Abstract

We consider three parametric relaxations of the 0-1 quadratic programming problem. These relaxations are to: quadratic maximization over simple box constraints, quadratic maximization over the sphere, and the maximum eigenvalue of a bordered matrix. When minimized over the parameter, each of the relaxations provides an upper bound on the original discrete problem. Moreover, these bounds are efficiently computable. Our main result is that, surprisingly, all three bounds are equal.
Key words: quadratic boolean programming, bounds, quadratic programming, trust region subproblems, minmax eigenvalue problems.

AMS 1991 Subject Classification:
Primary: 90C09, 90C25; Secondary: 90C27, 90C20.

1 INTRODUCTION

Consider the ±1 quadratic programming problem

\[(P) \quad \mu^* := \max_{x \in F} q(x) := x^t Q x + c^t x, \quad x \in F := \{-1,1\}^n, \quad \text{(1.1)}\]

where \(Q\) is an \(n \times n\) symmetric matrix and \(c \in \mathbb{R}^n\). Any problem with \(Q\) nonsymmetric can be reduced to \((P)\). Moreover, 0,1 quadratic programming is equivalent to \((P)\) via the transformation \(x = 2y - c\), where \(y\) is a \((0,1)\)-vector and \(c\) is the vector of ones. These problems have many applications, in particular in combinatorial optimization. However, they are NP-hard, see e.g. [9] pg 196, problem GT25, since \((P)\) is equivalent to the max-cut problem.

Various approaches have been used to solve or approximate ±1 or 0,1 programming problems. One of the possible techniques is to relax problem \((P)\) to a tractable nonlinear continuous problem in order to obtain upper bounds. This approach was used for graph partitioning problems in [7], and the maximum stable set problem in [16]. More recently it has been applied in e.g. [14, 25, 24, 5].

In this paper we study three different relaxations which yield three bounds. We replace \((P)\) by a relaxed problem:

\[(RP) \quad f(u) = \max_{x \in K} q_u(x) = x^t (Q - \text{diag}(u)) x + u^t e + e^t x, \quad \text{where } q_u \text{ is a parametrization of the quadratic function } q \text{ and is equivalent to it on the feasible set } F, e \text{ is the vector of ones, and } K \text{ is a relaxation of the feasible set } F. \text{ We then solve}\]

\[B := \min_{u \in L} f(u)\]

\[\text{The authors would like to thank DIMACS Center at Rutgers University for their support. This report is issued simultaneously at: the Department of Combinatorics and Optimization, University of Waterloo, Report CORR 93-05; the Department of Civil Engineering and Operations Research, Princeton University, Report SOR 93-5; and DIMACS Center, Rutgers University, Report 93-18.}\]
to find the best bound over all values of the parameter \( u \) that yield a tractable problem. In each case \( f(u) \) is a convex function and \( L \) is a convex set. Therefore, finding the bound \( B \) can be done in polynomial time, see [19]. The relaxations are to:

- quadratic maximization over simple box constraints
- quadratic maximization over the sphere
- the maximum eigenvalue of a bordered matrix.

Our contribution is to show that, surprisingly, these three seemingly unrelated relaxations all yield the same bound.

In Section 2 we present the first bound \( B_1 \), which is based on diagonal shifting \( Q \) to obtain a tractable concave quadratic programming problem with simple box constraints, i.e. the constraints are relaxed to \(-1 \leq z \leq 1\).

In Section 3 we present bound \( B_2 \) which again involves a diagonal shift but the \( \pm 1 \) constraints are relaxed to a sphere constraint. These relaxations are called trust region subproblems. We show that \( B_1 = B_2 \). In Section 4 we present bound \( B_3 \) which consists in minimizing the maximum eigenvalue of an enlarged shifted matrix of dimension \( n + 1 \). We show that \( B_1 = B_2 = B_3 \).

We include a final relaxation of a quadratic programming problem with no linear term and show that the bound obtained, \( B_4 \), again equals the previous three bounds.

### 1.1 Preliminaries

We will use the following notations: \( \text{diag}(v) \) denotes the diagonal matrix formed from the vector \( v \) and conversely, \( \text{diag}(M) \) is the vector of the diagonal elements of the matrix \( M \); \( e \) is the vector of ones; the matrices \( M_1 \preceq M_2 \) \((M_1 < M_2)\) refers to the Loewner partial order, i.e. that \( M_1 - M_2 \) is negative semidefinite (negative definite); similarly, \( v \leq w \), \((v < w)\) refers to coordinatewise ordering of the vectors; \( \text{conv}(S) \) is the convex hull of the set \( S \); \( \lambda_{\max}(M) \) denotes the largest eigenvalue of a symmetric matrix \( M \). The space of symmetric matrices is considered with the trace inner product \( \langle M, N \rangle = \text{trace} MN \).

For a convex function \( f : \mathbb{R}^n \to \mathbb{R} \), the vector \( \phi \in \mathbb{R}^n \) is a subgradient at the vector \( v \) if \( \phi^t(y-v) \leq f(y) - f(v), \forall y \in \mathbb{R}^n \). The set of all subgradients is a convex, compact, set called the subdifferential, and is denoted by \( \partial f(v) \). The directional derivative of \( f \) at \( v \) in the direction \( z \) is
The relationship between directional derivative and subdifferential is given by

\[ f'(v; z) = \max_{\phi \in \partial f(v)} \phi^t z. \]  \hspace{1cm} (1.2)

This means that \( f'(v; \cdot) \) is a positively homogeneous, sublinear functional, and it is the support function of \( \partial f(v) \). (For more details see e.g. [26].)

2 BOUND 1 - Convex Quadratic Programming

Consider the shifted function

\[ q_v(x) := x^t(Q - \text{diag}(v))x + v^t e + c^t x, \]  \hspace{1cm} (2.1)

and the relaxed problem

\[ (RP^1_v) \quad f_1(v) := \max_{-1 \leq x \leq 1} q_v(x). \]  \hspace{1cm} (2.2)

Then a bound for \((P)\) is

\[ B_1 := \min_{Q - \text{diag}(v) \leq 0} f_1(v). \]  \hspace{1cm} (2.3)

We now present some properties for the bound \( B_1 \).

**Lemma 2.1**

1. \( B_1 \) is an upper bound for \((P)\), i.e.

\[ \mu^* \leq B_1. \]

2. The function \( f_1 \) is convex and finite valued, with subdifferential

\[ \partial f_1(v) = \text{conv} \{ z = (1-x_i^2) \in \mathbb{R}^n : x = (x_i) \in \mathbb{R}^n, -1 \leq x \leq 1, f_1(v) = q_v(x) \}. \]  \hspace{1cm} (2.4)

3. The bound \( B_1 \) is attained for some \( v \in \mathbb{R}^n \) such that \( \lambda_{\text{max}}(Q - \text{diag}(v)) = 0 \).

4. \( B_1 = \inf_{Q - \text{diag}(v) < 0} f_1(v). \)
Proof: 1. For each \( v, \) \( x^t \text{diag}(v)x = v^t e \) on the feasible set \( F. \) Therefore, \( q_v(x) = q_0(x) = q(x) \) on \( F. \) This implies \( f_1(v) \) is an upper bound for \( (P) \), for each \( v. \)

2. For each fixed \( x, \) the function \( q_v(x) \) is linear in \( v. \) Therefore \( f_1(v) \) is the maximum of a set of linear functions. This implies that \( f_1 \) is convex in \( v. \) Moreover, compactness and continuity imply that \( f_1 \) is finite valued, which further implies that \( f_1 \) is continuous and subdifferentiable. In [4] (See also pg 26 in [8] or pg 188 in [6].) it is shown that at any point \( v \) and any direction \( z, \) the directional derivative of \( f_1 \) exists and is a support function given by

\[
f'_1(v; z) = \max_{\{ -1 \leq x \leq 1, f_1(v) = q_v(x) \}} z^t \nabla_v q_v(x).
\]

The relationship between directional derivative and subdifferential, see (1.2), yields the desired subdifferential formula, i.e. \( f'_1(v; z) \) is the support function of the convex hull of gradients \( \nabla q_v(x), \) at optimal points \( x, \) thus defining the set \( \partial f_1(v) \) in (2.4).

3. Note that \( q_v(x) = q_0(x) + \sum_i v_i(1 - x_i^2) \) so that

\[
v \geq w \text{ implies } q_v(x) \geq q_w(x), \quad \forall -1 \leq x \leq 1.
\]

Therefore

\[
v \geq w \text{ implies } f_1(v) \geq f_1(w). \tag{2.5}
\]

So, if \( B_1 = f_1(v) \), then we can decrease \( v \) until we lose negative definiteness. Similarly, if \( B_1 = \inf_j f(v^{(j)}) \), then we can assume that the sequence \( v^{(j)} \) is bounded above by e.g. \( e^t |Q| e + e^t |e| \), where \( | | \) replaces all the elements by their absolute values, and bounded below by \( \text{diag}(Q) \). This sequence must have a convergent subsequence and we can apply the previous argument to a cluster point to again get \( Q - \text{diag}(v) \) singular.

4. This follows from continuity and attainment.

For every \( v, \) the function \( q_v(x) = x^t (Q - \text{diag}(v))x + e^t e + e^t x \) yields the same values on the feasible set \( F. \) Therefore, we can replace \( q_0(x) \) in \( (P) \) with a concave function by restricting \( Q - \text{diag}(v) \leq 0. \) We then have a tractable problem to solve. In fact, these problems can be solved in polynomial time, see e.g. [15]. Efficient numerical algorithms for these problems are described in e.g. [2].

The Lagrangian for (2.3) is

\[
L_1(v, \Lambda) := f_1(v) + \text{trace} \Lambda (Q - \text{diag}(v)),
\]
where $\Lambda$ is a symmetric, positive semidefinite Lagrange multiplier matrix. The Slater constraint qualification holds for (2.3), i.e., there exists $v$ such that $Q - \text{diag}(v) < 0$. Therefore, $B_1 = f_1(v)$ if and only if the following optimality conditions hold for some $\Lambda$, see e.g. [17]:

$$
0 \in \partial f_1(v) - \text{diag}(\Lambda) \quad \text{(stationarity)}
$$

$$
\text{trace}\Lambda(Q - \text{diag}(v)) = 0 \quad \text{(complementary slackness)}
$$

$$
\Lambda = \Lambda^t \geq 0 \quad \text{(multiplier sign)}.
$$

Bound $B_1$ was first considered in Körner [14] for constrained problems. In a weaker form, which corresponds to setting $v = \lambda_{\text{max}}(Q)e$, where $\lambda_{\text{max}}$ denotes the largest eigenvalue, it was proposed by Hammer and Rubin [11].

3 BOUND 2 - Optimization Over Sphere

Now, for $u^e = 0$, we again consider the shifted function

$$
q_u(y) := y^t(Q - \text{diag}(u))y + u^te + d^ty,
$$

and the second relaxed problem

$$
(RP_u^2) \quad f_2(u) := \max_{\|y\|^2 = n} q_u(y).
$$

Now a bound for (P) is

$$
B_2 := \min_{u^e = 0} f_2(u) = \min_u f_2(u).
$$

We can restrict $u^e = 0$, since we can always replace $u$ by $u - \sum_i u_i e$ without changing the values of $q_u$ in $(RP_u^2)$. Note that $(RP_u^2)$ is not linearly constrained. These quadratically constrained problems are called trust region subproblems and are also tractable and can be solved in polynomial time, see [31]. This can also be seen from the fact that there is a dual problem to $(RP_u^2)$ that minimizes a convex function over an interval, see [29]. These trust region subproblems can be classified into: the easy case if the Hessian of the Lagrangian (see below) is positive definite; and the hard case otherwise. (For the theory and efficient algorithms, see e.g. [10, 28, 18].)

We now present some properties for the bound $B_2$.

**Lemma 3.1**

1. $B_2$ is an upper bound for (P), i.e.

$$
\mu^* \leq B_2.
$$
2. The function $f_2$ is convex and finite valued, with subdifferential
\[ \partial f_2(u) = \text{conv} \{ z = (1 - y_i^2) \in \mathbb{R}^n : y = (y_i) \in \mathbb{R}^n, \| y \|^2 = n, f_2(u) = q_u(y) \}. \]

3. The bound $B_3$ is attained for some $u \in \mathbb{R}^n$. Moreover, if $B_2 > \mu^*$, then the hard case holds for $(RP_u^2)$.

Proof: The statements follow similarly to the results in Lemma 2.1. Attainment for the hard case follows from the fact that the optimum for $(RP_u^2)$ cannot be unique if $f_2(u) = B_2$. (See the proof of Theorem 3.1 below for details.)

\[ B_1 = B_2 = \mu^*. \]

Proposition 3.1 If $Q$ is diagonal, then

Proof: If $Q = \text{diag}(q)$ for some $q \in \mathbb{R}^n$, then we can set the optimum of $(P)$ as $y_i = \pm 1$ coinciding with the sign of the corresponding component of the linear term $c_i$. We can now find the unique solution of the system of equations
\[ q_i - u_i - \lambda = -|c_i|, \quad \forall i, \]
with $\sum_i u_i = 0$, i.e., $\lambda = \frac{1}{n} (\sum_i |c_i| + \sum_i q_i)$ and $u_i = |c_i| + q_i - \lambda$, $i = 1, \ldots, n - 1$. These equations show that the exact solution of $(P)$ also solves $(RP_u^2)$ and $(RP_u^{1+\lambda e})$.

\[ f_2(u) = q_u(y_u) \]
\[ = \max_y q_u(y) + \lambda (n - \|y\|^2) \]
\[ = \max_y q_{(u + \lambda e)}(y) \]
\[ \geq \max_{-1 \leq y \leq 1} q_{(u + \lambda e)}(y) \]
\[ = f_1(u + \lambda e) \]
\[ \geq B_1, \]

Theorem 3.1 The bound $B_1 = B_2$.

Proof: Fix $u$ with $u^e = 0$ and set $y_u$ so that $\| y_u \|^2 = n$ and $f_2(u) = q_u(y_u)$. Then the optimality conditions for $(RP_u^2)$, see e.g., [10, 28], imply that there exists $\lambda$ such that
\[ Q - \text{diag}(u) - \lambda I \leq 0, \quad 2(Q - \text{diag}(u) - \lambda I)y_u = -e, \quad (3.4) \]
i.e. the Hessian of the Lagrangian is negative semidefinite and $y_u$ is a stationary point of the Lagrangian. Therefore, $y_u$ is a global unconstrained maximum for the Lagrangian. We get
\[ f_2(u) = q_u(y_u) \]
\[ = \max_y q_u(y) + \lambda (n - \|y\|^2) \]
\[ = \max_y q_{(u + \lambda e)}(y) \]
\[ \geq \max_{-1 \leq y \leq 1} q_{(u + \lambda e)}(y) \]
\[ = f_1(u + \lambda e) \]
\[ \geq B_1, \]

7
where we have used the facts: $\sum_i u_i = 0$ implies $\lambda n = \sum_i (u_i + \lambda)$; and $Q - \text{diag}(u) - \lambda I \leq 0$ to bound $B_1$. This shows that $B_2 \geq B_1$.

Now suppose that $B_2 = f_2(u)$. Then by definition of $B_2$ and convexity of $f_2$, we conclude that $0 \in \partial f_2(u) - \alpha e$, for some Lagrange multiplier $\alpha \in \mathbb{R}$. Therefore, there exists $z = (z_i) \in \partial f_1(u)$, with $z_i = \alpha$, $\forall i$. The subdifferential of $f_2$ consists of the convex hull of vectors with components $(1 - y_i^2)$, for optimal solutions $y = (y_i)$ of $(RP^2_u)$, see Lemma 3.1. By Caratheodory’s Theorem, see e.g. [26], for elements in the convex hull we need only consider $k \leq n + 1$ optimal solutions $y^{(j)}$. Therefore,

$$y := \sum_{j=1}^k \theta_j y^{(j)}, \quad \theta_j > 0, \quad \sum_{j=1}^k \theta_j = 1,$$

is the corresponding convex combination of optimal solutions for the components of the subdifferential $z_i = \alpha = \sum_{j=1}^k \theta_j (1 - (y_i^{(j)})^2)$. Now let $Y$ be the $n \times k$ matrix with components $(y_i^{(j)})$, $\Phi$ be the $n \times k$ matrix with squared components $((y_i^{(j)})^2)$ and let $\theta = (\theta_j) \in \mathbb{R}^k$. Then $\theta^T \Phi e = n \theta^T e = n$ and this equals $e^T \Phi \theta = (1 - \alpha)e^T e = (1 - \alpha)n$, i.e. $\alpha = 0$. Now let $\theta^* := (\sqrt{\theta_j}) \in \mathbb{R}^k$ and for fixed $i$, $z_i^* := (\sqrt{\theta_j} y_i^{(j)}) \in \mathbb{R}^k$. Then, the positivity of $\theta$ and the Cauchy-Schwartz inequality implies that $|y_i| \leq \sum_{j=1}^k \theta_j |y_i^{(j)}| = (\theta^*)^T z_i^* \leq \sqrt{\sum_{j=1}^k \theta_j (y_i^{(j)})^2} = 1$, since $\alpha = 0$. Therefore, the components of $y$ must satisfy $|y_i| \leq 1$. Moreover, the conditions for equality in Cauchy-Schwartz now yield

$$|y_i| \leq 1 \text{ with equality iff } |y_i^{(j)}| = 1 \forall j.$$  

(3.6)

In fact, $y_i = 1$ implies that $y_i^{(j)} = 1 \forall j$. (Similarly for $y_i = -1$.) Now each vector $y^{(j)}$ satisfies (3.4) and so the convex combination $y = Y \theta$ satisfies (3.4) as well. Therefore, $y$ solves the maximization problem in (3.5) with the inequality, $\max_{-1 \leq y_i\leq 1} q_i(y) + \lambda n$, and it also solves the one in the line above the inequality, $\max_y q_i(u + \lambda e) = \max_y q_i(y) + \lambda n$. Therefore, equality holds in (3.5) except perhaps at the end we might have $\theta > B_1$. So we have shown that

$$B_2 = f_2(u) \implies f_2(u) = f_1(u + \lambda e).$$  

(3.7)

Now let $v := u + \lambda e$ and $Q_\epsilon := Q - \text{diag}(v)$. From the optimality of the columns of $Y$, we know that

$$Q_\epsilon(Y - ye^T) = 0.$$  

(3.8)

Define the vectors $p^{(j)} := y^{(j)} - y$, i.e. the columns of the matrix $P := Y - ye^T$. Now define the symmetric positive semidefinite Lagrange multiplier matrix
\[
\Lambda = \sum \theta_j p^{(j)}(p^{(j)})^t. \quad \text{Since } Q_u p^{(j)} = 0, \quad \forall j, \text{ by (3.8), we see that } Q_u \Lambda = 0, \text{ i.e. complementary slackness holds in (2.6) for (2.3). Moreover, the stationarity conditions in (2.6) hold as well, i.e. the diagonal elements of } \Lambda \text{ satisfy }
\]

\[
\Lambda_{ii} = \sum \theta_j (y_i^{(j)} - y_i)^2 = \sum \theta_j (y_i^{(j)})^2 - y_i^2 = 1 - y_i^2,
\]

by the facts that \( y = Y \theta \) and \( \Phi \theta = c \). Therefore, we have satisfied the optimality conditions for finding \( B_1 \), i.e. we now have \( B_1 = f_1(v) = f_2(u) = B_2 \).

\[\square\]

**Corollary 3.1** \( B_2 = f_2(u) \), with Lagrange multiplier \( \lambda \) for \( (RP^2_u) \) if and only if \( B_2 = f_2(u) = f_1(u + \lambda e) = B_1 \).

**Proof.** The results follow directly from the proof of the Theorem. In particular, see (3.7).

\[\square\]

**Example 3.1** We now illustrate that equality holds in Theorem 3.1 and see how the various subgradients are calculated. Let \( Q = \begin{bmatrix} -2 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \) and \( c = 4 \left( \frac{\sqrt{2}}{\sqrt{3}} \right) \) in \( (P) \). Set \( Y = \begin{bmatrix} \sqrt{\frac{2}{3}} & \sqrt{\frac{1}{3}} \\ 0 & \sqrt{\frac{1}{3}} \end{bmatrix} \) with squared components of \( Y \) in \( \Phi = \begin{bmatrix} 2 & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} \end{bmatrix} \). Then the columns of \( Y \) satisfy the stationarity condition for optimality of \( (RP^2_u) \) with \( u = 0 \) and Lagrange multiplier \( \lambda = 0 \), i.e. \( 2QY = -ce \). Moreover, \( Q \) is negative semidefinite and \( \Phi e = 2e \) which implies that the columns of \( Y \) are optimal solutions of \( (RP^2_u) \). Now let \( \theta = (\frac{1}{2} \frac{1}{2})^t \). Then \( \Phi \theta = e \) which means that we have found a subgradient in \( \delta f_2(u) \) with equal components, i.e. \( u = 0 \) yields the minimum for \( f_2 \). (See the proof of Theorem 3.1.)

Now set \( y = Y \theta = \left( \frac{2\sqrt{2}}{\sqrt{3}} \right) \). Then the components satisfy \( |y_i| < 1 \).

Moreover, \( 2Qy = -c \). Therefore, \( y \) is an optimal solution for \( (RP^2_u) \) and, by Corollary 3.1, \( f_1(0) = f_2(0) \). Now choose the Lagrange multiplier matrix for \( B_1 \) in (2.3) to be \( \Lambda = \frac{1}{9} \begin{bmatrix} 1 & -\sqrt{3} & -\sqrt{3} \\ -\sqrt{3} & 3 & -\sqrt{3} \\ -\sqrt{3} & -\sqrt{3} & 3 \end{bmatrix} \). Then we see that complementary
slackness holds, i.e. \( \Lambda \geq 0 \), trace(\( \Lambda(Q - \text{diag}(u)) \)) = 0. Moreover, the diagonal of \( \Lambda \) equals the subgradient with components \( (1 - y_i^2) \), i.e. the bound \( B_1 \) is attained at \( u = 0 \).

If the problem \( (P) \) is without the linear term, i.e. \( c = 0 \), bound \( B_2 \) becomes much simpler. We have

\[
B_2 = \min_{u' e = 0} n\lambda_{\text{max}}(Q - \text{diag}(u)). \tag{3.9}
\]

For \( Q = L \), the Laplacian matrix of a graph, (3.9) provides an upper bound on the max-cut problem, see [5]. We will show in the next section that every problem \( (P) \) can be reduced to the form without a linear term by increasing the size of the matrix, and moreover, the transformation does not change the relaxed bounds.

4 BOUND 3 - Minimize \( \lambda_{\text{max}} \)

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

\[
q^c_{00} = 0 \\
q^c_{0i} = q^c_{i0} = \frac{1}{2}c_i \quad \text{for } i > 0 \\
q^c_{ij} = q_{ij} \quad \text{for } i, j > 0,
\]

i.e.

\[
Q^c := \begin{bmatrix}
0 & \frac{1}{2}c_e \\
\frac{1}{2}c_e & Q
\end{bmatrix}.
\]

In order to have analogous functions \( q^c_u(y) \) and \( f_3(u) \) as in the previous cases, let us introduce

\[
q^c_u(y) := y^t(Q^c - \text{diag}(u))y + u^t e, \tag{4.1}
\]

and the equivalent relaxed problem

\[
(RP^3_u) \quad f_3(u) := \max_{\|y\|_2 = n+1} q^c_u(y) = (n + 1)\lambda_{\text{max}}(Q^c - \text{diag}(u)) + u^t e. \tag{4.2}
\]

Just as for \( f_2 \), we can restrict \( u^t e = 0 \). Now a bound for \( (P) \) is

\[
B_3 := \min_{u' e = 0} f_3(u) = \min_u f_3(u). \tag{4.3}
\]

We now present some properties for the bound \( B_3 \).
Lemma 4.1  
1. $B_3$ is an upper bound for $(P)$, i.e.
\[ \mu^* \leq B_3. \]

2. The function $f_3$ is convex and finite valued, with subdifferential
\[ \partial f_3(u) = \text{conv}\{z = (1-y_t^2) \in \mathbb{R}^{n+1} : y \in \text{eigenspace of } \lambda_{\max}(Q^c - \text{diag}(u)), \|y\|^2 = n+1\}. \]

3. The bound $B_3$ is attained for some $u \in \mathbb{R}^{n+1}$. Moreover, if $B_3 > \mu^*$, then the eigenspace corresponding to $\lambda_{\max}(Q^c - \text{diag}(u))$ has dimension $\geq 1$ for $(RP_u^2)$.

4. $B_3 = \min_{u \in \mathcal{U}} (n+1) \lambda_{\max}(Q^c - \text{diag}(u))$.

Proof: The proof follows similarly as that for Lemma 3.1, as this relaxation is a trust region subproblem with no linear term.

\[ \square \]

Theorem 4.1 The bounds $B_1 = B_2 = B_3$.

Proof. If we restrict the first component of $y$ in (4.2) to be 1, then we get a value for $f_2$, i.e. this shows that $B_3 \geq B_2$.

To show the reverse inequality, suppose that $u$ is given with $u^t e = 0$ and $x$ solves $(RP_u^2)$, i.e. $f_2(u) = q_u(x)$. Therefore, the stationarity conditions and negative semidefiniteness conditions hold with Lagrange multiplier $\lambda$, see (3.4). Let $Q_u := Q - \text{diag}(u)$, $t := \lambda - \frac{1}{2} c^t x$, $D(t) := \begin{bmatrix} t & \frac{1}{2} c^t \\ \frac{1}{2} c & Q_u \end{bmatrix}$, and $y := \frac{1}{n+1}(1 x)^t$. Then $\lambda$ is an eigenvalue of $D(t)$ with eigenvector $y$, since $(Q_u - \lambda I)x = -\frac{1}{2} c$. Moreover, the optimality conditions $Q_u - \lambda I \leq 0$ implies $\lambda \geq \lambda_{\max}(Q_u)$. Therefore, by the interlacing theorem for eigenvalues, e.g. [12], $\lambda = \lambda_{\max}(D(t))$. This implies that
\[ \lambda - \frac{t}{n+1} = \frac{n \lambda - \frac{1}{2} c^t x}{n+1} \]
is the largest eigenvalue of the shifted matrix
\[ D(t) - \frac{t}{n+1} I = Q^c - \text{diag}(u^c), \]
thereby defining the $n+1$ dimensional vector $u^c := \left( \frac{n}{n+1} t \right)$. Therefore
\[ f_3(u^c) = n \lambda - \frac{1}{2} c^t x. \]
To complete the proof of the theorem, we need only show that this also equals $f_2(u)$. By the stationarity condition for $(RP_n^1)$, we can substitute $(Q_u - \lambda I)x = -\frac{1}{2}c$ and see that the objective value $f_2(u) = q_u(x) - \lambda(x^t x - n) = \lambda n - \frac{1}{2}c^t x$.

□

Similar relations between trust region subproblems and eigenvalue problems are presented in [30, 23]. The problem (4.3) is equivalent to minimizing the maximum eigenvalue of a matrix. These type of problems are treated in e.g. [20, 21], where efficient algorithms are presented as well as optimality conditions. The above theorem shows that these problems can also be treated using efficient trust region subproblem algorithms.

We can now combine the above equivalences between the three given bounds with a fourth bound to get:

**Corollary 4.1** Suppose that $f_4(v) := \max_{-1 \leq z \leq 1} q^z(x)$ and

$$B_4 := \min_{Q^e = \text{diag}(v) \leq 0} f_4(v).$$

Then

$$B_1 = B_2 = B_3 = B_4.$$

**Proof:** Problem $B_4$ corresponds to $B_3$ just as $B_1$ corresponds to $B_2$. Therefore the result follows from Theorems 3.1 and 4.1.

□

5 CONCLUSION

In this paper we considered three different relaxations of the $\pm 1$ quadratic programming problem and showed that all three, surprisingly, yield the same bound. Thus, our results provide a theoretical framework to alternatively use and combine different computational methods to get the bounds $B_1 = B_2 = B_3$. At present, it seems that the eigenvalue bound $B_3$ might be the one most efficiently computable. However, there is a lot of ongoing research to improve algorithms for all three problems used in our relaxations.

Bound $B_1$ corresponds to applying parametric programming to a quadratic programming problem with simple or box constraints. Efficient algorithms for this problem are given in e.g. [3]. These correspond to trust region type algorithms over the box or infinity norm rather than 2-norm. Bound $B_3$ corresponds to applying parametric programming to trust region subproblems

12
with the 2-norm. Efficient algorithms are given in [10, 18], where the sub-problems are solved to near optimality in typically 1-2 iterations. There is ongoing recent research to develop more efficient algorithm in particular for large problems. Finally $B_3$ corresponds to minimizing the maximum eigenvalue. Theory and efficient algorithms for this problem are surveyed in [20]. In fact, the theory shows that these algorithms have quadratic convergence properties, which is surprising for possibly nondifferentiable problems. An interior point algorithm to compute $B_3$ was given in [13].

There are several interesting questions that our equivalences raise, e.g. which is the most efficient way to solve the various problems. In particular, we see that we can solve min-max eigenvalue problems by applying known trust region subproblem algorithms or even quadratic programming combined with some subdifferential calculus like a bundle trust subgradient approach, see e.g. [27]. Another question is to study the performance of the relaxations in the presence of additional constraints. Problem (P) is an unconstrained $\pm 1$ quadratic programming problem. However, many combinatorial optimization problems naturally lead to constrained $\pm 1$ quadratic programming problems. From this point of view, the quadratic programming bound $B_1$ seems to be the most tractable, since it immediately allows adding additional linear constraints. However, one may add certain constraints to the other bounds. For example, in [22] an eigenvalue relaxation with additional polyhedral constraints is considered for the graph bisection problem.

Since the bounds $B_1 = B_2 = B_3$ are introduced in order to approximate the original discrete problem (P), it is important to study the connection between the original combinatorial problem and its relaxation. This has been done in e.g. [16] for the stable set problem, in Boppana [1] for the graph bisection problem, and [5] for the max-cut problem. The quality of the approximation may vary with different combinatorial optimization problems. However, in general, it seems that the nonlinear relaxations provide better bounds more often than the linear ones.

References


