MATHEMATICS OF OPERATIONS RESEARCH
Vol. 34, No. 4, November 2009, pp. 1008–1022
ISSN 0364-765X | EISSN 1526-5471 | 09 | 3404 | 1008

informs®

DOI 10.1287/moor.1090.0419

A Low-Dimensional Semidefinite Relaxation for the Quadratic Assignment Problem

Yichuan Ding
Department of Management Science and Engineering, School of Engineering, Stanford University, Stanford, California 94305, y7ding@standford.edu

Henry Wolkowicz
Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, hwolkowicz@uwaterloo.ca, http://orion.uwaterloo.ca/~hwolkowi

The quadratic assignment problem (QAP) is arguably one of the hardest NP-hard discrete optimization problems. Problems of dimension greater than 25 are still considered to be large scale. Current successful solution techniques use branch-and-bound methods, which rely on obtaining *strong and inexpensive* bounds. In this paper, we introduce a new semidefinite programming (SDP) relaxation for generating bounds for the QAP in the trace formulation. We apply majorization to obtain a relaxation of the orthogonal similarity set of the quadratic part of the objective function. This exploits the matrix structure of QAP and results in a relaxation with much smaller dimension than other current SDP relaxations. We compare the resulting bounds with several other computationally inexpensive bounds such as the convex quadratic programming relaxation (QPB). We find that our method provides stronger bounds on average and is adaptable for branch-and-bound methods.

Key words: quadratic assignment problem; semidefinite programming relaxations; interior point methods; large-scale problems
MSC2000 subject classification: Primary: 90C26, 90C22; secondary: 90C09, 65K10
OR/MS subject classification: Primary: Programming; secondary: quadratic
History: Received October 27, 2006; revised April 22, 2008, June 2, 2009, and August 10, 2009. Published online in *Articles in Advance* October 20, 2009.

1. Introduction. In this paper, we introduce a new efficient bound for the quadratic assignment problem (QAP). We use the *Koopmans-Beckmann* trace formulation

$$\text{(QAP)} \quad \mu^*_{\text{QAP}} := \min_{X \in \Pi} \operatorname{trace} AXBX^T + CX^T,$$

where A, B, and C are n by n real matrices and Π denotes the set of n by n permutation matrices. Throughout this paper, we assume the symmetric case, i.e., that both A and B are symmetric matrices. The QAP is considered to be one of the hardest NP-hard problems to solve in practice. Many important combinatorial optimization problems can be formulated as a QAP. Examples include the traveling salesman problem, VLSI design, keyboard design, and the graph-partitioning problem. The QAP is well-described by the problem of allocating a set of n facilities to a set of n locations while minimizing the quadratic objective arising from the distance between the locations in combination with the flow between the facilities. Recent surveys include Pardalos and Wolkowicz [29], Wolkowicz [34], Zhao et al. [36], Pardalos et al. [30], Karisch et al. [20], Hadley et al. [13, 14, 15], and Rendl and Wolkowicz [33].

Solving QAP to optimality usually requires a branch-and-bound (B&B) method. Essential for these methods are strong, inexpensive bounds at each node of the branching tree. In this paper, we study a new bound obtained from a semidefinite programming (SDP) relaxation. This relaxation uses only $O(n^2)$ variables and $O(n^2)$ constraints but yields a bound provably better than the so-called *projected eigenvalue bound* (PB) (Hadley et al. [14]), and is competitive with the recently introduced *quadratic programming bound* (QPB) (Anstreicher and Brixius [2]).

1.1. Outline. In §1.2, we continue with preliminary results and notation. In §1.3, we review some of the known bounds in the literature. Our main results appear in §2 where we compare relaxations that use a *vector lifting* of the matrix X into the space of $n^2 \times n^2$ matrices with a *matrix lifting* that remains in $\mathcal{S}^n$, the space of $n \times n$ symmetric matrices. We then parameterize and characterize the *orthogonal similarity set of B*, $\mathcal{O}(B)$, using majorization results on the eigenvalues of B (see Theorem 2.1). This results in three SDP relaxations, MSDR_1 to MSDR_3 (see §2.2). We conclude with numerical tests in §3.

1.2. Notation and preliminaries. For two real $m \times n$ matrices $A, B \in \mathcal{M}^{mn}$, $\langle A, B\rangle = \operatorname{trace} A^T B$ is the trace inner product. $\mathcal{M}^{nn} = \mathcal{M}^n$ denotes the set of n by n square real matrices and $\mathcal{S}^n$ denotes the space of $n \times n$ symmetric matrices. $\mathcal{S}^n_+$ (resp. $\mathcal{S}^n_{++}$) denotes the cone of positive semidefinite (resp. positive definite) matrices in $\mathcal{S}^n$. We let $A \succeq B$ (resp. $A \succ B$) denote the Löwner partial order $A - B \in \mathcal{S}^n_+$ (resp. $A - B \in \mathcal{S}^n_{++}$).

The linear transformation $\operatorname{diag} M$ denotes the vector formed from the diagonal of the matrix M and the adjoint linear transformation is $\operatorname{diag}^* v = \operatorname{Diag} v$, i.e., the diagonal matrix formed from the vector v. We use $A \otimes B$ to denote the Kronecker product of A and B and use $x = \operatorname{vec}(X)$ to denote the vector in $\mathbb{R}^{n^2}$ obtained from the columns of X. Then (see, e.g., Horn and Johnson [19]),

$$\operatorname{trace} AXBX^T = \langle AXB, X\rangle = \langle \operatorname{vec}(AXB), x\rangle = x^T(B \otimes A)x. \tag{1}$$

We let $\mathcal{N}$ denote the cone of nonnegative (elementwise) matrices $\mathcal{N} := \{X \in \mathcal{M}^n \colon X \geq 0\}$. $\mathcal{E}$ denotes the set of matrices with row and column sums 1, $\mathcal{E} := \{X \in \mathcal{M}^n \colon Xe = X^T e = e\}$, where e is the vector of ones. $E = ee^T$ is the matrix of ones and $\mathcal{D}$ denotes the set of doubly stochastic matrices $\mathcal{D} = \mathcal{E} \cap \mathcal{N}$. The *minimal product* of two vectors is

$$\langle x, y\rangle_- := \min_{\sigma,\, \pi} \sum_{i=1}^n x_{\sigma(i)} y_{\pi(i)},$$

where the minimum is over all permutations, σ, π, of the indices $\{1, 2, \dots, n\}$. Similarly, we define the maximal product of x, y, $\langle x, y\rangle_+ := \max_{\sigma,\pi} \sum_{i=1}^n x_{\sigma(i)} y_{\pi(i)}$. We denote the vector of eigenvalues of a matrix A by $\lambda(A)$.

Definition 1.1. Let $x, y \in \mathbb{R}^n$. By abuse of notation, we denote *x majorizes y* or *y is majorized by x* with $x \succeq y$ or $y \preceq x$. Let the components of both vectors be sorted in nonincreasing order, i.e., $x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(n)}$, $y_{\pi(1)} \geq y_{\pi(2)} \geq \cdots \geq y_{\pi(n)}$. Following, e.g., Marshall and Olkin [23], $x \succeq y$ if and only if

$$\sum_{i=1}^p x_{\sigma(i)} \geq \sum_{i=1}^p y_{\pi(i)}, \quad p = 1, 2, \dots, n-1,$$
$$\sum_{i=1}^n x_{\sigma(i)} = \sum_{i=1}^n y_{\pi(i)}.$$

In Marshall and Olkin [23], it is shown that $x \succeq y$ if and only if there exists $S \in \mathcal{D}$ with $Sx = y$. Note that for fixed y, the constraint $x \succeq y$ is not a convex constraint; however, $x \preceq y$ is a convex constraint and it has an equivalent LP formulation (e.g., Hardy et al. [18]).

1.3. Known relaxations for QAP. One of the earliest and least expensive relaxations for QAP is the Gilmore-Lawler bound (GLB), which is based on a linear programming (LP) formulation (see, e.g., Gilmore [12], Drezner [9]). Related dual-based LP bounds such as the Karisch-Çela-Clausen-Espersen bound (KCCEB), are discussed in Karisch et al. [21], Pardalos et al. [31], Drezner [9], and Hahn and Grant [16]. These formulations are currently able to handle problems with moderate size n (approximately 20) (Gilmore [12], Lawler [22]). Formulations based on nonlinear optimization include eigenvalue and parametric eigenvalue bounds (EB) (Finke et al. [11], Rendl and Wolkowicz [33]), projected eigenvalue bounds (PB) (Hadley et al. [14], Falkner et al. [10]), convex quadratic programming (QPB) bounds (Anstreicher and Brixius [2]), and SDP bounds (Rendl and Sotirov [32], Zhao et al. [36]). For recent numerical results that use these bounds, see, e.g., Anstreicher and Brixius [2], Rendl and Sotirov [32]. A summary and comparison of many of these bounds is given in Anstreicher [1].

Note that $\Pi = \mathcal{O} \cap \mathcal{E} \cap \mathcal{N}$, i.e., the addition of the orthogonal constraints changes the doubly stochastic matrices to permutation matrices. This illustrates the power of nonlinear quadratic constraints for QAP. Using the quadratic constraints, we can see that SDP arises naturally from Lagrangian relaxation (see, e.g., Nesterov et al. [27]). Alternatively, one can *lift* the problem using the positive semidefinite matrix $\binom{1}{\operatorname{vec}(X)}\binom{1}{\operatorname{vec}(X)}^T$ into the symmetric matrix space $\mathcal{S}^{n^2+1}$. One then obtains deep cuts for the convex hull of the lifted permutation matrices. However, this vector-lifting SDP relaxation requires $O(n^4)$ variables and, hence, is expensive to use. Problems with $n > 25$ become impractical for branch-and-bound methods.

It has been proved in Anstreicher and Wolkowicz [3] that strong (Lagrangian) duality holds for the following quadratic program with orthogonal constraints:

$$\mu^*_{\mathrm{EB}} = \min_{XX^T = X^TX = I} \operatorname{trace}(AXBX^T).$$

The optimal value μ^*_{EB} yields the so-called eigenvalue bound, denoted EB. The Lagrangian dual is

$$\mu^*_{\mathrm{EB}} = \max_{S,\,T\in\mathcal{S}^n} \min_{x\in\mathbb{R}^{n^2}} \{\mathrm{trace}(S) + \mathrm{trace}(T) + x^T(B\otimes A - I\otimes S - T\otimes I)x\}. \tag{2}$$

The inner minimization problem results in the *hidden semidefinite constraint*

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

Under this constraint, the inner minimization program is attained at $x=0$. As a result of strong duality, the equivalent dual program

$$\mu^*_{\mathrm{EB}} = \max_{S,\,T\in\mathcal{S}^n} \{\mathrm{trace}(S) + \mathrm{trace}(T)\colon\ B\otimes A - I\otimes S - T\otimes I \succeq 0\} \tag{3}$$

has the same value as the primal program, i.e., both yield the eigenvalue bound EB. One can then add the constant row and column sum linear constraints $Xe = X^Te = e$ to obtain the projected eigenvalue bound PB in Hadley et al. [14]. In Anstreicher and Brixius [2], the authors strengthen PB to get a (parametric) convex quadratic programming bound (QPB). This new bound QPB is inexpensive to compute and, under some mild assumptions, is strictly stronger than PB. QPB is a highly competitive bound if we take into account the trade-off between the quality of the bound and the expense in the computation. The use of QPB along with the Condor high-throughput computing system has resulted in the solution for the first time of several large QAP problems from the QAPLIB library (Burkard et al. [7], Anstreicher and Brixius [2], Anstreicher et al. [4]).

In this paper, we propose a new relaxation for QAP, which has comparable complexity to QPB. Moreover, our numerical tests show that this new bound usually obtains better bounds than QPB when applied to problem instances from the QAPLIB library.

2. SDP relaxation and quadratic matrix programming.

2.1. Vector-lifting SDP relaxation (VSDR). Consider the following quadratic constrained quadratic program:

$$\begin{aligned} (\mathrm{QCQP})\quad \mu^*_{\mathrm{QCQP}} := \min\ & (x^TQ_0x + c_0^Tx) + \beta_0\\ \text{s.t.}\ & (x^TQ_jx + c_j^Tx) + \beta_j \le 0, \quad j = 1,\ldots,m,\\ & x\in\mathbb{R}^n, \end{aligned}$$

where for all j, we have $Q_j\in\mathcal{S}^n$, $c_j\in\mathbb{R}^n$, and $\beta_j\in\mathbb{R}$. To find approximate solutions to QCQP, one can homogenize the quadratic functions to get the equivalent quadratic forms $q_j(x,x_0) = x^TQ_jx + c_j^Txx_0 + \beta_jx_0^2$ along with the additional constraint $x_0^2 = 1$. The homogenized forms can be linearized using the vector $\binom{x_0}{x}\in\mathbb{R}^{n+1}$, i.e.,

$$\begin{aligned} q_j(x,x_0) &= \begin{pmatrix} x_0\\ x\end{pmatrix}^T \begin{pmatrix} \beta_j & \frac{1}{2}c_j^T\\ \frac{1}{2}c_j & Q_j\end{pmatrix} \begin{pmatrix} x_0\\ x\end{pmatrix}\\ &= \mathrm{trace}\begin{pmatrix} \beta_j & \frac{1}{2}c_j^T\\ \frac{1}{2}c_j & Q_j\end{pmatrix}\begin{pmatrix} 1 & x^T\\ x & Y\end{pmatrix}, \end{aligned} \tag{4}$$

where Y represents xx^T and the constraint $Y = xx^T$ is relaxed to $xx^T \preceq Y$. Equivalently, we can use the Schur complement and get the lifted linear constraint

$$Z = \begin{pmatrix} 1 & x^T\\ x & Y\end{pmatrix} \succeq 0, \tag{5}$$

i.e., we can identify $y = x$. The objective function is now linear:

$$\mathrm{trace}\begin{pmatrix} \beta_0 & \frac{1}{2}c_0^T\\ \frac{1}{2}c_0 & Q_0\end{pmatrix} Z$$

and the constraints in QCQP are relaxed to linear inequality constraints:

$$\operatorname{trace}\begin{pmatrix} \beta_j & \frac{1}{2}c_j^T \\ \frac{1}{2}c_j & Q_j \end{pmatrix} Z \le 0, \quad j = 1, \ldots, m.$$

In this paper, we call this a *vector-lifting semidefinite relaxation* (VSDR) and we note that the unknown variable $Z \in \mathcal{S}^{n+1}$.

2.2. Matrix-lifting SDP relaxation (MSDR). Consider QCQP with matrix variables:

$$\begin{aligned} \text{(MQCQP)} \quad \mu^*_{\text{MQCQP}} := \min \;& \operatorname{trace}(X^T Q_0 X + C_0 X^T) + \beta_0 \\ \text{s.t.} \;& \operatorname{trace}(X^T Q_j X + C_j X^T) + \beta_j \le 0, \quad j = 1, \ldots, m, \\ & X \in \mathcal{M}^{nr}. \end{aligned}$$

Let $x := \operatorname{vec}(X)$, $c := \operatorname{vec}(C)$, δ_{ij} denote the Kronecker delta, and $E_{ij} = e_i e_j^T \in \mathcal{M}^n$ be the zero matrix except with one at the (i, j) position. Note that if $r = n$, then the orthogonality constraint $XX^T = I$ is equivalent to $x^T(I \otimes E_{ij})x = \delta_{ij}$, $\forall\, i, j$. $X^T X = I$ is equivalent to $x^T(E_{ij} \otimes I)x = \delta_{ij}$, $\forall\, i, j$. Using *both* of the redundant constraints $XX^T = I$ and $X^T X = I$ strengthens the SDP relaxation (see Anstreicher and Wolkowicz [3]). We can now rewrite QAP using the Kronecker product and see that it is a special case of MQCQP with linear and quadratic equality constraints and with nonnegativity constraints (recall that $\Pi = \mathcal{O} \cap \mathcal{E} \cap \mathcal{N}$).

$$\begin{aligned} \mu^*_{\text{QAP}} = \min \;& x^T(B \otimes A)x + c^T x \\ \text{s.t.} \;& x^T(I \otimes E_{ij})x = \delta_{ij}, \quad \forall\, i, j, \\ & x^T(E_{ij} \otimes I)x = \delta_{ij}, \quad \forall\, i, j, \\ & Xe = X^T e = e, \\ & x \ge 0. \end{aligned} \tag{6}$$

Note that in the case of QAP, we have $r = n$ and $x = \operatorname{vec}(X)$ from (6) is in $\mathbb{R}^{n^2}$. Relaxing the quadratic objective function and the quadratic orthogonality constraints results in a linearized/lifted constraint (5) and we end up with $Z = \begin{pmatrix} 1 & x^T \\ x & Y \end{pmatrix} \in \mathcal{S}^{n^2+1}$, a prohibitively large matrix. However, we can use a different approach and exploit the structure of the problem. We can replace the constraint $Y = xx^T$ with the constraint $Y = XX^T$ and then relax it to $Y \succeq XX^T$. This is equivalent to the linear semidefinite constraint $\begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0$. The size of this constraint is significantly smaller. We call this a *matrix-lifting semidefinite relaxation* and denote it MSDR. The relaxation for MQCQP with $X \in \mathcal{M}^{nr}$ is

$$\begin{aligned} \text{(MSDR)} \quad \mu^*_{\text{MSDR}} := \min \;& \operatorname{trace}(Q_0 Y + C_0 X^T) + \beta_0 \\ \text{s.t.} \;& \operatorname{trace}(Q_j Y + C_j X^T) + \beta_j \le 0, \quad j = 1, \ldots, m, \\ & \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0, \\ & X \in \mathcal{M}^{nr}, \quad Y \in \mathcal{S}^n. \end{aligned}$$

If $r \le n$ and the Slater constraint qualification holds, then MSDR solves MQCQP, $\mu^*_{\text{MQCQP}} = \mu^*_{\text{MSDR}}$ (see Beck [5], Beck and Teboulle [6]). However, the bound from MSDR is not tight in general.

To apply this to the QAP formulation in (6), we first reformulate it as MQCQP by removing B from the objective using the constraint $R = XB$:

$$\begin{aligned} \mu^*_{\mathrm{QAP}} = \min \quad & \operatorname{trace} \begin{pmatrix} X \\ R \end{pmatrix}^T \begin{pmatrix} 0 & \frac{1}{2}A \\ \frac{1}{2}A & 0 \end{pmatrix} \begin{pmatrix} X \\ R \end{pmatrix} + \operatorname{trace} CX^T \\ \text{s.t.} \quad & R = XB \\ & XX^T - I = X^T X - I = 0, \\ & Xe = X^T e = e, \\ & X \geq 0, \quad X \in \mathcal{M}^n. \end{aligned} \tag{7}$$

To linearize the objective function, we use

$$\operatorname{trace} \begin{pmatrix} X \\ R \end{pmatrix}^T \begin{pmatrix} 0 & \frac{1}{2}A \\ \frac{1}{2}A & 0 \end{pmatrix} \begin{pmatrix} X \\ R \end{pmatrix} = \operatorname{trace} \begin{pmatrix} 0 & \frac{1}{2}A \\ \frac{1}{2}A & 0 \end{pmatrix} \begin{pmatrix} X \\ R \end{pmatrix} \begin{pmatrix} X \\ R \end{pmatrix}^T$$

and the lifting

$$\begin{pmatrix} X \\ R \end{pmatrix} \begin{pmatrix} X \\ R \end{pmatrix}^T = \begin{pmatrix} XX^T & XR^T \\ RX^T & RR^T \end{pmatrix} = \begin{pmatrix} I & Y \\ Y & Z \end{pmatrix}. \tag{8}$$

This defines the symmetric matrices $Y, Z \in \mathcal{S}^n$, where we see $Y = RX^T = X(X^T R)X^T = XBX^T \in \mathcal{S}^n$. We can then relax this to get the convex quadratic constraint

$$G(X, R, Y, Z) := \begin{pmatrix} XX^T & XR^T \\ RX^T & RR^T \end{pmatrix} - \begin{pmatrix} I & Y \\ Y & Z \end{pmatrix} \preceq 0. \tag{9}$$

A Schur complement argument shows that the convex quadratic constraint (9) is equivalent to the linear conic constraint[1]

$$\begin{pmatrix} I & X^T & R^T \\ X & I & Y \\ R & Y & Z \end{pmatrix} \succeq 0. \tag{10}$$

The above discussion yields the MSDR relaxation for QAP:

$$\begin{aligned} (\mathrm{MSDR}_{\mathbf{0}}) \quad \mu^*_{\mathrm{QAP}} \geq \min \quad & \operatorname{trace} AY + \operatorname{trace} CX^T \\ \text{s.t.} \quad & R = XB, \\ & Xe = X^T e = e, \\ & \begin{pmatrix} I & X^T & R^T \\ X & I & Y \\ R & Y & Z \end{pmatrix} \succeq 0, \quad X \geq 0, \\ & X, R \in \mathcal{M}^n, \quad Y, Z \in \mathcal{S}^n, \end{aligned} \tag{11}$$

where Y represents or approximates $RX^T = XBX^T$ and Z represents or approximates $RR^T = XB^2X^T$. Because X is a *permutation matrix*, we conclude that the diagonal of Y is the X permutation of the diagonal of B (and, similarly, for the diagonals of Z and B^2):

$$\operatorname{diag}(Y) = X \operatorname{diag}(B), \qquad \operatorname{diag}(Z) = X \operatorname{diag}(B^2). \tag{12}$$

[1] Note that the linearized conic constraint is not onto, which suggests that it is more ill-conditioned than the convex quadratic constraint. Empirical tests in Ding et al. [8] confirm this.

Also, given that $Xe = X^T e = e$ and $Y = XBX^T$, $Z = XB^2X^T$ for all X, Y, Z feasible for the original QAP, we conclude that

$$Ye = XBe, \qquad Ze = XB^2e.$$

We may add these additional constraints to the above MSDR. These constraints essentially replace the orthogonality constraints. We get the first version of our SDP relaxation:

$$\begin{aligned}
(\mathrm{MSDR}_1) \quad \mu^*_{\mathrm{MSDR}_1} := \min \quad & \operatorname{trace} AY + \operatorname{trace} CX^T \\
\text{s.t.} \quad & Xe = X^T e = e, \\
& \left\{\begin{array}{l} \operatorname{diag}(Y) = X \operatorname{diag}(B) \\ \operatorname{diag}(Z) = X \operatorname{diag}(B^2) \\ Ye = XBe \\ Ze = XB^2 e \end{array}\right\}, \\
& \begin{pmatrix} I & X^T & (XB)^T \\ X & I & Y \\ XB & Y & Z \end{pmatrix} \succeq 0, \quad X \geq 0, \\
& X \in \mathcal{M}^n, \quad Y, Z \in \mathcal{S}^n.
\end{aligned}$$

Proposition 2.1. *Let B be nonsingular. In addition, suppose that (X, Y, Z) solves* MSDR_1 *and satisfies $Z = XB^2X^T$. Then, X is optimal for* QAP.

Proof. Via the Schur complement, we know that the semidefinite constraint in MSDR_1 is equivalent to

$$\begin{pmatrix} I - XX^T & Y - XBX^T \\ Y - XBX^T & Z - XB^2X^T \end{pmatrix} \succeq 0. \tag{13}$$

Therefore, $XX^T \preceq I$, $X^TX \preceq I$. Moreover, X satisfies $Xe = X^Te = e$, $X \geq 0$. Now, multiplying both sides of $\operatorname{diag}(Z) = X \operatorname{diag}(B^2)$ from the left by e^T yields $\operatorname{trace} Z = \operatorname{trace} B^2$. Because $Z = XB^2X^T$, we conclude that $\operatorname{trace} Z = \operatorname{trace} XB^2X^T = \operatorname{trace} B^2$, i.e., $\operatorname{trace} B^2(I - X^TX) = 0$. Because B is nonsingular, we conclude that $B^2 \succ 0$. Therefore, $I - X^TX \succeq 0$ implies that $I = X^TX$. Thus, the optimizer X is orthogonal and doubly stochastic ($X \in \mathcal{E} \cap \mathcal{N}$). Hence, X is a permutation matrix.

Moreover, (13) and $Z - XB^2X^T = 0$ imply the off-diagonal block $Y - XBX^T = 0$. Thus, we conclude that the bound μ^*_{MSDR1} from (MSDR_1) is tight. □

Remark 2.1. The assumption that B is nonsingular is made without loss of generality because we could shift B by a *small* positive multiple of the identity matrix, say ϵI, while simultaneously subtracting $\epsilon(\operatorname{trace} A)$, i.e.,

$$\begin{aligned}
\operatorname{trace}(AXBX^T + CX^T) &= \operatorname{trace}(AX(B + \epsilon I)X^T - \epsilon AXX^T + CX^T) \\
&= \operatorname{trace}(AX(B + \epsilon I)X^T + CX^T) - \epsilon \operatorname{trace} A.
\end{aligned}$$

2.2.1. The orthogonal similarity set of B. In this section, we include additional constraints in order to strengthen MSDR_1. Using majorization given in Definition 1.1, we now characterize the *convex hull of the orthogonal similarity set of* B, denoted $\operatorname{conv} \mathcal{O}(B)$.

Theorem 2.1. *Let*

$$\begin{aligned}
S_1 &:= \operatorname{conv} \mathcal{O}(B) = \operatorname{conv}\{Y \in \mathcal{S}^n \colon Y = XBX^T, X \in \mathcal{O}\}, \\
S_2 &:= \{Y \in \mathcal{S}^n \colon \operatorname{trace} \bar{A}Y \geq \langle \lambda(\bar{A}), \lambda(B) \rangle_-, \ \forall \bar{A} \in \mathcal{S}^n\}, \\
S_3 &:= \{Y \in \mathcal{S}^n \colon \operatorname{diag}(X^TYX) \preceq \lambda(B), \forall X \in \mathcal{O}\}, \\
S_4 &:= \{Y \in \mathcal{S}^n \colon \lambda(Y) \preceq \lambda(B)\}.
\end{aligned} \tag{14}$$

Then, S_1 is the convex hull of the orthogonal similarity set of B, and $S_1 = S_2 = S_3 = S_4$.

Proof. (i) $S_1 \subseteq S_2$: Let $Y \in S_1$, $\bar{A} \in \mathcal{S}^n$. Then,

$$\operatorname{trace} \bar{A}Y \geq \min_{Y \in \operatorname{conv}\mathcal{O}(B)} \operatorname{trace} \bar{A}Y = \min_{X \in \mathcal{O}} \operatorname{trace} \bar{A}XBX^T = \langle \lambda(\bar{A}), \lambda(B)\rangle_-$$

by the well-known minimal inner-product result (e.g., Rendl and Wolkowicz [33]), Finke et al. [11].

(ii) $S_2 \subseteq S_3$: Let $U \in \mathcal{O}$, $p \in \{1, 2, \ldots, n-1\}$, and let Γ_p denote the index set corresponding to the p smallest entries of $\operatorname{diag}(U^TYU)$. Define the support vector $\Delta^p \in \mathbb{R}^n$ of Γ_p by

$$(\Delta^p)_i = \begin{cases} 1 & \text{if } i \in \Gamma_p, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for $A_p := U \operatorname{Diag}(\Delta^p)U^T$, we get

$$\begin{aligned} \langle \Delta^p, \operatorname{diag}(U^TYU)\rangle &= \langle \operatorname{Diag}(\Delta^p), U^TYU\rangle \\ &= \langle U \operatorname{Diag}(\Delta^p)U^T, Y\rangle \\ &= \langle A_p, Y\rangle \\ &\geq \langle \Delta^p, \lambda(B)\rangle_- \end{aligned}$$

by definition of S_2. Because choosing $\bar{A} = \pm I$ implies $\operatorname{trace} Y = \operatorname{trace} B$, the inclusion follows.

(iii) $S_3 \subseteq S_4$: Let $Y \in S_3$ and let $Y = V \operatorname{Diag}(\lambda(Y))V^T$, $V \in \mathcal{O}$, be its spectral decomposition. Because $U \in \mathcal{O}$ implies that $\operatorname{diag}(U^TYU) \preceq \lambda(B)$, we may take $U = V$ and deduce

$$\lambda(Y) = \operatorname{diag}(V^TYV) \preceq \lambda(B).$$

(iv) $S_4 \subseteq S_1$: To obtain a contradiction, suppose $\lambda(\widehat{Y}) \preceq \lambda(B)$ but $\hat{Y} \notin \operatorname{conv}\mathcal{O}(B)$. Because $\mathcal{O}$ is a compact set, we conclude that the continuous image $\mathcal{O}(B) = \{Y\colon Y = XBX^T, X \in \mathcal{O}\}$ is compact. Hence, its convex hull $\operatorname{conv}\mathcal{O}(B)$ is compact as well. Therefore, a standard hyperplane separation argument implies that there exists $\bar{A} \in \mathcal{S}^n$ such that

$$\langle \bar{A}, \hat{Y}\rangle < \min_{Y \in \operatorname{conv}(\mathcal{O}(B))} \langle \bar{A}, Y\rangle = \min_{Y \in \mathcal{O}(B)} \langle \bar{A}, Y\rangle = \langle \lambda(\bar{A}), \lambda(B)\rangle_-.$$

As a result,

$$\langle \lambda(\bar{A}), \lambda(\hat{Y})\rangle_- \leq \langle \bar{A}, \hat{Y}\rangle < \langle \lambda(\bar{A}), \lambda(B)\rangle_-.$$

Without loss of generality, suppose that the eigenvalues $\lambda(\cdot)$ are in nondecreasing order. Then, the above minimum product inequality could be written as

$$\sum_{i=1}^n \lambda_i(\bar{A})\lambda_{n-i+1}(\hat{Y}) < \sum_{i=1}^n \lambda_i(\bar{A})\lambda_{n-i+1}(B),$$

which implies

$$0 > \sum_{i=1}^n \lambda_i(\bar{A})(\lambda_{n-i+1}(\hat{Y}) - \lambda_{n-i+1}(B)).$$

Because $\lambda_i(\bar{A}) = \sum_{j=1}^{i-1} (\lambda_{j+1}(\bar{A}) - \lambda_j(\bar{A})) + \lambda_1(\bar{A})$, we can rewrite the above inequality as

$$\begin{aligned} 0 &> \sum_{i=1}^n \left(\sum_{j=1}^{i-1} (\lambda_{j+1}(\bar{A}) - \lambda_j(\bar{A})) + \lambda_i(\bar{A})\right)(\lambda_{n-i+1}(\hat{Y}) - \lambda_{n-i+1}(B)) \\ &= \sum_{j=1}^{n-1} (\lambda_{j+1}(\bar{A}) - \lambda_j(\bar{A})) \sum_{i=j+1}^n (\lambda_{n-i+1}(\hat{Y}) - \lambda_{n-i+1}(B)) \\ &\quad + \lambda_1(\bar{A}) \sum_{i=1}^n (\lambda_i(\hat{Y}) - \lambda_i(B)). \end{aligned}$$

Notice that $\lambda(\hat{Y}) \preceq \lambda(B)$ implies $e^T\lambda(\hat{Y}) = e^T\lambda(B)$, so $\lambda_1(\bar{A}) \sum_{i=1}^n (\lambda_i(\hat{Y}) - \lambda_i(B)) = 0$. Thus, we have the following inequality:

$$0 > \sum_{j=1}^{n-1} (\lambda_{j+1}(\bar{A}) - \lambda_j(\bar{A})) \sum_{i=j+1}^n (\lambda_{n-i+1}(\hat{Y}) - \lambda_{n-i+1}(B)). \tag{15}$$

However, by assumption $\lambda_{j+1}(\bar{A}) \geq \lambda_j(\bar{A})$ and by the definition of $\lambda(\hat{Y})$ majorized by $\lambda(B)$,

$$\sum_{i=j+1}^{n} \lambda_{n-i+1}(\hat{Y}) = \sum_{t=1}^{n-j} \lambda_t(\hat{Y}) \geq \sum_{t=1}^{n-j} \lambda_t(B) = \sum_{i=j+1}^{n} \lambda_{n-i+1}(B),$$

which contradicts (15). $\square$

Remark 2.2. Based on our Theorem 2.1,[2] Xia [35] recognized that the equivalent sets S_1 to S_4 in (14) admit a semidefinite formulation, i.e.,

$$S_1 = S_5 := \left\{ Y \in S^n \colon Y = \sum_{i=1}^{n} \lambda_i(B) Y_i,\ \sum_{i=1}^{n} Y_i = I_n,\ \operatorname{trace} Y_i = 1,\ Y_i \succeq 0, i = 1, \ldots, n \right\}.$$

Xia [35] then proposed an orthogonal bound, denoted OB2, from the optimal value of the SDP

$$\mu^*_{\mathrm{OB2}} := \min_{X \geq 0,\, Xe = X^T e = e,\, Y \in S_5} \operatorname{trace}(AY + CX^T).$$

Note that this orthogonal bound OB2 can be applied to the projected version PQAP (given in §2.2.3), and then it is provably stronger than the convex quadratic programming bound QPB.

We failed to recognize this point in our initial work. Instead, motivated by Theorem 2.1, we now propose an inexpensive bound that is stronger than QPB for most of the problem instances we tested.

2.2.2. Strengthened MSDR bound. Suppose that $A = U_A \operatorname{Diag}(\lambda(A)) U_A^T$ denotes the orthogonal diagonalization of A with the vector of eigenvalues $\lambda(A)$ in nonincreasing order. We assume that the vector of eigenvalues $\lambda(B)$ is in nondecreasing order. Let

$$\delta^p := \{\overbrace{1, 1, \ldots, 1}^{p}, 0, 0, \ldots, 0\}, \quad p = 1, 2, \ldots, n-1.$$

We add the following cuts to MSDR_1:

$$\langle \delta^p, \operatorname{diag}(U_A^T Y U_A) \rangle \geq \langle \delta^p, \lambda(B) \rangle, \quad p = 1, 2, \ldots, n-1. \tag{16}$$

These are valid cuts because $\langle \delta^p, \operatorname{diag}(U_A^T Y U_A) \rangle \geq \langle \delta^p, \operatorname{diag}(U_A^T Y U_A) \rangle_- \geq \langle \delta^p, \lambda(B) \rangle_-$, for $Y \in S_1$ by part (ii) of the proof of Theorem 2.1.

Hence, we get the following relaxation:

$$\begin{array}{lll}
(\mathbf{MSDR_2}) & \mu^*_{\mathrm{MSDR}_2} := \min & \langle A, Y \rangle + \langle C, X \rangle \\
 & \text{s.t.} & Xe = X^T e = e \\
 & & \operatorname{diag}(Y) = X \operatorname{diag}(B) \\
 & & \operatorname{diag}(Z) = X \operatorname{diag}(B^2) \\
 & & Ye = XBe \\
 & & Ze = XB^2 e \\
 & & \langle \delta^p, \operatorname{diag}(U_A^T Y U_A) \rangle \geq \langle \delta^p, \lambda(B) \rangle, \quad p = 1, 2, \ldots, n-1 \\
 & & \begin{pmatrix} I & X^T & B^T X^T \\ X & I & Y \\ XB & Y & Z \end{pmatrix} \succeq 0, \quad X \geq 0 \\
 & & X \in \mathcal{M}^n, \quad Y, Z \in \mathcal{S}^n.
\end{array}$$

The cuts (16) approximate the majorization constraint

$$\operatorname{diag}(U_A^T Y U_A) \preceq \lambda(B) \tag{17}$$

and yield a comparison between the bounds MSDR_2 and EB.

[2] Xia [35] references our Theorem 2.1 from an earlier version of our paper.

LEMMA 2.1. *The bound from* MSDR$_2$ *is stronger than the eigenvalue bound* EB, *i.e.*,

$$\mu^*_{\mathrm{MSDR}_2} \geq \langle \lambda(A), \lambda(B)\rangle_- + \min_{Xe=X^Te=e,\, X\geq 0} \langle C, X\rangle.$$

PROOF. It is enough to show that the first terms on both sides of the inequality satisfy

$$\langle A, Y\rangle \geq \langle \lambda(A), \lambda(B)\rangle_-$$

for any Y feasible in MSDR$_\mathbf{2}$. Note that

$$\langle A, Y\rangle = \langle U_A \operatorname{Diag}(\lambda(A)) U_A^T, Y\rangle = \langle \lambda(A), \operatorname{diag}(U_A^T Y U_A)\rangle.$$

Because $\lambda(A)$ is a nonincreasing vector and $\lambda(B)$ is nondecreasing, we have $\langle \lambda(B), \lambda(A)\rangle = \langle \lambda(B), \lambda(A)\rangle_-$. Also,

$$\lambda(A) = \sum_{p=1}^{n-1} (\lambda_p(A) - \lambda_{p+1}(A))\delta^p + \lambda_n(A)e.$$

Therefore, because $\operatorname{diag}(Y) = X \operatorname{diag}(B)$ and $e^T X = e^T$, we have

$$\langle A, Y\rangle = \sum_{p=1}^{n-1} (\lambda_p(A) - \lambda_{p+1}(A))\langle \delta^p, \operatorname{diag}(U_A^T Y U_A)\rangle + \lambda_n(A)\langle e, \lambda(B)\rangle.$$

Because $\langle \delta^p, \operatorname{diag}(U_A^T Y U_A)\rangle \geq \langle \delta^p, \lambda(B)\rangle$ holds for any feasible Y, we have

$$\begin{aligned}
\langle A, Y\rangle &\geq \sum_{p=1}^{n-1} \big(\lambda_p(A) - \lambda_{p+1}(A)\big)\langle \delta^p, \lambda(B)\rangle + \lambda_n(A)\langle e, \lambda(B)\rangle \\
&= \sum_{p=1}^{n-1} \left(\big(\lambda_p(A) - \lambda_{p+1}(A)\big) \sum_{i=1}^{p} \lambda_i(B)\right) + \lambda_n(A) \sum_{i=1}^{n} \lambda_i(B) \\
&= \sum_{i=1}^{n} \lambda_i(B) \left(\sum_{p=i}^{n-1} \big(\lambda_p(A) - \lambda_{p+1}(A)\big) + \lambda_n(A)\right) \\
&= \sum_{i=1}^{n} \lambda_i(B)\lambda_i(A) \\
&= \langle \lambda(B), \lambda(A)\rangle_-. \quad \square
\end{aligned}$$

2.2.3. Projected bound. The row and column sum equality constraints of QAP, $\mathscr{E} = \{X \in \mathscr{M}^n\colon Xe = X^T e = e\}$, can be eliminated using a nullspace method. (In the following proposition, $\mathscr{O}$ refers to the orthogonal matrices of appropriate dimension.)

PROPOSITION 2.2 (HADLEY ET AL. [14]). *Let $V \in \mathscr{M}^{n,n-1}$ be full column rank and satisfy $V^T e = 0$. Then, $X \in \mathscr{E} \cap \mathscr{O}$ if and only if*

$$X = \frac{1}{n}E + V\widehat{X}V^T \quad \textit{for some } \widehat{X} \in \mathscr{O}. \quad \square$$

After substituting for X and using $\widehat{A} = V^T AV$, $\widehat{B} = V^T BV$, the QAP can now be reformulated as the projected version

$$\begin{aligned}
\text{(PQAP)} \quad \min \ & \operatorname{trace}\left(\widehat{A}\widehat{X}\widehat{B}\widehat{X}^T + \frac{1}{n}\widehat{A}\widehat{X}\widehat{B}E + \frac{1}{n}\widehat{A}E\widehat{B}\widehat{X}^T + \frac{1}{n^2}\widehat{A}E\widehat{B}E\right) \\
\text{s.t.} \ & \widehat{X}\widehat{X}^T = \widehat{X}^T\widehat{X} = I, \\
& X(\widehat{X}) = \frac{1}{n}E + V\widehat{X}V^T \geq 0.
\end{aligned}$$

We now define $\widehat{Y} = \widehat{X}\widehat{B}\widehat{X}^T$ and $\widehat{Z} = \widehat{Y}\widehat{Y} = \widehat{X}\widehat{B}\widehat{B}\widehat{X}^T$ and we replace X with $(1/n)E + V\widehat{X}V^T$. Then, the two terms XBX and $XBVV^TBX^T$ admit the representations

$$XBX^T = V\widehat{X}\widehat{B}\widehat{X}^T V^T + \frac{1}{n}EBV\widehat{X}^T V^T + \frac{1}{n}V\widehat{X}V^T BE + \frac{1}{n^2}E\widehat{B}E$$

and

$$XBVV^TBX^T = V\hat{Z}V^T + \frac{1}{n}EBVV^TBVX^TV^T + \frac{1}{n}VXV^TBVV^TBE + \frac{1}{n^2}EBVV^TBE,$$

respectively. In MSDR_2, we use Y to represent/approximate XBX^T and use Z to represent/approximate $XBBX^T$. However, $XBBX^T$ cannot be represented with $\hat{X}$ and $\hat{Y}$. Therefore, in the projected version we have to let Z represent $XBVV^TBX^T$ instead of $XBBX^T$, and we replace the corresponding diagonal constraint with $\operatorname{diag}(Z) = X\operatorname{diag}(BVV^TB)$.

Based on these definitions, PQAP has the following quadratic matrix programming formulation:

$$\begin{aligned}
\min\ & \operatorname{trace}(AY + CX^T)\\
\text{s.t. } & \operatorname{diag} Y = X\operatorname{diag}(B),\\
& \operatorname{diag} Z = X\operatorname{diag}(BVV^TB),\\
& X(\hat{X}) = V\hat{X}V^T + \frac{1}{n}E,\\
& Y(\hat{X},\hat{Y}) = V\hat{Y}V^T + \frac{1}{n}EBV\hat{X}^TV^T + \frac{1}{n}V\hat{X}V^TBE + \frac{1}{n^2}E\hat{B}E,\\
& Z(\hat{X},\hat{Z}) = V\hat{Z}V^T + \frac{1}{n}EBVV^TBVX^TV^T + \frac{1}{n}VXV^TBVV^TBE + \frac{1}{n^2}EBVV^TBE,\\
& \hat{R} = \hat{X}\hat{B},\\
& \begin{pmatrix} I & \hat{Y}\\ \hat{Y} & \hat{Z}\end{pmatrix} = \begin{pmatrix} \hat{X}\hat{X}^T & \hat{X}\hat{R}^T\\ \hat{R}\hat{X}^T & \hat{R}\hat{R}^T\end{pmatrix},\\
& X(\hat{X}) \geq 0,\\
& \hat{X}, \hat{R} \in \mathcal{M}^{n-1}, \quad \hat{Y}, \hat{Z} \in S^{n-1}.
\end{aligned} \tag{18}$$

We can now relax the quadratic constraint

$$\begin{pmatrix} I & \hat{Y}\\ \hat{Y} & \hat{Z}\end{pmatrix} = \begin{pmatrix} \hat{X}\hat{X}^T & \hat{X}\hat{R}^T\\ \hat{R}\hat{X}^T & \hat{R}\hat{R}^T\end{pmatrix}$$

with the convex constraint

$$\begin{pmatrix} I & \hat{X}^T & \hat{R}^T\\ \hat{X} & I & \hat{Y}\\ \hat{R} & \hat{Y} & \hat{Z}\end{pmatrix} \succeq 0.$$

As in MSDR_2, we now add the following cuts for $\hat{Y} \in \operatorname{conv}\mathcal{O}(\hat{X})$:

$$\langle \delta^p, \operatorname{diag}(U_{\hat{A}}^T\hat{Y}U_{\hat{A}})\rangle \geq \langle \delta^p, \lambda(\hat{B})\rangle, \quad p = 1, 2, \ldots, n-2,$$

where $\hat{A} = U_{\hat{A}}\operatorname{Diag}(\lambda(\hat{A}))U_{\hat{A}}^T$ is the spectral decomposition of $\hat{A}$ and $\lambda_1(\hat{A}) \leq \lambda_2(\hat{A}) \leq \cdots \leq \lambda_{n-1}(\hat{A})$. δ^p follows the definition in §2.2.1, i.e., $\delta^p \in R^{n-1}$, $\delta^p = \{0, 0, \ldots, 0, 1, \ldots, 1\}$. Our final projected relaxation MSDR_3 is

$$\begin{aligned}
(\mathrm{MSDR}_3)\quad \mu^*_{\mathrm{MSDR}_3} := \min\ & \langle A, Y(\hat{X},\hat{Y})\rangle + \langle C, X(\hat{X})\rangle\\
\text{s.t. } & \operatorname{diag}(Y(\hat{X},\hat{Y})) = X(\hat{X})\operatorname{diag}(B),\\
& \operatorname{diag}(Z(\hat{X},\hat{Z})) = X(\hat{X})\operatorname{diag}(BVV^TB),\\
& \langle \delta^p, \operatorname{diag}(U_{\hat{A}}^T\hat{Y}U_{\hat{A}})\rangle \geq \langle \delta^p, \lambda(\hat{B})\rangle, \quad p = 1, 2, \ldots, n-2,\\
& X(\hat{X}) \geq 0,\\
& \begin{pmatrix} I & \hat{X}^T & \hat{B}^T\hat{X}^T\\ \hat{X} & I & \hat{Y}\\ \hat{X}\hat{B} & \hat{Y} & \hat{Z}\end{pmatrix} \succeq 0,\\
& \hat{X} \in \mathcal{M}^{n-1}, \hat{Y}, \hat{Z} \in S^{n-1},
\end{aligned}$$

where

$$X(\widehat{X}) = \frac{1}{n}E + V\widehat{X}V^T,$$

$$Y(\widehat{X}, \widehat{Y}) = V\widehat{Y}V^T + \frac{1}{n}EBV\widehat{X}^TV^T + \frac{1}{n}V\widehat{X}V^TBE + \frac{1}{n^2}E\widehat{B}E,$$

$$Z(\widehat{X}, \widehat{Z}) = V\widehat{Z}V^T + \frac{1}{n}EBVV^TBVX^TV^T + \frac{1}{n}VXV^TBVV^TBE + \frac{1}{n^2}EBVV^TBE.$$

Note that the constraints $Ye = XBe$, $Ze = XB^2e$ are no longer needed in MSDR_3.

In MSDR_3, all the constraints act on the lower dimensional space obtained after the projection. The strategy of adding cuts after the projection has been successfully used in the projected eigenvalue bound PB and the quadratic programming bound QPB. For this reason, we propose MSDR_3 instead of $\mathrm{MSDR}_\mathbf{2}$.

Lemma 2.2. *Let μ^*_{PB} denote the projected eigenvalue bound. Then,*

$$\mu^*_{\mathrm{MSDR}_3} \geq \mu^*_{\mathrm{PB}}.$$

Proof. Because MSDR_3 has constraints

$$\langle \delta^p, \operatorname{diag}(U_{\widehat{A}}^T \widehat{Y} U_{\widehat{A}})\rangle \geq \langle \delta^p, \lambda(\widehat{B})\rangle, \quad p = 1, 2, \ldots, n-2,$$

we need only prove that $\operatorname{trace} \widehat{A}\widehat{Y} \geq \langle \lambda(\widehat{A}), \lambda(\widehat{B})\rangle_-$. This proof is the same as the proof for $\operatorname{trace} AY \geq \langle \lambda(A), \lambda(B)\rangle_-$ in Lemma 2.1. □

Remark 2.3. Every feasible solution to the original QAP satisfies $Y = XBX^T, X \in \Pi$. This implies that Y could be obtained from a permutation of the entries of B. Moreover, the diagonal entries of B remain on the diagonal after a permutation. Denote the off-diagonal entries of B by $0\mathrm{ffDiag}(B)$. We see that, for each $i, j = 1, 2, \ldots, n, i \neq j$, the following cuts are valid for any feasible Y:

$$\min[0\mathrm{ffDiag}(B)] \leq Y_{ij} \leq \max[0\mathrm{ffDiag}(B)]. \tag{19}$$

It is easy to verify that if the elements of $0\mathrm{ffDiag}(B)$ are all equal, then QAP can be solved by MSDR_1, $\mathrm{MSDR}_\mathbf{2}$, or MSDR_3 using the constraints in (19).

If B is diagonally dominant, than for any permutation X, we have that $Y = XBX^T$ is diagonally dominant. This property generates another series of cuts. These results could be used to add cuts for $Z = XB^2X^T$ as well.

3. Numerical results.

3.1. QAPLIB problems. In Table 1, we present a comparison of MSDR_3 with several other bounds applied to instances from QAPLIB (Burkard et al. [7]). The first column (OPT) denotes the exact optimal value. The following columns contain the Gilmore-Lawler bound (GLB) (Gilmore [12]); dual linear programming bound (KCCEB) (Karisch et al. [21], Hahn and Grant [16], Hahn and Grant [17]); projected eigenvalue bound (PB) (Hadley et al. [14]); convex quadratic programming bound (QPB) (Anstreicher and Brixius [2]); and the vector-lifting semidefinite relaxation bounds (SDR1, SDR2, and SDR3) (Zhao et al. [36]) computed by the bundle method (Rendl and Sotirov [32]). The last column is our MSDR_3 bound. All output values are rounded up to the nearest integer.

To solve QAP, the minimization of $\operatorname{trace} AXBX^T$ and $\operatorname{trace} BXAX^T$ are equivalent. However, in terms of the relaxation MSDR_3, exchanging the roles of A and B results in two different formulations and bounds. In our tests, we use the maximum of the two formulations for MSDR_3. When considering branching, we stay with the better formulation throughout to avoid doubling the computational work.

From Table 1, we see that the relative performance of the various bounds can vary on different instances. The *average* performance of the bounds can be ranked as follows:

$$\mathrm{PB} < \mathrm{QPB} < \mathrm{MSDR}_3 \approx \mathrm{SDR1} < \mathrm{SDR2} < \mathrm{SDR3}.$$

In Table 2, we present the number of variables and constraints used in each of the relaxations. Our bound MSDR_3 uses only $O(n^2)$ variables and only $O(n^2)$ constraints. If we solve MSDR_3 with an interior point method,

Table 1. Comparison of bounds for QAPLIB instances.

Problem	OPT	GLB	KCCEB	PB	QPB	SDR1	SDR2	SDR3	$MSDR_3$
Esc16a	68	38	41	47	55	47	49	59	50
Esc16b	292	220	274	250	250	250	275	288	276
Esc16c	160	83	91	95	95	95	111	142	123
Esc16d	16	3	4	−19	−19	−19	−13	8	1
Esc16e	28	12	12	6	6	6	11	23	14
Esc16g	26	12	12	9	9	9	10	20	13
Esc16h	996	625	704	708	708	708	905	970	906
Esc16i	14	0	0	−25	−25	−25	−22	9	0
Esc16j	8	1	2	−6	−6	−6	−5	7	0
Had12	1,652	1,536	1,619	1,573	1,592	1,604	1,639	1,643	1,595
Had14	2,724	2,492	2,661	2,609	2,630	2,651	2,707	2,715	2,634
Had16	3,720	3,358	3,553	3,560	3,594	3,612	3,675	3,699	3,587
Had18	5,358	4,776	5,078	5,104	5,141	5,174	5,282	5,317	5,153
Had20	692	6,166	6,567	6,625	6,674	6,713	6,843	6,885	6,681
Kra30a	88,900	68,360	75,566	63,717	68,257	69,736	68,526	77,647	72,480
Kra30b	91,420	69,065	76,235	63,818	68,400	70,324	71,429	81,156	73,155
Nug12	578	493	521	472	482	486	528	557	502
Nug14	1,014	852	N/a	871	891	903	958	992	918
Nug15	1,150	963	1,033	973	994	1,009	1,069	1,122	1,016
Nug16a	1,610	1,314	1,419	1,403	1,441	1,461	1,526	1,570	1,460
Nug16b	1,240	1,022	1,082	1,046	1,070	1,082	1,136	1,188	1,082
Nug17	1,732	1,388	1,498	1,487	1,523	1,548	1,619	1,669	1,549
Nug18	1,930	1,554	1,656	1,663	1,700	1,723	1,798	1,852	1,726
Nug20	2,570	2,057	2,173	2,196	2,252	2,281	2,380	2,451	2,291
Nug21	2,438	1,833	2,008	1,979	2,046	2,090	2,244	2,323	2,099
Nug22	3,596	2,483	2,834	2,966	3,049	3,140	3,372	3,440	3,137
Nug24	3,488	2,676	2,857	2,960	3,025	3,068	3,217	3,310	3,061
Nug25	3,744	2,869	3,064	3,190	3,268	3,305	3,438	3,535	3,300
Nug27	5,234	3,701	N/a	4,493	N/a	N/a	4,887	4,965	4,621
Nug30	6,124	4,539	4,785	5,266	5,362	5,413	5,651	5,803	5,446
Rou12	235,528	202,272	223,543	200,024	205,461	208,685	219,018	223,680	207,445
Rou15	354,210	298,548	323,589	296,705	303,487	306,833	320,567	333,287	303,456
Rou20	725,522	599,948	641,425	597,045	607,362	615,549	641,577	663,833	609,102
Scr12	31,410	27,858	29,538	4,727	8,223	11,117	23,844	29,321	18,803
Scr15	51,140	44,737	48,547	10,355	12,401	17,046	41,881	48,836	39,399
Scr20	110,030	86,766	94,489	16,113	23,480	28,535	82,106	94,998	50,548
Tai12a	224,416	195,918	220,804	193,124	199,378	203,595	215,241	222,784	202,134
Tai15a	388,214	327,501	351,938	325,019	330,205	333,437	349,179	364,761	331,956
Tai17a	491,812	412,722	441,501	408,910	415,576	419,619	440,333	451,317	418,356
Tai20a	703,482	580,674	616,644	575,831	584,938	591,994	617,630	637,300	587,266
Tai25a	1,167,256	962,417	1,005,978	956,657	981,870	974,004	908,248	1,041,337	970,788
Tai30a	1,818,146	1,504,688	1,565,313	1,500,407	1,517,829	1,529,135	1,573,580	1,652,186	1,521,368
Tho30	149,936	90,578	99,855	119,254	124,286	125,972	134,368	136,059	122,778

the complexity of computing the Newton direction in each iteration is $O(n^6)$ and the number of iterations of an interior point method is bounded by $O(n\ln(1/\epsilon))$ (Monteiro and Todd [26]). Therefore, the complexity of computing $MSDR_3$ with an interior point method is $O(n^7\ln(1/\epsilon))$. Note that the computational complexity for the most expensive SDP formulation, SDR3, is $O(n^{14}\ln(1/\epsilon))$ where ϵ is the desired accuracy. Thus, $MSDR_3$ is significantly less expensive than SDR3. Though QPB is less expensive than $MSDR_3$ in practice the complexity as a function of n is the same.

Table 2. Complexity of relaxations.

7 Methods	GLB	KCCEB	PB	QPB	SDR1	SDR2	SDR3	$MSDR_3$
Variables	$O(n^4)$	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^4)$	$O(n^4)$	$O(n^4)$	$O(n^2)$
Constraints	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^3)$	$O(n^4)$	$O(n^2)$

Table 3. CPU time and iterations for computing MSDR_3 on the Nugent problems.

Instances	Nug12	Nug15	Nug18	Nug20	Nug25	Nug27	Nug30
CPU time(s)	15.1	57.6	203.9	534.9	3,236.4	5,211.3	12,206.0
Number of iterations	18	19	22	26	27	25	29

Table 3 lists the CPU time (in seconds) for MSDR_3 for several of the *Nugent instances* (Nugent et al. [28]). (We used a SUN SPARC 10 and the SeDuMi[3] SDP package. For a rough comparison, note that the results in Anstreicher et al. [4] were done on a C3000 computer and took 3.2 CPU seconds for the Nug20 instance and 9 CPU seconds for the Nug25 instance for the QPB bound.)

3.2. MSDR_3 in a branch-and-bound framework. When solving general discrete optimization problems using B&B methods, one rarely has advance knowledge that helps in branching decisions. We now see that MSDR_3 helps in choosing a row and/or column for branching in our B&B approach for solving QAP.

If X is a permutation matrix, then the diagonal entries $\operatorname{diag}(Z) = X \operatorname{diag}(BVV^TB)$ are a permutation of the diagonal entries of BVV^TB. In fact, the converse is true under a mild assumption.

Proposition 3.1. *Assume the n entries of* $\operatorname{diag}(BVV^TB)$ *are all distinct. If* (X^*, Y^*, Z^*) *is an optimal solution to* MSDR_3 *that satisfies* $\operatorname{diag}(Z^*) = P\operatorname{diag}(BVV^TB)$ *for some* $P \in \Pi$, *then* (X^*, Y^*, Z^*) *solves QAP exactly.*

Proof. Without loss of generality, assume the entries of $b := \operatorname{diag}(BVV^TB)$ are strictly increasing, i.e., $b_1 < b_2 < \cdots < b_n$. By the feasibility of X^*, Z^*, we have $\operatorname{diag}(Z^*) = X^*b$. Also, we know $\operatorname{diag}(Z^*) = Pb$ for some $P \in \Pi$. Therefore, $X^*b = Pb$ holds as well. Now, assume $P_{i1} = 1$. Then, $\sum_{j=1}^n X_{ij}^* b_j = b_1$. Because $\sum_{j=1}^n X_{ij}^* = 1$ and $X_{ij}^* \geq 0$, $j = 1, 2, \ldots, n$, we conclude that b_1 is a convex combination of $b_1, b_2, \ldots, b_n$. However, b_1 is the strict minimum in $b_1, b_2, \ldots, b_n$. This implies that $X_{i1}^* = 1$. The conclusion follows for $P = X^*$ by finite induction after we delete column one and row i of X. □

As a consequence of Proposition 3.1, we may consider the original QAP problem in order to determine an optimal assignment of entries of $\operatorname{diag}(BVV^TB)$ to $\operatorname{diag}(Z)$, where each entry of $\operatorname{diag}(BVV^TB)$ requires a branch-and-bound process to determine its assigned position. For entries with a large difference from the mean of $\operatorname{diag}(BVV^TB)$, the assignments are particularly important because a change of their assigned positions usually leads to significant differences in the corresponding objective value. Therefore, in order to fathom more nodes early, our B&B strategy first processes those entries with large differences from the mean of $\operatorname{diag}(BVV^TB)$.

Branch-and-Bound Strategy 3.1. *Let* $b := \operatorname{diag}(BVV^TB)$. *Branch on the ith column of X where i corresponds to the element b_i that has the largest deviation from the mean of the elements of b.* (*If this strategy results in several elements close in value, then we randomly pick one of them.*)

For example, Nug12 yields

$$\operatorname{diag}(BVV^TB)^T = \begin{pmatrix} 23 & 14 & 14 & 23 & 17.67 & 8.67 & 8.67 & 17.67 & 23 & 14 & 14 & 23 \end{pmatrix}.$$

Therefore, the sixth or seventh entry has value 8.67; this has the largest difference from the mean value 16.72. Table 4 presents the MSDR_3 bounds in the first level of the branching tree for Nug12. The first and second columns of Table 4 present the results for branching on elements from the sixth column of X first. The other columns provide a comparison with branching from other columns first. On average, branching with the sixth column of X first generates tighter bounds and should lead to descendant nodes in the branch-and-bound tree that was fathomed earlier.

4. Conclusion. We have presented new bounds for QAP that are based on a matrix-lifting (rather than a vector-lifting) semidefinite relaxation. By exploiting the special doubly stochastic and orthogonality structure of the constraints, we obtained a series of cuts to further strengthen the relaxation. The resulting relaxation MSDR_3 is provably stronger than the projected eigenvalue bound PB and is comparable with the SDR1 bound and the quadratic programming bound QPB in our empirical tests. Moreover, due to the matrix-lifting property of the bound, it only uses $O(n^2)$ variables and $O(n^2)$ constraints. Hence, the complexity is comparable with that of QPB.

[3] Information is available at http://sedumi.ie.lehigh.edu.

Table 4. Results for the first level branching for Nug12.

Nodes	Bounds	Nodes	Bounds	Nodes	Bounds
$X_{1,6}=1$	523	$X_{1,1}=1$	508	$X_{1,2}=1$	512
$X_{2,6}=1$	528	$X_{2,1}=1$	509	$X_{2,2}=1$	513
$X_{3,6}=1$	520	$X_{3,1}=1$	507	$X_{3,2}=1$	508
$X_{4,6}=1$	517	$X_{4,1}=1$	515	$X_{4,2}=1$	510
$X_{5,6}=1$	537	$X_{5,1}=1$	512	$X_{5,2}=1$	519
$X_{6,6}=1$	529	$X_{6,1}=1$	517	$X_{6,2}=1$	513
$X_{7,6}=1$	507	$X_{7,1}=1$	516	$X_{7,2}=1$	507
$X_{8,6}=1$	519	$X_{8,1}=1$	524	$X_{8,2}=1$	513
$X_{9,6}=1$	522	$X_{9,1}=1$	524	$X_{9,2}=1$	514
$X_{10,6}=1$	527	$X_{10,1}=1$	514	$X_{10,2}=1$	513
$X_{11,6}=1$	506	$X_{11,1}=1$	527	$X_{11,2}=1$	510
$X_{12,6}=1$	504	$X_{12,1}=1$	510	$X_{12,2}=1$	516
Mean	519.9	Mean	515.3	Mean	512.3

Subsequent work has shown that our MSDR_3 relaxation and bound are particularly efficient for matrices with special structures, for example, if B is a Hamming distance matrix of a hypercube or a Manhattan distance matrix from rectangular grids (see, e.g., Mittelmann and Peng [24]). Additional new relaxations based on our work have been proposed (see, e.g., the bound OB2 in Xia [35]). Another recent application is decoding in multiple antenna systems (see Mobasher and Khandani [25]).

Acknowledgments. The authors are indebted to the associate editor and two anonymous referees for helping make numerous significant improvements to the paper.

References

[1] Anstreicher, K. M. 2003. Recent advances in the solution of quadratic assignment problems. *Math. Programming Ser. B* **97**(1-2) 27–42.
[2] Anstreicher, K. M., N. W. Brixius. 2001. A new bound for the quadratic assignment problem based on convex quadratic programming. *Math. Programming Ser. A* **89**(3) 341–357.
[3] Anstreicher, K. M., H. Wolkowicz. 2000. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.* **22**(1) 41–55.
[4] Anstreicher, K. M., N. W. Brixius, J.-P. Goux, J. Linderoth. 2002. Solving large quadratic assignment problems on computational grids. *Math. Programming Ser. A* **91**(3) 563–588.
[5] Beck, A. 2006. Quadratic matrix programming. *SIAM J. Optim.* **17**(4) 1224–1238.
[6] Beck, A., M. Teboulle. 2000. Global optimality conditions for quadratic optimization problems with binary constraints. *SIAM J. Optim.* **11**(1) 179–188.
[7] Burkard, R. E., S. Karisch, F. Rendl. 1991. QAPLIB—A quadratic assignment problem library. *Eur. J. Oper. Res.* **55** 115–119. (http://www.opt.math.tu-graz.ac.at/qaplib/.)
[8] Ding, Y., N. Krislock, J. Qian, H. Wolkowicz. 2008. Sensor network localization, Euclidean distance matrix completions, and graph realization. *Internat. Conf. Mobile Comput. Networking.* ACM, New York, 129–134.
[9] Drezner, Z. 1995. Lower bounds based on linear programming for the quadratic assignment problem. *Comput. Optim. Appl.* **4**(2) 159–165.
[10] Falkner, J., F. Rendl, H. Wolkowicz. 1994. A computational study of graph partitioning. *Math. Programming Ser. A* **66**(2) 211–239.
[11] Finke, G., R. E. Burkard, F. Rendl. 1987. Quadratic assignment problems. *Ann. Discrete Math.* **31** 61–82.
[12] Gilmore, P. C. 1962. Optimal and suboptimal algorithms for the quadratic assignment problem. *SIAM J. Appl. Math.* **10** 305–313.
[13] Hadley, S. W., F. Rendl, H. Wolkowicz. 1990. Bounds for the quadratic assignment problems using continuous optimization. *Integer Programming and Combinatorial Optimization*. University of Waterloo Press, Waterloo, ON, Canada, 237–248.
[14] Hadley, S. W., F. Rendl, H. Wolkowicz. 1992. A new lower bound via projection for the quadratic assignment problem. *Math. Oper. Res.* **17**(3) 727–739.
[15] Hadley, S. W., F. Rendl, H. Wolkowicz. 1992. Symmetrization of nonsymmetric quadratic assignment problems and the Hoffman-Wielandt inequality. *Linear Algebra Appl.* **167** 53–64.
[16] Hahn, P., T. Grant. 1998. Lower bounds for the quadratic assignment problem based upon a dual formulation. *Oper. Res.* **46**(6) 912–922.
[17] Hahn, P., T. Grant, N. Hall. 1998. A branch-and-bound algorithm for the quadratic assignment problem based on the Hungarian method. *Eur. J. Oper. Res.* **108**(3) 629–640.
[18] Hardy, G. H., J. E. Littlewood, G. Polya. 1934. *Inequalities*. Cambridge University Press, London and New York.
[19] Horn, R. A., C. R. Johnson. 1994. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, MA. (Corrected reprint of the 1991 original.)
[20] Karisch, S. E., F. Rendl, H. Wolkowicz. 1994. Trust regions and relaxations for the quadratic assignment problem. P. Pardalos, H. Wolkowicz, eds. *Quadratic Assignment and Related Problems*. American Mathematical Society, Providence, RI, 199–219.
[21] Karisch, S. E., E. Çela, J. Clausen, T. Espersen. 1999. A dual framework for lower bounds of the quadratic assignment problem based on linearization. *Computing* **63**(4) 351–403.

[22] Lawler, E. 1963. The quadratic assignment problem. *Management Sci.* **9** 586–599.

[23] Marshall, A. W., I. Olkin. 1979. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

[24] Mittelmann, H., J. Peng. 2007. Estimating bounds for quadratic assignment problems associated with Hamming and Manhattan distance matrices based on semidefinite programming. Technical report, University of Illinois at Urbana-Champaign, Urbana, IL.

[25] Mobasher, A., A. K. Khandani. 2007. Matrix-lifting semi-definite programming for decoding in multiple antenna systems. *Proc. 10th Canadian Workshop Inform. Theory* (*CWIT'07*)*, Edmonton, Alberta, Canada*, 136–139.

[26] Monteiro, R. D. C., M. J. Todd. 2000. Path-following methods. *Handbook of Semidefinite Programming*. Kluwer Academic Publishers, Boston, 267–306.

[27] Nesterov, Y. E., H. Wolkowicz, Y. Ye. 2000. Semidefinite programming relaxations of nonconvex quadratic optimization. *Handbook of Semidefinite Programming*, Vol. 27. *International Series Operation Research Management Science.* Kluwer Academic Publishers, Boston, 361–419.

[28] Nugent, C. E., T. E. Vollman, J. Ruml. 1968. An experimental comparison of techniques for the assignment of facilities to locations. *Oper. Res.* **16** 150–173.

[29] Pardalos, P. M., H. Wolkowicz, eds. 1994. *Quadratic Assignment and Related Problems*. American Mathematical Society, Providence, RI.

[30] Pardalos, P., F. Rendl, H. Wolkowicz. 1994. The quadratic assignment problem: A survey and recent developments. P. Pardalos, H. Wolkowicz, eds. *Quadratic Assignment and Related Problems*. American Mathematical Society, Providence, RI, 1–42.

[31] Pardalos, P. M., K. G. Ramakrishnan, M. G. C. Resende, Y. Li. 1997. Implementation of a variance reduction-based lower bound in a branch-and-bound algorithm for the quadratic assignment problem. *SIAM J. Optim.* **7**(1) 280–294.

[32] Rendl, F., R. Sotirov. 2007. Bounds for the quadratic assignment problem using the bundle method. *Math. Programming Ser. B* **109**(2–3) 505–524.

[33] Rendl, F., H. Wolkowicz. 1992. Applications of parametric programming and eigenvalue maximization to the quadratic assignment problem. *Math. Programming Ser. A* **53**(1) 63–78.

[34] Wolkowicz, H. 2000. Semidefinite programming approaches to the quadratic assignment problem. *Nonlinear Assignment Problems*, Vol. 7. *Combinatorial Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 143–174.

[35] Xia, Y. 2007. New semidefinite relaxations for the quadratic assignment problem. Personal communication.

[36] Zhao, Q., S. E. Karisch, F. Rendl, H. Wolkowicz. 1998. Semidefinite programming relaxations for the quadratic assignment problem. *J. Comb. Optim.* **2**(1) 71–109.