SIAM J. MATRIX ANAL. APPL.
Vol. 22, No. 1, pp. 41–55

ON LAGRANGIAN RELAXATION OF QUADRATIC MATRIX CONSTRAINTS*

KURT ANSTREICHER† AND HENRY WOLKOWICZ‡

Abstract. Quadratically constrained quadratic programs (QQPs) play an important modeling role for many diverse problems. These problems are in general NP hard and numerically intractable. Lagrangian relaxations often provide good approximate solutions to these hard problems. Such relaxations are equivalent to semidefinite programming relaxations.

For several special cases of QQP, e.g., convex programs and trust region subproblems, the Lagrangian relaxation provides the exact optimal value, i.e., there is a zero duality gap. However, this is not true for the general QQP, or even the QQP with two convex constraints, but a nonconvex objective.

In this paper we consider a certain QQP where the quadratic constraints correspond to the matrix orthogonality condition $XX^T = I$. For this problem we show that the Lagrangian dual based on relaxing the constraints $XX^T = I$ *and* the seemingly redundant constraints $X^TX = I$ has a zero duality gap. This result has natural applications to quadratic assignment and graph partitioning problems, as well as the problem of minimizing the weighted sum of the largest eigenvalues of a matrix. We also show that the technique of relaxing quadratic matrix constraints can be used to obtain a strengthened semidefinite relaxation for the max-cut problem.

Key words. Lagrangian relaxations, quadratically constrained quadratic programs, semidefinite programming, quadratic assignment, graph partitioning, max-cut problems

AMS subject classifications. 49M40, 52A41, 90C20, 90C27

PII. S0895479898340299

1. Introduction.

Quadratically constrained quadratic programs (QQPs) play an important modeling role for many diverse problems. They often provide a much improved model compared to the simpler linear relaxation of a problem. However, very large linear models can be solved efficiently, whereas QQPs are in general NP-hard and numerically intractable. Lagrangian relaxations often provide good approximate solutions to these hard problems. Moreover these relaxations can be shown to be equivalent to semidefinite programming (SDP) relaxations, and SDP problems can be solved efficiently, i.e., they are polynomial time problems; see, e.g., [31].

SDP relaxations provide a tractable approach for finding good bounds for many hard combinatorial problems. The best example is the application of SDP to the max-cut problem, where a 87% performance guarantee exists [11, 12]. Other examples include matrix completion problems [23, 22], as well as graph partitioning problems and the quadratic assignment problem (references given below).

In this paper we consider several quadratically constrained quadratic (nonconvex) programs arising from hard combinatorial problems. In particular, we look at the orthogonal relaxations of the quadratic assignment and graph partitioning problems. We show that the resulting well-known eigenvalue bounds for these problems can be obtained from the Lagrangian dual of the orthogonally constrained relaxations,

*Received by the editors June 9, 1998; accepted for publication (in revised form) by P. Van Dooren July 30, 1999; published electronically May 31, 2000.

http://www.siam.org/journals/simax/22-1/34029.html

†Department of Management Sciences, University of Iowa, Iowa City, IA 52242-1000 (kurt-anstreicher@uiowa.edu).

‡University of Waterloo, Department of Combinatorics and Optimization, Waterloo, Ontario N2L 3G1, Canada (henry@orion.uwaterloo.ca). This author's research was supported by Natural Sciences and Engineering Research Council of Canada.

but only if the seemingly redundant constraint $X^TX = I$ is explicitly added to the orthogonality constraint $XX^T = I$. Our main analytical tool is a strong duality result for a certain nonconvex QQP, where the quadratic constraints correspond to the orthogonality conditions $XX^T = I$, $X^TX = I$. We also show that the technique of applying Lagrangian relaxation to quadratic matrix constraints can be used to obtain a strengthened SDP relaxation for the max-cut problem.

Our results show that current tractable (nonconvex) relaxations for the quadratic assignment and graph partitioning problems can, in fact, be found using Lagrangian relaxations. (A converse statement is well known, i.e., the Lagrangian dual is equivalent to an (tractable) SDP relaxation.) Our results here provide further evidence to the following conjecture: *the Lagrangian relaxation of an appropriate QQP provides the strongest tractable relaxation for QQPs.*

1.1. Outline. We complete this section with the notation used in this paper.

In section 2, we present several known results on QQPs. We start with convex QQPs where a zero duality gap always holds. Then we look at the minimum eigenvalue problem and the trust region subproblem, where strong duality continues to hold. We conclude with the two trust region subproblem, the max-cut problem, and general nonconvex QQPs where nonzero duality gaps can occur.

The main results are in section 3. We show that strong duality holds for a class of orthogonally constrained quadratic programs if we add seemingly redundant constraints before constructing the Lagrangian dual.

In section 4 we apply this result to several problems, i.e., relaxations of quadratic assignment and graph partitioning problems, and a weighted sum of eigenvalue problem. In section 5 we present strengthened semidefinite relaxations for the max-cut problem. In section 6 we summarize our results and describe some promising directions for future research.

1.2. Notation. We now describe the notation used in the paper.

Let $\mathcal{S}_n$ denote the space of $n \times n$ symmetric matrices equipped with the trace inner product, $\langle A, B \rangle = \operatorname{tr} AB$, and let $A \succeq 0$ (resp., $A \succ 0$) denote positive semidefiniteness (resp., positive definiteness) and $A \succeq B$ denote $A - B \succeq 0$, i.e., $\mathcal{S}_n$ is equipped with the Löwner partial order. We let $\mathcal{P}$ denote the cone of symmetric positive semidefinite matrices; $\mathcal{M}_{m,n}$ denotes the space of general $m \times n$ matrices also equipped with the trace inner product, $\langle A, B \rangle = \operatorname{tr} A^TB$, while $\mathcal{M}_m$ denotes the space of general $m \times m$ matrices; $\mathcal{O}$ denotes the set of orthonormal (orthogonal) matrices; Π denotes the set of permutation matrices.

We let $\operatorname{Diag}(v)$ be the diagonal matrix formed from the vector v; its adjoint operator is $\operatorname{diag}(M)$, which is the vector formed from the diagonal of the matrix M. For $M \in \mathcal{M}_{m,n}$, the vector $m = \operatorname{vec}(M) \in \Re^{mn}$ is formed (columnwise) from M.

The Kronecker product of two matrices is denoted $A \otimes B$, and the Hadamard product is denoted $A \circ B$.

We use e to denote the vector of all ones, and $E = ee^T$ to denote the matrix of all ones.

2. Some known results. The general QQP is

$$\text{QQP} \qquad \begin{array}{l} \min q_0(x) \\ \text{s.t. } q_k(x) \leq 0 \text{ (or } = 0), \quad k = 1, \ldots, m, \end{array}$$

where $q_i(x) = x^TQ_ix - 2g_i^Tx$. We now present several QQP 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.

2.1. Convex quadratic programs. Consider the convex quadratic program

$$\text{CQP} \qquad \mu^* := \min q_0(x) \quad \text{s.t. } q_k(x) \le 0, \quad k = 1, \ldots, m,$$

where all $q_i(x)$ are convex quadratic functions. The dual is

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

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., 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., [44, 36, 34, 35] and, more recently, [26]. However, the dual may not be attained, e.g., consider the convex program

$$0 = \min\{x : x^2 \le 0\}$$

and its dual

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

Algorithmic approaches based on Lagrangian duality appear in, e.g., [19, 25, 31].

2.2. 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.,

$$\lambda_1 = \min\{x^T A x : x^T x = 1\}. \tag{2.1}$$

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.,

$$\lambda_1 = \max_\lambda \min_x \; x^T A x - \lambda(x^T x - 1). \tag{2.2}$$

To see this, note that the inner minimization problem in (2.2) is unconstrained. This implies that the outer maximization problem has the hidden semidefinite constraint (an ongoing theme in the 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 (2.1) has the special norm constraint and a homogeneous quadratic objective.

2.3. 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.

$$\text{TRS} \qquad \mu^* := \min q_0(x) \quad \text{s.t. } x^T x - \delta^2 \le 0 \ (\text{or } = 0).$$

Assuming that the constraint in TRS is written "≤," the Lagrangian dual is

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

This is equivalent to (see [43]) the (concave) nonlinear semidefinite program

$$\begin{aligned} \text{DTRS} \qquad \nu^* := \max\, & g_0^T (Q + \lambda I)^\dagger g_0 - \lambda \delta^2 \\ \text{s.t. } & Q + \lambda I \succeq 0, \\ & \lambda \ge 0. \end{aligned}$$

where $\cdot^\dagger$ denotes Moore–Penrose inverse. It is shown in [43] that strong duality holds for TRS, i.e., there is a zero duality gap $\mu^* = \nu^*$, and both the primal and dual are attained. Thus, as in the eigenvalue case, we see that this is an example of a nonconvex program where strong duality holds.

Extensions of this result to a two-sided general, possibly nonconvex constraint are discussed in [43, 28]. An algorithm based on Lagrangian duality appears in [40] and (implicitly) in [29, 41]. These algorithms are extremely efficient for the TRS problem, i.e., they solve this problem almost as quickly as they can solve an eigenvalue problem.

2.4. Two trust region subproblem. The two trust region subproblem (TTRS) consists of 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 Celis–Dennis–Tapia (CDT) problem; see [4].

In contrast to the above single TRS, the TTRS can have a nonzero duality gap; see, e.g., [33, 47, 48, 49]. This is closely related to quadratic theorems of the alternative, e.g., [5]. In addition, if the constraints are not convex, then the primal may not be attained; see, e.g., [26].

In [27], Martinez 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 close the 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.

2.5. Max-cut problem. Suppose that $G = (V, E)$ is an undirected graph with vertex set $V = \{v_i\}_{i=1}^n$ and weights w_{ij} on the edges $(v_i, v_j) \in E$. The *max-cut problem* consists of finding the index set $\mathcal{I} \subset \{1, 2, \ldots, n\}$, in order 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 the following discrete optimization problem with a quadratic objective:

$$\text{MC} \qquad \max \ \tfrac{1}{2} \textstyle\sum_{i<j} w_{ij}(1 - x_i x_j), \quad x \in \{\pm 1\}^n.$$

We equate $x_i = 1$ with $i \in \mathcal{I}$ and $x_i = -1$ otherwise. Define the homogeneous quadratic objective

$$q(x) := x^T Qx,$$

where Q is an $n \times n$ symmetric matrix. Then the MC problem is equivalent to the QQP

$$\begin{array}{rl} \mu^*_{MC} := & \max q(x) \\ & \text{s.t. } x_j^2 = 1, \quad j = 1, \ldots, n. \end{array}$$

This problem is NP-hard, i.e., intractable.

Since the above QQP has many nonconvex quadratic constraints, a duality gap for the Lagrangian relaxation is expected and does indeed occur most of the time. However, the Lagrangian dual is equivalent to the SDP relaxation (upper bound)

$$\mu^*_{MC} \le \mu^*_{MCSDP} := \begin{array}{ll} \max & \operatorname{tr} QX \\ \text{s.t.} & \operatorname{diag}(X) = e, \\ & X \succeq 0, \end{array} \tag{2.3}$$

which has proven to have very strong theoretical and practical properties, i.e., the bound has an 87% performance guarantee for the problem MC and a 97% performance in practice; see, e.g., [12, 18, 15]. Other theoretical results for general objectives and further relaxed constraints appear in [30, 46].

In [38], several unrelated, though tractable, bounds for MC are shown to be equivalent. These bounds include the box relaxation $-e \le x \le e$, the trust region relaxation $\sum_i x_i^2 = n$, and an eigenvalue relaxation. Furthermore, these bounds are all shown to be equivalent to the Lagrangian relaxation; see [37]. Thus we see that the Lagrangian relaxation is equivalent to the best of these tractable bounds.

2.6. General QQP. The general, possibly nonconvex QQP has many applications in modeling and approximation theory; see, e.g., the applications to SQP methods in [21]. Examples of approximations to QQPs also appear in [9].

The Lagrangian relaxation of a QQP is equivalent to the SDP relaxation and is sometimes referred to as the Shor relaxation; see [42]. 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 $x^T Qx = \operatorname{tr} x^T Qx = \operatorname{tr} Qxx^T$ and relaxing xx^T to a semidefinite matrix X.

One can relate the geometry of the original feasible set of QQP 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 [10, 20].

3. Orthogonally constrained programs with zero duality gaps. Consider the *orthonormal constraint*

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

(The set of such X is sometimes known as the Stiefel manifold; see, e.g., [7]. Applications and algorithms for optimization on orthonormal sets of matrices are discussed in [7].) 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 result of this section can be viewed as a matrix generalization of the strong duality result for the Rayleigh quotient problem (2.1).

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

$$\text{(3.1)} \qquad \mathrm{QQP_O} \qquad \mu^O := \begin{array}{ll} \min & \operatorname{tr} AXBX^T \\ \text{s.t.} & XX^T = I. \end{array}$$

This problem can be solved exactly using Lagrange multipliers (see, e.g., [14]) or using the classical Hoffman–Wielandt inequality (see, e.g., [3]). We include a simple proof for completeness.

PROPOSITION 3.1. *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* $\mu^O = \operatorname{tr} \Sigma\Lambda$ *and the optimal solution is obtained using the orthogonal matrices that yield the diagonalizations, i.e.,* $X^* = VU^T$.

Proof. The constraint $G(X) := XX^T - I$ maps $\mathcal{M}_n$ to $\mathcal{S}_n$. The Jacobian of the constraint at X acting on the direction h is $J(X)(h) = Xh^T + hX^T$. The adjoint of the Jacobian acting on $S \in \mathcal{S}_n$ is $J^*(X)(S) = 2SX$, since

$$\operatorname{tr} SJ(X)(h) = \operatorname{tr} h^T J^*(X)(S).$$

But $J^*(X)(S) = 0$ implies $S = 0$, i.e., J^* is one-one for all X orthogonal. Therefore, J is onto, i.e., the standard constraint qualification holds at the optimum. It follows that the necessary conditions for optimality are that the gradient of the Lagrangian

$$\text{(3.2)} \qquad L(X, S) = \operatorname{tr} AXBX^T - \operatorname{tr} S(XX^T - I)$$

is 0, i.e.,

$$AXB - SXI = 0.$$

Therefore,

$$AXBX^T = S = S^T,$$

i.e., $AXBX^T$ is symmetric, which means that A and XBX^T commute and so are mutually diagonalizable by the orthogonal matrix U. Therefore, we can assume that both A and B are diagonal and we choose X to be a product of permutations that gives the correct ordering of the eigenvalues. ☐

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

$$\text{(3.3)} \qquad \max_{S=S^T} \min_X \operatorname{tr} AXBX^T - \operatorname{tr} S(XX^T - I).$$

However, there can be a nonzero duality gap for the Lagrangian dual; see [50] for an example. The inner minimization in the dual problem (3.3) is an unconstrained quadratic minimization in the variables $\operatorname{vec}(X)$, with Hessian

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

Clearly this minimization is unbounded if the Hessian is not positive semidefinite. In order to close the duality gap, we need a larger class of quadratic functions.

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

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

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

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

THEOREM 3.2. *Strong duality holds for* $\mathrm{QQP_{OO}}$ *and* $\mathrm{DQQP_{OO}}$, *i.e.,* $\mu^D = \mu^O$ *and both primal and dual are attained.*

Proof. Let $A = V\Sigma V^T, B = U\Lambda U^T$, where V and U are orthonormal matrices whose columns are the eigenvectors of A and B, respectively, σ and λ are the corresponding vectors of eigenvalues, and $\Sigma = \operatorname{Diag}(\sigma)$, $\Lambda = \operatorname{Diag}(\lambda)$. Then for any S and T,

$$(B \otimes A) - (I \otimes S) - (T \otimes I) = (U \otimes V)\left[(\Lambda \otimes \Sigma) - (I \otimes \bar{S}) - (\bar{T} \otimes I)\right](U^T \otimes V^T),$$

where $\bar{S} = V^TSV$, $\bar{T} = U^TTU$. Since $U \otimes V$ is nonsingular, $\operatorname{tr} S = \operatorname{tr} \bar{S}$, and $\operatorname{tr} T = \operatorname{tr} \bar{T}$, the dual problem $\mathrm{DQQP_{OO}}$ is equivalent to

$$\begin{array}{rl} \mu^D = & \max \operatorname{tr} S + \operatorname{tr} T \\ & \text{s.t. } (\Lambda \otimes \Sigma) - (I \otimes S) - (T \otimes I) \succeq 0, \\ & \quad S = S^T,\ T = T^T. \end{array} \tag{3.4}$$

However, since Λ and Σ are diagonal matrices, (3.4) is equivalent to the ordinary linear program:

$$\mathrm{LD} \qquad \begin{array}{l} \max e^Ts + e^Tt \\ \text{s.t. } \lambda_i\sigma_j - s_j - t_i \geq 0, \quad i, j = 1, \dots, n. \end{array}$$

But LD is the dual of the linear assignment problem:

$$\mathrm{LP} \qquad \begin{array}{l} \min \sum_{i,j} \lambda_i \sigma_j x_{ij} \\ \text{s.t. } \sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n, \\ \qquad \sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n, \\ \qquad x_{ij} \geq 0, \quad i, j = 1, \dots, n. \end{array}$$

Assume without loss of generality that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. Then LP can be interpreted as the problem of finding a permutation $\pi(\cdot)$ of $\{1, \dots, n\}$ so that $\sum_{i=1}^n \lambda_i \sigma_{\pi(i)}$ is minimized. But the minimizing permutation is then $\pi(i) = i$, $i = 1, \dots, n$, and from Proposition 3.1 the solution value μ^D is exactly μ^O. $\square$

4. Applications. We now present three applications of the above strong duality result.

4.1. Quadratic assignment problem. Let A and B be $n \times n$ symmetric matrices, and consider the homogeneous quadratic assignment problem (QAP) (see, e.g., [32]),

$$\text{QAP} \qquad \begin{array}{l} \min \operatorname{tr} AXBX^T \\ \text{s.t. } X \in \Pi, \end{array}$$

where Π is the set of $n \times n$ permutation matrices. The set of orthonormal matrices contains the permutation matrices, and the orthonormally constrained problem (3.1) is an important relaxation of QAP. The bounds obtained are usually called the eigenvalue bounds for QAP; see [8, 13]. Theorem 3.2 shows that the eigenvalue bounds are in fact obtained from a Lagrangian relaxation of (3.1) after adding the seemingly redundant constraint $XX^T = I$.

4.2. Weighted sums of eigenvalues. Consider the problem of minimizing the weighted sum of the k largest eigenvalues of an $n \times n$ symmetric matrix Y, subject to linear equality constraints. An SDP formulation for this problem involving $2k$ semidefiniteness constraints on $n \times n$ matrices is given in [1, section 4.3]. We will show that the result of section 3 can be applied to obtain a new SDP formulation of the problem having only $k + 1$ semidefiniteness constraints on $n \times n$ matrices.

For convenience we consider the equivalent problem of maximizing the weighted sum of the k minimum eigenvalues of Y. Let $w_1 \geq w_2 \geq \cdots \geq w_k > w_{k+1} = w_{k+2} = \cdots = w_n = 0$, and let $W = \operatorname{Diag}(w)$. We are interested in the problem

$$\text{WEIG} \qquad \begin{array}{l} \max \displaystyle\sum_{i=1}^{k} w_i \lambda_i(Y) \\ \text{s.t. } \mathcal{A} \operatorname{vec}(Y) = b, \\ \qquad Y = Y^T, \end{array}$$

where $\lambda_1(Y) \leq \lambda_2(Y) \leq \cdots \leq \lambda_n(Y)$ are the eigenvalues of Y, and $\mathcal{A}$ is a $p \times n^2$ matrix. From Proposition 3.1 it is clear that, for any Y,

$$\sum_{i=1}^{k} w_i \lambda_i(Y) = \min_{X^T X = I} \operatorname{tr} YXWX^T,$$

and therefore from Theorem 3.2 the problem WEIG is equivalent to the problem

$$\begin{array}{ll} \max & \operatorname{tr} S + \operatorname{tr} T \\ \text{s.t.} & (W \otimes Y) - (I \otimes S) - (T \otimes I) \succeq 0, \\ & \mathcal{A} \operatorname{vec}(Y) = b, \\ & S = S^T, \ T = T^T, \ Y = Y^T. \end{array} \tag{4.1}$$

Note that, for any Y, the matrix $W \otimes Y$ is block diagonal, with the final $n - k$ blocks identically zero. Since $I \otimes S$ is also block diagonal, and $\operatorname{tr} T$ is a function of the diagonal of T only, it is obvious that T can be assumed to be a diagonal matrix $T = \operatorname{Diag}(t)$. Writing the problem (4.1) in terms of t, and separating the block diagonal constraints, results in the SDP

$$\begin{aligned} &\max \operatorname{tr} S + \sum_{i=1}^{k} t_i + (n-k)t_{k+1} \\ &\text{s.t. } w_i Y - S - t_i I \succeq 0, \quad i = 1\dots,k, \\ &\quad -S - t_{k+1} I \succeq 0, \\ &\quad \mathcal{A} \operatorname{vec} Y = b, \\ &\quad S = S^T. \end{aligned}$$

We have thus obtained an SDP representation for the problem WEIG with $k+1$ semidefiniteness constraints on $n \times n$ matrices, as claimed.

4.3. Graph partitioning problem. Let $G = (N, E)$ be an edge-weighted undirected graph with node set $N = \{1, \dots, n\}$, edge set E, and weights $w_{ij},\ ij \in E$. The graph partitioning (GP) problem consists of partitioning the node set N into k disjoint subsets $S_1, \dots, S_k$ of specified sizes $m_1 \geq m_2 \geq \cdots \geq m_k,\ \sum_{j=1}^{k} m_j = n$, so as to minimize the total weight of the edges connecting nodes in distinct subsets of the partition. This problem is well known to be NP-hard. GP can be modeled as a quadratic problem

$$\begin{aligned} z := \min \operatorname{tr} X^T L X \\ \text{s.t. } X \in P, \end{aligned}$$

where L is the Laplacian of the graph and P is the set of $n \times k$ partition matrices (i.e., each column of X is the indicator function of the corresponding set; $X_{ij} = 1$ if node i is in set j and 0 otherwise).

The well-known *Donath–Hoffman* bound [6] $z_{DH} \leq z$ for GP is

$$z_{DH} := \max_{e^T u = 0} \sum_{i=1}^{k} m_i \lambda_i (L + U),$$

where $U = \operatorname{Diag}(u)$, and $\lambda_1(L+U) \leq \lambda_2(L+U) \leq \cdots \leq \lambda_n(L+U)$ are the eigenvalues of $L + U$. We will now show that the Donath–Hoffman bound can be obtained by applying Lagrangian relaxation to an appropriate QQP relaxation of GP. (An SDP formulation for this bound is given in [1].) Clearly, if P is a partition matrix, then $x_i^T x_i = 1,\ i = 1, \dots, n$, where x_i^T is the ith row of X. Moreover, the columns of X are orthogonal with one another, and the norm of the jth column of X is $\sqrt{m_j}$. It follows that if X is a partition matrix, there is an $n \times n$ orthogonal matrix $\bar{X}$ such that

$$X = \bar{X} \begin{pmatrix} M^{1/2} \\ 0 \end{pmatrix},$$

where M is the $k \times k$ matrix $M = \operatorname{Diag}(m)$, and therefore

$$XX^T = \bar{X} \begin{pmatrix} M^{1/2} \\ 0 \end{pmatrix} (M^{1/2}, 0) \bar{X}^T = \bar{X} \bar{M} \bar{X}^T, \quad \text{where} \quad \bar{M} = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix}.$$

In addition, note that $x_i^T x_i$ is the ith diagonal element of XX^T, so the constraint $x_i^T x_i = 1$ is equivalent to $\bar{x}_i^T \bar{M} \bar{x}_i = 1$, where $\bar{x}_i^T$ is the ith row of $\bar{X}$. Since $\operatorname{tr} X^T L X =$

$\operatorname{tr} LXX^T$, a lower bound $z_1 \le z$ can be defined by

$$
\begin{aligned}
z_1 := & \min \operatorname{tr} L\bar{X}\bar{M}\bar{X}^T \\
& \text{s.t. } \bar{X}^T\bar{X} = I,\ \bar{X}\bar{X}^T = I, \\
& \quad \bar{x}_i^T\bar{M}\bar{x}_i = 1, i = 1, \dots, n.
\end{aligned} \tag{4.2}
$$

We will now obtain a second bound $z_2 \le z_1$ by applying a Lagrangian procedure to all of the constraints in (4.2). Using symmetric matrices S and T for the constraints $\bar{X}\bar{X}^T = I$ and $\bar{X}^T\bar{X} = I$, respectively, and a vector of multipliers u_i for the constraints $\bar{x}_i^T\bar{M}\bar{x}_i = 1$, $i = 1, \dots, n$, we obtain

$$
\begin{aligned}
z_2 := \max_{u,S,T} \min_{\bar{X}} \ & \operatorname{tr} L\bar{X}\bar{M}\bar{X}^T + \operatorname{tr} S(I - \bar{X}\bar{X}^T) + \operatorname{tr} T(I - \bar{X}^T\bar{X}) \\
& + \sum_{i=1}^n u_i(\bar{x}_i^T\bar{M}\bar{x}_i - 1).
\end{aligned}
$$

THEOREM 4.1. $z_2 = z_{DH}$.

Proof. Rearranging terms and using Kronecker product notation, the definition of z_2 can be rewritten as

$$
\begin{aligned}
z_2 = \max_{u,S,T} \quad & \operatorname{tr} S + \operatorname{tr} T - e^T u \\
& + \min_{\bar{X}} \ \operatorname{vec}(\bar{X})^T([\bar{M} \otimes (L+U)] - (I \otimes S) - (T \otimes I))\operatorname{vec}(\bar{X}),
\end{aligned}
$$

where $U = \operatorname{Diag}(u)$, and we are using the fact that

$$
\sum_{i=1}^n u_i\bar{x}_i^T\bar{M}\bar{x}_i = \operatorname{tr} U\bar{X}\bar{M}\bar{X}^T.
$$

Clearly if $[\bar{M} \otimes (L+U)] - (I \otimes S) - (T \otimes I) \succeq 0$, then $\bar{X} = 0$ solves the implicit minimization problem in the definition of z_2, and if this constraint fails to hold, the minimum is $-\infty$. Using this hidden semidefinite constraint, we can write

$$
\begin{aligned}
z_2 = & \max \operatorname{tr} S + \operatorname{tr} T - e^T u \\
& \text{s.t. } [\bar{M} \otimes (L+U)] - (I \otimes S) - (T \otimes I) \succeq 0, \\
& \quad S = S^T,\ T = T^T.
\end{aligned}
$$

Note that if $u' = u + \lambda e$ and $T' = T + \lambda\bar{M}$ for any scalar λ, then

$$
\begin{aligned}
\bar{M} \otimes (L + U') &= \bar{M} \otimes (L+U) + \lambda(\bar{M} \otimes I), \\
T' \otimes I &= T \otimes I + \lambda(\bar{M} \otimes I).
\end{aligned}
$$

In addition, $\operatorname{tr} T' = \operatorname{tr} T + \lambda n$ and $e^T u' = e^T u + \lambda n$. It follows that we may choose any normalization for $e^T u$ without affecting the value of z_2. Choosing $e^T u = 0$, we arrive at

$$
\begin{aligned}
z_2 = & \max \operatorname{tr} S + \operatorname{tr} T \\
& \text{s.t. } [\bar{M} \otimes (L+U)] - (I \otimes S) - (T \otimes I) \succeq 0, \\
& \quad e^T u = 0,\ S = S^T,\ T = T^T.
\end{aligned}
$$

However, as in the previous section, Proposition 3.1 and Theorem 3.2 together imply that for any U, the solution value in the problem

$$\begin{aligned} &\max \operatorname{tr} S + \operatorname{tr} T \\ &\text{s.t. } [\bar{M} \otimes (L+U)] - (I \otimes S) - (T \otimes I) \succeq 0, \\ &\qquad S = S^T,\ T = T^T, \end{aligned}$$

is exactly $\sum_{i=1}^{k} m_i \lambda_i(L+U)$. Therefore, we immediately have $z_{DH} = z_2$. ∎

SDP relaxations for the GP problem are obtained via Lagrangian relaxation in [45]. A useful corollary of Theorem 4.1 is that any Lagrangian relaxation based on a more tightly constrained problem than (4.2) will produce bounds that dominate the Donath–Hoffman bounds.

A problem closely related to the orthogonal relaxation of GP is the orthogonal Procrustes problem on the Stiefel manifold; see [7, section 3.5.2]. This problem has a linear term in the objective function, and there is no known analytic solution for the general case.

5. A strengthened relaxation for max-cut. As discussed above, the SDP relaxation for MC performs very well in practice and has strong theoretical properties. There have been attempts at further strengthening this relaxation. For example, a copositive relaxation is presented in [39]. Adding cuts to the SDP relaxation is discussed in [15, 16, 17, 18]. These improvements all involve heuristics, such as deciding which cuts to choose or solving a copositive problem, which is NP-hard in itself.

The relaxation in (2.3) is obtained by *lifting* the vector x into matrix space using $X = xx^T$. Though the matrix X in the lifting is not an orthogonal matrix, it is a *partial isometry* up to normalization, i.e.,

$$X^2 - nX = 0. \tag{5.1}$$

We will now show that we can improve the semidefinite relaxation presented in section 2.5 by considering Lagrangian relaxations using the matrix quadratic constraint (5.1). In particular, consider the relaxation of MC

$$\begin{aligned} \mu_1 := \max \quad & \operatorname{tr} \quad QX \\ \text{s.t.} \quad & \operatorname{diag}(X) = e, \\ & X^2 \quad -nX = 0, \end{aligned}$$

where X is a symmetric matrix. Note that if $X^2 = nX$, then $\operatorname{tr} QX = (1/n) \operatorname{tr} QX^2$, and $\operatorname{diag}(X^2) = ne$. As a result, the above relaxation is equivalent to the relaxation

$$\begin{aligned} \mu_1 = \max\ & \frac{1}{n} \operatorname{tr} QX^2 \\ \text{s.t. } & x_i^T x_i = n, \quad i = 1, \dots, n, \\ & X^2 - n x_0 X = 0, \\ & x_0^2 = 1, \end{aligned} \tag{5.2}$$

where x_i^T, $i = 1, \dots, n$, denotes the ith row of X, and x_0 is a scalar. (Note that if $x_0 = -1$, then changing x_0 to 1 and replacing X with $-X$ leaves the objective and constraints in (5.2) unchanged.) We will obtain an upper bound $\mu_2 \geq \mu_1$ by applying a Lagrangian procedure to all of the constraints in (5.2). Using multipliers u_i for the constraints $x_i^T x_i = n$, $i = 1, \ldots, n$, u_0 for the constraint $x_0^2 = 1$, and a symmetric matrix S for the matrix equality $X^2 - nX = 0$, we obtain a Lagrangian problem

$$\mu_2 := \min_{u_0,u,S} u_0 + nu^T e + \max_{x_0,X} \frac{1}{n} \operatorname{tr} QX^2 - \operatorname{tr} UX^2 + \operatorname{tr} SX^2 - nx_0 \operatorname{tr} SX - u_0 x_0^2,$$

where $U = \operatorname{Diag}(u)$. Letting $\bar{x}^T = (x_0, \operatorname{vec}(X)^T)$, this problem can be written in Kronecker product form as

$$\mu_2 = \min_{u_0,u,S} \; u_0 + ne^T u + \max_{\bar{x}} \; \bar{x}^T \bar{Q} \bar{x},$$

where

$$\bar{Q} = \begin{pmatrix} -u_0 & -\frac{n}{2}\operatorname{vec}(S)^T \\ -\frac{n}{2}\operatorname{vec}(S) & I \otimes \left(\frac{1}{n}Q - U + S\right) \end{pmatrix}.$$

Applying the hidden semidefinite constraint $\bar{Q} \preceq 0$, we obtain an equivalent problem,

$$\begin{aligned} \mu_2 = & \min u_0 + ne^T u \\ & \text{s.t.} \begin{pmatrix} u_0 & \frac{n}{2}\operatorname{vec}(S)^T \\ \frac{n}{2}\operatorname{vec}(S) & I \otimes \left(-\frac{1}{n}Q + U - S\right) \end{pmatrix} \succeq 0, \\ & S = S^T. \end{aligned} \tag{5.3}$$

Note that if we take $S = 0$ in (5.3), then $u_0 = 0$ is clearly optimal and the problem reduces to

$$\begin{aligned} & \min e^T u \\ & \text{s.t. } -Q + U \succeq 0, \end{aligned}$$

which is exactly the dual of (2.3), the usual SDP relaxation for MC. It follows that we have obtained an upper bound μ_2 which is a strengthening of the usual SDP bound, i.e., $\mu_2 \leq \mu^*_{MCSDP}$.

The strengthened relaxation (5.3) involves a semidefiniteness constraint on a $(n^2+1) \times (n^2+1)$ matrix, as opposed to an $n \times n$ matrix in the usual SDP relaxation (2.3). This dimensional increase can be mitigated by taking note of the fact that X in (5.2) must be a symmetric matrix, and therefore (5.2) can actually be written as a problem over a vector x of dimension $n(n+1)/2$. In addition, alternative relaxations can be obtained by not making the substitutions based on (5.1) used to obtain the problem (5.2). The effect of these alternatives on the performance of strengthened SDP bounds for MC is the topic of ongoing research; for up-to-date developments, see the URL http://orion.uwaterloo.ca/˜hwolkowi/henry/reports/strngthMC.ps.gz.

6. Conclusion. In this paper we have shown that a class of nonconvex quadratic problems with orthogonal constraints can satisfy strong duality if certain seemingly redundant constraints are added before the Lagrangian dual is formed. As applications of this result we showed that well-known eigenvalue bounds for QAP and GP problems can actually be obtained from the Lagrangian dual of QQP relaxations of these problems. We also showed that the technique of relaxing quadratic matrix constraints can be used to obtain strengthened SDP relaxations for the max-cut problem.

Adding constraints to close the duality gap is akin to adding valid inequalities in cutting plane methods for discrete optimization problems. In [2, 24] this approach, in combination with a lifting procedure, is used to solve discrete optimization problems. In our case we add quadratic constraints. The idea of quadratic valid inequalities has been used in [10]; and closing the duality gap has been discussed in [20].

Our success in closing the duality gap for the QQP_O problem considered in section 3, where we have the special Kronecker product in the objective function, raises several interesting questions. For example, can the strong duality result for QQP_O be extended to the same problem with an added linear term in the objective, or are there some other special classes of objective functions where this is possible? Another outstanding question is whether it is possible to add quadratic constraints to close the duality gap for the TTRS.

REFERENCES

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[2] E. Balas, S. Ceria, and G. Cornuejols, *A lift-and-project cutting plane algorithm for mixed* 0-1 *programs*, Math. Programming, 58 (1993), pp. 295–324.

[3] R. Bhatia, *Perturbation Bounds for Matrix Eigenvalues*, Pitman Res. Notes Math. Ser. 162, Longman Scientific and Technical, Harlow, UK, 1987.

[4] M. Celis, J. Dennis Jr., and R. Tapia, *A trust region strategy for nonlinear equality constrained optimization*, in Proceedings of the SIAM Conference on Numerical Optimization, Boulder, CO, 1984 , pp. 71–82. Also available as Technical Report TR84-1, Rice University, Houston, TX.

[5] J.-P. Crouzeix, J.-E. Martinez-Legaz, and A. Seeger, *An alternative theorem for quadratic forms and extensions*, Linear Algebra Appl., 215 (1995), pp. 121–134.

[6] W. Donath and A. Hoffman, *Lower bounds for the partitioning of graphs*, IBM J. Res. Develop., 17 (1973), pp. 420–425.

[7] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303–353.

[8] G. Finke, R. Burkhard, and F. Rendl, *Quadratic assignment problems*, Ann. Discrete Math., 31 (1987), pp. 61–82.

[9] M. Fu, Z. Luo, and Y. Ye, *Approximation algorithms for quadratic programming*, J. Combinatorial Optim., 2 (1998), pp. 29–50.

[10] T. Fujie and M. Kojima, *Semidefinite programming relaxation for nonconvex quadratic programs*, J. Global Optim., 10 (1997), pp. 367–380.

[11] M. X. Goemans, *Semidefinite programming in combinatorial optimization*, Math. Programming, 79 (1997), pp. 143–162.

[12] M. Goemans and D. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach., 42 (1995), pp. 1115–1145.

[13] S. Hadley, F. Rendl, and H. Wolkowicz, *Bounds for the quadratic assignment problems using continuous optimization*, in Integer Programming and Combinatorial Optimization, University of Waterloo Press, Waterloo, Ontario, Canada, 1990, pp. 237–248.

[14] S. Hadley, F. Rendl, and H. Wolkowicz, *A new lower bound via projection for the quadratic assignment problem*, Math. Oper. Res., 17 (1992), pp. 727–739.

[15] C. Helmberg, *An Interior Point Method for Semidefinite Programming and Max-Cut Bounds*, Ph.D. thesis, Graz University of Technology, Austria, 1994.

[16] C. Helmberg, *Fixing Variables in Semidefinite Relaxations*, Lecture Notes in Comput. Sci. 1284, Springer, Berlin, 1997.

[17] C. Helmberg and F. Rendl, *A spectral bundle method for semidefinite programming*, SIAM J. Optim, 10 (2000), pp. 673–696.

[18] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM J. Optim., 6 (1996), pp. 342–361.

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

[20] M. Kojima and L. Tunçel, *Cones of Matrices and Successive Convex Relaxations of Nonconvex Sets*, Technical Report B-338, Tokyo Institute of Technology, Tokyo, Japan, 1998.

[21] S. Kruk and H. Wolkowicz, *SQ^2P, sequential quadratic constrained quadratic programming*, in Advances in Nonlinear Programming, Y. Xiang Yuan, ed., Appl. Optim. 14, Kluwer Academic Publishers, Dordrecht, 1998, pp. 177–204.
[22] M. Laurent, *Cuts, matrix completions and graph rigidity*, Math. Programming, 79 (1997), pp. 255–284.
[23] M. Laurent, *A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems*, in Topics in Semidefinite and Interior-Point Methods, Fields Inst. Commun. 18, AMS, Providence, RI, 1998, pp. 51–76.
[24] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and* 0-1 *optimization*, SIAM J. Optim., 1 (1991), pp. 166–190.
[25] Z.-Q. Luo and J. Sun, *An Analytic Center Based Column Generation Algorithm for Convex Quadratic Feasibility Problems*, Technical report, McMaster University, Hamilton, Ontario, Canada, 1995; also available from ftp://ftp.nus.sg/pub/NUS/opt/qcut.ps.gz.
[26] Z.-Q. Luo and S. Zhang, *On the Extension of Frank-Wolfe Theorem*, Technical report, Erasmus University Rotterdam, The Netherlands, 1997.
[27] J. M. Martinez, *Local minimizers of quadratic functions on Euclidean balls and spheres*, SIAM J. Optim., 4 (1994), pp. 159–176.
[28] J. J. Moré, *Generalizations of the trust region problem*, Optim. Methods Softw., 2 (1993), pp. 189–209.
[29] J. Moré and D. Sorensen, *Computing a trust region step*, SIAM J. Sci. Stat. Comput., 4 (1983), pp. 553–572.
[30] Y. E. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, Optim. Methods Softw., 9 (1998), pp. 141–160.
[31] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Stud. Appl. Math. 13, SIAM, Philadelphia, PA, 1994.
[32] P. Pardalos, F. Rendl, and H. Wolkowicz, *The quadratic assignment problem: A survey and recent developments*, in Proceedings of the DIMACS Workshop on Quadratic Assignment Problems, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 16, AMS, Providence, RI, 1994, pp. 1–41.
[33] J.-M. Peng and Y.-X. Yuan, *Optimality conditions for the minimization of a quadratic with two quadratic constraints*, SIAM J. Optim., 7 (1997), pp. 579–594.
[34] 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, Princeton University Press, Princeton, NJ, 1970, pp. 445–480.
[35] 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 (1969), pp. 317–340.
[36] 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 (1970), pp. 365–383.
[37] S. Poljak, F. Rendl, and H. Wolkowicz, *A recipe for semidefinite relaxation for* (0,1)-*quadratic programming*, J. Global Optim., 7 (1995), pp. 51–73.
[38] S. Poljak and H. Wolkowicz, *Convex relaxations of* 0-1 *quadratic programming*, Math. Oper. Res., 20 (1995), pp. 550–561.
[39] A. Quist, E. D. Klerk, C. Roos, and T. Terlaky, *Copositive relaxation for general quadratic programming*, Optim. Methods Softw., 9 (1998), pp. 185–208.
[40] F. Rendl and H. Wolkowicz, *A semidefinite framework for trust region subproblems with applications to large scale minimization*, Math. Programming, 77 (1997), pp. 273–299.
[41] S. Santos and D. Sorensen, *A New Matrix-Free Algorithm for the Large-Scale Trust-Region Subproblem*, Technical Report TR95-20, Rice University, Houston, TX, 1995.
[42] N. Shor, *Quadratic optimization problems*, Izv. Akad. Nauk SSSR Tekhn. Kibernet., 222 (1987), pp. 128–139.
[43] R. J. Stern and H. Wolkowicz, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, SIAM J. Optim., 5 (1995), pp. 286–313.
[44] T. Terlaky, *On l_p programming*, European J. Oper. Res., 22 (1985), pp. 70–100.
[45] H. Wolkowicz and Q. Zhao, *Semidefinite relaxations for the graph partitioning problem*, Discrete Appl. Math., 96–97 (1999), pp. 461–479.
[46] Y. Ye, *Approximating quadratic programming with bound and quadratic constraints*, Math. Programming, 84 (1999), pp. 219–226.
[47] Y. Yuan, *Some Properties of Trust Region Algorithms for Nonsmooth Optimization*, Technical Report DAMTP 1983/NA14, University of Cambridge, Cambridge, UK, 1983.

[48] Y. Yuan, *On a subproblem of trust region algorithms for constrained optimization*, Math. Programming, 47 (1990), pp. 53–63.
[49] Y. Yuan, *A dual algorithm for minimizing a quadratic function with two quadratic constraints*, J. Comput. Math., 9 (1991), pp. 348–359.
[50] Q. Zhao, S. Karisch, F. Rendl, and H. Wolkowicz, *Semidefinite programming relaxations for the quadratic assignment problem*, J. Combin. Optim., 2 (1998), pp. 71–109.