DIMACS Series in Discrete Mathematics
and Theoretical Computer Science
Volume 00, 0000

TRUST REGIONS and RELAXATIONS for the QUADRATIC ASSIGNMENT PROBLEM

STEFAN E. KARISCH, FRANZ RENDL, AND HENRY WOLKOWICZ

ABSTRACT. General quadratic matrix minimization problems, with orthogonal constraints, arise in continuous relaxations for the (discrete) quadratic assignment problem (QAP). Currently, bounds for QAP are obtained by treating the quadratic and linear parts of the objective function, of the relaxations, separately. This paper handles general objectives as one function. The objectives can be both nonhomogeneous and nonconvex. The constraints are orthogonal or Loëwner partial order (positive semidefinite) constraints. Comparisons are made to standard trust region subproblems. Numerical results are obtained using a parametric eigenvalue technique.

CONTENTS

1991 *Mathematics Subject Classification.* Primary 90B80, 90C20, 90C35, 90C27; Secondary 65H20, 65K05.

Key words and phrases. Combinatorial optimization, quadratic assignment problem, trust regions, optimality conditions, algorithms.

This paper is in final form and no version of it will be submitted for publication elsewhere.

This report is issued simultaneously at: Institut für Mathematik, Technische Universität Graz; the Department of Combinatorics and Optimization, University of Waterloo, Report CORR 93-15; the Department of Civil Engineering and Operations Research, Princeton University, Report SOR 93-10 and DIMACS Center, Rutgers University, Report 93-34.

1. Introduction

Consider the general, equality constrained, matrix quadratic programming problem

$$(EP) \quad \begin{array}{ll} \min & \operatorname{tr}(AXBX^t + 2CX^t) \\ \text{s.t.} & XX^t = I, \end{array}$$

where I denotes the identity matrix, $\cdot^t$ means *transpose*, tr stands for *trace*, and A, B and C are real $n \times n$ matrices. This problem can be viewed as a matrix version of the well known trust region subproblem for unconstrained minimization, i.e.,

$$(TR) \ \min q(x) \text{ subject to } x^t x \leq \delta, \ x \in \Re^n,$$

where $q(x) = x^t Q x + c^t x$ is a quadratic, not necessarily convex, function on $\Re^n$. Characterizations of optimality and efficient numerical algorithms exist for (TR). In addition, in the homogeneous case ($C = 0, \ c = 0$), both problems reduce to eigenvalue problems. But the important problem of efficiently solving the general nonhomogeneous (EP) is still open.

After scaling (TR), the unit ball $\delta \leq 1$, can be assumed. Moreover, the hard part of the trust region subproblem is dealing with the case $x^t x = 1$, since the trivial < 1 case occurs only when q is convex with optimum in the interior of the ball. A characterization of optimality for (TR) holds without any gap between necessity and sufficiency, even in the absence of convexity. Eigenvalue type algorithms can be applied to quickly and efficiently solve the problem. (See e.g., [**10, 25**].) This is due to the fact that these problems are implicit convex problems. In fact, a dual program exists that consists in the maximization of a concave function over an interval [**33**]. However, the addition of a second trust region can create great difficulties both in the theory and the algorithms; see e.g., [**15, 24, 34, 35, 36**].

The general matrix quadratic programming problem (EP) has a quadratic objective function and $(n^2 + n)/2$ quadratic constraints. We further relax the orthogonal constraint $XX^t - I = 0$ to $XX^t - I = N$, where N is negative semidefinite; equivalently

$$XX^t - I \preceq 0,$$

is the partial ordering given by the positive semidefinite matrices, i.e., the Loëwner partial order. (We denote this latter problem by (P).) These problems resemble (TR) visually. Motivated by this, we extend the existing theory for (TR) to (P) and (EP). We also present a parametric eigenvalue approach to bound the solution of (P) and (EP).

The paper is organized as follows. In Section 2 we first present preliminary definitions and concepts. We also include a motivation for studying the matrix quadratic problems. We will show that the problems we investigate are relaxations of the *quadratic assignment problem* denoted QAP. We survey the theory of eigenvalue bounds for QAP that led to this relaxation in Section 2.2. In Section 3 we present first and second order optimality conditions for (P) and (EP). This includes conditions that guarantee orthogonality of the optimum in (P). In Section 4 we present perturbations of the matrices A, B using the full rank factorization of the linear term matrix C. This yields a parametric homogeneous eigenvalue problem, which is used to approximate the optimal solution of the general problem (P). Thereby, we obtain new lower bounds for QAP. Numerical tests are included.

2. Preliminary Notations and Motivation

2.1. Notations. We will use the following notation throughout the paper.

If S is symmetric then it has an orthogonal diagonalization $S = UDU^t$, where the eigenvalues of S are ordered

$$\lambda_1(S) \geq \lambda_2(S) \geq \ldots \geq \lambda_n(S).$$

Recall that the spectrum of S is real since S is symmetric. We will use $\lambda(S) := (\lambda_i(S))$ to denote the vector of eigenvalues. The set of positive semidefinite matrices is denoted $\mathcal{P}$ or psd. We let $A \preceq B$ denote the Loëwner partial order, i.e., $A-B$ is negative semidefinite (nsd). Similar definitions hold for $A \prec B$, $A \succeq B$, $A \succ B$.

We denote the vector of all ones of size n by $u := (1, \ldots, 1)^t \in \Re^n$. The vector of *row sums* of the matrix $S \in \Re^{n \times n}$ is $r(S)$, and the *sum of all the entries* of S is $s(S)$, that is $r(S) := Su$ and $s(S) := u^t Su$.

The vector u_k denotes the k-th unit vector, i.e., the k-th column of the identity matrix; while the matrix $E_k := u_k u_k^t$. We let $\operatorname{diag}(v)$ denote the diagonal matrix formed from the vector v and conversely, $\operatorname{diag}(S)$ is the vector of the diagonal elements of the matrix S. For a set K, we let $\operatorname{int}(K)$ denote interior and $\overline{K}$ denote closure.

We also need to define the *minimal scalar product* of two vectors α and $\beta \in \Re^n$. It is given by

$$\langle \alpha, \beta \rangle_- := \min\{\sum_{i=1}^n \alpha_i \beta_{\phi(i)} : \phi \text{ permutation }\}.$$

The *maximal scalar product* $\langle \alpha, \beta \rangle_+$ is defined analogously.

2.2. A Survey on Eigenvalue Bounds for the QAP. The *quadratic assignment problem (QAP)* consists in minimizing a quadratic over the set of permutations. In the *trace formulation* QAP is the problem

$$\mathbf{Q}AP \quad \min_{X \in \Pi} \; f(X) = \mathrm{tr}(AXB^t + C)X^t,$$

where Π is the set of permutation matrices. We assume in addition that A and B are real symmetric $n \times n$ matrices and $C \in \Re^{n \times n}$.

The QAP is an *NP*-hard combinatorial optimization problem since the special case of QAP, the travelling salesman problem, is well known to be *NP*-hard. QAP even belongs to the hard core of *NP*-hard problems since finding an ε-approximation of the optimal solution proves to be *NP*-hard. For the complexity proofs, see [**31**]. QAP is a very hard problem in practice as well, since instances of size $n \geq 15$ can prove to be intractable; see e.g., [**30**].

Since current solution techniques employ branch and bound methods, one has to improve the quality of lower bounds in order to be able to solve larger problems. The remainder of this section will show that (P) can also be seen as a relaxation of QAP, which therefore yields bounds for QAP. A different approach for obtaining bounds is presented in this proceedings in [**19**].

Eigenvalue bounds employ the trace formulation of QAP and use the observation that the set of permutation matrices satisfies

$$\Pi = \mathcal{O} \cap \mathcal{E} \cap \mathcal{N}. \tag{2.1}$$

Here $\mathcal{O} := \{X : X^tX = I\}$ is the set of *orthogonal* matrices, while $\mathcal{E} := \{X : Xu = X^tu = u\}$ is the set of all matrices having *row and column sums equal to one*, and $\mathcal{N} := \{X : X \geq 0\}$ is the set of *nonnegative* matrices.

The basic idea for eigenvalue bounds is to enlarge the feasible set Π to get a tractable problem. The orthogonal relaxation (i.e., optimizing over $\mathcal{O}$ instead of Π) was applied to the QAP in [**7, 29**] and makes use of the following fact, which can be viewed as a variant of a classical inequality commonly referred to as the "Hoffman- Wielandt Inequality".

THEOREM 2.1 ([**5, 7, 29**]). *Let A and B be real symmetric $n \times n$ matrices. Then*

$$\langle \lambda(A), \lambda(B) \rangle_- \leq trAXBX^t \leq \langle \lambda(A), \lambda(B) \rangle_+, \; \forall X \in \mathcal{O}.$$

Moreover, the lower (upper) bound is attained for $X = PQ^t$, where $P, Q \in \mathcal{O}$ contain the eigenvectors of A and B, respectively, in the order prescribed by the minimal (maximal) scalar product of the eigenvalues.

This result can be used as the basis for the calculation of eigenvalue bounds. The basic bound proposed in [**7**] was obtained by bounding the quadratic part of QAP by Theorem 2.1 and by solving the linear part separately. This yields the following lower bound for QAP

$$QAP(A, B, C) \geq \langle \lambda(A), \lambda(B) \rangle_- + LSAP(C). \tag{2.2}$$

LSAP(C) stands for the solution of the *linear sum assignment problem* with cost matrix C.

The quality of this bound is in general rather poor and was further improved by transformations of the objective function, called "shifts" and "reductions". These transformations consist in adding constants to the quadratic part and in appropriately modifying the linear part in order to keep the objective function value unchanged over the permutation matrices. A simple way to select these transformations was proposed in [**7**] and led to the bound labeled EVB2; see Table 1 below.

In [**29**] an iterative improvement technique was developed to find shifts and reductions for making the sum of the quadratic and linear bounds in (2.2) as large as possible. The parametric programming approach that was given in [**29**] resulted in a new bound EVB3. EVB3 is in most cases the best eigenvalue bound available. But since it is computationally very expensive, it is not suitable for branch and bound methods. Again we refer to Table 1.

Recently in [**14**], the relaxation of the feasible set was strengthened by optimizing over the smaller set $X \in \mathcal{O} \cap \mathcal{E}$. This is done by elimination of the constraints $\mathcal{E}$ and yields an equivalent projected problem PQAP. For the projection, an $n \times (n-1)$ matrix V such that $V^t u = 0$ and $V^t V = I_{n-1}$ was introduced. The new $(n-1)$-dimensional problem was then

$$\begin{array}{lll} \textbf{PQAP} & \min & \operatorname{tr}(\hat{A} Y \hat{B}^t + \hat{C}) Y^t + \text{const} \\ & \text{s.t.} & Y \in \mathcal{O}_{n-1},\ V Y V^t \geq -uu^t/n. \end{array}$$

where $\hat{A} = V^t A V$, $\hat{B} = V^t B V$, $\hat{C} = \frac{2}{n} V^t r(A) r^t(B) V + V^t C V$, and const $= \left[s(A)s(B)/n + s(C)\right]/n$.

By ignoring the projected nonnegativity constraints $VYV^t \geq -uu^t/n$, we obtain the eigenvalue bound EVB4

$$QAP(A,B,C) = PQAP(\hat{A},\hat{B},\hat{C}) \geq \left\langle \lambda(\hat{A}), \lambda(\hat{B}) \right\rangle_- + LSAP(\hat{C}) + \text{const}. \tag{2.3}$$

Before closing this section we look at the rank of the linear term of PQAP. Both $r(A)$ and $r(B)$ are at most rank 1, therefore $V^t r(A) r^t(B) V$ is also at most rank 1. The rank of the second part in $\hat{C}$ depends strongly on C, since

$$\text{rank } C - 2 \leq \text{rank } (V^t C V) \leq \min(\text{rank } V, \text{rank } C).$$

We can see that if QAP is pure quadratic, that is $C = 0$, the linear part will become at most rank 1. Below we exploit this rank 1 property when finding bounds using perturbed problems. The problem is further simplified if $\hat{C} = 0$, which occurs if $C = 0$ and either $r(A)$ or $r(B)$ are constant vectors.

2.3. Loëwner Partial Order. The above provides motivation for the study of (EP). In the context of QAP we can think of (EP) as being the relaxation of the projected instance PQAP of an $(n+1) \times (n+1)$ dimensional QAP. By treating the quadratic and linear parts of (EP) together one should expect better bounds, as shown in [**29**] for EVB3.

We now further relax the orthogonal constraint $X \in \mathcal{O}$ to $XX^t - I = N$ where N is a negative semidefinite matrix. For notional convenience we add a factor of 2 to the linear part. We then get the following relaxation to QAP

$$\begin{array}{lll} (P) & \min & f(X) = \mathrm{tr} AXBX^t + 2CX^t \\ & \text{s.t.} & g(X) = XX^t - I \preceq 0, \end{array} \tag{2.4}$$

where $\preceq$ refers to the Loëwner partial order defined above, see e.g., [**16**].

Let us now look at some of the properties of the constraint $g(X)$ of (P), which defines an operator from $\Re^{n\times n}$ to the space of $n \times n$ symmetric matrices. Note that g is $\mathcal{P}$-convex, i.e., for any $X_1, X_2 \in \Re^{n\times n}$ and any $\lambda \in [0,1]$ we have

$$\lambda g(X_1) + (1-\lambda) g(X_2) \succeq g(\lambda X_1 + (1-\lambda)X_2)$$

or equivalently

$$\lambda g(X_1) + (1-\lambda) g(X_2) - g(\lambda X_1 + (1-\lambda)X_2) \in \mathcal{P}.$$

When the eigenvalues of S are ordered

$$\lambda_1(S) \geq \lambda_2(S) \geq \ldots \geq \lambda_n(S),$$

then we can see that

$$S \preceq 0 \text{ if and only if } \lambda_1(S) \leq 0.$$

Thus we can replace the cone constraint in (P) by the scalar constraint

$$\lambda_1 = \lambda_1(g(X)) \leq 0.$$

The function $\lambda_1(g(X))$ is a convex function of X. We can find the derivative of $\lambda_k(g(X))$ using the corresponding normalized eigenvector v_k of $g(X)$. If λ_k is simple, see e.g., [**17**], then

$$\begin{aligned} \frac{\partial \lambda_k(g(X))}{\partial X_{ij}} &= v_k^t \frac{\partial g(X)}{\partial X_{ij}} v_k \\ &= \frac{1}{2} v_k^t (E_{ij} X^t + X E_{ij}^t) v_k \\ &= v_k(i)(X_{.j}^t v_k), \end{aligned} \tag{2.5}$$

where E_{ij} is the zero matrix with 1 in the i,j-position, $X_{.j}$ is the j-th column of X and $v_k(i)$ is the i-th component of v_k. If λ_k is not simple, then the function λ_k may not be differentiable. Since we want X orthogonal at the solution, we can expect multiple eigenvalues of 1. It is well known that the largest eigenvalue is convex and so we can obtain expressions for the subdifferentials of the largest eigenvalue if its multiplicity is > 1, see e.g., [**8**], [**13**],[**27, 28**]. These subdifferentials are just the convex hull, of the expression given above for the derivative, over all normalized eigenvectors v_k.

Note that the differentials at X in the direction h of the above functions in (P) are:

$$\begin{aligned} df(X;h) &= \operatorname{tr} A(XBh^t + hB^tX^t) + 2Ch^t \\ dg(X;h) &= Xh^t + hX^t. \end{aligned} \tag{2.6}$$

Moreover

$$df(X;h) = \operatorname{tr}\, 2AXBh^t + 2Ch^t,$$

since A and B are symmetric.

We use the inner-product

$$\langle S, T\rangle = \operatorname{tr} ST^t$$

on the space of $n \times n$ matrices. In the space of symmetric matrices with this inner-product, $\mathcal{P}$ is a closed convex cone with nonempty interior, $\operatorname{int}\mathcal{P} \neq \emptyset$. Moreover, $\mathcal{P}$ is *self polar*, i.e., the *polar cone*

$$\mathcal{P}^+ = \{T = T^t : \operatorname{tr} ST \geq 0,\ \forall S \in \mathcal{P}\} = \mathcal{P}$$

see e.g., [**20**].

We will also need the *singular values* of an $n \times n$ matrix E which we denote by

$$\sigma_1(E) \geq \ldots \geq \sigma_n(E).$$

The corresponding singular value decomposition of E is $U\Sigma V^t = E$, where U and V are orthogonal $n \times n$ matrices and Σ is a diagonal matrix containing the singular values of E.

We should mention that optimization over a partial order like the Loëwner order is an important problem. There are many applications which occur for example in control theory and combinatorial problems. Some applications were presented at the Fourth SIAM Conference on Optimization in Chicago in May 1992; see e.g., [**1, 3**].

3. Optimality Conditions

3.1. First Order Conditions. In this section we present the first order optimality conditions for the relaxed matrix quadratic programming problem

$$\begin{array}{lll} (P) & \min & f(X) = \operatorname{tr} AXBX^t + 2CX^t \\ & \text{s.t.} & g(X) = XX^t - I \preceq 0, \end{array}$$

with $A, B, C \in \Re^{n \times n}$; A and B are symmetric.

This relaxation provides, under certain circumstances, conditions which guarantee that the optimal solution is orthogonal. Thus we will see that under controlled assumptions, we do not weaken our bound by relaxing the constraint $XX^t - I = 0$ to $XX^t - I \preceq 0$. First we show the following.

LEMMA 3.1. *Y is an extreme point of the feasible set*

$$F = \{X : XX^t - I \preceq 0\},$$

if and only if it is an orthogonal matrix.

Proof. Suppose that the $n \times n$ matrix $Y \in F$ is not orthogonal. Let $Y = U\Sigma V^t$ be its singular value decomposition with the diagonal matrix Σ containing the singular value $0 \leq \sigma_k < 1$. Then we get

$$0 \preceq (Y \pm \epsilon U E_k V^t)(Y \pm \epsilon U E_k V^t) \preceq I,$$

for some ϵ, i.e., Y is not an extreme point.

Conversely, suppose Y is orthogonal, but $Y = \theta Y_1 + (1-\theta)Y_2$, for some $0 < \theta < 1$, and $Y_i \in F,\ i = 1,2$. Then the singular value decomposition of Y satisfies $\Sigma = I = \theta Z_1 + (1-\theta)Z_2$, where $Z_i = U^t Y_i V,\ i = 1,2$. Since $Z_i,\ i = 1,2$, are still feasible and so have norm ≤ 1, we conclude that $Z_i = I,\ i = 1,2$, i.e., $Y = Y_i,\ i = 1,2$. Thus Y is an extreme point.

□

Equivalent formulations of Lemma 3.1 can be found in [**6, 27**]. Although the formulations are slightly different, the resulting feasible set and its extreme points are the same.

We define the *Lagrangian* of (P)

$$\mathcal{L}(X,S) = f(X) + \mathrm{tr} S g(X)$$

and the *first order optimality conditions*

$$\begin{array}{lrl} & (i) & AXB + C = -SX \\ (3.1) & (ii) & \mathrm{tr}S(XX^t - I) = 0 \\ & (iii) & XX^t \preceq I \\ & (iv) & S \succeq 0. \end{array}$$

Note that the Lagrange multiplier is a psd symmetric matrix S, since $g(X)$ is symmetric and $\mathcal{P} = \mathcal{P}^+$.

THEOREM 3.1. *Suppose that X is a local minimizer of (P). Then (3.1) holds for some S.*

Proof. Since $I \in \mathrm{int}\mathcal{P}$, the zero matrix satisfies the Slater constraint qualification, i.e., $g(0) \prec 0$. Since X is a local minimizer of (P), the standard Lagrange multiplier theorem, see e.g., [**21**], states that there exists $S \in \mathcal{P}^+$ such that complementary slackness $\mathrm{tr} S(XX^t - I) = 0$ holds and, for all $n \times n$ matrices h

$$\begin{aligned} 0 &= \langle \nabla \mathcal{L}(X,S), h \rangle \\ &= \mathrm{tr} A(XBh^t + hBX^t) + S(Xh^t + hX^t) + 2Ch^t \\ &= 2\mathrm{tr}(AXB + C + SX)h^t, \end{aligned}$$

i.e., $AXB + C + SX = 0$.

□

From Lemma 3.1, we see that the solution X is orthogonal if it is an extreme point of the feasible set. This can be guaranteed by perturbations which make the objective function f concave. However, the above first order conditions provide us with a better means to guarantee orthogonality.

THEOREM 3.2. *Suppose that A and B are nonsingular and the smallest singular value*

$$\sigma_n(A^{-1}CB^{-1}) > 1. \tag{3.2}$$

Then, if X solves (P), it is orthogonal and the associated Lagrange multiplier

$$S \succ 0.$$

Proof. Suppose that X solves (P). To prove that S and X are nonsingular, the first order conditions (3.1)(i) implies that we need only show

$$X + A^{-1}CB^{-1} \text{ is nonsingular,} \tag{3.3}$$

since nonsingularity of S and X implies $SX = AXB + C$ is also nonsingular. Let

$$U(A^{-1}CB^{-1})V^t = \Sigma$$

be the singular value decomposition of $A^{-1}CB^{-1}$. Then (3.3) holds if and only if

$$Y + \Sigma \text{ is nonsingular,}$$

where $Y = UXV^t$. Now (3.1)(iii) implies that

$$YY^t = UXX^tU^t \preceq UIU^t = I, \tag{3.4}$$

since U and V are orthogonal. Now for $||x|| = 1$ we get

$$\begin{aligned} x^t(Y+\Sigma)x &\geq x^tYx + \sigma_n \\ &> x^t\frac{(Y+Y^t)}{2}x + 1 \\ &\geq -\sigma_1(Y) + 1, \text{ (e.g., [\textbf{23}] pg. 240])} \\ &\geq 0, \end{aligned}$$

by (3.4). This yields (3.3) and so $S \succ 0$. Now 3.1(ii) implies that

$$\begin{aligned} 0 &= \mathrm{tr} S(X^tX - I) \\ &= \mathrm{tr} S^{1/2}(X^tX - I)S^{1/2} \end{aligned}$$

and 3.1(iii) implies that

$$S^{1/2}(X^tX - I)S^{1/2} \preceq 0.$$

Thus, $X^tX - I = 0$.

□

In the pure quadratic case, i.e., $C = 0$, we do not have to apply the relaxation $XX^t \preceq I$ but can use $XX^t = I$. This yields the eigenvalue decomposition bounds of Theorem (2.1).

In the case that $A = B = 0$, the problem (P) can be solved explicitly, see e.g., [**16**], pg. 429.

COROLLARY 3.1. *Suppose that $A = B = 0$. Then the optimum X for problem (P) is obtained from the polar decomposition of C,*

$$C = SX, \tag{3.5}$$

where X is orthogonal and $S \succeq 0$. Moreover, the optimal value

$$trCX^t = \sum_{i=1}^{n} \sigma_i(C).$$

Proof. Since $f(X)$ is linear, the optimum X is an extreme point of the feasible set F, i.e., it is orthogonal. Thus (3.5) follows from the first order optimality condition for (P).

□

Note that if $f(X)$ is convex, i.e., $A \otimes B$ is psd, then Theorem 3.1 yields a first order characterization of optimality.

3.2. Second Order Conditions. We now present optimality conditions for (P) using second order information. We also present a conjecture that a characterization of optimality exists that has no gap between necessity and sufficiency independent of convexity of the objective function f. This would extends known results on trust region methods and methods for quadratic objectives with a single quadratic constraint. Note that the relaxation has $(n^2+n)/2$ constraints.

We first present a test for optimality in (P) which compares different solutions of the first order optimality conditions (3.1). This extends the result in [**9**] which deals with a single real valued quadratic constraint.

THEOREM 3.3. *Suppose that $X_i, S_i, i = 1, 2$, are solutions of the first order optimality conditions (3.1) with $S_i \succ 0, i = 1, 2$. Then*

$$4(f(X_2) - f(X_1)) = tr(X_1 - X_2)(S_1 - S_2)(X_1 - X_2)^t.$$

Proof. From the first order conditions (3.1) we get

$$\mathrm{tr} AX_iBX_i^t + CX_i^t = -\ \mathrm{tr} S_iX_iX_i^t,\ i=1,2,$$

and, after subtracting,

$$\text{(3.6)}\ \mathrm{tr} AX_2BX_2^t - \ \mathrm{tr} AX_1BX_1^t + C(X_2-X_1)^t = \ \mathrm{tr}(S_1X_1X_1^t - S_2X_2X_2^t).$$

Also, (3.1) implies that

$$\begin{aligned} \mathrm{tr} AX_1BX_2^t + CX_2^t &= -\ \mathrm{tr} S_1X_1X_2^t \\ \mathrm{tr} AX_2BX_1^t + CX_1^t &= -\ \mathrm{tr} S_2X_2X_1^t = -\ \mathrm{tr} S_2X_1X_2^t \end{aligned}$$

which, after subtracting, yields

$$\mathrm{tr} C(X_2-X_1)^t = -\ \mathrm{tr} S_1X_1X_2^t + \ \mathrm{tr} S_2X_1X_2^t. \tag{3.7}$$

Subtracting (3.6) and (3.7) we get

$$\begin{aligned} 2(f(X_2)-f(X_1)) &= \mathrm{tr} S_1X_1X_1^t - \ \mathrm{tr} S_2X_2X_2^t - \mathrm{tr}(S_1-S_2)X_1X_2^t \\ &= \mathrm{tr}(S_1-S_2)(I-X_1X_2^t) \\ &= \frac{\mathrm{tr}(S_1-S_2)}{2}(I-X_1X_2^t+I-X_2X_1^t), \end{aligned}$$

since $S \succ 0$ implies $X_1X_1^t = X_2X_2^t = I$.

□

We now characterize the feasible directions at a feasible point X.

LEMMA 3.2. *Let $F = \{X : XX^t \preceq I\}$ denote the feasible set of (P). Then F is a convex set. Moreover, let $X \in F$ and denote the set of feasible directions at X by*

$$\mathcal{D}_X = \{V \in \Re^{n\times n} : \exists\bar\theta > 0 \ \textit{with}\ (X+\bar\theta V)(X+\bar\theta V)^t \le I\}.$$

Then:

$$a)\ \emptyset \ne \mathit{int}\, \mathcal{D}_X = \{V : XV^t + VX^t \ \textit{is nd on}\ \mathcal{N}(XX^t - I)\};$$

and

$$b)\ \overline{\mathcal{D}}_X = \{V : XV^t + VX^t \ \textit{is nsd on}\ \mathcal{N}(XX^t - I)\},$$

where $\bar{\cdot}$ denotes closure.

Proof. Note that, for each $S \in \mathcal{P}$, the Hessian of $\mathrm{tr} SXX^t = I \otimes S$ is psd and so the constraint g is $\mathcal{P}$-convex and the feasible set F is a convex set. Moreover, $g(0) \prec 0$ i.e., $X = 0$ is in the interior of F. Therefore $\mathcal{D}_0 = \Re^{n\times n}$. Thus the result holds in the trivial case $X = 0$. Now suppose that $0 \ne XX^t \preceq I$. Then the direction $V = -X$ points into the interior of the feasible set, i.e.,

$$(X+\theta V)(X+\theta V)^t = (1-\theta)^2XX^t \prec I,$$

for $0 < \theta \le 1$. Thus $V = -X$ is in the interior of the convex cone $\mathcal{D}_X$.

Suppose h is a feasible direction. Then,

$$V_\lambda = \lambda h + (1-\lambda)(-X),$$

for $0 \leq \lambda < 1$, is a feasible direction pointing into the interior of the feasible set i.e., $V_\lambda \in \text{int } \mathcal{D}_X$. Therefore,

$$(X + \theta V_\lambda)(X + \theta V_\lambda)^t - I = \theta(XV_\lambda^t + V_\lambda X^t) + \theta^2 V_\lambda V_\lambda^t + XX^t - I$$

is nsd for small, $\theta > 0$. Since $V_\lambda V_\lambda^t$ is psd, this implies that $XV_\lambda^t + V_\lambda X^t$ is nd on $\mathcal{N}(XX^t - I)\}$ and thus $Xh^t + hX^t$ is nsd on $\mathcal{N}(XX^t - I)\}$. Thus we have shown that int $\mathcal{D}_X \neq \phi$ and

$$\mathcal{D}_X \subset \{V : XV^t + VX^t \text{ is nd on } \mathcal{N}(XX^t - I)\}.$$

Conversely, suppose that $Xh^t + hX^t$ is nd on $\mathcal{N}(XX^t - I)\}$. Then $\theta(Xh^t + hX^t) + \theta^2 hh^t$ is nd on $\mathcal{N}(XX^t - I)\}$, for small $\theta > 0$. This implies that

$$(X + \theta h)(X + \theta h)^t = \theta(Xh^t + hX^t) + \theta^2 hh^t + XX^t - I$$

is nsd, for small $\theta > 0$, see e.g., [**2**] or [**22**], i.e., $h \in \mathcal{D}_X$. We can perturb h and still maintain that $Xh^t + hX^t$ is nd on $\mathcal{N}(XX^t - I)\}$. Therefore $h \in \text{int}\mathcal{D}_X$. This proves a). Since $\text{int}\mathcal{D}_X \neq \emptyset$ and we are dealing with convex sets, b) follows from a continuity argument.

□

Corollary 3.2. *If $XX^t \preceq I$, then*

$$\emptyset \neq \{V : XV^t + VX^t \textit{ is nd on } \mathcal{N}(XX^t - I)\}.$$

We now present second order optimality conditions for (P). Note that (3.8) differs from the standard conditions in the literature and allows for sufficiency for a global optimum to hold. In this respect, it is close to the standard trust region results.

Theorem 3.4. *Suppose X is feasible for (P). Define the optimality conditions*

$$\begin{array}{lll} & (i) & S \succeq 0 \\ (3.8) & (ii) & AXB + C + SX = 0 \\ & (iii) & trS(XX^t - I) = 0 \\ & (iv) & trAhBh^t + Shh^t \geq 0, \ \textit{if } Xh^t + hX^t \textit{ is nsd on } \mathcal{N}(XX^t - I)\}. \end{array}$$

Then the following holds:

a) if (3.8) holds for some S, then X is a global minimum for (P) ;

b) if (3.8) holds for some $S \succ 0$, then X is a global minimum of (P) and $XX^t = I$.

Proof. Now suppose (3.8) holds. If Y is feasible for (P) and $f(Y) < f(X)$, then

$$\begin{array}{rcl} \mathcal{L}(Y,S) & = & f(Y) + \operatorname{tr} S(YY^t - I) \\ & < & f(X) \\ & = & f(X) + \operatorname{tr} S(XX^t - I) \\ & = & \mathcal{L}(X,S). \end{array} \tag{3.9}$$

We can assume that $YY^t - I$ is negative definite. (Use a small perturbation of Y into the interior of the feasible set and maintain $f(Y) < f(X)$.) Let $V = X - Y$ and $Y_\theta = X - \theta V$, and so Y_θ is a convex combination of X and Y for $0 \leq \theta \leq 1$. Then $Y_\theta Y_\theta^t - I = -\theta(XV^t + VX^t) + \theta^2 VV^t + XX^t - I$ is negative definite for small $\theta > 0$. Therefore, $XV^t + VX^t$ is positive definite on $\mathcal{N}(XX^t - I)$ which, by hypothesis, implies that

$$\langle V, \nabla^2 \mathcal{L}(X,S)V \rangle = \operatorname{tr} AVBV^t + SVV^t \geq 0.$$

But by (3.9) and the stationarity condition $\nabla \mathcal{L}(X,S) = 0$ in (3.8), we have

$$\mathcal{L}(Y,S) = \mathcal{L}(X,S) + \langle V, \nabla^2 \mathcal{L}(X,S)V \rangle < \mathcal{L}(X,S),$$

a contradiction. This proves (a).

(b) follows from complementary slackness, i.e., $S \succ 0$, $XX^t - I \preceq 0$ and $\operatorname{tr} S(XX^t - I) = 0$ imply $XX^t - I = 0$.

□

COROLLARY 3.3. *Suppose that*

$$\sigma_n(A^{-1}CB^{-1}) > 1. \tag{3.10}$$

Then the second order conditions (3.8) characterize optimality of X for (P), with X orthogonal, i.e., X is orthogonal and solves (P) if and only if the first order conditions (3.8) hold and

$$trAhBh^t + Shh^t \geq 0 \ if \ Xh^t + hX^t \ is \ nsd. \tag{3.11}$$

Proof. Necessity follows directly from the previous theorem and holds without (3.10). By Theorem 3.4, the first order conditions and (3.11) are sufficient for X to be a global optimum for (P). But Theorem 3.2 and (3.10) imply that if X solves (P) then it is orthogonal and the associated Lagrange multiplier $S \succ 0$. This proves sufficiency.

□

We now present a conjecture that the above sufficient conditions are in fact necessary. This would provide a characterization of optimality that parallels the one for the standard trust region subproblem TR. Recall that the standard second order necessary conditions differ from (3.8) in that the Hessian of the

Lagrangian is positive semidefinite on the tangent space, i.e. for all h such that $Xh^t + hX^t = 0$.

CONJECTURE 3.1. *X is a global minimum for (P) only if (3.8) holds for some S.*

4. Parametric Trust Region Bounds

In this section we present a parametric approach for solving (EP). We form bordered matrices by augmenting the matrices A and B using the full rank factorization of the linear term. We obtain a (larger) pure quadratic problem, which enables us to apply the eigenvalue bounds discussed above. Preliminary numerical results are given in Section 4.2.

We consider the matrix quadratic programming problem (EP) throughout this section, i.e.,

$$\begin{array}{lll} (EP) & \min & f(X) = \mathrm{tr} AXBX^t + 2CX^t \\ & \text{s.t.} & XX^t = I, \end{array}$$

where A, B are (real) symmetric $n \times n$ matrices and $C \in \Re^{n \times n}$ with $r := \mathrm{rank} C$. With respect to QAP this is again the orthogonal relaxation of PQAP, that is we assume again that the elimination of the linear equality constraints (the projection) was already done.

4.1. Symmetric Border Perturbations. Let $C = ab^t$ be a full rank factorization of C with $a, b \in \Re^{n \times r}$. With this factorization of C we define the following symmetric border perturbations of A, B

$$\bar{A} := \begin{bmatrix} \alpha & a^t \\ a & A \end{bmatrix} \text{ and } \bar{B} := \begin{bmatrix} \beta & b^t \\ b & B \end{bmatrix},$$

where α, β are symmetric matrices in $\Re^{r \times r}$. We will describe below how we choose these matrices α and β. Furthermore, we partition

$$Z = \begin{pmatrix} w & u^t \\ v & Y \end{pmatrix} \tag{4.12}$$

with $u, v \in \Re^{n \times r}$ and $w \in \Re^{r \times r}$. We now introduce the following $(n + r)$ dimensional eigenvalue problem

$$\begin{array}{lll} (P_{\alpha,\beta}) & \min & q_{\alpha,\beta}(Z) := \mathrm{tr} \bar{A} Z \bar{B} Z^t \\ & \text{s.t.} & ZZ^t = I_{n+r}. \end{array}$$

Since Z depends on the choice of α and β we will denote the optimal solution by $Z(\alpha, \beta)$. The following theorem shows that the optimal solution can be used to bound (EP).

THEOREM 4.1. *Suppose $Z(\alpha,\beta)$ solves $(P_{\alpha,\beta})$ and $\bar{X}$ solves (EP). Then*

$$f(\bar{X}) \geq q_{\alpha,\beta}(Z(\alpha,\beta)) - tr\alpha\beta. \tag{4.13}$$

Proof. Since $\bar{X}$ is the solution of (EP) with $\bar{X}\bar{X}^t = I$,

$$Z = \begin{pmatrix} I_r & 0 \\ 0 & \bar{X} \end{pmatrix}$$

is feasible but not necessarily optimal for $(P_{\alpha,\beta})$ and so

$$q_{\alpha,\beta}(Z) = f(\bar{X}) + \mathrm{tr}\alpha\beta \geq q_{\alpha,\beta}(Z(\alpha,\beta))$$

by optimality of $Z(\alpha,\beta)$. This proves the desired inequality.

□

We are now interested in choosing good values for α and β in order to maximize the lower bound

$$h(\alpha,\beta) := \min\{\mathrm{tr}\bar{A}Z\bar{B}Z^t : ZZ^t = I\} - \mathrm{tr}\alpha\beta.$$

This is a parametric programming problem, i.e., we want to

$$\max_{\alpha,\beta} h(\alpha,\beta) \tag{4.14}$$

or equivalently

$$\max_{\alpha,\beta}\{\left\langle\lambda(\bar{A}),\lambda(\bar{B})\right\rangle_{-} - \mathrm{tr}\alpha\beta\}. \tag{4.15}$$

We can use the techniques from [**29**] to maximize this function, i.e., we are maximizing the sum of two functions on $\Re^{r^2+r}$, where the first one is the minimal scalar product, while the second is a simple quadratic with the Hessian being a matrix of ones except for a zero diagonal. Both functions are in general not concave, and the first function does not have to be differentiable when there are multiple eigenvalues. However, we can still apply subdifferentiable optimization and ignore the lack of concavity. (In [**29**], the first function was the minimal scalar product for the bound for the quadratic part; while the second function was the optimal value of the LSAP, i.e., the bound for the linear part.)

For completeness we now include the differentials of the bound $h(\alpha,\beta)$. As mentioned in Section 2.3 the bound is differentiable if the eigenvalues are simple. We assume simple eigenvalues and suppose that $\lambda(\bar{A})$ and $\lambda(\bar{B})$ are ordered nondecreasingly and nonincreasingly, respectively. Then the differentials of $h(\alpha,\beta)$ with respect to α and β become

$$\begin{aligned} \frac{\partial h(\alpha,\beta)}{\partial \alpha_{ij}} &= \mathrm{diag}(P^t E_{ij} P)^t \lambda(\bar{B}) - \beta_{ji} \\ \frac{\partial h(\alpha,\beta)}{\partial \beta_{ij}} &= \mathrm{diag}\lambda(\bar{A})^t (Q^t E_{ij} Q) - \alpha_{ji}, \end{aligned} \tag{4.16}$$

where P and Q contain the eigenvectors of $\bar{A}$ and $\bar{B}$ in appropriate order, respectively. For multiple eigenvalues subgradients are used instead of gradients.

The numerics showed that there do not occur problems with using subgradient directions in practice.

The above theorem provides a lower bound for (EP). However, we do not necessarily obtain a feasible solution for (EP) from the optimal solution of $(P_{\alpha,\beta})$. But, if the matrices α, β are diagonal, ordered appropriately, and large in absolute value, then the trace in (4.15) will essentially cancel with part of the minimal scalar product.

To accomplish this, we choose

$$\alpha = \operatorname{diag}((rt_1, (r-1)t_1, \cdots, 2t_1, t_1)^t), \tag{4.17}$$

and

$$\beta = \operatorname{diag}((rt_2, (r-1)t_2, \cdots, 2t_2, t_2)^t), \tag{4.18}$$

with $t_1, t_2 \in \Re$. Recall that $r = \operatorname{rank} C$. Then, for

$$t_1 \to -\infty, \ t_2 \to +\infty$$

the solution Z of $(P_{\alpha,\beta})$ takes the form

$$Z(\alpha, \beta) = \begin{pmatrix} I_r & 0 \\ 0 & Y \end{pmatrix},$$

where Y is orthogonal and therefore feasible for (EP) in the limit. We leave it to the interested reader to work out the details.

Let us now consider the objective function of $(P_{\alpha,\beta})$.

$$\begin{aligned}
q_{\alpha,\beta}(Z) &= \operatorname{tr} \bar{A} Z \bar{B} Z^t \\
&= \operatorname{tr} \begin{pmatrix} \alpha & a^t \\ a & A \end{pmatrix} \begin{pmatrix} w & u^t \\ v & Y \end{pmatrix} \begin{pmatrix} \beta & b^t \\ b & B \end{pmatrix} \begin{pmatrix} w^t & v^t \\ u & Y^t \end{pmatrix} \\
&= \operatorname{tr} \alpha w \beta w^t + \alpha w b^t u + a^t v \beta w^t + a^t v b^t u + \\
&\quad \alpha u^t b w^t + \alpha u^t B u + a^t Y b w^t + a^t Y B u + \\
&\quad a w \beta v^t + a w b^t Y^t + A v \beta v^t + A v b^t Y^t + \\
&\quad a u^t b v^t + a u^t B Y^t + A Y b v^t + A Y B Y^t \\
&= \operatorname{tr} A Y B Y^t + 2 a w b^t Y^t + \alpha w \beta w^t + \\
&\quad 2 \alpha w b^t u + 2 a^t v \beta w + 2 a^t v b^t u + \alpha u^t B u + \\
&\quad A v \beta v^t + 2 a^t Y B u + 2 A Y b v^t \\
&= \operatorname{tr} A Y B Y^t + 2 a w b^t Y^t + \alpha w \beta w^t + g_{\alpha,\beta}(w, u, v, Y)
\end{aligned}$$

with

$$\begin{aligned}
g_{\alpha,\beta}(w, u, v, Y) \ :=& \ \operatorname{tr} 2 \alpha w b^t u + 2 a^t v \beta w + 2 a^t v b^t u + \alpha u^t B u + \\
& \ A v \beta v^t + 2 a^t Y B u + 2 A Y b v^t,
\end{aligned}$$

which we call the error term of $(P_{\alpha,\beta})$. We can see that the first two terms in $q_{\alpha,\beta}$ are almost in the form of the objective function of (EP), $f(X) = \operatorname{tr} AXBX^t + 2CX^t$, if $w \to I_r$ and $C = ab^t$. So for α and β given as in (4.17) and (4.18) we can bound (EP) from above and below at the same time. These considerations lead to the following approximation for (EP).

LEMMA 4.1. *Let (EP) and $(P_{\alpha,\beta})$ be given as above with diagonal matrices α and β as in (4.17) and (4.18), respectively. Suppose $Z = \begin{pmatrix} w & u^t \\ v & Y \end{pmatrix}$ solves $(P_{\alpha,\beta})$ and $\bar{X}$ solves (EP), and let $t_1 \to -\infty$ and $t_2 \to \infty$. Then*

$$f(Y) \geq f(\bar{X}) \geq f(Y) + g_{\alpha,\beta}(w,u,v,Y) + \alpha(w - I_r)\beta(w + I_r)^t. \tag{4.19}$$

Proof. The proof for the lower bound is analogous to the proof of Theorem (4.1).

Let $Z = \begin{pmatrix} w & u^t \\ v & Y \end{pmatrix}$ be a solution of Q_S. By the previous observation we know that Y is orthogonal for $t_1 \to -\infty$ and $t_2 \to \infty$. Thus we have a feasible solution for (EP) and $f(Y) \geq f(\bar{X})$ follows.

□

So in order to find a good approximation to (EP) one has to make the gap between lower and upper bound, $g_{\alpha,\beta}(w,u,v,Y) + \alpha(w - I_r)\beta(w + I_r)^t$, small. One way would be a parametric approach equivalent to the one presented above. An alternative would be applying nonsymmetric border perturbations on A and B. We briefly want to present this idea in the remainder of this subsection.

Let the full rank decomposition of C be now given by

$$C = -ab^t.$$

We then perturb A and B nonsymmetrically that is

$$\hat{A}_N = \begin{pmatrix} \alpha & a^t \\ -a & A \end{pmatrix} \text{ and } \hat{B}_N = \begin{pmatrix} \beta & b^t \\ -b & B \end{pmatrix}$$

and define the $(n+r)$ dimensional minimization problem

$$\begin{array}{lll} (P_N) & \min & q_N(Z) = \mathrm{tr}\hat{A}_N Z \hat{B}_N^t Z^t \\ & \text{s.t.} & ZZ^t = I_{n+r} \end{array}$$

with Z defined as in (4.12).

Since $\hat{A}_N$ and $\hat{B}_N$ are not symmetric, (P_N) is not as tractable a problem as $(P_{\alpha,\beta})$. But the advantage of this approach is that the objective function of (P_N) becomes

$$\begin{aligned} q_N(Z) &= \mathrm{tr}\hat{A}_N Z \hat{B}_N^t Z^t \\ &= \mathrm{tr}AYBY^t + 2awb^tY^t + \alpha w\beta w^t + g_N(w,u,v,Y) \end{aligned}$$

with $g_N(w,u,v,Y) = \mathrm{tr}2a^tvb^tu + \alpha u^tBu + \beta v^tAv$.

Size	Sol.	GLB	EVB2	EVB3	EVB4	EVB5
12	578	439	446	498	472	469
15	1150	963	927	1002	973	980
20	2570	2057	2075	2286	2196	2217
30	6124	4539	4982	5443	5266	5312
10	4954	3586	2774	4541	4079	3978
10	8082	6139	6365	7617	7211	7099
10	8649	7030	6869	8233	7837	7708
10	8843	6840	7314	8364	8006	7526
10	9571	7627	8095	8987	8672	8480
10	936	878	885	895	887	886
12	1652	1536	1562	1589	1573	1575
14	2724	2492	2574	2630	2609	2611
16	3720	3358	3518	3594	3560	3572
18	5358	4776	5035	5150	5104	5110
20	6922	6166	6533	6678	6625	6619
42	15812	11311	n.a.	14202	13830	13938
49	23386	16161	n.a.	21230	20715	20897
56	34458	23321	n.a.	31496	30701	30857

TABLE 1. Lower Bounds for QAP

We see that some terms cancelled and so the error term g_N becomes at least smaller than $g_{\alpha,\beta}$ in number of terms. It also seems that we got rid of the dominating terms of $g_{\alpha,\beta}$. This fact was also shown by numerical experiments.

One approach for tackling (P_N), which uses the fact that $\hat{A}$ and $\hat{B}$ are almost symmetric, is the theory of indefinite inner products, e.g., Gohberg et. al. [**12**]. So the application of this theory would be a future research direction.

4.2. Numerical Results. The remainder of this paper discusses preliminary numerical experiments. We first used the bound derived in Theorem 4.1 and applied the parametric programming approach discussed above to achieve a new eigenvalue bound for QAP. This new bound will be denoted by EVB5. We calculated the bound for four groups of instances of QAP. The results are given in Table 1. Note that some of the results in the table are taken from [**14**].

The first group of instances is from to [**26**]. These data are the most widely used for QAP. The instances are pure quadratic, i.e., $C = 0$. The second group of problems comes from [**4**]. The problems all have linear term $C \neq 0$. The third and the fourth group contain again pure quadratic problem instances of QAP. The examples in the third group are taken from [**14**] while the examples of the last group are due to [**32**].

The table is structered as follows. The first and second columns give, respectively, the size and the best known solution of the problem instances. For problems of size ≤ 15 the solutions are optimal. The other columns compare the classical Gilmore–Lawler bound (GLB) [**11, 18**], the eigenvalue bounds EVB2 through EVB4 discussed in the introduction, and EVB5, the new bound. There were no solutions available for EVB2 for problems of the last group.

We should give some technical notes about the computation of EVB5. We calculated the lower bound given in Theorem (4.1) and then maximized $h(\alpha, \beta)$ with respect to α and β.

Specifically, we selected α and β to be diagonal matrices. For the pure quadratic problems (groups 1,2 and 4) we used as starting values $\alpha_0 = \min\{\lambda(A)\}$ and $\beta_0 = \max\{\lambda(B)\}$. For the examples in the second group it proved to be preferable to choose $|\alpha_0|$ and $|\beta_0|$ large. (Since the problems from this group typically have a full rank linear term, these were the hardest for the present approach.) Then we proceeded iteratively. In each iteration we calculated a subgradient and followed the subgradient direction with a fixed stepsize. If the objective function decreased, we reduced the stepsize- factor, and otherwise we increased it. We stopped after a fixed number of iterations. In each computation 100 iterations into the direction of the (sub)gradient were made. We point out that about 10 to 20 iterations are sufficient to calculate bounds that are close to the values of EVB5 that are reported in Table 1.

It proved that subgradient directions were sufficient to improve the bound. By choosing the stepsizes carefully at each iteration we were able to find an improvement in the following iterations.

The comparisons of the different bounds show that EVB5 is in general a competitive bound compared to EVB2 and EVB4. The only exception are problems that have nonzero linear parts. This comes from the fact that EVB4 solves the linear part over the set of permutation matrices while EVB5 only uses orthogonal matrices.

The main point to consider lies in the fact that EVB5 allows further improvements by shifts or reductions. This might make EVB5 also competitive to EVB3 for which shifts and reductions were already selected to maximize the lower bound. So these first numerical results are very promising and encouraging for future work into this direction.

We also did numerical experiments for the approximation of (P), the general matrix quadratic programming problem. The results are shown in Table 2. The first and second column in the table give the size of the example and the rank of the linear term, respectively. The third column contains a lower bound while the fourth column gives an upper bound of the given instance. The last two columns represent the gap between lower and upper bounds, where the second last column contains the absolute gap while the last column shows the relative gap.

The problem instances were generated as follows. The elements of A and B are uniformly distributed real numbers on the interval $[1, 10]$. C was constructed by $C = ab^t$, with a and b being $n \times r$ matrices whose elements where also generated uniformly on the interval $[1, 10]$.

n	r	Lower Bd.	Upper Bd.	Abs. Gap	Rel. Gap
10	1	-5865.50	-5770.21	95.29	1.66 %
10	2	-5099.63	-4770.10	329.53	6.91 %
20	1	-18601.73	-18304.06	297.67	1.63 %
20	2	-20056.09	-19240.64	815.45	4.24 %
20	3	-19143.17	-17886.03	1257.14	7.03 %
20	4	-20801.24	-18986.02	1815.22	9.57 %
30	1	-36190.12	-35704.43	485.69	1.37 %
30	2	-40985.36	-39875.45	1109.91	2.79 %
30	3	-38647.93	-36909.37	1738.56	4.72 %
30	4	-42747.06	-40267.14	2479.20	6.15 %
40	1	-64111.57	-63427.04	684.53	1.08 %
40	2	-60088.59	-58548.47	1540.12	2.64 %
40	3	-64844.81	-62810.69	2034.12	3.24 %
40	4	-62945.42	-59930.75	3014.67	5.04 %

TABLE 2. Lower and upper bounds for (P)

One can observe that the quality of the approximation depends on the influence of the linear term of the given instance. Problems with a linear term of small rank typically yield a much smaller interval containing the correct optimum, than problems with linear term of full rank.

Acknowledgements The first two authors gratefully acknowledge financial support by Christian Doppler Laboratorium für Diskrete Optimierun. Thanks to NSERC of Canada and DIMACS for their support. The third author thanks the Department of Civil Engineering and Operations Research, Princeton University for their support during his research leave, Sept/92-July/93.

REFERENCES

1. F. ALIZADEH, *Combinatorial optimization with semidefinite matrices*, Proceedings of the Second Annual Integer Programming and Combinatorial Optimization Conference (Carnegie-Mellon University), 1992.
2. D.P. BERTSEKAS, *Constrained optimization and lagrange multipliers*, Academic Press, New York, NY, 1982.
3. S. BOYD and L. EL GHAOUI, *Method of centers of minimizing generalized eigenvalues*, Tech. report, Stanford University, 1992, to appear in the special issue of LAA in Systems and Control.
4. P. CARRARESI and F. MALUCELLI, *A new lower bound for the quadratic assignment problem*, Tech. Report TR-7/88, Universita di Pisa, 1988.
5. K. FAN, *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. U.S.A. **35** (1951), 1951.
6. P.A. FILLMORE and J.P. WILLIAMS, *Some convexity theorems for matrices*, Glasgow Mathematical Journal **10** (1971), 110–117.
7. G. FINKE, R.E. BURKARD, and F. RENDL, *Quadratic assignment problems*, Annals of Discrete Mathematics **31** (1987), 61–82.
8. R. FLETCHER, *Semi-definite matrix constraints in optimization*, SIAM J. Control and Optimization **23** (1985), 493–513.

9. W. GANDER, *Least squares with a quadratic constraint*, Numer. Math. **36** (1981), 291–307, This is the abbreviated version of the technical report published in 1978.
10. D.M. GAY, *Computing optimal locally constrained steps*, SIAM J. Sci. Statist. Comput. **2** (1981), 186–197.
11. P.C. GILMORE, *Optimal and suboptimal algorithms for the quadratic assignment problem*, SIAM Journal on Applied Mathematics **10** (1962), 305–313.
12. I. GOHBERG, P. LANCASTER, and L. RODMAN, *Matrices and indefinite scalar products*, Birkhauser, Verlag Basel, 1983.
13. B. GOLLAN, *Eigenvalue perturbations and nonlinear parametric optimization*, Mathematical Programming Studies **30** (1987), 67–81.
14. S.W. HADLEY, F. RENDL, and H. WOLKOWICZ, *A new lower bound via projection for the quadratic assignment problem*, Mathematics of Operations Research **17** (1992), 727–739.
15. M. HEINKENSCHLOS, *On the solution of a two ball trust region subproblem*, Tech. Report Nr.92-16, Universitat Trier, Mathenatik/Informatik, 1992.
16. R. HORN and C. JOHNSON, *Matrix analysis*, Cambridge University Press, New York, 1985.
17. P. LANCASTER and M. TISMENTESKY, *Theory of matrices*, second ed., Academic Press, New York, 1985.
18. E. LAWLER, *The quadratic assignment problem*, Management Science **9** (1963), 586–599.
19. Y. Li, P.M. PARDALOS, and M.G.C. RESENDE, *A greedy randomized adaptive search procedure for the quadratic assignment problem*, Tech. report, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 1993.
20. R. LOEWY and H. SCHNEIDER, *Positive operators on the n-dimensional ice cream cone*, J. Math. Anal. Appl. **49** (1975), 375–392.
21. D.G. LUENBERGER, *Optimization by vector space methods*, John Wiley, 1969.
22. ———, *Linear and nonlinear programming*, , Reading, Addison-Wesley, Massachusetts, 1984.
23. A.W. MARSHALL and I. OLKIN, *Inequalities: Theory of majorization and its applications*, Academic Press, New York, NY, 1979.
24. A.A. MELKMAN and C.A. MICCHELLI, *Optimal estimation of linear operators in hilbert spaces from inaccurate data*, SIAM J. Numer. Anal. **16** (1979), no. 1, 87–105.
25. J.J. MORÉ and D.C. SORENSEN, *Computing a trust region step*, SIAM J. Sci. Statist. Comput. **4** (1983), 553–572.
26. C.E. NUGENT, T.E. VOLLMAN, and J. RUML, *An experimental comparison of techniques for the assignment of facilities to locations*, Operations Research **16** (1968), 150–173.
27. M.L. OVERTON and R.S. WOMERSLEY, *On the sum of the largest eigenvalues of a symmetric matrix*, SIAM J. Matrix Analysis and Applications **13** (1992), 41–45.
28. ———, *Second derivatives for optimizing eigenvalues of symmetric matrices*, Tech. Report 627, Computer Science Department, NYU, 1993.
29. F. RENDL and H. WOLKOWICZ, *Applications of parametric programming and eigenvalue maximization to the quadratic assignment problem*, Mathematical Programming **53** (1992), 63–78.
30. C. ROUCAIROL, *A parallel branch and bound algorithm for the quadratic assignment problem*, Discrete Applied Mathematics **15** (1987), 211–225.
31. S. SAHNI and T. GONZALES, *P-complete approximation problems*, Journal of ACM **23** (1976), 555–565.
32. J. SKORIN-KAPOV, *Tabu search applied to the quadratic assignment problem*, ORSA Journal on Computing **2(1)** (1990), 33–45.
33. R.J. STERN and H. WOLKOWICZ, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, Tech. Report CORR 92-38, University of Waterloo, Waterloo, Canada, 1992.
34. K.A. WILLIAMSON, *A robust trust region algorithm for nonlinear programming*, Ph.D. thesis, Department of Mathematical Sciences, Rice University, Houston, TX, 1990.

35. Y. YUAN, *A dual algorithm for minimizing a quadratic function with two quadratic constraints*, Journal of Computational Mathematics **9** (1991), 348–359.
36. Y. ZHANG, *Computing a Celis-Dennis-Tapia trust-region step for equality constrained optimization*, Mathematical Programming **55** (1992), 109–124.

(S.E. Karisch) TECHNISCHE UNIVERSITAT GRAZ, INSTITUT FUR MATHEMATIK, KOPERNIKUSGASSA 24, A-8010 GRAZ, AUSTRIA
E-mail address, S.E. Karisch: karisch@fstgds09.tu-graz.ac.at

(F. Rendl) TECHNISCHE UNIVERSITAT GRAZ, INSTITUT FUR MATHEMATIK, KOPERNIKUSGASSA 24, A-8010 GRAZ, AUSTRIA
E-mail address, F. Rendl: rendl@ftug.dnet.tu-graz.ac.at

(H. Wolkowicz) UNIVERSITY OF WATERLOO, DEPARTMENT OF COMBINATORICS AND OPTIMIZATION, WATERLOO, ONTARIO, CANADA
E-mail address, H. Wolkowicz: hwolkowicz@orion.uwaterloo.ca