Further Results on Approximating Nonconvex Quadratic Optimization by Semidefinite Programming Relaxation[1]

October 3, 2001

Paul Tseng[2]

Abstract

We study approximation bounds for the SDP relaxation of quadratically constrained quadratic optimization: $\min f^0(x)$ subject to $f^k(x) \le 0$, $k = 1, ..., m$, where $f^k(x) = x^T A^k x + (b^k)^T x + c^k$. In the special case of ellipsoid constraints with interior feasible solution at 0, we show that the SDP relaxation, coupled with a rank-1 decomposition result of Sturm and Zhang, yields a feasible solution of the original problem with objective value at most $(1-\gamma)^2/(\sqrt{m}+\gamma)^2$ times the optimal objective value, where $\gamma = \sqrt{\max_k f^k(0) + 1}$. If the ellipsoids have a common center, this improves on the estimate $1/\left(2\ln(2(m+1)^2)\right)$ of Nemirovski et al. when $m \le 11$. For the single trust-region problem, corresponding to $m = 1$, this yields an exact optimal solution. In the general case, we extend some bounds derived by Nesterov and Ye for the special case where A^k is diagonal and $b^k = 0$ for $k = 1, ..., m$. We also discuss the generation of approximate solutions with high probability.

Key words. Quadratically constrained quadratic optimization, semidefinite programming relaxation, approximation algorithm.

1 Introduction

Consider the quadratically constrained quadratic program (QP):

$$\begin{array}{rll} v_{\text{QP}} := & \min & f^0(x) \\ & \text{s.t.} & f^k(x) \le 0, \;\; k = 1, ..., m, \end{array} \tag{1}$$

where $f^k(x) = x^T A^k x + (b^k)^T x + c^k$, with $A^k \in \Re^{n \times n}$ symmetric, $b^k \in \Re^n$, $c^k \in \Re$ for $k = 0, 1, ..., m$. We assume $c^0 = 0$. If $c^0 \ne 0$, our results still hold by suitably replacing $f^0(x)$ with $f^0(x) - f^0(0)$. This problem is NP-hard.

It was known through the work of Lovász and Schrivjer [5] and others that certain NP-hard combinatorial optimization problems can be approximated by semidefinite programming (SDP) problems, for which efficient solution method exist [1, 10, 11]. This motivated an important work of Goemans and Williamson [4] showing that, for special cases of (1) corresponding to certain NP-hard problems like Maximum Cut, SDP relaxation yields very good (randomized) approximation algorithms. This work was subsequently extended

[1]This research is supported by the National Science Foundation Grant CCR-9731273.
[2]Department of Mathematics, University of Washington, Seattle, Washington 98195, U.S.A. (tseng@math.washington.edu)

by Nesterov [7, 11], Ye [11, 17, 18], Nemirovski et al. [6], and Zhang [19] to other cases of (1), as well as to other combinatorial optimization problems–see [2, 16] and references therein.

Motivated by the aforementioned work, in this paper we make further study of the SDP relaxation of (1). In particular, by defining

$$B^k := \begin{bmatrix} A^k & b^k/2 \\ (b^k)^T/2 & c^k \end{bmatrix}, \quad k = 0, 1, ..., m,$$

and introducing $x_{n+1} = 1$, we rewrite (1) equivalently as

$$\begin{array}{ll} \min & \sum_{i,j=1}^{n+1} B^0_{ij} x_i x_j \\ \text{s.t.} & \sum_{i,j=1}^{n+1} B^k_{ij} x_i x_j \leq 0, \ k = 1, ..., m, \\ & x_{n+1} = 1. \end{array}$$

By further making the transformation $X = xx^T = [x_i x_j]_{i,j=1}^{n+1}$ for $x \in \Re^{n+1}$ with $x_{n+1} = 1$, we write the above problem equivalently as

$$\begin{array}{ll} \min & \langle B^0, X\rangle \\ \text{s.t.} & \langle B^k, X\rangle \leq 0, \ k = 1, ..., m, \\ & X_{n+1 n+1} = 1, \ X \succeq 0, \ \text{Rank} X = 1. \end{array}$$

Relaxing the rank-1 constraint yields the SDP relaxation of (1):

$$\begin{array}{lll} v_{\text{SDP}} := & \min & \langle B^0, X\rangle \\ & \text{s.t.} & \langle B^k, X\rangle \leq 0, \ k = 1, ..., m, \\ & & X_{n+1 n+1} = 1, \ X \succeq 0. \end{array} \tag{2}$$

Clearly $v_{\text{SDP}} \leq v_{\text{QP}}$. Let ρ_{SDP} denote the optimal objective value of (2) but with minimization replaced by maximization. Below we discuss known upper bounds on v_{QP} in terms of v_{SDP} and ρ_{SDP}.

Goemans and Williamson [4] showed that if $m = n$, $A^k = e^k(e^k)^T$, $b^k = 0$, $c^k = -1$ for $k = 1, ..., m$, and $b^0 = 0$, $-A^0$ is positive semidefinite with nonpositive off-diagonals and zero row sums, then

$$v_{\text{QP}} \leq (0.87856...)\ v_{\text{SDP}}. \tag{3}$$

Nesterov [7] showed that if $-A^0$ is allowed to be any positive semidefinite matrix, then

$$v_{\text{QP}} \leq \frac{2}{\pi}\ v_{\text{SDP}}. \tag{4}$$

Ye [17] and Nesterov [8] showed that this still holds if $A^1, ..., A^m$ are further allowed to be diagonal (or mutually commute). Zhang [19] showed that the Goemans-Williamson bound (3) still holds if $A^1, ..., A^m$ are similarly allowed to be diagonal (or mutually commute) and $-A^0$ is allowed to have nonzero row sums. Zhang also showed that if instead $-A^0$ has nonnegative off-diagonals, then $\upsilon_{\rm QP} = \upsilon_{\rm SDP}$ and an optimal solution of (1) can be easily found from an optimal solution of (2).

Of special interest is the case of ellipsoid constraints:

$$A^k = (F^k)^T F^k,\ b^k = 2(F^k)^T g^k,\ c^k = \|g^k\|^2 - h^k,\ k = 1, ..., m, \tag{5}$$

where $F^k \in \Re^{r^k \times n}$, $g^k \in \Re^{r^k}$, $h^k \in \{0,1\}$, $r^k \geq 1$, and $\|\cdot\|$ denotes the Euclidean norm. Then $f^k(x) = \|F^k x + g^k\|^2 - h^k$, $k = 1, ..., m$. Nemirovski, Roos, and Terlaky [6] showed that if in addition the ellipsoids have a common center and nonempty interior (i.e., $g^k = 0$, $h^k = 1$ for all k) and $\sum_{k=1}^m A^k \succ 0$, then a feasible solution x satisfying

$$f^0(x) \leq \frac{1}{2\ \ln(2(m+1)\mu)}\ \upsilon_{\rm SDP},$$

with $\mu := \min\{m+1, \max_{k=1,\ldots,m} \mathrm{Rank} A^k\}$, can be found from the SDP relaxation using a randomization scheme and then derandomizing. Notice that $\sum_{k=1}^m A^k \succ 0$ implies $\max_k \mathrm{Rank} A^k \geq n/m$. Also, if (1) has a ball constraint, then $\mu = \min\{m+1, n\}$. This result was extended by Ye [18] to allow the ellipsoids not to have a common center but assuming $A^0 \preceq 0$, $b^0 = 0$, and the origin is an interior feasible solution. Ye showed that a feasible solution $\tilde{x}$ can be randomly generated such that

$$\mathrm{E}\left[f^0(\tilde{x})\right] \leq \frac{(1 - \max_k \|g^k\|)^2}{4\ \ln(4mn \cdot \max_k r^k)}\ \upsilon_{\rm SDP}.$$

Ye remarked that, for general A^0 and b^0, an additional term depending on $\upsilon_{\rm SDP}$ and $\rho_{\rm SDP}$ appears on both sides. We will show that if, in addition to (5), the origin is a relatively interior feasible solution and (2) has an optimal solution, then a feasible solution x satisfying

$$f^0(x) \leq \frac{(1-\gamma)^2}{(\sqrt{\kappa}+\gamma)^2}\ \upsilon_{\rm SDP}, \tag{6}$$

where $\kappa := \mathrm{Card}\{k \in \{1, ..., m\} : h^k = 1\}$ and $\gamma := \max_{k:h_k=1} \|g^k\|$, can be found using a rank-1 decomposition procedure of Sturm and Zhang [15, Procedure 1]. Thus, in contrast to the bound of Ye, no assumption is made on A^0 or b^0 and (6) does not involve expectation nor n. Also, unlike previous work on SDP relaxation, rank reduction does not involve randomization and the feasible set need not be bounded. In the homogeneous case of $g^k = 0$ and $h^k = 1$ for all k, (6) reduces to $f^0(x) \leq \frac{1}{m}\upsilon_{\rm SDP}$. For $m \leq 11$, this improves on the above bound of Nemirovski et al. For $m = 1$, (1) and (5) correspond to the single trust-region problem and (6) implies that an exact optimal solution can be found by solving the SDP relaxation (2). A similar result was obtained in [15] in a more general context.

The work of Nesterov [7, 8, 11] and Ye [11, 17] showed a more general result than (4), namely, if

$$\mathcal{K} := \{k \in \{1, ..., m\} : A^k \text{ is diagonal and } b^k = 0\} \tag{7}$$

equals $\{1,...,m\}$, then

$$v_{\rm QP} \le \frac{2}{\pi}\, v_{\rm SDP} + \left(1 - \frac{2}{\pi}\right)\rho_{\rm SDP}.$$

Our second result is an extension of the above bound to the general case where $\mathcal{K} \neq \{1,...,m\}$. In particular, we show that an $\tilde{x}$ can be randomly generated to satisfy $f^k(\tilde{x}) \le 0$, $k \in \mathcal{K}$, with probability 1 and

$$\mathrm{E}\left[f^0(\tilde{x})\right] \;\le\; \frac{2}{\pi}\, v_{\rm SDP} + \left(1 - \frac{2}{\pi}\right)\rho^0_{\rm SDP}, \tag{8}$$

$$\mathrm{E}\left[f^\ell(\tilde{x})\right] \;\le\; \left(1 - \frac{2}{\pi}\right)\rho^\ell_{\rm SDP}, \quad \ell \in \{1,...,m\} \setminus \mathcal{K}, \tag{9}$$

where $\rho^\ell_{\rm SDP}$, $\ell \notin \mathcal{K}$, are defined similarly as $\rho_{\rm SDP}$ but with B^0 replaced by B^ℓ and with the inequality constraints not indexed by $\mathcal{K}$ dropped–see (19). An alternative bound that seems generally sharper is also considered. By using a large deviation result, the above bounds holding in expectation can be replaced by bounds holding with high probability–see Section 4. In the case where the constraints not indexed by $\mathcal{K}$ are ellipsoid constraints, we discuss ways to randomly generate feasible solutions that, with high probability, satisfy related bounds on the objective value–see Theorem 4.

Other approximation results for special cases of (1), *not* based on SDP, are discussed in [3, 8, 9, 11, 17]. In particular, for ellipsoid constraints with feasible set having nonempty bounded interior, Fu, Luo, and Ye [3] showed that, for a fixed $\epsilon > 0$ near 0, a feasible solution x with

$$f^0(x) \le \frac{1-\epsilon}{m^2(1+\epsilon)^2}\, v_{\rm QP}$$

can be found by using a column generation method to find an inexact analytic center of the feasible set and then minimizing f^0 over a Dikin ellipsoid centered there. The computational effort depends on $\ln(1/\epsilon)$ and $\ln(1/\delta)$, where δ is the radius of an Euclidean ball contained in the feasible set. The bound (6) improves on the above bound by a factor of $O(m)$ provided that γ is uniformly bounded away from zero. Some results of Nesterov [9], [11, pages 376, 377] suggest that, for simplex-type constraints, approximation techniques not based on SDP relaxation might be preferrable.

Throughout, $\Re^n$ denotes the space of n-dimensional column vectors, $\mathcal{S}^n$ denotes the space of $n \times n$ real symmetric matrices, and ^{T} denotes transpose. For $x \in \Re^n$, x_j denotes jth component of x and $\|x\| = \sqrt{x^T x}$. Also, e^k denotes the kth coordinate vector. For $A \in \Re^{m\times n}$, A_{ij} denotes the (i,j)th entry of A. For $A \in \mathcal{S}^n$ with $|A_{ij}| \le 1$ for all i,j, $\arcsin(A)$ denotes the matrix in $\mathcal{S}^n$ with (i,j)th entry $\arcsin(A_{ij})$. For $A, B \in \mathcal{S}^n$, we denote $\langle A, B\rangle = \sum_{i,j} A_{ij}B_{ij}$ and $A \succeq B$ (respectively, $A \succ B$) means $A - B$ is positive semidefinite (respectively, positive definite). Also, ":=" means "define".

2 SDP Relaxation Bounds: Ellipsoid Constraints Case

In this section, we make in addition to (5) the following assumption.

Assumption 1 *The origin* $0 \in \Re^n$ *is a feasible solution of* (1) *and* $f^k(0) < 0$ *whenever* $h^k = 1$.

Assumption 1 is equivalent to $g^k = 0$ whenever $h^k = 0$ and $\|g^k\| = \sqrt{f^k(0)+1} < 1$ whenever $h^k = 1$. It implies 0 is in the relative interior of the feasible set of (1) but not conversely. To satisfy Assumption 1, it suffices to find a feasible solution of (1) satisfying strictly those constraints with $h^k = 1$ and then translate the origin there. Such a feasible solution can be found efficiently by solving

$$\begin{array}{ll} \min & \max_{k:h^k=1} f^k(x) \\ \text{s.t.} & f^k(x) \le 0 \quad \forall k \text{ with } h_k = 0 \end{array}$$

as a second-order cone programming problem [10, page 221]. Notice that those constraints with $h^k = 0$ are in effect linear constraints. We also make the following assumption.

Assumption 2 (2) *has an optimal solution* X^*.

It can be seen by using (5) that if the feasible set of (1) is bounded, then so is the feasible set of (2) so that Assumption 2 holds. In the footnote below, we show that if $\{u \in \Re^n : u^T A^0 u \le 0, F^1 u = 0, ..., F^m u = 0\} = \{0\}$, then (feasible set of (2)) $\cap \{X : \langle B^0, X\rangle \le 0\}$ is nonempty and bounded so that Assumption 2 again holds.

We show below that a feasible solution x satisfying (6) can be found efficiently from X^*. Our analysis is based on the following rank-1 decomposition result of Sturm and Zhang [15, Proposition 3].

Lemma 1 *Let* $X \in \mathcal{S}^{n+1}$ *be a positive semidefinite matrix of rank* r. *Let* $B \in \mathcal{S}^{n+1}$. *Then,* $\langle B, X\rangle \le 0$ *if and only if there exist* $w_j \in \Re^{n+1}$, $j = 1, ..., r$, *such that*

$$X = \sum_{j=1}^{r} w_j w_j^T \quad \text{and} \quad w_j^T B w_j \le 0, \ j = 1, ..., r.$$

The proof of Lemma 1 is constructive [15, Procedure 1]: Given X and B with $\langle B, X\rangle \le 0$, choose any $w_1, ..., w_r$ satisfying $X = \sum_{j=1}^r w_j w_j^T$. If $w_j^T B w_j > 0$ for some j, then there is some ℓ with $w_\ell^T B w_\ell < 0$ and we swap w_j and w_ℓ with the linear combinations $(w_j + \alpha w_\ell)/\sqrt{1+\alpha^2}$ and $(w_\ell - \alpha w_j)/\sqrt{1+\alpha^2}$, where α solves $(w_j + \alpha w_\ell)^T B (w_j + \alpha w_\ell) = 0$. Each swap increases the number of w_j with $w_j^T B w_j = 0$ by at least 1, so the desired $w_1, ..., w_r$ are found after at most $r - 1$ replacements.

For

$$B := \begin{bmatrix} A^0 & b^0/2 \\ (b^0)^T/2 & -v_{\text{SDP}} \end{bmatrix},$$

we have from $X^*_{n+1n+1} = 1$ and $c^0 = 0$ that $\langle B, X^*\rangle = \langle B^0, X^*\rangle - v_{\text{SDP}} = 0$. Applying Lemma 1 to X^* and B, we can find $w_j = (u_j, t_j) \in \Re^n \times \Re$, $j = 1, ..., n+1$, such that

$$X^* = \sum_{j=1}^{n+1} w_j w_j^T \quad \text{and} \quad w_j^T B w_j \le 0, \ j = 1, ..., n+1.$$

Since X^* is a feasible solution of (2), this and (5) yield

$$u_j^T A^0 u_j + t_j (b^0)^T u_j \leq v_{\text{SDP}} t_j^2, \quad j = 1, ..., n+1, \tag{10}$$

$$\sum_{j=1}^{n+1} \left(u_j^T A^k u_j + t_j (b^k)^T u_j + t_j^2 c^k \right) = \langle B^k, X^* \rangle \leq 0, \quad k = 1, ..., m, \tag{11}$$

$$\sum_{j=1}^{n+1} t_j^2 = X^*_{n+1 n+1} = 1. \tag{12}$$

Using (5), we obtain from (11) and (12) that

$$\sum_{j=1}^{n+1} \|F^k u_j + t_j g^k\|^2 \leq h^k, \quad k = 1, ..., m. \tag{13}$$

Notice that the above results can be generalized to any feasible solution of (2).[3] If $h^k = 0$, then $g^k = 0$ so (13) yields $\|F^k u_j\|^2 = 0$ for all j. Also, summing (13) over all k with $h^k = 1$ yields

$$\sum_{j=1}^{n+1} \sum_{k:h^k=1} \|F^k u_j + t_j g^k\|^2 \leq \kappa, \tag{14}$$

where $\kappa := \text{Card}\{k : h^k = 1\}$. We need the following fact.

Lemma 2 *For any scalars $\kappa \geq 0$, $\alpha_j \geq 0$ and $\beta_j \geq 0$, $j = 1, ..., \ell$ ($\ell \geq 1$), such that $\sum_{j=1}^{\ell} \alpha_j \leq \kappa$ and $\sum_{j=1}^{\ell} \beta_j = 1$, there exists $\bar{j} \in \{1, ..., \ell\}$ such that $\beta_{\bar{j}} > 0$ and $\alpha_{\bar{j}} / \beta_{\bar{j}} \leq \kappa$.*

Proof. If the assertion is false, then for every $j \in \{1, ..., \ell\}$ such that $\beta_j > 0$ we would have $\alpha_j / \beta_j > \kappa$ or, equivalently, $\alpha_j > \kappa \beta_j$. Then we would have

$$\kappa \geq \sum_{j=1}^{\ell} \alpha_j \geq \sum_{j:\beta_j > 0} \alpha_j > \sum_{j:\beta_j > 0} \kappa \beta_j = \kappa,$$

a clear contradiction. ∎

By (12) and (14), we can apply Lemma 2 to $\alpha_j = \sum_{k:h^k=1} \|F^k u_j + t_j g^k\|^2$ and $\beta_j = t_j^2$ to conclude the existence of $\bar{j} \in \{1, ..., n+1\}$ such that

$$t_{\bar{j}}^2 > 0 \quad \text{and} \quad \sum_{k:h^k=1} \|F^k u_{\bar{j}} + t_{\bar{j}} g^k\|^2 / t_{\bar{j}}^2 \leq \kappa.$$

[3] In particular, Assumption 1 implies $X = e^{n+1}(e^{n+1})^T$ is a feasible solution of (2) with $\langle B^0, X \rangle = 0$. Thus, $\mathcal{X}^0 :=$ (feasible set of (2)) $\cap \{X : \langle B^0, X \rangle \leq 0\}$ is nonempty. Then for any $X \in \mathcal{X}^0$, repeating the above argument with X and 0 in place of X^* and v_{SDP} yields $X = \sum_{j=1}^{n+1} w_j w_j^T$ for some $w_j = (u_j, t_j) \in \Re^n \times \Re$ satisfying $u_j^T A^0 u_j + t_j (b^0)^T u_j \leq 0$, (12), and (13). For each j, (12) implies t_j is bounded while (13) implies $\|F^k u_j + t_j g^k\|^2 \leq h^k$. If u_j is unbounded for some j, then dividing by $\|u_j\|^2$ and taking limit yields a cluster point u of $u_j / \|u_j\|$ satisfying $\|u\| = 1$, $u^T A^0 u \leq 0$ and $\|F^k u\|^2 \leq 0$, $k = 1, ..., m$. Thus, if $\{u \in \Re^n : u^T A^0 u \leq 0, F^1 u = 0, ..., F^m u = 0\} = \{0\}$, then $\mathcal{X}^0$ is bounded.

In particular, we can choose $\bar{j}$ to minimize the ratio α_j/β_j over all j with $\beta_j > 0$. Thus

$$\|F^k u_{\bar{j}}/t_{\bar{j}} + g^k\| \le \sqrt{\kappa} \quad \text{whenever } h^k = 1. \tag{15}$$

Let

$$\begin{aligned} \bar{x} &:= \begin{cases} u_{\bar{j}}/t_{\bar{j}} & \text{if } (b^0)^T u_{\bar{j}}/t_{\bar{j}} \le 0; \\ -u_{\bar{j}}/t_{\bar{j}} & \text{otherwise}, \end{cases} \\ \bar{\tau} &:= \max\{\tau \in [0,1] : f^k(\tau\bar{x}) \le 0,\ k = 1, ..., m\}. \end{aligned}$$

Using (10) and (15), we prove below the following result.

Theorem 1 *Under Assumptions 1, 2 and* (5)*, the above construction yields a feasible solution* $x = \bar{\tau}\bar{x}$ *of* (1) *satisfying* (6)*, where* $\kappa := \operatorname{Card}\{k \in \{1, ..., m\} : h^k = 1\}$ *and* $\gamma := \max_{k:h_k=1} \|g^k\|$.

Proof. We estimate $\bar{\tau}$. Fix any $k \in \{1, ..., m\}$. Suppose $h^k = 0$. Then we have from $\|F^k u_{\bar{j}}\|^2 = 0$ that $f^k(\tau\bar{x}) = 0$ for all $\tau \in [0,1]$. Suppose $h^k = 1$. Then we see from (15) that if $(b^0)^T u_{\bar{j}}/t_{\bar{j}} \le 0$, then $\|F^k\bar{x} + g^k\| \le \sqrt{\kappa}$; and otherwise

$$\|F^k\bar{x} + g^k\| = \| - (F^k u_{\bar{j}}/t_{\bar{j}} + g^k) + 2g^k\| \le \|F^k u_{\bar{j}}/t_{\bar{j}} + g^k\| + 2\|g^k\| \le \sqrt{\kappa} + 2\|g^k\|.$$

Thus for any $\tau \in [0,1]$ we have

$$\|F^k(\tau\bar{x}) + g^k\| = \|\tau(F^k\bar{x} + g^k) + (1-\tau)g^k\| \le \tau(\sqrt{\kappa} + 2\|g^k\|) + (1-\tau)\|g^k\|.$$

Using $\|g^k\| < 1$, the right-hand side is below 1 (i.e., $f^k(\tau\bar{x}) \le 0$) whenever $\tau \le (1 - \|g^k\|)/(\sqrt{\kappa} + \|g^k\|)$. Thus,

$$\bar{\tau} \ge \min_{k:h_k=1} \frac{1 - \|g^k\|}{\sqrt{\kappa} + \|g^k\|} = \frac{1 - \max_{k:h_k=1}\|g^k\|}{\sqrt{\kappa} + \max_{k:h_k=1}\|g^k\|}, \tag{16}$$

where the equality follows from $(1-\gamma)/(\sqrt{\kappa}+\gamma)$ being a decreasing function of $\gamma \in [0,1)$. Notice that $\bar{\tau}$ can be easily computed by solving the quadratic equation $\|\tau F^k\bar{x} + g^k\|^2 = 1$ in τ for each k such that $\|F^k\bar{x} + g^k\|^2 > 1$ and then taking the minimum of all the positive roots found.

Finally, our choice of $\bar{x}$ implies $(b^0)^T\bar{x} \le 0$ and $(b^0)^T\bar{x} \le (b^0)^T u_{\bar{j}}/t_{\bar{j}}$. Then for any $\tau \in [0,1]$ we have $\tau \ge \tau^2$ and hence

$$\begin{aligned} f^0(\tau\bar{x}) &= \tau^2\bar{x}^T A^0\bar{x} + \tau(b^0)^T\bar{x} \\ &\le \tau^2\bar{x}^T A^0\bar{x} + \tau^2(b^0)^T\bar{x} \\ &\le \tau^2\bar{x}^T A^0\bar{x} + \tau^2(b^0)^T u_{\bar{j}}/t_{\bar{j}} \\ &= \tau^2\left(u_{\bar{j}}^T A^0 u_{\bar{j}} + t_{\bar{j}}(b^0)^T u_{\bar{j}}\right)/t_{\bar{j}}^2 \\ &\le \tau^2 \upsilon_{\text{SDP}}, \end{aligned}$$

where the last inequality uses (10). Since 0 is a feasible solution of (1) so that $v_{\mathrm{SDP}} \leq v_{\mathrm{QP}} \leq f(0) = 0$, setting $\tau = \bar{\tau}$ in the above inequality and using (16) completes the proof. ∎

In the above construction, the main effort lies in solving the SDP relaxation (2), for which many efficient methods exist. Given that an exact optimal solution of (1) is constructed when $m = 1$, we may speculate that when $m = 2$, which is also of considerable interest for trust-region methods (see [2, 12] and references therein), a good approximate solution will generally be found. It is worthwhile to test this numerically. In particular, if the rank-1 decomposition given by Lemma 1 is not unique, can we choose one so that the corresponding $\bar{x}$ minimizes $\min_{\tau \in [0,\bar{\tau}]} f^0(\tau \bar{x})$?

3 SDP Relaxation Bounds: General Case

In this section, following Goemans and Williamson, Nesterov, and Ye, we derive approximation bounds for (1) based on the SDP relaxation (2) under Assumption 2. Notice that Assumption 2 does not guarantee feasibility of (1), which is NP-hard to check in general. Since $X^* \succeq 0$, we can express

$$X^* = V^T V = [v_i^T v_j]_{i,j=1}^{n+1},$$

for some $V \in \Re^{n+1 \times n+1}$. Here v_i denotes the ith column of V. Choose randomly (according to uniform distribution) v on the unit sphere in $\Re^{n+1}$. Since $\|v_{n+1}\|^2 = X^*_{n+1n+1} = 1$, v_{n+1} also lies on this unit sphere. If $v^T v_{n+1} \leq 0$, then set for $i = 1, ..., n+1$,

$$\tilde{x}_i = \begin{cases} \sqrt{X^*_{ii}} & \text{if } v^T v_i \leq 0 \\ -\sqrt{X^*_{ii}} & \text{else} \end{cases}.$$

If $v^T v_{n+1} > 0$, then set for $i = 1, ..., n+1$,

$$\tilde{x}_i = \begin{cases} -\sqrt{X^*_{ii}} & \text{if } v^T v_i \leq 0 \\ \sqrt{X^*_{ii}} & \text{else} \end{cases}.$$

The above choice and $X^*_{n+1n+1} = 1$ ensure that $\tilde{x}_{n+1} = 1$ always.

For each i, j we have that $|\tilde{x}_i \tilde{x}_j| = \sqrt{X^*_{ii} X^*_{jj}}$. If $X^*_{ii} X^*_{jj} \neq 0$, then $\tilde{x}_i \tilde{x}_j = \sqrt{X^*_{ii} X^*_{jj}}$ if and only if $v^T v_i \leq 0$, $v^T v_j \leq 0$ or $v^T v_i > 0$, $v^T v_j > 0$. As was shown by Goemans and Williamson [4], the probability that this event occurs is

$$p = 1 - \frac{1}{\pi} \arccos(v_i^T v_j / \|v_i\| \|v_j\|) = 1 - \frac{1}{\pi} \arccos(X^*_{ij} / \sqrt{X^*_{ii} X^*_{jj}}).$$

Thus, the expectation of $\tilde{x}_i \tilde{x}_j$ is

$$\mathrm{E}[\tilde{x}_i \tilde{x}_j] \quad = \quad \sqrt{X^*_{ii} X^*_{jj}}\, p + (-\sqrt{X^*_{ii} X^*_{jj}})(1 - p)$$

$$
\begin{aligned}
&= \frac{2}{\pi}\sqrt{X^*_{ii}X^*_{jj}}\left(\frac{\pi}{2} - \arccos(X^*_{ij}/\sqrt{X^*_{ii}X^*_{jj}})\right) \\
&= \frac{2}{\pi}\sqrt{X^*_{ii}X^*_{jj}}\arcsin(X^*_{ij}/\sqrt{X^*_{ii}X^*_{jj}}). \qquad (17)
\end{aligned}
$$

If $X^*_{ii}X^*_{jj} = 0$, then $X^*_{ij} = 0$ since $X^* \succeq 0$, so (17) still holds with the convention that $0/0 = 0$.

Thus, for $k = 0, 1, ..., m$, since $\tilde{x}_{n+1} = 1$ always, (17) yields

$$
\begin{aligned}
\mathrm{E}\left[f^k(\tilde{x})\right] &= \sum_{i,j=1}^{n+1} B^k_{ij}\mathrm{E}[\tilde{x}_i\tilde{x}_j] \\
&= \sum_{i,j=1}^{n+1} B^k_{ij}\frac{2}{\pi}\sqrt{X^*_{ii}X^*_{jj}}\arcsin(X^*_{ij}/\sqrt{X^*_{ii}X^*_{jj}}) \\
&= \frac{2}{\pi}\langle B^k, D\arcsin(D^{-1}X^*D^{-1})D\rangle, \qquad (18)
\end{aligned}
$$

where $D = \mathrm{diag}[\sqrt{X^*_{ii}}]_{i=1}^{n+1}$. Notice that $|(D^{-1}X^*D^{-1})_{ij}| \leq 1$ for all i, j, so $\arcsin(D^{-1}X^*D^{-1})$ is defined. Since $\tilde{x}_i^2 = X^*_{ii}$ always for all i, we have $f^k(\tilde{x}) = \langle B^k, X^*\rangle \leq 0$ always for $k \in \mathcal{K}$ (see (7)).

We now derive bounds on $\mathrm{E}\left[f^k(\tilde{x})\right]$, $k \notin \mathcal{K}$, by using (18) and extending an analysis of Nesterov and Ye. We will make in addition to Assumption 2 the following assumption.

Assumption 3 *$\{x \in \Re^n : f^k(x) \leq 0,\ k \in \mathcal{K}\}$ is bounded.*

Consider for each $\ell \notin \mathcal{K}$ the following SDP problem:

$$
\begin{aligned}
\rho^\ell_{\mathrm{SDP}} := \max\ & \langle B^\ell, X\rangle \\
\text{s.t.}\ & \langle B^k, X\rangle = c^k_*, \quad k \in \mathcal{K}, \\
& \langle B^{m+1}, X\rangle = 1, \quad X \succeq 0,
\end{aligned} \qquad (19)
$$

where $B^{m+1} := e^{n+1}(e^{n+1})^T$ and $c^k_* := \langle B^k, X^*\rangle \leq 0$. For $\ell \neq 0$, ρ^ℓ_{SDP} measures how much the ℓth inequality in (2) is violated by the feasible solutions of the inequalities indexed by $\mathcal{K}$. Here we use the tigher constraints $\langle B^k, X\rangle = c^k_*$ instead of $\langle B^k, X\rangle \leq 0$ used by Nesterov and Ye [11]. This yields a tigher upper bound.

Let $\mathcal{X}$ denote the feasible set of (19). Since $X^* \in \mathcal{X}$, $\mathcal{X}$ is nonempty. By Assumption 3, the diagonal entries of $X \in \mathcal{X}$ are bounded which, together with $X \succeq 0$, implies $\mathcal{X}$ is bounded. By a result of Rockafellar [14, Theorem 30.4(i)], strong duality holds between (19) and its dual:

$$
\begin{aligned}
\rho^\ell_{\mathrm{SDP}} = \inf\ & \textstyle\sum_{k\in\mathcal{K}} c^k_* y^k + y^{m+1} \\
\text{s.t.}\ & -B^\ell + \textstyle\sum_{k\in\mathcal{K}\cup\{m+1\}} B^k y^k \succeq 0.
\end{aligned} \qquad (20)
$$

In general, the infimum in (20) need not be attained. As in [7, 11, 17], we make use of the following result of Nesterov [7].

Lemma 3 *For any $Y \succeq 0$ with $Y_{ii} \leq 1$ for all i, we have* $\arcsin(Y) \succeq Y$.

Fix any $\ell \notin \mathcal{K}$. For each $\epsilon > 0$, let $(y^k)_{k\in\mathcal{K}\cup\{m+1\}}$ be any feasible solution of the dual problem (20) such that $\sum_{k\in\mathcal{K}} c_*^k y^k + y^{m+1} \leq \rho^\ell_{\text{SDP}} + \epsilon$. Let $D := \text{diag}[\sqrt{X^*_{ii}}]_{i=1}^{n+1}$ and $Y := D^{-1}X^*D^{-1}$, with the convention that $Y_{ii} = 1$ if $X^*_{ii} = 0$ and $Y_{ij} = 0$ if $X^*_{ii}X^*_{jj} = 0$ and $i \neq j$. Since $X^* \succeq 0$, then $Y \succeq 0$ and $Y_{ii} = 1$, $i = 1, ..., n+1$. Thus

$$
\begin{aligned}
& \langle B^\ell, D\arcsin(Y)D\rangle \\
= \; & \langle B^\ell - \sum_{k\in\mathcal{K}\cup\{m+1\}} B^k y^k, D\arcsin(Y)D\rangle + \sum_{k\in\mathcal{K}\cup\{m+1\}} y^k \langle B^k, D\arcsin(Y)D\rangle \\
\leq \; & \langle B^\ell - \sum_{k\in\mathcal{K}\cup\{m+1\}} B^k y^k, DYD\rangle + \sum_{k\in\mathcal{K}\cup\{m+1\}} y^k \langle B^k, D\arcsin(Y)D\rangle \\
= \; & \langle B^\ell, X^*\rangle + \left(\frac{\pi}{2} - 1\right) \sum_{k\in\mathcal{K}\cup\{m+1\}} y^k \langle B^k, X^*\rangle \\
= \; & \langle B^\ell, X^*\rangle + \left(\frac{\pi}{2} - 1\right) \left(\sum_{k\in\mathcal{K}} y^k c_*^k + y^{m+1}\right) \\
\leq \; & \langle B^\ell, X^*\rangle + \left(\frac{\pi}{2} - 1\right) (\rho^\ell_{\text{SDP}} + \epsilon),
\end{aligned}
\tag{21}
$$

where the first inequality uses dual feasibility, Lemma 3 and the fact that $\langle W, Z\rangle \geq 0$ whenever $W \succeq 0, Z \succeq 0$; the second equality uses $DYD = X^*$ and the observations that B^k is diagonal for $k \in \mathcal{K} \cup \{m+1\}$ and $D\arcsin(Y)D$ has diagonal entries $\frac{\pi}{2}X^*_{ii}$ for all i. Since (21) holds for every $\epsilon > 0$, taking the limit as $\epsilon \to 0$ and using (18) yields

$$
\begin{aligned}
\text{E}\left[f^\ell(\tilde{x})\right] &= \frac{2}{\pi}\langle B^\ell, D\arcsin(D^{-1}X^*D^{-1})D\rangle \\
&= \frac{2}{\pi}\langle B^\ell, D\arcsin(Y)D\rangle \\
&\leq \frac{2}{\pi}\langle B^\ell, X^*\rangle + \left(1 - \frac{2}{\pi}\right)\rho^\ell_{\text{SDP}}.
\end{aligned}
$$

Since X^* is an optimal solution of (2), this establishes the following result.

Theorem 2 *Under Assumptions 2 and 3, the bounds* (8) *and* (9) *hold.*

In the case where $\mathcal{K} = \{1, ..., m\}$, the above bounds slightly refine analogous bounds obtained by Nesterov [7, Theorem 3.3], [11, Theorem 13.2.1] and Ye [17, Theorem 2], [11, Theorem 13.3.2, part 2]. As was considered by Nesterov [11] (also see [19]), the quadratic inequalies $f^k(x) \in 0$, $k \in \mathcal{K}$, can be generalized to constraints of the form $[x_i^2]_{i=1}^n \in \mathcal{F}$, where $\mathcal{F}$ is a compact convex set intersecting the positive orthant. In this general case, however, the corresponding relaxation may no longer be an SDP problem.

We can also obtain lower bounds analogous to those obtained in the above references. Consider for each $\ell \notin \mathcal{K}$ the following QP:

$$\begin{array}{rll} v^\ell_{\rm QP} := & \min & f^\ell(x) \\ & \text{s.t.} & f^k(x) \le 0, \ k \in \mathcal{K}. \end{array} \tag{22}$$

Since $\mathcal{X}$ is nonempty and bounded, then so is the feasible set of this QP. Thus $v^\ell_{\rm QP}$ is finite. By the definition of $\mathcal{K}$, we can apply [11, Theorem 13.3.1] or [19, Theorem 1] to reformulate this QP as an equivalent nonlinear program for which X^* and Y, D defined above form a feasible solution with objective function value $\frac{2}{\pi}\langle B^\ell, D\arcsin(Y)D\rangle$. Thus,

$$v^\ell_{\rm QP} \le \frac{2}{\pi}\langle B^\ell, D\arcsin(Y)D\rangle = \mathrm{E}\left[f^\ell(\tilde{x})\right]. \tag{23}$$

Notice that $v^0_{\rm QP} \le v_{\rm QP}$, with equality holding when $\mathcal{K} = \{1, ..., m\}$. If constraints not indexed by $\mathcal{K}$ are ellipsoid constraints, an upper bound on $v_{\rm QP}$ in terms of $v^0_{\rm QP}$ will be derived in Section 4.

We can more generally replace $\mathcal{K}$ in (20) by any $\mathcal{K}' \subset \{1, ..., m\}$ containing $\mathcal{K}$. This would yield a lower $\rho^\ell_{\rm SDP}$, but then the right-hand side of (21) would have an additional term of the form $\sum_{k\in\mathcal{K}'\setminus\mathcal{K}} y^k\langle B^k, D\arcsin(D^{-1}X^*D^{-1})D - \frac{\pi}{2}X^*\rangle$. By Lemma 4 below, this term is at most

$$\left(\frac{\pi}{2} - 1\right)\sum_{k\in\mathcal{K}'\setminus\mathcal{K}} y^k \sum_{i\ne j}|B_{ij}X^*_{ij}|.$$

Thus, the resulting bound would involve a dual solution y^k, $k \in \mathcal{K}\setminus\mathcal{K}$, as well as X^*.

Below we consider alternative bounds that complement the bounds from Theorem 2. Since $\arcsin(t)$ is convex for $t \in (0, 1]$ and has derivative 1 at $t = 0$, we have that $1 \le \arcsin(t)/t \le \arcsin(1)/1 = \frac{\pi}{2}$. By symmetry, this holds for $t \in [-1, 0)$ as well, so that

$$1 \le \frac{\arcsin(t)}{t} \le \frac{\pi}{2} \quad \forall t \in [-1, 0) \cup (0, 1]. \tag{24}$$

By using (24), the following lemma readily follows.

Lemma 4 *For any $X \succeq 0$ and $B \in \mathcal{S}^{n+1}$, we have*

$$\left|\frac{2}{\pi}\langle B, D\arcsin(D^{-1}XD^{-1})D\rangle - \langle B, X\rangle\right| \le \left(1 - \frac{2}{\pi}\right)\sum_{i\ne j}|B_{ij}X_{ij}|,$$

where $D = \mathrm{diag}[\sqrt{X_{ii}}]_{i=1}^{n+1}$.

By using (18) and Lemma 4, we obtain

$$\mathrm{E}\left[f^\ell(\tilde{x})\right] \le \langle B^\ell, X^*\rangle + \left(1 - \frac{2}{\pi}\right)\sum_{i\ne j}|B^\ell_{ij}X^*_{ij}|, \quad \ell = 0, 1, ..., m.$$

Since X^* is an optimal solution of (2), the above inequalities yield the following bounds.

Theorem 3 *Under Assumption 2,*

$$\begin{aligned}\mathrm{E}\left[f^0(\tilde{x})\right] &\leq \upsilon_{\text{SDP}} + \left(1 - \frac{2}{\pi}\right)\delta^0, \\ \mathrm{E}\left[f^\ell(\tilde{x})\right] &\leq \left(1 - \frac{2}{\pi}\right)\delta^\ell, \quad \ell \in \{1, ..., m\} \setminus \mathcal{K},\end{aligned}$$

where $\delta^\ell := \sum_{i \neq j} |B^\ell_{ij} X^*_{ij}|$.

The bounds in Theorem 3 depend on X^* as well as the off-diagonal quadratic coefficients A^k_{ij}, $i \neq j$, and the linear coefficients b^k_i. While these bounds might look less attractive than the bounds in Theorem 2, they were found to be sharper in all the examples this author tried. For example, if

$$m = n = 2, \quad f^0(x) = x_1x_2 + x_1 + x_2, \quad f^1(x) = x_1^2 - 1, \quad f^2(x) = x_2^2 - 1,$$

then $\mathcal{K} = \{1, 2\}$ and it is straightforward to verify that

$$\upsilon_{\text{QP}} = -1, \quad X^* = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}, \quad \upsilon_{\text{SDP}} = -\frac{3}{2}, \quad \rho^0_{\text{SDP}} = 3, \quad \delta^0 = \frac{3}{2}.$$

Here δ^0 is smaller than $\rho^0_{\text{SDP}} - \upsilon_{\text{SDP}}$ by a factor of 3!

4 Generating Approximate Solutions: General Case

The results of Section 3 show that $\tilde{x}$ is an approximate solution of (1) in expectation only. In this section we refine this result to generate approximate solutions with high probability.

The following lemma, attributed to Bernstein, refines the Chebychev inequality for bounded random variables. Its proof can be inferred from the argument in [13, pages 385-386]. A similar result was used in [6]. We note that the probabilistic analysis of Nesterov [7, page 159] is not applicable here since υ_{SDP} need not be below υ^0_{QP} in Theorem 2, except in the case of $\mathcal{K} = \{1, ..., m\}$.

Lemma 5 *Let ξ be a random variable with standard deviation σ. Suppose $\sigma \leq C$ and $|\xi - \mathrm{E}[\xi]| \leq K$ always for some constants C and K. Then, for any $t \in (0, C/K]$,*

$$\text{Prob}\left[\xi - \mathrm{E}[\xi] \geq \frac{3}{2} tC\right] \leq e^{-t^2/2}.$$

Fix any tolerance $\epsilon > 0$. For each $k \notin \mathcal{K}$, if $f^k(\tilde{x}) - \mathrm{E}[f^k(\tilde{x})] > \epsilon$, then since $\tilde{x}_{n+1} = 1$ always we have

$$\sum_{i,j=1}^{n+1} B^k_{ij}(\tilde{x}_i\tilde{x}_j - \mathrm{E}[\tilde{x}_i\tilde{x}_j]) > \epsilon. \tag{25}$$

Let

$$\Delta^k := \sum_{i \neq j} |B^k_{ij}| \sqrt{X^*_{ii} X^*_{jj}}.$$

Also, let $\lambda^k_{ij} := |B^k_{ij}|\sqrt{X^*_{ii}X^*_{jj}}/\Delta^k$ for $i \neq j$. Then $\sum_{i\neq j} \lambda^k_{ij} = 1$. Since $\tilde{x}_i^2 = \mathrm{E}[\tilde{x}_i^2]$ for all i, (25) implies that $B^k_{ij}(\tilde{x}_i\tilde{x}_j - \mathrm{E}[\tilde{x}_i\tilde{x}_j]) > \lambda^k_{ij}\epsilon$ for some $i \neq j$. The variance σ^2 of $\tilde{x}_i\tilde{x}_j$ can be bounded above as follows:

$$\sigma^2 = \mathrm{E}[(\tilde{x}_i\tilde{x}_j)^2] - \mathrm{E}[\tilde{x}_i\tilde{x}_j]^2 \leq \mathrm{E}[(\tilde{x}_i\tilde{x}_j)^2] = X^*_{ii}X^*_{jj}.$$

Also, $|\tilde{x}_i\tilde{x}_j| = \sqrt{X^*_{ii}X^*_{jj}}$ and, by (17) and $|t| \leq |\arcsin(t)| \leq \frac{\pi}{2}|t|$ (see (24)), we have $|\mathrm{E}[\tilde{x}_i\tilde{x}_j]| \leq \sqrt{X^*_{ii}X^*_{jj}}$. Thus

$$|\tilde{x}_i\tilde{x}_j - \mathrm{E}[\tilde{x}_i\tilde{x}_j]| \leq 2\sqrt{X^*_{ii}X^*_{jj}}.$$

Then, applying Lemma 5 with $\xi = (B^k_{ij}/\lambda^k_{ij})\tilde{x}_i\tilde{x}_j$, $C = \max_{k\notin\mathcal{K}} \Delta^k$, $K = 2C$, and $t = \frac{2}{3}\epsilon/C$, we obtain that

$$\mathrm{Prob}\left[(B^k_{ij}/\lambda^k_{ij})(\tilde{x}_i\tilde{x}_j - \mathrm{E}[\tilde{x}_i\tilde{x}_j]) \geq \epsilon\right] \leq e^{-\frac{2}{9}\epsilon^2/C^2}$$

provided that $\epsilon \leq \frac{3}{4}C$. Thus, provided that $\epsilon \leq \frac{3}{4}C$, we obtain

$$\mathrm{Prob}\left[\max_{k\notin\mathcal{K}}\{f^k(\tilde{x}) - \mathrm{E}[f^k(\tilde{x})]\} > \epsilon\right] \leq m_0 e^{-\frac{2}{9}\epsilon^2/C^2},$$

where $m_0 := m + 1 - \mathrm{Card}\mathcal{K}$. For each $k \in \mathcal{K}$, we have $f^k(\tilde{x}) \leq 0$ with probability 1.

Thus, if we generate $\tilde{x}$ randomly and independently L times, the probability that one of these L samples satisfies

$$f^k(\tilde{x}) \leq \mathrm{E}[f^k(\tilde{x})] + \epsilon, \quad k = 0, 1, ..., m, \tag{26}$$

is at least $1 - \left(m_0 e^{-\frac{2}{9}\epsilon^2/C^2}\right)^L$. If $m_0 = 1$ and we choose L to be, say, $50 \cdot C^2/\epsilon^2$, then this probability is about 0.999985. However, $\tilde{x}$ is not likely to be a feasible solution of (1). Notice that since $X^* \succeq 0$ so that $|X^*_{ij}| \leq \sqrt{X^*_{ii}X^*_{jj}} = |\tilde{x}_i\tilde{x}_j|$ and $X^*_{n+1n+1} = \tilde{x}_{n+1} = 1$, we have $\delta^k \leq \Delta^k = \sum_{i\neq j} |A^\ell_{ij}\tilde{x}_i\tilde{x}_j| + \sum_i |b^\ell_i\tilde{x}_i|$.

To construct feasible solutions with probability 1, we consider the special case where the constraints not indexed by $\mathcal{K}$ are ellipsoid constraints, i.e.,

$$f^k(x) = \|F^k x + g^k\|^2 - 1 \quad \forall k \in \{1, ..., m\} \setminus \mathcal{K}, \tag{27}$$

where $F^k \in \Re^{r^k \times n}$, $g^k \in \Re^{r^k}$, $r^k \geq 1$. We also assume that the origin is a feasible solution of (1) satisfying strictly those constraints not indexed by $\mathcal{K}$. This is equivalent to

$$c^k \leq 0 \quad \forall k \in \mathcal{K} \qquad \text{and} \qquad \|g^k\| < 1 \quad \forall k \notin \mathcal{K}. \tag{28}$$

Then, by moving $\tilde{x}$ sufficiently close toward the origin as was done in Section 2, we will construct feasible solutions with certainty. We give more details below.

Suppose $\tilde{x}$ satisfies (26). Let

$$\begin{aligned} \bar{x} &:= \begin{cases} \tilde{x} & \text{if } (b^0)^T\tilde{x} \le 0; \\ -\tilde{x} & \text{otherwise,} \end{cases} \\ \bar{\tau} &:= \max\{\tau \in [0,1] : f^k(\tau\bar{x}) \le 0,\ k = 1, ..., m\}, \\ \check{\tau} &:= \arg\min\{f^0(\tau\bar{x}) : \tau \in [0,\bar{\tau}]\}. \end{aligned}$$

Notice that $\bar{\tau}$ and $\check{\tau}$ can be easily computed. By using (26)–(28), we obtain the following main result.

Theorem 4 *Under Assumption 2 and* (27), (28), *for any* $\epsilon > 0$ *and integer* $L \ge 1$ *and any* $\eta^k \ge \mathrm{E}[f^k(\tilde{x})]$ *for* $k \in \{1, ..., m\} \setminus \mathcal{K}$, *if we generate* $\tilde{x}$ *randomly and independently* L *times as described in Section 3 and construct* $x = \check{\tau}\bar{x}$ *as above, then* x *is a feasible solution of* (1) *with probability 1 and satisfies*

$$f^0(x) \le \min_{k \notin \mathcal{K}} \left(\frac{1 - \|g^k\|}{\sqrt{1 + \eta^k + \epsilon} + \|g^k\|} \right)^2 \left(\mathrm{E}[f^0(\tilde{x})] + \epsilon \right), \tag{29}$$

with probability of at least $1 - \left(m_0 e^{-\frac{2}{9}\epsilon^2/C^2} \right)^L$, *where* $C := \max_{k \notin \mathcal{K}} \Delta^k$ *and* $m_0 := m + 1 - \mathrm{Card}\mathcal{K}$.

Proof. For each $k \in \mathcal{K}$, since A^k is diagonal and $b^k = 0$, we see that

$$f^k(\tau\bar{x}) = f^k(\tau\tilde{x}) = \tau^2 f^k(\tilde{x}) + (1 - \tau^2)c^k \le 0$$

for all $\tau \in [0,1]$. For each $k \in \{1, ..., m\} \setminus \mathcal{K}$, we see from (26) and (27) that if $(b^0)^T\tilde{x} \le 0$, then $\|F^k\bar{x} + g^k\| \le \sqrt{\kappa^k}$; and otherwise

$$\|F^k\bar{x} + g^k\| = \| - (F^k\tilde{x} + g^k) + 2g^k\| \le \|F^k\tilde{x} + g^k\| + 2\|g^k\| \le \sqrt{\kappa^k} + 2\|g^k\|,$$

where $\kappa^k := 1 + \mathrm{E}[f^k(\tilde{x})] + \epsilon$. Thus, arguing identically as in the proof of Theorem 1, we obtain that

$$\bar{\tau} \ge \min_{k \notin \mathcal{K}} \frac{1 - \|g^k\|}{\sqrt{\kappa^k} + \|g^k\|}. \tag{30}$$

Moreover, for all $\tau \in [0, \bar{\tau}]$, $\tau\bar{x}$ is a feasible solution of (1) with probability 1. Since $\check{\tau} \in [0, \bar{\tau}]$, then $x = \check{\tau}\bar{x}$ is a feasible solution of (1) with probability 1.

For each $k \in \{1, ..., m\} \setminus \mathcal{K}$, since $\mathrm{E}[f^k(\tilde{x})] \le \eta^k$, then $\kappa^k \le 1 + \eta^k + \epsilon$ and it follows from (30) that

$$\bar{\tau} \ \ge \ \hat{\tau} \ := \min_{k \notin \mathcal{K}} \frac{1 - \|g^k\|}{\sqrt{1 + \eta^k + \epsilon} + \|g^k\|} > 0,$$

implying $\check{\tau} \in (0, \bar{\tau}]$. Finally, our choice of $\bar{x}$ implies $(b^0)^T\bar{x} \leq 0$ and $(b^0)^T\bar{x} \leq (b^0)^T\tilde{x}$. Then, arguing similarly as in the proof of Theorem 1, we obtain for any $\tau \in [0, \bar{\tau}]$ that

$$\begin{aligned} f^0(\check{\tau}\bar{x}) &\leq f^0(\tau\bar{x}) \\ &\leq \tau^2\left(\tilde{x}^T A^0 \tilde{x} + (b^0)^T\tilde{x}\right) \\ &\leq \tau^2\left(\mathrm{E}[f^0(\tilde{x})] + \epsilon\right), \end{aligned}$$

where the last inequality uses (26). Setting $\tau = \hat{\tau}$ completes the proof. ∎

By Theorem 3, we can choose $\eta^k = \left(1 - \frac{2}{\pi}\right)\delta^k$ in Theorem 4. Then (29) becomes

$$f^0(x) \leq \min_{k \notin \mathcal{K}} \left(\frac{1 - \|g^k\|}{\sqrt{1 + \left(1 - \frac{2}{\pi}\right)\delta^k + \epsilon} + \|g^k\|}\right)^2 \left(v_{\text{SDP}} + \left(1 - \frac{2}{\pi}\right)\delta^0 + \epsilon\right),$$

If Assumption 3 also holds, then by Theorem 2, we can choose $\eta^k = \left(1 - \frac{2}{\pi}\right)\rho^k_{\text{SDP}}$ in Theorem 4. Then (29) becomes

$$f^0(x) \leq \min_{k \notin \mathcal{K}} \left(\frac{1 - \|g^k\|}{\sqrt{1 + \left(1 - \frac{2}{\pi}\right)\rho^k_{\text{SDP}} + \epsilon} + \|g^k\|}\right)^2 \left(\frac{2}{\pi}\, v_{\text{SDP}} + \left(1 - \frac{2}{\pi}\right)\rho^0_{\text{SDP}} + \epsilon\right),$$

If Assumption 3 also holds, then (22) with $\ell = 0$ has an optimal solution, say x^0. Since $-x^0$ is also a feasible solution of (22), then $(b^0)^T x^0 \leq 0$. By an argument similar to the proof of Theorem 4, it can be shown that tx^0 is a feasible solution of (1) whenever

$$0 \leq t \leq \min_{k:\|F^k x^0 + g^k\| > 1} \frac{1 - \|g^k\|}{\|F^k x^0 + g^k\| - \|g^k\|}.$$

Moreover, $f^0(tx^0) \leq t^2 f^0(x^0) = t^2 v^0_{\text{QP}}$. Since tx^0 is a feasible solution of (1), this implies $v_{\text{QP}} \leq f^0(tx^0) \leq t^2 v^0_{\text{QP}}$. Thus, we obtain the following upper bound on v_{QP} in terms of v^0_{QP}:

$$v_{\text{QP}} \leq \min_{k:\|F^k x^0 + g^k\| > 1} \left(\frac{1 - \|g^k\|}{\|F^k x^0 + g^k\| - \|g^k\|}\right)^2 v^0_{\text{QP}}.$$

This can be combined with the lower bound (23) and the upper bound (8) to yield bounds involving mainly $v_{\text{QP}}, v_{\text{SDP}}, \rho^0_{\text{SDP}}$ and quantities depending on x^0.

References

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

[2] Anstreicher, K. and Wolkowicz, H., On Lagrangian relaxation of quadratic matrix constraints, SIAM J. Matrix Anal. Appl., 22 (2000), 41–55.

[3] Fu, M., Luo, Z.-Q., and Ye, Y., Approximation algorithms for quadratic programming, J. Comb. Optim., 2 (1998), 29–50.

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

[5] Lovász, L. and Schrijver, A., Cones of matrices and set-functions and 0-1 optimization, SIAM J. Optim., 1 (1991), 166–190.

[6] Nemirovski, A., Roos, C., and Terlaky, T., On maximization of quadratic form over intersection of ellipsoids with common center, Math. Program., 86 (1999), 463–473.

[7] Nesterov, Y., Semidefinite relaxation and nonconvex quadratic optimization, Optim. Methods Software, 9 (1998), 141–160.

[8] Nesterov, Y., Global quadratic optimization via conic relaxation, Working paper, CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, November 1998.

[9] Nesterov, Y., Global quadratic optimization on the sets with simplex structure, Working paper, CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, February 1999.

[10] Nesterov, Y. and Nemirovskii, A., *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, SIAM, Philadelphia, 1993.

[11] Nesterov, Y., Wolkowicz, H., and Ye, Y., Semidefinite programming relaxations of nonconvex quadratic optimization, in *Handbook of Semidefinite Programming*, edited by H. Wolkowicz, R. Saigal, and L. Vandenberghe, Kluwer, Boston, 2000, 360–419.

[12] Peng, J.-M. and Yuan, Y.-X., Optimality conditions for the minimization of a quadratic with two quadratic constraints, SIAM J. Optim., 7 (1997), 579–594.

[13] Rényi, A., *Probability Theory*, North-Holland, Amsterdam, 1970.

[14] Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, 1970.

[15] Sturm, J. F. and Zhang, S., On cones of nonnegative quadratic functions, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, December 2000.

[16] Wolkowicz, H., Saigal, R., and Vandenberghe, L., editors, *Handbook of Semidefinite Programming*, Kluwer, Boston, 2000.

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

[18] Ye, Y., Approximating global quadratic optimization with convex quadratic constraints, J. Global Optim., 15 (1999), 1–17.

[19] Zhang, S., Quadratic maximization and semidefinite relaxation, Math. Program., 87 (2000), 453–465.