A STRENGTHENED SDP RELAXATION via a SECOND LIFTING for the MAX-CUT PROBLEM

Miguel F. Anjos [*] Henry Wolkowicz [†]

October 27, 1999

University of Waterloo
Department of Combinatorics and Optimization
Waterloo, Ontario N2L 3G1, Canada

Research Report CORR 99-55

Key words: Max-cut problem, semidefinite relaxations.

Abstract

We present a strengthened semidefinite programming, SDP, relaxation for the Max-Cut problem, MC, and for the general quadratic boolean maximization problem. The well-known SDP relaxation can be obtained via Lagrangian relaxation and results in an SDP with variable $X \in \mathcal{S}^n$, the space of $n \times n$ symmetric matrices, and n constraints, $\operatorname{diag}(X) = e$, where e is the vector of ones. The strengthened bound is based on applying a *lifting procedure* to this well-known semidefinite relaxation after adding the nonlinear constraints $X^2 - nX = 0$ and $X \circ X = E$. The lifting procedure is again done via Lagrangian relaxation and results in an SDP with variable $Y \in \mathcal{S}^{t(n-1)+1}$, where $t(r) = r(r+1)/2$, and $2t(n-1)+1$ constraints. It is shown that the new bound obtained this way strictly improves the previous SDP bound, both empirically and theoretically.

[*]Research supported by an FCAR Ph.D. Research Scholarship. E-mail manjos@math.uwaterloo.ca

[†]Research supported by Natural Sciences Engineering Research Council Canada. E-mail hwolkowi@orion.math.uwaterloo.ca.

[0]This report is available by anonymous ftp at orion.math.uwaterloo.ca in directory pub/henry/reports or with URL:
http://orion.math.uwaterloo.ca/~hwolkowi/henry/reports/ABSTRACS.html

Contents

1 Introduction

Semidefinite programming, SDP, has become a very intense area of research in recent years; and, one main reason for this is its success in finding bounds for the Max-Cut problem, MC, and for general quadratic boolean maximization. The current bounds have proven to be very tight both theoretically and in numerical tests, see e.g. [9, 20, 13, 12, 11].

In this paper we present a strengthened SDP relaxation for MC, i.e. an SDP program that provides a strengthened bound, both empirically (see Tables 1 and 2) and theoretically (see Theorem 3.1), for MC relative to the current well-known SDP bound.

One approach to deriving the SDP relaxation is through the Lagrangian dual, see e.g. [23, 21], i.e. one forms the Lagrangian dual of the quadratic constrained quadratic model of MC. The dual of this Lagrangian dual yields the SDP relaxation, MCSDP, i.e. a convex program that consists of matrix inequality constraints. Thus we have lifted/linearized a nonlinear, nonconvex problem to the space of symmetric matrices. The result is a tractable convex problem. The strengthened bound is obtained by adding redundant constraints from the original MC to MCSDP and finding the dual of the Lagrangian dual again, i.e. applying a second lifting. Empirical tests and theory indicate a strict improvement in the strengthened bound.

The paper is organized as follows. The remaining part of this section contains notation on the various operators and their adjoints as well as the basics of MC. In Section 2 we derive the well-known semidefinite relaxation of MC, denoted MCSDP, using Lagrangian relaxation, see (2.1). The strengthened relaxation MCPSDP3 is derived in Section 3, see (3.10). The fact that one always gets a strict improvement is proved in Theorem 3.1. This follows from the surprising fact that the first row of the feasible matrices in the strengthened relaxation MCPSDP3 yields a feasible matrix for the relaxation MCSDP.

1.1 Background

1.1.1 Notation and Preliminaries

We work in the space of $n \times n$ symmetric matrices, $\mathcal{S}^n$, with the trace inner product $\langle A, B\rangle = \operatorname{trace} AB$, and dimension $t(n) = n(n+1)/2$. We let $A \circ B$ denote **Hadamard (elementwise) product**. For given $Y \in \mathcal{S}^{t(n)+1}$, the $t(n)$ vector $x = Y_{0,1:t(n)}$ denotes the first (zero-th) row of Y after the first element. We let e denote the vector of ones and $E = ee^T$ the matrix of ones; their dimensions will be clear from the context. We also let e_i denote the i^{th} unit vector and define the elementary matrices $E_{ij} = \frac{1}{2}(e_i e_j^T + e_j e_i^T)$.

Though it is true that a linear operator on a finite dimensional space can be expressed as a matrix, we use operator notation and operator adjoints. We find that this simplifies cumbersome notation in the long run.

For $S \in \mathcal{S}^n$, the vector $\operatorname{diag}(S) \in \Re^n$ is the **diagonal of** S, while the adjoint operator $\operatorname{Diag}(v) = \operatorname{diag}^*(v)$ is the **diagonal matrix** with diagonal formed from the vector $v \in \Re^n$. We use both $\operatorname{Diag}(v)$ and $\operatorname{Diag} v$ if the meaning is clear. (Similarly for diag and other operators.) Also, **the symmetric vectorizing operator** $s = \operatorname{svec}(S) \in \Re^{t(n)}$, is formed (columnwise) from S while ignoring the strictly lower triangular part of S. Its inverse is the **symmetrizing matrix operator** $S = \operatorname{sMat}(s)$. The adjoint of svec is the operator hMat $= \operatorname{svec}^*$ which forms a symmetric matrix where the off-diagonal terms are multiplied by a half, i.e. this satisfies

$$\operatorname{svec}(S)^T x = \operatorname{trace} S \operatorname{hMat}(x), \qquad \forall S \in \mathcal{S}^n, x \in \Re^{t(n)}.$$

The adjoint of sMat is the operator dsvec which works like svec except that the off diagonal elements are multiplied by 2, i.e. this satisfies

$$\operatorname{dsvec}(S)^T x = \operatorname{trace} S \operatorname{sMat}(x), \qquad \forall S \in \mathcal{S}^n, x \in \Re^{t(n)}.$$

For notational convenience, we define the **symmetrizing diagonal vector** $\operatorname{sdiag}(x) := \operatorname{diag}(\operatorname{sMat}(x))$ and the **vectorizing symmetric vector** $\operatorname{vsMat}(x) := \operatorname{vec}(\operatorname{sMat}(x))$, where vec is the n^2 dimensional vector formed from the complete columns of the matrix; the adjoint of vsMat is then given by

$$\operatorname{vsMat}^*(x) = \operatorname{dsvec}\left[\frac{1}{2}\left(\operatorname{Mat}(x) + \operatorname{Mat}(x)^T\right)\right].$$

In summary,

$$\begin{array}{rcl} \operatorname{diag}^* & = & \operatorname{Diag} \\ \operatorname{svec}^* & = & \operatorname{hMat} \\ \operatorname{svec}^{-1} & = & \operatorname{sMat} \\ \operatorname{dsvec}^* & = & \operatorname{sMat} \\ \operatorname{vsMat}^* & = & \operatorname{dsvec}\left[\frac{1}{2}\left(\operatorname{Mat}(\cdot) + \operatorname{Mat}(\cdot)^T\right)\right]. \end{array}$$

In this paper we will use the relationships between the following matrices and vectors:

$$X \cong vv^T \cong \operatorname{sMat}(x) \in \mathcal{S}^n, \text{ and } Y \cong \begin{pmatrix} y_0 \\ x \end{pmatrix} \begin{pmatrix} y_0 & x^T \end{pmatrix} \in \mathcal{S}^{t(n)+1}, \ y_0 \in \Re.$$

In particular, by abuse of notation, we will often use x and X interchangeably when the meaning is clear.

We will have need of several results on the operators. The following Lemma and Corollary follow directly from the definitions.

Lemma 1.1 *If $A, B \in \mathcal{S}^n$, then*

$$\operatorname{diag}(A(B \circ C)) = \operatorname{diag}(B(A \circ C)).$$

■

Corollary 1.1 *If $A, B \in \mathcal{S}^n$, then*

$$\operatorname{trace}(A(B \circ C)) = \operatorname{trace}(B(A \circ C)).$$

■

For these and other results on Hadamard products, see e.g. [14, 15].

1.1.2 Max-Cut Problem

The max-cut problem is the problem of partitioning the node set of an edge-weighted undirected graph into two parts so as to maximize the total weight of edges *cut* by the partition. We tacitly assume that the graph in question is complete (if not, nonexisting edges can be given weight 0 to complete the graph). Mathematically, the problem can be formulated as follows (see e.g [19]). Let the graph be given by its weighted adjacency matrix A. Define the matrix $L := \operatorname{Diag}(Ae) - A$ (L is called the *Laplacian matrix* associated with the graph.) If a cut S is represented by a vector v where $v_i \in \{-1, 1\}$ depending on whether or not $i \in S$, we get the following formulation for the max-cut problem:

$$\text{MC} \quad \begin{array}{rll} \mu^* := & \text{maximize} & \frac{1}{4} v^T L v \\ & \text{s.t.} & v \in \{-1,1\}^n. \end{array}$$

Using $X := vv^T$ and $v^T L v = \operatorname{trace} LX$, this is equivalent to

$$\begin{array}{rll} \mu^* = & \text{maximize} & \operatorname{trace} \frac{1}{4} LX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & \operatorname{rank}(X) = 1 \\ & & X \succeq 0. \end{array}$$

It is well known that MC is an NP-hard problem, see e.g. [17]. Dropping the rank condition and setting $Q = \frac{1}{4} L$, yields the SDP relaxation with the upper bound $\mu^* \leq \nu^*$, see MCSDP below. This relaxation of MC is now well known and studied in e.g. [7, 5, 8, 18]. Goemans and Williamson [8] have provided estimates for the quality of the SDP bound for MC. They have shown that the optimal value of this relaxation is at most 14% above the value of the maximum cut, provided there are no negative edge weights. In fact, by randomly rounding a solution to the SDP relaxation, they find a ρ-approximation algorithm, i.e. a solution with value at least ρ times the optimal value, where $\rho = .878$. However, It is NP-hard to find a ρ-approximation algorithm for Max-Cut with factor better than .9412, see [10]. Numerical tests are presented in e.g. [11, 13].

Improvements for special cases of MC are presented in e.g. [22]. Further results on problems with general quadratic objective functions are presented in [20, 27], e.g. Nesterov [20] uses the SDP bound to provide estimates of the optimal value of MC, with arbitrary $L = L^T$, with constant relative accuracy.

2 Lagrangian Relaxation

A quadratic model for MC with a general homogeneous quadratic objective function is

$$\text{MC} \quad \begin{array}{rcll} \mu^* & = & \max & v^TQv \\ & & \text{s.t.} & v_i^2 - 1 = 0, \quad i = 1, \dots, n. \end{array}$$

Note that if the objective function has a linear term, then we can homogenize using an additional variable similarly constrained (see the beginning of section 3.1). Furthermore, we assume $Q \neq 0$ (wlog) in the sequel.

2.1 SDP Relaxation of MC - First Lifting

The SDP relaxation comes from the Lagrangian dual of the Lagrangian dual of MC, see e.g. [23, 21]. For completeness we include the details of such a derivation. The Lagrangian dual to MC is

$$\mu^* \leq \nu^* := \min_y \max_v v^TQv - v^T(\operatorname{Diag} y)v + e^Ty.$$

Since the quadratic is bounded above only if its Hessian, $2Q - 2\operatorname{Diag} y$, is negative semidefinite, this is equivalent to the following SDP

$$\begin{array}{rll} \nu^* =: & \min & e^Ty \\ & \text{s.t.} & \operatorname{Diag} y \succeq Q. \end{array}$$

Slater's (strict feasibility) constraint qualification holds for this problem. Therefore its Lagrangian dual satisfies

$$\text{MCSDP} \quad \begin{array}{rcrll} \mu^* & \leq & \nu^* = & \max & \operatorname{trace} QX \\ & & & \text{s.t.} & \operatorname{diag}(X) = e \\ & & & & X \succeq 0. \end{array} \tag{2.1}$$

We get the same relaxation as above if we use the relationship or lifting $X = vv^T$ and $v^TQv = \operatorname{trace} QX$.

The above relaxation is equivalent to the Shor relaxation [23] and the S-procedure in Yakubovitch [25, 26]. For the case that the objective function or the constraints contain a linear term, extra work must be done to include the possibility of inconsistency of the stationarity conditions. Alternatively, this can be done by homogenization and using strong duality of the trust region subproblem [24]. The latter technique is used below.

3 Strengthened SDP Relaxation - Second Lifting

Suppose that we want to strengthen the above SDP relaxation. It is not clear what constraints one can add to MC to accomplish this. For example, one can add triangle inequalities, see e.g. [11, 22, 12, 13]. These triangle inequalities define the *metric polytope* M_n, i.e. the inequalities for X feasible for MCSDP are:

$$X_{ij} + X_{jk} + X_{ik} \geq -1 \quad X_{ij} - X_{jk} - X_{ik} \geq -1 \quad \forall 1 \leq i, j, k \leq n.$$

These inequalities model the fact that for any three connected nodes of the graph, either two or none of the edges are cut. The complexity of adding these cuts is described in e.g. [22]. But it is not the case that adding a certain subset of triangle inequalities will improve every instance of max-cut and there are too many such inequalities to add all of them. In fact, it is impossible to add valid linear constraints to improve the performance ratio, [16].

To strengthen MCSDP, we start with the lifted program MC2 below. It is equivalent to MC and is obtained by adding redundant constraints to MCSDP. This is motivated by the work in [3, 2] where it is shown that adding redundant constraints that involve terms of the type XX^T can lead to strong duality. First we add to MCSDP the constraint $X^2 - nX = 0$. This constraint is motivated by $X^2 = vv^Tvv^T$ and $v^Tv = n$. Note that this constraint implies $X \succeq 0$, so this latter constraint becomes redundant and may be removed. Moreover, we can simultaneously diagonalize X and X^2, therefore the eigenvalues of X must satisfy $\lambda^2 - n\lambda = 0$, which implies the only eigenvalues are 0 and n. Since the diagonal constraint implies that the trace of X is n, we conclude that X is rank one. Thus MC2 is equivalent to MC via the factorization $X = vv^T$ and $\operatorname{trace} QX = v^TQv$. We also add the redundant constraints $X \circ X = E$ to obtain MC2. Note that this constraint (together with $X \succeq 0$) also implies rank one, see Theorem 3.2.

Our starting program equivalent to MC is therefore

$$\text{MC2} \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & X \circ X = E \\ & & X^2 - nX = 0. \end{array} \tag{3.1}$$

Note that MC2 is itself a Max-Cut problem but with additional nonlinear constraints, $t(n)$ variables, and with the same optimal objective value as MC.

3.1 Second lifting via Lagrangian duality

We follow the procedure in e.g. [21, 28] and obtain the SDP relaxation by finding the Lagrangian dual of the Lagrangian dual. As a final step we remove any redundant constraints. (This illustrates the advantage of the double dual approach for finding SDP relaxations, i.e. redundant constraints are automatically removed at the end.)

In order to efficiently apply Lagrangian relaxation and not lose the information from the linear constraint we need to replace the constraint with the norm constraint $||\mathrm{diag}\,(X) - e||^2 = 0$ and homogenize the problem. We then lift this matrix problem into a higher dimensional matrix space. To keep the dimension as low as possible, we take advantage of the fact that $X = \mathrm{sMat}\,(x)$ is a symmetric matrix. We then express MC2 as

$$\text{MC2} \qquad \begin{array}{rll} \mu^* = & \max & \mathrm{trace}\;(Q\,\mathrm{sMat}\,(x))\,y_0 \\ & \text{s.t.} & \mathrm{sdiag}\,(x)^T \mathrm{sdiag}\,(x) - 2e^T \mathrm{sdiag}\,(x) y_0 + n = 0 \\ & & \mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x) = E \\ & & \mathrm{sMat}\,(x)^2 - n\,\mathrm{sMat}\,(x) y_0 = 0 \\ & & 1 - y_0^2 = 0 \\ & & x \in \Re^{t(n)}, y_0 \in \Re. \end{array} \tag{3.2}$$

Note that this problem is equivalent to the previous formulation since we can change X to $-X$ if $y_0 = -1$. An alternative homogenization would be to change the objective function to $\frac{1}{n}\mathrm{trace}\;\left(Q \mathrm{sMat}\,(x)^2\right)$. It appears that (the eigenvalues of) Q should determine which homogenization would be better, i.e. which would result in a larger class of Lagrange multipliers when taking the dual and therefore reduce the duality gap. (This should be looked at in the future.)

We now take the Lagrangian dual of this strengthened formulation, i.e. we use Lagrange multipliers $w \in \Re$ and $T, S \in \mathcal{S}^n$ and get

$$\begin{array}{rl} \mu^* \leq \nu_2^* := & \displaystyle\min_{w,T,S} \max_{x, y_0^2=1} \mathrm{trace}\;(Q \mathrm{sMat}\,(x))\,y_0 \\ & + w(\mathrm{sdiag}\,(x)^T \mathrm{sdiag}\,(x) - 2e^T \mathrm{sdiag}\,(x) y_0 + n) \\ & + \mathrm{trace}\, T(E - \mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x)) \\ & \quad + \mathrm{trace}\, S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x) y_0). \end{array} \tag{3.3}$$

We can now move the variable y_0 into the Lagrangian without increasing the duality gap, since this is a trust region subproblem and the Lagrangian relaxation of it is tight [24]. This yields

$$\begin{array}{c}\nu_2^* = \min_{t,w,T,S} \max_{x,y_0} \operatorname{trace}\,(Q\mathrm{sMat}\,(x))\,y_0 \\ + w(\mathrm{sdiag}\,(x)^T\mathrm{sdiag}\,(x) - 2e^T\mathrm{sdiag}\,(x)y_0 + n) \\ + \operatorname{trace} T(E - \mathrm{sMat}\,(x)\circ\mathrm{sMat}\,(x)) \\ + \operatorname{trace} S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x)y_0) \\ + t(1-y_0^2).\end{array} \tag{3.4}$$

The inner maximization of the above relaxation is an unconstrained pure quadratic maximization, i.e. the optimal value is infinity unless the Hessian is negative semidefinite in which case $x = 0, y_0 = 0$ is optimal. Thus we need to calculate the Hessian.

Using $\operatorname{trace} Q\mathrm{sMat}\,(x) = x^T\mathrm{dsvec}\,(Q)$, and pulling out a 2 for convenience later on, we get the constant part (no Lagrange multipliers) of the Hessian:

$$2H_c := 2\begin{pmatrix} 0 & \frac{1}{2}\mathrm{dsvec}\,(Q)^T \\ \frac{1}{2}\mathrm{dsvec}\,(Q) & 0\end{pmatrix}. \tag{3.5}$$

The nonconstant part of the Hessian is made up of a linear combination of matrices, i.e. it is a linear operator on the Lagrange multipliers. To make the quadratic forms in (3.4) easier to differentiate we note that

$$\mathrm{dsvec}\,\mathrm{Diag}\,\mathrm{diag}\,\mathrm{sMat} = \mathrm{sdiag}^*\mathrm{sdiag} \qquad (= \mathrm{Diag}\,\mathrm{svec}\,(I))$$

and rewrite the quadratic forms as follows:

$$\begin{array}{rcl} \mathrm{sdiag}\,(x)^T\mathrm{sdiag}\,(x) & = & x^T\left(\mathrm{dsvec}\,\mathrm{Diag}\,\mathrm{diag}\,\mathrm{sMat}\right)x; \\ e^T\mathrm{sdiag}\,(x) & = & \left(\mathrm{dsvec}\,\mathrm{Diag}\,e\right)^T x; \end{array}$$

$$\begin{array}{rcl} \operatorname{trace} S(\mathrm{sMat}\,(x))^2 & = & \langle \mathrm{sMat}\,(x), S\,\mathrm{sMat}\,(x)\rangle \\ & = & \mathrm{vsMat}\,(x)^T\mathrm{vec}\,(S\,\mathrm{sMat}\,(x)) \\ & = & x^T\mathrm{vsMat}^*\mathrm{vec}\,(S\mathrm{sMat}\,(x)) \\ & = & x^T\left[\mathrm{vsMat}^*\mathrm{vec}\,(S\mathrm{sMat}\,)\right]x \\ & = & x^T\left[\mathrm{vsMat}^*\mathrm{vec}\,S\mathrm{Mat}\,\mathrm{vsMat}\,\right]x \\ & = & x^T\left[(\mathrm{Mat}\,\mathrm{vsMat}\,)^*\,S\,(\mathrm{Mat}\,\mathrm{vsMat}\,)\right]x; \end{array}$$

$$\begin{array}{rcl} \operatorname{trace} T(\mathrm{sMat}\,(x)\circ\mathrm{sMat}\,(x)) & = & x^T\left\{\mathrm{dsvec}\,(T\circ\mathrm{sMat}\,(x))\right\} \\ & = & x^T\left(\mathrm{dsvec}\,(T\circ\mathrm{sMat}\,)\right)x, \end{array}$$

where the expression with S involves $\mathrm{vsMat}^*\mathrm{vec}$ instead of dsvec because $S\mathrm{sMat}\,(x)$ may not be symmetric. (It is easy to verify that $\mathrm{vsMat}^*\mathrm{vec}$

reduces to dsvec if S is symmetric.) However, the expression still is a congruence of S. The last expression follows from Corollary 1.1. For notational convenience, we let $\mathcal{H}(w,T,S,t)$ denote the *negative* of the nonconstant part of the Hessian, and we split it into four linear operators with the factor 2:

$$
\begin{array}{rcl}
2\mathcal{H}(w,T,S,t) & := & 2\mathcal{H}_1(w)+2\mathcal{H}_2(T)+2\mathcal{H}_3(S)+2\mathcal{H}_4(t) \\
& := & 2w\begin{pmatrix} 0 & (\mathrm{dsvec}\,\mathrm{Diag}\,e)^T \\ (\mathrm{dsvec}\,\mathrm{Diag}\,e) & -\mathrm{sdiag}^*\mathrm{sdiag} \end{pmatrix} \\
& & +\,2\begin{pmatrix} 0 & 0 \\ 0 & \mathrm{dsvec}\,(T\circ \mathrm{sMat}\,) \end{pmatrix} \\
& & +\,2\begin{pmatrix} 0 & \frac{n}{2}\mathrm{dsvec}\,(S)^T \\ \frac{n}{2}\mathrm{dsvec}\,(S) & (\mathrm{Mat}\,\mathrm{vsMat}\,)^*\, S\,(\mathrm{Mat}\,\mathrm{vsMat}\,) \end{pmatrix} \\
& & +\,2t\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
\end{array}
\tag{3.6}
$$

The elements of the above matrices may need clarification. The matrix $\mathrm{sdiag}^*\mathrm{sdiag} \in \mathcal{S}^{t(n)}$ is diagonal with elements determined using

$$
\begin{array}{rcl}
e_i^T\,(\mathrm{sdiag}^*\mathrm{sdiag}\,)\,e_j & = & \mathrm{sdiag}\,(e_i)^T\mathrm{sdiag}\,(e_j) \\
& = & \begin{cases} 1 & \text{if i=j=t(k), for some } k \\ 0 & \text{otherwise.} \end{cases}
\end{array}
$$

Similarly, letting $T=\sum_{ij} t_{ij}E_{ij}$, we have

$$
\mathrm{dsvec}\,\,(T\circ \mathrm{sMat}\,) = \sum_{ij} t_{ij}\mathrm{dsvec}\,\,(E_{ij}\circ \mathrm{sMat}\,).
$$

Then the matrix $\mathrm{dsvec}\,(E_{ij}\circ\mathrm{sMat}\,)$ is found from using

$$
e_k^T\left[\mathrm{dsvec}\,\,(E_{ij}\circ \mathrm{sMat}\,(e_l))\right].
$$

Similarly, we can find the elements of $(\mathrm{Mat}\,\mathrm{vsMat}\,)^*\,S\,(\mathrm{Mat}\,\mathrm{vsMat}\,)$ by cancelling vec and Mat and using

$$
e_k^T\mathrm{vsMat}\,^*\mathrm{vec}\,\,(S\mathrm{sMat}\,(e_l)).
$$

We can cancel the 2 in (3.6) and (3.5) and get the (equivalent to the Lagrangian dual) semidefinite program

$$
\text{MCDSDP2}\quad \nu_2^* = \begin{array}{cc} \min & nw+\mathrm{trace}\,ET+\mathrm{trace}\,0S+t \\ \text{s.t.} & \mathcal{H}(w,T,S,t)\succeq H_c. \end{array}
\tag{3.7}
$$

If we take T sufficiently positive definite and t sufficiently large, then we can guarantee Slater's constraint qualification. Therefore the dual of this

SDP has the same optimal value ν_2^* and it provides the strengthened SDP relaxation of MC:

$$
\text{MCPSDP2} \qquad \begin{array}{rcl} \nu_2^* = & \max & \operatorname{trace} H_c Y \\ & \text{s.t.} & \mathcal{H}_1^*(Y) = n \\ & & \mathcal{H}_2^*(Y) = E \\ & & \mathcal{H}_3^*(Y) = 0 \\ & & \mathcal{H}_4^*(Y) = 1 \\ & & Y \succeq 0, Y \in \mathcal{S}^{t(n)+1} \end{array} \tag{3.8}
$$

To help define the adjoint operators we partition Y as

$$
Y = \begin{pmatrix} Y_{00} & x^T \\ x & \bar{Y} \end{pmatrix}.
$$

It is straightforward to check that

$$
\mathcal{H}_2^*(Y) = \operatorname{sMat} \operatorname{diag}(\bar{Y}) \quad \text{and} \quad \mathcal{H}_4^*(Y) = Y_{00},
$$

so the constraints $\mathcal{H}_2^*(Y) = E$ and $\mathcal{H}_4^*(Y) = 1$ are equivalent to $\operatorname{diag}(Y) = e$. Also, $\mathcal{H}_1^*(Y)$ is twice the sum of the elements in the first row of Y corresponding to the positions of the diagonal of $\operatorname{sMat}(x)$ minus the sum of the same elements in the diagonal of $\bar{Y}$, i.e.

$$
\mathcal{H}_1^*(Y) = 2\operatorname{svec}(I_n)^T x - \operatorname{trace} \operatorname{Diag}(\operatorname{svec}(I_n))\bar{Y}.
$$

The constraint $\mathcal{H}_1^*(Y) = n$ effectively requires that $Y_{0,t(i)} = 1, \forall\, i = 1, \ldots, n$, as shown in the proof of Lemma 3.1 below. Finally, to find $\mathcal{H}_3^*(Y)$, recall that by definition,

$$
\langle \mathcal{H}_3(S), Y \rangle \quad = \quad n\operatorname{dsvec}(S)^T x - \left\langle (\operatorname{Mat} \operatorname{vsMat})^* S \, (\operatorname{Mat} \operatorname{vsMat}), \bar{Y} \right\rangle.
$$

Taking adjoints,

$$
\begin{aligned}
\langle S, \mathcal{H}_3^*(Y) \rangle &= \operatorname{trace} S n\operatorname{sMat}(x) - \langle S, (\operatorname{Mat} \operatorname{vsMat})\, \bar{Y}\, (\operatorname{Mat} \operatorname{vsMat})^* \rangle \\
&= \langle S, n\operatorname{sMat}(x) - (\operatorname{Mat} \operatorname{vsMat})\, \bar{Y}\, (\operatorname{Mat} \operatorname{vsMat})^* \rangle.
\end{aligned}
$$

Note that $(\operatorname{Mat} \operatorname{vsMat})^* = \operatorname{vsMat}^* \operatorname{vec}$ is essentially (and in the symmetric case reduces to) sMat^* except that it acts on possibly nonsymmetric matrices. Hence,

$$
\mathcal{H}_3^*(Y) = n\operatorname{sMat}(x) - (\operatorname{Mat} \operatorname{vsMat})\, \bar{Y}\, (\operatorname{Mat} \operatorname{vsMat})^*. \tag{3.9}
$$

Equivalently, $\mathcal{H}_3^*(Y)$ consists of the sums in MCPSDP3 below. The constraint $\mathcal{H}_3^*(Y) = 0$ is key to showing that for Y feasible in MCPSDP2, sMat (x) is always positive semidefinite (and in fact feasible for MCSDP). This is proved in Lemma 3.2.

We now prove that the feasible set of MCPSDP2 has no strictly feasible points.

Lemma 3.1 *If Y is feasible for MCPSDP2, then Y is singular.*

Proof. Let Y be feasible for MCPSDP2. The constraints $\mathcal{H}_2^*(Y) = E$ and $\mathcal{H}_4^*(Y) = 1$ together imply that $\operatorname{diag}(Y) = e$. The constraint $\mathcal{H}_1^*(Y) = n$ can be written as

$$2\operatorname{svec}(I_n)^T x - \operatorname{trace}\operatorname{Diag}(\operatorname{svec}(I_n))\bar{Y} = n,$$

with $Y = \begin{pmatrix} 1 & x^T \\ x & \bar{Y} \end{pmatrix}$. Since $\operatorname{diag}(Y) = e$, $\operatorname{trace}\operatorname{Diag}(\operatorname{svec}(I_n))\bar{Y} = n$ and so $\operatorname{svec}(I_n)^T x = n$, or equivalently $\sum_{i=1}^{n} Y_{0,t(i)} = n$. Now $Y \succeq 0$ implies every principal minor of Y is nonnegative, so $|Y_{0,t(i)}| \leq 1$ must hold (again because $\operatorname{diag}(Y) = e$). So $\sum_{i=1}^{n} Y_{0,t(i)} = n \Rightarrow Y_{0,t(i)} = 1, i = 1, \ldots, n$. Hence each of the 2×2 principal minors obtained from the subsets of rows and columns $\{0, t(i)\}, i = 1, \ldots, n$ equals zero. Hence Y is not positive definite. ∎

In section 4 we proceed to characterize the feasible set of MCPSDP2 in a lower dimensional space where it has strictly feasible points. However, we first discuss how to obtain the second lifting directly from MC2.

3.2 Alternative Derivation - Direct Second Lifting

We can see the SDP relaxation directly for MC2 using the relationship

$$Y \cong \begin{pmatrix} y_0 \\ x \end{pmatrix} \begin{pmatrix} y_0 & x^T \end{pmatrix}, \quad X = \operatorname{sMat}(x).$$

The advantage in this is that we can use the origin of X from MC to directly express the constraints that the elements of X are ± 1 and $\operatorname{diag}(X) = e$. Thus we get

$$\operatorname{diag}(Y) = e, \quad \text{and} \quad Y_{0,t(i)} = 1, \forall i = 1, \ldots, n.$$

We can also express the $t(n+1)$ constraints from $X^2 - nX = 0$. The constraints corresponding to equating the diagonal entries become redundant.

After they are removed, the result is the simplified SDP relaxation:

$$\text{MCPSDP3} \qquad \begin{array}{rll} \nu_2^* = & \max & \operatorname{trace} H_c Y \\ & \text{s.t.} & \operatorname{diag}(Y) = e \\ & & Y_{0,t(i)} = 1, \quad \forall i = 1, \ldots, n \\ & & \sum_{k=1}^{i} Y_{t(i-1)+k,t(j-1)+k} + \sum_{k=i+1}^{j} Y_{t(k-1)+i,t(j-1)+k} \\ & & \quad + \sum_{k=j+1}^{n} Y_{t(k-1)+i,t(k-1)+j} - nY_{0,t(j-1)+i} = 0 \\ & & \qquad \forall i, j \,\text{s.t.}\, 1 \leq i < j \leq n \\ & & Y \succeq 0, Y \in \mathcal{S}^{t(n)+1}. \end{array} \tag{3.10}$$

This problem has $2t(n) + 1$ constraints (and the constraints are full rank).

Before proceeding, it is worth noting that MCPSDP2 and MCPSDP3 are equivalent problems, i.e. their feasible sets are the same. This follows by using Theorem 9 of [21] and the discussion (preceeding the Theorem) therein.

For simplicity of notation, we also rewrite the constraints in MCPSDP3. For i, j such that $1 \leq i < j \leq n$, define the (symmetric) matrix Q_{ij} as:

$$Q_{ij} := \sum_{k=1}^{i} E_{t(i-1)+k,t(j-1)+k} + \sum_{k=i+1}^{j} E_{t(k-1)+i,t(j-1)+k} + \sum_{k=j+1}^{n} E_{t(k-1)+i,t(k-1)+j} - nE_{0,t(j-1)+i}\,.$$

Then MCPSDP3 can be written as:

$$\text{MCPSDP3} \qquad \begin{array}{rll} \nu_2^* = & \max & \operatorname{trace} H_c Y \\ & \text{s.t.} & \operatorname{diag} Y = e \\ & & \operatorname{trace} E_{0,t(i)} Y = 1, i = 1, \ldots, n \\ & & \operatorname{trace} Q_{ij} Y = 0, \forall\, i, j \in \{1, \ldots, n\}, i < j \\ & & Y \succeq 0, Y \in \mathcal{S}^{t(n)+1}. \end{array} \tag{3.11}$$

The first two sets of constraints imply that the 2×2 leading principal minor of any feasible Y for MCPSDP3 is all ones. Hence, every feasible Y for MCPSDP3 is singular, which is expected in light of Lemma 3.1 (recall that the feasible sets of MCPSDP2 and MCPSDP3 are equal).

For the remainder of this paper, we work with the formulation MCPSDP3 of our strengthened relaxation.

3.3 Properties of the Strengthened Relaxation

One surprising result is that the matrix obtained by applying sMat to the first row of a feasible Y is positive semidefinite, even though this nonlinear constraint was discarded in MC2.

Lemma 3.2 *Suppose that Y is feasible in MCPSDP3. Then*

$$\mathrm{sMat}\,\left(Y_{0,1:t(n)}\right) \succeq 0$$

and so is feasible in MCSDP.

Proof. For Y feasible for MCPSDP3, write

$$Y = \begin{pmatrix} 1 & x^T \\ x & \bar{Y} \end{pmatrix},$$

with $x = Y_{0,1:t(n)}$. Note that $\bar{Y}$ is a principal submatrix of Y and therefore $\bar{Y} \succeq 0$.

By (3.9), the constraint $\mathcal{H}_3^*(Y) = 0$ is equivalent to

$$\mathrm{sMat}\,(x) = \frac{1}{n}\,(\mathrm{Mat}\,\mathrm{vsMat}\,)\,\bar{Y}\,(\mathrm{Mat}\,\mathrm{vsMat}\,)^*$$

and thus $\mathrm{sMat}\,(x)$ is a congruence of the positive semidefinite matrix $\bar{Y}$. The result follows. ∎

The added nonlinear constraint $X^2 - nX = 0$ has the following interesting and useful properties.

Lemma 3.3 *Suppose that $X, \bar{X}$ are both feasible for MCSDP. Then*

$$\mathrm{trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) \geq 0. \tag{3.12}$$

Suppose, in addition, that both

$$(X^2 - nX) \neq 0, \quad (\bar{X}^2 - n\bar{X}) \neq 0,$$

and both $X, \bar{X} \in \mathcal{F}$, a face of $\mathcal{P}$, with $\bar{X} \in \mathrm{relint}\,\mathcal{F}$. Then

$$\mathrm{trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) > 0. \tag{3.13}$$

Proof. By pulling out a square root, we see that

$$\mathrm{trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) = \mathrm{trace}\,\{\sqrt{X}(X - nI)\sqrt{X}\}\{\sqrt{\bar{X}}(\bar{X} - nI)\sqrt{\bar{X}}\}.$$

This is an inner product of congruences of negative semidefinite matrices and so an inner product of negative semidefinite matrices. The first inequality follows by the fact that $\mathcal{P}$ is a self-polar cone, i.e.

$$\mathcal{P} = \mathcal{P}^+ := \{Z : \langle Z, X\rangle \geq 0, \quad \forall X \in \mathcal{P}\}.$$

(This can be shown using the square root of a positive semidefinite matrix, commutativity of the trace, and congruence.)

To prove the second (strict) inequality, let $U = [P|Q]$ be an orthogonal matrix such that the columns of P span the range space of $\bar{X}$, while the columns of Q span the null space of $\bar{X}$. A face can be characterized by either the range space or the null space of any matrix in its relative interior, see e.g. [4]. Therefore $P^T\bar{X}P = \bar{D} \succ 0$ and $P^TXP = D \succeq 0$, while $Q^T\bar{X}Q = 0$ and $Q^TXQ = 0$. This implies

$$U^T\bar{X}U = \begin{bmatrix} \bar{D} & 0 \\ 0 & 0 \end{bmatrix}, \qquad U^TXU = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.14}$$

Our hypothesis also implies that $nI - \bar{D} \succ 0$. Therefore,

$$\operatorname{trace}(X^2 - nX)(\bar{X}^2 - n\bar{X}) = \operatorname{trace}(D^2 - nD)(\bar{D}^2 - n\bar{D}) > 0.$$

■

We now prove that, unless there is no gap between MCSDP and MC, the relaxation MCPSDP3 **always** provides a strict improvement over MCSDP, as the following theorem shows.

Theorem 3.1 *The optimal values satisfy*

$$\nu_2^* \leq \nu^* \quad \textit{and} \quad \nu_2^* = \nu^* \Rightarrow \nu_2^* = \mu^*. \tag{3.15}$$

Proof. Suppose that

$$Y^* = \begin{pmatrix} 1 & x^{*T} \\ x^* & \bar{Y}^* \end{pmatrix}$$

solves MCPSDP3. From Lemma 3.2, it is clear that $\operatorname{sMat}(x^*)$ is feasible for MCSDP. Therefore,

$$\begin{aligned} \nu_2^* &= \operatorname{trace} H_c Y^* \\ &= (\operatorname{dsvec} Q)^T x^* \\ &= \operatorname{trace} Q \operatorname{sMat}(x^*) \\ &\leq \nu^*. \end{aligned}$$

This establishes the inequality in (3.15).

Now assume that we also have

$$\nu_2^* = \nu^*. \tag{3.16}$$

Then feasibility of $X^* := \mathrm{sMat}\,(x^*)$ implies that it must, in fact, be optimal for MCSDP. Recall that ν_2^* is defined in (3.4). Also, we can assume that $X^{*2} - nX^* \neq 0$, or we are done. Therefore, we can sandwich the optimal values and see that $X^* = \mathrm{sMat}\,(x^*)$ is also optimal for the min-max problem

$$\omega^* = \min_S \phi(S), \tag{3.17}$$

where

$$\begin{aligned}\phi(S) := \max_{\mathrm{diag}\,(\mathrm{sMat}\,(x))=e, \mathrm{sMat}\,(x)\succeq 0} \quad & F(S,x) := \mathrm{trace}\ (Q\mathrm{sMat}\,(x)) \\ & + \mathrm{trace}\, S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x)),\end{aligned} \tag{3.18}$$

i.e. since more Lagrange multipliers gives us a better bound, we get

$$\nu^* \geq \omega^* \geq \nu_2^*,$$

which then implies equality actually holds for all three values. For S optimal in (3.17), now define the feasible set of the inner maximization problem as

$$G := \{x : \mathrm{diag}\,(\mathrm{sMat}\,(x)) = e,\ \mathrm{sMat}\,(x) \succeq 0\}$$

and the optimal set for the given S

$$R(S) = \{x \in G : F(S,x) = \phi(S)\}.$$

It is clear that G is a convex compact set. Therefore, $R(S)$ is also compact by continuity of F. Moreover, $R(S)$ is a subset of the optimal set of MCSDP, a subset of a minimal face $\mathcal{F}$ of $\mathcal{P}$, and, in fact, a feasible subset for MCSDP. Let $\bar{X} \in R(S) \cap \mathrm{relint}\, \mathcal{F}$. We now get the strict inequality

$$\mathrm{trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) > 0, \quad \forall x \in R(S), \tag{3.19}$$

from Lemma 3.3.

We now will apply [6, Theorem 2.1, page 188]. We see that the directional derivative of $\phi(S)$ in the direction $g = -(\bar{X}^2 - n\bar{X})$ exists and is given by

$$\max_{x\in R(S)} \left\langle \frac{\partial F(S,x)}{\partial S}, g \right\rangle.$$

By (3.19) and compactness we see that this must be negative, i.e. the directional derivative is negative which contradicts the fact that the two optimal values are equal. ■

We can also look at the added constraint $X \circ X = E$. Even though it does not imply $X \succeq 0$, it is interesting to note that adding only this constraint to MCSDP yields a problem equivalent to MC. This follows from the following theorem which characterizes all the $\{1,-1\}$-matrices in the positive semidefinite cone: they are exactly the rank-one matrices formed by the outer product of some $\{1,-1\}$ n-vector with itself. (This theorem follows as a Corollary to [15, Theorem 5.3.4][1]. We include a simple independent proof.)

Theorem 3.2 *Let X be an $n \times n$ symmetric matrix. Then*

$$X \succeq 0, X \in \{1,-1\}^{n\times n} \quad \textit{if and only if} \quad X = xx^T, \ \textit{for some} \ \ x \in \{1,-1\}^n.$$

Proof. Showing sufficiency is straightforward: if $X = xx^T$ then for any $y \in \Re^n$, we have

$$y^T X y = \|x^T y\|^2 \geq 0,$$

hence X is positive semidefinite.

To prove necessity, first observe that if X is symmetric, $X \in \{1,-1\}^{n\times n}$, and $X \succeq 0$, then all the diagonal entries of X equal 1.

If $n = 2$, the possibilities for X are

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

and it is easily checked that both are positive semidefinite and rank one.

For $n \geq 3$, we argue by contradiction. Suppose $X \in \{1,-1\}^{n\times n}$ and $X \succeq 0$ but X is not rank-one. Let $X = (x_{ij})$ and let x_i denote the i^{th} column of X. Then wlog (permuting columns if necessary) the first two columns of X are linearly independent, therefore $x_1 \neq x_2$ and $x_1 \neq -x_2$. Hence wlog (permuting rows if necessary) $x_{11} = x_{12} = 1$, and $x_{31} = -x_{32} = -1$ or $x_{31} = -x_{32} = 1$.

Thus the top left 3×3 principal submatrix of X is either

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}.$$

Since the determinants of these matrices are negative, we have a contradiction to $X \succeq 0$. Hence X is rank-one. Since $\operatorname{diag}(X) = e$, the result follows. ∎

[1]The authors thank Yin Zhang, Rice University, for this reference.

4 Geometry of the Feasible Set

We now study the geometrical structure of the feasible set of our relaxation. Let $\mathcal{Y}$ be the set of $Y \in \mathcal{S}^{t(n)+1}$ that are feasible for MCPSDP3. From Lemma 3.1, we have observed that $\mathcal{Y}$ has no strictly feasible points; so we seek to express $\mathcal{Y}$ in a lower dimensional space.

From the discussion on the direct second lifting, we know that the 2^n points

$$Y_v := \begin{pmatrix} 1 \\ x_v \end{pmatrix} \begin{pmatrix} 1 \\ x_v \end{pmatrix}^T, x_v := \operatorname{svec}(vv^T), v \in \mathcal{V} := \{-1, 1\}^n$$

all belong to $\mathcal{Y}$. Furthermore, since these are all the points we are interested in, we want to optimize over $\mathcal{F}$, the minimal face of the positive semidefinite cone (in $\mathcal{S}^{t(n)+1}$) s.t. $Y_v \in \mathcal{F},\ \forall v \in \mathcal{V}$.

Consider the barycenter of the set of points Y_v:

$$\hat{Y} := 2^{-n} \sum_{v \in \mathcal{V}} Y_v.$$

By definition of $\mathcal{F}$, $\hat{Y} \in \operatorname{relint} \mathcal{F}$. Since $\mathcal{F}$ is a proper face, we can find a mapping from a lower dimensional positive semidefinite cone to $\mathcal{F}$. We construct this mapping using the results of the next theorem, which describes some of the structure of $\hat{Y}$.

Let $P_{i,j}$ denote the $(t(n)+1) \times (t(n)+1)$ permutation matrix equal to the identity matrix with the i^{th} and j^{th} columns permuted. We define the (permutation) matrix P as:

$$P := \prod_{j=0}^{n} P_{j,t(j)} \qquad (= \prod_{j=2}^{n} P_{j,t(j)}).$$

Theorem 4.1 *1. $\hat{Y}$ is a $\{0,1\}$-matrix and*

$$\hat{Y}_{ij} = \begin{cases} 1, & \text{if} \quad i = t(k), j = t(l), k, l \in \{1, \dots, n\}, k \neq l \\ 1, & \text{if} \quad i = j \in \{0, 1, \dots, t(n)\} \\ 0, & \text{elsewhere.} \end{cases}$$

2. $\operatorname{rank}(\hat{Y}) = t(n-1) + 1$ and the eigenvalues of $\hat{Y}$ are $(n+1, 1, 0)$ with multiplicities $(1, t(n-1), n)$, respectively.

3.

$$\mathcal{N}(\hat{Y}) = \mathcal{R}\left(P \quad \begin{bmatrix} V \\ 0 \end{bmatrix} \right)$$

and

$$\mathcal{R}(\hat{Y}) = \mathcal{R}\left(P \quad \begin{bmatrix} e & 0 \\ 0 & I_{t(n-1)} \end{bmatrix} \right),$$

where $V \in \Re^{(n+1)\times n}$ *is any matrix s.t.* $\begin{bmatrix} e & V \end{bmatrix}$ *is an orthogonal matrix.*

Proof.

1. Let $v \in \mathcal{V}$ and consider Y_v. The elements of x_v have the form $(x_v)_j = v_\alpha v_\beta$, where $j = t(\beta - 1) + \alpha$ for $\alpha, \beta \in \{1, \ldots, n\}$, $\alpha \neq \beta$, and furthermore

$$(x_v)_j = \begin{cases} 1, & \text{if} \quad v_\alpha = v_\beta \\ -1, & \text{otherwise.} \end{cases}$$

First consider the case when $\alpha = \beta = k$; here $j = t(k)$ and it is clear that $(x_v)_{t(k)} = 1, k = 1, \ldots, n$. This holds independently of the choice of v so we may conclude that

$$\begin{pmatrix} 1 \\ x_v \end{pmatrix}_{t(k)} = 1,\ k = 0, \ldots, n, \quad \forall\, v \in \mathcal{V}. \tag{4.1}$$

Now suppose $\alpha \neq \beta$; then $v_\alpha = v_\beta$ for exactly 2^{n-1} elements of $\mathcal{V}$ and $v_\alpha \neq v_\beta$ for the other 2^{n-1} choices of v. Hence,

$$\sum_{v \in \mathcal{V}} (x_v)_j = 0, \forall\, j \notin \{t(0), \ldots, t(n)\}. \tag{4.2}$$

Equations (4.1) and (4.2) together imply that the 0^{th} column of $\hat{Y}$ equals $\sum_{k=0}^{n} e_{t(k)}$, i.e.

$$\hat{Y}_{i,0} = \begin{cases} 1, & \text{if} \quad i \in \{t(0), \ldots, t(n)\} \\ 0, & \text{otherwise.} \end{cases}$$

By symmetry of $\hat{Y}$, $\hat{Y}_{0,j} = \hat{Y}_{j,0}$, so it only remains to examine $\hat{Y}_{i,j}$ for $i, j = 1, \ldots, t(n)$.

The remaining $t(n)$ columns of $\hat{Y}$ are:

$$\hat{Y}_{:,j} = 2^{-n} \sum_{v \in \mathcal{V}} (x_v)_j \begin{pmatrix} 1 \\ x_v \end{pmatrix},$$

for $j = 1, \ldots, t(n)$. If $i = j$ then $\hat{Y}_{i,i} = 2^{-n} \sum_{v \in \mathcal{V}} (x_v)_i^2 = 1$, so we now suppose $i \neq j$.

If $i = t(k)$ and $j = t(l)$ for some $k, l \in \{1, \ldots, n\}$, $k \neq l$, then

$$\hat{Y}_{i,j} = 2^{-n} \sum_{v \in \mathcal{V}} (x_v)_{t(k)} (x_v)_{t(l)} = 1,$$

using (4.1).

If $i \neq t(k)$, $\forall k$ but $j = t(l)$, then

$$\hat{Y}_{i,j} = 2^{-n} \sum_{v \in \mathcal{V}} (x_v)_i = 1,$$

using (4.1) and (4.2). The case $i = t(k)$ but $j \neq t(l)$, $\forall l$ is handled similarly.

Finally, if $i \neq t(k)$, $\forall k$, and $j \neq t(l)$, $\forall l$, then we need only observe that $((x_v)_i, (x_v)_j) = (1,1)$ in exactly 2^{n-2} elements of $\mathcal{V}$, and the same count also holds for each of the combinations $(1,-1), (-1,1), (-1,-1)$. Thus, $\sum_{v \in \mathcal{V}} (x_v)_i (x_v)_j = 0$. Hence $\hat{Y}_{i,j} = 0$.

2. Define

$$\hat{Y}_P := P^T \hat{Y} P = \begin{bmatrix} E & 0 \\ 0 & I_{t(n)-n} \end{bmatrix} \in \mathcal{S}^{(t(n)+1) \times (t(n)+1)}.$$

Since this is a similarity transformation, $\hat{Y}$ and $\hat{Y}_P$ have exactly the same eigenvalues, so it suffices to prove the result for $\hat{Y}_P$. Also, $\hat{Y}_P$ is block diagonal, so its eigenvalues are those of the blocks. The lower block has the eigenvalue 1 with multiplicity $t(n) - n = t(n-1)$ (we have the set of eigenvectors $e_{n+1}, \ldots, e_{t(n)}$). The upper block is clearly rank-1; since

$$\hat{Y}_P \begin{pmatrix} e \\ 0 \end{pmatrix} = (n+1) \begin{pmatrix} e \\ 0 \end{pmatrix},$$

$n+1$ is its only nonzero eigenvalue. For V as in the statement of the theorem, $\hat{Y}_P \begin{pmatrix} V \\ 0 \end{pmatrix} = 0$. So the columns of V (extended with zeros) give a set of eigenvectors for the zero eigenvalue, which has multiplicity n.

3. The result follows by the similarity of $\hat{Y}$ and $\hat{Y}_P$ and the proof of the previous part of the theorem.

■

Now define the matrix

$$W := \quad P \quad \begin{bmatrix} e & 0 \\ 0 & I_{t(n-1)} \end{bmatrix} \quad \in \Re^{(t(n)+1)\times(t(n-1)+1)},$$

with $e \in \Re^{n+1}$. Then $\mathcal{R}(\hat{Y}) = \mathcal{R}(W)$ and we can use W to provide a mapping from $\mathcal{S}^{t(n-1)+1}$ to the minimal face $\mathcal{F}$: if $Y \in \mathcal{F}$, $Y = WZW^T$ for $Z \in \mathcal{S}^{t(n-1)+1}$, and we require $Z \succeq 0$ to stay in the positive semidefinite cone of the smaller space.

The projected version of MCPSDP3 is thus:

$$\begin{array}{rll} \nu_2^* = & \max & \text{trace}\,(W^T H_c W)Z \\ & \text{s.t.} & \text{trace}\,(W^T E_{ii} W)Z = 1, i = 0,\ldots,t(n) \\ & & \text{trace}\,(W^T E_{0,t(i)} W)Z = 1, i = 1,\ldots,n \\ & & \text{trace}\,(W^T Q_{ij} W)Z = 0, \forall\, i,j \in \{1,\ldots,n\}, i < j \\ & & Z \succeq 0, Z \in \mathcal{S}^{t(n-1)+1}. \end{array}$$

We now remove all the redundant constraints in this program.

Let w_i denote the i$^{\text{th}}$ column of W^T. The construction of W implies $w_0^T = w_{t(i)}^T = e_0^T$, $\forall i \in \{1,\ldots,n\}$, and the remaining columns of W:

$$\{w_{t(j-1)+i}^T : i,j \in \{1,\ldots,n\}, i < j\} = \{e_1^T, e_2^T, \ldots, e_{t(n-1)}^T\}$$

form a linearly independent set. (Together with w_0^T, they form a basis for $\Re^{t(n-1)+1}$.)

Since $W^T E_{ii} W = w_i w_i^T$ and $W^T E_{0,t(i)} W = \frac{1}{2}(w_0 w_{t(i)}^T + w_{t(i)}^T w_0^T)$, we have

$$W^T E_{t(i),t(i)} W = w_0 w_0^T = W^T E_{00} W, \forall\, i \in \{1,\ldots,n\}$$

and

$$W^T E_{0,t(i)} W = w_0 w_0^T = W^T E_{00} W, \forall\, i \in \{1,\ldots,n\}.$$

These observations allow us to remove $2n$ redundant constraints and obtain the projected SDP:

$$\text{MCPSDP3}_\text{P} \quad \begin{array}{rll} \nu_2^* = & \max & \text{trace}\,(W^T H_c W)Z \\ & \text{s.t.} & \text{trace}\,(W^T E_{ii} W)Z = 1, i \in \{0,1,\ldots,t(n)\}\backslash\{t(1),\ldots,t(n)\} \\ & & \text{trace}\,(W^T Q_{ij} W)Z = 0, \forall\, i,j \in \{1,\ldots,n\}, i < j \\ & & Z \succeq 0, Z \in \mathcal{S}^{t(n-1)+1}. \end{array}$$

To verify that there is no more redundancy in the constraints, observe that the first set of equality constraints is now equivalent to the (linear) constraint $\operatorname{diag}(Z) = e$, that $\operatorname{diag}(W^T Q_{ij} W) = 0, \forall\, i, j$ above, and that the first column of each matrix $W^T Q_{ij} W$ is equal to $\frac{-n}{2} w_{t(j-1)+i}$. Since these columns are all linearly independent, we conclude all the constraints are indeed linearly independent.

5 Solving the strengthened relaxation

We now solve the projected strengthened relaxation MCPSDP3$_\mathrm{P}$. First we simplify our notation. We have the following primal-dual pair:

$$
\text{(P)} \qquad
\begin{array}{rll}
p^* = & \max & \operatorname{trace}(W^T H_c W) Z \\
& \text{s.t.} & \operatorname{diag} Z = e \\
& & \operatorname{trace}(W^T Q_{ij} W) Z = 0, \forall\, (i,j) \in \mathcal{J} \\
& & Z \succeq 0, Z \in \mathcal{S}^{t(n-1)+1}
\end{array}
$$

$$
\text{(D)} \qquad
\begin{array}{rll}
d^* = & \min & \sum_{i \in \mathcal{I}} x_i \\
& \text{s.t.} & S = \operatorname{Diag}(x) + \sum_{(i,j) \in \mathcal{J}} y_{ij} W^T Q_{ij} W - W^T H_c W \\
& & S \succeq 0, x \in \Re^{t(n-1)+1}, y \in \Re^{t(n-1)},
\end{array}
$$

where

$$
\begin{array}{l}
\mathcal{I} := \{1, \ldots, t(n-1)+1\} \\
\mathcal{J} := \{(i,j) : i, j \in \{1, \ldots, n\}, i < j\}.
\end{array}
$$

We begin by establishing that Slater's constraint qualification holds for the primal problem (P). This guarantees $p^* = d^*$, i.e. there is no duality gap.

Lemma 5.1 *Slater's constraint qualification holds for* (P).

Proof. We consider the matrix $\tilde{Z} := I_{t(n-1)+1}$. Since $\tilde{Z} \succeq 0$, we only need to verify that it satisfies the equality constraints.

Clearly, $\operatorname{trace} E_{ii} \tilde{Z} = 1, \ \forall\, i \in \mathcal{I}$. Now observe that

$$
\begin{aligned}
W W^T &= P \begin{bmatrix} e & 0 \\ 0 & I_{t(n-1)} \end{bmatrix} \begin{bmatrix} e^T & 0 \\ 0 & I_{t(n-1)} \end{bmatrix} P^T \\
&= P \begin{bmatrix} E & 0 \\ 0 & I_{t(n-1)} \end{bmatrix} P^T \\
&= P \hat{Y}_P P^T \\
&= \hat{Y},
\end{aligned}
$$

where $\hat{Y}_P$ is the matrix defined in the proof of Theorem 4.1.

Using this observation, the second set of equality constraints for $\tilde{Z}$ may be written as

$$\operatorname{trace} Q_{ij}\hat{Y} = 0, \ \forall (i,j) \in \mathcal{J},$$

and it is now straightforward to verify that these equalities hold:

$$\operatorname{trace} Q_{ij}\hat{Y} = 0 \quad \Leftrightarrow \quad \sum_{k=1}^{i} \hat{Y}_{t(i-1)+k,t(j-1)+k} + \sum_{k=i+1}^{j} \hat{Y}_{t(k-1)+i,t(j-1)+k} + \sum_{k=j+1}^{n} \hat{Y}_{t(k-1)+i,t(k-1)+j} - n\hat{Y}_{0,t(j-1)+i} = 0$$

and the right-hand side holds because for each choice of $(i,j) \in \mathcal{J}$, by Theorem 4.1(1), the entries of $\hat{Y}$ involved are all zero. ■

It is straightforward to prove that the same is true for the dual problem.

Lemma 5.2 *Slater's constraint qualification holds for* (D).

Proof. Since $\sum_{i \in \mathcal{I}} E_{ii} = I_{t(n-1)+1}$, if we choose $\tilde{y}_{ij} := 0 \ \forall (i,j) \in \mathcal{J}$ and $\tilde{x}_i := \|\text{dsvec}\,(Q)\|_1 + 1 \ \forall i \in \mathcal{I}$, the corresponding dual (slack) matrix is

$$\tilde{S} = (\|\text{dsvec}\,(Q)\|_1 + 1) I_{t(n-1)+1} - W^T H_c W$$

and clearly $\tilde{S}$ is strictly diagonal dominant and has all its diagonal entries positive. It follows that $\tilde{S}$ is positive definite. ■

5.1 Primal-dual interior-point algorithm

The relaxations MCSDP and MCPSDP3$_\text{P}$ were compared on a variety of problems using the software package SDPPACK (version 0.9 Beta) [1]. In the case of MCPSDP3$_\text{P}$, the matrix X in this section is obtained by

$$X = \text{sMat}\,((WZ^*W^T)_{0,1:t(n)}),$$

where Z^* is optimal for MCPSDP3$_\text{P}$. In all the test problems we used, the resulting matrix X was always found to be positive semidefinite. We also examined the numerical rank of X, i.e. the number of eigenvalues that appear to be nonzero. See Table 1 for some typical results. We also tested both bounds on random quadratic boolean problems (i.e. with negative weights allowed). Results are presented in Table 2.

n	Weight of optimal cut	MCSDP bound (% rel. error)	MCPSDP3_P bound (% rel. error)	Numerical rank of X
5	4	4.5225 (13.06%) $\rho = 0.8845$	4.2890 (7.22%) $\rho = 0.9326$	2
10	12	12.5 (4.17%) $\rho = 0.9600$	12.3781 (3.15%) $\rho = 0.9695$	4
7	56	56.4055 (0.72%) $\rho = 0.9928$	56.0954 (0.17%) $\rho = 0.9983$	3
8	30	30.2015 (0.67%) $\rho = 0.9933$	30.0000 (7.5e-09%) $\rho = 1.0000$	1
9	58	58.9361 (1.61%) $\rho = 0.9841$	58.1182 (0.20%) $\rho = 0.9980$	3
10	64	64.0811 (0.1268%) $\rho = 0.9987$	64 (0%) $\rho = 1.0000$	1
12	88	90.3919 (2.72%) $\rho = 0.9735$	89.5733 (1.79%) $\rho = 0.9824$	4
14	114	115.1679 (1.02%) $\rho = 0.9899$	114.5758 (0.51%) $\rho = 0.9950$	3
16	158	160.0201 (1.28%) $\rho = 0.9874$	159.1054 (0.70%) $\rho = 0.9931$	4

Table 1: The first line of results corresponds to solving both MC relaxations for a 5-cycle with unit edge-weights. The second line corresponds to the Petersen graph with unit edge-weights. All the other results come from randomly generated weighted graphs. For each bound, the rel. error is defined as the difference between the bound and the value of optimal cut divided by the value of the optimal cut, and ρ equals the value of the optimal cut divided by the bound.

Our results show the strengthened bound MCPSDP3_P yielding a strict improvement over MCSDP every time.

Because the matrix variable Y in MCPSDP3_P has $O(n^4)$ scalar variables, solving the relaxation using an interior-point method becomes very slow and requires a large amount of memory, even for moderate n. However, it is important to note that the strengthened relaxation MCPSDP3_P has a special sparsity structure. Future work will aim at exploiting this structure to develop a specialized algorithm that addresses the above mentioned limitations and allows to efficiently solve the relaxation for large instances of MC.

n	Weight of optimal cut	MCSDP bound (% rel. error)	MCPSDP3$_P$ bound (% rel. error)	Numerical rank of X
9	10	13.1744 (31.74%) $\rho = 0.7590$	11.4416 (14.42%) $\rho = 0.8740$	3
10	54	58.5410 (8.41%) $\rho = 0.9224$	56.1157 (3.92%) $\rho = 0.9623$	4
12	120	125.8493 (4.87%) $\rho = 0.9535$	122.7504 (2.29%) $\rho = 0.9776$	4
14	104	118.4940 (13.94%) $\rho = 0.8777$	113.1409 (8.79%) $\rho = 0.9192$	5
16	182	191.6495 (5.30%) $\rho = 0.9497$	187.7041 (3.13%) $\rho = 0.9696$	5

Table 2: Results for randomly generated quadratic boolean problems. For each bound, the rel. error is defined as the difference between the bound and the value of optimal cut divided by the value of the optimal cut, and ρ equals the value of the optimal cut divided by the bound.

6 Conclusion

We have presented an SDP relaxation that provides a strengthened bound for MC relative to the current well-known SDP bound for MC. Though the computation time required to solve this new SDP relaxation is large compared to the time for solving the well-known SDP relaxation, it is hoped that a specialized algorithm exploiting structure will improve this situation and that this new bound will be competitive both in time and in quality. In addition, provable quality estimates need to be shown.

References

[1] F. ALIZADEH, J.-P. HAEBERLY, M. V. NAYAKKANKUPPAM, M.L. OVERTON, and S. SCHMIETA. SDPpack user's guide – version 0.9 Beta. Technical Report TR1997–737, Courant Institute of Mathematical Sciences, NYU, New York, NY, June 1997.

[2] K.M. ANSTREICHER, X. CHEN, H. WOLKOWICZ, and Y. YUAN. Strong duality for a trust-region type relaxation of QAP. *Linear Algebra Appl.*, to appear, 1999.

[3] K.M. ANSTREICHER and H. WOLKOWICZ. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.*, to appear, 1999.

[4] G.P. BARKER and D. CARLSON. Cones of diagonally dominant matrices. *Pacific J. of Math.*, 57:15–32, 1975.

[5] C. DELORME and S. POLJAK. Laplacian eigenvalues and the maximum cut problem. *Math. Programming*, 62(3):557–574, 1993.

[6] V.F. DEM'JANOV and V.N. MALOZEMOV. *Introduction to Minimax.* Dover Publications, 1974. translated from Russian.

[7] M. X. GOEMANS. Semidefinite programming in combinatorial optimization. *Math. Programming*, 79:143–162, 1997.

[8] M.X. GOEMANS and D.P. WILLIAMSON. .878-approximation algorithms for MAX CUT and MAX 2SAT. In *ACM Symposium on Theory of Computing (STOC)*, 1994.

[9] M.X. GOEMANS and D.P. WILLIAMSON. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.

[10] J. HASTAD. Some optimal inapproximability results. Technical report, Royal Institute of Technology, Stockholm, Sweden, 1999.

[11] C. HELMBERG. *An interior point method for semidefinite programming and max-cut bounds.* PhD thesis, Graz University of Technology, Austria, 1994.

[12] C. HELMBERG and F. RENDL. A spectral bundle method for semidefinite programming. Technical Report ZIB Preprint SC-97-37, Konrad-Zuse-Zentrum Berlin, Berlin, Germany, 1997. to appear in SIOPT.

[13] C. HELMBERG, F. RENDL, R. J. VANDERBEI, and H. WOLKOWICZ. An interior point method for semidefinite programming. *SIAM J. Optim.*, pages 342–361, 1996.

[14] R.A. HORN and C.R. JOHNSON. *Matrix Analysis.* Cambridge University Press, New York, 1985.

[15] R.A. HORN and C.R. JOHNSON. *Topics in Matrix Analysis.* Cambridge University Press, New York, 1991.

[16] H. KARLOFF. How good is the Goemans-Williamson MAX CUT algorithm? *SIAM J. on Computing*, 29(1):336–350, 1999.

[17] R. M. KARP. Reducibility among combinatorial problems. In R. E. Miller and J.W. Thatcher, editors, *Complexity of Computer Computation*, pages 85–103. Plenum Press, New York, 1972.

[18] M. LAURENT, S. POLJAK, and F. RENDL. Connections between semidefinite relaxations of the max-cut and stable set problems. *Math. Programming*, 77:225–246, 1997.

[19] B. MOHAR and S. POLJAK. Eigenvalues in combinatorial optimization. In *Combinatorial Graph-Theoretical Problems in Linear Algebra*, IMA Vol. 50. Springer-Verlag, 1993.

[20] Y. NESTEROV. Quality of semidefinite relaxation for nonconvex quadratic optimization. Technical report, CORE, Universite Catholique de Louvain, Belgium, 1997.

[21] S. POLJAK, F. RENDL, and H. WOLKOWICZ. A recipe for semidefinite relaxation for (0,1)-quadratic programming. *J. Global Optim.*, 7:51–73, 1995.

[22] F. RENDL. Max-Cut approximations in graphs with triangles. Technical report, Technische Universitat Graz, Graz, Austria, 1997.

[23] N.Z. SHOR. Quadratic optimization problems. *Izv. Akad. Nauk SSSR Tekhn. Kibernet.*, 222(1):128–139, 222, 1987.

[24] R. STERN and H. WOLKOWICZ. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optim.*, 5(2):286–313, 1995.

[25] V. A. YAKUBOVICH. The S-procedure and duality theorems for nonconvex problems of quadratic programming. *Vestnik Leningrad. Univ.*, 1973(1):81–87, 1973.

[26] V. A. YAKUBOVICH. Nonconvex optimization problem: the infinite-horizon linear-quadratic control problem with quadratic constraints. *Systems Control Lett.*, 19(1):13–22, 1992.

[27] Y. YE. Approximating quadratic programming with bound and quadratic constraints. *mp*, 84:219–226, 1999.

[28] Q. ZHAO, S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Semidefinite programming relaxations for the quadratic assignment problem. *J. Comb. Optim.*, 2(1):71–109, 1998.