knapsack problem where one has even included the redundant constraints  $x_i (\mathbf{a}^T \mathbf{x} - b) = 0, i = 1, \dots, n$ . One observes that in most cases the lower bound  $\min \mathbf{Q}_+$  is slightly better than  $\min \hat{\mathbf{Q}}_+$ .

This is encouraging since the Lasserre-SOS hierarchy [6, 7] is known to produce good lower bounds in general, and especially at the first level of the hierarchy for MAX-CUT problems whose matrix  $\mathbf{Q}$  of the associated quadratic form has certain properties, e.g.,  $Q_{ij} \geq 0$  for all  $i, j$  or  $\mathbf{Q} \succeq 0$  (in the maximizing case); see e.g. Marshall [11].

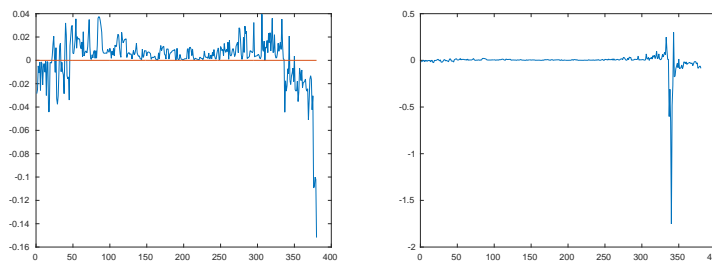


FIGURE 2. Difference  $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$  ( $n=10$ ) (left) and relative difference  $100 * (\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+) / \min \hat{\mathbf{Q}}_+$

**Example 2.5.** In a second sample of linear knapsack problems (2.11) with  $n = 10, 15$ , we have chosen the same vector  $\mathbf{a}$  as in Example 2.4 but now with

a cost criterion of the form  $c(i) = a(i) + s\eta$ ,  $i = 1, \dots, n$ , where  $\eta$  is a random variable uniformly distributed on  $[0, 1]$ , and  $s = 20, 10, 1$  is a weighting factor. The reason is that knapsack problems with ratios  $c(i)/a(i) \approx 1$  for all  $i$ , can be difficult to solve. As before the right-hand-side  $b$  is taken into  $[-|\mathbf{a}|, |\mathbf{a}|] \cap \mathbb{Z}$ . Figure 3 displays the results obtained for  $s = 20$ , and  $n = 10, 15$ . Figure 4 displays the same example for another sample with

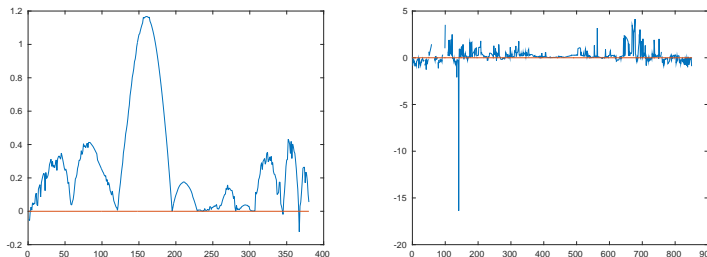


FIGURE 3. Difference  $\min \mathbf{Q}_+ - \min \text{LP}$ ,  $\mathbf{c} = \mathbf{a} + 20 * \eta$  ( $n=10,15$ )

cost  $\mathbf{c}$  and  $s = 10$ .

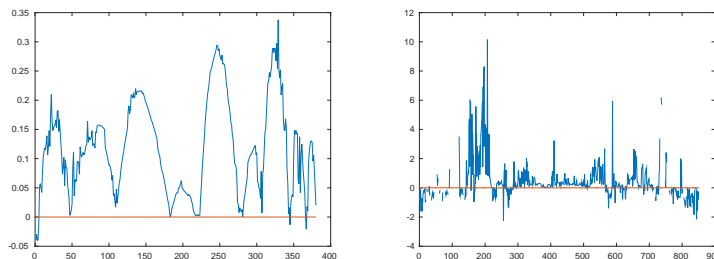


FIGURE 4. Difference  $\min \mathbf{Q}_+ - \min \text{LP}$ ,  $\mathbf{c} = \mathbf{a} + 10 * \eta$  ( $n=10,15$ )

Finally, as for Example 2.4, Figure 5 displays the difference  $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$  (where  $\min \hat{\mathbf{Q}}_+$  is the optimal value of the first SDP-relaxation of the Lasserre-SOS hierarchy applied to the initial formulation (2.11) of the knapsack problem where one has even included the redundant constraints  $x_i(\mathbf{a}^T \mathbf{x} - b) = 0$ ,  $i = 1, \dots, n$ ). Again one observes that in most cases the lower bound  $\min \mathbf{Q}_+$  is slightly better than  $\min \hat{\mathbf{Q}}_+$ .

**Example 2.6.** We next consider the same knapsack problems (2.11) as in Example 2.5 but now with quadratic criterion  $\mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x}$ , again with a cost criterion of the form  $c(i) = a(i) + s\eta$ ,  $i = 1, \dots, n$ , where  $\eta$  is a random variable uniformly distributed in  $[0, 1]$  and  $s$  is some weighting factor. The real symmetric matrix  $\mathbf{F}$  is also randomly generated and is not positive definite in general. Again in Figures 6 and 7 one observes that the lower



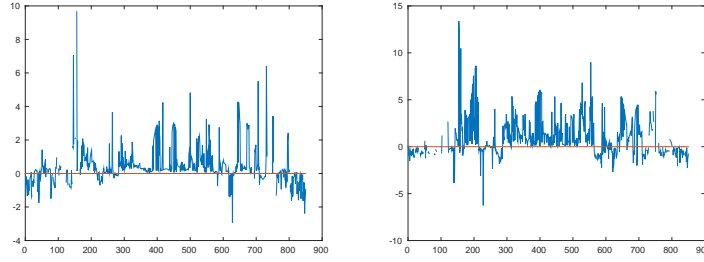


FIGURE 5. Example 2.5: Difference  $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$ ,  
( $n=15$ )  $\mathbf{c} = \mathbf{a} + 20\eta$  (left) and  $\mathbf{c} = \mathbf{a} + 10\eta$

bound  $\min \mathbf{Q}_+$  is almost always better than the optimal value  $\min \hat{\mathbf{Q}}_+$  of the first level of the Lasserre-SOS hierarchy applied to the original formulation (2.11) of the problem.

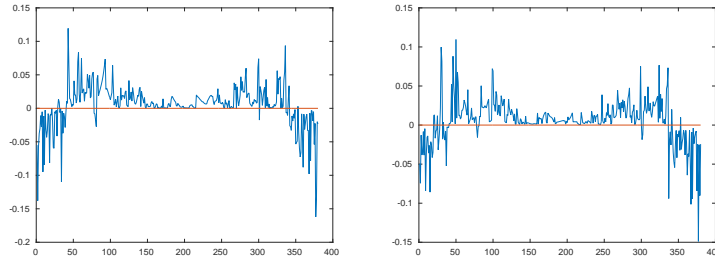


FIGURE 6. Example 2.6: Difference  $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$ ,  
( $n=10$ )  $\mathbf{c} = \mathbf{a} + 20\eta$  (left) and  $\mathbf{c} = \mathbf{a} + 10\eta$

**2.2. Extension to inequalities.** Let  $f(\mathbf{x}) := \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x}$  for some  $\mathbf{c} \in \mathbb{R}^n$  and some  $\mathbf{F}^T = \mathbf{F} \in \mathbb{R}^{n \times n}$ , and consider the problem:

$$(2.12) \quad \mathbf{P} : \quad f^* = \min_{\mathbf{x}} \{ f(\mathbf{x}) : \mathbf{A} \mathbf{x} \leq \mathbf{b}; \quad \mathbf{x} \in \{0, 1\}^n \},$$

for some cost vector  $\mathbf{c} \in \mathbb{Z}^n$ , some matrix  $\mathbf{A} \in \mathbb{Z}^{m \times n}$ , and some vector  $\mathbf{b} \in \mathbb{Z}^m$ . We may and will replace (2.12) with the equivalent pure integer program:

$$\mathbf{P}' : \quad f^* = \min_{\mathbf{x}, \mathbf{y}} \{ f(\mathbf{x}) : \mathbf{A} \mathbf{x} + \mathbf{y} = \mathbf{b}; \quad \mathbf{x} \in \{0, 1\}^n; \quad \mathbf{y} \in \mathbb{N}^m \}.$$

Next, as  $\mathbf{x} \in \{0, 1\}^n$  we can bound each integer variable  $y_j$  by  $M_j := b_j - \min\{\mathbf{A}_j \mathbf{x} : \mathbf{x} \in \{0, 1\}^n\}$ ,  $j = 1, \dots, m$ , where  $\mathbf{A}_j$  denotes the  $j$ -th row vector of the matrix  $\mathbf{A}$ ; and in fact  $M_j = b_j - \sum_i \min[0, \mathbf{A}_{ji}]$ ,  $j = 1, \dots, m$ . Then we may use the standard decomposition of  $y_j$  into a weighted sum of boolean

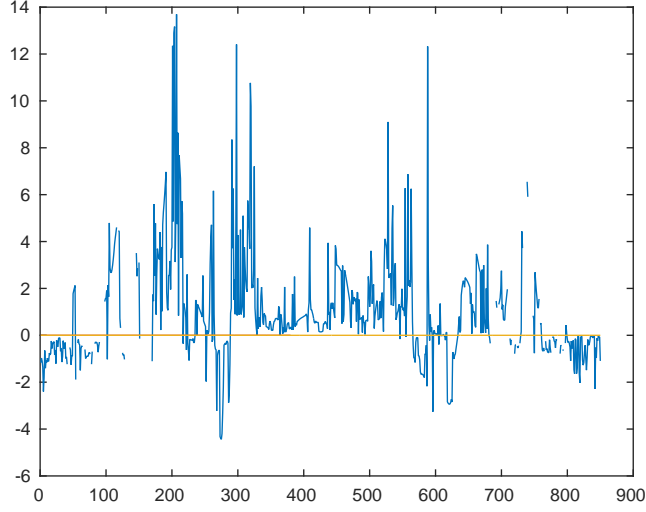


FIGURE 7. Example 2.6: Difference  $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$ ,  
( $n=15$ )  $\mathbf{c} = \mathbf{a} + 20\eta$  (left) and  $\mathbf{c} = \mathbf{a} + 10\eta$

variables :

$$y_j = \sum_{k=0}^{s_j} 2^k z_{jk}, \quad z_{jk} \in \{0, 1\}^n, \quad j = 1, \dots, m,$$

(where  $s_j := \lceil \log(M_j) \rceil$ ) and replace (2.12) with the equivalent 0/1 program:

$$f^* = \min_{\mathbf{x}, \mathbf{z}} \{ f(\mathbf{x}) : \mathbf{A}_j^T \mathbf{x} + \sum_{k=0}^{s_j} 2^k z_{jk} = b_j, j \leq m; (\mathbf{x}, \mathbf{z}) \in \{0, 1\}^{n+s} \},$$

(where  $s := \sum_j (1 + s_j)$ ) which is of the form (1.1).

**2.3. Extension to polynomial programs.** Let  $f \in \mathbb{R}[\mathbf{x}]$  be a polynomial of even degree  $d > 2$  and consider the polynomial program:

$$(2.13) \quad f^* = \min \{ f(\mathbf{x}) : \mathbf{A}\mathbf{x} = \mathbf{b}; \quad \mathbf{x} \in \{-1, 1\}^n \},$$

on the hyper cube  $\{-1, 1\}^n$ . Let  $d' := \lceil d/2 \rceil$ ,  $\mathbf{x} \mapsto g_j(\mathbf{x}) := 1 - x_j^2$ ,  $j = 1, \dots, n$ , and with  $f$  let us associate the scalars:

$$\begin{aligned} r_f^1 &= \min \{ L_{\mathbf{y}}(f) : \mathbf{M}_{d'}(\mathbf{y}) \succeq 0; \mathbf{M}_{d'-1}(g_j \mathbf{y}) = 0, j = 1, \dots, n \} \\ r_f^1 &= \max \{ L_{\mathbf{y}}(f) : \mathbf{M}_{d'}(\mathbf{y}) \succeq 0; \mathbf{M}_{d'-1}(g_j \mathbf{y}) = 0, j = 1, \dots, n \} \end{aligned}$$

where  $\mathbf{M}_{d'}(\mathbf{y})$  (resp.  $\mathbf{M}_{d'-1}(g_j \mathbf{y})$ ) is the moment matrix (resp. localizing matrix) of order  $d'$  associated with the real sequence  $\mathbf{y} = (y_\alpha)$ ,  $\alpha \in \mathbb{N}^n$ , (resp. with the sequence  $\mathbf{y}$  and the polynomial  $g_j$ ). It turns out that  $r_f^1$  (resp.  $r_f^2$ ) is the optimal value of the first SDP-relaxation of the Lasserre-SOS hierarchy associated with the optimization problem  $\min$  (resp.  $\max$ )  $\{ f(\mathbf{x}) : \mathbf{x} \in$

$\{-1, 1\}^n$  and so  $r_f^1 \leq \min\{f(\mathbf{x}) : \mathbf{x} \in \{-1, 1\}^n\}$  whereas  $r_f^2 \geq \max\{f(\mathbf{x}) : \mathbf{x} \in \{-1, 1\}^n\}$ . For more details, see e.g. [8, 9].

Next, if we define

$$(2.14) \quad \rho_f := \max_{i=1,2} |r_f^i|,$$

then it is straightforward to verify that  $\rho_f \geq \max\{|f(\mathbf{x})| : \mathbf{x} \in \{-1, 1\}^n\}$ . Then we have the following analogue of Lemma 2.1:

**Lemma 2.7.** *Let  $f^*$  be as (2.13) and let  $\rho_f$  be as in (2.14). Then  $f^*$  is the optimal value of the polynomial minimization problem:*

$$(2.15) \quad \min_{\mathbf{x} \in \{-1, 1\}^n} f(\mathbf{x}) + (2\rho_f + 1) \cdot \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2.$$

The proof being almost a verbatim copy of that of Lemma 2.1, is omitted.

As for the quadratic case and with same arguments, one may also show that if  $d$  is even, the polynomial optimization problem (2.15) is equivalent to minimizing the homogeneous polynomial  $\tilde{f}$  of degree  $d$  on the hypercube  $\{-1, 1\}^{n+1}$ , where

$$(\mathbf{x}, x_0) \mapsto \tilde{f}(\mathbf{x}) := x_0^d f(\mathbf{x}/x_0) + (2\rho_f + 1) x_0^{d-2} \cdot \|\mathbf{A}\mathbf{x} - x_0 \mathbf{b}\|^2.$$

But since it is not a MAX-CUT problem, to obtain a lower bound on  $f^*$  one may just as well consider solving the first level of the Lasserre-SOS hierarchy associated with (2.15) or even directly with (2.13). The advantage of using the formulation (2.15) is that one always minimizes on the hypercube  $\{-1, 1\}^n$  instead of minimizing on the subset  $\{-1, 1\}^n \cap \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}\}$  of the hypercube which is problem dependent.

**2.4. Comparing with the copositive formulation.** As already mentioned in the introduction, the 0/1 program (1.1) also has a copositive formulation. Namely, let  $e_i = (\delta_{i=j}) \in \mathbb{R}^n$ ,  $i = 1, \dots, n$ , and  $\mathbf{e} = (1, \dots, 1) \in \mathbb{R}^n$ . Following Burer [3, p. 481–482], introduce  $n$  additional variables  $\mathbf{z} = (z_1, \dots, z_n)$  and the  $n$  additional equality constraints  $x_i + z_i = 1$ ,  $i = 1, \dots, n$ , with  $\mathbf{z} \geq 0$  (which are necessary to obtain an equivalent formulation). So let  $\tilde{\mathbf{x}}^T = (\mathbf{x}^T, \mathbf{z}^T) \in \mathbb{R}^{2n}$  and with  $\mathbf{I} \in \mathbb{R}^{n \times n}$  being the identity matrix, introduce the real matrices

$$\tilde{\mathbf{F}} := \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{S} := \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{I} & \mathbf{I} \end{bmatrix}$$

and the real vectors  $\tilde{\mathbf{c}}^T := (\mathbf{c}^T, 0) \in \mathbb{R}^{2n}$  and  $\tilde{\mathbf{b}}^T = (\mathbf{b}^T, \mathbf{e}^T) \in \mathbb{R}^{m+n}$ . Let  $\mathbf{S}_i$  denote the  $i$ -th row vector of  $\mathbf{S}$ ,  $i = 1, \dots, 2n$ . Then the copositive

formulation of (1.1) reads:

$$\begin{aligned}
(2.16) \quad f^* = \min \quad & \tilde{\mathbf{c}}^T \tilde{\mathbf{x}} + \langle \tilde{\mathbf{F}}, \mathbf{X} \rangle \\
\text{s.t.} \quad & \mathbf{S}_i \tilde{\mathbf{x}} = \tilde{\mathbf{b}}_i, \quad i = 1, \dots, m+n \\
& \mathbf{S}_i \mathbf{X} \mathbf{S}_i^T = \tilde{\mathbf{b}}_i^2, \quad i = 1, \dots, m+n \\
& \mathbf{X}_{ii} = \tilde{\mathbf{x}}_i, \quad i = 1, \dots, 2n \\
& \begin{bmatrix} 1 & \tilde{\mathbf{x}}^T \\ \tilde{\mathbf{x}} & \mathbf{X} \end{bmatrix} \in \mathcal{C}_{2n+1}^*,
\end{aligned}$$

where  $\mathcal{C}_{2n+1}$  is the convex cone of  $(2n+1) \times (2n+1)$  copositive matrices and  $\mathcal{C}_{2n+1}^*$  is its dual, i.e., the convex cone of *completely positive matrices*.

The hard constraint being membership in  $\mathcal{C}_{2n+1}^*$ , a strategy is to use hierarchies of tractable approximations (of increasing size) of  $\mathcal{C}_{2n+1}^*$ , as described in e.g. Dürre [5]. In particular a possible choice for the first relaxation in such hierarchies is to replace (2.16) with the semidefinite program:

$$\begin{aligned}
(2.17) \quad f_{copo}^* = \min \quad & \tilde{\mathbf{c}}^T \tilde{\mathbf{x}} + \langle \tilde{\mathbf{F}}, \mathbf{X} \rangle \\
\text{s.t.} \quad & \mathbf{S}_i \tilde{\mathbf{x}} = \tilde{\mathbf{b}}_i, \quad i = 1, \dots, m+n \\
& \mathbf{S}_i \mathbf{X} \mathbf{S}_i^T = \tilde{\mathbf{b}}_i^2, \quad i = 1, \dots, m+n \\
& \mathbf{X}_{ii} = \tilde{\mathbf{x}}_i, \quad i = 1, \dots, 2n \\
& \begin{bmatrix} 1 & \tilde{\mathbf{x}}^T \\ \tilde{\mathbf{x}} & \mathbf{X} \end{bmatrix} \in \mathcal{S}_{2n+1}^+ \cap \mathcal{N}_{2n+1},
\end{aligned}$$

where  $\mathcal{S}_{2n+1}^+$  (resp.  $\mathcal{N}_{2n+1}$ ) is the convex cone of real symmetric positive semidefinite (resp. entrywise nonnegative) matrices. Then (2.17) is a semidefinite relaxation of (2.16) because  $\mathcal{C}_{2n+1}^* \subset \mathcal{S}_{2n+1}^+ \cap \mathcal{N}_{2n+1}$ , and so  $f_{copo}^* \leq f^*$ . In fact if one considers the problem

$$(2.18) \quad \min_{\mathbf{x} \in \{0,1\}^n} \{ \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} : \mathbf{A} \mathbf{x} = \mathbf{b}; (\mathbf{A}_i^T \mathbf{x})^2 = b_i^2, i = 1, \dots, m \},$$

which is clearly equivalent to (1.1), then the first SDP-relaxation of the Lasserre-SOS hierarchy associated with (2.18) reads:

$$\begin{aligned}
(2.19) \quad \min \quad & \mathbf{c}^T \mathbf{x} + \langle \mathbf{F}, \mathbf{X} \rangle \\
\text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\
& \mathbf{A}_i \mathbf{X} \mathbf{A}_i^T = \mathbf{b}_i^2, \quad i = 1, \dots, n \\
& \mathbf{X}_{ii} = x_i, \quad i = 1, \dots, n \\
& \begin{bmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{X} \end{bmatrix} \succeq 0,
\end{aligned}$$

which is of the same flavor as the semidefinite program (2.17) but with dimension twice as less than (2.17).

**Example 2.8.** We have compared the first SDP-relaxation of the MAX-CUT formulation with the first SDP-relaxation (2.17) of the copositive formulation for the linear 0/1 knapsack problems (1.1) with  $\mathbf{F} = \mathbf{0}$ ,  $n = 10, 15$  and vector  $\mathbf{a}$  as in Example 2.5.

We have kept the formulation on the hypercube  $\{0, 1\}^n$  rather than on the hypercube  $\{-1, 1\}^n$  and so in fact the first SDP relaxation is for problem

(2.3) with a quadratic cost function (and not a quadratic form as in the MAX-CUT formulation on  $\{-1, 1\}^n$ ).

In each case  $n = 10$  (resp.  $n = 15$ ), we have chosen 19 (resp. 18) values of the right-hand-side  $b = 10s$ ,  $s = 1, \dots, 19$  (resp.  $b = 20s$ ,  $s = 1, \dots, 18$ ), and for each problem we have run a sample of 10 problems with cost vector  $\mathbf{c} = \mathbf{a} + 10\eta$  where  $\eta$  is a random variable uniformly distributed in  $[0, 1]$ .

For  $n = 10$ , the lower bound  $f_{maxcut}^*$  from the MAX-CUT formulation dominates the lower bound  $f_{copo}^*$  in (2.17), in 111 out of 190 problems ( $\approx 58\%$ ) and the relative difference  $100 \cdot |f_{maxcut}^* - f_{copo}^*| / \max[f_{maxcut}^*, f_{copo}^*]$  never exceeds 0.05% over all 190 problems!

For  $n = 15$ ,  $f_{maxcut}^* > f_{copo}^*$  in 94 out of 180 problems ( $\approx 52\%$ ) and  $100 \cdot |f_{maxcut}^* - f_{copo}^*| / \max[f_{maxcut}^*, f_{copo}^*]$  never exceeds 0.55% in all 180 problems!

### 3. CONCLUSION

In this paper we have shown that a linear or quadratic 0/1 program has an equivalent MAX-CUT formulation and so the whole arsenal of approximation techniques for the latter can be applied. In particular, and as suggested by some preliminary tests on a (limited sample) of 0/1 knapsack examples, it is expected that the lower bound obtained from the Shor relaxation of MAX-CUT will be in general better than the one obtained from the standard LP-relaxation (for linear 0/1 programs) of the original problem. The situation might be even better for quadratic 0/1 programs since to obtain a lower bound there is no need to first compute a convex quadratic underestimator of the criterion before applying a convex quadratic relaxation.

### REFERENCES

- [1] I.M. Bomze. Copositive optimization - recent developments and applications, *European J. Oper. Res.* **216** (2012), 509–520.
- [2] I.M. Bomze, E. de Klerk. Solving standard quadratic optimization problems via linear, semidenite and copositive programming, *J. Global Optim.* **24** (2002), 163–185.
- [3] S. Burer. Copositive programming, in *Handbook of Semidefinite, Conic and Polynomial Optimization*, M. Anjos and J.B. Lasserre, Eds., Springer, New York, 2012, pp. 201–218.
- [4] E. de Klerk, D.V. Pasechnik. Approximation of the stability number of a graph via copositive programming, *SIAM J. Optim.* **12** (2002), 875–892.
- [5] M. Dür. Copositive Programming - a survey, In *Recent Advances in Optimization and its Applications in Engineering*, M. Diehl, F. Glineur, E. Jarlebring and W. Michiels Eds., Springer, New York, 2010, pp. 3 – 20.
- [6] J.B. Lasserre. Optimisation globale et théorie des moments, *C. R. Acad. Sci. Paris* **331** (2000), Série 1, pp. 929–934.
- [7] J.B. Lasserre. Global optimization with polynomials and the problem of moments, *SIAM J. Optim.* **11** (2001), pp. 796–817.
- [8] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, Imperial College Press, london, 2010.
- [9] J.B. Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization*, Cambridge University Press, Cambridge, 2015.

- [10] M. Laurent, Zhao Sun. Handelman's hierarchy for the maximum stable set problem, *J. Global Optim.* **60** (2014), pp. 393–423.
- [11] M. Marshall. Error estimates in the optimization of degree two polynomials on a discrete hypercube, *SIAM J. Optim.* **16** (2005), 297–309.
- [12] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization, *Optim. Methods Softw.* **9** (1998), 141–160.
- [13] A. Wigele, F. Rendl. Computational approaches to Max-Cut, in *Handbook of Semidefinite, Conic and Polynomial Optimization*, M. Anjos and J.B. Lasserre, Eds., Springer, New York, 2012, pp. 821–848.

LAAS-CNRS AND INSTITUTE OF MATHEMATICS, LAAS, 7 AVENUE DU COLONEL ROCHE, BP 54200, 31031 TOULOUSE CÉDEX 4, FRANCE.

*E-mail address:* `lasserre@laas.fr`