Semidefinite Programming for Discrete Optimization and Matrix Completion Problems*

Henry Wolkowicz † Miguel F. Anjos ‡

August 1, 2001

University of Waterloo
Department of Combinatorics & Optimization
Waterloo, Ontario N2L 3G1, Canada
Research Report CORR 2000-38

Key words: Semidefinite Programming, Discrete Optimization, Lagrangian Relaxation, Max-Cut Problem, Quadratic Assignment Problem, Graph Partitioning Problem, Max-Clique Problem, Stable Set Problem, Euclidean Distance Matrix, Matrix Completion Problem.

*Survey article for the proceedings of Discrete Optimization '99 where some of these results were presented as a plenary address.

†Research supported by The Natural Sciences and Engineering Research Council of Canada. E-mail hwolkowi@orion.math.uwaterloo.ca.

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

[0]This paper 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/ABSTRACTS.html

Abstract

Semidefinite Programming (SDP) is currently one of the most active areas of research in optimization. SDP has attracted researchers from a wide variety of areas because of its theoretical and numerical elegance as well as its wide applicability. In this paper we present a survey of two major areas of application for SDP, namely discrete optimization and matrix completion problems.

In the first part of this paper we present a recipe for finding SDP relaxations based on adding redundant constraints and using Lagrangian relaxation. We illustrate this with several examples. We first show that many relaxations for the Max-Cut problem (MC) are equivalent to both the Lagrangian and the well-known SDP relaxation. We then apply the recipe to obtain new strengthened SDP relaxations for MC as well as known SDP relaxations for several other hard discrete optimization problems.

In the second part of this paper we discuss two completion problems, the positive semidefinite and the Euclidean distance matrix completion problem. We present some theoretical results on the existence of such completions and then proceed to the application of SDP to find approximate completions. We conclude this paper with a new application of SDP to find approximate matrix completions for large and sparse instances of Euclidean distance matrices.

Contents

1 Introduction

There have been many survey articles written in the last few years on semidefinite programming (SDP) and its applicability to discrete optimization and matrix completion problems [2, 5, 37, 38, 39, 62, 77, 78, 106, etc.] This highlights the fact that SDP is currently one of the most active areas of research in optimization. In this paper we survey in depth two application areas where SDP research has recently made significant contributions. Several new results are also included.

The first part of the paper is based on the premise that Lagrangian relaxation is "best". By this we mean that good tractable bounds can always be obtained using Lagrangian relaxation. Since the SDP relaxation is equivalent to the Lagrangian relaxation, we explore approaches to obtain tight SDP relaxations for discrete optimization by applying a recipe for finding SDP relxations using Lagrangian duality. We begin by considering the Max-Cut problem (MC) in Section 2.2. We first present several different relaxations of MC that are equivalent to the SDP relaxation, including the Lagrangian relaxation, the relaxation over a sphere, the relaxation over a box, and the eigenvalue relaxation. This illustrates our theme on the strength of the Lagrangian relaxation. The question of which relaxation is most appropriate in practice for a given instance of MC remains open. Section 2.5 contains an overview of the main algorithms that have been proposed to compute the SDP bound, and Section 2.6 presents an overview of the known qualitative results about the quality of the SDP bound. We then proceed in Section 2.7 to derive new strengthened SDP relaxations for MC. To obtain these relaxations we apply the recipe for finding SDP relaxations presented in [101]. This recipe can be summarized as: add as many redundant quadratic constraints as possible; take the Lagrangian dual of the Lagrangian dual; remove redundant constraints and project the feasible set of the resulting SDP to guarantee strict feasibility. We also present several interesting properties of this tighter SDP relaxation. In particular, we show that it always improves on the well-known SDP relaxation whenever the latter is not optimal [7, 8]. In Section 3 we discuss the application of Lagrangian relaxation to general quadratically constrained quadratic problems and in Section 4 we present applications of the recipe to other discrete optimization problems, including the graph partitioning, quadratic assignment, max-clique and max-stable-set problems.

The second part of the paper presents several SDP algorithms for the positive semidefinite and the Euclidean distance matrix completion problems. The algorithms are shown to be efficient for large sparse problems. Section 5.1 presents some theoretical existence results for completions based on chordality. This follows the work in [42]. An approach to solving large sparse completion problems based on approximate completions [57] is outlined in Section 5.2. In Section 5.3 a similar approach for Euclidean distance matrix completions [1] is presented. However, the latter does not take advantage of sparsity and has difficulty solving large sparse problems. We conclude in Section 5.4 with a new characterization of Euclidean distance matrices from which we derive an algorithm that successfully exploits sparsity [3].

1.1 Notation and Preliminaries

We let $\mathcal{S}^n$ denote the space of $n \times n$ symmetric matrices. This space has dimension $t(n) := n(n+1)/2$ and is endowed with the trace inner product $\langle A, B\rangle = \operatorname{trace} AB$. We let $A \circ B$ denote the Hadamard (elementwise) matrix product and $A \succeq 0$ denote the Löwner partial order on $\mathcal{S}^n$, i.e. for $A \in \mathcal{S}^n$, $A \succeq 0$ if and only if A is positive semidefinite. We denote by $\mathcal{P}$ the cone of positive semidefinite matrices.

We also work with matrices in the space $\mathcal{S}^{t(n)+1}$. For given $Y \in \mathcal{S}^{t(n)+1}$, we index the rows and columns of Y by $0, 1, \ldots, t(n)$. We will be particularly interested in the vector x obtained from the first (0^{th}) row (or column) of Y with the first element dropped. Thus, in our notation, $x = Y_{0,1:t(n)}$.

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 $E_{ij} := \frac{1}{\sqrt{2}}(e_i e_j^T + e_j e_i^T), \quad \text{if } i \neq j$, $E_{ii} := \frac{1}{2}(e_i e_j^T + e_j e_i^T)$. For any vector $v \in \Re^n$, we let $\|v\| := \sqrt{v^T v}$ denote the ℓ_2 norm of v.

We use operator notation and operator adjoints. The adjoint of the linear operator $\mathcal{A}$ is denoted $\mathcal{A}^*$ and satisfies (by definition)

$$\langle \mathcal{A}\, x, y\rangle = \langle x, \mathcal{A}^* y\rangle\,,\ \forall x, y.$$

Given a matrix $S \in \mathcal{S}^n$, we now define several useful operators. The operator $\operatorname{diag}(S)$ returns a vector with the entries on the diagonal of S. Given $v \in \Re^n$, the operator $\operatorname{Diag}(v)$ returns an $n \times n$ diagonal matrix with the vector v on the diagonal. It is straightforward to check that Diag is the adjoint operator of diag. We use both $\operatorname{Diag}(v)$ and $\operatorname{Diag} v$ provided the meaning is clear, and the same convention applies to diag and all the other operators. The symmetric vectorizing operator svec satisfies $s = \operatorname{svec}(S) \in \Re^{t(n)}$ where s is formed column-wise from S and the strictly lower triangular part of S is ignored. Its inverse is the operator sMat , so $S = \operatorname{sMat}(s)$ if and only if $s = \operatorname{svec}(S)$. Note that the adjoint of svec is not sMat but $\operatorname{svec}^* = \operatorname{hMat}$ with $\operatorname{hMat}(s)$ being the operator that forms a symmetric matrix from s like sMat but also multiplies the off-diagonal terms by $\frac{1}{2}$ in the process. Similarly, the adjoint of sMat is the operator dsvec which acts like svec except that the off-diagonal elements are multiplied by 2.

For notational convenience, we also define the symmetrizing diagonal vector operator

$$\operatorname{sdiag}(x) := \operatorname{diag}(\operatorname{sMat}(x))$$

and the vectorizing symmetric vector operator

$$\operatorname{vsMat}(x) := \operatorname{vec}(\operatorname{sMat}(x)),$$

where $\operatorname{vec}(S)$ returns the n^2-dimensional vector formed column-wise from S like svec but with the complete columns of the matrix S. Note that the adjoint of vsMat is:

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

Let us summarize here some frequently used operators in this paper:

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

We will frequently use the following relationships between matrices and vectors:

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

2 The Max-Cut Problem

We begin our presentation with the study of one of the simplest NP-hard problems, albeit one for which SDP has been successful. The Max-Cut Problem (MC) is a discrete optimization problem on undirected graphs with weights on the edges. Given such a graph, the problem consists in finding a partition of the set of nodes into two parts, which we call shores, to maximize the sum of the weights on the edges that are cut by the partition (we say that an edge is cut if it has exactly one end on each shore of the partition). In this paper we shall assume that the graph in question is complete (if not, non-existing edges can be added with zero weight to complete the graph without changing the problem) and we require no restriction on the type of edge weights (so, in particular, negative edge weights are permitted).

Following [87], we can formulate MC as follows. Let the given graph G have node set $\{1, \ldots, n\}$ and let it be described by its weighted adjacency matrix $A = (w_{ij})$. Let $L := \operatorname{Diag}(Ae) - A$ denote the Laplacian matrix associated with the graph, where the linear operator Diag returns a diagonal matrix with diagonal formed from the vector given as its argument, and e denotes the vector of all ones. Let us also define the set $\mathcal{F}_n := \{\pm 1\}^n$, and let the vector $v \in \mathcal{F}_n$ represent any cut in the graph via the interpretation that the sets $\{i : v_i = +1\}$ and $\{i : v_i = -1\}$ form a partition of the node set of the graph. Then we can formulate MC as:

$$\text{(MC1)} \qquad \begin{array}{rll} \mu^* := & \max & \frac{1}{4} v^T L v \\ & \text{s.t.} & v \in \mathcal{F}_n, \end{array} \tag{2.1}$$

where here and throughout this paper μ^* denotes the optimal value of the MC problem.

It is straightforward to check that

$$\frac{1}{4} v^T L v = \sum_{i<j} w_{ij} \left(\frac{1 - v_i v_j}{2} \right)$$

and that the term multiplying w_{ij} in the sum equals one if the edge (i, j) is cut, and zero otherwise. Analogous quadratic terms having this property will be used in our formulations of MC.

We can view MC1 as the problem of maximizing a homogeneous quadratic function of v over the set $\mathcal{F}_n$. We show that this problem is equivalent to problem MCQ below which has a more general objective function. This equivalence shows that all the results about MC1 also extend to MCQ. Furthermore, the formulation MCQ will help us derive relaxations for MC in Sections 2.2 and 2.3.

Let us therefore consider the quadratic objective function

$$q_0(v) := v^T Q v - 2c^T v$$

(the meaning of the subscript 0 will become clear at the beginning of Section 2.2) and the corresponding ± 1-constrained quadratic problem MCQ:

$$\text{(MCQ)} \qquad \max_{v \in \mathcal{F}_n} q_0(v). \tag{2.2}$$

Clearly, MC1 corresponds to the choice $Q = \frac{1}{4}L$ and $c = 0$. Conversely, we can homogenize the problem MCQ by increasing the dimension by one. Indeed, given $q_0(v)$, define the $(n+1) \times (n+1)$ matrix Q^c obtained by adding a 0^{th} dimension to Q and placing the vector c in the new row and column, so that

$$Q^c := \begin{bmatrix} 0 & -c^T \\ -c & Q \end{bmatrix}. \tag{2.3}$$

If we consider the variable $\bar{v} = \begin{pmatrix} v_0 \\ v \end{pmatrix} \in \mathcal{F}_{n+1}$ and the new quadratic form

$$q_0^c(\bar{v}) := \bar{v}^T Q^c \bar{v} = v^T Q v - 2v_0(c^T v),$$

then we get an equivalent MC problem.

2.1 Higher-dimensional Embeddings of MC

We can express the feasible set $\mathcal{F}_n$ in several different ways by appropriately embedding all its points in spaces of varying dimensions. In this section we take a geometrical view of several such embeddings and the respective formulations of MC. Relaxations of these formulations will be considered in the remainder of Section 2.

1. If we define

$$\mathcal{F}_n(1) := \{v \in \Re^n : |v_i| = 1, i = 1, \ldots, n\}, \tag{2.4}$$

   then clearly $\mathcal{F}_n(1) = \mathcal{F}_n$ and the formulation MC1 corresponds to optimizing $\frac{1}{4}v^T L v$ over $\mathcal{F}_n(1)$.

For later reference, we note here that $\mathcal{F}_n(1)$ is the set of extreme points of the unit hypercube in $\Re^n$ (the ℓ_∞ norm unit ball). Furthermore, all the points $v \in \mathcal{F}_n(1)$ satisfy the constraints

$$\|v_i\|_2 = 1 \;\forall\, i, \qquad |v_i v_j| = |v_i^T v_j| = 1 \;\forall\, i < j.$$

We have deliberately added the transpose, even though the variables v_i are all scalars, to emphasize the similarity with the next embeddings.

2. For any given positive integer p, we can lift each of the variables v_i from a scalar to a vector of length p by defining

$$\begin{aligned} \mathcal{F}_n(p) \;\; := \;\; \{V \in \Re^{n \times p} : V = [v_1, \ldots, v_n]^T, \\ \|v_i\|_2 = 1 \;\forall\, i, |v_i^T v_j| = 1 \;\forall\, i < j\}. \end{aligned} \tag{2.5}$$

Note that if $p = 1$ then we simply recover $\mathcal{F}_n(1)$. For $p > 1$ the constraints on the inner products restrict the cosines of the angles between any two vectors v_i and v_j to equal ± 1. Hence for $i = 2, \ldots, n$, either $v_i = v_1$ or $v_i = -v_1$, and we can obtain a cut by choosing the sets $\{i : v_i^T v_1 = +1\}$ and $\{i : v_i^T v_1 = -1\}$ as the shores. Thus the objective function may be written as:

$$\sum_{i<j} w_{ij} \left(\frac{1 - v_i^T v_j}{2} \right)$$

or, in terms of the Laplacian, as:

$$\frac{1}{4} \operatorname{trace} V^T L V.$$

We have thus derived our second formulation of MC:

$$\text{(MC2)} \qquad \begin{array}{rll} \mu^* = & \max & \frac{1}{4}\operatorname{trace} V^T L V \\ & \text{s.t.} & v_i \in \Re^p, \|v_i\|_2 = 1 \quad \forall i \\ & & |v_i^T v_j| = 1 \quad \forall i < j \\ & & V = [v_1 \ldots v_n]^T. \end{array} \tag{2.6}$$

The commutativity of the arguments inside the trace means that

$$\frac{1}{4} \operatorname{trace} V^T L V = \operatorname{trace} \left(\frac{1}{4} L \right) V V^T,$$

and this observation leads us to an embedding of MC into $\mathcal{S}^n$, the space of symmetric $n \times n$ matrices, by rewriting MC2 in terms of the variable $X \in \mathcal{S}^n$ which is defined by

$$X_{ij} := v_i^T v_j, \quad \text{or equivalently,} \quad X := VV^T.$$

Then the constraint $\|v_i\|_2 = 1$ is equivalent to $\operatorname{diag}(X) = e$ and

$$|v_i^T v_j| = 1 \Leftrightarrow |X_{ij}| = 1, \quad \forall i < j.$$

Finally, $X = VV^T \Leftrightarrow X \succeq 0$, and our third formulation of MC is:

$$\text{(MC3)} \qquad \begin{array}{rll} \mu^* = & \max & \frac{1}{4}\operatorname{trace} LX \left(= \frac{1}{2} \sum\limits_{i<j} w_{ij}(1 - X_{ij}) \right) \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & |X_{ij}| = 1, \quad \forall i < j \\ & & X \succeq 0. \end{array} \qquad (2.7)$$

Note that although each X_{ij} can be interpreted as the cosine of the angle between some vectors v_i and v_j, the length p of these vectors does not appear explicitly in the formulation MC3.

Having derived MC3, we can obtain yet another formulation by applying the following Theorem:

Theorem 2.1 *([8, Theorem 3.2]) Let X be an $n \times n$ symmetric matrix. Then*

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

Thus we can replace the ± 1 constraint on the elements of X by the requirement that the rank of X be equal to one. Hence we obtain our fourth formulation of MC:

$$\text{(MC4)} \qquad \begin{array}{rll} \mu^* = & \max & \frac{1}{4}\operatorname{trace} LX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & \operatorname{rank}(X) = 1 \\ & & X \succeq 0, X \in \mathcal{S}^n. \end{array} \qquad (2.8)$$

3. We now introduce an embedding of MC in a space of even higher dimension. This embedding is interesting because of its connection to the strengthened SDP relaxations that we present in Section 2.7.

   For any given positive integer q, let us define the set:

$$\begin{array}{rl} \mathcal{F}_{t(n)+1}(q) := & \{U \in \Re^{(t(n)+1)\times q} : U = [u_0, u_1, \ldots, u_{t(n)}]^T, \\ & \|u_i\|_2 = 1 \ \forall\, i = 0, 1, \ldots, t(n), \\ & |u_0^T u_i| = 1 \ \forall\, i = 1, \ldots, t(n), \\ & \text{sMat}\ \left(u_0^T u_1, \ldots, u_0^T u_{t(n)}\right) \succeq 0\}, \end{array} \tag{2.9}$$

   where $t(i) = \frac{i(i+1)}{2}$. The constraints

$$|u_0^T u_i| = 1 \ \forall\, i \quad \text{and} \quad \text{sMat}\ \left(u_0^T u_1, \ldots, u_0^T u_{t(n)}\right) \succeq 0$$

   imply (by Theorem 2.1) that the matrix $X = \text{sMat}\ \left(u_0^T u_1, \ldots, u_0^T u_{t(n)}\right)$ has rank equal to one. By analogy with the previous embedding in $\mathcal{S}^n$, we can therefore write the following interpretation:

$$u_0^T u_{t(j-1)+i} = X_{ij} = v_i^T v_j, \quad \forall\, 1 \le i \le j \le n. \tag{2.10}$$

   This means that we can think of the cosines of the angles between u_0, the first row of U, and every other row $u_{t(j-1)+i}, 1 \le i \le j \le n$, as being equal to the cosines of the angles between the vectors v_i and v_j (corresponding to the indices i and j) in the previous embedding. We can thus write down the objective function in terms of the entries in the first row of U:

$$\sum_{i<j} w_{ij} \left(\frac{1 - u_0^T u_{t(j-1)+i}}{2} \right).$$

   Let us now define the matrix:

$$H_L := \begin{pmatrix} 0 & \frac{1}{2}\text{dsvec}\,(L)^T \\ \frac{1}{2}\text{dsvec}\,(L) & 0 \end{pmatrix}.$$

   Then, since

$$\sum_{i<j} w_{ij} \left(\frac{1 - u_0^T u_{t(j-1)+i}}{2} \right) = \frac{1}{4}\text{trace}\, H_L U U^T = \frac{1}{4}\text{trace}\, U^T H_L U,$$

we can write down our fifth formulation of MC:

$$\text{(MC5)} \qquad \begin{array}{rll} \mu^* = & \max & \frac{1}{4}\operatorname{trace} U^T H_L U \\ & \text{s.t.} & u_i \in \Re^q, \|u_i\|_2 = 1 \quad \forall i \\ & & |u_0^T u_i| = 1 \quad \forall i = 1, \dots, t(n) \\ & & \text{sMat}\left(u_0^T u_1, \dots, u_0^T u_{t(n)}\right) \succeq 0 \\ & & U = \left[u_0 \dots u_{t(n)}\right]^T. \end{array} \tag{2.11}$$

As for the remaining entries of U, we can interpret them as:

$$u_{t(j-1)+i}^T u_{t(l-1)+k} = (v_i^T v_j)(v_k^T v_l), \forall\, 1 \le i \le j \le n, \forall\, 1 \le k \le l \le n. \tag{2.12}$$

This interpretation is particularly interesting if we use again the analogy with the previous embedding as in (2.10). If $X = VV^T$ with $V \in \mathcal{F}_n(1)$ (so V is a column vector) then the elements of X always satisfy the equation $X = \frac{1}{n}X^2$, i.e. each entry of X is equal to the average of n products of entries of X. Using the interpretations (2.10) and (2.12), this is equivalent to the constraint that for each $k = 1, \dots, t(n)$, $u_0^T u_k$ be equal to the average of n specific elements $u_i^T u_j$ with $i, j \ge 1$. For the verification of the equation relating X and X^2 and a much detailed discussion of these interpretations, see Section 2.7.

We have thus embedded the feasible set $\mathcal{F}_n$ of MC in several different spaces and obtained corresponding formulations for MC. We now illustrate in the next three sections what we mean when we claim that the Lagrangian relaxation is "best". First, we introduce in Sections 2.2 and 2.3 a variety of (seemingly different) tractable relaxations obtained from these formulations. Then in Section 2.4 we present the Lagrangian relaxation and Theorem 2.3, which states that (surprisingly) the (upper) bounds on μ^* yielded by these relaxations, i.e. their optimal values, are all equal to the optimal value of the Lagrangian relaxation.

2.2 Relaxations for MC using $v_i \in \Re$

Let us begin our study of relaxations for MC by considering the embedding $\mathcal{F}_n(1)$ of the MC variables and the problem MCQ. We have already argued that MCQ is equivalent to MC. Before we continue, we show why it is helpful to allow a more general quadratic objective in this Section.

Consider the formulation MC1:

$$\begin{array}{rll} \mu^* := & \max & v^T Q v \\ & \text{s.t.} & v \in \mathcal{F}_n(1), \end{array}$$

where $Q = \frac{1}{4}L$, and recall that $\mathcal{F}_n(1)$ is the set of extreme points of the unit hypercube in $\Re^n$. One obvious relaxation is to optimize $v^T Q v$ over the entire hypercube. If we do so, the resulting relaxation falls into one of the following two cases:

1. If $Q \preceq 0$, i.e. Q is negative semidefinite, then the maximum over the hypercube is always equal to zero and is attained at the origin.

2. If Q is not negative semidefinite then (at least) one eigenvalue of Q is positive and Pardalos and Vavasis [99] showed that in this case the maximization of $v^T Q v$ over the hypercube is NP-hard. So the relaxation is no more tractable than the original problem.

Clearly we do not obtain a useful relaxation in either case.

By considering instead the problem MCQ, the objective function $q_0(v) = v^T Q v - 2c^T v$ has a linear term and this allows us to consider perturbations of $q_0(v)$ of the form

$$q_u(v) := v^T(Q + \text{Diag}\,(u))v - 2c^T v - u^T e, \tag{2.13}$$

with $u \in \Re^n$. It is important to note that if $v \in \mathcal{F}_n(1)$, then $v_i^2 = 1\ \forall\ i$ and therefore

$$q_u(v) = q_0(v) \qquad \forall\ v \in \mathcal{F}_n(1), \forall\ u \in \Re^n.$$

Hence,

$$\begin{array}{rll} \mu^* = & \max & q_u(v) \\ & \text{s.t.} & v \in \mathcal{F}_n(1) \\ & & u \in \Re^n. \end{array}$$

We now show how these perturbations help.

2.2.1 The Trivial Relaxation in $\Re^n$

For given $u \in \Re^n$ let us maximize the perturbed objective function without any of the constraints on v, i.e. let us consider the function

$$f_0(u) := \max_v q_u(v).$$

For any choice of u, this function gives us an upper bound on μ^*, since $\mu^* \leq f_0(u)$. Hence, minimizing $f_0(u)$ over all $u \in \Re^n$ gives us a (trivial) relaxation of MC:

$$\mu^* \leq B_0 := \min_u f_0(u). \tag{2.14}$$

Remark 2.2 *Note that $f_0(u)$ can take on the value $+\infty$. In particular, this will happen whenever the matrix $Q + \mathrm{Diag}\,(u)$ has at least one positive eigenvalue, since a quadratic function is unbounded above if the Hessian is not negative semidefinite. (In fact, a quadratic function is bounded above if and only if the Hessian is negative semidefinite and the stationarity equation is consistent. A proof of this well-known fact is given, for example, in [80, Lemma 3.6].) However, since (2.14) is a min-max problem, we can add the (hidden) semidefinite constraint $Q + \mathrm{Diag}\,(u) \preceq 0$ without changing the value of the bound B_0. The resulting problem is tractable since it consists of minimizing a convex function over a convex set. The trivial relaxation is thus equivalent to:*

$$B_0 = \min_{Q+\mathrm{Diag}\,(u)\preceq 0} f_0(u). \tag{2.15}$$

Furthermore, let us define the set

$$S := \left\{ u : u^T e = 0, \quad Q + \mathrm{Diag}\,(u) \preceq 0 \right\}.$$

Provided that $S \neq \emptyset$, it is shown in [102] that the optimality conditions for min-max problems imply that:

$$B_0 = \min_{u^T e = 0} f_0(u) = \min_{\substack{Q+\mathrm{Diag}\,(u)\preceq 0 \\ u^T e=0}} f_0(u).$$

2.2.2 The Trust-Region (Spherical) Relaxation in $\Re^n$

Next let us relax the feasible set $\mathcal{F}_n(1)$ to the sphere in $\Re^n$ of radius $\sqrt{n}$ and centered at the origin. (Note that all the points of $\mathcal{F}_n(1)$ are contained in this sphere.) If we define the function

$$f_1(u) := \max_{||v||^2=n} \; q_u(v), \tag{2.16}$$

then $\mu^* \leq f_1(u)$ for all u. This maximization problem is a trust-region subproblem and is tractable since its dual is a concave maximization problem over an interval [108, 111, 121]. Therefore we obtain the (tractable) trust-region relaxation:

$$\mu^* \leq B_1 := \min_u f_1(u). \tag{2.17}$$

2.2.3 The Box Relaxation in $\Re^n$

Alternatively, we can replace the spherical constraint with the box constraint or ℓ_∞ norm constraint (all the points of $\mathcal{F}_n(1)$ also lie in this unit box) and consider the function

$$\mu^* \leq f_2(u) := \max_{|v_i| \leq 1} q_u(v). \tag{2.18}$$

Since the maximization of a non-convex quadratic over the box constraint is NP-hard [99], we must add the hidden semidefinite constraint to make the calculation of $f_2(u)$ tractable and obtain the box relaxation:

$$\mu^* \leq B_2 := \min_{Q+\operatorname{Diag}(u) \preceq 0} f_2(u). \tag{2.19}$$

It is worth mentioning that it is precisely the addition of the hidden semidefinite constraint (to make the box relaxation tractable) that makes the bound B_2 equal to all the other bounds we are currently presenting (see Theorem 2.3).

2.2.4 The Eigenvalue Relaxation in $\Re^{n+1}$

We showed at the beginning of Section 2 how the problem MCQ can be homogenized at the price of increasing the dimension by 1. This homogenization yields three more bounds B_0^c, B_1^c and B_2^c via the same derivations used to obtain the bounds B_0, B_1 and B_2.

Given Q and c, recall the $(n+1) \times (n+1)$-matrix

$$Q^c = \begin{bmatrix} 0 & -c^T \\ -c & Q \end{bmatrix} \tag{2.20}$$

and the vector $\bar{v} = \begin{pmatrix} v_0 \\ v \end{pmatrix}$. By analogy with the previous relaxations, we define

$$q_u^c(\bar{v}) := \bar{v}^T(Q^c + \operatorname{Diag}(u))\bar{v} - u^T e \tag{2.21}$$

and the functions $f_i^c(u), i = 0, 1, 2$. Note that if v_0, the first component of $\bar{v}$, equals ± 1 then $q_u^c(\bar{v}) = q_u(v)$.

For brevity we discuss only the relaxation B_1^c analogous to the trust-region relaxation. This particular relaxation is interesting because it turns out to be equivalent to an eigenvalue bound for μ^*. Indeed, since

$$f_1^c(u) := \max_{||\bar{v}||^2 = n+1} q_u^c(\bar{v}),$$

it follows from the Courant-Fisher Theorem (e.g. [55], Theorem 4.2.11) that

$$f_1^c(u) = (n+1)\ \lambda_{\max}(Q^c + \mathrm{Diag}\,(u)) - u^T e,$$

where $\lambda_{\max}(\cdot)$ denotes the maximum eigenvalue of the matrix argument. Hence $f_1^c(u)$ is tractable and the (tractable) eigenvalue bound is

$$\mu^* \leq B_1^c := \min_u f_1^c(u). \tag{2.22}$$

2.3 Matrix Relaxations for MC using $v_i \in \Re^p$

We now introduce relaxations arising from the formulations of MC using the feasible set $\mathcal{F}_n(p)$ with $p > 1$.

2.3.1 The Goemans-Williamson Relaxation

This relaxation is obtained by considering the formulation MC2 and removing the (hard) ± 1 constraint on the inner products $v_i^T v_j$. The resulting relaxation gives us the bound B_3:

$$\begin{array}{rll} B_3 := & \max & \frac{1}{2}\sum_{i<j} w_{ij}(1 - v_i^T v_j) \\ & \text{s.t.} & \|v_i\|_2 = 1, \quad \forall i. \end{array} \tag{2.23}$$

We note that in their well-known qualitative analysis of the SDP relaxation (the next relaxation we present), Goemans and Williamson [40] proved and used the fact that this relaxation and the SDP relaxation are equivalent. (For more details on their qualitative analysis, see Section 2.6.)

2.3.2 The Semidefinite Relaxation

This relaxation can be derived in (at least) two different ways. One way is to relax the formulation MC3 by removing the ± 1 constraint on the elements of X. The result is the semidefinite programming problem

$$(\mathrm{SDP1}) \qquad \begin{array}{rll} B_4 := & \max & \mathrm{trace}\, QX \\ & \text{s.t.} & \mathrm{diag}\,(X) = e \\ & & X \succeq 0, \end{array} \tag{2.24}$$

where $Q = \frac{1}{4}L$. SDP1 is a convex programming problem and is therefore tractable [93].

Alternatively, this relaxation can be obtained from the formulation MC1 using the fact that the trace is commutative:

$$v^TQv = \operatorname{trace} v^TQv = \operatorname{trace} Qvv^T$$

and that for $v \in \mathcal{F}_n$, $X_{ij} = v_iv_j$ defines a symmetric, rank-one, positive semidefinite matrix X with diagonal elements 1. Therefore, we can lift the problem MC1 into the (higher dimensional) space $\mathcal{S}^n$ of symmetric matrices. This is an alternative way to derive the formulation MC4:

$$(\mathrm{MC4}) \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & \operatorname{rank}(X) = 1 \\ & & X \succeq 0, X \in \mathcal{S}^n\,. \end{array} \qquad (2.25)$$

Removing the rank-one constraint from MC4 yields the SDP1 relaxation and the bound B_4.

2.4 Strength of the Lagrangian Relaxation

Consider the problem MCQ and replace the constraint $v \in \mathcal{F}_n$ with the equivalent constraints $v_i^2 = 1, \forall\, i$. The result is yet another formulation of MC which we refer to as $\mathrm{MC_E}$:

$$(\mathrm{MC_E}) \qquad \begin{array}{rll} \mu^* = & \max & q_0(v) \\ & \text{s.t.} & v_i^2 = 1, \quad i = 1, \ldots, n. \end{array} \qquad (2.26)$$

It is straightforward to check that the Lagrangian dual of the problem $\mathrm{MC_E}$ is

$$\min_u \max_v q_0(v) + \sum_i u_i(v_i^2 - 1)$$

and that it yields precisely our first bound B_0.

It is shown in [102, 101] that all the above relaxations and bounds for MC are equivalent to the Lagrangian dual of $\mathrm{MC_E}$. The strong duality result for the trust-region subproblem [111] is the key for proving the following theorem:

Theorem 2.3 *All the bounds for MCQ discussed above are equal to the optimal value of the Lagrangian dual of the equivalent problem* $\mathrm{MC_E}$.

Hence our theme about the strength of the Lagrangian relaxation.

The application of Lagrangian relaxation to obtain quadratic bounds has been extensively studied and used in the literature, for example in [70] and more recently in [71]. The latter calls the Lagrangian relaxation the "best convex bound". Discussions on Lagrangian relaxation for non-convex problems also appear in [30]. More references are given throughout this paper.

2.5 Computing the Bounds

While it is true that all the relaxations we have presented so far yield the same bound, it is not necessarily true that all are equally efficient when it comes to computing bounds for MC. Since the qualitative analysis of Goemans and Williamson (see Section 2.6), a lot of research work has focused on the semidefinite relaxation SDP1. For this reason, and since all the bounds we have presented so far are equivalent to the SDP1 bound, we shall change our notation at this point and from now on denote the optimal value of SDP1 by ν_1^* (our subsequent SDP relaxations will be similarly indexed). It is also for that reason that this Section mostly focuses on algorithms for computing the bound ν_1^*. Nonetheless, we believe that it is still unclear at this time which are the best relaxations to use.

2.5.1 Computing the semidefinite programming bound

From a theoretical point of view, given a semidefinite programming problem, we can find in polynomial-time an approximate solution to within any (fixed) accuracy using interior-point methods. This follows from the seminal work of Nesterov and Nemirovskii much of which is summarized in [93]. They also implemented the first interior-point method for SDP in [92]. Independently, Alizadeh extended interior-point polynomial-time algorithms from linear programming to SDP and studied applications to discrete optimization [4, 5]. Non-smooth optimization methods for solving semidefinite programming problems have also been proposed (see e.g. [47]).

Before we proceed let us observe that $X = I$ is a strictly positive definite feasible point for SDP1 (usually refereed to as a Slater point) and therefore strong duality holds, i.e. both SDP1 and its dual DSDP1:

$$\text{(DSDP1)} \qquad \begin{array}{rcc} \nu_1^* := & \min & e^T y \\ & \text{s.t.} & Z = \operatorname{Diag}(y) - Q \\ & & Z \succeq 0, y \in \Re^n, \end{array} \qquad (2.27)$$

have the same optimal value ν_1^*. Hence to compute the bound it suffices to solve either one of these SDPs.

Most efficient interior-point methods available for solving SDPs (e.g. [49, 46, 114, 6, 94, 32, 89, 88, 20, 112]) are primal-dual methods that require solving a dense Newton system of dimension equal to the number of constraints. The solution of this system is then used as a search direction. Typically some form of line search is performed and it requires a few Cholesky factorizations of the matrix variables concerned to ensure the positive semidefinite constraints are not violated by the next iterate. Although current research is exploring ways to exploit sparsity in this framework (see e.g. [33, 34]), most interior-point approaches are still very slow when applied to large ($n \geq 1000$) instances of SDP1. (Practical applications typically have at least a few thousand variables.)

One important weakness of interior-point methods is that the matrix variables are usually dense even when the matrix Q and the linear constraints of the SDP are sparse and structured, as is the case for SDP1. Several researchers have therefore proposed alternative approaches to evaluate the bound ν_1^* which seek to exploit the structure of SDP1. We summarize here several promising approaches in this direction.

2.5.2 Solving the primal problem SDP1

A successful approach in this direction was introduced by Homer and Peinado [54] and improved on by Burer and Monteiro [21].

These algorithms can be interpreted as projected gradient methods applied to a constrained nonlinear reformulation of SDP1. More specifically, Homer and Peinado use the fact that the constraint $X \succeq 0, X \in \mathcal{S}^n$ is equivalently formulated as $X = VV^T, V \in \Re^{n \times n}$ (recall the connections between formulation MC2 with $p = n$ and formulation MC3). Burer and Monteiro improve on the efficiency of this approach by observing further that V can be restricted to be a lower triangular matrix, and hence simplify the computations involved in each iteration of the projected gradient method. We refer the reader to the above references for more details.

2.5.3 Solving the dual problem DSDP1

Another alternative to interior-point methods is the use of bundle methods for min-max eigenvalue optimization. As seen in Section 2.2.4, the MC problem is equivalent to the min-max eigenvalue problem

$$\begin{array}{ll} \min & e^T y + n\,\lambda_{\max}(Q - \mathrm{Diag}\,(y)) \\ \text{s.t.} & y \in \Re^n. \end{array} \tag{2.28}$$

Helmberg and Rendl [48] develop a suitable bundle method for solving this problem and report numerical results for relaxations of MC instances with up to $n = 3000$ nodes. A detailed survey of their work and related results appears in [47]. The min-max eigenvalue approach for more general SDPs is discussed in Section 3.1.

Finally, back in the realm of interior-point methods, Benson, Ye and Zhang [17] derived and implemented an efficient and promising potential-reduction affine scaling algorithm to solve DSDP1. This polynomial-time algorithm generates the Newton system very quickly by virtue of the special structure of the n linear constraints of SDP1. This approach is further improved by Choi and Ye [22] via the use of a preconditioned conjugate gradient method to accelerate the generation of an approximate solution for the Newton system. Computational results for problems of dimension up to $n = 14000$ are reported in [22].

2.5.4 Other relaxations

We conclude this section by recalling that there has been very little numerical experimentation with the other relaxations we have presented even though, for example, fast and efficient quadratic programming algorithms are available and could be used to compute the bound B_2 from the box relaxation. Furthermore it is possible that other relaxations could be better numerically in certain circumstances and therefore that the choice of tractable bound to use should dependent on the particular instance of the problem. We believe that more research is needed in this direction.

2.6 Qualitative Analysis of the Bounds

Several interesting results on the quality of the bound ν_1^*, and hence (by Theorem 2.3) on the quality of the Lagrangian relaxation, have been published in recent years. We have already mentioned the celebrated proof of Goemans and Williamson that, under the assumption that all the edge weights are non-negative, the SDP bound always satisfies

$$\mu^* \geq \alpha \nu_1^*, \tag{2.29}$$

where $\alpha = \min_{0 \le \theta \le \pi} \frac{2}{\pi} \frac{\theta}{1-\cos\theta} \approx 0.87856$. This immediately implies that $\nu_1^* \le 1.14\, \mu^*$, i.e. ν_1^* is guaranteed to overestimate μ^* by at most 14%.

Alternatively, we can state this result as follows. Let us define the quantity μ_* as the optimal value of the problem:

$$\begin{array}{rll} \mu_* := & \min & v^T Q v \\ & \text{s.t.} & v \in \mathcal{F}_n, \end{array} \tag{2.30}$$

where $Q = \frac{1}{4}L$. (Note that $\mu_* = 0$ in the absence of negative edge weights.) Goemans and Williamson [40] proved that

$$\frac{\mu^* - \nu_1^*}{\mu^* - \mu_*} \le (1-\alpha) \approx 0.1214. \tag{2.31}$$

Nesterov [91] proved that without any assumption on the matrix Q, the following result holds:

$$\frac{\mu^* - \nu_1^*}{\mu^* - \mu_*} \le \frac{4}{7}. \tag{2.32}$$

This line of analysis is extended further in [122, 95].

We now proceed to illustrating the application of Lagrangian relaxation to obtain tighter bounds for the MC problem.

2.7 Strengthened SDP Relaxations for MC

The results in Sections 2.2-2.4 may give the impression that we have the tightest possible tractable bound for MC. It turns out that this is not the case because adding redundant quadratic constraints to the MC formulations before applying Lagrangian relaxation makes it possible to obtain stronger bounds. In fact, the addition of redundant quadratic of the type that we use here was shown in [11, 10] to guarantee strong duality for certain problems where duality gaps can exist.

The process we employ to obtaining a strengthened SDP relaxation for MC is an illustration of the recipe to find SDP relaxations presented in [101]. This process also illustrates the power of using the Lagrangian relaxation to derive SDP relaxations. The recipe is roughly the following:

- Add redundant constraints to the MC formulation;

- Take the Lagrangian dual of the Lagrangian dual to obtain the SDP relaxation;

- Finally, remove all the redundant constraints in the SDP relaxation.

The first step of the recipe asks that we add redundant constraints to the MC formulation. In order to apply the recipe effectively, we shall make a particular choice of formulation from among the various formulations of MC presented in Section 2.1. Indeed, if we restrict ourselves to formulating MC over the feasible set $\mathcal{F}_n(1)$, then it is not clear what redundant constraints one can add. However, when MC is formulated over $\mathcal{F}_n(p)$, there are many constraints that can be added. Let us therefore recall the formulation MC3:

$$\text{(MC3)} \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} Q X \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & |X_{ij}| = 1, \quad \forall i < j \\ & & X \succeq 0 \end{array}$$

where $Q = \frac{1}{4}L$.

First we may consider adding linear constraints. Among the many linear inequalities we may add are the well-known triangle inequalities that define the metric polytope M_n [46, 48, 49]:

$$\begin{aligned} M_n := \{X \in \mathcal{S}^n \ : \operatorname{diag}(X) = e, \text{ and} \\ & X_{ij} + X_{ik} + X_{jk} \geq -1, X_{ij} - X_{ik} - X_{jk} \geq -1, \\ & - X_{ij} + X_{ik} - X_{jk} \geq -1, -X_{ij} - X_{ik} + X_{jk} \geq -1, \\ & \forall\, 1 \leq i < j < k \leq n\}. \end{aligned}$$

These inequalities model the easy observation that for any three mutually connected nodes of the graph, only two or none of the edges may be cut. There are $4\binom{n}{3}$ such inequalities, which is a rather large number of constraints to add to the SDP, and it is not the case that adding a certain subset of triangle inequalities will improve every instance of MC. Instead of adding these constraints to MC3, we will instead add certain quadratic constraints (see below) that are closely related to these inequalities.

Beyond the addition of linear constraints, the addition of redundant quadratic constraints can be particularly effective, as was already mentioned above. In fact, the appropriate choice of quadratic constraints will play an important role in our derivation of tighter bounds for MC.

Several interesting choices of quadratic constraints are available. One obvious possibility is to formulate the constraints $|X_{ij}| = 1$ in MC3 as quadratic constraints using the Hadamard product:

$$X \circ X = E,$$

where $E \in \mathcal{S}^n$ is the matrix of all ones. Let us in fact replace the absolute value constraints with these quadratic constraints and obtain the formulation:

$$\begin{array}{rll} \mu^* = & \max & \operatorname{trace} Q\,X \\ & \text{s.t.} & \operatorname{diag} X = e \\ & & X \circ X = E \\ & & X \succeq 0. \end{array} \tag{2.33}$$

Michel Goemans[1] recently suggested the following very interesting set of quadratic constraints:

$$X_{ij} = X_{ik}X_{kj}, \quad \forall\, 1 \leq i,j,k \leq n. \tag{2.34}$$

One interpretation for these constraints arises from the alternative derivation of MC4 in Section 2.3.2. If $X_{ij} = v_i v_j$ and $v_k^2 = 1$ for $k = 1, \ldots, n$ then

$$X_{ij} = v_i v_j = v_i v_k^2 v_j = v_i v_k \cdot v_k v_j = X_{ik} \cdot X_{kj}.$$

There is also a connection between these constraints and the triangle inequalities in the definition of the metric polytope above. This connection is used in the proof of Theorem 2.13.

For reasons that will be clear later, we do not add these constraints exactly as we have stated them. We shall instead add a weaker form of these constraints by virtue of the observation that

$$(X^2)_{ij} = \sum_{k=1}^{n} X_{ik}X_{kj}.$$

If, according to equation (2.34), each of the elements in the sum on the right equals X_{ij}, then $(X^2)_{ij} = n\,X_{ij}$ or equivalently

$$X^2 = n\,X.$$

This very useful quadratic constraint can alternatively be obtained by considering the formulation MC2 with $p = 1$ (or formulation MC4) and observing that if $X = vv^T, v \in \{\pm 1\}^n$, then

$$X^2 = (vv^T)(vv^T) = (v^T v)vv^T = nX.$$

[1]Presented at The 4th International Conference on High Performance Optimization Techniques, June 1999, Rotterdam, Netherlands.

Therefore we can add the redundant quadratic constraint $X^2 - nX = 0$ and obtain the formulation:

$$\begin{array}{rll} \mu^* = & \max & \operatorname{trace} Q X \\ & \text{s.t.} & \operatorname{diag} X = e \\ & & X \circ X = E \\ & & X^2 - nX = 0 \\ & & X \succeq 0. \end{array} \tag{2.35}$$

The constraint $X^2 - nX = 0$ will play a central role in the rest of this Section. In fact, we shall use it right away to argue that we can drop the constraint $X \succeq 0$ from our formulation (2.35) of MC. Indeed, because we can simultaneously diagonalize X and X^2, the constraint $X^2 - nX = 0$ implies that the eigenvalues of X must satisfy the equation $\lambda^2 - n\lambda = 0$. Therefore the only possible eigenvalues for X are 0 and n and we conclude that $X \succeq 0$ holds. (Let us note here that, by virtue of Lemma 2.8, we incur no loss by removing this constraint before proceeding.)

Hence after the first step of the recipe we have the following formulation of MC:

$$\text{(MC6)} \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} Q X \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & X \circ X = E \\ & & X^2 - nX = 0. \end{array} \tag{2.36}$$

The next step in the recipe is to form the Lagrangian dual of MC6 and then the dual of the dual. Before we construct the Lagrangian dual, we must pay special attention to the linear constraints $\operatorname{diag}(\operatorname{sMat}(x)) = e$ in order to avoid increasing the duality gap when we go to the dual. The following simple example illustrates what may happen.

Example 2.4 *Consider the problem*

$$\begin{array}{ll} \max & x^2 \\ s.t. & x = 0. \end{array}$$

Obviously the optimal value is 0*. However the Lagrangian dual has optimal value*

$$\inf_{\lambda} \max_{x} x^2 + \lambda x = +\infty,$$

so we have introduced a duality gap by lifting the linear constraint as it is. However, if we first replace the linear constraint by $x^2 = 0$ then the Lagrangian dual yields

$$\inf_{\lambda} \max_{x} x^2 + \lambda x^2 = 0.$$

Hence squaring the linear constraint eliminates the duality gap.

It is perhaps surprising that the trick illustrated in the Example works in general in our framework, i.e. replacing the constraints $Ax = b$ by $||Ax - b||^2 = 0$ before taking the dual ensures that $Ax = b$ holds in the dual. More precisely,

Theorem 2.5 *([101, Theorem 9]) Let $K \subset \Re^n$ be a finite set, let $q(x) = x^T Qx - 2c^T x$, $A \in \Re^{m\times n}$ and $b \in \Re^n$. Then there exists $\bar{\lambda} \in \Re$ such that*

$$\max_{x\in K}\{q(x) : ||Ax - b||^2 = 0\} = \max_{x\in K}\{q(x) - \lambda||Ax - b||^2 \forall\ \lambda \geq \bar{\lambda}.$$

Hence

$$\max_{x\in K}\{q(x) : ||Ax - b||^2 = 0\} = \min_{\lambda} \max_{x\in K} q(x) - \lambda||Ax - b||^2$$

and strong duality holds. ∎

In fact, the proof of the theorem shows that the quadratic penalty function is exact in the case that K is a finite set; thus by changing linear constraints to the norm squared constraint before lifting them we are ensuring that they hold after taking the dual [101]. This observation sheds some light on the success of SDP relaxation in discrete optimization. Finally, let us note that the effectiveness of this approach to lift the linear constraints can also be argued via the use of an augmented Lagrangian, i.e. the exactness can be obtained in this alternate way [80].

Let us now return to the application of the recipe. To reduce the number of variables by taking advantage of the symmetry in the problem, let us rewrite MC6 using the variable $x \in \Re^{t(n)}$ such that $x = \operatorname{svec}(X)$:

$$\begin{array}{rll} \mu^* = & \max & \operatorname{trace} Q \operatorname{sMat}(x) \\ & \text{s.t.} & \operatorname{diag}(\operatorname{sMat}(x)) = e \\ & & \operatorname{sMat}(x) \circ \operatorname{sMat}(x) = E \\ & & (\operatorname{sMat}(x))^2 - n\operatorname{sMat}(x) = 0 \\ & & x \in \Re^{t(n)}. \end{array}$$

Replacing the linear constraint by the norm constraint and homogenizing the problem using the scalar variable y_0, we have:

$$
\begin{array}{rll}
\mu^* = & \max & \operatorname{trace}\,(Q\,\mathrm{sMat}\,(x))\,y_0 \\
 & \text{s.t.} & \mathrm{sdiag}\,(x)^T \mathrm{sdiag}\,(x) - 2e^T \mathrm{sdiag}\,(x) y_0 + n = 0 \\
 & & E - \mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x) = 0 \\
 & & \mathrm{sMat}\,(x)^2 - n\,\mathrm{sMat}\,(x) y_0 = 0 \\
 & & 1 - y_0^2 = 0 \\
 & & x \in \Re^{t(n)}, y_0 \in \Re.
\end{array}
\tag{2.37}
$$

Note that this problem is equivalent to the previous formulation since we can change x to $-x$ if $y_0 = -1$.

We now write down the Lagrangian dual of 2.37 using Lagrange multipliers $w, t \in \Re$ and $T, S \in \mathcal{S}^n$:

$$
\begin{array}{c}
\mu^* \leq \nu_2^* := \min\limits_{t,w,T,S} \max\limits_{x,y_0} \operatorname{trace}\,(Q\mathrm{sMat}\,(x))\,y_0 \\
+ w(\mathrm{sdiag}\,(x)^T \mathrm{sdiag}\,(x) - 2e^T \mathrm{sdiag}\,(x) y_0 + n) \\
+ \operatorname{trace} T(E - \mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x)) \\
+ \operatorname{trace} S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x) y_0) \\
+ t(1 - y_0^2).
\end{array}
\tag{2.38}
$$

The inner maximization of the above relaxation is an unconstrained pure quadratic maximization whose optimal value is $+\infty$ unless the Hessian is negative semidefinite in which case $x = 0, y_0 = 0$ is optimal. Therefore let us calculate the Hessian.

Using $\operatorname{trace} Q\,\mathrm{sMat}\,(x) = x^T \mathrm{dsvec}\,(Q)$, and pulling out a 2 (for convenience later), we can express H_Q, the constant part (without Lagrange multipliers) of the Hessian as:

$$
2H_Q := 2 \begin{pmatrix} 0 & \frac{1}{2}\mathrm{dsvec}\,(Q)^T \\ \frac{1}{2}\mathrm{dsvec}\,(Q) & 0 \end{pmatrix}.
\tag{2.39}
$$

For notational convenience, we let $\mathcal{H}(w, T, S, t)$ denote the *negative* of the non-constant part of the Hessian, and we split it into four linear operators with the factor 2:

$$
\begin{array}{rcl}
2\mathcal{H}(w,T,S,t) & := & 2\mathcal{H}_1(w) + 2\mathcal{H}_2(T) + 2\mathcal{H}_3(S) + 2\mathcal{H}_4(t) \\
& = & 2w \begin{pmatrix} 0 & (\mathrm{dsvec}\,\mathrm{Diag}\,e)^T \\ (\mathrm{dsvec}\,\mathrm{Diag}\,e) & -\mathrm{sdiag}^*\mathrm{sdiag} \end{pmatrix} \\
& & \quad + 2 \begin{pmatrix} 0 & 0 \\ 0 & \mathrm{dsvec}\,(T \circ \mathrm{sMat}\,) \end{pmatrix} \\
& & \quad\quad + 2 \begin{pmatrix} 0 & \frac{n}{2}\mathrm{dsvec}\,(S)^T \\ \frac{n}{2}\mathrm{dsvec}\,(S) & (\mathrm{Mat}\,\mathrm{vsMat}\,)^* S\,(\mathrm{Mat}\,\mathrm{vsMat}\,) \end{pmatrix} \\
& & \quad\quad\quad + 2t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
\end{array}
\tag{2.40}
$$

We can cancel the 2 in (2.40) and (2.39) and get the (equivalent to the Lagrangian dual) semidefinite program DSDP2:

$$
(\mathrm{DSDP2}) \qquad \begin{array}{rcc} \nu_2^* = & \min & nw + \mathrm{trace}\,ET + \mathrm{trace}\,0S + t \\ & \text{s.t.} & \mathcal{H}(w,T,S,t) \succeq H_Q. \end{array} \tag{2.41}
$$

If we take T sufficiently positive definite and t sufficiently large, then we can guarantee Slater's constraint qualification. Therefore the dual of DSDP2 has the same optimal value ν_2^* and it provides a strengthened SDP relaxation of MC:

$$
(\mathrm{SDP2}) \qquad \begin{array}{rcc} \nu_2^* = & \max & \mathrm{trace}\,H_Q Y \\ & \text{s.t.} & \mathcal{H}_1^*(Y) = n \\ & & \mathcal{H}_2^*(Y) = E \\ & & \mathcal{H}_3^*(Y) = 0 \\ & & \mathcal{H}_4^*(Y) = 1 \\ & & Y \succeq 0, Y \in \mathcal{S}^{t(n)+1} \end{array} \tag{2.42}
$$

To help define the adjoint operators we partition Y as

$$
Y = \begin{pmatrix} Y_{00} & x^T \\ x & \bar{Y} \end{pmatrix}, \quad \bar{Y} \in \mathcal{S}^{t(n)}.
$$

It is straightforward to check that

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

so the constraints $\mathcal{H}_2^*(Y) = E$ and $\mathcal{H}_4^*(Y) = 1$ are equivalent to $\mathrm{diag}\,(Y) = e$. 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 $\mathrm{sMat}\,(x)$ minus the sum of the same elements in the diagonal of $\bar{Y}$, i.e.

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

The constraint $\mathcal{H}_1^*(Y) = n$ requires that $Y_{0,t(i)} = 1, \forall\, i = 1, \ldots, n$, as shown in the proof of Lemma 2.7 below.

Finally, to find $\mathcal{H}_3^*(Y)$, recall that by definition,

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

Taking adjoints,

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

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

$$\mathcal{H}_3^*(Y) = n\mathrm{sMat}\,(x) - (\mathrm{Mat}\,\mathrm{vsMat}\,)\,\bar{Y}\,(\mathrm{Mat}\,\mathrm{vsMat}\,)^*. \tag{2.43}$$

Equivalently, $\mathcal{H}_3^*(Y)$ consists of the sums in SDP2 below. The constraint $\mathcal{H}_3^*(Y) = 0$ is key to showing that for Y feasible for SDP2, $\mathrm{sMat}\,(x)$ is always positive semidefinite (and in fact feasible for SDP1).

The end result as an SDP with linear constraints. The last step of the recipe consists of removing the redundant constraints in this SDP. This is usually done using the structure of the problem. The result after deleting redundant constraints is the following SDP relaxation of MC (see [7] for details):

$$\text{(SDP2)} \qquad \begin{array}{rll} \nu_2^* = & \max & \mathrm{trace}\, H_Q\, Y \\ & \text{s.t.} & \mathrm{diag}\,(Y) = e \\ & & Y_{0,t(i)} = 1, i = 1, \ldots, n \\ & & Y_{0,T(i,j)} = \frac{1}{n} \sum\limits_{k=1}^{n} Y_{T(i,k),T(k,j)}, \;\; \forall\, i,j \text{ s.t. } 1 \leq i < j \leq n \\ & & Y \succeq 0, Y \in \mathcal{S}^{t(n)+1}, \end{array} \tag{2.44}$$

where

$$T(i,j) := \begin{cases} t(j-1) + i, \text{ if } i \leq j \\ t(i-1) + j, \text{ otherwise.} \end{cases} \tag{2.45}$$

Remark 2.6 *The indices for the linear constraints in SDP2 may be thought of as the entries of a matrix T constructed in the following way. Expanding the relationship $X = sMat(x)$ we have:*

$$X = \begin{pmatrix} x_1 & x_2 & x_4 & \dots \\ x_2 & x_3 & x_5 & \dots \\ & \dots & & x_{t(n)} \end{pmatrix}.$$

Let us now keep only the indices of the entries of x and thereby define the matrix T:

$$T = \begin{pmatrix} 1 & 2 & 4 & \dots \\ 2 & 3 & 5 & \dots \\ & \dots & & t(n) \end{pmatrix}.$$

In fact, there is still some redundancy in the constraints of SDP2 as we now show that Slater's constraint qualification does not hold.

Lemma 2.7 *If Y is feasible for SDP2, then Y is singular.*

Proof. Let Y be feasible for SDP2. 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 $\sum_{i=1}^{n} Y_{0,t(i)} = n \Rightarrow Y_{0,t(i)} = 1, i = 1, \dots, n$. Hence each of the 2×2 principal minors obtained from the subsets of rows and columns $\{0, t(i)\}, i = 1, \dots, n$ equals zero. Hence Y is not positive definite. ■

This result makes it possible to further reduce the number of constraints in SDP2 by projecting the problem onto the positive semidefinite cone of dimension $t(n-1)+1$. This is done in detail in [7].

2.7.1 Properties of the Strengthened Relaxation

We now state and prove some of the interesting properties of the relaxation SDP2.

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 not explicitly included in the formulation MC6.

Lemma 2.8 *Suppose that Y is feasible for SDP2. Then*

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

and so is feasible for SDP1.

Proof. For Y feasible for SDP2, 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 (2.43), the constraint $\mathcal{H}_3^*(Y) = 0$ is equivalent to

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

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

We now prove some of the interesting and useful properties of the quadratic constraint $X^2 - nX = 0$. These properties will help us in the proof of the strengthening result for SDP2 (Theorem 2.10).

Lemma 2.9 *Suppose that $X, \bar{X}$ are both feasible for SDP1. Then*

$$\text{trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) \geq 0. \tag{2.46}$$

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 the positive semidefinite cone $\mathcal{P}$, with $\bar{X} \in$ relint $\mathcal{F}$*. Then*

$$\text{trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) > 0. \tag{2.47}$$

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

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

Each of the matrices on the right-hand side is now a congruence of a positive semidefinite matrix, and hence is itself positive semidefinite. The first inequality now follows from the fact that $\mathcal{P}$ is a self-polar cone:

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

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. [14]). 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

$$U^T\bar{X}U = \begin{bmatrix} \bar{D} & 0 \\ 0 & 0 \end{bmatrix}, \qquad U^TXU = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.48}$$

Our hypothesis implies $(\bar{X}^2 - n\bar{X}) \neq 0$ and therefore, by Theorem 2.1, $\operatorname{rank}(\bar{X}) \geq 2$ which implies that $nI - \bar{D} \succ 0$. Similarly, $nI - D \succ 0$. Therefore,

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

■

We now have all the ingredients necessary to prove that, unless there is no gap between the relaxation SDP1 and MC, the relaxation SDP2 **always** provides a strict improvement over SDP1.

Theorem 2.10 *The optimal values satisfy*

$$\nu_2^* \leq \nu^* \quad \text{and} \quad \nu_2^* = \nu^* \Rightarrow \nu_2^* = \mu^*. \tag{2.49}$$

Proof. Suppose that

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

solves SDP2. From Lemma 2.8, it is clear that $\mathrm{sMat}\,(x^*)$ is feasible for SDP1. Therefore,

$$\begin{array}{rcl}\nu_2^* & = & \mathrm{trace}\, H_Q Y^* \\ & = & (\mathrm{dsvec}\, Q)^T x^* \\ & = & \mathrm{trace}\, Q \mathrm{sMat}\,(x^*) \\ & \leq & \nu^*.\end{array}$$

This establishes the inequality in (2.49).

Now assume that we also have

$$\nu_2^* = \nu^*. \tag{2.50}$$

Then feasibility of $X^* := \mathrm{sMat}\,(x^*)$ implies that it must, in fact, be optimal for SDP1. Recall that ν_2^* is defined in (2.38). 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

$$\omega^* = \min_S \phi(S), \tag{2.51}$$

where

$$\begin{array}{rl}\phi(S) := \displaystyle\max_{\mathrm{diag}\,(\mathrm{sMat}\,(x))=e, \mathrm{sMat}\,(x)\succeq 0} & F(S,x) := \mathrm{trace}\ (Q\mathrm{sMat}\,(x)) \\ & \qquad + \mathrm{trace}\, S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x)),\end{array} \tag{2.52}$$

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 (2.51), 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 SDP1, a subset of a minimal face $\mathcal{F}$ of $\mathcal{P}$, and, in fact, a feasible subset for SDP1. Let $\bar{X} \in R(S) \cap \operatorname{relint} \mathcal{F}$. We now get the strict inequality

$$\operatorname{trace}(X^2 - nX)(\bar{X}^2 - n\bar{X}) > 0, \quad \forall x \in R(S), \tag{2.53}$$

from Lemma 2.9.

We now will apply [26, 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 (2.53) 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. ■

2.7.2 A Further Strengthening of the Relaxation SDP2

We now examine SDP2 more closely and show how an even tighter relaxation can be obtained. It may be helpful to the reader at this point to reexamine the formulation MC5 since both SDP2 and the upcoming relaxation SDP3 have connections to that formulation.

Let us begin by recalling the alternative derivation of MC4 in Section 2.3.2 and the rank-one matrices $X = vv^T, v \in \{\pm 1\}^n$. We know that these matrices X have all their entries equal to ± 1. Hence the corresponding matrices Y feasible for SDP2 have all their entries in the first row and column equal to ± 1. Looking back to the formulation MC5 this statement corresponds to the (hard) constraints $|u_0^T u_i| = 1$ for all i.

Now let us consider the following constraints of SDP2:

$$Y_{0,T(i,j)} = \frac{1}{n} \sum_{k=1}^{n} Y_{T(i,k),T(k,j)}, \;\; \forall 1 \le i < j \le n, \tag{2.54}$$

for $Y = \begin{pmatrix} 1 & x^T \\ x & \bar{Y} \end{pmatrix}$ and $x = \operatorname{svec}(vv^T)$. The entry $Y_{0,T(i,j)}$ is in the first row of Y and therefore it is equal to 1 in magnitude. The corresponding constraint in (2.54) says that it must be equal to the average of n specific entries in the block $\bar{Y}$. But each of these n entries has magnitude at most

1, and hence for equality to hold, they must all have magnitude equal to 1, and in fact they must all equal $Y_{0,T(i,j)}$.

Let us state this observation in a different way. If Y and X are both rank-one, then the block $\bar{Y} = xx^T$ and $Y_{T(i,k),T(k,j)} = x_{T(i,k)}x_{T(k,j)} = v_i v_k \cdot v_k v_j$. But if $v_k^2 = 1$, then $Y_{T(i,k),T(k,j)} = v_i v_j = X_{ij} = Y_{0,T(i,j)}$.

There is yet another interpretation for this observation. Recall that we obtained the quadratic constraint $X^2 = n\,X$ by appropriately adding up (and thereby weakening) the quadratic constraints (2.34). What we are observing here is that the constraints (2.54) consist of sums that originate in this weakening of the constraints (2.34). Hence we can now "undo" these sums and retrieve a linearized version of the constraints (2.34) in terms of the entries of the matrix variable Y.

This discussion leads us to define the relaxation SDP3 as:

$$\text{(SDP3)} \qquad \begin{array}{rll} \nu_3^* = & \max & \operatorname{trace} H_Q Z \\ & \text{s.t.} & \operatorname{diag}(Z) = e \\ & & Z_{0,t(i)} = 1, i = 1, \ldots, n \\ & & Z_{0,T(i,j)} = Z_{T(i,k),T(k,j)}, \;\; \forall\, k, \forall\, 1 \le i < j \le n \\ & & Z \succeq 0, Z \in \mathcal{S}^{t(n)+1}. \end{array} \qquad (2.55)$$

We proceed to prove that SDP3 is a *strict* improvement on the addition of all the triangle inequalities to SDP1. First, let us define:

$$F_n := \{X \in \mathcal{S}^n \, : X = \operatorname{sMat}(Z_{0,1:t(n)}), Z \text{ feasible for SDP3}\}.$$

Since the feasible set of SDP3 is convex and compact, and since F_n is the image of that feasible set under a linear transformation, it follows that F_n is also convex and compact.

For completeness, we begin by proving that SDP3 is indeed a relaxation of MC. This is not guaranteed a priori since SDP3 is a strengthening of SDP2.

Lemma 2.11 *$C_n \subseteq F_n$.*

Proof. Consider an extreme point of C_n, $X = vv^T, v \in \{\pm 1\}^n$. Let $x = \operatorname{svec}(X)$ and $Z = \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T$. We show that Z is feasible for SDP3.

Clearly $Z \succeq 0$ and $Z_{0,0} = 1$. Since $x_{T(i,j)} = v_i v_j$, for $1 \le i \le j \le n$,

$$Z_{T(i,j),T(i,j)} = (x_{T(i,j)})^2 = v_i^2 v_j^2 = 1.$$

Therefore $\operatorname{diag}(Z) = e$. Also, $Z_{0,t(i)} = Z_{0,T(i,i)} = x_{T(i,i)} = {v_i}^2 = 1$. Finally, for $1 \leq i < j \leq n$,

$$\begin{aligned} Z_{T(i,k),T(k,j)} &= x_{T(i,k)}x_{T(k,j)} \\ &= v_i v_k v_k v_j \\ &= v_i v_j \\ &= x_{T(i,j)} \\ &= Z_{0,T(i,j)}. \end{aligned}$$

Hence, each $X = vv^T, v \in \{\pm 1\}^n$ has a corresponding Z feasible for SDP3, and so $X \in F_n$. Since C_n and F_n are convex, we are done. ∎

Clearly, every Z feasible for SDP3 is feasible for SDP2. Therefore, by Lemma 2.8 above, we have the inclusion:

Corollary 2.12 $F_n \subseteq \mathcal{E}_n$. ∎

Using Lemma 2.11, we observe that $\mu^* \leq \nu_3^* \leq \nu_2^* \leq \nu_1^*$. Furthermore the strengthening result of Theorem 2.10 also holds for SDP3.

We now exploit the fact that there is a strong connection between the quadratic constraints (2.34) and the triangle inequalities to prove the next theorem.

Theorem 2.13 $F_n \subseteq M_n$.

Proof. Suppose $X \in F_n$, then $X = \operatorname{sMat}(Z_{0,1:t(n)})$ for some Z feasible for SDP3. Since $Z_{0,t(i)} = 1\ \forall\, i$, it follows that $\operatorname{diag}(X) = e$ holds.

Given i, j, k such that $1 \leq i < j < k \leq n$, let $Z_{i,j,k}$ denote the 4×4 principal minor of Z corresponding to the indices $0, T(i,j), T(i,k), T(j,k)$. Let $a = X_{ij} = Z_{0,T(i,j)}, b = X_{ik} = Z_{0,T(i,k)}, c = X_{jk} = Z_{0,T(j,k)}$. Then

$$Z_{i,j,k} = \begin{pmatrix} 1 & a & b & c \\ a & 1 & c & b \\ b & c & 1 & a \\ c & b & a & 1 \end{pmatrix},$$

since $\operatorname{diag}(Z) = e$ and $Z_{0,T(i,j)} = Z_{T(i,k),T(k,j)}$, $Z_{0,T(i,k)} = Z_{T(i,j),T(j,k)}$ and $Z_{0,T(j,k)} = Z_{T(j,i),T(i,k)}$ all hold for Z feasible for SDP3. Now:

$$
\begin{aligned}
Z_{i,j,k} \succeq 0 \quad &\Leftrightarrow \quad \begin{pmatrix} 1 & c & b \\ c & 1 & a \\ b & a & 1 \end{pmatrix} - \begin{pmatrix} a \\ b \\ c \end{pmatrix} \begin{pmatrix} a & b & c \end{pmatrix} \succeq 0 \\
&\Leftrightarrow \quad \begin{pmatrix} 1-a^2 & c-ab & b-ac \\ c-ab & 1-b^2 & a-bc \\ b-ac & a-bc & 1-c^2 \end{pmatrix} \succeq 0 \\
&\Rightarrow \quad e^T \begin{pmatrix} 1-a^2 & c-ab & b-ac \\ c-ab & 1-b^2 & a-bc \\ b-ac & a-bc & 1-c^2 \end{pmatrix} e \geq 0.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
Z_{i,j,k} \succeq 0 \quad &\Rightarrow \quad 3-(a+b+c)^2+2(a+b+c) \geq 0 \\
&\Leftrightarrow \quad \gamma^2 - 2\gamma - 3 \leq 0, \text{where}\, \gamma := a+b+c \\
&\Leftrightarrow \quad (\gamma-3)(\gamma+1) \leq 0 \\
&\Leftrightarrow \quad -1 \leq \gamma \leq 3 \\
&\Rightarrow \quad a+b+c \geq -1.
\end{aligned}
$$

Therefore, $X_{ij} + X_{ik} + X_{jk} \geq -1$ holds for X.

Because multiplication of row and column i of $Z_{i,j,k}$ by -1 will not affect the positive semidefiniteness of $Z_{i,j,k}$, if we multiply the two rows and two columns of $Z_{i,j,k}$ with indices $T(i,k)$ and $T(j,k)$ and apply the same argument to the resulting matrix, we obtain the inequality $X_{ij} - X_{ik} - X_{jk} \geq -1$. Similarly, the inequalities $-X_{ij} + X_{ik} - X_{jk} \geq -1$ and $-X_{ij} - X_{ik} + X_{jk} \geq -1$ also hold. ■

We have thus proved the following:

Corollary 2.14 $C_n \subseteq F_n \subseteq \mathcal{E}_n \cap M_n$. ■

Appropriate examples are provided in [8] to prove the following theorem:

Theorem 2.15 $C_n \subsetneq F_n \subsetneq \mathcal{E}_n \cap M_n$ *for* $n \geq 5$. ■

2.7.3 Numerical Results

The relaxations SDP1, SDP2 and SDP3 were compared for several interesting problems using the software package SDPPACK (version 0.9 Beta) [6]. For completeness we also solved the linear relaxation over the metric polytope:

$$\begin{array}{ll} \max & \operatorname{trace} Q\,X \\ \text{s.t.} & X \in M_n. \end{array}$$

This relaxation is easily formulated as an LP and we solved it using the Matlab solver LINPROG. The results are summarized in Table 1. The value ρ equals the value of the optimal cut divided by the bound, and R.E. denotes the relative error with respect to the optimal cut.

Graph	μ^*	SDP1 bound	SDP2 bound	M_n bound	$\mathcal{E}_n \cap M_n$ bound	SDP3 bound
C_5	4	4.5225 $\rho = 0.8845$ R.E.: 13.06%	4.2889 $\rho = 0.9326$ R.E.: 7.22%	4.0000 $\rho = 1.0000$ R.E.: 0%	4.0000 $\rho = 1.0000$ R.E.: 0%	4.0000 $\rho = 1.0000$ R.E.: 0%
$K_5 \backslash e$	6	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.1160 $\rho = 0.9810$ R.E.: 1.93%	6.0000 $\rho = 1.0000$ R.E.: 0%	6.0000 $\rho = 1.0000$ R.E.: 0%	6.0000 $\rho = 1.0000$ R.E.: 0%
K_5	6	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.6667 $\rho = 0.9000$ R.E.: 11.11%	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.2500 $\rho = 0.9600$ R.E.: 4.17%
Given by $A(G)$	9.28	9.6040 $\rho = 0.9663$ R.E.: 3.49%	9.4056 $\rho = 0.9866$ R.E.: 1.35%	9.3867 $\rho = 0.9886$ R.E.: 1.15%	9.2961 $\rho = 0.9983$ R.E.: 0.17%	9.2800 $\rho = 1.0000$ R.E.: 0%
AW_9^2	12	13.5 $\rho = 0.8889$ R.E.: 12.50%	12.9827 $\rho = 0.9243$ R.E.: 8.19%	12.8571 $\rho = 0.9333$ R.E.: 7.14%	12.6114 $\rho = 0.9515$ R.E.: 5.10%	12.4967 $\rho = 0.9603$ R.E.: 4.14%
Pet.	12	12.5 $\rho = 0.9600$ R.E.: 4.17%	12.3781 $\rho = 0.9695$ R.E.: 3.15%	12.0000 $\rho = 1.0000$ R.E.: 0%	12.0000 $\rho = 1.0000$ R.E.: 0%	12.0000 $\rho = 1.0000$ R.E.: 0%
Rand. gen.	88	90.3919 $\rho = 0.9735$ R.E.: 2.72%	89.5733 $\rho = 0.9824$ R.E.: 1.79%	89.3333 $\rho = 0.9851$ R.E.: 1.52%	88.0029 $\rho = 1.0000$ R.E.: $3.3E-5$	88.0000 $\rho = 1.0000$ R.E.: $9.9E-7$

Table 1: Numerical comparison of all MC relaxations for small test problems

The test problems in Table 1 are as follows:

1. The first line of results corresponds to solving the three SDP relaxations for a 5-cycle with unit edge-weights.

2. The second line corresponds to the complete graph on 5 vertices with unit edge-weights on all edges except one, which is assigned weight zero.

3. The third line corresponds to the complete graph on 5 vertices with unit edge-weights. In this example, none of the four SDP relaxations attains the MC optimal value, and in fact they are not distinguishable. Only the linear relaxation M_n gives a noticeably weaker bound.

4. The fourth line corresponds to the graph defined by the weighted adjacency matrix

$$A(G) = \begin{pmatrix} 0 & 1.52 & 1.52 & 1.52 & 0.16 \\ 1.52 & 0 & 1.60 & 1.60 & 1.52 \\ 1.52 & 1.60 & 0 & 1.60 & 1.52 \\ 1.52 & 1.60 & 1.60 & 0 & 1.52 \\ 0.16 & 1.52 & 1.52 & 1.52 & 0 \end{pmatrix}.$$

   This problem is interesting because it shows a significant difference between SDP3 and all the other relaxations; in this case, SDP3 is the only relaxation that attains the MC optimal value.

5. The fifth line corresponds to the graph in Figure 1 with unit edge weights. This graph is the antiweb AW_9^2 and it is the hardest example[2] that the authors know for the relaxation $\mathcal{E}_n \cap M_n$. It is interesting that SDP3 performs better on this example than on the K_5 with unit edge weights.

6. The last two lines correspond to slightly larger graphs. The graph on 10 vertices is the Petersen graph with unit edge-weights. The graph on 12 vertices is a randomly generated graph that gives slightly different results for each relaxation (the exact description of the graph is in [8]).

In Table 1, a relative error equal to zero means that the relative error was below 10^{-11}, the value of the smallest default stopping criteria used by SDP-PACK.

[2] We thank Franz Rendl for suggesting this interesting example.

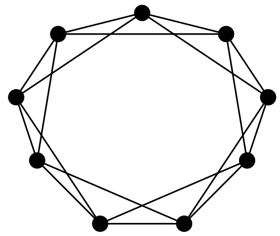

Figure 1: Antiweb AW_9^2

We conclude by pointing out that solving the relaxations SDP2 and SDP3 using an interior-point method becomes very time-consuming and requires large amounts of memory even for moderate values of n. Nonetheless their constraints are very sparse and have a special structure therefore it is hoped that research efforts like those mentioned in Section 2.5, or perhaps even entirely new approaches, will allow these relaxations to be solved efficiently for larger values of n.

3 SDP and Lagrangian Relaxation for Q^2Ps

We now move on to illustrate the Lagrangian relaxation approach for general quadratically constrained quadratic problems (Q^2P). In this Section we briefly outline the approach for the general Q^2P and specific instances are considered in some detail in Section 4. This general quadratic problem is also studied in e.g. [31, 68, 67, 69, 116] and [104, 74, 72, 84, 16]. The more general polynomial optimization problem is considered in [75] which presents a relaxation very similar to SDP3 but motivated by result results in the theory of moments and positive polynomials.

The quadratic problem we consider is the following Q^2P :

$$(\mathrm{Q}^2\mathrm{P}_x) \quad \begin{array}{rll} q^* := & \max & q_0(x) := x^TQ_0x + 2g_0^Tx + \alpha_0 \\ & \text{s.t.} & q_k(x) := x^TQ_kx + 2g_k^Tx + \alpha_k \le 0 \\ & & k \in \mathcal{I} := \{1,\dots,m\} \\ & & x \in \Re^n, \end{array} \tag{3.56}$$

where the matrices $Q_k \neq 0, k = 0,\dots,m$, are symmetric.

Let us define

$$P_k := \begin{bmatrix} \alpha_k & g_k^T \\ g_k & Q_k \end{bmatrix} \tag{3.57}$$

and (by abuse of notation)

$$q_k(y) := y^T P_k y, \ k = 0, 1, \ldots, m.$$

Using the technique for proving the equivalence of MC and MCQ at the beginning of Section 2, we obtain a homogenized formulation of $\mathrm{Q^2P}_x$ in terms of the new variable y and we denote it $\mathrm{Q^2P}_y$:

$$(\mathrm{Q^2P}_y) \quad \begin{array}{rll} q^* = & \max & q_0(y) \\ & \text{s.t.} & q_k(y) \le 0, k \in \mathcal{I} \\ & & y_0^2 = 1 \\ & & y = \begin{pmatrix} y_0 \\ x \end{pmatrix} \in \Re^{n+1}. \end{array} \tag{3.58}$$

If $y_0 = 1$ is optimal for $\mathrm{Q^2P}_y$, then y is optimal for $\mathrm{Q^2P}_x$ and if $y_0 = -1$ is optimal for $\mathrm{Q^2P}_y$, then $-y$ is optimal for $\mathrm{Q^2P}_x$. Hence the optimal values of $\mathrm{Q^2P}_x$ and $\mathrm{Q^2P}_y$ are equal.

The Lagrangian relaxation of the homogenized problem $\mathrm{Q^2P}_y$ provides a simpler path for obtaining the SDP relaxation. Indeed, the Lagrangian of $\mathrm{Q^2P}_y$ is

$$L(y, \mu, \lambda) := y^T P_0 y - \mu(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y$$

and therefore the Lagrangian relaxation of $\mathrm{Q^2P}_y$ is

$$(\mathrm{DQ^2P}_y) \quad d^* := \min_{\substack{\mu \\ \lambda \ge 0}} \max_y y^T P_0 y - \mu(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y.$$

Note that

$$\begin{aligned} d^* &= \min_{\lambda \ge 0} \min_{\mu} \max_y y^T P_0 y - \mu(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y \\ &= \min_{\lambda \ge 0} \max_{y_0^2 = 1} y^T P_0 y + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y, \end{aligned}$$

by the strong duality of the trust-region subproblem [111]. Therefore

$$(\mathrm{DQ^2P}_x) \quad d^* = \min_{\lambda \ge 0} \max_x q_0(x) + \sum_{k \in \mathcal{I}} \lambda_k q_k(x).$$

and we have shown the equivalence of the dual values for the problems in x and in y. (This is similar to the approaches in [120, 110].)

By weak duality, we have

$$d^* \geq q^* = \max_y \min_{\substack{\mu \\ \lambda \geq 0}} y^T P_0 y - \mu(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y.$$

If we can find the optimal values μ^* and λ^* for the dual variables, then we obtain a single quadratic function whose maximal value is an upper bound on q^*:

$$q^* \leq d^* = \max_y y^T P_0 y - \mu^*(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k^* y^T P_k y. \tag{3.59}$$

Furthermore the Lagrangian $L(y, \mu, \lambda)$ is a quadratic function of y and therefore we can add the following hidden semidefinite constraint to the outer minimization of $\mathrm{DQ^2P}_y$:

$$P_0 - \mu E_{00} + \sum_{k \in \mathcal{I}} \lambda_k P_k \preceq 0, \quad \lambda \geq 0, \tag{3.60}$$

where E_{00} is the zero matrix with 1 in the top left corner. The maximum of the maximization subproblem is attained for $y = 0$ and thus the dual problem $\mathrm{DQ^2P}_y$ is equivalent to the SDP

$$\begin{array}{rll} d^* = & \min & \mu \\ & \text{s.t.} & \mu E_{00} - \sum_{k \in \mathcal{I}} \lambda_k P_k \succeq P_0 \\ & & \lambda \geq 0. \end{array}$$

One important observation is that a greater number of quadratic constraints $q_k(y)$ means that we obtain a stronger dual. This is equivalent to our earlier claim that adding redundant quadratic constraints strengthens the SDP relaxation. An excellent illustration of the effectiveness of this strategy is presented in Section 4.4 where this approach achieves strong duality.

Another approach is presented in detail in Kojima and Tunçel [68, 67]. For problems that also have linear equality constraints the notion of copositivity can be used to strengthen the SDP relaxation [103]. However the result is not a tractable relaxation in general.

3.1 Solving SDPs arising from $\mathrm{Q^2P}$s

There are many existing packages for solving SDPs in the public domain (see e.g. Christoph Helmberg's SDP web page

http://www.zib.de/helmberg/semidef.html
or The Handbook of Semidefinite Programming [119].) However we already alluded in Section 2.5 to the limitations of many algorithms when it comes to solving large SDPs. We outlined in that Section several research directions that seek to exploit the sparsity and structure of SDP1 and/or DSDP1. For more general SDPs Kojima et al [66, 33, 34] have made promising advances and Borchers [20, 19] exploits the BLAS routines. Nonetheless, the question of efficiently exploiting sparsity is still very much an open question.

We briefly outline one approach that may help exploit structure and sparsity for SDPs in discrete optimization. Recall from Section 2 that the SDP relaxation SDP1 gives the same bound for MC as the eigenvalue bound (2.22). In fact this equivalence of the bounds holds for any SDP for which trace X is constant over all feasible matrices X [48, 47]. (Note that many SDPs that arise in applications satisfy this property.) In particular the constant trace condition holds for all SDPs that arise from problems which have a bounded feasible set. We can see this by homogenizing Q^2P as in (3.58) and then adding the redundant constraint $\|y\|^2 \leq K$ with K sufficiently large. Now the identity I is in the range of the linear operator $\mathcal{A}^*$ and this is precisely equivalent to the constant trace condition. Therefore Q^2Ps with bounded feasible set can all be phrased as min-max eigenvalue problems for which the inherent structure and sparsity can be exploited. Clearly all 0,1 or ± 1 problems satisfy this boundedness condition and in particular the Graph Partitioning and the Quadratic Assignment Problem that we study in the next Section fall into this class.

4 Specific Instances of SDP Relaxations

We now study in some detail four specific problems and show how to apply the recipe for SDP relaxations. In each case we derive a min-max eigenvalue problem from the Lagrangian dual of an appropriately chosen quadratically constrained problem. The dual of this min-max eigenvalue problem then provides an SDP relaxation for the original problem. Adding redundant quadratic constraints at the start helps in reducing the duality gap. Once we obtain the SDP relaxation, any remaining redundancy in the constraints is eliminated if we ensure that the linear constraints have full row rank and that Slater's condition holds. This illustrates again the strength of this Lagrangian approach.

4.1 The Graph Partitioning Problem

Let $G = (V, E)$ be an undirected graph as in the description of the MC problem. The graph partitioning problem is the problem of partitioning the node set V into k disjoint subsets of specified sizes to minimize the total weight of the edges connecting nodes in distinct subsets of the partition. Let $A = (a_{ij})$ be the weighted adjacency matrix of G, i.e.

$$a_{ij} = \begin{cases} w_{ij} & ij \in E \\ 0 & \text{otherwise.} \end{cases}$$

The graph partitioning problem can be described by the following 0,1 quadratic problem [107]:

$$\text{(GP)} \qquad \begin{array}{rc} w(E_{uncut}) = \max & \frac{1}{2}\text{trace}\, X^t A X \\ \text{s.t.} & X e_k = e_n \\ & X^T e_n = m \\ & X_{ij} \in \{0,1\}, \ \forall ij, \end{array}$$

where e_k is the vector of ones of appropriate size and m is the vector of ordered set sizes

$$m_1 \geq \ldots \geq m_k \geq 1 \quad \text{and} \quad k < n.$$

The columns of the 0,1 $n \times k$ matrices X are the indicator vectors for the sets. If we replace the 0,1 constraints by quadratic constraints and the linear constraints taking their norm squared, we obtain the equivalent problem:

$$\begin{array}{rc} w(E_{uncut}) = \max & \frac{1}{2}\text{trace}\, X^t A X \\ \text{s.t.} & ||X e_k - e_n||^2 + ||X^T e_n - m||^2 = 0 \\ & X_{ij}^2 - X_{ij} = 0, \ \forall ij. \end{array}$$

The Lagrangian relaxation yields the following bound.

$$\begin{array}{rl} B_{GP} := & \min_{\alpha, W} \max_{X} \text{trace}\, [\\ & \frac{1}{2} X^T A X + \alpha(e_k e_k^T X^T X + X^T e_n e_n^T X) + W^T (X \circ X) \\ & -2\alpha(e_k e_n^T X + m e_n^T X) - W^T X \,] \\ & +\alpha(n + \sum_i m_i^2). \end{array} \tag{4.61}$$

We can now homogenize the problem by adding a variable x.

$$\begin{array}{rl} B_{GP} := & \min_{\alpha, W} \max_{X, \, x^2=1} \text{trace}\, [\\ & \frac{1}{2} X^T A X + \alpha(e_k e_k^T X^T X + X^T e_n e_n^T X) + W^T (X \circ X) \\ & +x(-2\alpha(e_k e_n^T X + m e_n^T X) - W^T X) \,] \\ & +\alpha(n + \sum_i m_i^2). \end{array}$$

We now lift the variable x into the Lagrangian to get a min-max eigenvalue problem.

$$
\begin{array}{rcl}
B_{GP} & := & \min\limits_{\alpha,W,\delta} \max\limits_{X,x} \operatorname{trace} [\\
& & \frac{1}{2}X^T AX + \alpha(e_k e_k^T X^T X + X^T e_n e_n^T X) + W^T(X \circ X) + \delta x^2 \\
& & +x(-2\alpha(e_k e_n^T X + m e_n^T X) - W^T X)] \\
& & +\alpha(n + \sum_i m_i^2) - \delta.
\end{array}
$$

The above has a hidden semidefinite constraint.

$$
\begin{array}{ll}
\min & \alpha(n + \sum_i m_i^2) - \delta \\
\text{s.t.} & L_A + \operatorname{Arrow}(\delta, \operatorname{vec}(W)) + \alpha L_\alpha \preceq 0,
\end{array}
\tag{4.62}
$$

where we define the matrices

$$
L_A := \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} I \otimes A \end{bmatrix},
\tag{4.63}
$$

$$
v = \operatorname{vec} e_n m^T,
$$

$$
L_\alpha := \begin{bmatrix} 0 & -(e+v)^T \\ -(e+v) & (e_k e_k^T I \otimes I + I \otimes e_n e_n^T) \end{bmatrix},
\tag{4.64}
$$

and the linear operator

$$
\operatorname{Arrow}(\delta, \operatorname{vec}(W)) := \begin{bmatrix} \delta & -\frac{1}{2}(\operatorname{vec}(W))^T \\ -\frac{1}{2}(\operatorname{vec}(W)) & \operatorname{Diag}(\operatorname{vec}(W)) \end{bmatrix}.
\tag{4.65}
$$

The dual problem yields the semidefinite relaxation of (GP).

$$
\begin{array}{ll}
\max & \operatorname{trace} L_A Y \\
\text{s.t.} & \operatorname{diag}(Y) = (1, Y_{0,1:n})^T \\
& \operatorname{trace} Y L_\alpha = 0 \\
& Y \succeq 0.
\end{array}
\tag{4.66}
$$

4.2 The Quadratic Assignment Problem

While MC can be considered the simplest of the NP-hard problems, the quadratic assignment problem (QAP) can be considered the hardest. This is an area where $n = 30$ is a large-scale problem. We shall use the trace formulation of the QAP where the variable X is a permutation matrix, i.e.

X is a 0,1 matrix and all row and column sums are equal to one. The formulation is:

$$
(QAP_E) \qquad \begin{array}{rll} \mu^* := & \max \quad q(X) & = \operatorname{trace}(AXB - 2C)X^T \\ & \text{s.t.} & Xe = e \\ & & X^T e = e \\ & & X_{ij} \in \{0,1\} \quad \forall i,j. \end{array} \tag{4.67}
$$

(See [98] for applications and other formulations of the QAP.)

Let us apply the recipe. We first add redundant quadratic constraints to the model. Since the set of permutation matrices is equal to the intersection of the set of orthogonal matrices with the 0,1 matrices [123], we can add both of the following (equivalent) definitions of orthogonality: $XX^T = I$ and $X^T X = I$. The recipe also requires the application of Theorem 2.5 to the linear constraints before taking the Lagrangian dual. We thus obtain the following formulation for the QAP:

$$
(QAP_E) \qquad \begin{array}{rll} \mu^* := & \min & \operatorname{trace} AXBX^T - 2CX^T \\ & \text{s.t.} & XX^T = I \\ & & X^T X = I \\ & & \|Xe - e\|^2 = 0 \\ & & \|X^T e - e\|^2 = 0 \\ & & X_{ij}^2 - X_{ij} = 0, \quad \forall i,j. \end{array}
$$

Other relaxations and bounds can be obtained by adding redundant constraints such as $\operatorname{trace} XX^T = n$ or $0 \leq X_{ij} \leq 1$, $\forall i,j$.

It turns out that the squared linear constraints are eliminated by the projection later so we can add them together without any loss: $\|Xe - e\|^2 + \|X^T e - e\|^2 = 0$. We first add the 0,1 and row-column sum constraints to the objective function using Lagrange multipliers W_{ij} and u_0 respectively.

$$
\begin{array}{rl} \mu_{\mathcal{O}} = \min\limits_{XX^T = X^T X = I} \max\limits_{W, u_0} & \{\operatorname{trace} AXBX^T - 2CX^T + \sum_{ij} W_{ij}(X_{ij}^2 - X_{ij}) \\ & + u_0(\|Xe - e\|^2 + \|X^T e - e\|^2)\}. \end{array} \tag{4.68}
$$

Interchanging min and max yields

$$
\begin{array}{rl} \mu_{\mathcal{O}} \geq \mu_{\mathcal{L}} := \max\limits_{W, u_0} \min\limits_{XX^T = X^T X = I} & \{\operatorname{trace} AXBX^T - 2CX^T + \sum_{ij} W_{ij}(X_{ij}^2 - X_{ij}) \\ & + u_0(\|Xe - e\|^2 + \|X^T e - e\|^2)\}. \end{array} \tag{4.69}
$$

We now homogenize the objective function using the constrained scalar variable x_0 and increasing the dimension of the problem by 1. This simplifies the transition to an SDP:

$$\begin{aligned}\mu_{\mathcal{O}} \geq \mu_{\mathcal{L}} = \max_{W,u_0} \min_{XX^T=X^TX=I, x_0^2=1} & \{\operatorname{trace}[AXBX^T + W(X\circ X)^T \\ & + u_0(\|Xe\|^2+\|X^Te\|^2) - x_0(2C+W)X^T] \\ & - 2x_0u_0e^T(X+X^T)e + 2nu_0x_0^2\}.\end{aligned} \tag{4.70}$$

Introducing a Lagrange multiplier w_0 for the constraint on x_0 and Lagrange multipliers S_b for $XX^T = I$ and S_o for $X^TX = I$ we get the lower bound μ_R:

$$\begin{aligned}\mu_{\mathcal{O}} \geq \mu_{\mathcal{L}} \geq \mu_R := \max_{W,S_b,S_o,u_0,w_0} \min_{X,\ x_0} & \{\operatorname{trace}[AXBX^T + u_0(\|Xe\|^2+\|X^Te\|^2) \\ & + W(X\circ X)^T + w_0x_0^2 + S_bXX^T + S_oX^TX] \\ & - \operatorname{trace} x_0(2C+W)X^T - 2x_0u_0e^T(X+X^T)e \\ & - w_0 - \operatorname{trace} S_b - \operatorname{trace} S_o + 2nu_0x_0^2\}.\end{aligned} \tag{4.71}$$

Note that we grouped the quadratic, linear, and constant terms together in (4.71). Now we define $x := \operatorname{vec}(X)$, $y^T := (x_0, x^T)$ and $w^T := (w_0, \operatorname{vec}(W)^T)$ to obtain:

$$\begin{aligned}\mu_R = \max_{w,S_b,S_o,u_0} \min_{y} & \{y^T[L_Q + \operatorname{Arrow}(w) + \mathrm{B}^0\mathrm{Diag}(S_b) + \mathrm{O}^0\mathrm{Diag}(S_o) + u_0D]\,y \\ & - w_0 - \operatorname{trace} S_b - \operatorname{trace} S_o\},\end{aligned} \tag{4.72}$$

where L_Q is as above and we used the linear operators

$$\operatorname{Arrow}(w) := \begin{bmatrix} w_0 & -\frac{1}{2}w_{1:n^2}^T \\ -\frac{1}{2}w_{1:n^2} & \operatorname{Diag}(w_{1:n^2}) \end{bmatrix}, \tag{4.73}$$

$$\mathrm{B}^0\mathrm{Diag}(S) := \begin{bmatrix} 0 & 0 \\ 0 & I\otimes S_b \end{bmatrix}, \tag{4.74}$$

$$\mathrm{O}^0\mathrm{Diag}(S) := \begin{bmatrix} 0 & 0 \\ 0 & S_o\otimes I \end{bmatrix}, \tag{4.75}$$

and

$$D := \begin{bmatrix} n & -e^T\otimes e^T \\ -e\otimes e & I\otimes E \end{bmatrix} + \begin{bmatrix} n & -e^T\otimes e^T \\ -e\otimes e & E\otimes I \end{bmatrix}.$$

There is a hidden semidefinite constraint in (4.72): the inner minimization problem is bounded below only if the Hessian of the quadratic form is positive semidefinite. And in that case the quadratic form has minimum value 0. Hence we have the equivalent SDP:

$$(D_{\mathcal{O}}) \quad \begin{array}{ll} \max & -w_0 - \operatorname{trace} S_b - \operatorname{trace} S_o \\ \text{s.t.} & L_Q + \operatorname{Arrow}(w) + \mathrm{B}^0\operatorname{Diag}(S_b) + \mathrm{O}^0\operatorname{Diag}(S_o) + u_0 D \succeq 0. \end{array}$$

We now obtain our desired SDP relaxation of $(QAP_{\mathcal{O}})$ as the Lagrangian dual of $(D_{\mathcal{O}})$. We introduce the $(n^2+1) \times (n^2+1)$ dual matrix variable $Y \succeq 0$ and derive the dual problem to the SDP $(D_{\mathcal{O}})$.

$$(SDP_{\mathcal{O}}) \quad \begin{array}{ll} \min & \operatorname{trace} L_Q Y \\ \text{s.t.} & \mathrm{b}^0\operatorname{diag}(Y) = I, \quad \mathrm{o}^0\operatorname{diag}(Y) = I \\ & \operatorname{arrow}(Y) = e_0, \quad \operatorname{trace} DY = 0 \\ & Y \succeq 0, \end{array} \tag{4.76}$$

where the *arrow operator*, acting on the $(n^2+1) \times (n^2+1)$ matrix Y, is the adjoint operator to $\operatorname{Arrow}(\cdot)$ and is defined by

$$\operatorname{arrow}(Y) := \operatorname{diag}(Y) - \left(0, (Y_{0,1:n^2})^T\right), \tag{4.77}$$

i.e. the arrow constraint guarantees that the diagonal and the first (0^{th}) row (or column) are identical.

The *block-0-diagonal operator* and *off-0-diagonal operator* acting on Y are defined by

$$\mathrm{b}^0\operatorname{diag}(Y) := \sum_{k=1}^{n} Y_{(k,\cdot),(k,\cdot)} \tag{4.78}$$

and

$$\mathrm{o}^0\operatorname{diag}(Y) := \sum_{k=1}^{n} Y_{(\cdot,k),(\cdot,k)}. \tag{4.79}$$

These are the adjoint operators of $\mathrm{B}^0\operatorname{Diag}(\cdot)$ and $\mathrm{O}^0\operatorname{Diag}(\cdot)$, respectively. The block-0-diagonal operator guarantees that the sum of the diagonal blocks equals the identity. The off-0-diagonal operator guarantees that the trace of each diagonal block is 1, while the trace of the off-diagonal blocks is 0. These constraints come from the orthogonality constraints, $XX^T = I$ and $X^TX = I$, respectively.

We have expressed the orthogonality constraints with both $XX^T = I$ and $X^TX = I$. It is interesting to note that this redundancy adds extra constraints into the relaxation which are **not** redundant. These constraints reduce the size of the feasible set and so tighten the bounds.

Proposition 4.1 *Suppose that Y is feasible for the SDP relaxation (4.76). Then Y is singular.*

Proof. Note that $D \neq 0$ is positive semidefinite. Therefore Y has to be singular to satisfy the constraint $\operatorname{trace} DY = 0$. ■

This means that the feasible set of the primal problem $(SDP_{\mathcal{O}})$ has no interior. It is not difficult to find an interior-point for the dual $(D_{\mathcal{O}})$, which means that Slater's constraint qualification (strict feasibility) holds for $(D_{\mathcal{O}})$. Therefore $(SDP_{\mathcal{O}})$ is attained and there is no duality gap in theory, for the usual primal-dual pair. However if Slater's constraint qualification fails, then this is not the proper dual, since perturbations in the right-hand-side will not result in the dual value. This is because we cannot stay exactly feasible, since the interior is empty (see [105]). In fact we may never attain the supremum of $(D_{\mathcal{O}})$, which may cause instability when implementing any kind of interior-point method. Since Slater's constraint qualification fails for the primal, the set of optimal solutions of the dual is an unbounded set and an interior-point method may never converge. Therefore we have to express the feasible set of $(SDP_{\mathcal{O}})$ in some lower dimensional space. We study this below when we project the problem onto a face of the semidefinite cone.

However, if we add the rank-one condition, then the relaxation is exact.

Theorem 4.2 *Suppose that Y is restricted to be rank-one in $(SDP_{\mathcal{O}})$, i.e. $Y = \begin{pmatrix} 1 \\ x \end{pmatrix} (1 \; x^T)$, for some $x \in \Re^{n^2}$. Then the optimal solution of $(SDP_{\mathcal{O}})$ provides the permutation matrix $X = \operatorname{Mat}(x)$ that solves the QAP.*

Proof. The arrow-constraint in $(SDP_{\mathcal{O}})$ guarantees that the diagonal of Y is 0 or 1. The 0-diagonal and assignment constraint now guarantee that $\operatorname{Mat}(x)$ is a permutation matrix. Therefore the optimization is over the permutation matrices and so the optimum of QAP is obtained. ■

We now devote our attention to homogenization since that results in a min-max eigenvalue problem and an equivalent SDP. We have seen that we can homogenize by increasing the dimension of the problem by 1. We first add the 0,1 constraints to the objective function using Lagrange multipliers W_{ij}.

$$\min_{W} \max_{XX^T=I} \operatorname{trace}(AXB - 2C)X^T + \sum_{ij} W_{ij}(X_{ij}^2 - X_{ij}). \tag{4.80}$$

We now homogenize the objective function by multiplying by a constrained scalar x.

$$\min_{W} \max_{XX^T=I, x^2=1} \operatorname{trace}\left[AXBX^T + W(X \circ X)^T - x(2C+W)X^T\right]. \tag{4.81}$$

We can now use Lagrange multipliers to get a parameterized min-max eigenvalue problem in dimension n^2+1. We get the following bound. The parameters are: the symmetric $n \times n$ matrix $\Lambda = \Lambda^T$, the general $n \times n$ matrix W and the scalar α.

$$\begin{array}{rcl} B_{QAP} & := & \min\limits_{\Lambda, W, \alpha} \max\limits_{X} \operatorname{trace} [\\ & & AXBX^T + \Lambda XX^T + W^T(X \circ X) + \alpha x^2 \\ & & -x(2C+W)X^T \,] - \alpha - \operatorname{trace}\Lambda. \end{array} \tag{4.82}$$

We have grouped the quadratic, original linear, and constant terms together. The hidden semidefinite constraint now yields an SDP:

$$\begin{array}{ll} \min & -\operatorname{trace}\Lambda - \alpha \\ \text{s.t.} & L_Q + \operatorname{Arrow}(\alpha, \operatorname{vec}(W)) + \mathrm{B}^0\mathrm{Diag}(\Lambda) \preceq 0, \end{array} \tag{4.83}$$

where we define the matrix

$$L_Q := \begin{bmatrix} 0 & -\operatorname{vec}(C)^T \\ -\operatorname{vec}(C) & B \otimes A \end{bmatrix}, \tag{4.84}$$

and the linear operators

$$\operatorname{Arrow}(\alpha, \operatorname{vec}(W)) := \begin{bmatrix} \alpha & -\frac{1}{2}\operatorname{vec}(W)^T \\ -\frac{1}{2}\operatorname{vec}(W) & \operatorname{Diag}(\operatorname{vec}(W)) \end{bmatrix}, \tag{4.85}$$

$$\mathrm{B}^0\mathrm{Diag}(\Lambda) := \begin{bmatrix} 0 & 0 \\ 0 & I \otimes \Lambda \end{bmatrix}. \tag{4.86}$$

We can now introduce the $(n^2+1)\times(n^2+1)$ dual variable matrix $Y \succeq 0$ and derive the dual problem to this min-max eigenvalue problem, i.e.

$$\max_{Y\succeq 0}\min_{\Lambda,W,\alpha} -\mathrm{trace}\,\Lambda - \alpha + \mathrm{trace}\,Y(L_Q + \mathrm{Arrow}\,(\alpha, \mathrm{vec}\,(W)) + \mathrm{B^0Diag}\,(\Lambda)).$$

The inner minimization problem is unconstrained and linear in the variables. Therefore, after reorganizing the variables, we can differentiate to get the dual problem to this dual problem, or the semidefinite relaxation to the original QAP. (Recall that $Y_{i,j:k}$ refers to the i-th row and columns j to k of the matrix Y; and $\mathrm{b^0diag}\,(Y)$ is the block diagonal sum of Y which ignores the first row.) The derivatives with respect to α and W yields the first constraint and the derivative with respect to Λ yields the second constraint in the following problem. Equivalently, the constraints are the adjoints of the linear operators Arrow and $\mathrm{B^0Diag}$.

$$\begin{array}{cc} \max & \mathrm{trace}\,L_Q Y \\ \text{s.t.} & \mathrm{diag}\,(Y) = (1, Y_{0,1:n^2})^T \\ & \mathrm{b^0diag}\,(Y) = I \\ & Y \succeq 0. \end{array} \tag{4.87}$$

Another primal-dual pair can be obtained using a trust-region subproblem as the inner maximization problem, rather than homogenizing to an eigenvalue problem. This is done by adding the redundant trust-region constraint $\mathrm{trace}\,XX^T = n$. As mentioned above, we can also add the redundant constraint

$$||Xe - e||^2 + ||X^Te - e||^2 = 0.$$

A primal-dual interior-point method based on these types of dual pairs of problems, such as (4.87) and (4.83), is tested and studied in [123].

4.3 The Max-Clique and Max-Stable-Set Problems

Consider again the undirected graph $G = (E, V)$ defined above. The max-clique problem consists in finding the largest connected subgraph. We let $\omega(G)$ denote the size of the largest clique in G. A stable set is a subset of nodes of V such that no two nodes are adjacent. We denote the size of the largest stable set in $\bar{G}$, the complement of G, by $\alpha(\bar{G})$. Clearly

$$\alpha(\bar{G}) = \omega(G).$$

Bounds for these problems and relationships to the theta function, or Lovász number of the graph, are described in the expository paper of Knuth [65] (see also [109]).

In this section we show that the Lovász bound on $\omega(G)$ can be alternatively obtained from two distinct 0,1 problems (4.88) and (4.91) by Lagrangian relaxations. Let A be the incidence matrix of the graph, i.e. $A = (a_{ij})$ with $a_{ij} = 1$ if $ij \in E$ and 0 otherwise. If x is the indicator vector for the largest clique in G of size k, A then $x^T(I+A)x/x^Tx = k^2/k = k$. A quadratic formulation of the max-clique problem is the following 0,1 quadratic problem.

$$\begin{array}{rll} \omega(G) = & \max & \frac{x^T(I+A)x}{x^Tx} \\ & \text{s.t.} & x_ix_j = 0, \text{ if } ij \notin E, \ i \neq j \\ & & x_i \in \{0,1\}, \ \forall i. \end{array} \tag{4.88}$$

Therefore, a quadratic relaxation of the max-clique problem is the following quadratic constrained problem.

$$\begin{array}{rll} \omega(G) \leq \omega_1^* := & \max & x^T(I+A)x \\ & \text{s.t.} & x_ix_j = 0, \text{ if } ij \notin E, \ i \neq j \\ & & x^Tx = 1. \end{array} \tag{4.89}$$

The Lagrangian relaxation for this problem is the perturbed min-max eigenvalue problem and the equivalent SDP:

$$\begin{array}{rll} \omega_1^* & \leq & \min\limits_{W_{ij}=0, \text{ if } ij \in E, \text{ or } i=j} \quad \lambda_{\max}(I+A+W) - \alpha x^Tx + \alpha \\ & = & \min\limits_{w,\alpha} \max\limits_{x} x^T(I+A)x + \sum\limits_{ij \notin E, \ i \neq j} w_{ij}x_ix_j - \alpha x^Tx + \alpha \\ & = & \min\limits_{\substack{I+A+W \preceq \alpha I \\ W_{ij}=0, \text{ if } ij \in E, \text{ or } i=j}} \quad \alpha \end{array}$$

i.e. minimize the max eigenvalue over perturbations in the off-diagonal elements corresponding to disjoint nodes. This bound is equal to the Lovász theta function on the complementary graph:

$$\vartheta(\bar{G}) = \min_{A \in \mathcal{A}} \lambda_{\max}(A), \tag{4.90}$$

where

$$\mathcal{A} = \{A : A \text{ symmetric } \ n \times n \text{ matrix with } \ A_{ij} = 1, \text{ if } ij \in E, \text{ or } i = j\}.$$

By considering the (optimal) indicator vector for the largest clique, we see that the following 0,1 quadratic problem describes exactly the max-clique problem. Note that if node i is not in the largest clique, then necessarily, $x_i x_j = 0$ for some j with node j in the clique, i.e. necessarily $x_i = 0$ in the indicator vector.

$$\begin{array}{rll} \omega(G) = & \max & x^T x \\ & \text{s,t.} & x_i x_j = 0, \text{ if } ij \notin E, \; i \neq j \\ & & x_i^2 - x_i = 0, \; \forall i. \end{array} \tag{4.91}$$

The Lagrangian relaxation yields the bound

$$B_{\text{clique}} := \min_{W,\lambda} \max_x x^T x + \sum_{ij \notin E, \; i \neq j} w_{ij} x_i x_j + \sum_i \lambda_i (x_i^2 - x_i).$$

We let W be an $n \times n$ matrix with zeros in positions where $ij \in E$. We can homogenize by adding the constraint $y^2 = 1$ and then lifting it into the Lagrangian.

$$\min_{\alpha,W,\lambda} \max_{x,y} x^T x + \sum_{ij \notin E} w_{ij} x_i x_j + \sum_i \lambda_i x_i^2 + \alpha y^2 - y \sum_i \lambda_i x_i - \alpha.$$

We now exploit the hidden semidefinite constraint to obtain the SDP:

$$\begin{array}{rlc} B_{\text{clique}} = & \min\limits_{W,\lambda,\alpha} & -\alpha \\ & \text{s.t.} & L_A + L_W(W) + \text{Arrow}\,(\alpha, \lambda) \preceq 0 \\ & & W_{ij} = 0, \; \forall ij \in E, \text{ or } i = j, \end{array} \tag{4.92}$$

where the matrix

$$L_A := \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}, \tag{4.93}$$

and the linear operators

$$L_W(W) := \begin{bmatrix} 0 & 0 \\ 0 & W \end{bmatrix}, \tag{4.94}$$

$$\text{Arrow}\,(\alpha, \lambda) := \begin{bmatrix} \alpha & -\frac{1}{2}\lambda^T \\ -\frac{1}{2}\lambda & \text{Diag}\,(\lambda) \end{bmatrix}. \tag{4.95}$$

The dual of the above min-max eigenvalue problem yields the semidefinite relaxation for the max-clique problem with $Y \in \mathcal{S}_{n+1}$.

$$\begin{array}{ll} \max & \operatorname{trace} L_A Y \\ \text{s.t.} & \operatorname{diag}(Y) = (1, Y_{0,1:n})^T \\ & Y_{ij} = 0, \ \forall ij \notin E \\ & Y \succeq 0. \end{array} \tag{4.96}$$

The equivalence of the bounds (4.90) and (4.96) was shown in lemma 2.17 of [82].

Consider the problem (4.88) with an additional redundant constraint

$$x_i x_j \geq 0 \text{ for } ij \in E \tag{4.97}$$

That is

$$\begin{array}{lll} \omega(G) = & \max & \frac{x^T(I+A)x}{x^Tx} \\ & \text{s.t.} & x_i x_j = 0, \ \text{if } ij \notin E, \ i \neq j \\ & & x_i x_j \geq 0, \ \text{if } ij \in E, \\ & & x_i \in \{0,1\}, \ \forall i. \end{array} \tag{4.98}$$

A quadratic relaxation of the max-clique problem is the following quadratically constrained problem:

$$\begin{array}{lll} \omega(G) \leq \omega_1^* := & \max & x^T(I+A)x \\ & \text{s.t.} & x_i x_j = 0, \ \text{if } ij \notin E, \ i \neq j \\ & & x_i x_j \geq 0, \ \text{if } ij \in E, \\ & & x^T x = 1. \end{array} \tag{4.99}$$

The Lagrangian relaxation for this problem is equal to Schrijver's improvement [109] of the theta function on the complementary graph:

$$\vartheta'(\bar{G}) = \min_{A \in \mathcal{A}'} \lambda_{\max}(A),$$

where

$$\mathcal{A}' = \{A : A \text{ symmetric } \ n \times n \text{ matrix with } \ A_{ij} \geq 1, \text{ if } ij \in E, \text{ or } i = j\}.$$

Haemers [44] constructed graphs where $\vartheta'(\bar{G})$ is strictly smaller than $\vartheta(\bar{G})$.

Analogously, it is possible to modify the problem (4.91) by adding the constraint (4.97).

4.4 Orthogonally Constrained Problems: Achieving Zero Duality Gaps

As a final illustration of the strength of Lagrangian relaxation and the power of adding appropriate redundant quadratic constraints we consider the orthonormal type constraints:

$$X^T X = I, \qquad X \in \mathcal{M}_{m,n}.$$

(This set is sometimes known as the Stiefel manifold. Applications and algorithms for optimization over orthonormal sets of matrices are discussed in [27].) We also consider the trust-region type constraint

$$X^T X \preceq I, \qquad X \in \mathcal{M}_{m,n}.$$

We follow the approach in [11, 10, 9] and show that if $m = n$ then strong duality holds for certain (non-convex) quadratic problems defined over orthonormal matrices after adding some quadratic redundant constraints. Because of the similarity of the orthonormality constraint to the (vector) norm constraint $x^T x = 1$, the results of this section can be viewed as a matrix generalization of the strong duality result for the well-known Rayleigh Quotient problem [100].

Let A and B be $n \times n$ symmetric matrices, and consider the orthonormal constrained homogeneous problem:

$$(\mathrm{QQP_O}) \qquad \mu^O := \begin{array}{ll} \min & \operatorname{trace} AXBX^T \\ \text{s.t.} & XX^T = I. \end{array} \tag{4.99}$$

This problem can be solved exactly using Lagrange multipliers [43] or the classical Hoffman-Wielandt inequality [18].

Proposition 4.3 *Suppose that the orthogonal diagonalizations of A, B are $A = V\Sigma V^T$ and $B = U\Lambda U^T$, respectively, where the eigenvalues in Σ are ordered non-increasing, and the eigenvalues in Λ are ordered nondecreasing. Then the optimal value of* $\mathrm{QQP_O}$ *is $\mu^O = \operatorname{trace} \Sigma\Lambda$, and the optimal solution is obtained using the orthogonal matrices that yield the diagonalizations, i.e. $X^* = VU^T$.* ■

The Lagrangian dual of $\mathrm{QQP_O}$ is

$$\max_{S=S^T} \min_X \operatorname{trace} AXBX^T - \operatorname{trace} S(XX^T - I). \tag{4.100}$$

However, there can be a nonzero duality gap for the Lagrangian dual, see [123] for an example. The inner minimization in the dual problem (4.100) is an unconstrained quadratic minimization in the variables $\operatorname{vec}(X)$, with hidden constraint on the Hessian

$$B \otimes A - I \otimes S \succeq 0.$$

The first order stationarity conditions are equivalent to $AXB = SX$ or $AXBX^T = S$. One can easily construct examples where the semidefinite condition and the stationarity conditions are in conflict and thus a duality gap occurs. In order to close the duality gap, we need a larger class of quadratic functions.

Note that in $\mathrm{QQP_O}$ the constraints $XX^T = I$ and $X^TX = I$ are equivalent. Adding the redundant constraints $X^TX = I$, we arrive at

$$\mathrm{QQP_{OO}} \qquad \begin{array}{rll} \mu^O := & \min & \operatorname{trace} AXBX^T \\ & \text{s.t.} & XX^T = I,\ X^TX = I. \end{array}$$

Using symmetric matrices S and T to relax the constraints $XX^T = I$ and $X^TX = I$, respectively, we obtain a dual problem

$$\mathrm{DQQP_{OO}} \qquad \begin{array}{rll} \mu^O \geq \mu^D := & \max & \operatorname{trace} S + \operatorname{trace} T \\ & \text{s.t.} & (I \otimes S) + (T \otimes I) \preceq (B \otimes A) \\ & & S = S^T,\ T = T^T. \end{array}$$

Theorem 4.4 *Strong duality holds for* $\mathrm{QQP_{OO}}$ *and* $\mathrm{DQQP_{OO}}$, *i.e.,* $\mu^D = \mu^O$ *and both primal and dual are attained.* ■

A further relaxation of the above orthogonal relaxation is the trust-region relaxation studied in [64]:

$$\begin{array}{rll} \mu^*_{QAPT} := & \min & \operatorname{trace} AXBX^T \\ & \text{s.t.} & XX^T \preceq I. \end{array}$$

The constraints are convex with respect to the Löwner partial order and so it is hoped that solving this problem would be useful. The set

$$\{X : W = XX^T \preceq I\}$$

is studied in [97, 29] and is useful in eigenvalue variational principles. Furthermore the problem (4.100) is visually similar to the trust-region subproblem so we would like to find a characterization of optimality.

We study the matrix trust-region relaxation of QAP:

$$\begin{array}{rll} \mu^*_{SDPT} = & \min & \operatorname{trace} AXBX^T \\ & \text{s.t.} & XX^T \preceq I. \end{array}$$

The following generalization of the Hoffman-Wielandt inequality holds.

Theorem 4.5 *For any $XX^T \preceq I$, we have*

$$\sum_{i=1}^n \min\{\lambda_i \mu_{n-i+1}, 0\} \leq trAXBX^T \leq \sum_{i=1}^n \max\{\lambda_i \mu_i, 0\}$$

and the upper bound is attained if

$$X = P \operatorname{Diag}(\epsilon_1, \epsilon_2, \cdots, \epsilon_n) Q^T, \tag{4.100}$$

where

$$\varepsilon_i = \left\{ \begin{array}{ll} 1, & \lambda_i \mu_i > 0, \\ \alpha \in [0,1], & \lambda_i \mu_i = 0, \\ 0, & \lambda_i \mu_i < 0; \end{array} \right. \tag{4.100}$$

The lower bound is attained if

$$X = P \operatorname{Diag}(\epsilon_1, \epsilon_2, \cdots, \epsilon_n) Q^T, \tag{4.100}$$

where

$$\varepsilon_i = \left\{ \begin{array}{ll} 1, & \lambda_i \mu_{n-i+1} < 0, \\ \alpha \in [0,1], & \lambda_i \mu_{n-i+1} = 0, \\ 0, & \lambda_i \mu_{n-i+1} > 0. \end{array} \right. \tag{4.100}$$

■

The lower bound in the above theorem states that $\mu^*_{SDPT} = \sum_{i=1}^n [\lambda_i \mu_i]^-$. Since the Theorem provides the feasible point of attainment, i.e. an upper bound for the relaxation problem, we will prove the theorem by proving another theorem that shows that the value μ^*_{SDPT} is also attained by a Lagrangian dual problem. Note that since XX^T and X^TX have the same eigenvalues, $XX^T \preceq I$ if and only if $X^TX \preceq I$. Explicitly using both sets of constraints, as in [11], we obtain

$$\text{QAPTR} \qquad \mu^*_{QAPT} := \begin{array}{ll} \min & \operatorname{trace} AXBX^T \\ \text{s.t.} & XX^T \preceq I, \quad X^TX \preceq I. \end{array}$$

Next we apply Lagrangian relaxation to QAPTR, using matrices $S \succeq 0$ and $T \succeq$ to relax the constraints $XX^T \preceq I$ and $X^TX \preceq I$, respectively. This results in the dual problem

$$\text{DQAPTR} \qquad \mu^*_{QAPT} \geq \mu^D_{QAPT} := \begin{array}{ll} \max & -\operatorname{trace} S - \operatorname{trace} T \\ \text{s.t.} & (B \otimes A) + (I \otimes S) + (T \otimes I) \succeq 0 \\ & S \succeq 0, \ T \succeq 0. \end{array}$$

To prove that $\mu^*_{QAPT} = \mu^D_{QAPT}$ we will use the following simple result:

Lemma 4.6 *Let $\lambda \in \Re^n$, $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. For $\gamma \in \Re^n$ consider the problem*

$$\min \quad z_\pi := \sum_{i=1}^n [\lambda_i \gamma_{\pi(i)}]^-,$$

where $\pi(\cdot)$ is a permutation of $\{1, \ldots, n\}$. Then the permutation that minimizes z_π satisfies $\gamma_{\pi(1)} \geq \gamma_{\pi(2)} \geq \cdots \gamma_{\pi(n)}$. ■

Theorem 4.7 *Strong duality holds for* QAPTR *and* DQAPTR*:*

$$\mu^D_{QAPT} = \mu^*_{QAPT}$$

and both primal and dual optimal values are attained. ■

These results conclude the first part of the paper which illustrated the strength of the Lagrangian relaxation. We now proceed to our second application of SDP.

5 Matrix Completion Problems

Semidefinite programming problems arise in surprisingly many different areas of mathematics and engineering where they sometimes have different names. In engineering they are often referred to as linear matrix inequalities problems. In matrix theory, the class of problems called matrix completion problems is closely related to SDP. In this last Section we study application of SDP to this class of problems.

A *symmetric partial matrix* is a symmetric matrix where certain entries are fixed or specified while the remaining entries are unspecified or free. The symmetric matrix completion problem endeavors to specify the free elements in such a way that the resulting matrix satisfies certain required properties. For example, the positive semidefinite matrix completion problem (PSDM) consists of finding a completion so that the resulting matrix is symmetric positive semidefinite, while the Euclidean distance matrix completion problem (EDM) seeks a completion that forms a Euclidean distance matrix (a precise definition of this class of matrices is given below).

In this Section we show how successful SDP has been in solving matrix completion problems. We begin in Section 5.1 with theoretical existence results for completions based on chordality. This follows the work in [42]. We then present an efficient approach to solve PSDM completion problems [57]. This approach successfully solves large sparse problems. In Section 5.3 this approach is extended to the EDM completion problem (based on the work in [1]) but is shown to exhibit difficulties in the large sparse case. Hence we conclude by presenting in Section 5.4 a new characterization of Euclidean distance matrices and new algorithms that efficiently solve large sparse problems.

5.1 Existence Results

Both the PSDM and EDM problems have been extensively studied in the literature. Let us first phrase the completion problem using the graph of the matrix. Suppose that $\mathcal{G}(V, E)$ is a finite undirected graph. The edges of the graph correspond to fixed elements in the matrix, i.e. $A(\mathcal{G})$ is a *$\mathcal{G}$-partial matrix*

$$\text{if } a_{ij} \text{ is defined if and only if } \{i, j\} \in E.$$

$A(\mathcal{G})$ is a *$\mathcal{G}$-partial positive matrix* if $a_{ij} = \overline{a_{ji}}, \forall \{i, j\} \in E$ and all existing principal minors are positive. With $\mathcal{J} = (V, \bar{E}), E \subset \bar{E}$ a $\mathcal{J}$-partial matrix.

$B(\mathcal{J})$ *extends* the $\mathcal{G}$-partial matrix $A(\mathcal{G})$ if $b_{ij} = a_{ij}, \forall\{i,j\} \in E$, i.e. the missing (free) elements in the matrix are filled in.

$\mathcal{G}$ is *positive completable* if every $\mathcal{G}$-partial positive matrix can be extended to a positive definite matrix. With this definition we look at the pattern of fixed elements in the matrix rather than specific elements. The following is the key property to guarantee that a completion is possible.

Definition 5.1 *$\mathcal{G}$ is* chordal *if there are no minimal cycles of length ≥ 4. (every cycle of length ≥ 4 has a chord)*

Theorem 5.2 *([42]) $\mathcal{G}$ is positive completable if and only if $\mathcal{G}$ is chordal.* ■

When a positive definite completion is possible, then the one of maximum determinant is unique and can be characterized.

Theorem 5.3 *([42]) Let A be a partial symmetric matrix all of whose diagonal entries are specified, and suppose that A has a positive definite completion. Then, among all positive definite completions, there is a unique one with maximum determinant.*

The 1990 survey paper [56] presents many of the theoretical results for completion problems. Similar existence results are known for the EDM completion problem, see e.g. the comparison of the two problems [77], as well as the survey paper [78] and [12, 15, 36, 63]. Related results appear in [76, 81, 79, 36, 28, 61, 13, 35, 53, 52, 41, 58, 23, 25, 85, 15, 60, 96, 59].

One can use determinantal inequalities (e.g. [36]) or semidefinite programming techniques to find completions. For example, to find a positive semidefinite completion, with fixed elements $a_{ij}, \{i,j\} \in E$, one can solve the following (feasibility) problem.

$$\begin{array}{rl} \max & \operatorname{trace} CX \\ \text{subject to} & \operatorname{trace} E_{ij}P = a_{ij}, \quad \forall\{i,j\} \in E \\ & P \succeq 0 \end{array}$$

where $E_{ij} = e_i e_j^t + e_j e_k^t$, and C is an arbitrary symmetric matrix. To find the solution with maximum determinant, one can use the objective $\max \log \det(X)$ in the above problem, see e.g. [117]. (These problems can be solved very efficiently.)

It is not clear that finding completions this way is efficient for large sparse problems, since current SDP codes cannot yet handle the general large sparse case very well.

5.2 Approximate Positive Semidefinite Completions

However, one can reformulate the completion problem as an *approximate positive semidefinite completion problem.* This trick, which we now outline, allows efficient solution of the large sparse case, see [57].

Suppose we are given the real, nonnegative (element-wise) **symmetric matrix of weights** $H = H^t \geq 0$ with positive diagonal elements $H_{ii} > 0$, $\forall i$. The positive element H_{ij} provide a weighting on the importance of fixing the element a_{ij} in the **partial symmetric matrix** $A = A^t$ (For notational purposes, we assume that the free elements of A are set to 0 if not specified otherwise.)

Recall that $||A||_F = \sqrt{\operatorname{trace} A^*A}$ is the *Frobenius norm*, and $\circ$ denotes *Hadamard product.* Define the objective function

$$f(P) := ||H \circ (A - P)||_F^2.$$

This function weights the fixed elements while ignoring the free elements. The *weighted, best approximate, completion problem* is

$$(CM) \qquad \begin{array}{rcc} \mu^* := & \min & f(P) \\ & \text{subject to} & \mathcal{K}P = b \\ & & P \succeq 0, \end{array}$$

where $\mathcal{K} : \mathcal{S}^n \to \Re^m$ is a linear operator, and $b \in \Re^m$. We include the linear operator $\mathcal{K}$ to allow for additional constraints, e.g. when certain elements need to be exactly fixed.

To solve CM, we present a dual problem and a primal-dual interior-point algorithm that can exploit sparsity. Following is the Lagrangian for CM.

$$L(P, y, \Lambda) = f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda P.$$

The primal problem can be obtained from

$$\mu^* = \min_P \max_{\Lambda \succeq 0} L(P, \alpha, \Lambda, y),$$

while the dual problem comes from

$$\nu^* = \max_{\Lambda \succeq 0} \min_P L(P, \alpha, \Lambda, y),$$

i.e.

$$(DCM) \qquad \begin{array}{cc} \max & f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda P \\ \text{subject to} & \nabla f(P) - K^* y - \Lambda = 0 \\ & \Lambda \succeq 0. \end{array}$$

Theorem 5.4 *Suppose that Slater's constraint qualification holds, i.e. there exists a positive definite feasible solution $\hat{X}$. The matrix $\bar{P} \succeq 0$ and vector-matrix $\bar{y}, \bar{\Lambda} \succeq 0$ solve CM and DCM if and only if*

$$\begin{array}{cc} K\bar{P} = b & \textit{primal feas.} \\ 2H^{(2)} \circ (\bar{P} - A) - K^*\bar{y} - \bar{\Lambda} = 0 & \textit{dual feas.} \\ \operatorname{trace} \bar{\Lambda}\bar{P} = 0 & \textit{compl. slack.} \end{array}$$

■

Primal-dual interior-point methods are based on solving a perturbation of the above optimality conditions.

$$\begin{array}{cc} K\bar{P} = b & \text{primal feas.} \\ 2H^{(2)} \circ (\bar{P} - A) - K^*\bar{y} - \bar{\Lambda} = 0 & \text{dual feas.} \\ \bar{P} - \mu\bar{\Lambda}^{-1} = 0 & \text{pert. compl. slack.} \end{array} \tag{5.1}$$

Remark 5.5 *In fact, most algorithms use the*

$$\bar{\Lambda}\bar{P} - \mu I = 0$$

perturbed version of complementary slackness. We specifically use (5.1), since it allows us to exploit sparsity. (However, we pay for this with some loss of accuracy near the optimum.) See [113] for a discussion of the many different choices for search directions.

Two algorithms can be derived. The dual-step-first exploits sparsity if many elements are free; while the primal-step-first exploits sparsity if many elements are fixed. The details are given in [57]. We will follow a similar strategy below when deriving algorithms for the (approximate) EDM completion problem, see Sections 5.3 and 5.4.

Numerical tests show that large sparse problems can be solved very efficiently. We include a few tests done on a Sparc 20 using Matlab 5.3 . The time per iteration (though not included) was directly proportional to the number of fixed elements (non-zeros) in the dual-step-first method, e.g. for $n = 155$ this was typically 16 seconds cpu time. Similar results held for the primal-step-first algorithm, i.e. the time was proportional to the number of free elements. The details for several of the tests follow in Tables 2 and 3. (Each test appears on one line and includes 20 test problems.)

dim	toler	H dens. / infty	$A \succeq 0$	cond(A)	$H \succ 0$	min/max	iters
83	10^{-6}	.007/.001	no	235.1	no	24/29	25.5
85	10^{-5}	.008/.001	yes	94.7	no	11/17	13.1
85	10^{-6}	.0075/.001	no	299.9	no	23/27	25.2
87	10^{-6}	.006/.001	yes	74.2	yes	14/19	16.9
89	10^{-6}	.006/.001	no	179.3	no	23/28	15.2
110	10^{-6}	.007/.001	yes	172.3	yes	15/20	17.8
155	10^{-6}	.01/0	yes	643.9	yes	14/18	15.3
655	10^{-6}	.017/0	yes	1.4	no	13/16	14.
755	10^{-6}	.002/0	yes	1.5	no	14/17	15.

Table 2: PSD completion data for dual-step-first method (20 problems per test): dimension; tolerance for duality gap; density of non-zeros in H/ density of infinite values in H; positive semidefiniteness of A; positive definiteness of H; min and max number of iterations; average number of iterations.

5.3 Approximate EDM Completions

We now look at the EDM completion problem. We follow the successful approach above and use some known characterizations of EDMs. (The details can be found in [1].)

An $n \times n$ symmetric matrix $D = (d_{ij})$ with nonnegative elements and zero diagonal is called a *pre-distance matrix* (or dissimilarity matrix). A

dim	toler	H dens. / infty	$A \succeq 0$	cond(A)	$H \succ 0$	min/max	iters
85	10^{-5}	.0219/.02	yes	1374.5	no	16/23	18.9
95	10^{-5}	.0206/.02	yes	2.7	no	8/14	11.1
95	10^{-6}	1/.999	yes	196.	yes	14/18	16.8
145	10^{-6}	.01/.997	yes	658.5	yes	13/17	14.9

Table 3: PSD completion data for primal-step-first (20 problems per test): dimension; tolerance for duality gap; density of non-zeros in H/ density of infinite values in H; positive semidefiniteness of A; positive definiteness of H; min and max number of iterations; average number of iterations.

pre-distance matrix such that there exists points $x^1, x^2, \ldots, x^n$ in $\Re^r$ with

$$d_{ij} = \|x^i - x^j\|^2, \quad i,j = 1,2,\ldots,n$$

is called a *(squared) Euclidean distance matrix* (EDM). The smallest value of r is called *the embedding dimension* of D. (r is always $\leq n-1$.)

Given a partial symmetric matrix A with certain elements specified, the *Euclidean distance matrix completion problem* (EDMCP) consists in finding the unspecified elements of A that make A a EDM. In other words, we wish to determine the relative locations of points in Euclidean space, when we are only given a subset of the pairwise distances between the points.

There are surprisingly many applications for this problem, sometimes called the molecule problem. These applications include NMR data, determination of protein structure, surveying, satellite ranging, and molecular conformation; e.g. the survey [24] and the discussion in [51] and the related papers [50, 90, 115, 118, 45].

We now consider the approximate EDMCP and follow the approach in [1], where the reader will find all the proofs and details omitted here. Let A be a pre-distance matrix and let H be an $n \times n$ symmetric matrix with nonnegative elements (weights). Consider the objective function

$$f(D) := \|H \circ (A - D)\|_F^2,$$

where $\circ$ denotes *Hadamard product.* The *weighted, closest Euclidean distance matrix problem* is

$$(CDM_0) \qquad \mu^* := \begin{array}{cc} \min & f(D) \\ \text{subject to} & D \in \mathcal{E}, \end{array}$$

where $\mathcal{E}$ denotes the cone of EDMs.

5.3.1 EDM Model

The cone of EDM is homeomorphic to a face of the cone of positive semidefinite matrices. This can be seen from the fact that a pre-distance matrix D is a EDM if and only if D is negative semidefinite on

$$M := \left\{x \in \Re^n : x^t e = 0\right\},$$

where e is the vector of all ones. (For these and other related results see the development in [1].) Now, define

$$V \text{ is } n \times (n-1) \quad \text{full column rank with} \quad V^t e = 0. \tag{5.2}$$

Then

$$J := VV^{\dagger} = I - \frac{ee^t}{n} \tag{5.3}$$

is the orthogonal projection onto M, where $V^{\dagger}$ denotes the Moore-Penrose generalized inverse.

Define the *centered* and *hollow* subspaces

$$\begin{aligned} \mathcal{S}_C &:= \{B \in \mathcal{S}^n : Be = 0\}, \\ \mathcal{S}_H &:= \{D \in \mathcal{S}^n : \operatorname{diag}(D) = 0\}, \end{aligned}$$

and the two linear operators

$$\begin{aligned} \mathcal{K}(B) &:= \operatorname{diag}(B)\, e^t + e \operatorname{diag}(B)^t - 2B, \\ \mathcal{T}(D) &:= -\tfrac{1}{2} JDJ. \end{aligned}$$

The operator $-2\mathcal{T}$ is an orthogonal projection onto $\mathcal{S}_C$.

Theorem 5.6 *The linear operators satisfy*

$$\begin{aligned} \mathcal{K}(\mathcal{S}_C) = \mathcal{S}_H, \\ \mathcal{T}(\mathcal{S}_H) = \mathcal{S}_C, \end{aligned}$$

and $\mathcal{K}_{|\mathcal{S}_C}$ and $\mathcal{T}_{|\mathcal{S}_H}$ are inverses of each other. ■

Lemma 5.7 *The hollow matrix $D \in \mathcal{E}$ if and only if*

$$v^T e = 0 \quad \Rightarrow \quad v^T X v \leq 0.$$

From the above we see that a hollow matrix D is EDM if and only if it is negative semidefinite on the orthogonal complement of e, i.e. if and only if $B = \mathcal{T}(D) \succeq 0$ (positive semidefinite). Alternatively, D is EDM if and only if $D = \mathcal{K}(B)$, for some B with $Be = 0$ and $B \succeq 0$. In this case the embedding dimension r is given by the rank of B. Moreover if $B = XX^t$, then the coordinates of the points $x^1, x^2, \ldots, x^n$ that generate D are given by the rows of X and, since $Be = 0$, it follows that the origin coincides with the centroid of these points.

The cone of EDMs, $\mathcal{E}$, has empty interior. This can cause problems for interior-point methods. We can correct this by projection and moving to a smaller dimensional space [1]; note that

$$V \cdot V : \mathcal{S}_{n-1} \to \mathcal{S}_n$$

$$V \cdot V : \mathcal{P}_{n-1} \to \mathcal{P}_n$$

Define the composite operators

$$\mathcal{K}_V(X) \ := \ \mathcal{K}(VXV^t),$$

and

$$\mathcal{T}_V(D) \ := \ V^{\dagger}\mathcal{T}(D)(V^{\dagger})^t = -\tfrac{1}{2}V^{\dagger}D(V^{\dagger})^t.$$

Lemma 5.8

$$\mathcal{K}_V(\mathcal{S}_{n-1}) = \mathcal{S}_H,$$
$$\mathcal{T}_V(\mathcal{S}_H) = \mathcal{S}_{n-1},$$

and $\mathcal{K}_V$ *and* $\mathcal{T}_V$ *are inverses of each other on these two spaces.* ■

Corollary 5.9

$$\mathcal{K}_V(\mathcal{P}) \ = \ \mathcal{E},$$
$$\mathcal{T}_V(\mathcal{E}) \ = \ \mathcal{P}.$$

■

We can summarize the above and obtain the model used in [1] (Re)Define the closest EDM problem:

$$f_0(X) := \|H \circ (A - \mathcal{K}_V(X))\|_F^2$$
$$= \|H \circ \mathcal{K}_V(B - X)\|_F^2,$$

where $B = \mathcal{T}_V(A)$. ($\mathcal{K}_V$ and $\mathcal{T}_V$ are both linear operators)

$$
\begin{array}{lccl}
 & \mu_0^* := & \min & f_0(X) \\
(\mathrm{CDM}_0) & & \text{subject to} & \mathcal{A}X = b \\
 & & & X \succeq 0.
\end{array}
$$

The additional constraint using $\mathcal{A} : \mathcal{S}_{n-1} \longrightarrow \Re^m$, could represent some of the fixed elements in the given matrix A.

Numerical tests for this model are given in [1]. The number of iterations are comparable to those for the semidefinite completion problem (Section 5.2), though the time per iteration was much higher, i.e. sparsity was not exploited efficiently.

5.4 Alternate EDM Model for the Large Sparse Case

The above model appears to be quite efficient for solving the EDM completion problem. It handles the lack of interiority and actually reduces the dimension of the problem. There is one major difference between this model CDM_0 and the one used in Section 5.2. That is, the operator $H\circ$ in the objective function is replaced by $H \circ \mathcal{K}_V$. This change allows one to reduce the dimension of the problem and obtain Slater's constraint qualification for both the primal and dual problems. However, one cannot exploit sparsity as one did in CDM. As is often the case in modelling, a model that appears to be simpler is often not more efficient in computations. We now outline a different approach that increases the dimension of the problem but can exploit sparsity. The details can be found in [3]. (Recall that e denotes the vector of ones.)

Lemma 5.10 *Let*

$$
\begin{array}{lcl}
\mathcal{F} & := & \{X \in \mathcal{S}^n : v^T e = 0 \quad \Rightarrow \quad v^T X v \leq 0\}, \\
\mathcal{F}_0 & := & \{X \in \mathcal{S}^n : X - \alpha e e^t \preceq 0, \quad \textit{for some } \alpha \geq 0\}, \\
\mathcal{F}_1 & := & \{X \in \mathcal{S}^n : X - \alpha e e^t \preceq 0, \quad \forall\, \alpha \geq \bar{\alpha}, \textit{ for some } \bar{\alpha} \geq 0\}.
\end{array}
$$

Then

$$
\mathrm{ri}\,(\mathcal{F}) \subset \mathcal{F}_0 = \mathcal{F}_1 \subset \mathcal{F} \subset \overline{\mathcal{F}_0}. \tag{5.4}
$$

Proof. Suppose that $\bar{X} \in \mathrm{ri}\,(\mathcal{F})$ (i.e. $v^T e = 0, v \neq 0 \Rightarrow v^T \bar{X} v < 0$) but $\bar{X} \notin \mathcal{F}_0$. Then, for each $\alpha \geq 0$, there exists w_α with $||w_\alpha|| = 1$, such that $w_\alpha \to \bar{w}$, as $\alpha \to \infty$ and

$$
w_\alpha^T(\bar{X} - \alpha e e^t) w_\alpha > 0, \quad \forall\, \alpha \geq 0,
$$

i.e.

$$w_\alpha^T \bar{X} w_\alpha > \alpha w_\alpha^T e e^t w_\alpha, \quad \forall\, \alpha \geq 0.$$

Since w_α converges and the left-hand-side of the above inequality must be finite, this implies that $e^t \bar{w} = \bar{w}^T \bar{X} \bar{w} = 0$, a contradiction. Therefore, $\mathrm{ri}\,(\mathcal{F}) \subset \mathcal{F}_0$. That $\mathcal{F}_0 = \mathcal{F}_1$ is clear.

Now suppose that $\bar{X} - \alpha e e^t \preceq 0$, $\alpha \geq 0$. Let $v^T e = 0$. Then $0 \geq v^T(\bar{X} - \alpha e e^t)v = v^T \bar{X} v$, i.e. $\mathcal{F}_0 \subset \mathcal{F}$. The final inclusion comes from the first and the fact that $\mathcal{F}$ is closed. ■

Unfortunately, we cannot enforce equality in (5.4). This can be seen from the fact that $\mathcal{F}_0 = \mathcal{P} + \mathrm{span}\,\{ee^t\} = \mathcal{P} + \mathrm{span}\mathcal{F}$, where $\mathcal{F} = \mathrm{cone}\,\{ee^t\}$ is a face (actually a ray) of the positive semidefinite cone generated by ee^t. $\mathcal{P}$ and the sum of $\mathcal{P}$ and the span of a face is never closed, see [105, Lemma 2.2]. If we assume that X is hollow, then the same result holds. This is used in the algorithm for large problems.

Corollary 5.11 *Let*

$$\begin{aligned}
\mathcal{E} &:= \{X \in \mathcal{S}_H : v^T e = 0 \quad \Rightarrow \quad v^T X v \leq 0\}, \\
\mathcal{E}_0 &:= \{X \in \mathcal{S}_H : X - \alpha e e^t \preceq 0, \quad \textit{for some } \alpha\}, \\
\mathcal{E}_1 &:= \{X \in \mathcal{S}_H : X - \alpha e e^t \preceq 0, \quad \forall\, \alpha \geq \bar{\alpha}, \textit{ for some } \bar{\alpha}\}.
\end{aligned}$$

Then

$$\mathrm{ri}\,(\mathcal{E}) \subset \mathcal{E} = \mathcal{E} \subset \mathcal{E} \subset \overline{\mathcal{E}}. \tag{5.5}$$

Proof. The proof is similar to that in the above Lemma 5.10. We only include the details about the closure.

Suppose that $0 \neq X_k \in \mathcal{E}_0$, i.e. $\mathrm{diag}\,(X_k) = 0, X_k \preceq \alpha_k E$, for some α_k; and, suppose that $X_k \to \bar{X}$. Since X_k is hollow it has exactly one positive eigenvalue and this must be smaller than α_k. However, since X_k converges to $\bar{X}$, we conclude that $\bar{X} \leq \lambda_{\max}(\bar{X})E$, where $\lambda_{\max}(\bar{X})$ is the largest eigenvalue of $\bar{X}$. ■

We can now use a different simplified objective function to obtain a new model. We let $E = ee^t$ and

$$f(P) := \|H \circ (A - P)\|_F^2,$$

and

$$
\text{(CDM)} \qquad \begin{array}{rcc} \mu^* := & \min & f(P) \\ & \text{subject to} & \mathcal{K}P = b \\ & & \alpha E - P \succeq 0, \end{array}
$$

where $\mathcal{K}$ is a linear operator. We assume that this linear equality constraint contains the constraint $\operatorname{diag}(P) = 0$, i.e. that P is a hollow matrix.

We now derive the dual problem for CDM. For $\Lambda \in \mathcal{S}^n$ and $y \in \Re^m$, let

$$
L(P, \alpha, \Lambda, y) = f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda(\alpha E - P) \tag{5.6}
$$

denote the *Lagrangian* of CDM. It is easy to see that the primal problem CDM is equivalent to

$$
\mu^* = \min_{P,\alpha} \max_{\substack{y \\ \Lambda \succeq 0}} L(P, \alpha, \Lambda, y). \tag{5.7}
$$

We assume that the generalized Slater's constraint qualification,

$$
\exists\ \alpha, P \text{ with } P - \alpha E \prec 0,\ KP = b,
$$

holds for CDM.

Slater's condition implies that strong duality holds, i.e. this means

$$
\mu^* = \nu^* := \max_{\substack{y \\ \Lambda \succeq 0}} \min_{P,\alpha} L(P, \alpha, \Lambda, y) \tag{5.8}
$$

and ν^* is attained for some $\Lambda \succeq 0, y$ see e.g. [83]. The inner minimization of the convex, in P, Lagrangian is unconstrained and we can differentiate to get the equivalent problem

$$
\nu^* = \max_{\substack{\nabla L(P,\alpha,\Lambda,y)=0 \\ \Lambda \succeq 0}} f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda(\alpha E - P). \tag{5.9}
$$

We can now state the dual problem.

$$
\text{(DCDM)} \qquad \begin{array}{rcc} \nu^* := & \max & f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda(\alpha E - P) \\ & \text{subject to} & \nabla_P f(P) - \mathcal{K}^* y + \Lambda = 0 \\ & & -\operatorname{trace} \Lambda E = 0 \\ & & \Lambda \succeq 0. \end{array} \tag{5.10}
$$

The above pair of dual problems, CDM and DCDM, provide an optimality criteria in terms of feasibility and complementary slackness. This provides the basis for many algorithms including primal-dual interior-point algorithms. In particular, we see that the duality gap, in the case of primal and dual feasibility, is given by the difference of the primal and dual optimal values:

$$-\langle y, b - KP\rangle + \operatorname{trace}\Lambda(\alpha E - P) = \operatorname{trace}\Lambda(\alpha E - P). \tag{5.11}$$

Using the derivative $\nabla_P f(P) = 2H^{(2)} \circ (P - A)$, and primal-dual feasibility, we see that complementary slackness is given by

$$\operatorname{trace}(\alpha E - P)\left(-2H^{(2)} \circ (P - A) + K^* y\right) = 0. \tag{5.12}$$

Theorem 5.12 *The pair $\bar{P} \succeq 0, \bar{\alpha}$ and $\bar{\Lambda} \succeq 0, \bar{y}$ solve CDM and DCDM if and only if*

$$\begin{array}{cl} K\bar{P} = b & \textit{primal feasibility} \\ 2H^{(2)} \circ (\bar{P} - A) - K^*\bar{y} - \bar{\Lambda} = 0, \quad -\operatorname{trace}\Lambda E = 0 & \textit{dual feasibility} \\ \operatorname{trace}\bar{\Lambda}\left(\bar{\alpha}E - \bar{P}\right) = 0 & \textit{compl. slack.} \end{array}$$

■

The above yields an equation for the solution of CDM. (Recall that the primal feasibility constraint is assumed to include the fact that P is a hollow matrix.) However, we do not apply a Newton type method directly to this equation but rather to a perturbed equation which allows us to stay interior to $\mathcal{P}$ and $\Re_+$. We note that though the generalized Slater's constraint qualification holds for the primal, it fails for the dual since $\Lambda \succ 0 \Rightarrow \operatorname{trace}\Lambda E > 0$. Therefore, there is no duality gap between the optimal values, but numerical complications can arise. We address this later on.

5.4.1 Interior-point algorithms

We now present the interior-point algorithms for CDM. We present a dual-step-first. (A primal-step-first version can be similarly derived.) The difference in efficiency arises from the fact that the primal variable P does not change very much if few elements of A are free, while the dual variable Λ does not change very much if many elements of A are free.

Since we can increase the weights in H to try and fix certain elements of P, we restrict ourselves to the case where the only linear equality constraints are those that fix the diagonal at 0.

5.4.2 The Log-Barrier Approach

We now derive a primal-dual interior-point method using the log-barrier approach, see e.g. [49]. This is an alternative way of deriving the optimality conditions in Theorem 5.12. The log-barrier problem for CDM is

$$\min_{\substack{\operatorname{diag}(P)=0 \\ P\succ 0}} B_\mu(P) := f(P) - \mu \log\det(\alpha E - P),$$

where $\mu \downarrow 0$. For each $\mu > 0$ we take one Newton step toward minimizing the log-barrier function. The Lagrangian for this problem is

$$f(P) - y^t \operatorname{diag}(P) - \mu \log\det(\alpha E - P).$$

Therefore, we take one Newton step for solving the stationarity conditions

$$\begin{array}{rcl} \nabla_P & = & 2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + \mu\,(\alpha E - P)^{-1} = 0 \\ \nabla_\alpha & = & -\mu \operatorname{trace} E\,(\alpha E - P)^{-1} = 0 \\ \operatorname{diag}(P) & = & 0. \end{array} \tag{5.13}$$

After the substitution $-\mu\,(\alpha E - P)^{-1} = 2H^{(2)} \circ (P - A) - \operatorname{Diag}(y)$, the first two equations become the perturbed complementary slackness equations. The new optimality conditions are

$$\begin{array}{rcl} (\alpha E - P)\left(-2H^{(2)} \circ (P - A) + \operatorname{Diag}(y)\right) & = & \mu I, \\ \operatorname{trace} E\left(2H^{(2)} \circ (P - A) - \operatorname{Diag}(y)\right) & = & 0 \\ \operatorname{diag}(P) & = & 0. \end{array} \tag{5.14}$$

And, the estimate of the barrier parameter is

$$n\mu \;=\; \operatorname{trace}(\alpha E - P)\left(-2H^{(2)} \circ (P - A) + \operatorname{Diag}(y)\right). \tag{5.15}$$

The Newton direction is dependent on which of the equations (5.13),(5.14) we choose to solve. The equation (5.14) is shown to perform better in many applications. A discussion on various choices is given in [113]. (See also [73].) However we choose (5.13) below in order to exploit sparsity. The linearization to find the Newton direction is done below.

5.4.3 Primal-Dual Feasible Algorithm - Dual Step First

The algorithm essentially solves for the step h, w and backtracks to ensure both primal and dual strict feasibility. This yields the primal-step-first algorithm since we only solve for the step h, w for changes in the primal variables P, α. We do need to evaluate the dual variable to update the barrier parameter μ using the perturbed complementarity condition.

Alternatively, we can work with dual step and perturbed complementary slackness. (We follow the approach in [49]. See also [86].) We keep primal feasibility, identify Λ

$$\Lambda = \mu(\alpha E - P)^{-1}, \tag{5.16}$$

and replace equations (5.13) and (5.14). This yields

$$\begin{array}{cc} \operatorname{diag}(P) = 0 & \text{primal feasibility} \\ 2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + \Lambda = 0, \quad -\operatorname{trace} \Lambda E = 0 & \text{dual feasibility} \\ -(\alpha E - P) + \mu \Lambda^{-1} = 0, & \text{pert. compl. slack.} \end{array} \tag{5.17}$$

Remark 5.13 *Dual feasibility implies that* $\operatorname{trace} \Lambda E = 0$. *Therefore,*

$$\Lambda = V \hat{\Lambda} V^t, \ \hat{\Lambda} \succ 0,$$

where V is defined in (5.2). There are many choices for V. In particular, we can make a sparse choice, i.e. one with many zero elements. Therefore, in an interior-point approach we cannot maintain dual feasibility, e.g. during the algorithm $\operatorname{trace} \Lambda E > 0$ *with* $= 0$ *only in the limit.*

Alternatively, we could eliminate the troublesome equation in the dual to obtain the following equivalent characterization of optimality

$$\begin{array}{cc} \operatorname{diag}(P) = 0 & \textit{primal feasibility} \\ 2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + V \Lambda V^t = 0 & \textit{dual feasibility} \\ -(\alpha E - P) + \mu V \Lambda^{-1} V^t = 0 & \textit{pert. compl. slack.} \end{array} \tag{5.18}$$

We apply Newton's method to solve (5.17). We let

h	denote the step for	P
w	denote the step for	α
l	denote the step for	Λ
s	denote the step for	y.

(By abuse of notation, we use l as a matrix here and also as an index. The meaning is clear from the context.) We get

$$\operatorname{diag}(h) = -\operatorname{diag}(P), \tag{5.19}$$

i.e. the diagonal (linear) constraint will be satisfied if we take a full Newton step or if we start with the initial $\operatorname{diag}(P) = 0$. Therefore, we may as well start with $\operatorname{diag}(P) = 0$ and restrict $\operatorname{diag}(h) = 0$. Then linearization of the complementary slackness equation yields

$$-(\alpha + w)E + (P + h) + \mu\Lambda^{-1} - \mu\Lambda^{-1} l \Lambda^{-1} \;=\; 0,$$

or

$$(\alpha + w)E - P - h \;=\; \mu\Lambda^{-1} - \mu\Lambda^{-1} l \Lambda^{-1}, \tag{5.20}$$

where $\operatorname{diag}(P) = \operatorname{diag}(h) = 0$. We get

$$h \;=\; -\mu\Lambda^{-1} + \mu\Lambda^{-1} l \Lambda^{-1} - P + (\alpha + w)E. \tag{5.21}$$

and

$$l = \frac{1}{\mu}\Lambda \left\{P + h - (\alpha + w)E\right\}\Lambda + \Lambda. \tag{5.22}$$

The linearization of the dual feasibility equations yields

$$\begin{array}{rcl} 2H^{(2)} \circ h - \operatorname{Diag}(s) + l & = & -(2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + \Lambda), \\ -\operatorname{trace} lE & = & \operatorname{trace} \Lambda E, \end{array} \tag{5.23}$$

with $\operatorname{diag}(P) = \operatorname{diag}(h) = 0$. We assume that we start with an initial primal-dual feasible solution. However, we include the feasibility equation on the right-hand-side of (5.23), because roundoff error can cause loss of feasibility. (Since Newton directions maintain linear equations, we could theoretically substitute for h in this linearization with the right-hand side being 0. We do however forcibly maintain a zero diagonal.)

We can eliminate the primal step h and dual step s and solve for the dual step l, w. From the linearization of the dual in (5.23) and the expression for h in (5.21),

$$\begin{array}{rcl} -\operatorname{Diag}(s) + l & = & -2H^{(2)} \circ h - (2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + \Lambda) \\ & = & -2H^{(2)} \circ (-\mu\Lambda^{-1} + \mu\Lambda^{-1} l \Lambda^{-1} - P + (\alpha + w)E) \\ & & \quad - (2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + \Lambda) \\ \operatorname{diag}(h) & = & \operatorname{diag}(-\mu\Lambda^{-1} + \mu\Lambda^{-1} l \Lambda^{-1} - P + (\alpha + w)E) = 0 \\ \operatorname{trace}(lE) & = & -\operatorname{trace}(\Lambda E). \end{array} \tag{5.24}$$

Since we have the constraint $\operatorname{diag}(P) = 0$ in CDM, we can, without loss of generality, set the diagonal of the weight matrix H to zero, i.e. $\operatorname{diag}(H) = 0$. We can start with initial $\operatorname{diag}(P) = 0$ and $\operatorname{diag}(\Lambda) = y$. Therefore

$$s = \operatorname{diag}(l).$$

We can now eliminate s from the first equation.

$$\begin{array}{rcl} -\operatorname{Diag}\operatorname{diag}(l) + l & = & -2H^{(2)} \circ (-\mu\Lambda^{-1} + \mu\Lambda^{-1}l\Lambda^{-1} + (\alpha + w)E) \\ & & \quad - (2H^{(2)} \circ (-A) - \operatorname{Diag}(y) + \Lambda). \end{array} \tag{5.25}$$

and, assuming that $\operatorname{diag}(P) = 0$,

$$\begin{array}{rcl} 0 & = & \operatorname{diag}(-\mu\Lambda^{-1} + \mu\Lambda^{-1}l\Lambda^{-1} - P + (\alpha + w)E) \\ & = & \mu\operatorname{diag}(-\Lambda^{-1} + \Lambda^{-1}l\Lambda^{-1}) + (\alpha + w)e \end{array} \tag{5.26}$$

From this we already see that if Λ started sparse and H was similarly sparse, then Λ stays sparse and l is sparse.

We can now move the variables to the left and get the Newton equation

$$\begin{array}{rcl} 2H^{(2)} \circ (wE + \mu\Lambda^{-1}l\Lambda^{-1}) - \operatorname{Diag}\operatorname{diag}(l) + l & = & 2H^{(2)} \circ \{\mu\Lambda^{-1} + A - \alpha E\} \\ & & +\operatorname{Diag}(y) - \Lambda \\ \operatorname{diag}(\mu\Lambda^{-1}l\Lambda^{-1}) + we & = & \operatorname{diag}(\mu\Lambda^{-1}) - \alpha e \\ \operatorname{trace}(lE) & = & -\operatorname{trace}(\Lambda E). \end{array} \tag{5.27}$$

This system is square, order $1+t(n) = 1+\frac{n(n+1)}{2}$, since we need only consider the strictly upper triangular part in the first equation and Λ, l are symmetric matrices.

We can now solve this system for l, set $s = \operatorname{diag}(l), t = -\operatorname{trace}(\Lambda+l)E-\lambda$, and substitute to find h, w. We then take the primal-dual step and backtrack to ensure both primal and dual positive definiteness. Note that we cannot maintain dual positive definiteness if we maintain dual feasibility. However, we can maintain dual positive definiteness on the orthogonal complement of e, i.e. maintain $V^t\Lambda V \succ 0$.

Let nnz denote the number of nonzero, upper triangular, elements of H. We assume that the diagonal of H is zero and H is symmetric. Let F denote the $nnz + n \times 2$ matrix with row p denoting the indices of the p-th nonzero, upper triangular, element of $H + I$ ordered by columns, i.e. for $p = 1, \ldots, nnz + n$,

$$\{(F_{p1}, F_{p2})_{p=1,\ldots nnz+n}\} = \{ij : H_{ij} \neq 0, i \leq j, \text{ ordered by columns}\}. \tag{5.28}$$

Let δ_{ij} denote the *Kronecker delta function*, i.e. it is 1 if $i=j$ and 0 otherwise; δ_{ijkl} is 1 when all $i=j=k=l$ and 0 otherwise; $\delta_{(ij)(kl)}$ is 1 when $(ij)=(kl)$ and 0 otherwise. Let $E_{ij}=\left(e_ie_j^t+e_j^te_i\right)/\sqrt{2}$ denote the ij unit matrix in $\mathcal{S}^n$, where $E_{ij}=\left(e_ie_j^t+e_j^te_i\right)/2$ if $i=j$. (This set of matrices forms an orthonormal basis of $\mathcal{S}^n$.) Then $\operatorname{trace} EE_{ij}=\sqrt{2}$ (resp. 1) if $i\neq j$ (resp. $i=j$). From (5.27) the first $t(n)$ rows, with $w=0$, $k\neq l$, and $k=l$, components of the left-hand-side are, respectively,

$$
\begin{array}{rcl}
k\neq l, i\neq j \text{ LHS } (5.27) & = & \operatorname{trace} E_{kl}\left\{2H^{(2)}\circ(\mu\Lambda^{-1}E_{ij}\Lambda^{-1})\right.\\
& & \qquad \left.-\operatorname{Diag}\operatorname{diag}(E_{ij})+E_{ij}\right\}\\
& = & \mu\operatorname{trace}\left(e_ke_l^t+e_le_k^t\right)\left(H^{(2)}\right.\\
& & \qquad \left.\circ\Lambda^{-1}(e_ie_j^t+e_je_i^t)\Lambda^{-1}\right)+\delta_{(ij)(kl)}\\
& = & \mu\left[2e_l^t\left(H^{(2)}\circ\Lambda^{-1}_{:,i}\Lambda^{-1}_{j:}\right)e_k+\right.\\
& & \qquad \left.2e_k^t\left(H^{(2)}\circ\Lambda^{-1}_{:,i}\Lambda^{-1}_{j:}\right)e_l\right]+\delta_{(ij)(kl)};\\
k\neq l, i\neq j \text{ LHS } (5.27) & = & 2\mu H^{(2)}_{kl}\left(\Lambda^{-1}_{li}\Lambda^{-1}_{jk}+\Lambda^{-1}_{ki}\Lambda^{-1}_{jl}\right)+\delta_{(ij)(kl)};\\
k\neq l, i=j \text{ LHS } (5.27) & = & \operatorname{trace} E_{kl}\left\{2\mu H^{(2)}\circ[\Lambda^{-1}E_{jj}\Lambda^{-1}]\right.\\
& & \qquad \left.-\operatorname{Diag}\operatorname{diag}(E_{jj})+E_{jj}\right\}\\
& = & 2\sqrt{2}\mu\operatorname{trace} e_ke_l^t\left(H^{(2)}\circ\Lambda^{-1}e_je_j^t\Lambda^{-1}\right)\\
& = & 2\sqrt{2}\mu H^{(2)}_{kl}\left(\Lambda^{-1}_{lj}\Lambda^{-1}_{jk}\right);\\
k=l, i\neq j \text{ LHS } (5.27) & = & \sqrt{2}\mu\Lambda^{-1}_{ki}\Lambda^{-1}_{jk}, \quad k=1,\dots n;\\
k=l, i=j \text{ LHS } (5.27) & = & \mu\Lambda^{-1}_{ki}\Lambda^{-1}_{ik}, \quad k=1,\dots n.
\end{array}
\tag{5.29}
$$

The last column of LHS, with the matrix $l=0$ and $w=1$, is:

$$
\begin{array}{rcl}
w=1, k\neq l \text{ LHS } (5.27) & = & \operatorname{trace}\left(E_{kl}(2H^{(2)}\circ E)\right);\\
w=1, k=l \text{ LHS } (5.27) & = & 1.
\end{array}
\tag{5.30}
$$

While the last row of LHS is:

$$
\begin{array}{rcl}
i\neq j \text{ LHS } (5.27) & = & \operatorname{trace}\left(E_{ij}E\right)=\sqrt{2};\\
i=j \text{ LHS } (5.27) & = & 1.
\end{array}
\tag{5.31}
$$

Suppose that we represent the Newton system as

$$
\operatorname{sMat}\left[L(\operatorname{svec}(l))\right]=\operatorname{sMat}\left[\operatorname{svec}(RHS)\right],
\tag{5.32}
$$

where $\operatorname{svec}(S)$ denotes the vector formed from the nonzero elements of the columns of the upper triangular part of the symmetric matrix S, where the strict upper triangular part of S is multiplied by $\sqrt{2}$. This guarantees that

$\operatorname{trace} XY = \operatorname{svec}(X)^t \operatorname{svec}(Y)$, i.e. svec is an isometry; the operator sMat is its inverse, and RHS is the matrix on the right-hand-side of (5.27). The system is order $nnz + n$. From (5.32), we can write the system as a matrix and column vector equation with matrix L and vector of unknowns $\operatorname{svec}(l)$.

$$L(\operatorname{svec}(l)) = \operatorname{svec}(RHS). \tag{5.33}$$

Then for

$$p = kl, k \leq l, \;\; q = ij, i \leq j,$$

the pq component of the matrix L is

$$L_{pq} = \begin{cases} 2\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_1}} + \Lambda^{-1}_{F_{p_1},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_2}} \right) & \text{if } p \neq q, \; k \neq l, \; i \neq j; \\ 2\sqrt{2}\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_2}} \Lambda^{-1}_{F_{q_2},F_{p_1}} \right) & \text{if } p \neq q, \; k \neq l, \; i = j; \\ 2\sqrt{2}\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_2}} \Lambda^{-1}_{F_{q_2},F_{p_1}} \right) & \text{if } p = q, \; k \neq l, \; i = j; \\ 2\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_1}} + \Lambda^{-1}_{F_{p_1},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_2}} \right) + 1 & \text{if } p = q, k \neq l, \; i \neq j; \\ \sqrt{2}\mu \Lambda^{-1}_{F_{p_1},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_1}} & \text{if } k = l, \; i \neq j; \\ \mu \Lambda^{-1}_{F_{p_1},F_{q_1}} \Lambda^{-1}_{F_{q_1},F_{p_1}} & \text{if } k = l, \; i = j \\ 2\sqrt{(2)} H^{(2)}_{F_{p_2},F_{p_1}} & \text{if } w = 1, \; k \neq l \\ 1 & \text{if } w = 1, \; k = l. \end{cases} \tag{5.34}$$

The p-th row can be calculated using the Hadamard product of pairs of columns of Λ^{-1},

$$\Lambda^{-1}_{F_{p_2},F_{:,1}} \circ \Lambda^{-1}_{F_{p_1},F_{:,2}}. \tag{5.35}$$

This allows for complete vectorization and simplifies the construction of the linear system, especially in the large sparse case.

The $p = kl, k \leq l$, and last row, component of the right-hand-side of the system (5.32) is

$$RHS_p = \begin{cases} \sqrt{2} \left(2H^{(2)}_p \circ \left\{ \mu \Lambda^{-1}_p + A_p - \alpha \right\} - \Lambda_p \right), & \text{if } k \neq l \\ \mu \Lambda^{-1}_{kk} - \alpha & \text{if } k = l \\ -\operatorname{trace}(\Lambda E) & \text{last row} \end{cases}$$

The above provides a sparse system of linear equations for the search direction in a primal-dual interior-point algorithm. One would then take a step in this direction, backtrack to guarantee positive definiteness and then repeat the process with a new system, i.e. follow the standard paradigm for these algorithms.

5.4.4 Primal-Dual Feasible Algorithm - Primal Step First

Alternatively, if many elements of H are sufficiently large, i.e. if we fix (or specify) elements of A, then it is more efficient to eliminate l and solve for h first. The algorithm is similar to the dual-step-first one. The details can be found in [3].

6 Conclusion

In this paper we showed the strength of Lagrangian relaxation for obtaining semidefinite programming relaxations for several discrete optimization problems. We have presented a recipe for finding such relaxations based on adding redundant quadratic constraints and using Lagrangian duality and illustrated it with several examples, including the derivation of new strengthened SDP relaxations for MC. We also discussed the application of SDP to matrix completion problems. We showed how SDP can be used to find approximate positive semidefinite and Euclidean distance matrix completions and we concluded by presenting a new SDP algorithm which exploits sparsity and structure in large instances of Euclidean distance matrix completion problems.

References

[1] A. ALFAKIH, A. KHANDANI, and H. WOLKOWICZ. Solving Euclidean distance matrix completion problems via semidefinite programming. *Comput. Optim. Appl.*, 12(1-3):13–30, 1999. Computational optimization—a tribute to Olvi Mangasarian, Part I.

[2] A. ALFAKIH and H. WOLKOWICZ. Matrix completion problems. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms,*

and Applications, page 11. Kluwer Academic Publishers, Boston, MA, 2000.

[3] A. ALFAKIH and H. WOLKOWICZ. A new semidefinite programming model for large sparse Euclidean distance matrix completion problems. Technical Report CORR 2000-37, University of Waterloo, Waterloo, Canada, 2000. in progress.

[4] F. ALIZADEH. *Combinatorial optimization with interior point methods and semidefinite matrices.* PhD thesis, University of Minnesota, 1991.

[5] F. ALIZADEH. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.*, 5:13–51, 1995.

[6] 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.

[7] M.F. ANJOS and H. WOLKOWICZ. A strengthened SDP relaxation via a second lifting for the Max-Cut problem. Technical Report Research Report, CORR 99-55, University of Waterloo, Waterloo, Ontario, 1999. 28 pages.

[8] M.F. ANJOS and H. WOLKOWICZ. A strengthened SDP relaxation via a second lifting for the Max-Cut problem. *Special Issue of Discrete Applied Mathematics devoted to Foundations of Heuristics in Combinatorial Optimization*, to appear.

[9] K.M. ANSTREICHER. Eigenvalue bounds versus semidefinite relaxations for the quadratic assignment problem. Technical report, University of Iowa, Iowa City, IA, 1999.

[10] K.M. ANSTREICHER, X. CHEN, H. WOLKOWICZ, and Y. YUAN. Strong duality for a trust-region type relaxation of the quadratic assignment problem. *Linear Algebra Appl.*, 301(1-3):121–136, 1999.

[11] K.M. ANSTREICHER and H. WOLKOWICZ. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.*, 22(1):41–55, 2000.

[12] M. BAKONYI and C. R. JOHNSON. The Euclidean distance matrix completion problem. *SIAM J. Matrix Anal. Appl.*, 16(2):646–654, 1995.

[13] M. BAKONYI and G. NAEVDAL. On the matrix completion method for multidimensional moment problems. *Acta Sci. Math. (Szeged)*, 64(3-4):547–558, 1998.

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

[15] W.W. BARRETT, C.R. JOHNSON, and M. LUNDQUIST. Determinantal formulae for matrix completions associated with chordal graphs. *Linear Algebra Appl.*, 121:265–289, 1989. Linear algebra and applications (Valencia, 1987).

[16] A. BARVINOK. On convex properties of the quadratic image of the sphere. Technical report, University of Michigan, 1999.

[17] S. J. BENSON, Y. YE, and X. ZHANG. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM J. Optim.*, 10(2):443–461 (electronic), 2000.

[18] R. BHATIA. *Perturbation Bounds for Matrix Eigenvalues : Pitman Research Notes in Mathematics Series 162.* Longman, New York, 1987.

[19] B. BORCHERS. CSDP, a C library for semidefinite programming. *Optim. Methods Softw.*, 11/12(1-4):613–623, 1999. Interior point methods.

[20] B. BORCHERS. SDPLIB 1.2, a library of semidefinite programming test problems. *Optim. Methods Softw.*, 11(1):683–690, 1999. Interior point methods.

[21] S. BURER and R.D.C. MONTEIRO. An efficient algorithm for solving the MAXCUT SDP relaxation. Technical report, Georgia Tech, Atlanta, GA, 1999.

[22] C.C. CHOI and Y. YE. Application of semidefinite programming to circuit partitioning. Technical report, The University of Iowa, Iowa City, Iowa, 1999.

[23] N. COHEN and J. DANCIS. Maximal rank Hermitian completions of partially specified Hermitian matrices. *Linear Algebra Appl.*, 244:265–276, 1996.

[24] G.M. CRIPPEN and T.F. HAVEL. *Distance Geometry and Molecular Conformation.* Wiley, New York, 1988.

[25] J. DANCIS. Positive semidefinite completions of partial Hermitian matrices. *Linear Algebra Appl.*, 175:97–114, 1992.

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

[27] A. EDELMAN, T. ARIAS, and S.T. SMITH. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353 (electronic), 1999.

[28] S.M. FALLAT, C.R. JOHNSON, and R.L. SMITH. The general totally positive matrix completion problem with few unspecified entries. *Electron. J. Linear Algebra*, 7:1–20 (electronic), 2000.

[29] P.A. FILLMORE and J.P. WILLIAMS. Some convexity theorems for matrices. *Glasgow Math. J.*, 10:110–117, 1971.

[30] F. FORGÓ. *Nonconvex Programming.* Akadémiai Kiadó, Budapest, 1988.

[31] T. FUJIE and M. KOJIMA. Semidefinite programming relaxation for nonconvex quadratic programs. *J. Global Optim.*, 10(4):367–380, 1997.

[32] K. FUJISAWA, M. KOJIMA, and K. NAKATA. SDPA semidefinite programming algorithm. Technical report, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1995.

[33] K. FUJISAWA, M. KOJIMA, and K. NAKATA. Exploiting sparsity in primal-dual interior-point methods for semidefinite programming. *Math. Programming*, 79, 1997.

[34] K. FUKUDA, M. KOJIMA, K. MUROTA, and K. NAKATA. Exploiting sparsity in semidefinite programming via matrix completion I: general framework. Technical Report B-358, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1999.

[35] J.F. GEELEN. Maximum rank matrix completion. *Linear Algebra Appl.*, 288(1-3):211–217, 1999.

[36] W. GLUNT, T.L. HAYDEN, C.R. JOHNSON, and P. TARAZAGA. Positive definite completions and determinant maximization. *Linear Algebra Appl.*, 288(1-3):1–10, 1999.

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

[38] M.X. GOEMANS. Semidefinite programming and combinatorial optimization. *Documenta Mathematica*, Extra Volume ICM 1998:657–666, 1998. Invited talk at the International Congress of Mathematicians, Berlin, 1998.

[39] M.X. GOEMANS and F. RENDL. Combinatorial optimization. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications*. Kluwer Academic Publishers, Boston, MA, 2000.

[40] 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.

[41] I. GOHBERG, M.A. KAASHOEK, and F. van SCHAGEN. *Partially specified matrices and operators: classification, completion, applications*. Birkhäuser Verlag, Basel, 1995.

[42] B. GRONE, C.R. JOHNSON, E. MARQUES de SA, and H. WOLKOWICZ. Positive definite completions of partial Hermitian matrices. *Linear Algebra Appl.*, 58:109–124, 1984.

[43] S.W. HADLEY, F. RENDL, and H. WOLKOWICZ. A new lower bound via projection for the quadratic assignment problem. *Math. Oper. Res.*, 17(3):727–739, 1992.

[44] W. HAEMERS. On some problems of Lovász concerning the Shannon capacity of graphs. *IEEE Transactions on Information Theory*, 25:231–232, 1979.

[45] T.L. HAYDEN, R. REAMS, and J. WELLS. Methods for constructing distance matrices and the inverse eigenvalue problem. *Linear Algebra Appl.*, 295(1-3):97–112, 1999.

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

[47] C. HELMBERG and F. OUSTRY. Bundle methods to minimize the maximum eigenvalue function. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications.* Kluwer Academic Publishers, Boston, MA, 2000.

[48] C. HELMBERG and F. RENDL. A spectral bundle method for semidefinite programming. *SIAM J. Optim.*, 10(3):673 – 696, 2000.

[49] C. HELMBERG, F. RENDL, R. J. VANDERBEI, and H. WOLKOWICZ. An interior-point method for semidefinite programming. *SIAM J. Optim.*, 6(2):342–361, 1996.

[50] B. HENDRICKSON. *The molecule problem: Determining conformation from pairwise distances.* PhD thesis, Cornell University, Ithaca, New York, 1991.

[51] B. HENDRICKSON. The molecule problem: exploiting structure in global optimization. *SIAM J. Optim.*, 5(4):835–857, 1995.

[52] L. HOGBEN. Completions of inverse M-matrix patterns. *Linear Algebra Appl.*, 282(1-3):145–160, 1998.

[53] L. HOGBEN. Completions of M-matrix patterns. *Linear Algebra Appl.*, 285(1-3):143–152, 1998.

[54] S. HOMER and M. PEINADO. n the performance of polynomial-time clique algorithms on very large graphs. Technical report, Boston University, Boston, MA, 1994.

[55] R.A. HORN and C.R. JOHNSON. *Matrix analysis.* Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[56] C.R. JOHNSON. Matrix completion problems: a survey. In *Matrix theory and applications (Phoenix, AZ, 1989)*, pages 171–198. Amer. Math. Soc., Providence, RI, 1990.

[57] C.R. JOHNSON, B. KROSCHEL, and H. WOLKOWICZ. An interior-point method for approximate positive semidefinite completions. *Comput. Optim. Appl.*, 9(2):175–190, 1998.

[58] C.R. JOHNSON and B.K. KROSCHEL. The combinatorially symmetric P-matrix completion problem. *Electron. J. Linear Algebra*, 1:59–63 (electronic), 1996.

[59] C.R. JOHNSON and L. RODMAN. Inertia possibilities for completions of partial Hermitian matrices. *Linear and Multilinear Algebra*, 16(1-4):179–195, 1984.

[60] C.R. JOHNSON and L. RODMAN. Chordal inheritance principles and positive definite completions of partial matrices over function rings. In *Contributions to operator theory and its applications (Mesa, AZ, 1987)*, pages 107–127. Birkhäuser, Basel, 1988.

[61] C.R. JOHNSON and R.L. SMITH. The symmetric inverse M-matrix completion problem. *Linear Algebra Appl.*, 290(1-3):193–212, 1999.

[62] C.R. JOHNSON and P. TARAZAGA. Connections between the real positive semidefinite and distance matrix completion problems. *Linear Algebra Appl.*, 223/224:375–391, 1995. Special issue honoring Miroslav Fiedler and Vlastimil Pták.

[63] C.R. JOHNSON and P. TARAZAGA. Approximate semidefinite matrices in a subspace. *SIAM J. Matrix Anal. Appl.*, 19(4):861–871, 1999.

[64] S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Trust regions and relaxations for the quadratic assignment problem. In *Quadratic assignment and related problems (New Brunswick, NJ, 1993)*, pages 199–219. Amer. Math. Soc., Providence, RI, 1994.

[65] D.E. KNUTH. The sandwich theorem. *Electronic J. Combinatorics*, 1:48pp, 1994.

[66] M. KOJIMA, S. KOJIMA, and S. HARA. Linear algebra for semidefinite programming. Technical Report 1004, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1997. Linear matrix inequalities and positive semidefinite programming (Japanese) (Kyoto, 1996).

[67] M. KOJIMA and L. TUNÇEL. Discretization and localization in successive convex relaxation methods for nonconvex quadratic optimization problems. Technical Report CORR98-34, Dept. of Combinatorics and Optimization, University of Waterloo, 1998.

[68] M. KOJIMA and L. TUNÇEL. Cones of matrices and successive convex relaxations of nonconvex sets. *SIAM J. Optim.*, 1999. To appear.

[69] M. KOJIMA and L. TUNÇEL. On the finite convergence of successive convex relaxation methods. Technical Report CORR99-36, Dept. of Combinatorics and Optimization, University of Waterloo, 1999.

[70] F. KÖRNER. A tight bound for the Boolean quadratic optimization problem and its use in a branch and bound algorithm. *Optimization*, 19(5):711–721, 1988.

[71] F. KÖRNER. Remarks on a difficult test problem for quadratic Boolean programming. *Optimization*, 26:355–357, 1992.

[72] S. KRUK. Semidefinite programming applied to nonlinear programming. Master's thesis, University of Waterloo, 1996.

[73] S. KRUK, M. MURAMATSU, F. RENDL, R.J. VANDERBEI, and H. WOLKOWICZ. The Gauss-Newton direction in linear and semidefinite programming. *Optimization Methods and Software*, 15(1):1–27, 2001.

[74] S. KRUK and H. WOLKOWICZ. $\mathrm{sq}^2\mathrm{p}$, sequential quadratic constrained quadratic programming. In *Advances in nonlinear programming (Beijing, 1996)*, pages 177–204, Dordrecht, 1998. Kluwer Acad. Publ.

[75] J.B. LASSERRE. Optimality conditions and lmi relaxations for 0-1 programs. Laas research report, LAAS-CNRS, Toulouse, France, 2000.

[76] M. LAURENT. Cuts, matrix completions and graph rigidity. *Math. Programming*, 79:255–284, 1997.

[77] M. LAURENT. A connection between positive semidefinite and Euclidean distance matrix completion problems. *Linear Algebra Appl.*, 273:9–22, 1998.

[78] M. LAURENT. A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems. In *Topics in Semidefinite and Interior-Point Methods*, volume 18 of *The Fields Institute for Research in Mathematical Sciences, Communications Series*, pages 51–76, Providence, Rhode Island, 1998. American Mathematical Society.

[79] M. LAURENT. Polynomial instances of the positive semidefinite and Euclidean distance matrix completion problems. *SIAM J. Matrix Anal. Appl.*, to appear, 2000.

[80] C. LEMARÉCHAL and F. OUSTRY. Semidefinite relaxations and Lagrangian duality with application to combinatorial optimization. Technical report, Institut National de Recherche en Informatique et en Automatique, INRIA, St Martin, France, 1999.

[81] M. LITTMAN, D.F. SWAYNE, N. DEAN, and A. BUJA. Visualizing the embedding of objects in Euclidean space. Technical report, Bellcore, Morristown, NJ 07962-1910, 1999.

[82] L. LOVÁSZ and A. SCHRIJVER. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1(2):166–190, 1991.

[83] D.G. LUENBERGER. *Optimization by Vector Space Methods.* John Wiley, 1969.

[84] Z-Q. LUO and S. ZHANG. On the extension of Frank-Wolfe theorem. Technical report, Erasmus University Rotterdam, The Netherlands, 1997.

[85] R. MATHIAS. Matrix completions, norms and Hadamard products. *Proc. Amer. Math. Soc.*, 117(4):905–918, 1993.

[86] S. MIZUNO, M.J. TODD, and Y. YE. On adaptive-step primal-dual interior-point algorithms for linear programming. *Math. Oper. Res.*, 18(4):964–981, 1993.

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

[88] R.D.C. MONTEIRO. Primal-dual path-following algorithms for semidefinite programming. *SIAM J. Optim.*, 7(3):663–678, 1997.

[89] R.D.C. MONTEIRO and P.R. ZANJÁCOMO. Implementation of primal-dual methods for semidefinite programming based on Monteiro and Tsuchiya Newton directions and their variants. Technical report, Georgia Tech, Atlanta, GA, 1997.

[90] J.J. MORÉ and Z. WU. Distance geometry optimization for protein structures. *J. Global Optim.*, 15(3):219–234, 1999.

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

[92] Y.E. NESTEROV and A.S. NEMIROVSKI. *Optimization over positive semidefinite matrices: Mathematical background and user's manual.* USSR Acad. Sci. Centr. Econ. & Math. Inst., 32 Krasikova St., Moscow 117418 USSR, 1990.

[93] Y.E. NESTEROV and A.S. NEMIROVSKI. *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Publications. SIAM, Philadelphia, USA, 1994.

[94] Y.E. NESTEROV and M.J. TODD. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22(1):1–42, 1997.

[95] Y.E. NESTEROV, H. WOLKOWICZ, and Y. YE. Semidefinite programming relaxations of nonconvex quadratic optimization. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications*, page 34. Kluwer Academic Publishers, Boston, MA, 2000.

[96] M. NEWMAN. Matrix completion theorems. *Proc. Amer. Math. Soc.*, 94(1):39–45, 1985.

[97] M.L. OVERTON and R.S. WOMERSLEY. Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices. *Math. Programming*, 62(2, Ser. B):321–357, 1993.

[98] P. PARDALOS, F. RENDL, and H. WOLKOWICZ. The quadratic assignment problem: a survey and recent developments. In P. Pardalos and H. Wolkowicz, editors, *Quadratic assignment and related problems (New Brunswick, NJ, 1993)*, pages 1–42. Amer. Math. Soc., Providence, RI, 1994.

[99] P.M. PARDALOS and S.A. VAVASIS. Quadratic programming with one negative eigenvalue is NP-hard. *J. Global Optim.*, 1(1):15–22, 1991.

[100] B.N. PARLETT. *The symmetric eigenvalue problem.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Corrected reprint of the 1980 original.

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

[102] S. POLJAK and H. WOLKOWICZ. Convex relaxations of $(0,1)$-quadratic programming. *Math. Oper. Res.*, 20(3):550–561, 1995.

[103] A.J. QUIST, E. DE KLERK, C. ROOS, and T. TERLAKY. Copositive relaxation for general quadratic programming. *Optim. Methods Softw.*, 9:185–208, 1998. Special Issue Celebrating the 60th Birthday of Professor Naum Shor.

[104] M.V. RAMANA. *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems.* PhD thesis, Johns Hopkins University, Baltimore, Md, 1993.

[105] M.V. RAMANA, L. TUNÇEL, and H. WOLKOWICZ. Strong duality for semidefinite programming. *SIAM J. Optim.*, 7(3):641–662, 1997.

[106] F. RENDL. Semidefinite programming and combinatorial optimization. *Appl. Numer. Math.*, 29:255–281, 1999.

[107] F. RENDL and H. WOLKOWICZ. A projection technique for partitioning the nodes of a graph. *Ann. Oper. Res.*, 58:155–179, 1995. Applied mathematical programming and modeling, II (APMOD 93) (Budapest, 1993).

[108] F. RENDL and H. WOLKOWICZ. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Math. Programming*, 77(2, Ser. B):273–299, 1997.

[109] A. SCHRIJVER. A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Infor. Theory*, IT-25:425–429, 1979.

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

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

[112] J.F. STURM. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Technical report, Communications Research Laboratory, McMaster University, Hamilton, Canada, 1998. To appear in *Optimization Methods and Software*.

[113] M.J. TODD. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Optim. Methods Softw.*, 11&12:1–46, 1999.

[114] M.J. TODD, K.C. TOH, and R.H. TUTUNCU. A MATLAB software package for semidefinite programming. Technical report, School of OR and IE, Cornell University, Ithaca, NY, 1996.

[115] M.W. TROSSET. Applications of multidimensional scaling to molecular conformation. *Computing Science and Statistics*, 29:148–152, 1998.

[116] L. TUNÇEL and S. XU. Complexity analyses of discretized successive convex relaxation methods. Technical Report CORR 99-37, Department of Combinatorics and Optimization, Waterloo, Ont, 1999.

[117] L. VANDENBERGHE, S. BOYD, and S.-P. WU. Determinant maximization with linear matrix inequality constraints. *SIAM J. Matrix Anal. Appl.*, 19(2):499–533, 1998.

[118] C.P. WELLS. *An Improved Method for Sampling of Molecular Conformation Space.* PhD thesis, University of Kentucky, 1995.

[119] H. WOLKOWICZ, R. SAIGAL, and L. VANDENBERGHE, editors. *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications.* Kluwer Academic Publishers, Boston, MA, 2000. xxvi+654 pages.

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

[121] Y. YE. A new complexity result on minimization of a quadratic function with a sphere constraint. In *Recent Advances in Global Optimization*, pages 19–31. Princeton University Press, 1992.

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

[123] 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. Semidefinite programming and interior-point approaches for combinatorial optimization problems (Toronto, ON, 1996).