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 Natural Sciences Engineering Research Council Canada.

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 has 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, e.g. the people working on numerical issues for large scale problems are now using sophisticated techniques in convex analysis; 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 $x^2 - x = 0$, and $x^2 = 1$, respectively. Thus many hard combinatorial problems can be modelled using quadratically constrained quadratic programs, denoted $\mathrm{Q}^2\mathrm{P}$. 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). However, using the positive semidefinite matrix relationship $X = xx^T$, 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 explore the relationship between the SDP and Lagrangian relaxations of various classes of $\mathrm{Q}^2\mathrm{P}$s. (This continues on the work in e.g. [103, 91].) We then explore 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, see Section 2.1. 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 Q^2P in Section 2.2. This 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 is studied in Section 2.3. In every instance where one has a tractable bound, we find a Q^2P 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 Q^2P. 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 Q^2P

The Q^2P in x is

$$(Q^2P_x) \qquad \begin{array}{rll} \mu^* := & \min & q_0(x) := x^TQ_0x + 2g_0^Tx + \alpha_0 \\ & s.t. & q_k(x) := x^TQ_kx + 2g_k^Tx + \alpha_k \leq 0 \\ & & k \in \mathcal{I} := \{1, \ldots, m\} \\ & & x \in \Re^n, \end{array}$$

where $Q_k \in \mathcal{S}^n$, the space of symmetric matrices. The Lagrangian of Q^2P_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 Q^2P at the optimum x (e.g. the Mangasarian-Fromovitz constraint qualification), then the Karush-Kuhn-Tucker necessary conditions for optimality hold, i.e.

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

for some $0 \leq \lambda \in \Re^m$. Therefore, the optimum x can be searched among the points satisfying stationarity of the Lagrangian and complementary slackness. Moreover, if the Lagrangian is also convex, then this (and primal feasibility) is a sufficient condition for optimality.

Lagrange multipliers can also be used to derive the Lagrangian dual (or relaxation) of Q^{2}P$_x$

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

i.e. each inner unconstrained minimization problem provides a lower bound for Q^{2}P; and, we then choose the best of these lower bounds. A zero duality gap holds if $\mu^* = \nu^*$. This 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, e.g. 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 result of the inner product $\langle Q(x), \Lambda \rangle$. With a nonnegativity constraint, we will get a sign restriction on the Lagrange multiplier.*

1.2. SEMIDEFINITE PROGRAMMING PRELIMINARIES

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

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

where K is a convex cone ($K + K \subset K, \quad \alpha K \subset K, \ \forall \alpha \geq 0$) and $g(x) \succeq_K 0$ denotes the cone partial order $g(x) \in K$. If $K = \Re^n_+$, then we get the usual elementwise ordering $g(x) \geq 0$. This notation allows for 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 this is called the *Löwner partial order*, e.g. [69]. In fact, this is a very general mathematical program and includes e.g. standard inequality, equality, and semidefinite constraints, when $K = \Re^n_+ \otimes \{0\} \otimes \mathcal{P}$, and $\{0\}$ is the set containing the 0 vector. We now look at more of the similarities and differences 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 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 appears 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. Some sample results similar to LP are: the SDP cone is *self-polar*, i.e.

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

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

A face of a cone K is defined as

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

i.e. points $x \in \mathcal{F}$ can be convex combinations of points $y, z \in K$ only if the points $y, z \in \mathcal{F}$. The faces of $\Re^n_+$ 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

$$\{x \in \Re^n_+ : \langle x, z \rangle = 0\} = \{x \in \Re^n_+ : y_i = 0 \Rightarrow x_i = 0\}.$$

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

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

where $\mathcal{N}(\cdot)$ denotes nullspace. The faces of $\mathcal{P}$ are *exposed*, i.e. $\mathcal{F} = \mathcal{P} \cap \phi^\perp$, where $\phi \in \mathcal{P} \cap \mathcal{F}^\perp$ is a *conjugate face*. (Here $\cdot^\perp$ denotes orthogonal complement.) Moreover, they are *projectionally exposed*, i.e. there exists a projection matrix P such that $\mathcal{F} = P\mathcal{P}P$; 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 projectionally 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 the following property of the faces of $\mathcal{P}$.

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

On the other hand, these difficulties can be corrected by taking advantage of the following.

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

We amplify on this below, in Remark 1.3.

Now, consider the typical primal SDP

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

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, i.e. the components are defined as

$$(\mathcal{A}(X))_i = \operatorname{Trace} A_i X = b_i,$$

for some given symmetric matrices A_i.

We can define the dual and prove weak duality using the notion of a hidden constraint, i.e.

$$\begin{array}{rcl} \mu^* & = & \min\limits_{X \succeq 0} \max\limits_{y} C \bullet X + y^T(b - \mathcal{A}(X)) \\ & \geq & \max\limits_{y} \min\limits_{X \succeq 0} y^T b + (C - \mathcal{A}^* y) \bullet X \\ & = & \nu^*, \end{array} \tag{4}$$

where $\mathcal{A}^*$ is the adjoint operator of $\mathcal{A}$, i.e. it satisfies

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

and can be defined as $\mathcal{A}^*(y) = \sum_i y_i A_i$. (The symmetric matrices A_i are defined as above.) The dual program is

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

The first equality 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; while the last equality follows from the hidden constraint that guarantees the inner minimization is finite valued.

The duality theory gives rise to the elegant characterization of optimality

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

The variables $X, (y, Z)$ are called a primal-dual optimal pair, Z is the (dual) slack variable. In the case that complementary slackness may fail, the variables are called a primal-dual feasible pair. Note that complementary slackness should actually be written in the equivalent form $\operatorname{Trace} ZX = 0$. This follows from orthogonal diagonalization of Z, X. This is also equivalent to having a zero duality gap between primal and dual values for a feasible pair since

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

Remark 1.3 *We mentioned above that the optimality conditions require a CQ, e.g. strict feasibility, and that the 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 so that its range satisfies $\mathcal{R}(\mathcal{A}^*) = span(\mathcal{F})$*, for some face* $\mathcal{F}$ *of* $\mathcal{P}$*, then we can choose* C *in the set* $\mathcal{R}(\mathcal{A}^*) + \mathcal{P}$*, but not its closure. This would force failure of dual feasibility and a duality gap. For example, let* $A := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$*, and so* $\mathcal{A}^*y = yA$*. For* $n = 2$*, we can visualize symmetric matrices in* $\Re^3$ *and* $\mathcal{P}$ *as the* ice-cream cone*, the cone of vectors that make an angle less or equal to 45 degrees with the identity matrix (vector) in the trace inner-product. The faces are rays on the boundary of the cone. Then we see that* $\mathcal{R}(\mathcal{A}^*)$ *is the span of a face, i.e. the face through* A*, and* $\mathcal{R}(\mathcal{A}^*) + \mathcal{P}$ *is not closed. We choose* $C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$*, and* $b = 0$*. Since* $b = 0, X \succeq 0$*, we see that* X *feasible implies* $X_{22} = 0$*, which in turn implies* $X_{21} = 0$*, i.e. the optimal value* $C \bullet X = 0$*. Then, we can choose e.g.* $X = 0$ *primal feasible and optimal with optimal value 0. However, the dual is infeasible as we argued above. In fact, we can see this directly since* $C - yA$ *cannot be* *semidefinite. (The dual optimal value is* $\infty < 0$, *i.e. we have a duality gap.)*

The condition (3) can be used to correct this problem by allowing for 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]. For our model (P), 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.* $\mathcal{A}^* y \preceq_{\mathcal{F}^+} C$. *For our example, the minimal face* $\mathcal{F} = span \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ *which implies that* $C - (-1)A \in \mathcal{F}^+$, *i.e. the new dual optimal value is 0 and we have a zero duality gap.*

If we perturb the complementary slackness conditions,

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

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. 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; this is clear from the empirical evidence and the derivation of several public domain codes, e.g. [47, 1, 105, 106, 16]. SDP are convex programs and fall into the class of problems that can be approximately solved to a desired accuracy in polynomial time, i.e. these are tractable problems. 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. However, this is an ongoing area of research 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 explicitly evaluate the primal matrix X, see e.g. [13, 46, 61].

First and second order optimality conditions for SDP are given in e.g. [100, 101]; see also the survey article [113]. Nondegeneracy and strict complementarity are discussed in [2, 85]. Both nondegeneracy and strict complementarity can fail, though they are also both 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 follow through to SDP. Note that strict complementarity for SDP translates to $Z + X \succ 0$.

2. RELAXATIONS OF Q^2P

We now look at a particular instance of Q^2P, the quadratic model for the *Max-Cut problem.* In particular, we start with several different tractable relaxations for the max-cut 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. This includes problems where orthogonal constraints arise, e.g. orthogonal relaxations of the quadratic assignment and graph partitioning problems. In particular, this part of the chapter emphasizes the theme about the strength of the 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 edge set $V = \{v_i\}_{i=1}^n$ and weights w_{ij} on the edges $(v_i, v_j) \in E$. We want to find the index set $\mathcal{I} \subset \{1, 2, \dots n\}$, to maximize the weight of the edges with one end point with index in $\mathcal{I}$ and the other in the complement. This is equivalent to

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

where $\mathcal{F} := \{\pm 1\}^n$, and $x_i = 1$, if $i \in \mathcal{I}$, and -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

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

MCQ (and MC) is an intractable problem. There are many different ways to relax the problem and find approximate solutions and/or bounds. The simplest example is to relax the constraints to the $[-1, 1]$ interval. 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 bound relaxes the constraints to the unit ball of radius $\sqrt{n}$, while another relaxes the constraints to the convex hull, i.e. to the unit cube.

The relaxations yield bounds which are derived using the fact that we can perturb the objective function q_0 since $x_i^2 = 1$ on the feasible set $\mathcal{F}$, i.e. note that

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

where e is the vector of ones (of the appropriate dimension), and Diag (u) denotes the diagonal matrix formed from the vector u. For each u we get a trivial upper bound obtained from ignoring the constraints and allowing the diagonal perturbations, i.e. we have

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

We then get the following trivial bound.

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

Let

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

Note that if the set S is not empty, then we can minimize over the unconstrained parameter u or add the restriction to $u^T e = 0$, 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 perturbation analysis of min-max problems, e.g. [29]. For example, if 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 (or $u^T e = 0$, in the restricted case), in any direction h, $||h|| = 1$, and the directional derivative equals

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

where $\mathcal{X}(u)$ denotes the optimal set of the inner problem for given u. One can now compare stationarity for f_0 in the unconstrained case and then use 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 it is finite valued. (Details can be found in [92]. The comment about using $u^T e = 0$ or not, is true for the bounds that follow as well.)

But, the function f_0 can take on the value $+\infty$. We can restrict the parameters and avoid infinite values for the inner maximization problem by adding the hidden semidefinite constraint, i.e. we use the fact that a quadratic function is unbounded above if the Hessian is not negative semidefinite. (Note that a quadratic function is bounded above if and only if the Hessian is negative semidefinite and the stationarity equation is consistent.) The following is a tractable bound since we minimize a convex function over a convex set.

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

Next we relax the feasible set to the sphere of radius $\sqrt{n}$. We get

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

And 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, see e.g. Section 2.3 below. Thus we have our second tractable bound.

We can replace the spherical constraint with the box constraint.

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

After adding the semidefinite constraint to make the bound tractable, i.e. to make the calculation of f_2 tractable, we get our next bounds.

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

and

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

Given Q and c, define the $(n+1)\times(n+1)$-matrix Q^c by adding a $0-th$ row and column, so that

$$\begin{array}{ll} q^c_{00} = 0 & \\ q^c_{0i} = q^c_{i0} = -c_i & \text{for } i > 0 \\ q^c_{ij} = q_{ij} & \text{for } i,j > 0, \end{array}$$

i.e.

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

In order to have analogous functions $q_u^c(y)$ and $f_i(u)$ as in the previous cases, let us introduce

$$q_u^c(y) := y^T(Q^c + \mathrm{Diag}\,(u))y - u^Te, \tag{17}$$

where u, e are now in $\Re^{n+1}$. Note that q_u^c reduces to q_u if the first component y_0 is ± 1. The equivalent relaxed problem is

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

where $\lambda_{\max}$ denotes the maximum eigenvalue. Now another bound is

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

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 consider the special case that $c = 0$. We now look at the SDP bound. The relaxation comes from the fact that the trace is commutative, i.e.

$$x^TQx = \mathrm{Trace}\, x^TQx = \mathrm{Trace}\, Qxx^T$$

and, for $x \in \mathcal{F}$, $y_{ij} = x_ix_j$ defines a symmetric, rank one, positive semidefinite matrix Y with diagonal elements 1. Therefore, we can *lift* the problem into the higher dimensional space of symmetric matrices and relax the rank one constraint. This yields the following relaxation and our bound 3.

$$\begin{array}{rrl} & B_3 := \quad \max & \mathrm{Trace}\, QY \\ MCSDP & \text{s.t.} & \mathrm{diag}\,(Y) = e \\ & & Y \succeq 0, \end{array} \tag{20}$$

where $\mathrm{diag}\,(Y)$ 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 these different bounds. We now do something that may seem meaningless; we replace the ± 1 constraints with $x_i^2 = 1, \forall i$. This does not change the feasible set of the original problem. In [92, 91] it is shown that all the above relaxations and bounds for MC come from the Lagrangian dual of (P_E), the following equivalent problem to MCQ. Thus we enforce our theme about the strength of the Lagrangian relaxation. The strong duality result for the trust region subproblem is the key to the proofs.

$$
(P_E) \quad \begin{array}{ll} \max & q_0(x) = x^TQx - 2c^Tx \\ \text{s.t.} & x_i^2 = 1, \quad i = 1, \cdots, n. \end{array} \tag{21}
$$

Note that the Lagrangian dual of P_E yields precisely our trivial first bound B_0 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 Q^2P. The general Q^2P problem is also studied in e.g. [33, 52, 96, 58, 56, 68, 59].

Quadratic bounds using a Lagrangian relaxation have been extensively studied and 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, $Ax = b$, is transformed to a quadratic constraint via $||Ax - b||^2 = 0$. The reason for these transformations for linear equality constraints is discussed in [91], i.e. 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 the Q^2P in x.

$$
(Q^2P_x) \quad \begin{array}{rll} \mu^* := & \min & q_0(x) := x^TQ_0x + 2g_0^Tx + \alpha_0 \\ & s.t. & q_k(x) := x^TQ_kx + 2g_k^Tx + \alpha_k \leq 0 \\ & & k \in \mathcal{I} := \{1, \ldots, m\} \\ & & x \in \Re^n, \end{array} \tag{22}
$$

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

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

(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 are not convex, necessarily. Therefore the feasible set can be a very "nasty" set. This problem is a very hard problem to solve in general, see e.g. [81].

Let

$$P_k := \begin{bmatrix} \alpha_k & g_k^T \\ g_k & Q_k \end{bmatrix} \tag{23}$$

and, by abuse of notation, define

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

Then an equivalent homogenized formulation to (Q^2P_x) is

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

It is clear that the optimal values of the two equivalent formulations are equal. In fact, if $y_0 = -1$ is optimal, then we can replace y by $-y$. This is 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 correct reference will be clear from the context.

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

The Lagrangian Relaxation of a General $\mathbf{Q^2P}$. The Lagrangian relaxation of the homogenized problem $\mathrm{Q^2P}_y$ 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 both the SDP and Lagrangian relaxation are equal. The Lagrangian of $\mathrm{Q^2P}_y$ is

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

The Lagrangian relaxation of $\mathrm{Q^2P}_y$ is

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

Note that

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

from strong duality of the trust region subproblem, see [104]. Therefore, we get equivalence of the dual values for the problems in x and in y. (This is similar to the approaches in [116, 103].)

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

We immediately conclude that *weak duality* holds

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

Therefore, if the optimal σ^*, λ^* can be found, we have found a single quadratic function whose minimal value approximates the original minimal value μ^*, i.e.

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

Moreover, in the dual program, the Lagrangian is a quadratic function of y. Therefore, the outer maximization problem has the nonnegativity and an additional hidden semidefinite constraint

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

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

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

Valid Inequalities. Using the above approach we see that more constraints $q_k(y)$ means that we have a stronger dual. This can be phrased as adding redundant constraints to get new valid inequalities to 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, see [94].

Specific instances of these relaxations (quadratic assignment and graph partitioning problems) appear in [121, 112]. We will present the *recipe* for generating a relaxation below 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 on nonconvex quadratic programs where this holds as well. In this Section 2.3 we amplify on our theme that illustrates the strength of the Lagrangian relaxation, i.e. that a tractable bound implies a Lagrangian relaxation is at work.

Recall the general quadratic constrained quadratic program (22). 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

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

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

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

Thus we have defined our dual problem

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

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

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

i.e. we have found an optimum $\bar{x}$ and have a zero duality gap 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: *necessary conditions for the sufficiency conditions to hold*, i.e. we need stationarity of the Lagrangian and positive semidefiniteness of the Hessian of the Lagrangian. Thus, when these two conditions are incompatible we lose strong duality; we can even expect a duality gap.

We now present several Q^2P 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 where the duality gap may be nonzero. We continue to emphasize our theme that illustrates that the Lagrangian relaxation is best.

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

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

where all $q_i(x)$ are convex quadratic functions. We now see that Lagrangian duality can always solve this problem.

The dual is

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

If ν^* is attained at λ^*, x^*, then a *sufficient* condition for x^* to be optimal for CQP is primal feasibility and complementary slackness, i.e.

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

In addition, it is well known that the Karush-Kuhn-Tucker (KKT) conditions are sufficient for global optimality, and under an appropriate constraint qualification the KKT conditions are also necessary. Therefore strong duality holds if a constraint qualification is satisfied, i.e. in this case there is no duality gap and the dual is attained.

However, surprisingly, *if the primal value of CQP is bounded then it is attained and there is no duality gap*, see e.g. [109, 90, 89, 88]. (This can be considered to be an extension of the Frank-Wolfe Theorem, [68].) However, the dual may not be attained, e.g. consider the convex program

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

and its (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 e.g. [48, 67, 77].

Rayleigh Quotient. Suppose that $A = A^T \in \mathcal{S}^n$. It is well known that the smallest eigenvalue λ_1 of A is obtained from the Rayleigh quotient, i.e.

$$\lambda_1 = \min\{x^T Ax : x^T x = 1\}. \tag{27}$$

Since A is not necessarily positive semidefinite, this is the minimization of a nonconvex function on a nonconvex set. However, the Rayleigh quotient forms the basis for many algorithms for finding the smallest eigenvalue, and these algorithms are very efficient. In fact, it is easy to see that there is no duality gap for this nonconvex problem, i.e.

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

To see this note that the inner minimization problem in (28) is unconstrained. This implies that the outer maximization problem has the hidden semidefinite constraint (an ongoing theme in this paper)

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

i.e. λ is at most the smallest eigenvalue of A. With λ set to the smallest eigenvalue, the inner minimization yields the eigenvector corresponding to λ_1. Thus, we have an example of a *nonconvex problem for which strong duality holds*. Note that the problem (27) has the special norm constraint, and a homogeneous quadratic objective. Thus, this nonconvex problem can be solved efficiently. But, also, strong duality holds for the Lagrangian dual which supports our theme.

Trust Region Subproblem. We will next see that strong duality holds for a larger class of seemingly nonconvex problems. The trust region subproblem, TRS, is the minimization of a quadratic function subject to a norm constraint. No convexity or homogeneity of the objective function is assumed. We allow for a further extension, i.e. we do not assume convexity of the constraint and allow indefinite quadratic functions for both objective and constraint. (See e.g. [22] for applications of indefinite quadratic forms.) This problem is important in nonlinear programming, e.g. [75, 74].

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

or the generalized trust region subproblem [104, 73].

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

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

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

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

This is equivalent to (see [104]) the (concave) nonlinear semidefinite program

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

where $\cdot^\dagger$ denotes Moore-Penrose 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, see example (26) below.) Thus, as in the eigenvalue case, we see that this is an example of a nonconvex program where strong duality holds. In addition, this implies that this problem can be solved efficiently; polynomial time results are presented in [117].

We include a short proof of strong duality, for the inequality constrained case, based on the outline in [62], i.e. we fall back on the convex case 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 can assume that TRS is nonconvex. (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. Then TRS is equivalent to

$$
\begin{array}{rcll}
\mu^* & = & \min\limits_{x^Tx\leq\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^tx + \gamma x^Tx \\
& = & \min\limits_{x^Tx=\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^tx + \gamma x^tx, \quad (Q_0 \text{ is indefinite}) \\
& = & \min\limits_{x^Tx=\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^tx + \gamma\delta^2 \\
& = & \min\limits_{x^Tx\leq\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^tx + \gamma\delta^2, \quad (Q_0-\gamma I \text{ is singular}) \\
& = & \max\limits_{\lambda\geq 0}\min\limits_{x} & x^T(Q_0-\gamma I)x - 2c_0^tx + \lambda(x^Tx-\delta^2) + \gamma\delta^2 \\
& & & \hfill (\text{convex case}) \\
& = & \max\limits_{\lambda\geq 0}\min\limits_{x} & x^TQ_0x - 2c_0^tx + (\lambda-\gamma)(x^Tx-\delta^2) \\
& \leq & \max\limits_{\lambda\geq \gamma}\min\limits_{x} & x^TQ_0x - 2c_0^tx + (\lambda-\gamma)(x^Tx-\delta^2) \quad (\gamma<0) \\
& = & \nu^* \leq \mu^*. &
\end{array}
\tag{29}
$$

■

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]. These algorithms are extremely efficient for the TRS problem, i.e. they solve this problem almost as quickly as an eigenvalue problem.

The fact that we can solve the TRS efficiently even though the objective and constraint may be nonconvex is surprising. In fact, in [70] Martinez shows that the TRS can have at most one local and non-global optimum, and 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 0 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 using the constraint qualification that $\alpha < \beta$. In [73], necessary and sufficient optimality conditions are presented for GTRS using the constraint qualification that $\min q_0(x) < \max q_0(x)$. Using these results in combination with the extension of the Frank-Wolfe result (e.g. [68]) gives us the following.

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

Two Trust Region Subproblem. The two trust region subproblem, TTRS, consists in minimizing a (possibly nonconvex) quadratic function subject to a norm and a least squares constraint, i.e. two convex quadratic constraints. This problem arises in solving general nonlinear programs using a sequential quadratic programming approach, and is often called the CDT problem, see [21].

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

As mentioned above, Martinez [70] shows that the TRS can have at most one local and nonglobal optimum, and the Lagrangian at this point has one negative eigenvalue. Therefore, if we have such a case and add another ball constraint that contains the local, nonglobal, optimum in its interior and also makes this point the global optimum, we obtain a TTRS where we cannot have a zero duality gap due to the negative eigenvalue. It is uncertain what constraints could be added to close this duality gap. In fact, it is still an open problem whether TTRS is an NP-hard or a polynomial time problem.

General $\mathbf{Q^2P}$. The general, possibly nonconvex, Q^2P has many applications in modeling and approximation theory, see e.g. the applications to SQP methods in [58]. Examples of approximations to Q^2P also appear in [32].

The Lagrangian relaxation of a Q^2P is equivalent to the SDP relaxation, and is sometimes referred to as the Shor relaxation, see [103]. The Lagrangian relaxation can be written as an SDP if one takes into the account the hidden semidefinite constraint, i.e. a quadratic function is 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 fact that $x^TQx = \mathrm{Trace}\, x^TQx = \mathrm{Trace}\, Qxx^T$, and relaxing xx^T to a semidefinite matrix X.

One can relate the geometry of the original feasible set of Q^2P with the feasible set of the SDP relaxation. The connection is through *valid quadratic inequalities*, i.e. nonnegative (convex) combinations of the quadratic functions; 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 constraints*

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

(sometimes known as the Stiefel manifold, e.g. [28]) and the trust region type constraint

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

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 results of this section can be viewed as a matrix generalization of the strong duality result for the Rayleigh Quotient problem (27).

Let A and B be $n \times n$ symmetric matrices, and consider the orthonormally constrained homogeneous $\mathrm{Q}^2\mathrm{P}$

$$\mathrm{QQP_O} \qquad \begin{array}{rll} \mu^O := & \min & \operatorname{Trace} AXBX^T \\ & \text{s.t.} & XX^T = I. \end{array} \tag{30}$$

This problem can be solved exactly using Lagrange multipliers, see e.g. [43], or using the classical Hoffman-Wielandt inequality, e.g. [14].

Proposition 2.1 *Suppose that the orthogonal diagonalizations of A, B are $A = V\Sigma V^T$ and $B = U\Lambda U^T$, respectively, where the eigenvalues in Σ are ordered nonincreasing, and the eigenvalues in Λ are ordered nondecreasing. Then the optimal value of* $\mathrm{QQP_O}$ *is $\mu^O = \operatorname{Trace} \Sigma\Lambda$, and the optimal solution is obtained using the orthogonal matrices that yield the diagonalizations, i.e. $X^* = VU^T$.* ∎

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

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

However, there can be a nonzero duality gap for the Lagrangian dual, see [121] for an example. The inner minimization in the dual problem (31) is an unconstrained quadratic minimization in the variables $\operatorname{vec}(X)$ (the vector formed from the columns of the matrix 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 $\mathrm{QQP_O}$, are equivalent to $AXB = SX$ or, by orthogonality, $AXBX^T = S$, which yields mutual diagonalizability of A and XBX^T and the characterization of the optimum. One can easily construct examples where the semidefinite condition and the stationarity are in conflict and result in a duality gap. In order to close the duality gap, we need a larger class of quadratic functions.

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

$$\begin{array}{lrll} \mathrm{QQP_{OO}} & \mu^O := & \min & \operatorname{Trace} AXBX^T \\ & & \text{s.t.} & XX^T = I,\ X^TX = I. \end{array} \qquad (32)$$

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

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

Theorem 2.3 *Strong duality holds for* $\mathrm{QQP_{OO}}$ *and* $\mathrm{DQQP_{OO}}$, *i.e.,* $\mu^D = \mu^O$ *and both primal and dual are attained.* ■

A further relaxation of the above orthogonal relaxation is the trust region relaxation studied in [49]

$$\begin{array}{rll} \mu^*_{QAPT} := & \min & \operatorname{Trace} AXBX^T \\ & \text{s.t.} & XX^T \preceq I. \end{array} \qquad (33)$$

The constraints are convex with respect to the Löwner partial order, i.e. the partial order induced by the cone $\mathcal{P}$. (See (1).) Therefore, it is hoped that solving this problem would be useful. Also, this problem is visually similar to the TRS discussed above. And so we would like to find a characterization of optimality.

The set

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

is studied separately in [80, 30] and is useful in eigenvalue variational principles.

The following generalization of the Hoffman-Wielandt inequality holds, see [5].

Theorem 2.4 *Let $V^TAV = \Sigma$, $U^TBU = \Lambda$, be the orthogonal diagonalizations of A, B, respectively. Let the eigenvalues be in nonincreasing order, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then for any $XX^T \leq I$, we have*

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

The upper bound is attained for $X = V\mathrm{Diag}\,(\epsilon)U^T$, where $\epsilon_i = 1$ if $\sigma_i\lambda_i \geq 0$, and $\epsilon_i = 0$ otherwise. The lower bound is attained for $X = V\mathrm{Diag}\,(\epsilon)JU^T$, where $\epsilon_i = 1$ if $\sigma_i\lambda_{n+1-i} \leq 0$ and $\epsilon_i = 0$ otherwise. The matrix $J = (e_n, e_{n-1}, \cdots, e_1)$, where e_i is the ith elementary unit vector.

■

For a scalar ξ, let $\xi^- := \min\{0, \xi\}$. The lower bound in the above Theorem 2.4 states that $\mu^*_{QAPT} = \sum_{i=1}^{n}[\lambda_i\sigma_{n-i+1}]^-$. 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 constraint (33). Note that XX^T and X^TX have the same eigenvalues, which implies $XX^T \preceq I$ if and only if $X^TX \preceq I$. Explicitly using both sets of constraints, as in [6], we obtain

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

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

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

The following is proved in [5].

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

3. A STRENGTHENED BOUND FOR MC

From the results in Section 2.1 above, it would appear that we might have the strongest possible tractable bound for the max-cut problem MC (or its equivalent quadratic model MCQ). However, adding redundant constraints can strengthen bounds. 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 in final SDP*

We will use the above recipe to strengthen our bound for MC. First, we introduce the following notation for linear operators and adjoints: for $S \in \mathcal{S}^n$, the vector $s = \mathrm{svec}\,(S) \in \Re^{t(n)}$, is formed (columnwise) from S while ignoring the strictly lower triangular part of S. Its inverse is the operator $S = \mathrm{sMat}\,(s)$. The adjoint of svec is the operator $\mathrm{hMat}\,(v)$ which forms a symmetric matrix where the off-diagonal terms are multiplied by a half, i.e. this satisfies

$$\mathrm{svec}\,(S)^T v = \mathrm{Trace}\, S\, \mathrm{hMat}\,(v), \qquad \forall S \in \mathcal{S}^n, v \in \Re^{t(n)},$$

where $t(n) = n(n+1)/2$. The adjoint of sMat is the operator $\mathrm{dsvec}\,(S)$ which works like svec except that the off diagonal elements are multiplied by 2, i.e. this satisfies

$$\mathrm{dsvec}\,(S)^T v = \mathrm{Trace}\, S\, \mathrm{sMat}\,(v), \qquad \forall S \in \mathcal{S}^n, v \in \Re^{t(n)}.$$

For notational convenience, we define the vectors

$$\mathrm{sdiag}\,(s) := \mathrm{diag}\,(\mathrm{sMat}\,(s))$$

and

$$\mathrm{vsMat}\,(s) = \mathrm{vec}\,(\mathrm{sMat}\,(s));$$

the adjoint of vsMat is then given by

$$\mathrm{vsMat}^*(s) = \mathrm{dsvec}\,\left(\left(\mathrm{Mat}\,(v) + \mathrm{Mat}\,(v)^T\right)/2\right).$$

The following bound is motivated by the strong duality results presented in Section 2.3. (It is studied in depth in [6]. We refer the reader there for missing details.) The SDP bound (20) for MCQ arises from a lifting procedure, i.e. identifying

$$0 \preceq X = xx^T \text{ and } x^TQx = \operatorname{Trace} QX.$$

Discarding the rank one condition on X results in the tractable SDP bound. It is not clear what constraints one can add to P_E, (21), in order to strengthen the Lagrangian relaxation, i.e. 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), i.e. one can try and add the so-called triangle inequalities which are the standard inequalities used in branch and bound methods for MC. This has proven to be very successful in practice, e.g. [47]. However, one cannot guarantee an improvement, [50]. But, in the space of matrices, it is also true that

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

a constraint of type (32) and (33). Therefore we can use the following equivalent quadratic matrix model for MCQ.

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

where X is a symmetric matrix. (Note that a common diagonalization of X and X^2 shows that 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) and start with, MC2, an equivalent program to (34) and so to MC.

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

To efficiently apply Lagrangian relaxation and not lose information from the linear constraint, we replace the constraint with the norm constraint $||\operatorname{diag}(X) - e||^2 = 0$. Now homogenize by adding the variable y_0 and constraint $1 - y_0^2 = 0$. (Because of the homogenization, if $x, y_0 = -1$ is optimal, then we can just multiply both x and y by -1 without changing the optimal value or feasibility.) We use $X = \operatorname{sMat}(x)$ is a symmetric matrix and get the following equivalent program.

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

In fact, if we let $z = \begin{pmatrix} y_0 \\ x \end{pmatrix}$, then we see that this is itself a max-cut problem with the same objective function as MCQ, additional linear constraints $\mathrm{sdiag}\,x = e$ (these nodes must be grouped together), and extra nonlinear constraints given by $\mathrm{sMat}\,(x)^2 - n\,\mathrm{sMat}\,(x)y_0 = 0$.

As discussed in Remark 1.2, the Lagrange multipliers for symmetric matrix valued functions will be symmetric matrices. We now take the Lagrangian dual with the appropriate Lagrange multipliers w, T, S.

$$
\begin{array}{l}
\mu^* \leq \\
\nu_2^* := \min\limits_{w,T,S}\ \max\limits_{x, y_0^2=1}\ \text{Trace }(Q\mathrm{sMat}\,(x))\,y_0 \\
\qquad + w(\mathrm{sdiag}\,(x)^T\mathrm{sdiag}\,(x) - 2e^T\mathrm{sdiag}\,(x)y_0 + n) \\
\qquad\quad + \text{Trace}\,T(E - \mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x)) \\
\qquad\quad + \text{Trace}\,S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x)y_0).
\end{array}
$$

We can move the variable y_0 into the Lagrangian without increasing the duality gap since this is a trust region subproblem.

$$
\begin{array}{rl}
\nu_2^* = & \min\limits_{t,w,S}\ \max\limits_{x,y_0} \text{Trace }(Q\mathrm{sMat}\,(x))\,y_0 \\
& + w(\mathrm{sdiag}\,(x)^T\mathrm{sdiag}\,(x) - 2e^T\mathrm{sdiag}\,(x)y_0 + n) \\
& + \text{Trace}\,T(\mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x) - E) \\
& + \text{Trace}\,S((\mathrm{sMat}\,(x))^2 - n\,\mathrm{sMat}\,(x)y_0) \\
& + t(1 - y_0^2).
\end{array}
$$

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

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

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

We now illustrate and manipulate some of the other linear operators that appear in the Lagrangian. This facilitates the derivation of the Hessian and the adjoints.

The matrix sdiag *sdiag $\in \mathcal{S}^{t(n)}$ is diagonal with elements determined using

$$\begin{array}{rcl} e_i^T\,(\mathrm{sdiag}\,^*\mathrm{sdiag}\,)\,e_j &=& \mathrm{sdiag}\,(e_i)^T\mathrm{sdiag}\,(e_j) \\ &=& \left\{\begin{array}{ll} 1 & \text{if i=j=t(k)} \\ 0 & \text{otherwise.} \end{array}\right. \end{array} \tag{37}$$

Similarly, we find that, for $T = \sum_{ij} t_{ij}E_{ij}$, where the matrices E_{ij} are the elementary matrices $e_ie_j^T + e_je_i^T$, we have

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

Also

$$\mathrm{dsvec}\,\mathrm{Diag}\,\mathrm{diag}\,\mathrm{sMat}\, = \mathrm{sdiag}\,^*\mathrm{sdiag}\, = \mathrm{Diag}\,\mathrm{svec}\,(I). \tag{39}$$

We can use this to rewrite the quadratic forms in the Lagrangian as follows:

$$\begin{array}{rcl} \mathrm{sdiag}\,(x)^T\mathrm{sdiag}\,(x) &=& x^T\,(\mathrm{dsvec}\,\mathrm{Diag}\,\mathrm{diag}\,\mathrm{sMat}\,)\,x; \\ e^T\mathrm{sdiag}\,(x)y_0 &=& (\mathrm{dsvec}\,\mathrm{Diag}\,e)\,xy_0; \end{array} \tag{40}$$

$$\begin{array}{rcl} \mathrm{Trace}\,T(\mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x)) &=& x^T\,\{\mathrm{dsvec}\;(T \circ \mathrm{sMat}\,(x))\} \\ &=& x^T\,(\mathrm{dsvec}\;(T \circ \mathrm{sMat}\,))\,x. \end{array} \tag{41}$$

$$\begin{array}{rcl} \mathrm{Trace}\,S(\mathrm{sMat}\,(x))^2 &=& \mathrm{Trace}\,\mathrm{sMat}\,(x)\,S\,\mathrm{sMat}\,(x) \\ &=& x^T\mathrm{dsvec}\;(S\mathrm{sMat}\,(x)) \\ &=& x^T\,(\mathrm{dsvec}\,S\mathrm{sMat}\,)\,x. \end{array} \tag{42}$$

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. We split it into four linear operators with the factor 2:

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

After cancelling the 2, we get the (equivalent to the Lagrangian dual) semidefinite program

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

If we take T sufficiently positive definite and t sufficiently large, then we can guarantee Slater's constraint qualification. Therefore, there is no duality gap between this SDP and its dual, which is the strengthened SDP relaxation of MC.

$$
\text{MCPSDP2} \qquad \begin{array}{rcc} \nu_2^* = & \max & \operatorname{Trace} H_c Y \\ & \text{s.t.} & \mathcal{H}_1^*(Y) = n \\ & & \mathcal{H}_2^*(Y) = E \\ & & \mathcal{H}_3^*(Y) = 0 \\ & & \mathcal{H}_4^*(Y) = 1 \\ & & Y \succeq 0. \end{array}
$$

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

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

From (41) and

$$
e_i^T \left(\operatorname{dsvec}\ (T \circ \operatorname{sMat}\,)\right) e_j = \left(\operatorname{svec}\,(T) \circ e_j\right)^T e_i,
$$

we get

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

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

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

Also, $\mathcal{H}_1^*(Y)$ is twice the sum of the elements in the first row of Y corresponding to the positions of the diagonal of $\operatorname{sMat}(x)$ minus the sum of the same elements in the diagonal of $\bar{Y}$, i.e.

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

The constraint $\mathcal{H}_1^*(Y) = n$ effectively requires that

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

This comes from the $\operatorname{diag} X = e$ constraint.

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 in the quadratic form in (42), i.e. this block can be seen in two ways: as a quadratic form or as its $t(n) \times t(n)$ symmetric matrix representation $\operatorname{dsvec} S \operatorname{sMat}$. Thus $\operatorname{dsvec} \cdot \operatorname{sMat}$ is a linear operator that maps $n \times n$ symmetric matrices to

$t(n) \times t(n)$ symmetric matrices. We now use this and the equation that defines the adjoint.

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

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

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

Similar to the above expressions for $\mathcal{H}_3$, $\mathrm{sMat}\, \cdot \mathrm{vsMat}\,{}^*\mathrm{vec}$ is a linear operator that maps $t(n) \times t(n)$ symmetric matrices to $n \times n$ symmetric matrices. The $n \times n$ matrix $\mathrm{sMat}\,\bar{Y}\mathrm{vsMat}\,{}^*\mathrm{vec}$ can also be seen as the quadratic form on $v \in \Re^n$

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

The k,l element of this quadratic form can be found using

$$e_k^t\mathrm{sMat}\,\bar{Y}\mathrm{vsMat}\,{}^*\mathrm{vec}\, e_l = \left\langle \mathrm{dsvec}\, \left(\frac{e_k e_l^T + e_l e_K^T}{2}\right) \mathrm{sMat}\,, \bar{Y} \right\rangle .$$

Equivalently, $\mathcal{H}_3^*(Y)$ consists of the sums in MCPSDP3 below. This results in this simplified SDP relaxation MCPSDP3, where we have removed many of the redundant constraints. This problem has $2t(n) - 1$ constraints. The space has increased to $S^{t(n)+1}$, from $\mathcal{S}^n$ in (20). The constraints come from (44), (45), and the upper triangular part of (46). Note that this latter constraint can be obtained by looking at the original motivation $\mathrm{sMat}\,(x) = \frac{1}{n}\mathrm{sMat}\,(x)\mathrm{sMat}\,(x)$, and using the indices that correspond to $\bar{Y} = xx^T$. In comparison to the original SDP relaxation MCSDP, (20), we still have the diagonal all ones constraint. However, we now have additional constraints on the first row of Y.

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

Remark 3.1 *In addition, since the first row has elements equal to 1, and the diagonal is all 1, the Slater constraint qualification fails for this problem, i.e. we cannot have a positive definite feasible matrix. This means that 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 satisfies Slater's condition. (See [3] for the details.)*

The constraint $\mathcal{H}_3^*(Y) = 0$ is the key to showing the following useful result, i.e. for Y feasible in MCPSDP2 (or equivalently MCPSDP3), with first row x, (see (43)) sMat (x) is always positive semidefinite (and in fact feasible for MCSDP) even though this nonlinear constraint was discarded in MC2. We include a proof to illustrate the usefulness of using adjoints.

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

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

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

Proof. For Y feasible for MCPSDP3, write

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

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

We have previously established that

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

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

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

and thus 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 $X, \bar{X}$ are both feasible for MCSDP. Then*

$$\mathrm{Trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) \geq 0. \qquad (47)$$

Suppose, in addition, that both

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

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

$$\mathrm{Trace}\,(X^2 - nX)(\bar{X}^2 - n\bar{X}) > 0. \qquad (48)$$

Proof. See [3]. ■

The above two lemmas can be used to show the strength of this new bound, i.e. unless there is no gap between MCSDP and MC, the relaxation MCPSDP3 **always** provides a strict improvement over MCSDP. (The proof is given in [3].)

Theorem 3.1 *The optimal values satisfy*

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

■

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. A new primal-dual interior-point method for semidefinite programming. In J.G. Lewis, editor, *Proceedings of the Fifth SIAM Conference on Applied Linear Algebra*, pages 113–117. SIAM, 1994.

[2] F. ALIZADEH, J-P.A. HAEBERLY, and M.L. OVERTON. Complementarity and nondegeneracy in semidefinite programming. *Math. Programming*, 77:111–128, 1997.

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

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

[5] K.M. ANSTREICHER, X. CHEN, H. WOLKOWICZ, and Y. YUAN. Strong duality for a trust-region type relaxation of QAP. *Linear Algebra Appl.*, 301:121–136, 1999.

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

[7] E. BALAS. A modified lift-and-project procedure. *Math. Programming*, 79(1-3, Ser. B):19–31, 1997. Lectures on Mathematical Programming (ismp97) (Lausanne, 1997).

[8] E. BALAS and E. ZEMEL. Lifting and complementing yields all the facets of positive zero-one programming polytopes. In *Mathematical programming (Rio de Janeiro, 1981)*, pages 13–24. North-Holland, Amsterdam, 1984.

[9] G. P. BARKER, M. LAIDACKER, and G. POOLE. Projectionally exposed cones. *SIAM J. Algebraic Discrete Methods*, 8(1):100–105, 1987.

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

[11] R. BELLMAN and K. FAN. On systems of linear inequalities in Hermitian matrix variables. In *Proceedings of Symposia in Pure Mathematics, Vol 7, AMS*, 1963.

[12] A. BEN-ISRAEL, A. CHARNES, and K. KORTANEK. Duality and asymptotic solvability over cones. *Bulletin of American Mathematical Society*, 75(2):318–324, 1969.

[13] S. J. BENSON, Y. YE, and X. ZHANG. Solving large-scale sparse semidefinite programs for combinatorial optimization. Technical report, The University of Iowa, 1997. To appear in SIAM J. Optim.

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

[15] F. BOHNENBLUST. Joint positiveness of matrices. Manuscript, 1948.

[16] B. BORCHERS. SDPLIB 1.2, a library of semidefinite programming test problems. Technical report, New Mexico Tech, Soccorrow, NM, 1999. To appear in Optim. Methods Softw.

[17] J.M. BORWEIN and H. WOLKOWICZ. Characterization of optimality for the abstract convex program with finite-dimensional range. *J. Austral. Math. Soc. Ser. A*, 30(4):390–411, 1980/81.

[18] J.M. BORWEIN and H. WOLKOWICZ. Regularizing the abstract convex program. *J. Math. Anal. Appl.*, 83(2):495–530, 1981.

[19] J.M. BORWEIN and H. WOLKOWICZ. Characterizations of optimality without constraint qualification for the abstract convex program. *Math. Programming Stud.*, 19:77–100, 1982. Optimality and stability in mathematical programming.

[20] S. BOYD, V. BALAKRISHNAN, E. FERON, and L. El GHAOUI. Control system analysis and synthesis via linear matrix inequalities. *Proc. ACC*, pages 2147–2154, 1993.

[21] M.R. CELIS, J.E. DENNIS Jr., and R.A. TAPIA. A trust region strategy for nonlinear equality constrained optimization. In *Proceedings of the SIAM Conference on Numerical Optimization, Boulder, CO*, 1984. Also available as Technical Report TR84-1, Department of Mathematical Sciences, Rice University, Houston, TX.

[22] S. CHANDRASEKARAN, M. GU, and A. H. SAYED. The sparse basis problem and multilinear algebra. *SIAM J. Matrix Anal. Appl.*, 20(2):354–362, 1999.

[23] B. CRAVEN and J.J. KOLIHA. Generalizations of Farkas' thoerem. *SIAM J. Matrix Anal. Appl.*, 8:983–997, 1977.

[24] B. CRAVEN and B. MOND. Linear programming with matrix variables. *Linear Algebra Appl.*, 38:73–80, 1981.

[25] J.-P. CROUZEIX, J.-E. MARTINEZ-LEGAZ, and A. SEEGER. An alternative theorem for quadratic forms and extensions. *Linear Algebra Appl.*, 215:121–134, 1995.

[26] DEM'YANOV and MALOZEMOV. *Introduction to Minimax.* Keter Publishing House, Jerusalem, 1974.

[27] H. DYM and I. GOHBERG. Extensions of band matrices with band inverses. *Linear Algebra Appl.*, 36:1–24, 1981.

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

[29] A.V. FIACCO. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, volume 165 of *Mathematics in Science and Engineering.* Academic Press, 1983.

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

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

[32] M. FU, Z. LUO, and Y. YE. Approximation algorithms for quadratic programming. *J. Comb. Optim.*, 2(1):29–50, 1998.

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

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

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

[36] A. J. GOLDMAN and A. W. TUCKER. Theory of linear programming. In *Linear inequalities and related systems*, pages 53–97. Princeton University Press, Princeton, N.J., 1956. Annals of Mathematics Studies, no. 38.

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

[38] M. GRÖTSCHEL, L. LOVÁSZ, and A. SCHRIJVER. *Geometric Algorithms and Combinatorial Optimization.* Springer Verlag, Berlin-Heidelberg, 1988.

[39] B. GRÜNBAUM. On the facial structure of convex polytopes. *Bull. Amer. Math. Soc.*, 71:559–560, 1965.

[40] B. GRÜNBAUM. *Convex polytopes.* Interscience Publishers John Wiley & Sons, Inc., New York, 1967. With the cooperation of Victor Klee, M. A. Perles and G. C. Shephard. Pure and Applied Mathematics, Vol. 16.

[41] M. GUIGNARD. Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. *SIAM J. of Control*, 7:232–241, 1969.

[42] O. GÜLER and L. TUNCEL. Characterization of the barrier parameter of homogeneous convex cones. *Math. Programming*, 81:55–76, 1998.

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

[44] R. HAUSER. Self-scaled barrier functions: Decomposition and classification. Technical Report DAMTP 1999/NA13, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England, 1998.

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

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

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

[48] F. JARRE. On the convergence of the method of analytic centers when applied to convex quadratic programs. *Math. Programming*, 49(3):341–358, 1990.

[49] S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Trust regions and relaxations for the quadratic assignment problem. In

Quadratic assignment and related problems (New Brunswick, NJ, 1993), pages 199–219. Amer. Math. Soc., Providence, RI, 1994.

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

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

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

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

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

[55] K. O. KORTANEK and J. P. EVANS. Asymptotic Lagrange regularity for pseudoconcave programming with weak constraint qualification. *Operations Res.*, 16:849–857, 1968.

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

[57] S. KRUK, M. MURAMATSU, F. RENDL, R.J. VANDERBEI, and H. WOLKOWICZ. The Gauss-Newton direction in linear and semidefinite programming. Technical Report CORR 98-16, University of Waterloo, Waterloo, Canada, 1998.

[58] S. KRUK and H. WOLKOWICZ. SQ^2P, sequential quadratic constrained quadratic programming. In Ya xiang Yuan, editor, *Advances in Nonlinear Programming*, volume 14 of *Applied Optimization*, pages 177–204. Kluwer, Dordrecht, 1998. Proceedings of Nonlinear Programming Conference in Beijing in honour of Professor M.J.D. Powell.

[59] S. KRUK and H. WOLKOWICZ. Sequential, quadratic constrained, quadratic programming for general nonlinear programming. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications*, page 11. Kluwer Academic Publishers, Boston, MA, 2000.

[60] M. LAURENT. A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems. In *Topics in Semidefinite and Interior-Point Methods*, volume 18 of *The Fields*

Institute for Research in Mathematical Sciences, Communications Series, Providence, Rhode Island, 1998. American Mathematical Society.

[61] C. LEMARÉCHAL and F. OUSTRY. Nonsmooth algorithms to solve semidefinite programs. In L. EL Ghaoui and S-I. Niculescu, editors, *Recent Advances on LMI methods in Control*, Advances in Design and Control series. SIAM, 1999. To appear.

[62] A.S. LEWIS. *Take-home final exam, Course CO663 in Convex Analysis*. University of Waterloo, Ontario, Canada, 1994.

[63] A.S. LEWIS. Derivatives of spectral functions. *Math. Oper. Res.*, 21(3):576–588, 1996.

[64] A.S. LEWIS and M.L. OVERTON. Eigenvalue optimization. *Acta Numerica*, 5:149–190, 1996.

[65] L. LOVÁSZ. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:1–7, 1979.

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

[67] Z-Q. LUO and J. SUN. An analytic center based column generation algorithm for convex quadratic feasibility problems. Technical report, McMaster University, Hamilton, Ontario, 1995.

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

[69] A.W. MARSHALL and I. OLKIN. *Inequalities: Theory of Majorization and its Applications*. Academic Press, New York, NY, 1979.

[70] J.M. MARTINEZ. Local minimizers of quadratic functions on Euclidean balls and spheres. *SIAM J. Optim.*, 4(1):159–176, 1994.

[71] V.M. MATROSOV and A.I. MALIKOV. The development of the ideas of A. M. Lyapunov over one hundred years: 1892–1992. *Izv. Vyssh. Uchebn. Zaved. Mat.*, 4:3–47, 1993.

[72] R.D.C. MONTEIRO and Y. ZHANG. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Math. Programming*, 81(3, Ser. A):281–299, 1998.

[73] J. J. MORÉ. Generalizations of the trust region problem. *Optim. Methods Software*, 2:189–209, 1993.

[74] J.J. MORÉ and D.C. SORENSEN. Computing a trust region step. *SIAM J. Sci. Statist. Comput.*, 4:553–572, 1983.

[75] J.J. MORÉ and D.C. SORENSEN. Newton's method. In Gene H. Golub, editor, *Studies in Numerical Analysis*, volume 24 of *MAA Studies in Mathematics*. The Mathematical Association of America, 1984.

[76] Y.E. NESTEROV and A.S. NEMIROVSKI. Polynomial barrier methods in convex programming. *Èkonom. i Mat. Metody*, 24(6):1084–1091, 1988.

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

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

[79] J. W. NIEUWENHUIS. Another application of Guignard's generalized Kuhn-Tucker conditions. *J. Optim. Theory Appl.*, 30(1):117–125, 1980.

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

[81] P.M. PARDALOS. Quadratic programming with one negative eigenvalue is NP-hard. *J. Global Optim.*, 1:15–22, 1991.

[82] P. C. PARKS. A. M. Lyapunov's stability theory—100 years on. *IMA J. Math. Control Inform.*, 9(4):275–303, 1992.

[83] G. PATAKI. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358, 1998.

[84] G. PATAKI. Geometry of Semidefinite Programming. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications*. Kluwer Academic Publishers, Boston, MA, 2000.

[85] G. PATAKI and L. TUNCEL. On the generic properties of convex optimization problems in conic form. Technical Report CORR

97-16, Department of Combinatorics and Optimization, Waterloo, Ont, 1997.

[86] J. PENG and Y. YUAN. Optimality conditions for the minimization of a quadratic with two quadratic constraints. *SIAM J. Optim.*, 7(3):579–594, 1997.

[87] A.L. PERESSINI. *Ordered Topological Vector Spaces.* Harper & Row Publishers, New York, 1967.

[88] E.L. PETERSON and J.G. ECKER. Geometric programming: duality in quadratic programming and l_p-approximation. II. (Canonical programs). *SIAM J. Appl. Math.*, 17:317–340, 1969.

[89] E.L. PETERSON and J.G. ECKER. Geometric programming: duality in quadratic programming and l_p-approximation. I. In *Proceedings of the Princeton Symposium on Mathematical Programming (Princeton Univ., 1967)*, pages 445–480, Princeton, N.J., 1970. Princeton Univ. Press.

[90] E.L. PETERSON and J.G. ECKER. Geometric programming: duality in quadratic programming and l_p-approximation. III. Degenerate programs. *J. Math. Anal. Appl.*, 29:365–383, 1970.

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

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

[93] G.D. POOLE and M. LAIDACKER. Projectionally exposed cones in $\mathbf{r}^3$. *Linear Algebra Appl.*, 111:183–190, 1988.

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

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

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

[97] M.V. RAMANA. An exact duality theory for semidefinite programming and its complexity implications. *Math. Programming*, 77:129–162, 1997.

[98] F. RENDL and H. WOLKOWICZ. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Math. Programming*, 77(2, Ser. B):273–299, 1997.

[99] S.A. SANTOS and D.C. SORENSEN. A new matrix-free algorithm for the large-scale trust-region subproblem. Technical Report TR95-20, Rice University, Houston, TX, 1995.

[100] A. SHAPIRO. First and second order analysis of nonlinear semidefinite programs. *Math. Programming*, 77:301–320, 1997.

[101] A. SHAPIRO. Duality and optimality conditions. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications*. Kluwer Academic Publishers, Boston, MA, 2000.

[102] H.D. SHERALI and W.P. ADAMS. Computational advances using the reformulation-linearization technique (rlt) to solve discrete and continuous nonconvex problems. *Optima*, 49:1–6, 1996.

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

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

[105] J.F. STURM. *Primal-Dual Interior Point Approach to Semidefinite Programming.* PhD thesis, Erasmus University Rotterdam, 1997.

[106] J.F. STURM. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Technical report, Communications Research Laboratory, McMaster University, Hamilton, Canada, 1998. To appear in *Optimization Methods and Software.*

[107] C. H. SUNG and B.-S. TAM. A study of projectionally exposed cones. *Linear Algebra Appl.*, 139:225–252, 1990.

[108] O. TAUSSKY. Positive-definite matrices and their role in the study of the characteristic roots of general matrices. *Advances in Math.*, 2:175–186 (1968), 1968.

[109] T. TERLAKY. On l_p programming. *European J. Oper. Res.*, 22:70–100, 1985.

[110] M.J. TODD. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Optim. Methods Softw.*, 11&12:1–46, 1999.

[111] L. TUNCEL and S. XU. On homogeneous convex cones, Caratheodory number, and duality mapping semidefinite liftings.

Technical Report CORR 99-21, Department of Combinatorics and Optimization, Waterloo, Ont, 1999.

[112] H. WOLKOWICZ and Q. ZHAO. Semidefinite relaxations for the graph partitioning problem. *Discrete Applied Math.*, 96-97(1-3):461–479, 1999.

[113] S. WRIGHT. Recent developments in interior-point methods. In M.J.D. Powell, editor, *Proceedings of 19th IFIP TC7 Conference on System Modelling and Optimization, July, 1999, Cambridge.* Kluwer Academic Publishers, Boston, MA, 2000. To appear.

[114] V.A. YAKUBOVICH. The method of matrix inequalities in the stability theory of nonlinear control systems, *I*, *II*, *III*. *Automation and Remote Control*, 25-26(4):905–917, 577–592, 753–763, April 1967.

[115] V.A. YAKUBOVICH. 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):1039–1042, 1973.

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

[117] Y. YE. A new complexity result on minimization of a quadratic function with a sphere constraint. In *Recent Advances in Global Optimization*, pages 19–31. Princeton University Press, 1992.

[118] Y. YUAN. Some properties of trust region algorithms for nonsmooth optimization. Technical Report DAMTP 1983/NA14, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England, 1983.

[119] Y. YUAN. On a subproblem of trust region algorithms for constrained optimization. *Math. Programming*, 47:53–63, 1990.

[120] Y. YUAN. A dual algorithm for minimizing a quadratic function with two quadratic constraints. *Journal of Computational Mathematics*, 9:348–359, 1991.

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

[122] S. ZLOBEC. Optimality conditions for mathematical programming problems. *Glasnik Mat. Ser. III*, 6(26):187–192, 1971.