A MAX-CUT FORMULATION OF 0/1 PROGRAMS

Jean-Bernard Lasserre

HAL Id: hal-01154698
https://hal.science/hal-01154698v3

Submitted on 22 Dec 2015

A MAX-CUT FORMULATION OF 0/1 PROGRAMS

JEAN B. LASSERRE

ABSTRACT. We show that the linear or quadratic 0/1 program

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

can be formulated as a MAX-CUT problem whose associated graph is simply related to the matrices $\mathbf{F}$ and $\mathbf{A}^T\mathbf{A}$. Hence the whole arsenal of approximation techniques for MAX-CUT can be applied. We also compare the lower bound of the resulting semidefinite (or Shor) relaxation with that of the standard LP-relaxation and the first semidefinite relaxations associated with the Lasserre hierarchy and the copositive formulations of $\mathbf{P}$.

Keywords: linear and quadratic 0/1 programs; MAX-CUT problem; LP- and semidefinite relaxations.

1. INTRODUCTION

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

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

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 $\mathbf{x}^T\tilde{\mathbf{F}}\mathbf{x}$ 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 [4] and compute 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 [5], Dür [6], Bomze [2], and Bomze and de Klerk [3].

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. This idea had been already briefly mentioned in Rendl [16] in the context of partitioning and ordering problems. 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 $G = (V, E)$ with $n + 1$ nodes and $(i, j) \in E$ 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 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 obtains 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 [8, 9] would provide a monotone nondecreasing sequence of improved lower bounds $f_1^* \leq f_d^* \leq f^*$, $d = 2, \ldots$, 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 [12]. For more details on recent developments on computational approaches to MAX-CUT the interested reader is referred to Palagi et al. [15]. In addition one may also obtain performance guarantees *à la Nesterov* [14] or their improvements by Marshall [13]. 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 polynomial criterion $\tilde{f}(\mathbf{x})$ on the hypercube $\{-1, 1\}^n$ (and not a MAX-CUT problem any more as the degree of $\tilde{f}$ is larger than 2).

Numerical experiments. For 0/1 linear programs ($\mathbf{F} = 0$) 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. In fact f_1^* is almost always better than the lower bound obtained from the first SDP-relaxation of the Lasserre-SOS hierarchy applied to the initial formulation (1.1) of the problem (an SDP of same size which is a basic quadratic relaxation of the initial problem). 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.) For quadratic $0/1$ knapsacks, f_1^* is also better than the lower bound obtained by relaxing $\{0,1\}^n$ to $[0,1]^n$, replacing $\mathbf{F}$ with a convex quadratic underestimator, and solving the resulting convex quadratic relaxation.

We have also considered the MAX-CUT formulation of the k-cluster problem $\rho = \min\{\mathbf{x}^T \mathbf{A}\mathbf{x} : \mathbf{e}^T\mathbf{x} = k; \mathbf{x} \in \{0,1\}^n\}$ where $k \in \mathbb{N}$ and $\mathbf{A}$ is some real symmetric matrix, for $n = 70$ and $n = 100$ variables and with various percentages of zero entries in the matrix $\mathbf{A}$. In all examples the optimal value of the Shor relaxation was almost indistinguishable of the first semidefinite relaxation of the Lasserre-SOS hierarchy applied to the initial formulation. We have also compared the Shor relaxation with a quadratic convex relaxation of the problem. The former always provides better lower bounds (and sometimes significantly better). At last we have also compared the Shor relaxations respectively associated with the MAX-CUT and copositive formulations. Again the respective optimal values are very close (less than 1% relative difference and in a few cases the former is significantly better).

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:

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

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:

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

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

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

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

It is straightforward to verify that

$$(2.3) \qquad \rho(\mathbf{c},\mathbf{F}) \geq \max\{\,|\mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{F}\mathbf{x}| : \mathbf{x} \in \{-1,1\}^n\,\},$$

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 [8, 9]) 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 P.

Lemma 2.1. *Let $\mathbf{P}$ be as (2.1) and let $\rho(\mathbf{c}, \mathbf{F})$ be as in (2.2). If $f^* < +\infty$ then f^* is the optimal value of the quadratic minimization problem, i.e.,*

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

Moreover $\mathbf{P}$ has no feasible solution ($f^* = +\infty$) if and only if $\theta > \rho(\mathbf{c}, \mathbf{F})$.

Proof. Let $\Delta := \{\mathbf{x} \in \{-1, 1\}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}$ be the feasible set of $\mathbf{P}$ defined in (2.1), and let $f : \mathbb{R}^n \to \mathbb{R}$ be the function

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

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,

$$(2.6) \quad f(\mathbf{x}) \begin{cases} = \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} \text{ on } \Delta \\ \geq \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} + 2\,\rho(\mathbf{c}, \mathbf{F}) + 1 > \rho(\mathbf{c}, \mathbf{F}) \text{ on } \{-1, 1\}^n \setminus \Delta. \end{cases}$$

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

Next, let $Q : \mathbb{R}^{n+1} \to \mathbb{R}$ 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:

$$(2.7) \quad Q(\mathbf{x}, x_0) = x_0 \mathbf{c}^T \mathbf{x} + \mathbf{x}^T \mathbf{F} \mathbf{x} + (2\,\rho(\mathbf{c}, \mathbf{F}) + 1) \cdot \|\mathbf{A}\mathbf{x} - x_0 \mathbf{b}\|^2.$$

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 (2.7). If $f^* < +\infty$ then*

$$(2.8) \quad f^* = \theta := \min_{(\mathbf{x}, x_0) \in \{-1,1\}^{n+1}} Q(\mathbf{x}, x_0),$$

that is, f^* is the optimal value of the MAX-CUT problem associated with the quadratic form Q. Moreover $f^* = +\infty$ if and only if $\theta > \rho(\mathbf{c}, \mathbf{F})$.

Proof. Let f be as in (2.5). By definition of Q,

$$(2.9) \quad \min_{\mathbf{x} \in \{-1,1\}^n} f(\mathbf{x}) = \min_{(\mathbf{x}, x_0) \in \{-1,1\}^{n+1}} \{Q(\mathbf{x}, x_0) : x_0 = 1\}.$$

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,

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

4

which combined with (2.9) and Lemma 2.1 yields the desired result. $\qquad\square$

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

$$(2.10) \qquad \min_{\mathbf{X}} \{\langle \mathbf{Q}, \mathbf{X} \rangle : \mathbf{X} \succeq 0;\ X_{ii} = 1, \quad i = 1, \dots, n+1\}$$

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

$$(2.11) \qquad \max_{\mathbf{X}} \{\langle \mathbf{Q}, \mathbf{X} \rangle : \mathbf{X} \succeq 0;\ X_{ii} = 1, \quad i = 1, \dots, n+1\}$$

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^* *and let* $\rho(\mathbf{c}, \mathbf{F})$ *be as in (2.2). If* $f^* < +\infty$ *then*

$$(2.12) \qquad \min \mathbf{Q}_+ \leq f^* \leq \frac{2}{\pi} \min \mathbf{Q}_+ + (1 - \frac{2}{\pi}) \max \mathbf{Q}^+,$$

where $\mathbf{Q}$ *is the real symmetric matrix asociated with the quadratic form (2.7) and* $\min \mathbf{Q}_+$ *(resp.* $\max \mathbf{Q}^+$*) is the optimal value of the semidefinite program (2.10) (resp. (2.11)). Moreover, if* $\min \mathbf{Q}_+ > \rho(\mathbf{c}, \mathbf{F})$ *then* $\mathbf{P}$ *has no feasible solution (i.e.,* $f^* = +\infty$*).*

Proof. The bounds in (2.12) are from Nesterov [14]. In addition, one may also use the bounds provided in Marshall [13] which sometimes improve those in (2.12). Next, let $\Delta := \{\mathbf{x} \in \{-1, 1\}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}$ and assume that $\Delta \neq \emptyset$. Then by (2.3) and (2.6), $\min \mathbf{Q}_+ \leq f^* \leq \rho(\mathbf{c}, \mathbf{F})$. Therefore $\Delta = \emptyset$ if $\min \mathbf{Q}_+ > \rho(\mathbf{c}, \mathbf{F})$. $\qquad\square$

The quality of the upper bound in (2.12) depends strongly on the magnitude of the "penalty coefficient" $2\rho(\mathbf{c}, \mathbf{F}) + 1$ in the definition of the function f in (2.5). Whether or not the penalty parameter $\rho(\mathbf{c}, \mathbf{F})$ could be taken smaller (at least in some cases) has not been investigated and is beyond the scope of this paper. 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 $\mathbf{Q} \not\succeq 0$ or $\mathbf{Q} \not\preceq 0$). 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.

2.2. Sparsity. Hence to each 0/1 program (1.1) one may associate a graph $G = (V, E)$ with $n + 1$ nodes and an arc $(i, j) \in E$ connects the nodes $i, j \in V$ 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 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 $k = 1, \dots, m$, that is, in any of the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$, at most one of the two variables (x_i, x_j) appears (a structural condition). A weaker

condition (non-structural as it depends on values of entries of $\mathbf{A}$) is that $\sum_k \mathbf{A}_{ki} \mathbf{A}_{kj} = 0$.

2.3. 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:

$$(2.13) \qquad \mathbf{P}: \qquad f^* = \min_{\mathbf{x}} \{ f(\mathbf{x}) : \mathbf{A}\, \mathbf{x} \leq \mathbf{b}; \quad \mathbf{x} \in \{0,1\}^n \},$$

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 (2.13) with the equivalent pure integer program:

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

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

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

(where $s_j := \lceil \log(M_j) \rceil$) and replace (2.13) 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).

2.4. Extension to polynomial programs.
Let $f \in \mathbb{R}[\mathbf{x}]$ be a polynomial of even degree $d > 2$, written as

$$\mathbf{x} \mapsto f(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha \mathbf{x}^\alpha,$$

for some vector of coefficients $\mathbf{f} = (f_\alpha) \in \mathbb{R}^{s(d)}$ (with $s(d) = \binom{n+d}{n}$). Given a sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, define the Riesz functional $L_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ by:

$$f \mapsto L_{\mathbf{y}}(f) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha\, y_\alpha, \qquad f \in \mathbb{R}[\mathbf{x}].$$

Consider the polynomial program:

$$(2.14) \qquad f^* = \min \{ f(\mathbf{x}) : \mathbf{A}\mathbf{x} = \mathbf{b}; \quad \mathbf{x} \in \{-1,1\}^n \},$$

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

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

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. [10, 11]. Next, if we define

$$(2.15) \qquad \rho_f := \max_{i=1,2} |r_f^i|,$$

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 2.4. *Let f^* be as (2.14) and ρ_f as in (2.15). If $f^* < +\infty$ then*

$$(2.16) \qquad f^* = \rho := \min_{\mathbf{x} \in \{-1,1\}^n} f(\mathbf{x}) + (2\,\rho_f + 1) \cdot \|\mathbf{A}\,\mathbf{x} - \mathbf{b}\|^2$$

(a polynomial optimization problem on$\{-1,1\}^n$). Moreover $f^* = +\infty$ if and only if $\rho > \rho_f$.

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 (2.16) 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 (2.16) or even directly with (2.14). The advantage of using the formulation (2.16) 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.

3. Some numerical experiments

All semidefinite programs were solved by using the GloptiPoly software dedicated to solving the Generalized Problem of Moments and which implements the Lasserre-SOS hierarchy of semidefinite relaxations. For more details the interested reader is referred to Henrion et al. [7]. Typically, for a MAX-CUT problem of size $n = 100$, solving the Shor relaxation with GloptiPoly (and calling the semidefinite solver MOSEK [1]) takes approximately 42s on a Macbook-pro lap-top with Intel Core i5 processor at 2.6 GHz.

3.1. **Some 0/1-knapsack examples.** To test the efficiency of the MAX-CUT formulation we have first considered 0/1-knapsack problems of small size (up to $n = 15$ variables) and compared the first semidefinite relaxation (or Shor-relaxation) of the MAX-CUT formulation (2.4) with some other relaxations, and in particular with:

- The LP-relaxation of (2.1) (when $\mathbf{F} = 0$) which consists of replacing the constraint $\mathbf{x} \in \{-1, 1\}^n$ with $\mathbf{x} \in [-1, 1]^n$, and
- The first semidefinite relaxation of (2.1) in the Lasserre-SOS hierarchy.

Example 3.1. To evaluate the quality of the lower bound obtained with the MAX-CUT formulation consider the following simple linear knapsack-type examples:

$$(3.1) \qquad \min \{\mathbf{c}^T \mathbf{x} : \ \mathbf{a}^T \mathbf{x} = b; \ \mathbf{x} \in \{-1, 1\}^n \},$$

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 $\mathbf{a} = (3,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 37)$ and for $n = 15$

$$\mathbf{a} = (3,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 37,\ 41,\ 43,\ 47,\ 51,\ 53)$$

and $\mathbf{c} = (53,\ 51,\ 47,\ 43,\ 41,\ 37,\ 31,\ 29,\ 23,\ 19,\ 17,\ 13,\ 11,\ 7,\ 3)$.

The right-hand-side b is taken into $[-|\mathbf{a}|, |\mathbf{a}|] \cap \mathbb{Z}$. Figure 1 displays the difference $\min \mathbf{Q}_+ - \min \mathrm{LP}$ where the lower bound $\min \mathrm{LP}$ 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. On the x-axis one reads $b + |\mathbf{a}|$. As expected the lower bound $\min \mathbf{Q}_+$ is much better than $\min \mathrm{LP}$. 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^*.
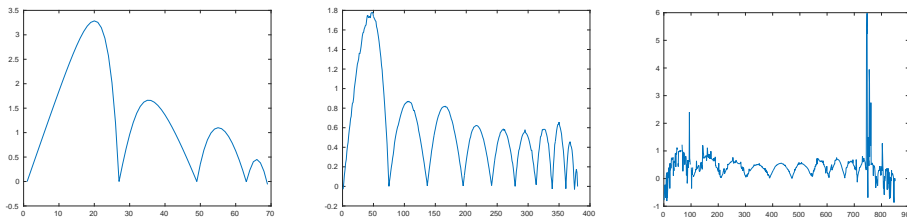

FIGURE 1. Difference $\min \mathbf{Q}_+ - \min \mathrm{LP}$ with n=(4,10,15); One reads $b + |\mathbf{a}|$ on the x-axis

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 (3.1) of the knapsack problem where one has even included the redundant constraints $x_i (\mathbf{a}^T \mathbf{x} - b) = 0$, $i = 1, \ldots, n$. 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 [8, 9] 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 [13].
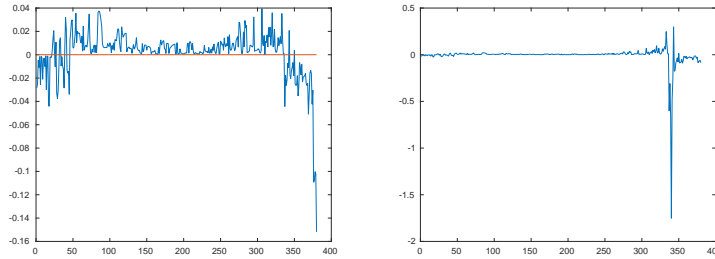
8

FIGURE 2. Difference $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$ (n=10) (left) and relative difference $100 * (\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+)/\min \hat{\mathbf{Q}}_+$; One reads $b + |\mathbf{a}|$ on the x-axis

Example 3.2. In a second sample of linear knapsack problems (3.1) with $n = 10, 15$, we have chosen the same vector $\mathbf{a}$ as in Example 3.1 but now with a cost criterion of the form $c(i) = a(i) + s\,\eta$, $i = 1, \ldots, n$, where η is a random variable uniformly distributed on $[0, 1]$, and s is a weighting factor. The reason is that knapsack problems with ratii $c(i)/a(i) \approx 1$ 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 results obtained for $s = 10$, and $n = 10, 15$.
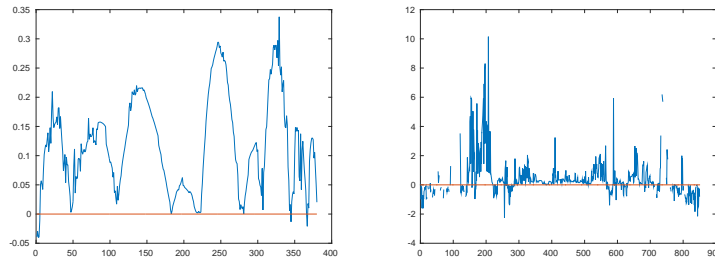

FIGURE 3. Difference $\min \mathbf{Q}_+ - \min \mathrm{LP}$, $\mathbf{c} = \mathbf{a} + 10 * \eta$ (n=10,15); $b + |\mathbf{a}|$ on the x-axis

Finally, as for Example 3.1, Figure 4 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 (3.1) of the knapsack problem where one has even included the redundant constraints $x_i\,(\mathbf{a}^T\mathbf{x} - b) = 0$, $i = 1, \ldots, n$). Again one observes that in most cases the lower bound $\min \mathbf{Q}_+$ is slightly better than $\min \hat{\mathbf{Q}}_+$.

Example 3.3. We next consider the same knapsack problems (3.1) as in Example 3.2 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 $c(i) = a(i) + s\,\eta$, $i = 1, \ldots, n$, where η is a random
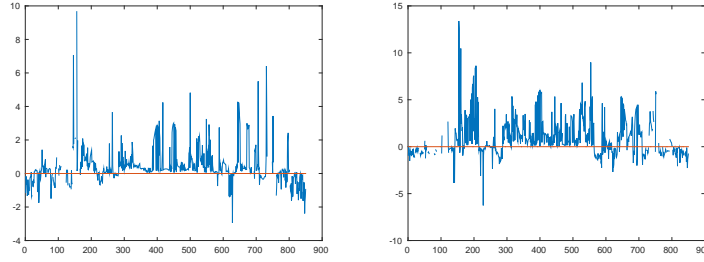
FIGURE 4. Example 3.2: Difference $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$, (n=15) $\mathbf{c} = \mathbf{a} + 20\eta$ (left) and $\mathbf{c} = \mathbf{a} + 10\eta$; $b + |\mathbf{a}|$ on the x-axis

variable uniformly distributed in $[0, 1]$. The real symmetric matrix $\mathbf{F}$ is also randomly generated and is not positive definite in general. Again in Figure 5 one observes that the lower 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 (3.1) of the problem.
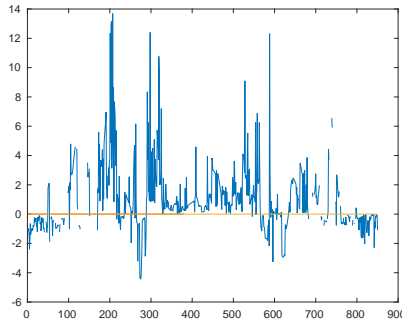

FIGURE 5. Example 3.3: Difference $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$, (n=15) $\mathbf{c} = \mathbf{a} + 10\eta$; $b + |\mathbf{a}|$ on the x-axis

3.2. **Larger size problems: The k-cluster problem.** As generating data $(\mathbf{c}, \mathbf{a}, b)$ for difficult 0/1-knapsack problems of larger size is not obvious, we next consider the so-called k-cluster problem in $n = 70$ and $n = 80$ variables:

$$(3.2) \qquad \min \{ \mathbf{x}^T \mathbf{A} \mathbf{x} : \mathbf{e}^T \mathbf{x} = k; \mathbf{x} \in \{0, 1\}^n \},$$

where $k \in \mathbb{N}$ is a fixed parameter, $\mathbf{e} = (1, 1, \ldots, 1) \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ with $\mathbf{A} = \mathbf{A}^T$ and $\mathbf{A}_{ii} = 0$ for all $i = 1, \ldots, n$. Indeed the constraint $\mathbf{e}^T \mathbf{x} = k$ of the k-cluster problem is structural and does not depend on a vector $\mathbf{a}$ as in knapsack examples. After the change of variables $\mathbf{x} \to (\mathbf{x} + \mathbf{e})/2$ the k-cluster problem reads:

$$(3.3) \qquad \mathbf{P}: \quad \min \{ 2\mathbf{e}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{A} \mathbf{x} : \mathbf{e}^T \mathbf{x} = 2k - n; \mathbf{x} \in \{-1, 1\}^n \}.$$

10

We have compared the optimal value f^*_{maxcut} for the Shor relaxation of the MAX-CUT formulation (2.4) of (3.3) with the optimal value f^*_{Shor} for the first semidefinite relaxation of (3.3) in the Lasserre-SOS hierarchy. The latter is a rather basic quadratic relaxation of the initial problem. In all our experiments with $n = 70$ and $n = 80$, both values are almost identical since the relative difference is no more than 0.014% and 0.06% respectively.

Of course when $k > n$ the k-cluster problem (3.3) has no feasible solution to **P**. In view of the special structure of the k-cluster problem it is straightforward to check that the first semidefinite relaxation of (3.3) (or equivalently of (3.2)) in the Lasserre-SOS hierarchy is infeasible, hence certifying that $f^* = +\infty$. On the other hand the Shor relaxation of the MAX-CUT formulation of (3.3) has always a finite value. However, one may prove that if $k > n$ then the optimal value is larger than $\rho(\mathbf{c}, \mathbf{F})$, which by Proposition 2.3 also certifies that $f^* = +\infty$. That is, for the k-cluster problem (3.3) the Shor relaxation of its MAX-CUT formulation (2.4) also detects infeasibility.

Comparing with a quadratic convex relaxation. Decompose **A** as $\mathbf{U}^T \Lambda \mathbf{U}$ with $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ (the identity matrix) and Λ is the diagonal matrix of eigenvalues of **A**. Let $\theta := |\lambda_{min}(\mathbf{A})| + 1$ where $\lambda_{min}(\mathbf{A})$ is the smallest eigenvalue of **A**. Replace **A** with $\tilde{\mathbf{A}} := \mathbf{U}^T(\Lambda + \theta\mathbf{I})\mathbf{U}$ so that the resulting the quadratic form $\mathbf{x} \mapsto \mathbf{x}^T \tilde{\mathbf{A}} \mathbf{x}$ is convex. Then observe that

$$\min \{ 2\mathbf{e}^T\mathbf{A}\mathbf{x} + \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} : \mathbf{e}^T\mathbf{x} = 2k - n; \mathbf{x} \in \{-1, 1\}^n \} = f^* + n\theta$$

because $\mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x} + \theta\|\mathbf{x}\|^2$ and $x_i^2 = 1$ for all $i = 1, \ldots, n$. Therefore one may compare the optimal value f^*_{maxcut} of the Shor relaxation of (3.3) with $\tilde{\psi} = \psi - n\theta$, where

$$\psi = \min \{ 2\mathbf{e}^T\mathbf{A}\mathbf{x} + \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} : \mathbf{e}^T\mathbf{x} = 2k - n; x_i^2 \le 1, i = 1, \ldots, n \},$$

which is a convex quadratic relaxation. Results displayed in Table 1 show that the Shor relaxation is always better, and sometimes significantly better. However this convex relaxation is rather weak.

Next we do the same comparison for a k-cluster problem where some sparsity is introduced in the cost matrix **A**. Namely once **A** has been generated then for each non-diagonal entry $\mathbf{A}_{ij}$ we decide $\mathbf{A}_{ij} = \mathbf{A}_{ji} := 0$ with probability 0.4, 0.7 and 0.8. Results displayed in Table 2 for $n = 100$ show that again the Shor relaxation is always better (and some times significantly better) than the quadratic convex relaxation. The sparser is **A** the more efficient is the Shor relaxation compared with the quadratic convex relaxation.

n=70	k=50	k=45	k=40	k=35	k=30	k=25
f^*_{maxcut}	2324.68	1350.74	488.83	-261.78	-900.55	-1429.66
$\tilde{\psi}$	2260.00	1278.08	410.48	-343.95	-985.57	-1514.77
$\frac{100(f^*_{maxcut}-\tilde{\psi})}{\tilde{\psi}}$	2.86%	5.68%	19.08%	23.89%	8.62%	5.61%

TABLE 1. k-cluster: Shor relaxation (f^*_{maxcut}) vs convex quadratic relaxation ($\tilde{\psi} = \psi - n\theta$) with $n = 70$

n=100	k=70	k=60	k=50	k=40	k=20
	$\mathbf{A}_{ij} = 0$ with probability 0.8				
f^*_{maxcut}	513.32	-60.83	-450.91	-781.84	-1084.98
$\tilde{\psi}$	419.07	-147.77	-572.27	-894.72	-1188.59
$\frac{100(f^*_{maxcut}-\tilde{\psi})}{\tilde{\psi}}$	22.4%	58.8%	21.2%	12.6%	8.7%
	$\mathbf{A}_{ij} = 0$ with probability 0.6				
f^*_{maxcut}	1278.01	256.43	-627.72	-1248.39	-1988.94
$\tilde{\psi}$	1192.99	155.44	-723.34	-1372.94	-2095.66
$\frac{100(f^*_{maxcut}-\tilde{\psi})}{\tilde{\psi}}$	7.1%	64.9%	13.2%	9.0%	5.0%

TABLE 2. k-cluster: Shor relaxation (f^*_{maxcut}) vs convex quadratic relaxation ($\tilde{\psi} = \psi - n\theta$) with $n = 100$. $\mathbf{A}_{ij} = 0$ with probability $p = 0.8, 0.6$.

3.3. Comparing with the copositive formulation.

As already mentioned in the introduction, the 0/1 program (1.1) also has a copositive formulation. Namely, let $e_i = (\delta_{i=j}) \in \mathbb{R}^n$, $i = 1, \ldots, n$, and $\mathbf{e} = (1, \ldots, 1) \in \mathbb{R}^n$. Following Burer [4, p. 481–482], introduce n additional variables $\mathbf{z} = (z_1, \ldots, z_n)$ and the n additional equality constraints $x_i + z_i = 1$, $i = 1, \ldots, n$, 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

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

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 $\mathbf{S}$, $i = 1, \ldots, 2n$. Then the copositive formulation of (1.1) reads:

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

(3.4)

12

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. Dür [6]. In particular a possible choice for the first relaxation in such hierarchies is to replace (3.4) with the semidefinite program:
(3.5)
$$
\begin{aligned}
f_{copo}^* = \min \quad & \tilde{\mathbf{c}}^T \tilde{\mathbf{x}} + \langle \tilde{\mathbf{F}}, \mathbf{X} \rangle \\
\text{s.t.} \quad & \mathbf{S}_i\, \tilde{\mathbf{x}} = \tilde{\mathbf{b}}_i; \quad \mathbf{S}_i\, \mathbf{X}\, \mathbf{S}_i^T = \tilde{\mathbf{b}}_i^2, \quad i = 1, \ldots, m + n \\
& \mathbf{X}_{ii} = \tilde{\mathbf{x}}_i, \quad i = 1, \ldots, 2n; \quad \begin{bmatrix} 1 & \tilde{\mathbf{x}}^T \\ \tilde{\mathbf{x}} & \mathbf{X} \end{bmatrix} \in \mathcal{S}_{2n+1}^+ \cap \mathcal{N}_{2n+1},
\end{aligned}
$$

where $\mathcal{S}_{2n+1}^+$ (resp. $\mathcal{N}_{2n+1}$) is the convex cone of real symmetric positive semidefinite (resp. entrywise nonnegative) matrices. Then (3.5) is a semidefinite relaxation of (3.4) because $\mathcal{C}_{2n+1}^* \subset \mathcal{S}_{2n+1}^+ \cap \mathcal{N}_{2n+1}$, and so $f_{copo}^* \leq f^*$.

Example 3.4. We have considered the k-cluster problem (3.2) with $n = 50$ variables and where $\mathbf{F}$ is randomly generated with 80% of zero entries in $\mathbf{F}$. Then we compared the optimal values f_{maxcut}^* and f_{copo}^* of the respective Shor-relaxations of the MAX-CUT formulation (2.4) and the copositive formulation of (3.2). For most values of k the lower bound f_{maxcut}^* was almost identical to f_{copo}^* (the relative difference $\Delta := 100\,(f_{maxcut}^* - f_{copo}^*)/f_{copo}^*$ being at most 0.1%), and was better ($\Delta = 68\%$ and 44%) for small $k = 12$ and $k = 14$ respectively. However notice that the size of the semidefinite constraint in the latter is twice as large as the one in the former.

Remark 3.5. Notice that in all examples the bounds f_{maxcut}^* and f_{Shor}^* (respectively denoted $\min \mathbf{Q}_+$ and $\min \hat{\mathbf{Q}}_+$ in Section 3.1) are almost identical. It might be tempting to conjecture that they are indeed identical and the error is coming from numerical roundoffs. However results displayed in Figure 5 show that the difference $\min \mathbf{Q}_+ - \min \hat{\mathbf{Q}}_+$ is not uniformly distributed.

4. Conclusion

In this paper we have shown that a linear or quadratic 0/1 program **P** 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 and k-cluster examples, it is expected that the lower bound obtained from the Shor relaxation of this MAX-CUT formulation 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 when the Shor relaxation of the MAX-CUT formulation is compared with a convex quadratic relaxation.

13

References

[1] E. D. Andersen and K. D. Andersen. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In *High Performance Optimization* (H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds.), Kluwer Academic Publishers, Dordrecht, the Netherlands, 2000, pp. 197–232.

[2] I.M. Bomze. Copositive optimization - recent developments and applications, European J. Oper. Res. **216** (2012), 509–520.

[3] I.M. Bomze, E. de Klerk. Solving standard quadratic optimization problems via linear, semidenite and copositive programming, J. Global Optim. **24** (2002), 163–185.

[4] 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.

[5] E. de Klerk, D.V. Pasechnik. Approximation of the stability number of a graph via copositive programming, SIAM J. Optim. **12** (2002), 875–892.

[6] 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.

[7] D. Henrion, J.B. Lasserre, J. Lofberg. GloptiPoly 3: moments, optimization and semidefinite programming, Optim. Methods & Softwares **24** (2009), pp. 761–779.

[8] J.B. Lasserre. Optimisation globale et théorie des moments, C. R. Acad. Sci. Paris **331** (2000), Série 1, pp. 929–934.

[9] J.B. Lasserre. Global optimization with polynomials and the problem of moments, SIAM J. Optim. **11** (2001), pp. 796–817.

[10] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, Imperial College Press, London, 2010.

[11] J.B. Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization*, Cambridge University Press, Cambridge, 2015.

[12] M. Laurent, Zhao Sun. Handelman's hierarchy for the maximum stable set problem, *J. Global Optim.* **60** (2014), pp. 393–423.

[13] M. Marshall. Error estimates in the optimization of degree two polynomials on a discrete hypercube, *SIAM J. Optim.* **16** (2005), 297–309.

[14] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization, *Optim. Methods Softw.* **9** (1998), 141–160.

[15] L. Palagi, V. Piccialli, F. Rendl, G. Rinaldi, A. Wiegele. 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. 821–847.

[16] F. Rendl. Semidefinite relaxations for partitioning, assignment and ordering problems, 4OR **10**, pp. 321–346, 2012.

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