

SEMIDEFINITE AND LAGRANGIAN RELAXATIONS FOR HARD COMBINATORIAL PROBLEMS*

Henry Wolkowicz

Department of Combinatorics and Optimization

University of Waterloo

Waterloo, Ontario, Canada

hwolkowi@orion.math.uwaterloo.ca

Abstract Semidefinite Programming is currently a very exciting and active area of research. Semidefinite relaxations generally provide very tight bounds for many classes of numerically hard problems. In addition, these relaxations can be solved efficiently by interior-point methods.

In this paper we study these semidefinite relaxations using the equivalent Lagrangian relaxations. In particular, the theme of the paper is to show that the Lagrangian relaxation is, in some respects, best. In all instances we consider, we show that whenever we have a tractable bound (relaxation), then the same bound can be obtained from a Lagrangian relaxation.

Keywords: Semidefinite Programming, Lagrangian Duality, Relaxations, Quadratic Constrained Quadratic Programs, Hard Combinatorial Problems.

1. INTRODUCTION

Semidefinite Programming (denoted SDP and sometimes called linear matrix inequalities, LMIs) is a generalization of linear programming (denoted LP), where the nonnegativity constraints on vector variables are replaced by positive semidefinite constraints on symmetric matrix variables. These problems have an old history dating back more than 100 years to Lyapunov's theory for stability of differential equations, e.g. [82, 71]. They were studied and applied in engineering applications as early as the 1960s, e.g. [114, 116, 115] and continued into the 1980's (see e.g. the historical outline in [20] and the work on matrix completion

*Research supported by The Natural Sciences and Engineering Research Council of Canada.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35514-6_15](https://doi.org/10.1007/978-0-387-35514-6_15)

problems in [27, 37, 60]). In addition, SDP is a special case of *optimization over cone constraints* (generalized linear programming), which dates back more than 30 years to e.g. Bellman and Fan [11], and was an ongoing active area of research, e.g. [12, 41, 23, 24, 55, 122, 79, 19].

The last ten years have seen an enormous interest in the SDP area, due to many new and important applications in, e.g. combinatorial optimization, engineering (systems and control), statistics, etc. This interest increased greatly due to the fact that SDP problems can be solved efficiently (are tractable) by the new interior-point methods, e.g. [77]. One of the interesting side effects of the activity is that it has brought various different areas of research into contact. For example, the people working on numerical issues for large scale problems are now using sophisticated techniques in convex analysis, and using adjoint operators rather than matrix representations is becoming common. We are all benefiting from the elegance and applicability of this area.

Combinatorial and discrete optimization problems often involve binary (0, 1 or ± 1) decision variables. These can be modelled using quadratic constraints $x^2 - x = 0$, or $x^2 = 1$, respectively. Thus many hard combinatorial problems can be modelled using quadratically constrained quadratic programs, denoted Q²P. However, these latter problems can be just as hard (intractable) to solve. Therefore, relaxations are used to find approximate bounds and solutions. One can use linear approximations and obtain models that can be solved efficiently (tractable models). Moreover, using the positive semidefinite matrix construction $X = xx^T$, we can *lift* the problem into matrix space and obtain a *Semidefinite Programming Relaxation* by ignoring the rank one restriction on X , see e.g. [65, 38, 66, 102, 8, 7, 78]. This lifting process provides surprisingly stronger bounds, both empirically and theoretically, than have previously been found, e.g. [35, 47, 3]. Thus SDP provides a means of finding an approximate solution to quadratic models for hard problems.

In this paper we study SDP relaxations as well as exploring the relationship between the SDP and Lagrangian relaxations of various classes of Q²Ps. (This continues on the work in e.g. [103, 91].) We then consider the strength of these relaxations, in the sense of strong duality. We see that in the simplest case of one constraint (the so-called trust region subproblem, TRS), strong duality holds. However, even two convex constraints (the CDT problem) can result in a duality gap. Therefore, it is surprising that there is a class of matrix problems with orthogonal constraints for which there is a zero duality gap. This motivates adding certain nonlinear redundant constraints in order to derive a strengthened SDP relaxation for the max-cut problem.

Throughout this paper we emphasize the theme (or conjecture) that Lagrangian relaxation is somehow best. Though this question is very vague, so perhaps an answer may not be available, it does give the flavour of the approach used in the paper.

The paper is organized as follows. We begin in Section 2.1 with a well known problem in this area, the Max-Cut problem. We present several different relaxations. Following our theme, all these bounds, including the SDP bound, end up being equivalent to the Lagrangian relaxation. We then discuss the TRS and the CDT problem and the difference in strong duality for them. This is followed by finding the SDP relaxation for general Q²P in Section 2.2. It includes descriptions of the relationships between the SDP relaxation and the Lagrangian relaxation via convex quadratic valid inequalities, following [33, 52]. Occurrences of strong duality for nonconvex quadratic programs are studied in Section 2.3. In every instance where one has a tractable bound, we find a Q²P such that the bound is attained by the Lagrangian relaxation. This follows the work in [6, 5].

We conclude with a strengthened SDP bound based on a second lifting procedure. This illustrates a *recipe* for constructing semidefinite relaxations using the Lagrangian dual of the Lagrangian dual of the original Q²P. In addition, we see the advantages of this approach as redundant constraints added at the start provide strengthened bounds, but do not result in redundancy in the final SDP relaxation.

1.1. LAGRANGE MULTIPLIERS FOR Q²P

The Q²P in x is

$$\begin{aligned}
 (\text{Q}^2\text{P}_x) \quad \mu^* := \min_{x \in \mathbb{R}^n} \quad & q_0(x) := x^T Q_0 x + 2g_0^T x + \alpha_0 \\
 \text{s.t.} \quad & q_k(x) := x^T Q_k x + 2g_k^T x + \alpha_k \leq 0, \\
 & k \in \mathcal{I} := \{1, \dots, m\},
 \end{aligned}$$

where $Q_k \in \mathcal{S}^n$, the space of symmetric matrices. The Lagrangian of Q²P _{x} is

$$L(x, \lambda) := q_0(x) + \sum_{k \in \mathcal{I}} \lambda_k q_k(x),$$

where $\lambda = (\lambda_k) \geq 0$ are nonnegative Lagrange multipliers.

Lagrange multipliers can be used in two ways. First, if a constraint qualification holds for Q²P _{x} at the optimum x (e.g. the Mangasarian-Fromovitz constraint qualification), then the Karush-Kuhn-Tucker necessary conditions for optimality are satisfied, i.e.

$$\nabla L(x, \lambda) = 0 \quad \text{and} \quad \lambda_k q_k(x) = 0, \quad \forall k \in \mathcal{I},$$

for some $0 \leq \lambda \in \mathfrak{R}^m$. Therefore the search for the optimum x can be restricted to the points satisfying stationarity of the Lagrangian and complementary slackness. Moreover, if the Lagrangian is also convex, then these (and primal feasibility) are sufficient conditions for optimality.

Lagrange multipliers can also be used to derive the Lagrangian dual (or relaxation) of Q^2P_x

$$(\text{DQ}^2\text{P}_x) \quad \mu^* \geq \nu^* := \max_{\lambda \geq 0} \min_x q_0(x) + \sum_{k \in \mathcal{I}} \lambda_k q_k(x),$$

i.e. each inner unconstrained minimization problem provides a lower bound for Q^2P_x ; and we then choose the best of these lower bounds. A zero duality gap holds if $\mu^* = \nu^*$, but this condition can fail in the nonconvex case. Strong duality holds if $\mu^* = \nu^*$ and also ν^* is attained.

Remark 1.1 *Unfortunately, the term strong duality is ambiguous in the literature as it is sometimes used to define a zero duality gap with both primal and dual attainment.*

Remark 1.2 *Let $q := (q_k)$. Then the sum in the Lagrangian can be rewritten as $\langle \lambda, q(x) \rangle = \lambda^T q(x)$. This appears to be too trivial to mention. However, it does emphasize how Lagrange multipliers arise when one is dealing with matrix valued constraints. In particular, if one has a constraint $Q(x) = 0$, where the image of Q is a symmetric matrix, then the Lagrange multiplier will be a symmetric matrix, say $\Lambda = \Lambda^T$, and the term in the Lagrangian will be the inner product $\langle Q(x), \Lambda \rangle$. With a nonnegativity constraint, there will be a sign restriction on the Lagrange multiplier.*

1.2. SEMIDEFINITE PROGRAMMING PRELIMINARIES

SDP is an extension of LP in that matrix variables replace vector variables and nonnegativity elementwise is replaced by positive semidefiniteness. In addition, it is a special case of the cone programming problem

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) \succeq_K 0, \end{aligned} \tag{1}$$

where K is a convex cone ($K+K \subset K$, $\alpha K \subset K$, $\forall \alpha \geq 0$) and $g(x) \succeq_K 0$ denotes the cone partial order $g(x) \in K$. If $K = \mathfrak{R}_+^n$, then we get the usual elementwise ordering $g(x) \geq 0$. This notation allows comparisons with many results that hold for the elementwise ordering, e.g. Jensen's inequality for the definition of convexity. In the case that $K = \mathcal{P}$, the cone of positive semidefinite matrices, then we have the *Löwner partial*

order, e.g. [69]. In fact, this is a very general mathematical program and includes standard inequality, equality, and semidefinite constraints, when $K = \mathfrak{R}_+^n \otimes \{0\} \otimes \mathcal{P}$, where $\{0\}$ is the set containing the 0 vector. We now look at some of the relations between LP and SDP.

Geometry. Much of the elegant geometry of polyhedral sets developed for LP can be extended to SDP. This was studied as early as 1948 by Bohnenblust [15] and later by Tausky [108] and also by Barker and Carlson [10]. More recently, motivated by the high interest in SDP, many new results on differentiability and multiplicity of eigenvalues have appeared, see e.g. Lewis [63, 64] and Pataki [83, 84], respectively. In addition, results on characterizing different types of homogeneous and self-scaled cones appear in [42, 111, 44].

Let \mathcal{P}^n (or \mathcal{P} when the meaning is clear) denote the cone of positive semidefinite matrices in \mathcal{S}^n , the space of $n \times n$ symmetric matrices. The following remarks have analogues in LP. The SDP cone is *self-polar*, i.e.

$$\mathcal{P} = \mathcal{P}^+ := \{Y : X \bullet Y \geq 0, \forall X \in \mathcal{P}\},$$

where $X \bullet Y = \text{Trace } XY$ is the trace inner product. Also, \mathcal{P} is *homogeneous*, i.e., for any $X, Y \in \text{int}(\mathcal{P})$, there exists an invertible linear operator \mathcal{A} from \mathcal{S}^n to \mathcal{S}^n that leaves \mathcal{P} invariant and that has the property $\mathcal{A}(X) = Y$.

A face of a cone K is defined as

$$\mathcal{F} = \{x \in K : y, z \in K, x = \alpha y + (1 - \alpha)z, 0 < \alpha < 1 \Rightarrow y, z \in \mathcal{F}\},$$

i.e. $x \in \mathcal{F}$ can be an interior point of the line segment that joins $y, z \in K$ only if $y, z \in \mathcal{F}$. The faces of \mathfrak{R}_+^n can be characterized using any vector in its relative interior, i.e. for y in the relative interior, we define the vector z by $z_i = 1$ if $y_i = 0$, and $z_i = 0$ if $y_i > 0$. Thus z is in the relative interior of the complementary face. Then the corresponding face that contains y in its relative interior is

$$\{x \in \mathfrak{R}_+^n : \langle x, z \rangle = 0\} = \{x \in \mathfrak{R}_+^n : y_i = 0 \Rightarrow x_i = 0\}.$$

The faces \mathcal{F} of \mathcal{P} can be similarly characterized using any $Y \in \text{relint } \mathcal{F}$, i.e.

$$\mathcal{F} = \{X : \mathcal{N}(X) \supset \mathcal{N}(Y)\},$$

where $\mathcal{N}(\cdot)$ denotes nullspace. The faces of \mathcal{P} are *exposed*, i.e. $\mathcal{F} = \mathcal{P} \cap \varphi^\perp$, where $\varphi \in \mathcal{P} \cap \mathcal{F}^\perp$ is a *conjugate face*. (Here the subscript \perp denotes the orthogonal complement.) Moreover, they are *projectionally exposed*, i.e. there exists a projection matrix P such that $\mathcal{F} = PPP$; note that $P \cdot P$ is a projection on \mathcal{S}^n . In fact, we can choose P (and so

$P \cdot P$) to be an orthogonal projection. See e.g. [87, 40, 39] for results on faces and [18, 9, 93, 107] for results on projectively exposed cones.

Duality and Optimality Conditions. Extensions of the optimality conditions and duality theory from LP to SDP appeared in [11]. However, unlike the LP case, strong duality theorems required a Slater-type constraint qualification (denoted CQ). This CQ is strict feasibility, i.e. there exists a feasible point in the interior of \mathcal{P} . In [95] it is shown that difficulties in the duality theory can arise due to a property of the faces of \mathcal{P} , namely that

$$\mathcal{P} + \text{span}(\mathcal{F}) \text{ is never closed.} \tag{2}$$

On the other hand, these difficulties can be corrected by taking advantage of the fact that

$$\mathcal{P} + \mathcal{F}^\perp \text{ is always closed.} \tag{3}$$

We amplify these statements in Remark 1.3 below.

Now, consider the typical primal SDP

$$(P) \quad \begin{aligned} \mu^* &:= \min C \bullet X \\ \text{s.t. } &\mathcal{A}(X) = b \text{ and } X \succeq 0, \end{aligned}$$

where $C \in \mathcal{S}^n$, $b \in \mathbb{R}^m$, and $\mathcal{A} : \mathcal{S}^n \rightarrow \mathbb{R}^m$, is a linear operator. Thus the components of the constraint equations are

$$(\mathcal{A}(X))_i = \text{Trace } A_i X = b_i, \quad i = 1, 2, \dots, m,$$

for some given symmetric matrices A_i .

We can define the dual and prove weak duality using the notion of a hidden constraint. Specifically we use the form

$$\begin{aligned} \mu^* &= \min_{X \succeq 0} \max_y C \bullet X + y^T(b - \mathcal{A}(X)) \\ &\geq \max_y \min_{X \succeq 0} y^T b + (C - \mathcal{A}^*(y)) \bullet X = \nu^*, \end{aligned} \tag{4}$$

where \mathcal{A}^* is the adjoint operator of \mathcal{A} , so it satisfies

$$\mathcal{A}(X)^T y = X \bullet \mathcal{A}^*(y), \quad \forall X, \forall y,$$

which implies $\mathcal{A}^*(y) = \sum_i y_i A_i$, the symmetric matrices A_i being defined above. The dual program is

$$(D) \quad \begin{aligned} \nu^* &:= \max b^T y \\ \text{s.t. } &\mathcal{A}^*(y) \preceq C. \end{aligned}$$

The first equation in (4) follows from the fact that the inner maximization is finite (the hidden constraint) if and only if $\mathcal{A}(X) = b$. The inequality follows from taking the minimization first. The last equation follows from the hidden constraint, because y has to be such that the inner minimization is finite valued.

The duality theory gives rise to the elegant characterization of optimality

$$\begin{array}{rcl}
 \mathcal{A}^*(y) + Z - C & = & 0 \quad \text{dual feasibility} \\
 b - \mathcal{A}(X) & = & 0 \quad \text{primal feasibility} \\
 ZX & = & 0 \quad \text{complementary slackness} \\
 Z, X & \succeq & 0 \quad \text{positive semidefiniteness.}
 \end{array} \tag{5}$$

The variables $X, (y, Z)$ are called a primal-dual optimal pair, Z being the (dual) slack variable. If only complementary slackness fails, then the variables are called a primal-dual feasible pair. Note that complementary slackness can be written in the equivalent form $\text{Trace } ZX = 0$, which follows from an orthogonal diagonalization of Z and X . This is also equivalent to having a zero duality gap between primal and dual values for a feasible pair since

$$\text{Trace } ZX = \text{Trace } (C - \mathcal{A}^*(y))X = \text{Trace } CX - y^T b.$$

Remark 1.3 *We mentioned above that the optimality conditions require a CQ, i.e. strict feasibility, and that condition (2) can cause difficulties. Consider the SDP P and its dual D . Dual feasibility can be written in the form*

$$\mathcal{A}^*(y) + Z = C, \quad Z \succeq 0.$$

If we choose the linear operator \mathcal{A}^ so that its range satisfies $\mathcal{R}(\mathcal{A}^*) = \text{span}(\mathcal{F})$, for some face \mathcal{F} of \mathcal{P} , then we can choose any C in the set $\mathcal{R}(\mathcal{A}^*) + \mathcal{P}$, but not its closure, because choices in the closure can force failure of dual feasibility and a duality gap. For example, let $A := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, and so $\mathcal{A}^*(y) = yA$. Then $C = \begin{pmatrix} \varepsilon & 1 \\ 1 & 0 \end{pmatrix}$ is admissible if and only if ε is positive. Therefore we address $C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $b = 0$. In this case, the primal problem constraint $\mathcal{A}(X) = b$ implies $X_{22} = 0$, and then the other constraint $X \succeq 0$ implies $X_{12} = X_{21} = 0$. Thus $\mu^* = C \bullet X = 0$ is optimal in P . We have noted, however, that the dual is infeasible for the given choice of C .*

The condition (3) can correct this problem by allowing a larger set of dual variables. Modified optimality conditions without any CQ can be obtained. This was done in [17] and also in [97, 95]. In our primal problem

P , we let \mathcal{F} be the minimal face, i.e. the smallest face containing the feasible set. Then the strengthened dual is $\mu^* = \max b^T y$ s.t. $A^* y \preceq_{\mathcal{F}^+} C$. For our example, the minimal face is $\mathcal{F} = \text{span} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, which implies that $C - (-1)A$ is in \mathcal{F}^+ . Hence the new dual optimal value is 0, and we have a zero duality gap.

If we perturb the complementary slackness conditions,

$$ZX = \mu I, \quad \mu > 0,$$

where μ is a parameter, then (5) becomes the (modified) optimality conditions of a log-barrier problem. These are the equations that are used in interior-point methods. However, unlike linear programming, we have an interesting subtle complication that is also discussed in [113]. One cannot apply Newton's method directly since ZX is not necessarily symmetric, and so we end up with an overdetermined system of equations. There are various ways of modifying this system in order to get good search directions, see e.g. [110, 72, 57]. Many of these directions work very well in practice, which is clear from empirical evidence and the derivation of several public domain codes, e.g. [47, 1, 105, 106, 16]. The SDP problems are tractable because they are convex programs and fall into the class of problems that can be approximately solved to a desired accuracy in polynomial time. This follows from the seminal work of Nesterov and Nemirovski [76, 77]. The algorithms that currently work well are the primal-dual interior-point algorithms. This area of research is ongoing, however, and there are many classes of problems with special structure where dual algorithms based on a bundle trust approach perform better; this is especially true if it is too expensive to evaluate the primal matrix X explicitly, see e.g. [13, 46, 61].

First and second order optimality conditions for SDP are given in [100, 101] and in the survey article [113], for example. Nondegeneracy and strict complementarity are discussed in [2, 85]. Both nondegeneracy and strict complementarity can fail, though they are generic conditions. In addition the theorem of Goldman and Tucker [36], about the existence of an optimal primal-dual pair that satisfies strict complementary slackness, does not apply to SDP. Note that strict complementarity for SDP is the condition $Z + X \succ 0$.

2. RELAXATIONS OF Q²P

We now look at a particular instance of Q²P, namely the quadratic model for the *Max-Cut problem*. We start with several different tractable relaxations for this problem that have appeared in the literature. We

show that, surprisingly, they are all equal to the Lagrangian (and SDP) relaxation.

We then consider trust region type problems and discuss when strong duality holds. We include problems where orthogonal constraints arise, e.g. orthogonal relaxations of the quadratic assignment and graph partitioning problems. Thus this part of the chapter emphasizes the theme about the strength of Lagrangian relaxation.

2.1. RELAXATIONS FOR THE MAX-CUT PROBLEM

One of the problems for which the SDP relaxation has been particularly successful, both empirically and theoretically, is the *Max-Cut Problem*, e.g. [45, 35, 34]. Let $G = (V, E)$ be an undirected graph with vertex set $V = \{v_i\}_{i=1}^n$ and weights w_{ij} on the edges $(v_i, v_j) \in E$. We seek the index set $\mathcal{I} \subset \{1, 2, \dots, n\}$, that maximizes the sum of the weights of the edges with one end point with index in \mathcal{I} and the other in the complement. This is equivalent to

$$(MC) \quad \max \frac{1}{2} \sum_{i < j} w_{ij} (1 - x_i x_j), \quad x \in \mathcal{F},$$

where $\mathcal{F} := \{\pm 1\}^n$, and $x_i = 1$ if $i \in \mathcal{I}$, and $x_i = -1$ otherwise. The objective function is a (homogeneous) quadratic form, $x^T Q x$.

Several Different Relaxations. We rewrite MC as the more general problem

$$(MCQ) \quad \mu^* := \max_{x \in \mathcal{F}} q_0(x), \quad \text{where } q_0(x) := x^T Q x - 2c^T x. \quad (6)$$

The MCQ (and MC) problem is intractable, but there are many different ways to relax the problem and find approximate solutions and/or bounds. The simplest way is to relax the constraints to the interval conditions $x \in [-1, 1]^n$. This *bound constrained quadratic problem* is NP-hard if Q is not negative semidefinite [81]; while, if Q is negative semidefinite and $c = 0$, then the solution is the trivial 0 solution.

Other relaxations are also geometric in nature and involve perturbations of the objective function. For example, one method relaxes the constraints to the unit ball of radius \sqrt{n} , while another relaxes the constraints to their convex hull, i.e. to the unit cube.

The relaxations yield bounds which are derived by making changes to the objective function q_0 that are zero when $x_i^2 = 1$, i.e. on the feasible set \mathcal{F} . In particular, for every $u \in \mathbb{R}^n$ we have

$$\begin{aligned} q_u(x) &:= x^T (Q + \text{Diag}(u)) x - 2c^T x - u^T e \\ &= q_0(x), \quad \forall x \in \mathcal{F}, \end{aligned} \quad (7)$$

where e is the vector of ones (of the appropriate dimension), and $\text{Diag}(u)$ denotes the diagonal matrix formed from the vector u . For each u we get an upper bound by ignoring the constraints and by allowing the diagonal perturbations, i.e. we have

$$\mu^* \leq f_0(u) := \max_x q_u(x). \quad (8)$$

We then find the bound

$$\mu^* \leq B_0 := \min_u f_0(u). \quad (9)$$

Let

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

Note that, if the set S is not empty, then we can minimize over the unconstrained parameter u , or we can add the restriction $u^T e = 0$, i.e.

$$B_0 = \min_{u^T e = 0} f_0(u), \text{ if } S \neq \emptyset.$$

This can be seen from the equivalence of the optimality conditions for min-max problems, e.g. [26, Pg 188, Theorem 2.1], or from a perturbation analysis of min-max problems, e.g. [29]. For example, if the solution of the inner maximization problem is attained at a unique x , then f_0 is differentiable and $\nabla f_0(u) = x \circ x - e$, where $x \circ x$ is the *Hadamard* (elementwise) product. In general, the function $f_0(u)$ is directionally differentiable at each point u (with $u^T e = 0$ in the restricted case), in any direction h . Assuming $\|h\| = 1$, the directional derivative has the value

$$\partial f_0(u; h) = \max_{x \in \mathcal{X}(u)} \left\langle \frac{\partial q_u(x)}{\partial u}, h \right\rangle = \max_{x \in \mathcal{X}(u)} \langle x \circ x - e, h \rangle,$$

where $\mathcal{X}(u)$ denotes the set of values of x that solve the inner problem (8) for given u . One can now compare stationarity for f_0 in the unconstrained case with the use of a Lagrange multiplier for the constraint $u^T e = 0$. This involves the subgradient of the convex function f_0 and its domain, i.e. the set where f_0 is finite valued. Details can be found in [92]. The comment about the possible use of $u^T e = 0$ is also true for the bounds that follow.

But the function f_0 can take on the value $+\infty$. We can avoid these infinite values by restricting the parameter u , using a hidden semidefinite constraint. Specifically, this constraint depends on the remark that a quadratic function is bounded above if and only if its Hessian is negative semidefinite and its stationarity equation is consistent. Thus we

obtain the following bound, which is tractable since we minimize a convex function over a convex set:

$$\mu^* \leq B_0 = \min_{Q + \text{Diag}(u) \leq 0} f_0(u). \tag{10}$$

Next we relax the conditions on x to the sphere of radius \sqrt{n} , which provides

$$\mu^* \leq f_1(u) := \max_{\|x\|^2=n} q_u(x). \tag{11}$$

Hence our next bound is

$$\mu^* \leq B_1 := \min_u f_1(u). \tag{12}$$

The inner maximization problem is called a trust region subproblem and is tractable, as shown in Section 2.3 below. Thus we have our second tractable bound.

We can replace the spherical constraint with a box constraint, which gives

$$\mu^* \leq f_2(u) := \max_{|x_i| \leq 1} q_u(x). \tag{13}$$

After adding the semidefinite constraint on u to make the calculation of f_2 tractable, we obtain our next bounds

$$\mu^* \leq \min_u f_2(u) \tag{14}$$

and

$$\mu^* \leq B_2 := \min_{Q + \text{Diag}(u) \leq 0} f_2(u). \tag{15}$$

Given Q and c of the function (7), we define the $(n + 1) \times (n + 1)$ matrix Q^c by adding a leading row and column, so Q^c has the elements

$$\begin{aligned} q_{00}^c &= 0, & q_{0i}^c &= q_{i0}^c = -c_i \text{ for } i > 0, \text{ and} \\ q_{ij}^c &= q_{ij} \text{ for } i, j > 0, \end{aligned}$$

i.e.

$$Q^c := \begin{bmatrix} 0 & -c^T \\ -c & Q \end{bmatrix}. \tag{16}$$

Further, in order to have functions $q_u^c(y)$ and $f_i(u)$ that are analogous to the previous cases, we introduce

$$q_u^c(y) := y^T(Q^c + \text{Diag}(u))y - u^T e, \quad y \in \mathfrak{R}^{n+1}, \tag{17}$$

where u and e are also in \Re^{n+1} . Note that q_u^c reduces to q_u if the first component of y is $y_0 = 1$. The equivalent relaxed problem is

$$\mu^* \leq f_1^c(u) := \max_{\|y\|^2=n+1} q_u^c(y) = (n+1) \lambda_{\max}(Q^c + \text{Diag}(u)) - u^T e, \quad (18)$$

where $\lambda_{\max}(A)$ denotes the maximum eigenvalue of A , say. Thus another bound on μ^* is

$$\mu^* \leq B_1^c := \min_u f_1^c(u). \quad (19)$$

Similarly, we get equivalent bounds B_0^c and homogenized bounds for the other models.

The above argument shows that we can homogenize the problem by moving into a higher dimension. Therefore we can assume the special case $c = 0$. We now look at the SDP bound. The relaxation comes from the fact that the trace has the commutative property

$$x^T Q x = \text{Trace } x^T Q x = \text{Trace } Q x x^T,$$

and, for $x \in \mathcal{F}$, the matrix Y with elements $y_{ij} = x_i x_j$ is symmetric, rank one and positive semidefinite, with ones on the diagonal. Therefore we can *lift* the problem into the higher dimensional space of symmetric matrices and relax the rank one constraint. Thus we obtain a relaxation that gives the bound

$$\text{(MCSDP)} \quad B_3 := \max_{Y \in \mathcal{S}^n} \text{Trace } QY \quad (20)$$

s.t. $\text{diag}(Y) = e$ and $Y \succeq 0$,

where $\text{diag}(Y)$ is the vector formed from the diagonal of Y . This SDP is a convex programming problem and is tractable.

We have presented several different tractable bounds that have simple geometric interpretations. It is not at all clear which bounds are better or how to compare them. We now do something that may seem meaningless; we replace the ± 1 constraints with $x_i^2 = 1, i = 0, \dots, n$, which does not change the feasible set of the original problem. In [92, 91] it is shown, however, that all the above relaxations and bounds for MC come from the Lagrangian dual of the following equivalent problem to MCQ:

$$\text{(PE)} \quad \max \quad q_0(x) = x^T Q x - 2c^T x \quad (21)$$

s.t. $x_i^2 = 1, i = 1, \dots, n$.

Thus we enforce our theme about the strength of Lagrangian relaxation. The strong duality result for the trust region subproblem is the key to the proofs. Note that the Lagrangian dual of P_E yields precisely our first bound B_0 on μ^* , given in (9).

Theorem 2.1 *All the bounds for MCQ discussed above are equal to the optimal value of the Lagrangian dual of the equivalent program P_E . ■*

2.2. GENERAL Q²P

We now move on to applying the Lagrangian relaxation to *general* quadratic constrained quadratic problems, denoted Q²P. The general Q²P problem is also studied in [33, 52, 96, 58, 56, 68, 59], for example.

Quadratic bounds using a Lagrangian relaxation have received much attention and been applied in the literature, for example in [53] and, more recently, in [54]. The latter calls the Lagrangian relaxation the “best convex bound”. Discussions on Lagrangian relaxation for nonconvex programs also appear in [31]. More references are given throughout this paper.

Remark 2.1 *Any equality constraints are written as two inequality constraints; any linear equality constraints, $Ax = b$, are transformed to the quadratic constraint $\|Ax - b\|^2 = 0$. The reason for these transformations for linear equality constraints is discussed in [91]. It is that the Lagrangian dual essentially ignores linear constraints, as can be seen from: $-\infty = \max_\lambda \min_x -x^2 + \lambda x$, which is the dual of the problem $\min\{-x^2 : x = 0\}$.*

We now recall from Section 1.1 the Q²P in x :

$$\begin{aligned} \mu^* := \min_{x \in \mathbb{R}^n} \quad & q_0(x) := x^T Q_0 x + 2g_0^T x + \alpha_0 \\ (\text{Q}^2\text{P}_x) \quad & \text{s.t.} \quad q_k(x) := x^T Q_k x + 2g_k^T x + \alpha_k \leq 0, \\ & k \in \mathcal{I} := \{1, \dots, m\}, \end{aligned} \quad (22)$$

where the matrices Q_k are symmetric. The *feasible set* is

$$\mathcal{F} := \{x \in \mathbb{R}^n : q_k(x) \leq 0, \forall k \in \mathcal{I}\}.$$

Note that, though the feasible set \mathcal{F} may be empty, the feasible set of the relaxation may not be. The objective function and the constraints need not be convex. Therefore the feasible set can have “nasty” features that cause the problem to be very hard to solve in general, see e.g. [81].

Let

$$P_k := \begin{bmatrix} \alpha_k & g_k^T \\ g_k & Q_k \end{bmatrix}, \quad (23)$$

and, by abuse of notation, define

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

Then an equivalent formulation of Q^2P_x is the homogenized problem

$$(Q^2P_y) \quad \begin{aligned} \mu^* &= \min q_0(y) \\ \text{s.t.} \quad & q_k(y) \leq 0, \quad k \in \mathcal{I} \\ & y_0^2 = 1, \quad y = \begin{pmatrix} y_0 \\ x \end{pmatrix} \in \mathfrak{R}^{n+1}. \end{aligned}$$

We see that the optimal values of Q^2P_x and Q^2P_y are equal. Further, if $y_0 = -1$ is optimal, then we can replace y by $-y$, because the objective function and all but the last constraint are homogeneous.

We will refer to both equivalent formulations of Q^2P in the sequel. The relevant one will be clear from the context.

Remark 2.2 *Note that we could replace the constraint $y_0^2 = 1$ by $y_0 = 1$ as in [33]. In the $y_0 = 1$ case, the feasible sets of the two formulations coincide exactly, while in the former case they can differ by a sign. Specifically, $x \in \mathcal{F}$ implies that both $\begin{pmatrix} -1 \\ -x \end{pmatrix}$ and $\begin{pmatrix} 1 \\ x \end{pmatrix}$ are in \mathcal{F}_y , where \mathcal{F}_y is the feasible region of the homogenized problem Q^2P_y .*

The Lagrangian Relaxation of a General Q^2P . The Lagrangian relaxation of the homogenized problem Q^2P_y provides a simple technique for obtaining the SDP relaxation. In addition, an application of the strong duality result for the trust region subproblem shows that the SDP and Lagrangian relaxation are equal. The problem Q^2P_y has the Lagrangian

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

and the Lagrangian relaxation of Q^2P_y is

$$(DQ^2P_y) \quad \nu^* := \max_{\lambda \geq 0, \sigma} \min_y y^T P_0 y - \sigma(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y.$$

Note that

$$\begin{aligned} \nu^* &= \max_{\lambda \geq 0} \max_{\sigma} \min_y y^T P_0 y - \sigma(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y \\ &= \max_{\lambda \geq 0} \min_{y: y_0^2=1} y^T P_0 y + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y, \end{aligned}$$

from strong duality of the trust region subproblem [104]. Therefore, we get equivalence of the dual values for the problems DQ^2P_y and

$$(DQ^2P_x) \quad \nu^* = \max_{\lambda \geq 0} \min_x q_0(x) + \sum_{k \in \mathcal{I}} \lambda_k q_k(x),$$

which is similar to the approaches in [116, 103]. It follows that *weak duality*

$$\nu^* \leq \mu^* = \min_y \max_{\lambda \geq 0, \sigma} y^T P_0 y - \sigma(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k y^T P_k y$$

holds. Therefore, if the optimal σ^* and λ^* can be found, we have derived a single quadratic function whose minimal value approximates the original minimal value μ^* of $Q^2 P_y$, i.e.

$$\mu^* \geq \nu^* = \min_y y^T P_0 y - \sigma^*(y_0^2 - 1) + \sum_{k \in \mathcal{I}} \lambda_k^* y^T P_k y. \quad (24)$$

Moreover, in the dual program $DQ^2 P_y$, the Lagrangian is a quadratic function of y . Therefore the outer maximization problem has not only nonnegativity constraints but also the hidden semidefinite constraint

$$P_0 - \sigma E_{00} + \sum_{k \in \mathcal{I}} \lambda_k P_k \succeq 0, \quad \lambda \geq 0, \quad (25)$$

where E_{00} is the zero matrix except for 1 in the top left corner. The solution of the minimization subproblem is attained at $y = 0$. Therefore the Lagrangian dual is equivalent to the SDP problem

$$\begin{aligned} \nu^* &:= \max \quad \sigma \\ \text{(DSDP)} \quad &\text{s.t.} \quad \sigma E_{00} - \sum_{k \in \mathcal{I}} \lambda_k P_k \preceq P_0 \\ &\quad \lambda \geq 0. \end{aligned}$$

Valid Inequalities. Using the above approach, we see that more constraints $q_k(y)$ give a stronger dual. Therefore the addition of redundant constraints to get new valid inequalities can strengthen the relaxation. We will see how this occurs when we look at orthogonally constrained problems below. Another approach is also specified in detail in [33] and [52, 51].

For problems that also have linear inequality constraints, one can use the notion of copositivity to strengthen the SDP relaxation. However, this does not result in a tractable relaxation in general [94].

Specific instances of these relaxations (graph partitioning and quadratic assignment problems) appear in [121, 112]. We will present a *recipe* for generating a relaxation in Section 3.

2.3. STRONG DUALITY

In the case of strong duality (zero duality gap and dual attainment), our bounds are exact. As expected, this holds (generically) in the general

convex case. Surprisingly, there are several cases of nonconvex quadratic programs where it holds as well. In this Section 2.3 we amplify on our theme that confirms the strength of Lagrangian relaxation, namely that a tractable bound implies a Lagrangian relaxation is at work.

We recall the general quadratically constrained quadratic program (22), where for simplicity we have replaced each equality constraint by two inequality constraints. We will use equality constraints when absolutely required. We let \mathcal{F} denote the feasible set.

We define the Lagrangian

$$L(x, \lambda) := q_0(x) + \sum_{k=1}^m \lambda_k q_k(x),$$

and the dual functional

$$\varphi(\lambda) := \min_x L(x, \lambda).$$

The Lagrangian is linear in λ , and so the dual functional is a concave function of λ . Thus the calculation of the maximum of this concave function is a tractable problem if $\varphi(\lambda)$ can be evaluated efficiently. For each $\lambda \geq 0$, we have the lower bound

$$\begin{aligned} \mu^* &= \min_{x \in \mathcal{F}} q_0(x) \geq \min_{x \in \mathcal{F}} L(x, \lambda) \\ &\geq \min_x L(x, \lambda) = \varphi(\lambda). \end{aligned}$$

Thus we deduce the dual problem

$$\mu^* \geq \nu^* = \max_{\lambda \geq 0} \varphi(\lambda),$$

which provides a lower bound for the primal problem. If, in addition, we find a feasible $\bar{x} \in \mathcal{F}$ with attainment in the Lagrangian $\bar{x} \in \operatorname{argmin}_x L(x, \bar{\lambda})$ and with complementary slackness $\sum_k \bar{\lambda}_k q_k(\bar{x}) = 0$, then

$$\begin{aligned} \mu^* &\geq \nu^* \geq L(\bar{x}, \bar{\lambda}) \\ &= q_0(\bar{x}) \geq \mu^*. \end{aligned}$$

Therefore \bar{x} is optimal and the duality gap is zero when these sufficiency conditions (feasibility, attainment, complementary slackness) hold. Note that, since we are dealing with an unconstrained minimum of a quadratic Lagrangian, we obtain the interesting statement: the given conditions hold only if the Lagrangian is stationary and its Hessian is positive

semidefinite. Further, when these two conditions are incompatible, we lose strong duality, and can even expect a duality gap.

We now present several Q²P problems where the Lagrangian relaxation is important and well known. In all these cases, the Lagrangian dual provides an important theoretical tool for algorithmic development, even if the duality gap may be nonzero. We continue to emphasize our theme that the Lagrangian relaxation is best.

Convex Quadratic Programs. We start with the easy case; consider the convex quadratic program

$$\begin{aligned} \text{(CQP)} \quad \mu^* &:= \min q_0(x) \\ \text{s.t.} \quad &q_k(x) \leq 0, \quad k = 1, \dots, m, \end{aligned}$$

where all the functions $q_i(x)$ are convex and quadratic. The following remarks show that Lagrangian duality can always solve this problem.

The dual is

$$\text{(DCQP)} \quad \nu^* := \max_{\lambda \geq 0} \min_x q_0(x) + \sum_{k=1}^m \lambda_k q_k(x).$$

If ν^* is attained at $\lambda = \lambda^*$ and $x = x^*$, then *sufficient* conditions for x^* to be optimal for CQP are primal feasibility and complementary slackness

$$\sum_{k=1}^m \lambda_k^* q_k(x^*) = 0.$$

It is also well known that the Karush–Kuhn–Tucker (KKT) conditions are sufficient for global optimality, and, under an appropriate constraint qualification, they are also necessary. Therefore strong duality holds if a constraint qualification is satisfied, and then there is no duality gap and the dual is attained.

Further, *if the primal value of CQP is bounded then it is attained and there is no duality gap*, see [109, 89, 90, 88]. This assertion can be regarded as an extension of the Frank–Wolfe Theorem [68]. Surprisingly, however, the dual may not be attained. For example, the convex program

$$0 = \min\{x : x^2 \leq 0\} \tag{26}$$

has the (unattained) dual

$$0 = \max_{\lambda \geq 0} \min_x x + \lambda x^2 = \max_{\lambda > 0} \min_x x + \lambda x^2.$$

Algorithmic approaches based on Lagrangian duality appear in [48, 67, 77], for instance.

Rayleigh Quotients. Suppose that $A = A^T \in \mathcal{S}^n$. It is well known that the smallest eigenvalue λ_1 of A is the Rayleigh quotient

$$\lambda_1 = \min\{x^T A x : x^T x = 1\}. \quad (27)$$

Since A is not necessarily positive semidefinite, this form may require the minimization of a nonconvex function on a nonconvex set. The Rayleigh quotient, however, forms the basis of many very efficient algorithms for finding the smallest eigenvalue. It is easy to deduce that there is no duality gap for this nonconvex problem, by using the equation

$$\lambda_1 = \max_{\lambda} \min_x x^T A x - \lambda(x^T x - 1) = \max_{A - \lambda I \succeq 0} \lambda. \quad (28)$$

Specifically, the inner minimization problem in (28) is unconstrained. Therefore the outer maximization problem has the hidden semidefinite constraint (an ongoing theme in this paper)

$$A - \lambda I \succeq 0,$$

which requires λ to be at most the smallest eigenvalue of A . With λ set to the smallest eigenvalue, the inner minimization yields either $x = 0$ or the eigenvector corresponding to λ_1 , the corresponding value of the inner objective function being λ_1 in both cases. Thus, we have an example of a *nonconvex problem for which strong duality holds*. Note that the problem (27) has a special norm constraint, and a homogeneous quadratic objective. Thus, this nonconvex problem can be solved efficiently. Furthermore, strong duality holds for the Lagrangian dual, which supports our theme.

Trust Region Subproblems. We are going to see that strong duality holds for a larger class of apparently nonconvex problems. The trust region subproblem, TRS, is the minimization of a quadratic function subject to a norm constraint, where no convexity or homogeneity of the objective function is assumed. We make a further extension, i.e. we do not assume convexity of the constraint, so both the objective and constraint functions are allowed to be indefinite quadratics. Some applications of indefinite quadratic forms are given in [22]. This problem is important in nonlinear programming, e.g. [75, 74]. Specifically, TRS has the form

$$\begin{aligned} \text{(TRS)} \quad \mu^* &:= \min q_0(x) = x^T Q_0 x - 2c_0^T x \\ &\text{s.t. } x^T x - \delta^2 \leq 0 \text{ (or } = 0), \end{aligned}$$

and the generalized trust region subproblem [104, 73] is

$$\begin{aligned} \text{(GTRS)} \quad \mu^* &:= \min q_0(x) = x^T Q_0 x - 2c_0^T x \\ &\text{s.t. } q_1(x) \leq 0 \text{ (or } = 0), \end{aligned}$$

where q_1 is another quadratic function. Further, one can have two sided constraints $\alpha \leq q_1(x) \leq \beta$, which occur in some trust region algorithms as well.

For TRS, assuming that the constraint is written " \leq ," the Lagrangian dual is:

$$(DTRS) \quad \nu^* := \max_{\lambda \geq 0} \min_x q_0(x) + \lambda(x^T x - \delta^2).$$

An equivalent problem (see [104]) is the (concave) nonlinear semidefinite program

$$(DTRS) \quad \begin{aligned} \nu^* := \sup \quad & c_0^T (Q_0 + \lambda I)^\dagger c_0 - \lambda \delta^2 \\ \text{s.t.} \quad & Q_0 + \lambda I \succ 0, \quad \lambda \geq 0, \end{aligned}$$

where the superscript \dagger denotes the Moore–Penrose generalized inverse. It is shown in [104] that there is a zero duality gap for TRS, $\mu^* = \nu^*$. (The primal is attained though the dual may not be, as in example (26) below.) Thus, as in the eigenvalue calculation, we have an example of a nonconvex program where strong duality holds. Therefore this problem can be solved efficiently, polynomial time results being presented in [117].

We include a short proof of strong duality for the inequality constrained case, based on the outline in [62], which applies a convex case result after a perturbation. Note that the key to the proof is being able to pass between the inequality and equality constraints.

Proof.

Without loss of generality, we assume that TRS is nonconvex, because otherwise we apply the convex results discussed above. Therefore μ^* is attained on the boundary of the feasible set and the smallest eigenvalue of Q_0 , denoted γ , is negative. Thus TRS has the required property

$$\begin{aligned} \mu^* &= \min_{x^T x \leq \delta^2} x^T (Q_0 - \gamma I)x - 2c_0^T x + \gamma x^T x \\ &= \min_{x^T x = \delta^2} x^T (Q_0 - \gamma I)x - 2c_0^T x + \gamma x^T x \quad (Q_0 \text{ is indefinite}) \\ &= \min_{x^T x = \delta^2} x^T (Q_0 - \gamma I)x - 2c_0^T x + \gamma \delta^2 \\ &= \min_{x^T x \leq \delta^2} x^T (Q_0 - \gamma I)x - 2c_0^T x + \gamma \delta^2 \quad (Q_0 - \gamma I \text{ is singular}) \\ &= \max_{\lambda \geq 0} \min_x x^T (Q_0 - \gamma I)x - 2c_0^T x + \lambda(x^T x - \delta^2) + \gamma \delta^2 \\ &= \max_{\lambda \geq 0} \min_x x^T Q_0 x - 2c_0^T x + (\lambda - \gamma)(x^T x - \delta^2) \quad (\text{convex case}) \\ &\leq \max_{\lambda \geq \gamma} \min_x x^T Q_0 x - 2c_0^T x + (\lambda - \gamma)(x^T x - \delta^2) \quad (\gamma < 0) \\ &= \nu^* \leq \mu^*. \end{aligned} \tag{29}$$



As mentioned above, extensions of this result to a two-sided, general, possibly nonconvex, constraint are discussed in [104, 73]. An algorithm based on Lagrangian duality appears in [98] and (implicitly) in [74, 99]. Such algorithms are highly efficient for the TRS problem, since they solve it almost as quickly as an eigenvalue problem.

This efficiency when the objective and constraint may be nonconvex is surprising. In fact, Martinez [70] shows that the TRS can have at most one local and nonglobal optimum, and that the Lagrangian at this point has one negative eigenvalue. Therefore, it is even more surprising that the Lagrangian dual (relaxation) allows one to find the global minimum without ever getting trapped near the local minimum.

In fact, for GTRS we still have a zero duality gap, though strong duality may fail, as shown in (26). The results in [104] provide strong duality for GTRS with a two sided constraint $\alpha \leq q_1(x) \leq \beta$, the constraint qualification being $\alpha < \beta$. In [73], necessary and sufficient optimality conditions are presented for GTRS with the constraint qualification $\min q_0(x) < \max q_0(x)$. A combination of these results with the extension of the Frank–Wolfe theorem [68] gives the following properties.

Theorem 2.2 *Consider GTRS: a zero duality gap always holds and, if the optimal value is finite, then it is attained.* ■

Two Trust Region Subproblem. The two trust region subproblem, TTRS, is the minimization of a (possibly nonconvex) quadratic function subject to a norm and a least squares constraint, i.e. two convex quadratic constraints. This problem arises in algorithms for general nonlinear programs that use a sequential quadratic programming approach, and is often called the CDT problem, because it was introduced by Celis, Dennis and Tapia [21].

In contrast to the above single TRS, the TTRS can have a nonzero duality gap [86, 118, 119, 120], which is closely related to quadratic theorems of the alternative [25]. In addition, if the constraints are not convex, then the primal may not be attained (see [68]).

As mentioned above, Martinez [70] shows that the TRS can have at most one local optimum that is nonglobal, the Lagrangian there having one negative eigenvalue. Therefore, if we have such a case and add another ball constraint that contains the local, nonglobal, optimum in its interior, so that this point becomes the global optimum, then we obtain a TTRS that does not have a zero duality gap due to the negative eigenvalue. It is uncertain what additional constraints are successful at

closing this duality gap. In fact, it is still an open problem whether TTRS is an NP-hard or a polynomial time problem.

General Q²P. The general, possibly nonconvex, Q²P has many applications in modeling and approximation theory, see e.g. the applications to SQP methods in [58]. Examples of approximations to Q²P also appear in [32].

The Lagrangian relaxation of a Q²P is equivalent to the SDP relaxation, and is sometimes called the Shor relaxation [103]. The Lagrangian relaxation can be written as an SDP if one takes into the account the hidden semidefinite constraint, the inner quadratic objective function being bounded below only if the Hessian is positive semidefinite. The SDP relaxation is then the Lagrangian dual of this semidefinite program. It can also be obtained directly by *lifting* the problem into matrix space using the identity $x^T Q x = \text{Trace } x^T Q x = \text{Trace } Q x x^T$, and relaxing $x x^T$ to a semidefinite matrix X .

The geometry of the original feasible set of Q²P can be related to the feasible set of the SDP relaxation. The connection is through *valid quadratic inequalities*, i.e. nonnegative (convex) combinations of the quadratic constraints; see [33, 52] and our Section 2.2.

Orthogonally Constrained Programs with Zero Duality Gaps.

Let $\mathcal{M}_{m,n}$ denote the space of $m \times n$ real matrices. We now follow the approach in [6, 5, 4] and consider the *orthonormal type constraint*

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

sometimes known as the Stiefel manifold [28], and the trust region type constraint

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

Applications and algorithms for optimization on orthonormal sets of matrices are discussed in [28]. In this section we will show that, for $m = n$, strong duality holds for a certain (nonconvex) quadratic program defined over orthonormal matrices. Because of the similarity of the orthonormality constraint to the norm constraint $x^T x = 1$, the given results can be viewed as a matrix generalization of the strong duality property (27) for the Rayleigh Quotient problem.

Let A and B be $n \times n$ symmetric matrices, and consider the orthonormally constrained homogeneous Q²P

$$\begin{aligned} (\text{QQP}_0) \quad \mu^O := \min \quad & \text{Trace } A X B X^T \\ \text{s.t.} \quad & X X^T = I. \end{aligned} \quad (30)$$

This problem can be solved exactly using Lagrange multipliers [43], or using the classical Hoffman–Wielandt inequality [14], the solution being as follows.

Proposition 2.1 *Let the orthogonal diagonalizations of A and B be $A = V\Sigma V^T$ and $B = U\Lambda U^T$, respectively, where the eigenvalues of Σ and Λ are in nonincreasing and nondecreasing order, respectively. Then the optimal value of QQP_O is $\mu^O = \text{Trace } \Sigma\Lambda$, and the optimal matrix X is the product VU^T , which is obtained from the orthogonal matrices of the diagonalizations. ■*

The Lagrangian dual of QQP_O is

$$\max_{S=S^T} \min_X \text{Trace } AXBX^T - \text{Trace } S(XX^T - I). \tag{31}$$

There can be a nonzero duality gap for this Lagrangian dual, however, as shown by example in [121]. The inner minimization of problem (31) is an unconstrained quadratic minimization in the elements of X , with hidden constraint on the Hessian

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

On the other hand, the first order stationarity conditions at an optimum X , for the original problem QQP_O , are equivalent to $AXB = SX$ or, by orthogonality, $AXBX^T = S$, which yields not only mutual diagonalizability of A and XBX^T but also the characterization of the optimum. One can easily construct examples with a duality gap, caused by a conflict between the semidefinite condition and stationarity. In order to close the duality gap, we need more constraints on X .

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

$$\begin{aligned} (\text{QQP}_{OO}) \quad \mu^O := \min \quad & \text{Trace } AXBX^T \\ \text{s.t.} \quad & XX^T = I, \quad X^T X = I. \end{aligned} \tag{32}$$

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

$$\begin{aligned} (\text{DQQP}_{OO}) \quad \mu^O \geq \mu^D := \max \quad & \text{Trace } S + \text{Trace } T \\ \text{s.t.} \quad & (I \otimes S) + (T \otimes I) \preceq (B \otimes A) \\ & S = S^T, \quad T = T^T, \end{aligned}$$

because, by homogeneity, the inner minimization over X is achieved at $X = 0$.

Theorem 2.3 *Strong duality holds for QQP_{OO} and 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 [49],

$$\begin{aligned} (\text{QAPT}) \quad \mu_{\text{QAPT}}^* := \quad & \min \quad \text{Trace } AXBX^T \\ & \text{s.t. } \quad XX^T \preceq I. \end{aligned} \quad (33)$$

The constraints are convex with respect to the Löwner partial order, i.e. the partial order induced by the cone of positive semidefnite matrices. This problem is visually similar to the TRS problem discussed above, and we hope that methods for its solution will be useful. Therefore we would like to find a characterization of optimality. The set $\{X : XX^T \preceq I\}$ is studied separately in [80, 30], and is useful in eigenvalue variational principles.

It is shown in [5] that the following generalization of the Hoffman-Wielandt inequality holds,

Theorem 2.4 *Let $V^T AV = \Sigma$ and $U^T BU = \Lambda$ be the orthogonal diagonalizations of A and B , respectively, and let their eigenvalues be in nonincreasing order, say $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Then, for any $XX^T \preceq I$, we have the inequality*

$$\sum_{i=1}^n \min\{\lambda_i \sigma_{n-i+1}, 0\} \leq \text{Trace } AXBX^T \leq \sum_{i=1}^n \max\{\lambda_i \sigma_i, 0\}.$$

The upper bound is attained for $X = V \text{Diag}(\varepsilon) U^T$, where $\varepsilon_i = 1$ if $\sigma_i \lambda_i \geq 0$, and $\varepsilon_i = 0$ otherwise. The lower bound is attained for $X = V \text{Diag}(\varepsilon) J U^T$, where $\varepsilon_i = 1$ if $\sigma_i \lambda_{n+1-i} \leq 0$ and $\varepsilon_i = 0$ otherwise. Further, J is the permutation matrix $(e_n, e_{n-1}, \dots, e_1)$, where e_i is the i -th coordinate vector. ■

For a scalar ξ , let $[\xi]^- := \min\{0, \xi\}$. The lower bound in the above Theorem 2.4 establishes $\mu_{\text{QAPT}}^* = \sum_{i=1}^n [\lambda_i \sigma_{n-i+1}]^-$. Since the theorem also provides the feasible point of attainment, i.e. an upper bound for the relaxation problem, the theorem can be proved by showing that the value μ_{QAPT}^* is also attained by a Lagrangian dual program. However, there can be a duality gap if we use the Lagrangian dual of the trust region type relaxation with the constraint (33). Note that XX^T and $X^T X$ have the same eigenvalues, which implies $XX^T \preceq I$ if and only if $X^T X \preceq I$.

Explicitly, using both sets of constraints as in [6], we write QAPT in the form

$$\begin{aligned} \text{(QAPTR)} \quad \mu_{QAPT}^* &:= \min \text{Trace } AXBX^T \\ \text{s.t.} \quad &XX^T \preceq I, \quad X^T X \preceq I. \end{aligned}$$

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

$$\begin{aligned} \text{(DQAPTR)} \quad \mu_{QAPT}^* \geq \mu_{QAPT}^D &:= \max -\text{Trace } S - \text{Trace } T \\ \text{s.t.} \quad &-(I \otimes S) - (T \otimes I) \preceq (B \otimes A) \\ &S \succeq 0, \quad T \succeq 0. \end{aligned}$$

The following properties are proved in [5].

Theorem 2.5 *Strong duality holds for QAPTR and DQAPTR, i.e. $\mu_{QAPT}^D = \mu_{QAPT}^*$, and both primal and dual are attained. ■*

3. A STRENGTHENED BOUND FOR MC

From the results in Section 2.1, it appears that we might have the strongest possible tractable bound for the max-cut problem MC (or its equivalent quadratic model MCQ). Bounds can be strengthened, however, by adding redundant constraints. We now present a strategy (recipe) for constructing relaxations.

Algorithm 3.1 (Recipe for SDP relaxations)

- *add redundant constraints*
- *homogenize*
- *take Lagrangian dual*
- *use hidden semidefinite constraint to obtain SDP equivalent*
- *take Lagrangian dual of SDP*
- *check Slater's CQ — project if it fails*
- *delete redundant constraints from final SDP*

We will use the above recipe to strengthen our bound for MC. Therefore we introduce the following notation for linear operators and adjoints. For $S \in \mathcal{S}^n$, the vector $s = \text{svec}(S) \in \mathfrak{R}^{t(n)}$ is formed (columnwise) from

S , ignoring the strictly lower triangular part of S , so $t(n) = n(n + 1)/2$. The inverse operator, which constructs $S \in \mathcal{S}^n$ from $s \in \mathfrak{R}^{t(n)}$ is $S = \text{sMat}(s)$. The adjoint of svec is the operator hMat , where $\text{hMat}(v)$ is like $\text{sMat}(v)$, $v \in \mathfrak{R}^{t(n)}$, except that the off-diagonal elements are halved, in order to provide the adjoint equation

$$\text{svec}(S)^T v = \text{Trace } S \text{ hMat}(v), \quad \forall S \in \mathcal{S}^n, \quad v \in \mathfrak{R}^{t(n)}.$$

The adjoint of sMat is the operator $\text{dsvec}(S)$, which works like svec , except that the off diagonal elements are multiplied by 2, in order to satisfy

$$\text{dsvec}(S)^T v = \text{Trace } S \text{ sMat}(v), \quad \forall S \in \mathcal{S}^n, \quad v \in \mathfrak{R}^{t(n)}.$$

For notational convenience, we define the vectors

$$\text{sdiag}(s) := \text{diag}(\text{sMat}(s)) \in \mathfrak{R}^n, \quad \forall s \in \mathfrak{R}^{t(n)},$$

and

$$\text{vsMat}(s) = \text{vec}(\text{sMat}(s)) \in \mathfrak{R}^{n^2}, \quad \forall s \in \mathfrak{R}^{t(n)}.$$

The adjoint of vsMat is then given by

$$\text{vsMat}^*(v) = \text{dsvec}((\text{Mat}(v) + \text{Mat}(v)^T)/2) \in \mathfrak{R}^{t(n)}, \quad \forall v \in \mathfrak{R}^{n^2}.$$

The following bound is motivated by the strong duality results of Section 2.3, and is studied in depth in [6]. We recall that the SDP bound (20) for MCQ arises from a lifting procedure that employs

$$0 \preceq X = xx^T \quad \text{and} \quad x^T Q x = \text{Trace } Q X.$$

Discarding the rank one condition on X provides the tractable SDP bound, but it is not clear what constraints one can add to the P_E problem (21), in order to strengthen the Lagrangian relaxation. Linear combinations of the constraints will not help since they are already included in the Lagrangian. Nor is it clear what linear constraints to add to the SDP relaxation (20). One can try including the so-called triangle inequalities, which are standard in branch and bound methods for MC. This choice is sometimes very successful in practice [47], but one cannot guarantee an improvement [50]. Instead, we see that the matrices $X = xx^T$ have the property

$$X^2 = xx^T xx^T = nX$$

in the case $x_i = \pm 1, i = 1, \dots, n$. Therefore the quadratic matrix model

$$\begin{aligned} \mu^* := \max \quad & \text{Trace } Q X \\ \text{s.t.} \quad & \text{diag}(X) = e \\ & X^2 - nX = 0, \end{aligned} \tag{34}$$

where X is a symmetric matrix, is equivalent to MCQ. A common diagonalization of X and X^2 shows that the only eigenvalues of X are 0 and n , while (34) shows $\text{Trace } X = n$. Therefore the rank of X is one and $X \succeq 0$.

To illustrate the recipe (Algorithm 3.1), we add an additional redundant constraint, $X \circ X = E$, where E is the matrix of ones, obtaining the program

$$\begin{aligned}
 \text{(MC2)} \quad \mu^* = \max \quad & \text{Trace } QX \\
 \text{s.t.} \quad & \text{diag}(X) = e \\
 & X \circ X = E \\
 & X^2 - nX = 0,
 \end{aligned} \tag{35}$$

which is equivalent to (34) and so to MC. To apply Lagrangian relaxation efficiently, without losing information from the linear constraint, we replace $\text{diag}(X) = e$ by the norm constraint $\|\text{diag}(X) - e\|^2 = 0$. Then we introduce homogeneity by adding the variable y_0 and the constraint $1 - y_0^2 = 0$. (Because of the homogenization, if x and $y_0 = -1$ are optimal, then we can just multiply both x and y by -1 without changing the optimal value or feasibility.) Thus, because $X = \text{sMat}(x)$ is a symmetric matrix, we get the equivalent program

$$\begin{aligned}
 \max_{x, y_0} \quad & y_0 \text{Trace}(Q \text{sMat}(x)) \\
 \text{s.t.} \quad & \text{sdiag}(x)^T \text{sdiag}(x) - 2y_0 e^T \text{sdiag}(x) + n = 0 \\
 & \text{sMat}(x) \circ \text{sMat}(x) = E \\
 & \text{sMat}(x)^2 - ny_0 \text{sMat}(x) = 0 \\
 & 1 - y_0^2 = 0.
 \end{aligned} \tag{36}$$

In fact, letting $z = \begin{pmatrix} y_0 \\ x \end{pmatrix}$, we see that we have another max-cut problem with the same objective function as MCQ, additional linear constraints $\text{sdiag } x = e$ (these nodes must be grouped together), and the extra nonlinear constraints $\text{sMat}(x)^2 - ny_0 \text{sMat}(x) = 0$.

As discussed in Remark 1.2, the Lagrange multipliers for symmetric matrix valued constraint functions are symmetric matrices. Therefore problem (36) has the Lagrangian dual

$$\begin{aligned}
 \mu^* \leq \nu_2^* := \min_{w, T, S} \max_{x, y_0^2=1} \quad & y_0 \text{Trace}(Q \text{sMat}(x)) \\
 & + w(\text{sdiag}(x)^T \text{sdiag}(x) - 2y_0 e^T \text{sdiag}(x) + n) \\
 & + \text{Trace } T(E - \text{sMat}(x) \circ \text{sMat}(x)) \\
 & + \text{Trace } S((\text{sMat}(x))^2 - ny_0 \text{sMat}(x)),
 \end{aligned}$$

where w , T and S are the Lagrange multipliers, the matrices T and S being symmetric. We can move the constraint on y_0 into the Lagrangian

without increasing the duality gap, which provides the form

$$\begin{aligned} \nu_2^* = \min_{w,T,S,t} \max_{x,y_0} & y_0 \text{Trace} (Q \text{sMat} (x)) \\ & + w (\text{sdiag} (x)^T \text{sdiag} (x) - 2y_0 e^T \text{sdiag} (x) + n) \\ & + \text{Trace} T (E - \text{sMat} (x) \circ \text{sMat} (x)) \\ & + \text{Trace} S ((\text{sMat} (x))^2 - ny_0 \text{sMat} (x)) + t(1 - y_0^2). \end{aligned}$$

The inner maximization of the above relaxation is an unconstrained pure quadratic maximization, so the optimal value is infinity unless the Hessian is negative semidefnite (**hidden constraint**), in which case $x = 0$ and $y_0 = 0$ are optimal. Therefore we need to evaluate the Hessian of the Lagrangian to find the first SDP in the recipe.

Using $\text{Trace} (Q \text{sMat} (x)) = x^T \text{dsvec} (Q)$, and adding a 2 for convenience, we find that the constant part (no Lagrange multipliers) of the Hessian is

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

We now specify and manipulate some of the other linear operators that appear in the Lagrangian, which facilitates the derivation of the Hessian and the adjoints.

The matrix $\text{sdiag}^* \text{sdiag} \in \mathcal{S}^{t(n)}$, where sdiag^* is the adjoint of sdiag , is diagonal and has the elements

$$\begin{aligned} e_i^T (\text{sdiag}^* \text{sdiag}) e_j &= \text{sdiag} (e_i)^T \text{sdiag} (e_j) \\ &= \begin{cases} 1 & \text{if } i = j = t(k) \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{37}$$

Moreover, if $T = \sum_{ij} t_{ij} E_{ij}$, where the matrices $E_{ij} \in \mathcal{S}^n$ are the elementary matrices $e_i e_j^T + e_j e_i^T$, then linearity gives

$$\text{dsvec} (T \circ \text{sMat}) = \sum_{ij} t_{ij} \text{dsvec} (E_{ij} \circ \text{sMat}). \tag{38}$$

Also we have

$$\text{dsvec} \text{Diag} \text{diag} \text{sMat} = \text{sdiag}^* \text{sdiag} = \text{Diag} \text{svec} (I_n), \tag{39}$$

where the first two terms can be equated to a matrix because they are linear operators from $\mathfrak{R}^{t(n)}$ to $\mathfrak{R}^{t(n)}$, and where I_n is the $n \times n$ unit matrix. Using these relations, we rewrite the quadratic forms in the Lagrangian as

$$\begin{aligned} \text{sdiag} (x)^T \text{sdiag} (x) &= x^T (\text{dsvec} \text{Diag} \text{diag} \text{sMat}) x, \\ y_0 e^T \text{sdiag} (x) &= y_0 (\text{dsvec} I_n)^T x, \end{aligned} \tag{40}$$

$$\begin{aligned} \text{Trace } T(\text{sMat}(x) \circ \text{sMat}(x)) &= x^T \{ \text{dsvec}(T \circ \text{sMat}(x)) \} \\ &= x^T (\text{dsvec}(T \circ \text{sMat})) x, \end{aligned} \tag{41}$$

$$\begin{aligned} \text{Trace } S(\text{sMat}(x))^2 &= \text{Trace } \text{sMat}(x) S \text{sMat}(x) \\ &= x^T \text{dsvec}(S \text{sMat}(x)) \\ &= x^T (\text{dsvec } S \text{sMat}) x. \end{aligned} \tag{42}$$

Because $S \text{sMat}(x)$ may not be symmetric in (42), we let $\text{dsvec}(B)$ denote $\text{dsvec}(\frac{1}{2}B + \frac{1}{2}B^T)$ if B is any unsymmetric $n \times n$ matrix. It is easy to find the Hessian of a quadratic form written as $x^T \cdot x$. Therefore, we can now write down the *negative* of the nonconstant part of the Hessian of the inner maximization of ν_2^* . Splitting it into four linear operators with the factor 2, we find the formula

$$\begin{aligned} 2\mathcal{H}(w, T, S, t) &:= 2\mathcal{H}_1(w) + 2\mathcal{H}_2(T) + 2\mathcal{H}_3(S) + 2\mathcal{H}_4(t) \\ &:= 2w \begin{pmatrix} 0 & (\text{dsvec } I_n)^T \\ (\text{dsvec } I_n) & -\text{sdiag}^* \text{sdiag} \end{pmatrix} \\ &\quad + 2 \begin{pmatrix} 0 & 0 \\ 0 & \text{dsvec}(T \circ \text{sMat}) \end{pmatrix} \\ &\quad + 2 \begin{pmatrix} 0 & \frac{n}{2} \text{dsvec}(S)^T \\ \frac{n}{2} \text{dsvec}(S) & -\text{dsvec } S \text{sMat} \end{pmatrix} + 2t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

After cancelling the 2, it follows that our calculation is the semidefinite program

$$\begin{aligned} \text{(MCSDP2)} \quad \nu_2^* &= \min_{w, T, S, t} \quad nw + \text{Trace } ET + \text{Trace } 0S + t \\ &\text{s.t.} \quad \mathcal{H}(w, T, S, t) \succeq H_c. \end{aligned}$$

By taking T to be sufficiently positive definite and t to be sufficiently large, we can guarantee Slater’s constraint qualification. Therefore there is no duality gap between this SDP and its dual, which is the following strengthened SDP relaxation of MC

$$\begin{aligned} \text{(MCPSDP2)} \quad \nu_2^* &= \max_{Y \succeq 0} \quad \text{Trace } H_c Y \\ &\text{s.t.} \quad \mathcal{H}_1^*(Y) = n, \quad \mathcal{H}_2^*(Y) = E, \\ &\quad \mathcal{H}_3^*(Y) = 0, \quad \mathcal{H}_4^*(Y) = 1, \end{aligned}$$

where the superscripts * denote adjoint operators as usual.

To find the explicit form of the SDP relaxation, we now derive the adjoint operators explicitly. We partition Y as

$$Y = \begin{pmatrix} Y_{00} & x^T \\ x & \bar{Y} \end{pmatrix}. \tag{43}$$

From (41) and

$$e_i^T (\text{dsvec} (T \circ \text{sMat})) e_j = (\text{dsvec} (T) \circ e_j)^T e_i, \quad 1 \leq i, j \leq t(n),$$

we deduce

$$\mathcal{H}_2^*(Y) = \text{sMat} \text{diag} (\bar{Y}).$$

Therefore, $\mathcal{H}_2^*(Y) = E$ and $\mathcal{H}_4^*(Y) = 1$ are equivalent to

$$\text{diag} (Y) = e. \tag{44}$$

Also, $\mathcal{H}_1^*(Y)$ is twice the sum of the elements in the first row of Y whose positions correspond to the diagonal elements of $\text{sMat} (\cdot)$ minus the sum of the elements in these positions in the diagonal of \bar{Y} , so the notation (43) provides

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

It follows from equation (44) that the constraint $\mathcal{H}_1^*(Y) = n$ requires

$$x_{t(i)} = Y_{0,t(i)} = 1, \quad \forall i = 1, \dots, n. \tag{45}$$

Finally, we find $\mathcal{H}_3^*(Y)$. Recall that the definition of \mathcal{H}_3 is made up of four blocks, with the bottom right block defined by minus the quadratic form (42). Thus this block can be seen in two ways: either as a quadratic form or as its $t(n) \times t(n)$ symmetric matrix representation $-\text{dsvec} S \text{sMat}$. We regard $\text{dsvec} \cdot \text{sMat}$ as the linear operator that maps any $n \times n$ symmetric matrix S to the $t(n) \times t(n)$ symmetric matrix $\text{dsvec} S \text{sMat}$. Further, we write $\text{vsMat}^* \text{vec}$ instead of $\text{dsvec} (= \text{sMat}^*)$, in order to avoid applying dsvec to unsymmetric matrices, as mentioned after (42). Therefore the definition of $\mathcal{H}_3(S)$ gives

$$\begin{aligned} \langle \mathcal{H}_3(S), Y \rangle &= n \text{dsvec} (S)^T x - \langle \text{vsMat}^* \text{vec} S \text{sMat}, \bar{Y} \rangle \\ &= n \text{Trace} S \text{sMat} (x) - \langle S, \text{sMat} \bar{Y} \text{vsMat}^* \text{vec} \rangle \\ &= \langle S, n \text{sMat} (x) - \text{sMat} \bar{Y} \text{vsMat}^* \text{vec} \rangle, \end{aligned}$$

which provides the adjoint

$$\mathcal{H}_3^*(Y) = n \text{sMat} (x) - \text{sMat} \bar{Y} \text{vsMat}^* \text{vec}. \tag{46}$$

We now have another linear operator that maps symmetric matrices into symmetric matrices. Specifically, $\text{sMat} \cdot \text{vsMat}^* \text{vec}$ maps any $t(n) \times t(n)$ symmetric matrix \bar{Y} to $\text{sMat} \bar{Y} \text{vsMat}^* \text{vec}$, which is construed as an $n \times n$ symmetric matrix by invoking the quadratic form

$$\begin{aligned} v^T \text{sMat} \bar{Y} \text{vsMat}^* \text{vec} v &= \langle vv^T, \text{sMat} \bar{Y} \text{vsMat}^* \text{vec} \rangle \\ &= \text{Trace} \{ \text{sMat} \bar{Y} \text{dsvec} (vv^T) \}. \end{aligned}$$

Indeed, the k, l element of this quadratic form, which becomes the k, l element of $\text{sMat } \bar{Y} \text{ vsMat}^* \text{vec}$, has the value

$$e_k^T \text{sMat } \bar{Y} \text{ vsMat}^* \text{vec } e_l = \frac{1}{2} \text{Trace} \{ \text{sMat } \bar{Y} \text{ dsvec} (e_k e_l^T + e_l e_k^T) \}.$$

The explicit form of $\mathcal{H}_3^*(Y)$ is given in the sums of MCPSDP3 below, and also we have removed many of the redundant constraints. We see that the space of the variables of the new problem has increased to $S^{t(n)+1}$, from S^n in (20), and that there are now $2t(n) - 1$ constraints. They come from (44), (45), and the upper triangular part of (46). The latter constraints can be related to the original condition $\text{sMat}(x) = \frac{1}{n} \text{sMat}(x) \text{sMat}(x)$. As in the original SDP relaxation MCSDP (20), we still require the matrix of variables to have ones on the diagonal, but we now have additional constraints on the first row of Y . Specifically, the new problem has the form

$$\begin{aligned}
 & \max_{Y \in S^{t(n)+1}} && \text{Trace } H_c Y \\
 & \text{s.t.} && \text{diag}(Y) = e \\
 & && Y_{0,t(i)} = 1, \quad \forall i = 1, \dots, n \\
 \text{(MCPSDP3)} & && Y_{0,t(j-1)+i} = \frac{1}{n} \left\{ \sum_{k=1}^i Y_{t(i-1)+k,t(j-1)+k} \right. \\
 & && \quad \left. + \sum_{k=i+1}^j Y_{t(k-1)+i,t(j-1)+k} \right. \\
 & && \quad \left. + \sum_{k=j+1}^n Y_{t(k-1)+i,t(k-1)+j} \right\}, \\
 & && \forall 1 \leq i < j \leq n \\
 & && Y \succeq 0.
 \end{aligned}$$

Remark 3.1 *Since the first row of Y has some off-diagonal elements equal to 1, and since all the diagonal elements are 1, the Slater constraint qualification fails for this problem, i.e. we cannot have a positive definite feasible matrix. Therefore we can project this problem onto a face of the semidefinite cone, and then project into a smaller dimensional space to get a new further simplified problem. The final program has constraints that are linearly independent and that satisfy Slater’s condition. Details are given in [3].*

The constraint $\mathcal{H}_3^*(Y) = 0$ is the key to proving the following useful result. It is remarkable because the nonlinear constraint that $\text{sMat}(x)$ be positive semidefinite (and in fact feasible for MCSDP) was discarded in MC2. A proof is included to illustrate the usefulness of using adjoints.

Lemma 3.1 *Suppose that Y is feasible in MCPSDP3. Then its first row satisfies*

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

and so is feasible in the original SDP relaxation MCSDP (20).

Proof. Let Y be feasible for MCPSP3, and write

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

The fact that \bar{Y} is a principal submatrix of Y implies $\bar{Y} \succeq 0$. We have previously established that

$$\mathcal{H}_3^*(Y) = n \text{sMat}(x) - \text{sMat} \bar{Y} \text{vsMat}^* \text{vec},$$

where $\text{vsMat}^* \text{vec}$ is essentially sMat^* except that it acts on possibly non-symmetric matrices. Therefore the constraint $\mathcal{H}_3^*(Y) = 0$ is equivalent to

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

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

The term $X^2 - nX$, from the added nonlinear constraint $X^2 - nX = 0$ in the original max-cut problem (34) or (35), has the following interesting and useful properties in the SDP relaxation.

Lemma 3.2 *Suppose that both X and \bar{X} are feasible for MCSDP. Then*

$$\text{Trace}(X^2 - nX)(\bar{X}^2 - n\bar{X}) \geq 0. \tag{47}$$

Suppose, in addition, that

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

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

$$\text{Trace}(X^2 - nX)(\bar{X}^2 - n\bar{X}) > 0. \tag{48}$$

Proof. See [3]. ■

The above two lemmas can be used to show the strength of this new bound. Indeed, unless there is no gap between MCSDP and MC, the relaxation MCPSP3 **always** provides a strict improvement over MCSDP. The proof is given in [3].

Theorem 3.1 *The optimal values satisfy*

$$\nu_2^* \leq \nu^* \quad \text{and} \quad \nu_2^* = \nu^* \Rightarrow \nu_2^* = \mu^*. \tag{49}$$

■

Thus we see that this new bound is always strictly better than the previous one. Numerical results on small problems are presented in [3]. They consistently provide a significant improvement over the already strong original SDP bound.

4. CONCLUSION

We have looked at several problems where strong relaxations exist. In each case we have shown that our theme holds; one cannot do better than the Lagrangian relaxation. In particular, this has led us to a recipe for finding strong relaxations for hard, discrete optimization problems and a strengthened relaxation for the max-cut problem.

Acknowledgments

The author would like to thank Mike Powell for many helpful comments that significantly improved the readability of this paper.

References

- [1] F. Alizadeh, J-P.A. Haeberly and M.L. Overton (1994), A new primal-dual interior-point method for semidefinite programming, in J.G. Lewis, ed., *Proceedings of the Fifth SIAM Conference on Applied Linear Algebra*, SIAM, pp. 113–117.
- [2] F. Alizadeh, J-P.A. Haeberly and M.L. Overton (1997), Complementarity and nondegeneracy in semidefinite programming, *Math. Programming*, 77, pp. 111–128.
- [3] M. Anjos and H. Wolkowicz (1999), A strengthened SDP relaxation via a second lifting for the Max-Cut problem, Technical Report Research Report, CORR 99-55, University of Waterloo, Waterloo, Canada.
- [4] K.M. Anstreicher (1999), Eigenvalue bounds versus semidefinite relaxations for the quadratic assignment problem, Technical report, University of Iowa, Iowa City, IA.
- [5] K.M. Anstreicher, X. Chen, H. Wolkowicz and Y. Yuan (1999), Strong duality for a trust-region type relaxation of QAP, *Linear Algebra Appl.*, 301, pp. 121–136.
- [6] K.M. Anstreicher and H. Wolkowicz (1999), On Lagrangian relaxation of quadratic matrix constraints, to appear in *SIAM J. Matrix Anal. Appl.*

- [7] E. Balas (1997), A modified lift-and-project procedure, *Math. Programming*, 79(1-3, Ser. B), pp. 19–31. Lectures on Mathematical Programming ismp97 (Lausanne, 1997).
- [8] E. Balas and E. Zemel (1984), Lifting and complementing yields all the facets of positive zero-one programming polytopes, in *Mathematical programming (Rio de Janeiro, 1981)*, North-Holland, Amsterdam, pp. 13–24.
- [9] G. P. Barker, M. Laidacker and G. Poole (1987), Projectionally exposed cones, *SIAM J. Algebraic Discrete Methods*, 8(1), pp. 100–105.
- [10] G.P. Barker and D. Carlson (1975), Cones of diagonally dominant matrices, *Pacific J. of Math.*, 57, pp. 15–32.
- [11] R. Bellman and K. Fan (1963), On systems of linear inequalities in Hermitian matrix variables, in *Proceedings of Symposia in Pure Mathematics, Vol 7, AMS*.
- [12] A. Ben-Israel, A. Charnes and K. Kortanek (1969), Duality and asymptotic solvability over cones, *Bull. Amer. Math. Soc.*, 75(2), pp. 318–324.
- [13] S. J. Benson, Y. Ye, and X. Zhang (1997), Solving large-scale sparse semidefinite programs for combinatorial optimization, Technical report, The University of Iowa.
- [14] R. Bhatia (1987), *Perturbation Bounds for Matrix Eigenvalues: Pitman Research Notes in Mathematics Series 162*, Longman, New York.
- [15] F. Bohnenblust (1948), Joint positiveness of matrices, Manuscript.
- [16] B. Borchers (1999), SDPLIB 1.2, a library of semidefinite programming test problems, Technical report, New Mexico Tech, Soccorrow, NM, to appear in *Optim. Methods Softw.*
- [17] J.M. Borwein and H. Wolkowicz (1980/81), Characterization of optimality for the abstract convex program with finite-dimensional range, *J. Austral. Math. Soc. Ser. A*, 30(4), pp. 390–411.
- [18] J.M. Borwein and H. Wolkowicz (1981), Regularizing the abstract convex program, *J. Math. Anal. Appl.*, 83(2), pp. 495–530.
- [19] J.M. Borwein and H. Wolkowicz (1982), Characterizations of optimality without constraint qualification for the abstract convex program, *Math. Programming Stud.*, 19, pp. 77–100. Optimality and stability in mathematical programming.
- [20] S. Boyd, V. Balakrishnan, E. Feron, and L. El Ghaoui (1993), Control system analysis and synthesis via linear matrix inequalities, *Proc. ACC*, pp. 2147–2154.

- [21] M.R. Celis, J.E. Dennis Jr. and R.A. Tapia (1984), A trust region strategy for nonlinear equality constrained optimization, in *Proceedings of the SIAM Conference on Numerical Optimization, Boulder, CO*. Also available as Technical Report TR84-1, Department of Mathematical Sciences, Rice University, Houston, TX.
- [22] S. Chandrasekaran, M. Gu and A. H. Sayed (1999), The sparse basis problem and multilinear algebra, *SIAM J. Matrix Anal. Appl.*, 20(2), pp. 354–362.
- [23] B. Craven and J.J. Koliha (1977), Generalizations of Farkas' theorem, *SIAM J. Matrix Anal. Appl.*, 8, pp. 983–997.
- [24] B. Craven and B. Mond (1981), Linear programming with matrix variables, *Linear Algebra Appl.*, 38, pp. 73–80.
- [25] J.-P. Crouzeix, J.-E. Martinez-Legaz and A. Seeger (1995), An alternative theorem for quadratic forms and extensions, *Linear Algebra Appl.*, 215, pp. 121–134.
- [26] Dem'yanov and Malozemov (1974), *Introduction to Minimax*, Keter Publishing House, Jerusalem.
- [27] H. Dym and I. Gohberg (1981), Extensions of band matrices with band inverses, *Linear Algebra Appl.*, 36, pp. 1–24.
- [28] A. Edelman, T. Arias and S.T. Smith (1999), The geometry of algorithms with orthogonality constraints, *SIAM J. Matrix Anal. Appl.*, 20(2), pp. 303–353 (electronic).
- [29] A.V. Fiacco (1983), *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, volume 165 of *Mathematics in Science and Engineering*, Academic Press.
- [30] P.A. Fillmore and J.P. Williams (1971), Some convexity theorems for matrices, *Glasgow Mathematical Journal*, 10, pp. 110–117.
- [31] F. Forgó (1988). *Nonconvex Programming*, Akadémiai Kiadó, Budapest.
- [32] M. Fu, Z. Luo and Y. Ye (1998), Approximation algorithms for quadratic programming, *J. Comb. Optim.*, 2(1), pp. 29–50.
- [33] T. Fujie and M. Kojima (1997), Semidefinite programming relaxation for nonconvex quadratic programs. *J. Global Optim.*, 10(4), pp. 367–380.
- [34] M.X. Goemans (1997), Semidefinite programming in combinatorial optimization, *Math. Programming*, 79, pp. 143–162.
- [35] M.X. Goemans and D.P. Williamson (1995), Improved approximation algorithms for maximum cut and satisfiability problems

- using semidefinite programming, *J. Assoc. Comput. Mach.*, 42(6), pp. 1115–1145.
- [36] A. J. Goldman and A. W. Tucker (1956), Theory of linear programming, in *Linear inequalities and related systems*, Annals of Mathematics Studies, no. 38, Princeton University Press, pp. 53–97..
- [37] B. Grone, C.R. Johnson, E. Marques de Sa and H. Wolkowicz (1984), Positive definite completions of partial Hermitian matrices, *Linear Algebra Appl.*, 58, pp. 109–124.
- [38] M. Grötschel, L. Lovász and A. Schrijver (1988), *Geometric Algorithms and Combinatorial Optimization*, Springer Verlag, Berlin.
- [39] Branko Grünbaum (1965), On the facial structure of convex polytopes, *Bull. Amer. Math. Soc.*, 71, pp. 559–560.
- [40] Branko Grünbaum (1967), *Convex polytopes*, Pure and Applied Mathematics, Vol. 16, Interscience Publishers John Wiley & Sons, Inc., New York. With the cooperation of Victor Klee, M. A. Perles and G. C. Shephard.
- [41] M. Guignard (1969), Generalized Kuhn–Tucker conditions for mathematical programming problems in a Banach space, *SIAM J. of Control*, 7, pp. 232–241.
- [42] O. Güler and L. Tuncel (1998), Characterization of the barrier parameter of homogeneous convex cones, *Math. Programming*, 81, pp. 55–76.
- [43] S.W. Hadley, F. Rendl and H. Wolkowicz (1992), A new lower bound via projection for the quadratic assignment problem, *Math. Oper. Res.*, 17(3), pp. 727–739.
- [44] R. Hauser (1999), Self-scaled barrier functions: Decomposition and classification, Technical Report DAMTP 1999/NA13, Department of Applied Mathematics and Theoretical Physics, University of Cambridge.
- [45] C. Helmberg (1994), *An interior point method for semidefinite programming and max-cut bounds*, PhD thesis, Graz University of Technology, Austria.
- [46] C. Helmberg and F. Rendl (1997), A spectral bundle method for semidefinite programming, Technical Report ZIB Preprint SC-97-37, Konrad-Zuse-Zentrum Berlin, Berlin, Germany, to appear in *SIAM J. Optim.*
- [47] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz (1996), An interior-point method for semidefinite programming, *SIAM J. Optim.*, 6(2), pp. 342–361.

- [48] F. Jarre (1990), On the convergence of the method of analytic centers when applied to convex quadratic programs, *Math. Programming*, 49(3), pp. 341–358.
- [49] S.E. Karisch, F. Rendl and H. Wolkowicz (1994), Trust regions and relaxations for the quadratic assignment problem, in *Quadratic assignment and related problems (New Brunswick, NJ, 1993)*, Amer. Math. Soc., Providence, RI, pp. 199–219.
- [50] H. Karloff (1999), How good is the Goemans-Williamson MAX CUT algorithm?, *SIAM J. Comput.*, 29(1), pp. 336–350.
- [51] M. Kojima and L. Tunçel (1998), Discretization and localization in successive convex relaxation methods for nonconvex quadratic optimization problems, Technical Report CORR98-34, University of Waterloo, Waterloo, Canada.
- [52] M. Kojima and L. Tunçel (1999), Cones of matrices and successive convex relaxations of nonconvex sets, to appear in *SIAM J. Optim.*.
- [53] F. Körner (1988), A tight bound for the Boolean quadratic optimization problem and its use in a branch and bound algorithm, *Optimization*, 19(5), pp. 711–721.
- [54] F. Körner (1992), Remarks on a difficult test problem for quadratic Boolean programming, *Optimization*, 26, pp. 355–357.
- [55] K. O. Kortanek and J. P. Evans (1968), Asymptotic Lagrange regularity for pseudoconcave programming with weak constraint qualification, *Operations Res.*, 16, pp. 849–857.
- [56] S. Kruk (1996), Semidefinite programming applied to nonlinear programming, Master's thesis, University of Waterloo, Waterloo, Canada.
- [57] S. Kruk, M. Muramatsu, F. Rendl, R.J. Vanderbei and H. Wolkowicz (1998), The Gauss–Newton direction in linear and semidefinite programming, Technical Report CORR 98-16, University of Waterloo, Waterloo, Canada.
- [58] S. Kruk and H. Wolkowicz (1998), SQ²P, sequential quadratic constrained quadratic programming, in Ya-xiang Yuan, ed., *Advances in Nonlinear Programming*, volume 14 of *Applied Optimization*, Kluwer, Dordrecht, pp. 177–204. Proceedings of Nonlinear Programming Conference in Beijing in honour of Professor M.J.D. Powell.
- [59] S. Kruk and H. Wolkowicz (2000), Sequential, quadratic constrained, quadratic programming for general nonlinear programming, in H. Wolkowicz, R. Saigal and L. Vandenbergh, eds,

Handbook of Semidefinite Programming: Theory, Algorithms and Applications, Kluwer Academic Publishers, Boston, MA.

- [60] M. Laurent (1998), 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*, Amer. Math. Soc., Providence, Rhode Island.
- [61] C. Lemaréchal and F. Oustry (1999), Nonsmooth algorithms to solve semidefinite programs, in L. EL Ghaoui and S-I. Niculescu, eds, *Recent Advances on LMI methods in Control*, Advances in Design and Control series, SIAM, to appear.
- [62] A.S. Lewis (1994), *Take-home final exam, Course CO663 in Convex Analysis*, University of Waterloo, Waterloo, Canada.
- [63] A.S. Lewis (1996), Derivatives of spectral functions, *Math. Oper. Res.*, 21(3), pp. 576–588.
- [64] A.S. Lewis and M.L. Overton (1996), Eigenvalue optimization, *Acta Numerica*, 5, pp. 149–190.
- [65] L. Lovász (1979), On the Shannon capacity of a graph, *IEEE Transactions on Information Theory*, 25, pp. 1–7.
- [66] L. Lovász and A. Schrijver (1991), Cones of matrices and set-functions and 0-1 optimization, *SIAM J. Optim.*, 1(2), pp. 166–190.
- [67] Z-Q. Luo and J. Sun (1995), An analytic center based column generation algorithm for convex quadratic feasibility problems, Technical report, McMaster University, Hamilton, Ontario.
- [68] Z-Q. Luo and S. Zhang (1997), On the extension of Frank–Wolfe theorem, Technical report, Erasmus University, Rotterdam, The Netherlands.
- [69] A.W. Marshall and I. Olkin (1979), *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York.
- [70] J.M. Martinez (1994), Local minimizers of quadratic functions on Euclidean balls and spheres, *SIAM J. Optim.*, 4(1), pp. 159–176.
- [71] V.M. Matrosova and A.I. Malikov (1993), The development of the ideas of A. M. Lyapunov over one hundred years: 1892–1992, *Izv. Vyssh. Uchebn. Zaved. Mat.*, 4, pp. 3–47.
- [72] R.D.C. Monteiro and Y. Zhang (1998), A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming, *Math. Programming*, 81, pp. 281–299.

- [73] J. J. Moré (1993), Generalizations of the trust region problem, *Optim. Methods Software*, 2, pp. 189–209.
- [74] J.J. Moré and D.C. Sorensen (1983), Computing a trust region step, *SIAM J. Sci. Statist. Comput.*, 4, pp. 553–572.
- [75] J.J. Moré and D.C. Sorensen (1984), Newton’s method, in G.H. Golub, ed., *Studies in Numerical Analysis*, vol. 24 of *MAA Studies in Mathematics*, The Mathematical Association of America.
- [76] Y.E. Nesterov and A.S. Nemirovski (1988), Polynomial barrier methods in convex programming, *Ėkonom. i Mat. Metody*, 24(6), pp. 1084–1091.
- [77] Y.E. Nesterov and A.S. Nemirovski (1994), *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Publications, SIAM, Philadelphia, USA.
- [78] Y.E. Nesterov, H. Wolkowicz and Y. Ye (2000), Semidefinite programming relaxations of nonconvex quadratic optimization, in H. Wolkowicz, R. Saigal and L. Vandenberghe, eds, *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*, Kluwer Academic Publishers, Boston, MA.
- [79] J. W. Nieuwenhuis (1980), Another application of Guignard’s generalized Kuhn–Tucker conditions, *J. Optim. Theory Appl.*, 30(1), pp. 117–125.
- [80] M.L. Overton and R.S. Womersley (1993), Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices, *Math. Programming*, 62, pp. 321–357.
- [81] P.M. Pardalos (1991), Quadratic programming with one negative eigenvalue is NP-hard, *J. Global Optim.*, 1, pp. 15–22.
- [82] P. C. Parks (1992), A. M. Lyapunov’s stability theory—100 years on, *IMA J. Math. Control Inform.*, 9(4), pp. 275–303.
- [83] G. Pataki (1998), On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues, *Math. Oper. Res.*, 23(2), pp. 339–358.
- [84] G. Pataki (2000), Geometry of Semidefinite Programming, in H. Wolkowicz, R. Saigal and L. Vandenberghe, eds, *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*, Kluwer Academic Publishers, Boston, MA.
- [85] G. Pataki and L. Tuncel (1997), On the generic properties of convex optimization problems in conic form, Technical Report CORR 97-16, University of Waterloo, Waterloo, Canada.

- [86] J. Peng and Y. Yuan (1997), Optimality conditions for the minimization of a quadratic with two quadratic constraints, *SIAM J. Optim.*, 7(3), pp. 579–594.
- [87] A.L. Peressini (1967), *Ordered Topological Vector Spaces*, Harper & Row Publishers, New York.
- [88] E.L. Peterson and J.G. Ecker (1969), Geometric programming: duality in quadratic programming and l_p -approximation, II. Canonical programs, *SIAM J. Appl. Math.*, 17, pp. 317–340.
- [89] E.L. Peterson and J.G. Ecker (1970), Geometric programming: duality in quadratic programming and l_p -approximation, III. Degenerate programs, *J. Math. Anal. Appl.*, 29, pp. 365–383.
- [90] E.L. Peterson and J.G. Ecker (1970), Geometric programming: duality in quadratic programming and l_p -approximation. I., in *Proceedings of the Princeton Symposium on Mathematical Programming (Princeton Univ., 1967)*, Princeton Univ. Press, pp. 445–480.
- [91] S. Poljak, F. Rendl and H. Wolkowicz (1995), A recipe for semidefinite relaxation for $(0, 1)$ -quadratic programming, *J. Global Optim.*, 7(1), pp. 51–73.
- [92] S. Poljak and H. Wolkowicz (1995), Convex relaxations of $(0, 1)$ -quadratic programming, *Math. Oper. Res.*, 20(3), pp. 550–561.
- [93] G.D. Poole and M. Laidacker (1988), Projectionally exposed cones in \mathbf{r}^3 , *Linear Algebra Appl.*, 111, pp. 183–190.
- [94] A.J. Quist, E. De Klerk, C. Roos and T. Terlaky (1998), Copositive relaxation for general quadratic programming, *Optim. Methods Softw.*, 9, pp. 185–208. Special Issue Celebrating the 60th Birthday of Professor Naum Shor.
- [95] M. Ramana, L. Tunçel and H. Wolkowicz (1997), Strong duality for semidefinite programming, *SIAM J. Optim.*, 7(3), pp. 641–662.
- [96] M.V. Ramana (1993), *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems*, PhD thesis, Johns Hopkins University, Baltimore.
- [97] M.V. Ramana (1997), An exact duality theory for semidefinite programming and its complexity implications, *Math. Programming*, 77, pp. 129–162.
- [98] F. Rendl and H. Wolkowicz (1997), A semidefinite framework for trust region subproblems with applications to large scale minimization, *Math. Programming*, 77, pp. 273–299.
- [99] S.A. Santos and D.C. Sorensen (1995), A new matrix-free algorithm for the large-scale trust-region subproblem, Technical Report TR95-20, Rice University, Houston, TX.

- [100] A. Shapiro (1997), First and second order analysis of nonlinear semidefinite programs, *Math. Programming*, 77, pp. 301–320.
- [101] A. Shapiro (2000), Duality and optimality conditions, in H. Wolkowicz, R. Saigal and L. Vandenberghe, eds, *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*, Kluwer Academic Publishers, Boston, MA.
- [102] H.D. Sherali and W.P. Adams (1996), Computational advances using the reformulation-linearization technique (rlt) to solve discrete and continuous nonconvex problems, *Optima*, 49, pp. 1–6.
- [103] N.Z. Shor (1987), Quadratic optimization problems, *Izv. Akad. Nauk SSSR Tekhn. Kibernet.*, 222(1), pp. 128–139.
- [104] R. Stern and H. Wolkowicz (1995), Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations, *SIAM J. Optim.*, 5(2), pp. 286–313.
- [105] J.F. Sturm (1997), *Primal-Dual Interior Point Approach to Semidefinite Programming*, PhD thesis, Erasmus University, Rotterdam, The Netherlands.
- [106] J.F. Sturm (1998), Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Technical report, Communications Research Laboratory, McMaster University, Hamilton, Canada. To appear in *Optimization Methods and Software*.
- [107] C. H. Sung and B.-S. Tam (1990), A study of projectionally exposed cones, *Linear Algebra Appl.*, 139, pp. 225–252.
- [108] O. Taussky (1968), Positive-definite matrices and their role in the study of the characteristic roots of general matrices, *Advances in Math.*, 2, pp. 175–186.
- [109] T. Terlaky (1985), On l_p programming, *European J. Oper. Res.*, 22, pp. 70–100.
- [110] M.J. Todd (1999), A study of search directions in primal-dual interior-point methods for semidefinite programming, *Optim. Methods Softw.*, 11&12, pp. 1–46.
- [111] L. Tuncel and S. Xu (1999), On homogeneous convex cones, Caratheodory number and duality mapping semidefinite liftings, Technical Report CORR 99-21, University of Waterloo, Waterloo, Canada.
- [112] H. Wolkowicz and Q. Zhao (1999), Semidefinite relaxations for the graph partitioning problem, *Discrete Applied Math.*, 96-97(1-3), pp. 461–479.

- [113] S. Wright (2000), Recent developments in interior-point methods, in M.J.D. Powell and S. Scholtes, eds, *System Modelling and Optimization: Proceedings of 19th IFIP TC7 Conference, July, 1999*, Cambridge, Kluwer Academic Publishers, Boston, MA.
- [114] V.A. Yakubovich (1967), The method of matrix inequalities in the stability theory of nonlinear control systems, *I, II, III*, *Automation and Remote Control*, 25-26(4), pp. 905–917, 577–592, 753–763.
- [115] V.A. Yakubovich (1973), Minimization of quadratic functionals under the quadratic constraints and the necessity of a frequency condition in the quadratic criterion for absolute stability of nonlinear control systems, *Dokl. Akad. Nauk SSSR*, 209(14), pp. 1039–1042.
- [116] V.A. Yakubovich (1973), The S-procedure and duality theorems for nonconvex problems of quadratic programming, *Vestnik Leningrad Univ.*, 1973(1), pp. 81–87.
- [117] Y. Ye (1992), A new complexity result on minimization of a quadratic function with a sphere constraint, in *Recent Advances in Global Optimization*, Princeton University Press, pp. 19–31.
- [118] Y. Yuan (1983), Some properties of trust region algorithms for nonsmooth optimization, Technical Report DAMTP 1983/NA14, Department of Applied Mathematics and Theoretical Physics, University of Cambridge.
- [119] Y. Yuan (1990), On a subproblem of trust region algorithms for constrained optimization, *Math. Programming*, 47, pp. 53–63.
- [120] Y. Yuan (1991), A dual algorithm for minimizing a quadratic function with two quadratic constraints, *Journal of Computational Mathematics*, 9, pp. 348–359.
- [121] Q. Zhao, S.E. Karisch, F. Rendl and H. Wolkowicz (1998), Semidefnite programming relaxations for the quadratic assignment problem, *J. Comb. Optim.*, 2(1), pp. 71–109.
- [122] S. Zlobec (1971), Optimality conditions for mathematical programming problems, *Glasnik Mat. Ser. III*, 6(26), pp. 187–192.