SIAM J. Optim.
Vol. 21, No. 3, pp. 706–729

ON THE REDUCTION OF DUALITY GAP IN BOX CONSTRAINED NONCONVEX QUADRATIC PROGRAM*

YONG XIA†, XIAOLING SUN‡, DUAN LI§, and XIAOJIN ZHENG¶

Abstract. In this paper, we investigate in this paper the reduction of the duality gap between box constrained nonconvex quadratic programming and its semidefinite programming (SDP) relaxation (or Lagrangian dual). Characterizing the zero duality gap by a set of saddle-point-type conditions, we propose a parameterized distance measure $\delta(\theta)$ between a polyhedral set C and a perturbed nonconvex set $\Lambda(\theta)$ to measure the dissatisfaction degree of the optimality conditions for zero duality gap. An underestimation of the duality gap is then derived which leads to a reduction of the duality gap proportional to $\delta^2(\theta^*)$ for the identified best parameter θ^*. This reduction of duality gap can be extended to the cases with both box and linear equality constraints. We demonstrate that the computation of $\delta(\theta^*)$ can be reduced to the cell enumeration of hyperplane arrangement in discrete geometry. In particular, we show that the reduction of duality gap can be achieved in polynomial time for a fixed degeneracy degree of the modified coefficient matrix determined from SDP relaxation.

Key words. box constrained quadratic program, semidefinite relaxation, reduction of duality gap, cell enumeration of hyperplane arrangement, linear equality constraints

AMS subject classifications. 90C10, 90C22, 90C46

DOI. 10.1137/100802153

1. Introduction.

In this paper, we consider the following nonconvex quadratic program with linear equality and box constraints:

$$
\begin{aligned}
(P_e) \quad & \min\ f(x) := x^T Qx + 2c^T x \\
& \text{s.t. } Ax = b, \\
& \quad\ x \in [-1, 1]^n,
\end{aligned}
$$

where Q is an $n \times n$ indefinite or negative semidefinite symmetric matrix, $c \in \mathbb{R}^n$, A is an $m \times n$ full rank matrix, and $b \in \mathbb{R}^n$. When the set of linear constraints is absent, problem (P_e) reduces to the following box constrained nonconvex quadratic program:

$$
(P) \quad \min\{f(x) = x^T Qx + 2c^T x | x \in [-1, 1]^n\}.
$$

*Received by the editors July 14, 2010; accepted for publication (in revised form) May 2, 2011; published electronically August 4, 2011. This research was supported by Research Grants Council of Hong Kong under grant 414207; by the National Natural Science Foundation of China under grants 10971034, 70832002, and 11001006; by the Joint NSFC/RGC grants under grant 71061160506; by the fund of State Key Laboratory of Software Development Environment under grant SKLSDE-2011ZX-15; and by fundamental research funds for the central universities under grant YWF-11-03-Q-051.

http://www.siam.org/journals/siopt/21-3/80215.html

†State Key Laboratory on Software Development Environment; LMIB of the Ministry of Education, School of Mathematics and System Sciences, Beihang University, Beijing, 100191, People's Republic of China (dearyxia@gmail.com).

‡Department of Management Science, School of Management, Fudan University, Shanghai 200433, People's Republic of China (xls@fudan.edu.cn).

§Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong (dli@se.cuhk.edu.hk).

¶School of Economics and Management, Tongji University, Shanghai 200092, People's Republic of China (xjzheng@fudan.edu.cn).

As revealed in [19], (P) is equivalent to the minimization of $q(x) = x^T Qx + 2c^T x$ on $\{-1, 1\}^n$ if all the diagonal elements of Q are zero. Thus, problem formulation (P) includes the binary quadratic program as a special case and is in general NP-hard.

Global optimization methods for (P) are primarily based on various convex relaxations and branching rules. An and Tao [2] proposed a branch-and-bound method for (P) using a difference of two convex functions (D.C.) optimization approach and ellipsoidal relaxation for the box constraints. Hansen et al. [13] presented a finite branch-and-bound method for (P) where branching is based on the sign of the first-order derivative. Vandenbussche and Nemhauser [24], [25] developed a linear programming based branch-and-cut scheme for (P) which uses valid inequalities derived from the convex hull of the first-order KKT conditions. Yajima and Fujie [26] also proposed cutting plane methods based on valid inequalities. More recently, Burer and Vandenbussche [8] proposed a finite branch-and-bound method for (P) in which SDP relaxations of the first-order KKT conditions of (P) are used in finite KKT-branching. A survey of global optimization methods for (P) can be found in [3]. Solution methods for quadratic programs with linear constraints were summarized in [10], [18].

Semidefinite programming (SDP) relaxations for combinatorial optimization problems and nonconvex quadratic programs have recently attracted much attention due to their tractability in computation (see, e.g., [17], [23]). Goemans and Williamson [11] showed that for a max-cut problem, SDP relaxation yields a 0.878 approximate solution. Nesterov [16] and Ye [27] established $2/\pi$ approximation results of SDP relaxations for quadratic programs with box constraints and simple quadratic constraints. Malik et al. [15] investigated the gap between the max-cut problem and its SDP relaxation and showed that the gap can be reduced by computing a reduced-rank binary quadratic problem. Recently, Ben–Ameur and Neto [6] derived spectral bounds for the max-cut problem which are tighter than the SDP bound of a max-cut problem.

In this paper, we focus on estimating the gap between a box constrained quadratic program and its SDP relaxation. We first derive an estimation of the duality gap between (P) and its SDP relaxation, which leads to a reduction of the duality gap and an improved bound for (P). The results are then extended to problem (P_e).

Rewriting $x \in [-1, 1]^n$ as $x_i^2 - 1 \le 0$ for $i = 1, \dots, n$ gives rise to the following Lagrangian relaxation problem of (P):

$$(L_\lambda) \quad d(\lambda) = \inf_{x \in \mathbb{R}^n} \{L(x, \lambda) := x^T[Q + \mathrm{diag}(\lambda)]x + 2c^T x - e^T \lambda\},$$

where $\lambda \ge 0$, $e = (1, \dots, 1)^T$, and $\mathrm{diag}(\lambda)$ denote the diagonal matrix with λ_i being its ith diagonal element. The dual problem of (P) can be expressed as

$$(D) \quad \max_{\lambda \ge 0} d(\lambda).$$

Let $v(\cdot)$ denote the optimal value of problem $(\cdot)$. Then $v(P) - v(D)$ measures the duality gap between (P) and (D). Using Shor's relaxation scheme [20], the dual problem (D) can be reduced to the following SDP:

$$\begin{aligned} (D_s) \quad & \max -\tau - e^T \lambda \\ & \text{s.t.} \begin{pmatrix} Q + \mathrm{diag}(\lambda) & c \\ c^T & \tau \end{pmatrix} \succeq 0, \\ & \tau \in \mathbb{R}, \quad \lambda \ge 0. \end{aligned}$$

(D) is equivalent to (D_s) in the sense that $v(D_s) = v(D)$ and λ^* is optimal to (D) if and only if (λ^*, τ^*) is optimal to (D_s) with $\tau^* = -v(D) - e^T\lambda^*$.

The idea explored in this paper for estimating the duality gap is based on measuring the degree of dissatisfaction of the optimality conditions by a certain distance between a polyhedral set and a perturbed nonconvex set derived from the saddle-point conditions. More precisely, we define a parameterized distance measure, $\delta(\theta)$, between the set $C = \{x \in \mathbb{R}^n | (Q + \mathrm{diag}(\lambda^*))x + c = 0\}$, with λ^* being the optimal dual solution to (D) and a perturbed set $\Lambda(\theta)$ that corresponds to the complementarity conditions in the saddle-point conditions, where $\theta \in [0, 1]$ is a perturbation parameter. We derive a lower bound of the duality gap given by $\xi_{r+1}\delta^2(\theta^*)$ for the identified best $\theta^* \in [0, 1]$, where $r = \mathrm{rank}(Q + \mathrm{diag}(\lambda^*))$ and ξ_{r+1} is the smallest positive eigenvalue of $Q + \mathrm{diag}(\lambda^*)$, which leads further to an improved lower bound $\nu^* = v(D_s)+ \xi_{r+1}\delta^2(\theta^*)$ for (P). In addition, we demonstrate that the computation of $\delta(\theta^*)$ can be reduced to cell enumeration of hyperplane arrangement in discrete geometry. As the complexity of cell enumeration is $O(n^{r+1})$ (see [4], [21]), where r is the dimension of the null space of $Q + \mathrm{diag}(\lambda^*)$, we deduce that $\delta(\theta^*)$ can be computed in polynomial time for fixed r. In particular, we discuss the special case of $r = 1$ and show that $\delta(\theta^*)$ can be computed efficiently in such a case.

The distance measure of the dissatisfaction of the optimality conditions can be extended to the linear equality constrained problem (P_e). An estimation of the duality gap is established for (P_e) using a similar approach as for (P). We show that a lower bound of the duality gap between (P_e) and its SDP relaxation is given by $1/(1+\kappa)\xi_{r+1}\delta_e^2(\theta^*)$, where $\delta_e^2(\theta^*)$ is a certain distance measure, and $\kappa > 0$ is the maximum eigenvalue of certain positive semidefinite matrix derived from A. The gap reduction technique proposed in this paper can be viewed as a generalization of the results in [22], [29] where binary quadratic programs are considered.

The paper is organized as follows. We discuss the basic properties of the SDP relaxations and optimality conditions for (P) in section 2. We establish the main result of the estimation of the duality gap for (P) in section 3. In section 4, we investigate the computational issues of the estimation and present a procedure for computing the distance $\delta(\theta^*)$. We then discuss the partition of set C and its relationship to the cell enumerations of hyperplane arrangement in section 5. In section 6, we extend the main results to problem (P_e). Finally, we give some concluding remarks in section 7.

Throughout the paper, we denote by $\mathbb{R}^n_+$ the nonnegative orthant of $\mathbb{R}^n$ and $\mathcal{S}^n$ the set of all $n \times n$ symmetric matrices. Notation $A \succcurlyeq B$ implies that matrix $A - B$ is positive semidefinite. We adopt the standard inner product in symmetric matrix space, i.e., $A \bullet B = \mathrm{Tr}(AB) = \sum_{i,j=1}^n a_{ij}b_{ij}$. We also denote by I_q the identity matrix of order q.

2. SDP relaxations and optimality conditions. In this section, we first introduce some basic properties of the Lagrangian dual problem (D) and its SDP reformulations. We then establish the optimality conditions for the zero duality gap between (P) and (D). Finally, we discuss the strict complementarity for (D_s) and its conic dual problem.

2.1. SDP relaxations. The problem (D_s) in section 1 can be viewed as the dual form of SDP relaxation for (P). The primal form of SDP relaxation for (P) can be obtained by lifting $x \in \mathbb{R}^n$ to $Y = xx^T \in \mathcal{S}^n$. Note that $Y = xx^T$ for some $x \in [-1, 1]^n$ implies $Y_{ii} = x_i^2 \le 1$, $i = 1, \ldots, n$. Relaxing $Y = xx^T$ to $Y \succcurlyeq xx^T$ yields the following SDP relaxation problem:

$$(P_s) \quad \min \begin{pmatrix} Q & c \\ c^T & 0 \end{pmatrix} \bullet \begin{pmatrix} Y & x \\ x^T & 1 \end{pmatrix}$$
$$\text{s.t. } Y_{ii} \le 1, \quad i = 1, \ldots, n,$$
$$\begin{pmatrix} Y & x \\ x^T & 1 \end{pmatrix} \succcurlyeq 0,$$

which is the conic dual of (D_s). It is easy to verify the strict feasibility (Slater condition) for both (P_s) and (D_s). Therefore, both SDP problems (P_s) and (D_s) can be solved efficiently by interior-point methods (see, e.g., [17], [23]).

By the conic duality theorem (see, e.g., [7], [17]), the strict feasibility of (P_s) and (D_s) implies the solvability of both (P_s) and (D_s), $v(P_s) = v(D_s)$ and the satisfaction of the complementary condition $X \bullet H(\lambda, \tau) = 0$ for any optimal solutions (Y, x) to (P_s) and (λ, τ) to (D_s), where

$$X = \begin{pmatrix} Y & x \\ x^T & 1 \end{pmatrix}, \qquad H(\lambda, \tau) = \begin{pmatrix} Q + \text{diag}(\lambda) & c \\ c^T & \tau \end{pmatrix}. \tag{2.1}$$

Since $X \succcurlyeq 0$ and $H(\lambda, \tau) \succcurlyeq 0$, $X \bullet H(\lambda, \tau) = 0$ is equivalent to $XH(\lambda, \tau) = H(\lambda, \tau)X = 0$. To see this, let $A, B \in \mathcal{S}^n$ be such that $A \succcurlyeq 0$, $B \succcurlyeq 0$, and $A \bullet B = 0$. Then $0 = A \bullet B = \text{Tr}(AB) = \text{Tr}(A^{1/2}A^{1/2}B^{1/2}B^{1/2}) = \text{Tr}(A^{1/2}B^{1/2}(A^{1/2}B^{1/2})^T)$. It follows that $A^{1/2}B^{1/2} = 0$, and hence $AB = BA = 0$.

The KKT optimality conditions for (P_s) and (D_s) can be described as follows.

LEMMA 2.1. *Let (Y, x) and (λ, τ) be feasible solutions to (P_s) and (D_s), respectively. Then they are optimal if and only if*

$$XH(\lambda, \tau) = 0, \tag{2.2}$$

$$\lambda_i(Y_{ii} - 1) = 0, \qquad i = 1, \ldots, n, \tag{2.3}$$

where X and $H(\lambda, \tau)$ are defined in (2.1).

Note that condition $XH(\lambda, \tau) = 0$ implies X and $H(\lambda, \tau)$ are commutative. Therefore, condition (2.2) is equivalent to

$$[Q + \text{diag}(\lambda)]Y + cx^T = 0, \tag{2.4}$$

$$[Q + \text{diag}(\lambda)]x + c = 0, \tag{2.5}$$

$$c^T Y + \tau x^T = 0, \tag{2.6}$$

$$c^T x + \tau = 0. \tag{2.7}$$

PROPOSITION 2.1. *The optimal solution to (D_s) is unique.*

Proof. Let $(\bar{\lambda}, \bar{\tau})$ and $(\tilde{\lambda}, \tilde{\tau})$ be two optimal solutions to (D_s). By (2.2) in Lemma 2.1, there is an optimal solution (Y, x) to (P_s) such that

$$H(\bar{\lambda}, \bar{\tau})X = 0, \qquad H(\tilde{\lambda}, \tilde{\tau})X = 0.$$

Thus,

$$[H(\bar{\lambda}, \bar{\tau}) - H(\tilde{\lambda}, \tilde{\tau})]X = 0,$$

which gives rise to

$$\begin{pmatrix} \mathrm{diag}(\bar{\lambda}-\tilde{\lambda}) & 0 \\ 0 & \bar{\tau}-\tilde{\tau} \end{pmatrix} \begin{pmatrix} Y & x \\ x^T & 1 \end{pmatrix} = 0.$$

Therefore,

$$(\bar{\lambda}_i - \tilde{\lambda}_i) Y_{ii} = 0, \quad i = 1, \ldots, n, \quad \bar{\tau} - \tilde{\tau} = 0. \tag{2.8}$$

On the other hand, by (2.3) in Lemma 2.1, we have

$$\bar{\lambda}_i (Y_{ii} - 1) = 0, \quad \tilde{\lambda}_i (Y_{ii} - 1) = 0, \quad i = 1, \ldots, n. \tag{2.9}$$

For any $i = 1, \ldots, n$, if $Y_{ii} \neq 0$, then by (2.8), $\bar{\lambda}_i = \tilde{\lambda}_i$. Otherwise, if $Y_{ii} = 0$, then by (2.9), $\bar{\lambda}_i = \tilde{\lambda}_i = 0$. Therefore, $\bar{\lambda} = \tilde{\lambda}$ and $\bar{\tau} = \tilde{\tau}$. ∎

2.2. Strong duality. Next, we discuss the optimality conditions for the strong duality between (P) and (D).

LEMMA 2.2 (see [5]). *For any $\lambda \in \mathbb{R}^n_+$, $d(\lambda) > -\infty$ with x solving (L_λ) if and only if*

(i) $Q + \mathrm{diag}(\lambda) \succcurlyeq 0$*;*

(ii) $[Q + \mathrm{diag}(\lambda)]x + c = 0$*.*

The following condition of a saddle point type characterizes the zero duality gap between (P) and (D).

LEMMA 2.3. *Let $x^* \in [-1, 1]^n$ and $\lambda^* \in \mathbb{R}^n_+$. Then x^* solves (P), λ^* solves (D), and $v(P) = v(D)$ if and only if*

$$Q + \mathrm{diag}(\lambda^*) \succcurlyeq 0, \tag{2.10}$$

$$[Q + \mathrm{diag}(\lambda^*)]x^* + c = 0, \tag{2.11}$$

$$\lambda_i^*[(x_i^*)^2 - 1] = 0, \qquad i = 1, \ldots, n. \tag{2.12}$$

Proof. Suppose that (2.10)–(2.12) hold for some $x^* \in [-1, 1]^n$ and $\lambda^* \in \mathbb{R}^n_+$. From conditions (i) and (ii) in Lemma 2.2, we know that x^* solves (L_{λ^*}). It then follows from (2.12) that

$$d(\lambda^*) = \min_{x \in \mathbb{R}^n} L(x, \lambda^*) = L(x^*, \lambda^*) = f(x^*) \geq v(P).$$

By the weak duality, λ^* solves (D) and $v(D) = v(P)$.

Conversely, if $x^* \in [-1, 1]^n$ and $\lambda^* \in \mathbb{R}^n_+$ solve (P) and (D), respectively, and $v(P) = v(D)$, then

$$L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^n \lambda_i^*[(x_i^*)^2 - 1] \leq f(x^*) = d(\lambda^*) = \min_{x \in \mathbb{R}^n} L(x, \lambda^*).$$

Thus, x^* solves (L_{λ^*}) and, by Lemma 2.2, (2.10) and (2.11) hold. Furthermore, $L(x^*, \lambda^*) = f(x^*)$ implies (2.12). ∎

Let λ^* be the optimal solution to (D). Define

$$\begin{aligned} Q^* &= Q + \mathrm{diag}(\lambda^*), \\ J^* &= \{i \in \{1, \ldots, n\} | \lambda_i^* > 0\}, \\ I^* &= \{1, \ldots, n\} \setminus J^*. \end{aligned}$$

We also define

$$C = \{x \in \mathbb{R}^n | Q^* x + c = 0\}, \tag{2.13}$$

$$\Lambda^* = \{x \in \mathbb{R}^n | x_i \in [-1, 1] \text{ for } i \in I^*; x_i \in \{-1, 1\} \text{ for } i \in J^*\}. \tag{2.14}$$

PROPOSITION 2.2. *Let λ^* be the optimal solution to (D). Then*

(i) $C \cap [-1, 1]^n \neq \varnothing$;

(ii) $v(P) = v(D)$ *if and only if* $C \cap \Lambda^* \neq \varnothing$. *Furthermore, any* $x^* \in C \cap \Lambda^*$ *is an optimal solution to (P).*

Proof.

(i) Let (Y, x) be an optimal solution to (P_s). Since $Y_{ii} \leq 1$ and $Y - xx^T \succcurlyeq 0$, we have $1 \geq Y_{ii} \geq x_i^2$. On the other hand, (2.5) implies that $x \in C$. Thus, $x \in C \cap [-1, 1]^n$.

(ii) Since λ^* solves (D), by Lemma 2.2, we have $Q + \mathrm{diag}(\lambda^*) \succcurlyeq 0$. Also, $x^* \in C \cap \Lambda^*$ is equivalent to conditions (2.11) and (2.12). Part (ii) then follows from Lemma 2.3. □

PROPOSITION 2.3. *The duality gap $v(P) - v(D) = 0$ if and only if there exists an optimal solution (Y, x) to (P_s) satisfying $Y = xx^T$.*

Proof. The "if" part is obvious since (P_s) is resulted from relaxing (P) by replacing $Y = xx^T$ with $Y \succcurlyeq xx^T$. Suppose that $v(P) - v(D) = 0$. Let $x \in [-1, 1]^n$ and λ be the optimal solutions to (P) and (D), respectively. By Lemma 2.3, the saddle point conditions (2.10)–(2.12) hold. Let $Y = xx^T$ and $\tau = -c^T x$. Since $Q + \mathrm{diag}(\lambda) \succcurlyeq 0$, it follows from (2.11) that

$$\begin{pmatrix} I_n & 0 \\ x^T & 1 \end{pmatrix} \begin{pmatrix} Q + \mathrm{diag}(\lambda) & c \\ c^T & \tau \end{pmatrix} \begin{pmatrix} I_n & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} Q + \mathrm{diag}(\lambda) & 0 \\ 0 & 0 \end{pmatrix} \succcurlyeq 0.$$

Thus, (λ, τ) satisfies the semidefiniteness condition in (D_s). Also, $Y_{ii} = x_i^2 \leq 1$ for $i = 1, \ldots, n$ and (2.12) implies that $\lambda_i(Y_{ii} - 1) = 0$. Thus, (Y, x) and (λ, τ) are, respectively, feasible to (P_s) and (D_s). Moreover, it is easy to verify that the complementarity conditions (2.4)–(2.7) hold. Thus, by Lemma 2.1, (Y, x) is an optimal solution to (P_s) satisfying $Y = xx^T$. □

COROLLARY 2.1. *Assume that the optimal solution λ^* to (D) satisfies $Q^* = Q + \mathrm{diag}(\lambda^*) \succ 0$. Then $x = -(Q^*)^{-1} c$ is the unique optimal solution to (P) and $v(P) = v(D)$.*

Proof. Let (Y, x) be an optimal solution to (P_s). By Lemma 2.1, the complementarity conditions (2.4)–(2.7) hold. Since $Q^* \succ 0$, (2.5) has a unique solution $x = -(Q^*)^{-1} c$. Substituting $x = -(Q^*)^{-1} c$ into (2.4) yields $Y = (Q^*)^{-1} c c^T (Q^*)^{-1} = xx^T$. The conclusion then follows from Proposition 2.3. □

Example 2.1. Consider an instance of (P) with

$$Q = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 1 & -3 \\ 2 & -3 & 0 \end{pmatrix}, \qquad c = \begin{pmatrix} -5 \\ -5 \\ -3 \end{pmatrix}.$$

For this example, the optimal solution to (D_s) is $\lambda^* = (0, 19/3, 16/3)^T$ with $v(D) = v(D_s) = -64/3$. As

$$Q^* = Q + \mathrm{diag}(\lambda^*) = \begin{pmatrix} 3 & 2 & 2 \\ 2 & \frac{22}{3} & -3 \\ 2 & -3 & \frac{16}{3} \end{pmatrix} \succ 0,$$

$x^* = -(Q^*)^{-1}c = (1/3, 1, 1)^T$ is the optimal solution to this example. It can be verified that $f(x^*) = v(D) = -64/3$.

2.3. Strict complementarity. We now turn to discuss the strict complementarity relation between (P_s) and (D_s) which will be used later in analyzing the complexity of computing the parameterized distance measure in section 5.

Let (Y^*, x^*) and (λ^*, τ^*) be optimal solutions to (P_s) and (D_s), respectively. The complementarity condition (2.2) in Lemma 2.1 implies that $X^* = [Y^*, x; (x^*)^T, 1]$ and $H(\lambda^*, \tau^*)$ commute. So by the simultaneous diagonalization theorem (see, e.g., Horn and Johnson [14, Theorem 4.5.15]), they share a common eigenvector system, i.e., there exists Γ with $\Gamma^T\Gamma = I_{n+1}$ such that

$$X^* = \Gamma \mathrm{diag}(\alpha_1, \ldots, \alpha_{n+1})\Gamma^T,$$
$$H(\lambda^*, \tau^*) = \Gamma \mathrm{diag}(\beta_1, \ldots, \beta_{n+1})\Gamma^T,$$

where $\alpha_i \geq 0$ and $\beta_i \geq 0$ for $i = 1, \ldots, n+1$. The complementary condition (2.2) then implies

$$\alpha_i \beta_i = 0, \qquad i = 1, \ldots, n+1.$$

Therefore,

$$\mathrm{rank}(X^*) + \mathrm{rank}(H(\lambda^*, \tau^*)) \leq n + 1. \tag{2.15}$$

The following strict complementarity was introduced in Alizadeh, Haeberly, and Overton [1].

Definition 2.1. *A strict complementarity holds for (P_s) and (D_s) if the equality holds in* (2.15) *for any optimal solutions (Y^*, x^*) to (P_s) and (λ^*, τ^*) to (D_s), i.e.,*

$$\mathrm{rank}(X^*) + \mathrm{rank}(H(\lambda^*, \tau^*)) = n + 1. \tag{2.16}$$

It was shown in Alizadeh, Haeberly, and Overton [1] that the strict complementarity holds *generically* for general SDP problems in the sense that the set of SDP problems for which the strict complementarity fails to hold has measure 0 in the space of problem parameters.

Next, we discuss the relation between the rank of Q^* and the rank of X^*. Let (λ^*, τ^*) be the optimal solution to (D_s). Let $\mathrm{rank}(Q^*) = n - r$ with $0 < r < n$. Let $0 = \xi_1 = \cdots = \xi_r < \xi_{r+1} \leq \cdots \leq \xi_n$ be the eigenvalues of Q^*. Then there exists an orthogonal matrix $U = (U_1, \ldots, U_n)$ such that

$$U^T Q^* U = \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}, \tag{2.17}$$

where $W = \mathrm{diag}(\xi_{r+1}, \ldots, \xi_n)$. It is easy to see that the null space of Q^* is spanned by $U_1, \ldots, U_r$.

LEMMA 2.4. *Let λ^* be the optimal solution to (D) and let U be defined in* (2.17). *Then*
(i) $c^T U_i = 0$, *for* $i = 1, \dots, r$;
(ii) $v(D) = -e^T\lambda^* - \sum_{i=r+1}^n (c^T U_i)^2/\xi_i$.

Proof.
(i) Since $d(\lambda^*) > -\infty$, by Lemma 2.2, there exists x such that $Q^* x = -c$. By (2.17), for $i = 1, \dots, r$, we have

$$\begin{aligned} -c^T U_i &= x^T Q^* U_i \\ &= x^T U \cdot \mathrm{diag}(0, \dots, 0, \xi_{r+1}, \dots, \xi_n) U^T U_i \\ &= x^T U \cdot \mathrm{diag}(0, \dots, 0, \xi_{r+1}, \dots, \xi_n) e_i \\ &= 0. \end{aligned}$$

(ii) By the definition of the dual problem, we have

$$\begin{aligned} v(D) = d(\lambda^*) &= \min_{x\in\mathbb{R}^n} [x^T Q^* x + 2c^T x - e^T \lambda^*] \\ &= \min_{y\in\mathbb{R}^n} \left[\sum_{i=r+1}^n \xi_i \left(y_i + \frac{c^T U_i}{\xi_i} \right)^2 - e^T\lambda^* - \sum_{i=r+1}^n \frac{(c^T U_i)^2}{\xi_i} \right] \\ &= -e^T\lambda^* - \sum_{i=r+1}^n \frac{(c^T U_i)^2}{\xi_i}, \end{aligned}$$

where the relation $x = Uy$ is used in the above derivation. ∎

The following result establishes the relation between $r = n - \mathrm{rank}(Q^*)$ and the rank of X^*.

PROPOSITION 2.4. *Let (Y^*, x^*) and (λ^*, τ^*) be optimal solutions to (P_s) and (D_s), respectively. Under the strict complementarity condition for (P_s) and (D_s), it holds that*

$$r = \mathrm{rank}(X^*) - 1 = \mathrm{rank}(Y^* - x^*(x^*)^T). \tag{2.18}$$

In particular, if $c = 0$, then $r = \mathrm{rank}(Y^*)$.

Proof. Let $\omega = (c^T U_{r+1}, \dots, c^T U_n)^T$. It follows from (2.17) and Lemma 2.4 (i) that

$$\begin{aligned} \begin{pmatrix} U^T & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Q^* & c \\ c^T & \tau^* \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} U^T Q^* U & U^T c \\ c^T U & \tau^* \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & W & \omega \\ 0 & \omega^T & \tau^* \end{pmatrix}. \end{aligned} \tag{2.19}$$

Also, by the form of (D_s) and Lemma 2.4 (ii), we have

$$\tau^* = \sum_{i=r+1}^n \frac{(c^T U_i)^2}{\xi_i} = \sum_{i=r+1}^n \frac{(\omega_{i-r})^2}{\xi_i}. \tag{2.20}$$

We then deduce from (2.19) and (2.20) that

$$\mathrm{rank}(H(\lambda^*, \tau^*)) = \mathrm{rank}(W) = \mathrm{rank}(Q^*) = n - r.$$

This, together with the strict complementarity condition (2.16), gives rise to

$$r = n - \mathrm{rank}(H(\lambda^*, \tau^*)) = n - (n + 1 - \mathrm{rank}(X^*)) = \mathrm{rank}(X^*) - 1.$$

The equality $\mathrm{rank}(X^*) - 1 = \mathrm{rank}(Y^* - x^* x^{*T})$ holds because

$$\begin{pmatrix} Y^* - x^*(x^*)^T & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & -x^* \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Y^* & x^* \\ (x^*)^T & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ -(x^*)^T & 1 \end{pmatrix}.$$

This proves (2.18).

Now, if $c = 0$, then the complementarity condition (2.2) reduces to $Y^* Q^* = 0$. Again, by the simultaneous diagonalization theorem, we have

$$\mathrm{rank}(Y^*) + \mathrm{rank}(Q^*) \leq n. \tag{2.21}$$

By the strict complementarity condition (2.16) and the fact that $\mathrm{rank}(H(\lambda^*, \tau^*)) = \mathrm{rank}(Q^*)$ and $\mathrm{rank}(X^*) = \mathrm{rank}(Y^* - x^*(x^*)^T) + 1$, we obtain from (2.21) that

$$\mathrm{rank}(Y^* - x^*(x^*)^T) \geq \mathrm{rank}(Y^*).$$

On the other hand, we always have $\mathrm{rank}(Y^* - x^*(x^*)^T) \leq \mathrm{rank}(Y^*)$. To see this, let $\sigma_i(A)$ denote the ith largest eigenvalue of a symmetric matrix A. By Weyl's inequality (see [14, Theorem 4.3.1]), we have

$$\sigma_i(Y^* - x^*(x^*)^T) \leq \sigma_i(Y^* - x^*(x^*)^T) + \sigma_n(x^*(x^*)^T) \leq \sigma_i(Y^*). \tag{2.22}$$

Since both $Y^* - x^*(x^*)^T \succcurlyeq 0$ and $Y^* \succcurlyeq 0$, we infer from (2.22) that $\mathrm{rank}(Y^* - x^*(x^*)^T) \leq \mathrm{rank}(Y^*)$. Thus,

$$\mathrm{rank}(Y^* - x^*(x^*)^T) = \mathrm{rank}(Y^*),$$

and $r = \mathrm{rank}(Y^*)$ by (2.18). □

Remark 2.1. Let λ^* be the optimal solution to (D). We see from (2.21) that in the homogeneous case of (P) with $c = 0$, $r = n - \mathrm{rank}(Q^*) = 1$ implies $\mathrm{rank}(Y^*) = 1$ or 0. In both cases, we have $Y^* = x^*(x^*)^T$ for some $x^* \in \mathbb{R}^n$ and hence the strong duality $v(P) = v(D)$ holds by Proposition 2.3. However, this is not true in general cases when $c \neq 0$. In fact, by (2.18), $r = 1$ implies $\mathrm{rank}(Y^* - x^*(x^*)^T) = 1$ which does not imply the strong duality. The following small-scale instance of (P) illustrates such a case. Consider

$$\min\{8x_1x_2 + 3x_1 + 3x_2 \mid -1 \leq x_i \leq 1,\ i = 1, 2\}.$$

The global optimal solution of this example is $x^* = (1, -1)^T$ with $v(P) = -8$. It can be verified that the optimal solution to (D) is $\lambda^* = (4, 4)^T$ with $v(D) = -8 - 9/4 < v(P) = -8$. A nonzero duality gap exists. Note that in this example

$$Q^* = Q + \mathrm{diag}(\lambda^*) = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}.$$

Thus, $r = n - \mathrm{rank}(Q^*) = 2 - 1 = 1$.

3. Estimation of duality gap. In this section, we discuss how to estimate the duality gap between (P) and (D) when $Q^* = Q + \mathrm{diag}(\lambda^*) \not\succcurlyeq 0$. Using a distance measure between C and a parametric nonconvex set corresponding to the complementarity condition (2.12), we obtain an underestimation of the duality gap and thus an improved lower bound of (P).

For any $\theta \in [0,1]$, we define the following perturbed set of Λ^*:

$$\Lambda(\theta) = \{x \in \mathbb{R}^n \mid -1 \le x_i \le 1 \text{ for } i \in I^*;\ 1-\theta \le x_i^2 \le 1 \text{ for } i \in J^*\}, \tag{3.1}$$

and define the distance between C and $\Lambda(\theta)$ as

$$\delta(\theta) = \operatorname{dist}(C, \Lambda(\theta)) = \min\{\|x-y\| \mid x \in C,\ y \in \Lambda(\theta)\}. \tag{3.2}$$

The following basic properties of $\delta(\theta)$ can be easily derived:

(i) $\Lambda(0) = \Lambda^*$ and $\Lambda(1) = [-1,1]^n$, where Λ^* is defined in (2.14). Thus, we conclude from item (i) of Proposition 2.2 that $\delta(1) = 0$.

(ii) $C \cap \Lambda(0) \neq \varnothing \Leftrightarrow \delta(0) = 0$. Thus, by Proposition 2.2, $v(P) = v(D)$ if and only if $\delta(0) = 0$.

(iii) If $J^* = \varnothing$, then $\Lambda^* = [-1,1]^n$, which implies $\delta(0) = 0$ by item (i) of Proposition 2.2.

We need the following lemma.

LEMMA 3.1. *Let λ^* be the optimal solution to (D) and U be defined in* (2.17). *Then for any $x \in \mathbb{R}^n$, it holds that*

$$q(x) + \sum_{i=1}^n \lambda_i^*(x_i^2 - 1) = v(D) + \sum_{i=r+1}^n \xi_i \left(y_i + \frac{c^T U_i}{\xi_i} \right)^2, \tag{3.3}$$

where $y = U^T x$.

Proof. For any $x \in \mathbb{R}^n$, let $y = U^T x$. It follows from (2.17) and Lemma 2.4 (ii) that

$$\begin{aligned} q(x) + \sum_{i=1}^n \lambda_i^*(x_i^2 - 1) &= x^T Q x + 2c^T x + \sum_{i=1}^n \lambda_i^*(x_i^2 - 1) \\ &= x^T[Q + \operatorname{diag}(\lambda^*)]x + 2c^T x - e^T \lambda^* \\ &= y^T(U^T Q^* U)y + 2b^T U y - e^T \lambda^* \\ &= \sum_{i=r+1}^n \xi_i y_i^2 + 2c^T U y - e^T \lambda^* \\ &= \sum_{i=r+1}^n \xi_i \left(y_i + \frac{c^T U_i}{\xi_i} \right)^2 - e^T \lambda^* - \sum_{i=r+1}^n \frac{(c^T U_i)^2}{\xi_i} \\ &= v(D) + \sum_{i=r+1}^n \xi_i \left(y_i + \frac{c^T U_i}{\xi_i} \right)^2, \end{aligned}$$

which is (3.3). ∎

We are now in position to give the main result of this section.

THEOREM 3.1. *Let λ^* be the optimal solution to (D). Assume that $J^* \neq \varnothing$ and $\delta(0) > 0$. Then*

$$v(P) - v(D) \ge \max_{\theta \in [0,1]} \phi(\theta), \tag{3.4}$$

where

$$\phi(\theta) = \min\{\theta \min_{i\in J^*} \lambda_i^*,\ \xi_{r+1}\delta^2(\theta)\}. \tag{3.5}$$

Moreover, there is a unique $\theta^* \in (0,1)$ *such that*

$$\max_{\theta\in[0,1]} \phi(\theta) = \xi_{r+1}\delta^2(\theta^*) = \theta^* \min_{i\in J^*} \lambda_i^*. \tag{3.6}$$

Proof. By the definition of C (cf. (2.13)), for any w satisfying $Uw \in C$, we have $Q^*Uw + c = 0$. It follows from (2.17) that

$$U \cdot \mathrm{diag}(0, \ldots, 0, \xi_{r+1}, \ldots, \xi_n) U^T U w = -c,$$

which in turn gives rise to

$$\mathrm{diag}(0, \ldots, 0, \xi_{r+1}, \ldots, \xi_n) w = -U^T c.$$

Thus, $Uw \in C$ if and only if

$$w_i \in \mathbb{R}, \quad i = 1, \ldots, r, \quad w_i = -\frac{c^T U_i}{\xi_i}, \quad i = r+1, \ldots, n. \tag{3.7}$$

Let $\tilde{w} = (w_1, \ldots, w_r)^T$. Using (3.7) and the orthogonality of U, we have the following for any $\theta \in [0,1]$:

$$\begin{aligned}
\delta^2(\theta) &= \min\{\|z - x\|^2 | z \in C,\ x \in \Lambda(\theta)\} \\
&= \min\{\|Uw - Uy\|^2 | Uw \in C,\ Uy \in \Lambda(\theta)\} \\
&= \min\left\{\sum_{i=1}^{r} (w_i - y_i)^2 + \sum_{i=r+1}^{n} \left(y_i + \frac{c^T U_i}{\xi_i}\right)^2 | \tilde{w} \in \mathbb{R}^r;\ Uy \in \Lambda(\theta)\right\} \\
&= \min\left\{\sum_{i=r+1}^{n} \left(y_i + \frac{c^T U_i}{\xi_i}\right)^2 | Uy \in \Lambda(\theta)\right\}. \qquad (3.8)
\end{aligned}$$

Thus, for any y with $Uy \in \Lambda(\theta)$, it holds that

$$\delta^2(\theta) \leq \sum_{i=r+1}^{n} \left(y_i + \frac{c^T U_i}{\xi_i}\right)^2. \tag{3.9}$$

For any $x \in [-1,1]^n$, we consider the following two cases:
Case (a): $x = Uy \in \Lambda(\theta)$. It then follows from Lemma 3.1 and (3.9) that

$$\begin{aligned}
q(x) &= v(D) + \sum_{i=r+1}^{n} \xi_i \left(y_i + \frac{c^T U_i}{\xi_i}\right)^2 - \sum_{i=1}^{n} \lambda_i^*(x_i^2 - 1) \\
&\geq v(D) + \xi_{r+1} \sum_{i=r+1}^{n} \left(y_i + \frac{c^T U_i}{\xi_i}\right)^2 \\
&\geq v(D) + \xi_{r+1}\delta^2(\theta). \qquad (3.10)
\end{aligned}$$

Case (b): $x \in [-1,1]^n \setminus \Lambda(\theta)$. By (3.1), there is an $i_0 \in J^*$ such that $1 - x_{i_0}^2 > \theta$. It follows from Lemma 3.1 that

$$\begin{aligned}
q(x) &= v(D) + \sum_{i=r+1}^{n} \xi_i \left(y_i + \frac{c^T U_i}{\xi_i} \right)^2 - \sum_{i=1}^{n} \lambda_i^*(x_i^2 - 1) \\
&\geq v(D) + \sum_{i=1}^{n} \lambda_i^*(1 - x_i^2) \\
&> v(D) + \lambda_{i_0}^* \theta \\
&\geq v(D) + \theta \min_{i \in J^*} \lambda_i^*. \qquad (3.11)
\end{aligned}$$

Combining (3.10) and (3.11) yields

$$v(P) - v(D) \geq \max_{\theta \in [0,1]} \phi(\theta), \qquad (3.12)$$

where $\phi(\theta)$ is defined in (3.5). Notice that $\delta^2(\theta)$ is a continuous and decreasing function of θ on $[0,1]$ with $\delta(0) > 0$ and $\delta(1) = 0$. Also, $\min_{i \in J^*} \lambda_i^* > 0$. So $\phi(\theta)$ has a unique maximizer θ^* on $[0,1]$ with $0 < \theta^* < 1$ such that

$$\max_{\theta \in [0,1]} \phi(\theta) = \theta^* \min_{i \in J^*} \lambda_i^* = \xi_{r+1} \delta^2(\theta^*),$$

which is (3.6). ∎

Remark 3.1. Combining (3.4) with (3.6) in Theorem 3.1, we know that there is a $\theta^* \in (0,1)$ such that $v(P) - v(D) \geq \xi_{r+1}\delta^2(\theta^*)$. This estimation of duality gap clearly leads to an improved lower bound of (P):

$$\nu^* = v(D) + \xi_{r+1} \delta^2(\theta^*). \qquad (3.13)$$

We point out that the improved bound in (3.13) for the box-constrained quadratic program (P) is a generalization of the results in [22] where binary quadratic programs are considered. Consider the max-cut problem

$$\max_{x \in \{-1,1\}^n} x^T Q x, \qquad (3.14)$$

where Q is an $n \times n$ symmetric matrix (not necessarily nonnegative). Notice that for the max-cut problem (3.14), the complementarity condition (2.12) is automatically fulfilled in the saddle point conditions. Consequently, the distance measure $\delta(\theta)$ in (3.2) reduces to $\delta = dist(\{-1,1\}^n, C)$, where

$$C = \{x \in \mathbb{R}^n | (\mathrm{diag}(\lambda^*) - Q)x = 0\}.$$

The derivation of the estimation of duality gap in Theorem 3.1 can be then simplified and an upper bound for the max-cut problem (3.14) is obtained as

$$\nu_s = \gamma^* - \xi_{r+1} \delta^2.$$

It has been shown in [22, Proposition 3] that ν_s is identical to the improved upper bound derived in [15] where a reduced-rank binary quadratic program is used to measure the duality gap between the max-cut problem and its SDP relaxation.

4. Computation of the duality gap estimate. We discuss in this section how to compute the duality gap estimate $\max_{\theta \in [0,1]} \phi(\theta)$ in (3.4), or, equivalently, $\xi_{r+1}\delta^2(\theta^*)$

in (3.6). We first establish the relationship between the computation of $\delta(\theta^*)$ and the partition of the set C. We then describe a procedure of computing $\max_{\theta\in[0,1]} \phi(\theta)$.

For any $\theta \in [0,1]$, by the definition of $\delta(\theta)$ (cf. (3.2)), we have

$$\begin{aligned}\delta^2(\theta) &= \min\{\|x-y\|^2 | x \in C,\ y \in \Lambda(\theta)\}\\ &= \min \sum_{i\in I^*}(x_i-y_i)^2 + \sum_{i\in J^*}(x_i-y_i)^2\\ &\quad \text{s.t. } x \in C, \quad y_i \in [-1,1], \quad i \in I^*,\\ &\qquad\quad y_i \in [-1,-\sqrt{1-\theta}] \cup [\sqrt{1-\theta},1], \quad i \in J^*.\end{aligned} \tag{4.1}$$

We notice in the above expression that for any $x \in C$, the sign of the variable y_i $(i \in J^*)$ that achieves the minimization solely depends on the sign of x_i $(i \in J^*)$. Namely, $y_i \in [-1,-\sqrt{1-\theta}]$ if $x_i < 0$ and $y_i \in [\sqrt{1-\theta},1]$ if $x_i > 0$ in the term $\sum_{i\in J^*}(x_i-y_i)^2$.

Suppose that we can partition the set C into m polyhedral subsets, C_k $(k = 1,\dots,m)$ in such a way that the sign of each x_i $(i \in J^*)$ does not change within C_k. Let

$$C = \bigcup_{k=1}^{m} C_k. \tag{4.2}$$

Suppose also that a relative interior-point π^k of C_k can be computed for each k. For each $k = 1,\dots,m$, let $\omega^k = (\omega_i^k)_{i\in J^*}$ be the *sign vector* of C_k, where

$$\omega_i^k = \begin{cases} -1, & \pi_i^k < 0,\\ 1, & \pi_i^k \geq 0.\end{cases} \tag{4.3}$$

Define

$$\Lambda_k(\theta) = \{y \in \mathbb{R}^n | y_i \in [-1,1],\ i \in I^*;\ \omega_i^k y_i \in [\sqrt{1-\theta},\ 1],\ i \in J^*\}, \tag{4.4}$$

$$\delta_k^2(\theta) = \min\left\{\sum_{i\in I^*}(x_i-y_i)^2 + \sum_{i\in J^*}(x_i-y_i)^2 | x \in C,\ y \in \Lambda_k(\theta)\right\}. \tag{4.5}$$

Since C and $\Lambda_k(\theta)$ are both polyhedral sets, $\delta_k^2(\theta)$ can be calculated by solving a convex quadratic program.

LEMMA 4.1. *For any $\theta \in [0,1]$, it holds that*

$$\delta^2(\theta) = \min_{k=1,\dots,m} \delta_k^2(\theta). \tag{4.6}$$

Proof. By (4.1)–(4.5), we have

$$\begin{aligned}\delta^2(\theta) &= \min\{\|x-y\|^2 | x \in C,\ y \in \Lambda(\theta)\}\\ &= \min_{k=1,\dots,m} \min\{\|x-y\|^2 | x \in C,\ y \in \Lambda_k(\theta)\}\\ &= \min_{k=1,\dots,m} \delta_k^2(\theta),\end{aligned}$$

which is exactly (4.6). ∎

Now, for any $\theta \in [0,1]$ and $k = 1, \ldots, m$, define

$$\phi_k(\theta) = \min\{\theta \min_{i\in J^*} \lambda_i^*,\ \xi_{r+1}\delta_k^2(\theta)\}. \tag{4.7}$$

By Lemma 4.1, $\delta(0) > 0$ implies $\delta_k(0) > 0$ for $k = 1, \ldots, m$. Define

$$K = \{k \in \{1, \cdots, m\} | \min_{i\in J^*} \lambda_i^* > \xi_{r+1}\delta_k^2(1)\}.$$

For $k \in K$, let θ^k be the unique solution of the following equation:

$$\theta^k \min_{i\in J^*} \lambda_i^* = \xi_{r+1}\delta_k^2(\theta^k). \tag{4.8}$$

Then θ^k is the maximizer of $\phi_k(\theta)$ on $[0,1]$. On the other hand, for $k \in \{1, \ldots, m\} \setminus K$, $\theta^k = 1$ is the unique maximizer of $\phi_k(\theta)$ on $[0,1]$.

THEOREM 4.1. *Let θ^k be defined in* (4.8) *when $k \in K$ or $\theta^k = 1$ when $k \notin K$. Then we have*

$$\max_{\theta\in[0,1]} \phi(\theta) = \min_{k=1,\ldots,m} \min_{\theta\in[0,1]} \phi_k(\theta) = (\min_{i\in J^*} \lambda_i^*)(\min_{k=1,\ldots,m} \theta^k). \tag{4.9}$$

Proof. We first notice that if $\delta(0) = 0$, then all terms in (4.9) are equal to zero. Suppose that $\delta(0) > 0$. By Lemma 4.1 and (4.7), we have

$$\begin{aligned}
\max_{\theta\in[0,1]} \phi(\theta) &= \max_{\theta\in[0,1]} \min\{\theta \min_{i\in J^*} \lambda_i^*,\ \xi_{r+1}\delta^2(\theta)\} \\
&= \max_{\theta\in[0,1]} \min\{\theta \min_{i\in J^*} \lambda_i^*,\ \xi_{r+1} \min_{k=1,\ldots,m} \delta_k^2(\theta)\} \\
&= \max_{\theta\in[0,1]} \min_{k=1,\ldots,m} \min\{\theta \min_{i\in J^*} \lambda_i^*,\ \xi_{r+1}\delta_k^2(\theta)\} \\
&= \max_{\theta\in[0,1]} \min_{k=1,\ldots,m} \phi_k(\theta) \\
&\le \min_{k=1,\ldots,m} \max_{\theta\in[0,1]} \phi_k(\theta). \qquad (4.10)
\end{aligned}$$

On the other hand, from the proof of Theorem 3.1, $\max_{\theta\in[0,1]} \phi(\theta)$ has a unique optimal solution $\theta^* \in (0,1)$ such that $\phi(\theta^*) = \theta^* \min_{i\in J^*} \lambda_i^* = \xi_{r+1}\delta^2(\theta^*)$. By Lemma 4.1, there exists $k_0 \in \{1, \ldots, m\}$ such that

$$\max_{\theta\in[0,1]} \phi(\theta) = \theta^* \min_{i\in J^*} \lambda_i^* = \xi_{r+1}\delta^2(\theta^*) = \min_{k=1,\ldots,m} \xi_{r+1}\delta_k^2(\theta^*) = \xi_{r+1}\delta_{k_0}^2(\theta^*).$$

Since the solution of (4.8) is unique, we deduce from the above equation that $\theta_{k_0} = \theta^*$. Hence,

$$\min_{k=1,\ldots,m} \max_{\theta\in[0,1]} \phi_k(\theta) = (\min_{i\in J^*} \lambda_i^*)(\min_{k=1,\ldots,m} \theta^k) \le \theta^{k_0} \min_{i\in J^*} \lambda_i^* = \theta^* \min_{i\in J^*} \lambda_i^* = \max_{\theta\in[0,1]} \phi(\theta),$$

which, together with (4.10), implies (4.9). ∎

Next, we discuss how to compute $\max_{\theta\in[0,1]} \phi(\theta)$ efficiently. Theorem 4.1 suggests that $\max_{\theta\in[0,1]} \phi(\theta)$ can be obtained by calculating the minimum value of θ^k $(k = 1, \ldots, m)$, which is determined by solving

$$\max_{\theta\in[0,1]} \phi_k(\theta). \tag{4.11}$$

Notice that if $\xi_{r+1}\delta_k^2(1) > \min_{i\in J^*}\lambda_i^*$, then $\theta^k = 1$ is the unique maximizer of $\phi_k(\theta)$ on $[0, 1]$. Otherwise, (4.11) can be reduced to a second-order cone program as is shown in the following theorem.

THEOREM 4.2. *Assume that $\xi_{r+1}\delta_k^2(1) \le \min_{i\in J^*}\lambda_i^*$. Let σ^k be the optimal value of the following second-order cone program:*

$$\begin{aligned}
&\max \sigma\\
&\text{s.t.} \left\| \begin{pmatrix} \sqrt{\xi_{r+1}}\cdot(x-y) \\ \sqrt{\min_{i\in J^*}\lambda_i^*}\cdot\sigma \end{pmatrix} \right\| \le \sqrt{\min_{i\in J^*}\lambda_i^*},\\
&\quad Q^*x + c = 0,\\
&\quad -1 \le y_i \le 1, \quad i \in I^*,\\
&\quad \sigma \le \omega_i^k y_i \le 1, \quad i \in J^*,\\
&\quad 0 \le \sigma \le 1,
\end{aligned} \tag{4.12}$$

then $\theta^k = 1 - \sigma_k^2$ solves (4.11).

Proof. We first notice that if $\xi_{r+1}\delta_k^2(1) \le \min_{i\in J^*}\lambda_i^*$, then θ^k solves (4.11) if and only if (4.8) holds, while condition (4.8) can be expressed as

$$\theta^k = \min\{\theta|\xi_{r+1}\delta_k^2(\theta) \le \theta\min_{i\in J^*}\lambda_i^*, 0 \le \theta \le 1\}. \tag{4.13}$$

By the definition of $\delta_k^2(\theta)$ (cf. (4.5)), it is easy to see that (4.13) is equivalent to

$$\begin{aligned}
&\min \theta\\
&\text{s.t. } \xi_{r+1}\|x-y\|^2 \le \theta\min_{i\in J^*}\lambda_i^*,\\
&\quad Q^*x + c = 0,\\
&\quad -1 \le y_i \le 1, \quad i \in I^*,\\
&\quad \sqrt{1-\theta} \le \omega_i^k y_i \le 1, \quad i \in J^*,\\
&\quad 0 \le \theta \le 1.
\end{aligned} \tag{4.14}$$

Let $\sigma = \sqrt{1-\theta}$ for $\theta \in [0, 1]$. Then (4.14) can be expressed as

$$\begin{aligned}
&\min 1-\sigma^2\\
&\text{s.t. } \xi_{r+1}\|x-y\|^2 \le (1-\sigma^2)\min_{i\in J^*}\lambda_i^*,\\
&\quad Q^*x + c = 0,\\
&\quad -1 \le y_i \le 1, \quad i \in I^*,\\
&\quad \sigma \le \omega_i^k y_i \le 1, \quad i \in J^*,\\
&\quad 0 \le \sigma \le 1.
\end{aligned} \tag{4.15}$$

Since σ^2 is strictly increasing on $[0, 1]$, (4.15) has the same optimal solution as the following problem:

$$
\begin{aligned}
&\max \sigma \\
&\text{s.t. } \xi_{r+1}\|x-y\|^2 + \sigma^2 \min_{i\in J^*} \lambda_i^* \le \min_{i\in J^*} \lambda_i^*, \\
&\quad Q^* x + c = 0, \\
&\quad -1 \le y_i \le 1, \quad i \in I^*, \\
&\quad \sigma \le \omega_i^k y_i \le 1, \quad i \in J^*, \\
&\quad 0 \le \sigma \le 1. \qquad (4.16)
\end{aligned}
$$

Notice that the first constraint in (4.16) can be reformulated as the second-order cone constraint in (4.12). Therefore, (4.16) is equivalent to (4.12) and hence $\theta^k = 1 - \sigma_k^2$ solves (4.11), where σ_k is the optimal value of (4.12). □

Theorems 4.1 and 4.2 suggest that $\max_{\theta\in[0,1]} \phi(\theta) = \xi_{r+1}\delta^2(\theta^*)$ in Theorem 3.1 can be computed in the following three steps:

(i) Partition C into polyhedral sets C_k $(k = 1, \ldots, m)$ such that x_i $(i \in J^*)$ does not change sign within each C_k. Calculate the sign vector ω^k defined in (4.3) for each C_k.
(ii) Compute θ^k by solving (4.12) or $\theta^k = 1$ if $\xi_{r+1}\delta_k^2(1) > \min_{i\in J^*} \lambda_i^*$.
(iii) Calculate $\max_{\theta\in[0,1]} \phi(\theta)$ by (4.9).

Remark 4.1. Notice that apart from the improved lower bound $\nu^* = v(D) + \xi_{r+1}\delta^2(\theta^*)$ in (3.4), we can also obtain a feasible solution related to the distance measure $\delta(\theta^*)$ between C and $\Lambda(\theta^*)$ when computing $\max_{\theta\in[0,1]} \phi(\theta)$ in the above steps (ii) and (iii). Indeed, let (x^*, y^*) be the optimal solution that achieves the distance measure $\delta(\theta^*)$. From the above discussions, we see that either (x^*, y^*) is an optimal solution to the second-order cone program (4.12) or it is an optimal solution to the convex quadratic program in (4.5) when $\xi_{r+1}\delta_k^2(1) > \min_{i\in J^*} \lambda_i^*$. Note that $y^* \in \Lambda(\theta^*) \subseteq [0,1]^n$. To satisfy the complementarity condition for (P), y_i^* $(i \in J^*)$ can be rounded off to -1 or 1 to obtain a new solution $x_s \in \Lambda^* \subseteq [0,1]^n$ (cf. (2.14)). The feasible solution x_s generated by the above heuristic method can be viewed as an approximate solution associated with the distance measure $\delta(\theta^*)$ and the improved lower bound ν^*. It appears that the feasible solution generated in such a way is of a good quality as evidenced in the numerical examples in section 5.3.

5. Partition of C and cell enumeration.

In this section, we discuss how to partition C into a number of polyhedral subsets C_k's such that x_i $(i \in J^*)$ does not change sign within C_k. It turns out that the partition of the set C is related to the cell enumeration of hyperplane arrangement in discrete geometry.

5.1. Cell enumeration.

Notice that the set C can be expressed as

$$C = \left\{ x \in \mathbb{R}^n | x = x^0 + \sum_{k=1}^{r} z_k U_k,\ z \in \mathbb{R}^r \right\}, \qquad (5.1)$$

where $Q^* x^0 + c = 0$ and $U_1, \ldots, U_r$ are the basis of the null space of Q^*. By (2.17) and item (i) of Lemma 2.4, a special choice of solution x^0 is given by

$$x^0 = Uw, \quad \text{where } w_i = 0, \quad i = 1, \ldots, r, \quad w_i = -\frac{c^T U_i}{\xi_i}, \quad i = r+1, \ldots, n.$$

Let

$$g_j(z) = x_j^0 + \sum_{k=1}^{r} U_{jk} z_k, \qquad j \in J^*,$$

where $U_k = (U_{1k}, \ldots, U_{nk})^T$. Let $t = |J^*|$. Define the following t hyperplanes in $\mathbb{R}^r$:

$$H_j = \{z \in \mathbb{R}^r | g_j(z) = 0\}, \qquad j \in J^*. \tag{5.2}$$

These t hyperplanes partition C into a number of r-dimensional convex polyhedral sets. All faces of these partitioned convex polyhedral sets define an *arrangement* of C. Each r-dimensional convex polyhedral set resulted from this partition is called a *cell* of the hyperplane arrangement. The reader is referred to [9] for more details of hyperplane arrangement.

Define

$$H_j^+ = \{z \in \mathbb{R}^r | g_j(z) \geq 0\}, \qquad j \in J^*,$$
$$H_j^- = \{z \in \mathbb{R}^r | g_j(z) < 0\}, \qquad j \in J^*.$$

Let φ^k be a cell generated from the hyperplane arrangement defined in (5.2) and let ϖ^k be an interior point of φ^k in $\mathbb{R}^r$. Associate the cell φ^k with a sign vector $\chi^k \in \{-1, 1\}^t$ defined by

$$\chi_j^k = \begin{cases} 1 & \text{if } \varpi^k \in H_j^+, \\ -1 & \text{if } \varpi^k \in H_j^-. \end{cases} \tag{5.3}$$

Since the sign vector of a cell is invariant for any interior point of φ^k, we can represent a cell φ^k by its sign vector χ^k.

A key observation is that there is a one-to-one mapping between the cells of the hyperplane arrangement defined by (5.2) and the sets C_k in (4.2). More precisely, each sign vector χ^k of a cell defined in (5.3) corresponds to the sign vector ω^k of the set C_k defined in (4.3) and vice versa. Therefore, ω^k $(k = 1, \ldots, m)$ can be obtained by enumerating all the cells of the hyperplane arrangement defined by (5.2) and computing an interior point for each cell. Efficient search methods for enumerating all the cells of a hyperplane arrangement are proposed in [4], [21], which also find an interior point for each cell of the hyperplane arrangement during the enumeration process. It is known that the number of cells generated from the hyperplane arrangement specified by (5.2) is $m = O(n^r)$ (see, e.g., [28]). Moreover, the complexity of the cell enumeration method in [21] for the hyperplane arrangement defined by (5.2) is $O(m|J^*|) = O(n^{r+1})$. Therefore, the sign vectors ω^k $(k = 1 \ldots, m)$ can be computed efficiently when r is small.

PROPOSITION 5.1. *If r is fixed, then $\delta(\theta^*)$ can be computed in polynomial time.*

Proof. Using the cell enumeration method in [21], the complexity of computing all the sign vectors ω^k $(k = 1, \ldots, m)$ is $O(n^{r+1})$. Also, we notice from Theorem 4.2 that for a given ω^k, θ^k can be calculated in polynomial time. By (4.9), $\delta(\theta^*)$ is calculated by calling $m = O(n^r)$ times of θ^k. All these facts together imply that $\delta(\theta^*)$ can be computed in polynomial time. ∎

We have seen from (2.18) in Proposition 2.4 that $r = \text{rank}(X^*) - 1$ or $r = \text{rank}(Y^*)$ if $c = 0$. So a low rank optimal solution to (P_s) implies a small value of r.

The following result gives an upper bound for the value of r.

PROPOSITION 5.2. *Suppose that there exists a positive definite $p \times p$ principal submatrix of Q, then*

$$r \leq n - p. \tag{5.4}$$

Proof. Let Q_1 be a positive definite $p \times p$ principal submatrix of Q. Without loss of generality, we assume that Q_1 is a $p \times p$ leading principal submatrix of Q. Then the $p \times p$ leading principal submatrix of $Q^* = Q + \mathrm{diag}(\lambda^*)$ is

$$Q_1^* = Q_1 + \mathrm{diag}(\lambda_1^*, \ldots, \lambda_p^*).$$

For any nonzero $v \in \mathbb{R}^p$, we have

$$v^T Q_1^* v = v^T Q_1 v + \sum_{i=1}^{p} \lambda_i^* v_i^2 \geq v^T Q_1 v > 0.$$

Thus, Q_1^* is also positive definite, and hence $\mathrm{rank}(Q^*) \geq p$. Therefore, $r = n - \mathrm{rank}(Q^*) \leq n - p$. This is (5.4). ∎

Remark 5.1. Proposition 5.2 provides an upper bound for the value of r which depends on the order of the positive definite principal submatrix of Q. This upper bound also gives a sufficient condition to guarantee a small r: If there exists a positive definite principal submatrix of Q with an order close to n, then r must be small.

5.2. The case *r* = 1. When $r = 1$, C is a one-dimensional line in $\mathbb{R}^n$ with the following expression:

$$C = \{x \in \mathbb{R}^n | x = x^0 + zU_1,\ z \in \mathbb{R}\},$$

where $U_1 = (U_{11}, \ldots, U_{n1})^T$.

For each $j \in J^*$ with $U_{j1} \neq 0$, the sign of x_j changes at a breaking point $\alpha_j = -x_j^0 / U_{j1}$. It is possible that some α_j's take the same value. Rank all different α_j's for $j \in J^*$ in the following ascending order:

$$\alpha_{j_1} < \alpha_{j_2} < \cdots < \alpha_{j_q},$$

where $1 \leq q \leq |J^*| \leq n$. Let $\alpha_{j_0} = -\infty$ and $\alpha_{j_{q+1}} = +\infty$. Then C is partitioned into $q + 1$ line segments

$$C_k = \{x \in \mathbb{R}^n | x = x^0 + \alpha U_1,\ \alpha \in [\alpha_{j_{k-1}}, \alpha_{j_k}]\},\ k = 1, \ldots, q + 1.$$

The relative interior points π^k of C_k can be taken as follow:

$$\begin{aligned}
\pi^1 &= x^0 + (\alpha_{j_1} - 1) U_1, \\
\pi^k &= x^0 + \frac{1}{2}(\alpha_{j_{k-1}} + \alpha_{j_k}) U_1, \qquad k = 2, \ldots, q, \\
\pi^{q+1} &= x^0 + (\alpha_{j_q} + 1) U_1.
\end{aligned}$$

Therefore, the sign vectors ω^k of C_k in (4.3) can be calculated efficiently via the above procedure.

5.3. Illustrative examples. We now give two illustrative examples to demonstrate the computation of the duality gap estimation. The SDP problems and the second-order cone program problems involved in these examples were modeled by CVX 1.2 [12] and solved by SDPT3 within CVX.

Example 5.1. Consider an instance of (P) with

$$Q = \begin{pmatrix} 13 & 10 & 3 & 7 & 13 \\ 10 & 8 & -9 & 2 & -1 \\ 3 & -9 & -10 & 13 & 19 \\ 7 & 2 & 13 & 13 & 3 \\ 13 & -1 & 19 & 3 & -4 \end{pmatrix}, \qquad c = \begin{pmatrix} -4 \\ 21 \\ 24 \\ 9 \\ 14 \end{pmatrix}.$$

The optimal value of this example is $v(P) = -153.31$. Solving the SDP relaxation (D_s), we obtain the optimal dual solution $\lambda^* = (15.63, 27.86, 43.01, 0, 26.70)^T$ with $v(D_s) = -163.59$. Furthermore, the vector of eigenvalues of $Q^* = Q + \mathrm{diag}(\lambda^*)$ is $\xi = (0, 12.51, 18.55, 45.32, 56.83)^T$. Thus, $r = n - \mathrm{rank}(Q^*) = 5 - 4 = 1$, $I^* = \{4\}$, and $J^* = \{1, 2, 3, 5\}$. It can be calculated that the breaking points are $\alpha_{j_0} = -2.54$, $\alpha_{j_1} = -0.78$, $\alpha_{j_2} = 1.54$, and $\alpha_{j_3} = 16.05$. Thus, C can be partitioned into five line segments. The sign vectors of C_k $(k = 1, 2, 3, 4, 5)$ are $\omega^1 = (-1, -1, -1, 1)^T$, $\omega^2 = (1, -1, -1, 1)^T$, $\omega^3 = (1, -1, -1, -1)^T$, $\omega^4 = (1, -1, 1, -1)^T$, and $\omega^5 = (1, 1, 1, -1)^T$, respectively. Using Theorems 4.1 and 4.2, we can calculate that $\theta^* = 0.1472$ and $\delta(\theta^*) = 0.43$, which leads to a nonzero duality gap. By Theorem 3.1, a lower bound of the duality gap is $\xi_{r+1}\delta^2(\theta^*) = 12.51 \times 0.43^2 = 2.31$. An improved lower bound is given by $\nu^* = -163.59 + 2.31 = -161.28$. So the original gap is reduced by about 22%.

In this example, the feasible solution associated with the distance measure $\delta(\theta^*)$ (see Remark 4.1) is $x_s = (1, -1, -1, -0.27, -1)^T$ with $f(x_s) = -150.99$. The relative error of this solution is $(f(x_s) - v(P))/|v(P)| = (-150.99 + 153.31)/153.31 = 1.51\%$.

Example 5.2. Consider an instance of (P) with

$$Q = \begin{pmatrix} 0 & 3 & 6 & -3 & 6 & 3 & -5 & -4 & 8 & -1 \\ 3 & 0 & -6 & -10 & -3 & 8 & -6 & 8 & 5 & 7 \\ 6 & -6 & 0 & -7 & 7 & 7 & 4 & 9 & 2 & -4 \\ -3 & -10 & -7 & 0 & 9 & 8 & 6 & 9 & 2 & -4 \\ 6 & -3 & 7 & 9 & 0 & -4 & -5 & 4 & -3 & 1 \\ 3 & 8 & 7 & 8 & -4 & 0 & -4 & -10 & -8 & -5 \\ -5 & -6 & 4 & 6 & -5 & -4 & 0 & 9 & -10 & 6 \\ -4 & 8 & 9 & 9 & 4 & -10 & 9 & 0 & -1 & 4 \\ 8 & 5 & 2 & -6 & -3 & -8 & -10 & -1 & 0 & 6 \\ -1 & 7 & -4 & -5 & 1 & -5 & 6 & 4 & 6 & 0 \end{pmatrix}, \qquad c = \begin{pmatrix} -5 \\ -4 \\ -5 \\ 4 \\ 2 \\ 2 \\ -4 \\ -3 \\ 3 \\ 4 \end{pmatrix}.$$

The optimal value of this example is $v(P) = -290$. Solving the SDP relaxation (D_s), we obtain

$$\lambda^* = (10.02, 29.41, 39.64, 41.61, 27.30, 46.01, 20.21, 25.63, 31.38, 15.87)^T$$

and $v(D_s) = -298.64$. Furthermore, the vector of eigenvalues of $Q^* = Q + \mathrm{diag}(\lambda^*)$ is

$$\xi = (0, 0, 6.05, 16.11, 25.02, 36.99, 38.19, 48.61, 55.46, 60.65)^T.$$

Thus, $r = n - \mathrm{rank}(Q^*) = 10 - 8 = 2$, $I^* = \{1, 2\}$, and $J^* = \{3, 4, \ldots, 10\}$. Applying the cell enumeration procedure in [21], we can partition the set C into 56 polyhedral subsets and obtain the corresponding sign vectors ω^k. By Theorems 4.1 and 4.2, we can calculate that $\theta^* = 0.1860$ and $\delta(\theta^*) = 0.55$. An underestimation of the duality gap is $\xi_{r+1}\delta^2(\theta^*) = 6.05 \times 0.55^2 = 1.83$. The corresponding improved lower bound is given by $\nu^* = -298.64 + 1.83 = -296.81$. So the original gap is reduced by about 21%.

In this example, the feasible solution associated with the distance measure $\delta(\theta^*)$ is $x_s = (1, 1, 1, 1, -1, -1, -1, -1, -1, 1)^T$ with $f(x_s) = -290$, which is exactly the optimal solution.

6. Extension to problem with box and equality constraints. In this section, we extend our main results for (P) to problem

$$(P_e) \quad \min\{x^T Qx + 2c^T x | Ax = b,\ x \in [-1, 1]^n\}.$$

We assume that the feasible set $X = \{x \in [-1, 1]^n | Ax = b\}$ has nonempty relative interior which ensures the polynomial solvability of the SDP relaxation of (P_e). For simplicity, we describe only the main results on the estimation of duality gap.

The Lagrangian dual problem of (P_e) is

$$(D_e) \quad \max_{\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^n_+} d(\mu, \lambda),$$

where

$$d(\mu, \lambda) = \inf_{x \in \mathbb{R}^n} \{x^T [Q + \operatorname{diag}(\lambda)]x + (2c + A^T \mu)^T x - e^T \lambda - \mu^T b\}.$$

The SDP reformulation of (D_e) is

$$\begin{aligned} &\max -\tau - e^T\lambda - b^T\mu \\ &\text{s.t.} \begin{pmatrix} Q + \operatorname{diag}(\lambda) & c + \frac{1}{2}A^T\mu \\ (c + \frac{1}{2}A^T\mu)^T & \tau \end{pmatrix} \succcurlyeq 0, \\ &\quad \tau \in \mathbb{R}, \quad \lambda \in \mathbb{R}^n_+, \quad \mu \in \mathbb{R}^m. \end{aligned}$$

Let (μ^*, λ^*) be an optimal solution to (D_e). Define

$$\begin{aligned} &Q^* = Q + \operatorname{diag}(\lambda^*), \quad J^* = \{i \in \{1, \ldots, n\} | \lambda_i^* > 0\}, \quad I^* = \{1, \ldots, n\} \setminus J^*, \\ &C_e = \left\{x \in \mathbb{R}^n | Q^*x + c + \frac{1}{2}A^T\mu^* = 0\right\}, \quad S = \{x \in \mathbb{R}^n | Ax = b\}, \\ &\Lambda^* = \{x \in \mathbb{R}^n | x_i \in [-1, 1] \quad \text{for } i \in I^*, \quad x_i \in \{-1, 1\} \quad \text{for } i \in J^*\}. \end{aligned}$$

For $0 \le \theta \le 1$, define the perturbed set

$$\Lambda(\theta) = \{x \in \mathbb{R}^n | -1 \le x_i \le 1 \text{ for } i \in I^*,\ 1 - \theta \le x_i^2 \le 1 \text{ for } i \in J^*\}.$$

The distance between $C_e \cap S$ and $\Lambda(\theta)$ is

$$\delta_e(\theta) = \operatorname{dist}(C_e \cap S, \Lambda(\theta)) = \min\{\|x - z\| | x \in C_e \cap S,\ z \in \Lambda(\theta)\}.$$

By the saddle point conditions of (P_e), we can verify that $v(P_e) = v(D_e)$ if and only if $\delta_e(0) = 0$. Also, $J^* = \varnothing$ implies $\delta_e(0) = 0$, and hence $v(P_e) = v(D_e)$.

LEMMA 6.1. *Let A be an $m \times n$ matrix with full row rank. Let V be an $n \times r$ matrix with $V^T V = I_r$. Let $b \in \mathbb{R}^m$ be such that $\{x \in \mathbb{R}^r | AVx = b\} \neq \varnothing$. Then for any $y \in \mathbb{R}^r$*

$$\min\{\|y-x\|^2 | AVx=b\} = (AVy-b)^T(AVV^TA^T)^+(AVy-b), \tag{6.1}$$

where $\|\cdot\|$ *is the Euclidean norm and* B^+ *denotes the Moore–Penrose generalized inverse of* B.

Proof. Let x^* be the optimal solution of (6.1). By the KKT conditions, there exists $\lambda^* \in \mathbb{R}^m$ such that

$$2(y-x^*) - V^TA^T\lambda^* = 0, \qquad AVx^* = b.$$

Thus, $\lambda^* = 2(AVV^TA^T)^+AV(y-x^*) = 2(AVV^TA^T)^+(AVy-b)$. Therefore, $y - x^* = (1/2)V^TA^T\lambda^* = V^TA^T(AVV^TA^T)^+(AVy-b)$, which implies (6.1). ☐

In what follows, we denote by $\lambda_{\max}(B)$ the maximum eigenvalue of matrix B. Suppose that

$$U^TQ^*U = \mathrm{diag}(0,\ldots,0,\xi_{r+1},\ldots,\xi_n).$$

THEOREM 6.1. *Assume that* $J^* \neq \emptyset$ *and* $\delta_e(0) > 0$. *Let* $V_0 = (U_1,\ldots,U_r)$ *and* $V_1 = (U_{r+1},\ldots,U_n)$. *Then there exists a* $\theta^* \in (0,1)$ *such that*

$$v(P_e) - v(D_e) \geq \frac{1}{1+\kappa}\xi_{r+1}\delta_e^2(\theta^*), \tag{6.2}$$

where

$$\kappa = \lambda_{\max}(V_1^TA^T(AV_0V_0^TA^T)^+AV_1).$$

Proof. Let

$$\vartheta_i = -\frac{(c+A^T\mu^*)^TU_i}{\xi_i}, \qquad i = r+1,\ldots,n.$$

Similar to (3.8), we have

$$\delta_e^2(\theta) = \min_{Uy\in\Lambda(\theta)}\left\{\psi_e(y) + \sum_{i=r+1}^{n}(y_i-\vartheta_i)^2\right\}, \tag{6.3}$$

where

$$\psi_e(y) = \min\left\{\sum_{i=1}^{r}(w_i-y_i)^2 \Big| \sum_{i=1}^{r}(AU_i)w_i = b - \sum_{i=r+1}^{n}(AU_i)\vartheta_i\right\}.$$

Let $w^0 = (w_1,\ldots,w_r)^T$, $\vartheta = (\vartheta_{r+1},\ldots,\vartheta_n)$, $y^0 = (y_1,\ldots,y_r)^T$, $y^1 = (y_{r+1},\ldots,y_n)^T$. Notice that A is of full row rank and $V_0^TV_0 = I_r$. By Lemma 6.1, we have

$$\begin{aligned}\psi_e(y) &= \min\{\|w^0-y^0\|^2 | AV_0w^0 = b - AV_1\vartheta\}\\ &= (AV_0y^0 - b + AV_1\vartheta)^T(AV_0V_0^TA^T)^+(AV_0y^0 - b + AV_1\vartheta).\end{aligned} \tag{6.4}$$

Suppose now that $Uy \in S$. Then $AUy = b$, i.e., $AV_0y^0 = b - \sum_{i=r+1}(AU_i)y_i$. We obtain from (6.4) that

$$
\begin{aligned}
\psi_e(y) &= \left(\sum_{i=r+1}^{n}(AU_i)(y_i-\vartheta_i)\right)^T (AV_0V_0^TA^T)^+\left(\sum_{i=r+1}^{n}(AU_i)(y_i-\vartheta_i)\right)\\
&= (AV_1(y^1-\vartheta))^T(AV_0V_0^TA^T)^+(AV_1(y^1-\vartheta))\\
&= (y^1-\vartheta)^TV_1^TA^T(AV_0V_0^TA^T)^+AV_1(y^1-\vartheta)\\
&\le \lambda_{\max}(V_1^TA^T(AV_0V_0^TA^T)^+AV_1)(y^1-\vartheta)^T(y^1-\vartheta)\\
&= \kappa\sum_{i=r+1}^{n}(y_i-\vartheta_i)^2. \qquad (6.5)
\end{aligned}
$$

Therefore, for any y with $Uy\in\Lambda(\theta)\cap S$, we deduce from (6.3) and (6.5) that

$$
\delta_e^2(\theta)\le(1+\kappa)\sum_{i=r+1}^{n}(y_i-\vartheta_i)^2. \qquad (6.6)
$$

Using (6.6) and the similar arguments as in the proof of Theorem 3.1, we can show that

$$
v(P_e)-v(D_e)\ge\max_{\theta\in[0,1]}\phi_e(\theta), \qquad (6.7)
$$

where

$$
\phi_e(\theta)=\min\left\{\theta\min_{i\in J_e}\lambda_i^*,\ \frac{1}{1+\kappa}\xi_{r+1}\delta_e^2(\theta)\right\}. \qquad (6.8)
$$

Note that $\delta_e^2(\theta)$ is a continuous and decreasing function of θ on $[0,1]$ and $\min_{i\in J_e}\lambda_i^*>0$. Also, $\delta_e(0)>0$ and $\delta_e(1)=0$. Therefore, (6.7) and (6.8) imply that $\phi_e(\theta)$ has a unique maximizer $\theta^*\in(0,1)$ such that (6.2) holds. ∎

The computation of the distance $\delta_e(\theta^*)$ in Theorem 6.1 is similar to that of $\delta(\theta^*)$ in section 4. Let

$$
E=\begin{pmatrix}Q^*\\A\end{pmatrix},\qquad \varepsilon=\begin{pmatrix}-c-A^T\mu^*\\b\end{pmatrix}.
$$

Then $C_e\cap S=\{x\in\mathbb{R}^n|Ex=\varepsilon\}$. Notice that $n-r\le\operatorname{rank}(E)\le n-r+m$. Thus, the dimension of $C_e\cap S$, denoted by $s=\dim(C_e\cap S)=n-\operatorname{rank}(E)$, satisfies $\max(0,r-m)\le s\le r$.

We partition $C_e\cap S$ into a number of s-dimensional convex polyhedral subsets C_k's such that x_i $(i\in J^*)$ does not change sign within C_k. Denote by ω^k the sign vector of the set C_k. For each k, we solve the following second-order cone program:

$$
\begin{aligned}
&\max\ \sigma \\
&\text{s.t.}\ \left\| \begin{pmatrix} \sqrt{\frac{\xi_{r+1}}{1+\kappa}} \cdot (x-y) \\ \sqrt{\min_{i\in J^*} \lambda_i^*} \cdot \sigma \end{pmatrix} \right\| \le \sqrt{\min_{i\in J^*} \lambda_i^*}, \\
&\quad (Q+\operatorname{diag}(\lambda^*))x + c + A^T\mu^* = 0, \\
&\quad Ax = b, \\
&\quad -1 \le y_i \le 1, \qquad i \in I^*, \\
&\quad \sigma \le \omega_i^k y_i \le 1, \qquad i \in J^*, \\
&\quad 0 \le \sigma \le 1.
\end{aligned} \tag{6.9}
$$

If (6.9) is infeasible, set $\theta^k = 1$. Otherwise, let (σ^k, x^k, y^k) be any maximizer of problem (6.9) and set $\theta^k = 1 - (\sigma^k)^2$. Then $\theta^* = \min_{k=1,\dots,m} \theta^k$.

7. Conclusions. A tight lower bound is a key to the success of any exact solution algorithm for nonconvex minimization optimization problems. SDP relaxations have been known to generate tight lower bounds for both combinatorial optimization and nonconvex continuous quadratic programs. A natural yet challenging question is whether or not we are able to achieve bounds better than SDP bounds. Although several duality gap reduction schemes have been derived for binary quadratic programming [22], [29], especially for max-cut problems [6], [15], the current paper, to the best of our knowledge, represents the first attempt in devising a duality gap reduction scheme for nonconvex quadratic programs with box and linear equality constraints.

We have shown in this paper that a reduction of the duality gap between the box constrained quadratic optimization problem and its Lagrangian dual or SDP relaxation can be achieved using a parameterized distance measure related to the dissatisfaction degree of the saddle point conditions. We have also applied similar strategy to cases with additional linear equality constraints. We have demonstrated that the computation of the estimation is related to the cell enumeration of hyperplane arrangement. Furthermore, the computational complexity of such a cell enumeration is mainly determined by the degeneracy degree of the modified coefficient matrix determined from SDP relaxation.

As seen from both Examples 5.1 and 5.2, the feasible solution x_s extracted from the computation of the distance measure $\delta(\theta^*)$ (see Remark 4.1) appears to be of a good quality. An interesting question is how to analyze the theoretical quality of such a feasible solution, which will be one of our future research topics.

REFERENCES

[1] F. Alizadeh, J. P. A. Haeberly, and M. L. Overton, *Complementarity and nondegeneracy in semidefinite programming*, Math. Program., 77 (1997), pp. 111–128.

[2] L. T. H. An and P. D. Tao, *A branch and bound method via D. C. optimization algorithms and ellipsoidal technique for box constrained nonconvex quadratic problems*, J. Global Optim., 13 (1998), pp. 171–206.

[3] P. De Angelis, P. M. Pardalos, and G. Toraldo, *Quadratic programming with box constraints*, in Developments in Global Optimization, I. M. Bomze, T. Csendes, R. Horst, and P. M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, 1997, pp. 73–93.

[4] D. Avis and K. Fukuda, *Reverse search for enumeration*, Discrete Appl. Math., 65 (1996), pp. 21–46.

[5] A. Beck and M. Teboulle, *Global optimality conditions for quadratic optimization problems with binary constraints*, SIAM J. Optim., 11 (2000), pp. 179–188.

[6] W. Ben-Ameur and J. Neto, *Spectral bounds for the maximum cut problem*, Networks, 52 (2008), pp. 8–13.

[7] A. Ben-Tal, *Conic and Robust Optimization*, Lecture Notes, Universita di Roma La Sapienzia, Rome, Italy, 2002, available online at http://ie.technion.ac.il/Home/Users/morbt/rom.pdf.

[8] S. Burer and D. Vandenbussche, *Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound*, Comput. Optim. Appl., 43 (2009), pp. 181–195.

[9] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, Springer-Verlag, Heidelberg, 1987.

[10] C. A. Floudas and V. Visweswaran, *Quadratic optimization*, in Handbook of Global Optimization, R. Horst and P. M. Pardalos, eds., Kluwer Academic Publishers, 1995, pp. 217–269.

[11] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM, 42 (1995), pp. 1115–1145.

[12] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming*, version 1. 21, 2011, available online at http://cvxr.com/cvx.

[13] P. Hansen, B. Jaumard, M. L. Ruiz, and J. Xiong, *Global minimization of indefinite quadratic functions subject to box constraints*, Naval Res. Logist., 40 (1993), pp. 373–392.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[15] U. Malik, I. M. Jaimoukha, G. D. Halikias, and S. K. Gungah, *On the gap between the quadratic integer programming problem and its semidefinite relaxation*, Math. Program., 107 (2006), pp. 505–515.

[16] Y. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, Optim. Methods Softw., 9 (1998), pp. 141–160.

[17] Y. Nesterov and A. Nemirovsky, *Interior-Point Polynomial Methods in Convex Programming*, SIAM, Philadelphia, PA, 1994.

[18] P. M. Pardalos and J. B. Rosen, *Constrained Global Optimization: Algorithms and Applications*, Springer-Verlag, Berlin, 1987.

[19] I. G. Rosenberg, 0-1 *optimization and nonlinear programming*, RAIRO Oper. Res., 6 (1972), pp. 95–97.

[20] N. Z. Shor, *Quadratic optimization problems*, Sov. J. Comput. Syst. Sci., 25 (1987), pp. 1–11.

[21] N. Sleumer, *Output-sensitive cell enumeration in hyperplane arrangements*, Nordic J. Comput., 6 (1999), pp. 137–161.

[22] X. L. Sun, C.L. Liu, D. Li, and J. J. Gao, *On duality gap in binary quadratic optimization*, J. Global Optim., to appear.

[23] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.

[24] D. Vanderbussche and G. L. Nemhauser, *A branch-and-cut algorithm for nonconvex quadratic programs with box constraints*, Math. Program., 102 (2005), pp. 371–405.

[25] D. Vanderbussche and G. L. Nemhauser, *A polyhedral study of nonconvex quadratic programs with box constraints*, Math. Program., 102 (2005), pp. 531–557.

[26] Y. Yajima and T. Fujie, *A polyhedral approach for nonconvex quadratic programming problems with box constraints*, J. Global Optim., 13 (1998), pp. 151–170.

[27] Y. Ye, *Approximating quadratic programming with bound and quadratic constraints*, Math. Program., 84 (1999), pp. 219–226.

[28] T. Zaslavsky, *Facing up to arrangements: face-count formulas for partitions of space by hyperplanes*, Mem. Am. Math. Soc., 1 (1975), pp. 1–101.

[29] X. J. Zheng, X. L. Sun, D. Li, and Y. Xia, *Duality gap estimation of linear equality constrained binary quadratic programming*, Math. Oper. Res., 35 (2010), pp. 864–880.