J Glob Optim (2011) 49:293–311
DOI 10.1007/s10898-010-9545-5

Semidefinite relaxation bounds for bi-quadratic optimization problems with quadratic constraints

Xinzhen Zhang · Chen Ling · Liqun Qi

Received: 5 July 2009 / Accepted: 19 March 2010 / Published online: 8 April 2010

Abstract This paper studies the relationship between the so-called bi-quadratic optimization problem and its semidefinite programming (SDP) relaxation. It is shown that each r-bound approximation solution of the relaxed bi-linear SDP can be used to generate in randomized polynomial time an $\mathcal{O}(r)$-approximation solution of the original bi-quadratic optimization problem, where the constant in $\mathcal{O}(r)$ does not involve the dimension of variables and the data of problems. For special cases of maximization model, we provide an approximation algorithm for the considered problems.

Keywords Bi-quadratic optimization · Semidefinite programming relaxation · Approximation solution · Probabilistic solution

Mathematics Subject Classification (2000) 15A69 · 90C22 · 90C26 · 90C59

Xinzhen Zhang work is supported by the National Natural Science Foundation of China (10771120).
Chen Ling work is supported by Chinese NSF Grants 10871168 and 10971187, and a Hong Kong Polytechnic University Postdoctoral Fellowship.
Liqun Qi work is supported by the Hong Kong Research Grant Council.

X. Zhang · L. Qi
Department of Applied Mathematics, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong
e-mail: xzzhang@yahoo.cn

L. Qi
e-mail: maqilq@polyu.edu.hk

C. Ling (✉)
School of Science, Hangzhou Dianzi University, 310018 Hangzhou, China
e-mail: linghz@163.com; cling_zufe@sina.com

Present Address:
C. Ling
School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou, China

1 Introduction

In this paper, we consider the following two bi-quadratic polynomial optimization problems

$$\begin{aligned} \min\ & f(x, y) := \mathcal{B}xxyy \\ \text{s.t.}\ & x^\top A_p x \geq 1,\ \ p = 1, \ldots, m_1, \\ & y^\top B_q y \geq 1,\ \ q = 1, \ldots, n_1, \end{aligned} \tag{1}$$

and

$$\begin{aligned} \max\ & f(x, y) = \mathcal{B}xxyy \\ \text{s.t.}\ & x^\top A_p x \leq 1,\ \ p = 0, 1, \ldots, m_1, \\ & y^\top B_q y \leq 1,\ \ q = 1, \ldots, n_1, \end{aligned} \tag{2}$$

where $\mathcal{B}xxyy = \sum_{i,j=1}^{m} \sum_{k,l=1}^{n} \mathcal{B}_{ijkl} x_i x_j y_k y_l$, $A_0 \in \Re^{m\times m}$ is symmetric indefinite matrix, whereas the matrices $A_p \in \Re^{m\times m}$ $(p = 1, 2, \ldots, m_1)$ and $B_q \in \Re^{n\times n}$ $(q = 1, 2, \ldots, n_1)$ are symmetric positive semidefinite. Let $f_{\min}$ and $f_{\max}$ be the optimal values of (1) and (2), respectively. Obviously, $f_{\max} \geq 0$. Furthermore, throughout this paper, we assume that the optimal values $f_{\min}$ and $f_{\max}$ are attainable, which implies that $f_{\min} \geq 0$.

The bi-quadratic optimization problems (1) and (2) are natural generalizations of bi-quadratic optimization over unit spheres, studied in Ling et al. [17] and Wang et al. [31], which arises from the strong ellipticity condition problem in solid mechanics (for the minimization model with $n = m = 3, n_1 = m_1 = 1, A_1 = I_m$ and $B_1 = I_n$) and the entanglement problem in quantum physics; see [5,8,11,15,24,26,30] and the references therein. In fact, bi-quadratic optimization over unit spheres also has another application, such as the best rank-one approximation to a tensor. The best rank-one approximation problem has wide applications in signal and image processing, wireless communication systems, data analysis, higher-order statistics, as well as independent component analysis [3,4,6,7,10,16,21,33]. Furthermore, the bi-quadratic optimization problems (1) and (2) can be regarded as generalizations of general quadratic optimization problems studied in He et al. [12]. The reason for this is that if there exist matrices $C \in \Re^{m\times m}$ and $D \in \Re^{n\times n}$ such that $\mathcal{B} = C \otimes D$ where $\otimes$ denotes the standard Kronecker product, then the minimization model (1) will be equivalent to solving the following two quadratic optimization problems:

$$\begin{aligned} \min\ & x^\top C x \\ \text{s.t.}\ & x^\top A_p x \geq 1,\ \ p = 1, \ldots, m_1 \end{aligned} \tag{3}$$

and

$$\begin{aligned} \min\ & y^\top D y \\ \text{s.t.}\ & y^\top B_q y \geq 1, q = 1, \ldots, n_1, \end{aligned} \tag{4}$$

which were shown to be NP-hard even when C and D are positive definite due to Luo et al. [18]. It is evident that the bi-quadratic optimization problem (1) is more general than the quadratic optimization problems (3) and (4). Hence, it is NP-hard and more difficult to solve.

Analogously, the bi-quadratic maximization model (2) can be equivalently reformulated to solve two maximization quadratic optimization when $\mathcal{B} = C \otimes D$. It is well-known from Nemirovski et al. [20] that the problem of maximizing a homogeneous quadratic form over the unit cube is NP-hard to solve even in the case of positive semidefinite matrix appearing in its objective function. Therefore, the bi-quadratic maximization problem (2) is also NP-hard to solve.

For general quadratic optimization problem, a popular approach to approximately solving the considered problem is to use their SDP relaxations, which can be solved in polynomial time and have received much attention, e.g., [9,13,22,25]. We wonder whether their corresponding SDP relaxations can be used for approximately solving the original bi-quadratic problem (1) and (2). The answer is positive.

Motivated by He et al. [12], in this paper we solve approximately bi-quadratic optimization problems by their corresponding SDP relaxations. In the SDP relaxation of quadratic optimization, $x^\top Ax$ is rewritten as $A \bullet X$ with $X = xx^\top$, $X \succeq 0$, and then discard the rank restriction. By a similar technique to that used in quadratic optimization, the bi-quadratic optimization problems (1) and (2) are relaxed to the following bi-linear SDP problems with linear matrix inequality constraints

$$\begin{array}{ll} \min & g(X,Y) := (\mathcal{B}X) \bullet Y \\ \text{s.t.} & A_p \bullet X \geq 1, \ p = 1, \ldots, m_1, \\ & B_q \bullet Y \geq 1, \ q = 1, \ldots, n_1, \\ & X \succeq 0, \ Y \succeq 0 \end{array} \tag{5}$$

and

$$\begin{array}{ll} \max & g(X,Y) = (\mathcal{B}X) \bullet Y \\ \text{s.t.} & A_p \bullet X \leq 1, \ p = 0, 1, \ldots, m_1, \\ & B_q \bullet Y \leq 1, \ q = 1, \ldots, n_1, \\ & X \succeq 0, \ Y \succeq 0, \end{array} \tag{6}$$

respectively. Here, $\mathcal{B}X$ stands for a symmetric $n \times n$ matrix with

$$(\mathcal{B}X)_{kl} = \sum_{i,j=1}^{m} \mathcal{B}_{ijkl} X_{ij}.$$

Denote by $g_{\min}^{\text{sdp}}$ and $g_{\max}^{\text{sdp}}$ the optimal values of (5) and (6), respectively. Without loss of generality, we assume the optimal values are attainable, which implies that $g_{\min}^{\text{sdp}} \geq 0$. It is easy to see that, for the optimization problem (1) with $m_1 = n_1 = 1$, $A_1 = I_m$ and $B_1 = I_n$, if its optimal value is attainable, then the original problem can be equivalently reformulated as the problem studied by Ling et al. [17]. In this case, from Ling et al. [17], we know that its bi-linear SDP relaxation is tight for the problem (1). For a general quadratic/bi-quadratic problem, its SDP relaxation is not tight for the original problem. In fact, for the quadratic optimization problem (3), its SDP relaxation does not always provide a tight approximation in general. However, it does lead to provably approximation solutions for certain type of quadratic optimization problems, see [1,12,20], which motivates us to extend the existing methods for quadratic optimization problems to bi-quadratic optimization problems.

The paper is organized as follows. In Sect. 2, we analyze the approximation ratio of the SDP relaxations for bi-quadratic optimization problems. In Sect. 3, we present a polynomial time approximation algorithm for the bi-quadratic maximization model. In Sect. 4, we extend the approximation bound results obtained in Sect. 2 to the complex cases.

Notation Throughout this paper, the spaces of n-dimensional real and complex vectors are denoted by $\Re^n$ and $\mathcal{C}^n$, respectively. The spaces of $n \times n$ real symmetric and complex Hermitian matrices are denoted by $\mathcal{S}^n$ and $\mathcal{H}^n$, respectively. Matrix $Z \in \mathcal{H}^n$ means that $\text{Re}(Z)$ is symmetric and $\text{Im}(Z)$ is skew-symmetric, where $\text{Re}(Z)$ and $\text{Im}(Z)$ stand for the real and imaginary part of Z, respectively. For two real matrices A and B with the same dimension, $A \bullet B$ stands for usual matrix inner product, i.e., $A \bullet B = \text{tr}(A^\top B)$, where $\text{tr}(\cdot)$ denotes the trace

of a matrix. In addition, $\|A\|_F$ denotes the Frobenius norm of A, i.e., $\|A\|_F = (A \bullet A)^{1/2}$, and I_n stands for the $n \times n$ identity matrix. For two complex matrices A and B, their inner product

$$A \bullet B = \mathrm{Re}(\mathrm{tr}(A^H B)) = \mathrm{tr}\left(\mathrm{Re}(A)^\top \mathrm{Re}(B) + \mathrm{Im}(A)^\top \mathrm{Im}(B)\right),$$

where A^H denotes the conjugate transpose of matrix A. The notation $A \succeq 0$ ($\succ 0$) means that A is positive semidefinite (positive definite).

2 Bi-linear SDP relaxation bounds for the bi-quadratic optimization model

In this section, we study the approximation solutions for the bi-quadratic optimization models (1) and (2), based upon the approximation solutions for their bi-linear SDP relaxations. We first introduce the following definitions, which characterize the quality measure of approximation ratio.

Definition 1 The problem has an r-bound approximation solution for the given minimization model, if there is an algorithm $\mathfrak{A}$ whose complexity is polynomial such that when applied to the problem, it returns a feasible solution with objective value p such that

$$\begin{cases} rp \le p_{\min} \le p, & \text{if } p_{\min} \ge 0, \\ p_{\min} \le p \le r p_{\min}, & \text{if } p_{\min} < 0, \end{cases}$$

where $p_{\min}$ is the minimum value of the problem and $0 < r \le 1$. The feasible solution is said to be an r-bound approximation solution of the minimization model. The algorithm $\mathfrak{A}$ is said to be an r-bound approximation algorithm.

Consider the special case of (1), in which $m_1 = n_1 = 1$, A_1 and B_1 are positive definite. It is easy to see that the optimal solution pair must satisfy the constraints with equality. In this case, there exists an appropriate tensor $\bar{\mathcal{B}}$ such that (1) is equivalent to

$$\begin{aligned} \min\ & \bar{\mathcal{B}} x x y y \\ \text{s.t.}\ & x^\top x = 1, \\ & y^\top y = 1, \end{aligned}$$

which has no polynomial time algorithm $\mathfrak{A}$ to get a positive bound approximation solution for every instance of (1), see Theorem 2.2 in Ling et al. [17]. That is, a constant r-bound approximation solution may not exist for (1), so that we present a weaker notation of $(1-\epsilon)$-relative approximation solution, which is defined as follows.

Definition 2 Let $1 > \epsilon \ge 0$ and $\mathfrak{A}$ be an approximation algorithm for the given minimization model. If for any instance of the given minimization model, the algorithm $\mathfrak{A}$ returns a feasible solution with objective value p such that

$$p - p_{\min} \le \epsilon(p_{\max} - p_{\min}),$$

where $p_{\min}(p_{\max})$ is the minimum (maximun) value of the problem. The feasible solution is said to be $(1-\epsilon)$-relative approximation solution of the minimization model.

Similarly, we can give the definitions of r-bound and $(1-\epsilon)$-relative approximation solution for the maximization problem.

Based on Definition 1, we argue that there is a finite and data-independent approximation bound between the optimal values of (1) and its SDP relaxation. To this end, we need some probability estimation results which play important roles in what follows. Lemma 1 (a) comes from He et al. [12], Lemma 2 comes from Luo et al. [18] and has been used in Luo et al. [19], and Lemma 3 comes from So et al. [27]. In addition, Lemma 1 (b) can be proved easily by Lemma 1 (a) and symmetry.

Lemma 1 *Let A and Z be two real symmetric $n \times n$ matrices with $Z \succeq 0$ and $\mathrm{tr}(AZ) \geq 0$. Let $\xi \sim \mathcal{N}(0, Z)$ be a normal random vector with zero mean and covariance matrix Z. Then the following probability estimation hold.*

(*a*) *For any $0 \leq \gamma \leq 1$ we have*

$$\mathrm{Prob}\left\{\xi^\top A\xi < \gamma E[\xi^\top A\xi]\right\} < 1 - \frac{3}{100}.$$

(*b*) *For $\beta \geq 1$, we have*

$$\mathrm{Prob}\left\{\xi^\top A\xi > \beta E[\xi^\top A\xi]\right\} < 1 - \frac{3}{100}.$$

Lemma 2 *Let A and Z be two real symmetric $n \times n$ matrices with $A \succeq 0$ and $Z \succeq 0$. Suppose $\xi \sim \mathcal{N}(0, Z)$ is a normal random vector with zero mean and covariance matrix Z. Then, for any $\gamma > 0$,*

$$\mathrm{Prob}\{\xi^\top A\xi < \gamma E[\xi^\top A\xi]\} \leq \max\left\{\sqrt{\gamma}, \frac{2(r-1)\gamma}{\pi - 2}\right\},$$

where $r := \min\{\mathrm{rank}(A), \mathrm{rank}(Z)\}$.

Lemma 3 *Let A and Z be two real symmetric $n \times n$ matrices with $A \succeq 0$ and $Z \succeq 0$. Suppose $\xi \sim \mathcal{N}(0, Z)$ is a normal random vector with zero mean and covariance matrix Z. Then, for any $\gamma > 0$,*

$$\mathrm{Prob}\left\{\xi^\top A\xi > \gamma E[\xi^\top A\xi]\right\} \leq e^{\frac{1}{2}(1-\gamma+\ln\gamma)}.$$

Let $\gamma = \frac{1}{\rho^2}$. It holds that

$$\mathrm{Prob}\left\{\rho^2\xi^\top A\xi > E[\xi^\top A\xi]\right\} \leq e^{\frac{1}{2}\left(1-\frac{1}{\rho^2}-2\ln\rho\right)}.$$

Now we are ready to establish the first main result in this section, which characterizes the approximation ratio for the bi-linear SDP relaxation to (1). Our argumentation is similar to those of He et al. [12] and Luo et al. [18].

Theorem 1 *Suppose that the optimal value of (5) is nonnegative. Let $(\bar{X}, \bar{Y})$ be an r-bound approximation solution of (5). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (1) and the probability that*

$$\frac{r}{10^8 m_1^2 n_1^2} f(\bar{x}, \bar{y}) \leq f_{\min} \leq f(\bar{x}, \bar{y})$$

is at least $\dfrac{1}{2500}$.

Proof Consider the semidefinite programming of the following form

$$\begin{aligned} \min\ & (\bar{Y}\mathcal{B}) \bullet X \\ \text{s.t.}\ & A_p \bullet X \geq 1,\ \ p = 1, \ldots, m_1, \\ & X \succeq 0, \end{aligned} \tag{7}$$

where $\bar{Y}\mathcal{B}$ is a symmetric $m \times m$ matrix with

$$(\bar{Y}\mathcal{B})_{kl} = \sum_{k,l=1}^{n} \mathcal{B}_{ijkl}\bar{Y}_{kl}.$$

It is well-known that there exists an optimal solution X^* of (7) with rank r_{X^*} satisfying $\frac{r_{X^*}(r_{X^*}+1)}{2} \leq m_1$, which can be found in polynomial time; cf. [23] and [14]. Clearly, it holds that

$$(\bar{Y}\mathcal{B}) \bullet X^* \leq (\mathcal{B}\bar{X}) \bullet \bar{Y}.$$

Based upon X^*, we further consider the following standard SDP problem

$$\begin{aligned} \min\ & (\mathcal{B}X^*) \bullet Y \\ \text{s.t.}\ & B_q \bullet Y \geq 1,\ \ q = 1, \ldots, n_1, \\ & Y \succeq 0. \end{aligned} \tag{8}$$

We can find an optimal solution Y^* of (8) with rank r_{Y^*} satisfing $\frac{r_{Y^*}(r_{Y^*}+1)}{2} \leq n_1$. Since X^* and Y^* are the optimal solutions of (7) and (8), respectively, the matrix pair (X^*, Y^*) satisfies

$$0 \leq (\mathcal{B}X^*) \bullet Y^* \leq (\mathcal{B}\bar{X}) \bullet \bar{Y} \tag{9}$$

and

$$r_{X^*} \leq \sqrt{2m_1},\ \ r_{Y^*} \leq \sqrt{2n_1}. \tag{10}$$

Let $\xi \sim \mathcal{N}(0, X^*)$ and $\eta \sim \mathcal{N}(0, Y^*)$ be two independent normal random vectors, whose covariance matrices are X^* and Y^*, respectively. From the process of the proof of Theorem 3.3 in He et al. [12], it follows by Lemma 1 (b) and Lemma 2 that

$$\text{Prob}(\Omega) \geq \frac{3}{100} - m_1 \max\left\{\sqrt{\gamma_1}, \frac{2(r_{X^*}-1)\gamma_1}{\pi - 2}\right\}, \tag{11}$$

where

$$\Omega = \left\{\min_{1 \leq p \leq m_1} \xi^\top A_p \xi \geq \gamma_1,\ \xi^\top (Y^*\mathcal{B})\xi \leq \mu_1 (Y^*\mathcal{B}) \bullet X^*\right\},$$

$\gamma_1 > 0$ and $\mu_1 \geq 1$. By the assumption that the optimal value of (5) is nonnegative, we can see that $(\mathcal{B}xx^\top) \bullet Y^* \geq 0$ for any given sample value x of ξ in Ω. Hence, by Lemma 1 (b), we have

$$\text{Prob}\left\{\eta^\top(\mathcal{B}xx^\top)\eta > \mu_2(\mathcal{B}xx^\top) \bullet Y^*\right\} < 1 - \frac{3}{100}$$

for every sample value x of ξ in Ω, where $\mu_2 \geq 1$. Note that the above estimation is independent with the sample value x of ξ. Consequently, it is easy to prove that

$$\text{Prob}\left(\left\{\eta^\top(\mathcal{B}\xi\xi^\top)\eta > \mu_2(\mathcal{B}\xi\xi^\top) \bullet Y^*\right\} \cap \Omega\right) \leq \left(1 - \frac{3}{100}\right)\text{Prob}(\Omega),$$

which implies that the conditional probability

$$\text{Prob}\left\{\eta^\top(\mathcal{B}\xi\xi^\top)\eta > \mu_2(\mathcal{B}\xi\xi^\top)\bullet Y^* \mid \Omega\right\} \le 1 - \frac{3}{100}. \tag{12}$$

On the other hand, from the independence of the random variables ξ and η, it follows from Lemma 2 that for any $\gamma_2 > 0$,

$$\begin{aligned}\text{Prob}\left\{\min_{1\le q\le n_1} \eta^\top B_q\eta < \gamma_2 \mid \Omega\right\} &= \text{Prob}\left\{\min_{1\le q\le n_1} \eta^\top B_q\eta < \gamma_2\right\}\\ &\le \sum_{q=1}^{n_1}\text{Prob}\left\{\eta^\top B_q\eta < \gamma_2 E[\eta^\top B_q\eta]\right\}\\ &\le n_1\max\left\{\sqrt{\gamma_2}, \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}\right\}.\end{aligned}$$

This implies, together with (12), that

$$\begin{aligned}&\text{Prob}\left\{\eta^\top(\mathcal{B}\xi\xi^\top)\eta \le \mu_2(\mathcal{B}\xi\xi^\top)\bullet Y^*, \min_{1\le q\le n_1} \eta^\top B_q\eta \ge \gamma_2 \mid \Omega\right\}\\ &\ge 1 - \text{Prob}\left\{\eta^\top(\mathcal{B}\xi\xi^\top)\eta > \mu_2(\mathcal{B}\xi\xi^\top)\bullet Y^* \mid \Omega\right\} - \text{Prob}\left\{\min_{1\le q\le n_1} \eta^\top B_q\eta < \gamma_2 \mid \Omega\right\}\\ &\ge \frac{3}{100} - n_1\max\left\{\sqrt{\gamma_2}, \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}\right\},\end{aligned} \tag{13}$$

where the first inequality comes from the fact that

$$\text{Prob}(U\cap V) \ge 1 - \text{Prob}(U^c) - \text{Prob}(V^c)$$

for any two random events U and V, where U^c stands for the contrary event of U, etc. Noticing the relation that

$$\begin{aligned}&\left\{\min_{1\le p\le m_1} \xi^\top A_p\xi \ge \gamma_1, \min_{1\le q\le n_1} \eta^\top B_q\eta \ge \gamma_2, \eta^\top(\mathcal{B}\xi\xi^\top)\eta \le \mu_1\mu_2(\mathcal{B}X^*)\bullet Y^*\right\}\\ &\supseteq \left\{\eta^\top(\mathcal{B}\xi\xi^\top)\eta \le \mu_2(\mathcal{B}\xi\xi^\top)\bullet Y^*, \min_{1\le q\le n_1} \eta^\top B_q\eta \ge \gamma_2,\right\}\bigcap\Omega,\end{aligned}$$

it follows from (11) and (13) that

$$\begin{aligned}&\text{Prob}\left\{\min_{1\le p\le m_1} \xi^\top A_p\xi \ge \gamma_1, \min_{1\le q\le n_1} \eta^\top B_q\eta \ge \gamma_2, \eta^\top(\mathcal{B}\xi\xi^\top)\eta \le \mu_1\mu_2(\mathcal{B}X^*)\bullet Y^*\right\}\\ &\ge \left(\frac{3}{100} - m_1\max\left\{\sqrt{\gamma_1}, \frac{2(r_{X^*}-1)\gamma_1}{\pi-2}\right\}\right)\left(\frac{3}{100} - n_1\max\left\{\sqrt{\gamma_2}, \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}\right\}\right).\end{aligned}$$

Let $\gamma_1 = \frac{1}{10^4 m_1^2}$, $\gamma_2 = \frac{1}{10^4 n_1^2}$, $\mu_1 = 1$ and $\mu_2 = 1$. By (10), we have

$$\sqrt{\gamma_1} \ge \frac{2(r_{X^*}-1)\gamma_1}{\pi-2} \quad\text{and}\quad \sqrt{\gamma_2} \ge \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}.$$

Thus, it holds that

$$\begin{aligned}&\text{Prob}\left\{\min_{1\le p\le m_1} \xi^\top A_p\xi \ge \gamma_1, \min_{1\le q\le n_1} \eta^\top B_q\eta \ge \gamma_2, \eta^\top(\mathcal{B}\xi\xi^\top)\eta \le \mu_1\mu_2(\mathcal{B}X^*)\bullet Y^*\right\}\\ &\ge \frac{1}{2500},\end{aligned}$$

which implies that there exists a vector pair $(x, y) \in \Re^m \times \Re^n$ such that

$$\min_{1\le p\le m_1} x^\top A_p x \ge \gamma_1, \quad \min_{1\le q\le n_1} y^\top B_q y \ge \gamma_2 \tag{14}$$

and

$$y^\top(\mathcal{B}xx^\top)y \le \mu_1\mu_2(\mathcal{B}X^*) \bullet Y^*. \tag{15}$$

Let $\bar{x} = \dfrac{x}{\sqrt{\gamma_1}}$ and $\bar{y} = \dfrac{y}{\sqrt{\gamma_2}}$. Then, by (14), we know that $(\bar{x}, \bar{y})$ is a feasible solution pair of (1), i.e.

$$\bar{x}^\top A_p\bar{x} \ge 1 \ (p = 1, \ldots, m_1) \ \text{ and } \ \bar{y}^\top B_q\bar{y} \ge 1 \ (q = 1, \ldots, n_1).$$

Furthermore, by (9) and (15), we have

$$f(\bar{x}, \bar{y}) \le \frac{\mu_1\mu_2}{\gamma_1\gamma_2}(\mathcal{B}X^*) \bullet Y^* \le \frac{\mu_1\mu_2}{\gamma_1\gamma_2}(\mathcal{B}\bar{X}) \bullet \bar{Y}. \tag{16}$$

Since $(\bar{X}, \bar{Y})$ is an r-bound approximation solution of (5), one has

$$(\mathcal{B}\bar{X}) \bullet \bar{Y} \le \frac{1}{r} g_{\min}^{\text{sdp}} \le \frac{1}{r} f_{\min},$$

where the second inequality due to the fact that (5) is a relaxation of (1). This implies, together with (16), that

$$f(\bar{x}, \bar{y}) \le \frac{\mu_1\mu_2}{\gamma_1\gamma_2}(\mathcal{B}\bar{X}) \bullet \bar{Y} \le \frac{10^8 m_1^2 n_1^2}{r} f_{\min}.$$

Thus the desired result follows. □

In the case where $m_1, n_1 \le 2$, we have the following result, which is a generalization of Theorem 2.4 in Ling et al. [17].

Proposition 1 *Suppose that $m_1, n_1 \le 2$. Then, the bi-quadratic optimization problem* (1) *and its bi-linear SDP relaxation* (5) *are equivalent.*

Proof Without loss of generality, we assume that $m_1 = n_1 = 2$. Suppose that $(\bar{X}, \bar{Y})$ is an optimal solution pair of (5). Similar to the proof of the theorem above, we can find a matrix pair (X^*, Y^*) such that

$$(\mathcal{B}X^*) \bullet Y^* \le (\mathcal{B}X^*) \bullet \bar{Y} \le (\mathcal{B}\bar{X}) \bullet \bar{Y} \tag{17}$$

and

$$\frac{r_{X^*}(r_{X^*} + 1)}{2} \le 2, \quad \frac{r_{Y^*}(r_{Y^*} + 1)}{2} \le 2. \tag{18}$$

By (17) and (18), we know that (X^*, Y^*) is an optimal solution matrix pair of (5), which satisfies $r_{X^*} = r_{Y^*} = 1$. Hence, there exist $x^* \in \Re^m$ and $y^* \in \Re^n$ such that $X^* = x^*(x^*)^\top$ and $Y^* = y^*(y^*)^\top$. Then, we have

$$(x^*)^\top A_p x^* \ge 1 \ (p = 1, 2), \quad (y^*)^\top B_q y^* \ge 1 \ (q = 1, 2) \tag{19}$$

and

$$f(x^*, y^*) = g(X^*, Y^*). \tag{20}$$

By (19), we know that (x^*, y^*) is feasible for (1). Furthermore, by (20), it follows that

$$f(x^*, y^*) = f_{\min}.$$

We obtain the desired result and complete the proof. □

In the rest of this section, we discuss the approximation bound for the maximization problem (2).

Theorem 2 *Suppose that $(\bar{X}, \bar{Y})$ is an r-bound approximation solution of (6). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (2) such that*

$$\frac{r}{4\left(1 + 2\ln(100m_1^2)\right)\ln(100n_1)} f_{\max} \le f(\bar{x}, \bar{y}) \le f_{\max}.$$

Proof Without loss of generality, we assume that the ranks of matrices $\bar{X}$ and $\bar{Y}$ satisfy $r_{\bar{X}} \le \sqrt{2(m_1+1)}, r_{\bar{Y}} \le \sqrt{2n_1}$, respectively. Let $\bar{X} = ZZ^\top$ with $Z \in \Re^{m \times r_{\bar{X}}}$. Since $Z^\top(\bar{Y}\mathcal{B})Z$ is symmetric, there exists an orthogonal matrix Q such that $Q^\top Z^\top(\bar{Y}\mathcal{B})ZQ$ is diagonal. Let $\xi_k, k = 1, 2, \ldots, r_{\bar{X}}$ be i.i.d random variables taking values -1 and 1 with equal probabilities, and let

$$x(\xi) := \frac{1}{\sqrt{\max\limits_{0 \le p \le m_1} (\xi^\top \bar{A}_p \xi + 1)}} ZQ\xi,$$

where $\bar{A}_p = Q^\top Z^\top A_p ZQ$ $(p = 0, 1, \ldots, m_1)$ and $\xi = (\xi_1, \ldots, \xi_{r_{\bar{X}}})^\top$.

It is easy to see that the random vector $x(\xi)$ is always well-defined from the positive semidefinition of A_i for $i = 1, 2, \ldots, m_1$, and $x(\xi)^\top A_p x(\xi) \le 1$ for all $p = 0, 1, \ldots, m_1$. From the definition of $x(\xi)$, it holds that

$$\begin{aligned} x(\xi)^\top(\bar{Y}\mathcal{B})x(\xi) &= \frac{1}{\max\limits_{0 \le p \le m_1} (\xi^\top \bar{A}_p \xi + 1)} \xi^\top Q^\top Z^\top(\bar{Y}\mathcal{B})ZQ\xi \\ &= \frac{1}{\max\limits_{0 \le p \le m_1} (\xi^\top \bar{A}_p \xi + 1)} (\bar{Y}\mathcal{B}) \bullet \bar{X}. \end{aligned}$$

It is ready to verify that $\mathrm{tr}(\bar{A}_p) = A_p \bullet \bar{X} \le 1$ $(p = 0, 1, \ldots, m_1)$ and $\bar{A}_p \succeq 0$ for $p = 1, \ldots, m_1$. Therefore, from the process of the proof of Theorem 4.2, Lemma 4.1 in He et al. [12] and (12) in Nemirovski et al. [20], it follows that for any $\alpha > 2$,

$$\mathrm{Prob}(\Theta) \ge \frac{3}{100} - 2m_1^2 e^{-\frac{\alpha-1}{2}}, \tag{21}$$

where

$$\Theta = \left\{ x(\xi)^\top(\bar{Y}\mathcal{B})x(\xi) \ge \frac{1}{\alpha}(\bar{Y}\mathcal{B}) \bullet \bar{X} \right\}.$$

Let $\eta \sim \mathcal{N}(0, Y^*)$ be an normal random variable with the covariance matrix Y^*. From the fact that $x(\xi)$ and η are independent, by a similar way to that used in the proof of Theorem 1, we can prove that the conditional probability

$$\mathrm{Prob}\left\{ \eta^\top(\mathcal{B}x(\xi)x(\xi)^\top)\eta < \nu(\mathcal{B}x(\xi)x(\xi)^\top) \bullet \bar{Y} \mid \Theta \right\} < 1 - \frac{3}{100} \tag{22}$$

for any $0 \le \nu \le 1$.

On the other hand, since $E[\eta^\top B_q \eta] = B_q \bullet \bar{Y} \le 1$ for $q = 1, \ldots, n_1$, it is ready to see that $\{\eta^\top B_q \eta > \beta\} \subseteq \{\eta^\top B_q \eta > \beta E[\eta^\top B_q \eta]\}$, where $\beta > 0$. Consequently, by Lemma 3, we have that for $q = 1, \ldots, n_1$,

$$\text{Prob}\left\{\eta^\top B_q \eta > \beta\right\} \le \text{Prob}\left\{\eta^\top B_q \eta > \beta E[\eta^\top B_q \eta]\right\} \le e^{\frac{1}{2}(1-\beta+\ln\beta)}.$$

Therefore, from the independence of $x(\xi)$ and η, we have

$$\begin{aligned}\text{Prob}\left\{\max_{1\le q\le n_1} \eta^\top B_q \eta > \beta \mid \Theta\right\} &= \text{Prob}\left(\bigcup_{q=1}^{n_1}\left\{y^\top B_q y > \beta\right\}\right)\\ &\le n_1 e^{\frac{1}{2}(1-\beta+\ln\beta)}.\end{aligned} \tag{23}$$

By (22) and (23), it follows that

$$\begin{aligned}&\text{Prob}\left\{\max_{1\le q\le n_1} \eta^\top B_q \eta \le \beta, \eta^\top(\mathcal{B}x(\xi)x(\xi)^\top)\eta \ge \nu(\mathcal{B}x(\xi)x(\xi)^\top)\bullet\bar{Y} \mid \Theta\right\}\\ &\ge \frac{3}{100} - n_1 e^{\frac{1}{2}(1-\beta+\ln\beta)}.\end{aligned} \tag{24}$$

Noticing that

$$\begin{aligned}&\left\{\eta^\top(\mathcal{B}x(\xi)x(\xi)^\top)\eta \ge \frac{1}{\alpha}\nu(\mathcal{B}\bar{X})\bullet\bar{Y}, \max_{1\le q\le n_1} \eta^\top B_q \eta \le \beta\right\}\\ &\supseteq \left\{\eta^\top(\mathcal{B}x(\xi)x(\xi)^\top)\eta \ge \nu(\mathcal{B}x(\xi)x(\xi)^\top)\bullet\bar{Y}, \max_{1\le q\le n_1} \eta^\top B_q \eta \le \beta\right\}\bigcap\Theta,\end{aligned}$$

it follows from (21) and (24) that

$$\begin{aligned}&\text{Prob}\left\{\eta^\top(\mathcal{B}x(\xi)x(\xi)^\top)\eta \ge \frac{1}{\alpha}\nu(\mathcal{B}\bar{X})\bullet\bar{Y}, \max_{1\le q\le n_1} \eta^\top B_q \eta \le \beta\right\}\\ &\ge \left(\frac{3}{100} - 2m_1^2 e^{-\frac{\alpha-1}{2}}\right)\left(\frac{3}{100} - n_1 e^{\frac{1}{2}(1-\beta+\ln\beta)}\right).\end{aligned}$$

Let $\alpha = 1 + 2\ln(100m_1^2)$ and $\beta = 4\ln(100n_1)$, we have

$$\text{Prob}\left\{\eta^\top(\mathcal{B}x(\xi)x(\xi)^\top)\eta \ge \frac{1}{\alpha}\nu(\mathcal{B}\bar{X})\bullet\bar{Y}, \max_{1\le q\le n_1} \eta^\top B_q \eta \le \beta\right\} \ge \frac{1}{10^4} > 0,$$

which implies that there exist vectors $\bar{x} = x(\xi) \in \Re^m$ and $y \in \Re^n$ such that

$$\bar{x}^\top A_p \bar{x} \le 1 \ (p = 0, 1, \ldots, m_1), \quad y^\top B_q y \le \beta \ (q = 1, \ldots, n_1)$$

and

$$y^\top(\mathcal{B}\bar{x}\bar{x}^\top)y \ge \frac{1}{\alpha}\nu(\mathcal{B}\bar{X})\bullet\bar{Y}.$$

Let $\bar{y} = \dfrac{y}{\sqrt{\beta}}$ and $\nu = 1$. Then $(\bar{x}, \bar{y})$ is a feasible solution of (2) satisfying

$$\frac{1}{\alpha\beta}(\mathcal{B}\bar{X})\bullet\bar{Y} \le \bar{y}^\top(\mathcal{B}\bar{x}\bar{x}^\top)\bar{y} \le f_{\max}.$$

Furthermore, by the definition of r-bound approximation solution, we obtain the desired result and complete the proof. □

Similar to Proposition 1, we have

Proposition 2 *Suppose that the numbers of constraints on x and y are not larger than 2, respectively. Then, the bi-quadratic optimization problem (2) and its bi-linear SDP relaxation (6) are equivalent.*

Remark Notice that the computational effort required for solving the bi-linear SDP relaxations of (1) and (2) can be significantly large. Therefore, it is very interesting to analyze the size of the resulted SDP relaxations, which will be our future research topic.

3 Approximation solution of the bi-quadratic problem

Our main goal in this paper is to design polynomial time approximation algorithms for (1) and (2). Theorems 1 and 2 show that this task depends strongly on our ability to approximately solve the relaxed problems (5) and (6), which by themselves are also NP-hard. However, it is possible to derive approximation solution of the relaxed problems. In this section, we consider some forms of optimization problems whose approximation solution of their SDP relaxation problem can be solved in polynomial time. We first give an approximation result for the general model (2) under some mild assumptions. Then we investigate the bi-quadratic optimization problems with two constraints.

3.1 The bi-quadratic maximization model

In this subsection, we consider the maximization problem (2). To this end, we make the following assumptions.

(A1) $|\mathrm{tr}(A_0)| < m, \mathrm{tr}(A_p) < m$ for every $p = 1, \ldots, m_1$, and $\mathrm{tr}(B_q) < n$ for every $q = 1, \ldots, n_1$.

(A2) There exist nonnegative numbers α_p $(p = 0, 1, \ldots, m_1)$ with $\sum_{p=0}^{m_1} \alpha_p = 1$ and β_q $(q = 1, \ldots, n_1)$ with $\sum_{q=1}^{n_1} \beta_q = 1$, such that

$$\sum_{p=0}^{m_1} \alpha_p A_p - I_m \succeq 0 \ \text{ and } \ \sum_{q=1}^{n_1} \beta_q B_q - I_n \succeq 0.$$

(A3) $A_0 + I_m \succeq 0$.

We further need the following lemma, which generalizes the result used in Ling et al. [17].

Lemma 4 *For any $X \in \mathcal{S}^m$, the following statements hold.*

(1) *If $\|X\|_F \le \frac{1}{m}$, then $\bar{X} := X + \frac{1}{m} I_m \succeq 0$.*

(2) *Suppose $m \ge 2$. If $\mathrm{tr}(X) \le 0$ and $X \succeq -\frac{1}{m} I_m$, then $\|X\|_F \le \sqrt{1 - \frac{1}{m}}$.*

Proof (1) Since $\|X\|_F \le \frac{1}{m}$, it follows that $|x_{ii}| \le \frac{1}{m}$ for every $i = 1, \ldots, m$. This implies that

$$\mathrm{tr}(\bar{X}) = \mathrm{tr}(X) + 1 = \sum_{i=1}^{m} x_{ii} + 1 \ge 0. \tag{25}$$

To show that $\bar{X} \succeq 0$, by Lemma 2.1 in Berkelaar et al. [2], we only need to show that

$$\sqrt{m-1}\|\bar{X}\|_F \le \mathrm{tr}(\bar{X}). \tag{26}$$

It is easy to see that

$$\|\bar{X}\|_F^2 = \|X\|_F^2 + \frac{2}{m}\mathrm{tr}(X) + \frac{1}{m}, \tag{27}$$

which implies, together with (25), that

$$\begin{aligned}(m-1)\|\bar{X}\|_F^2 - (\mathrm{tr}(\bar{X}))^2 &= (m-1)\left[\|X\|_F^2 + \frac{2}{m}\mathrm{tr}(X) + \frac{1}{m} - \frac{1}{m-1}(\mathrm{tr}(X)+1)^2\right] \\ &\le (m-1)\left[\frac{1}{m^2} + \frac{2}{m}\mathrm{tr}(X) + \frac{1}{m} - \frac{1}{m-1}(\mathrm{tr}(X)+1)^2\right] \\ &= -\left(\mathrm{tr}(X) + \frac{1}{m}\right)^2 \le 0.\end{aligned}$$

Therefore, (26) holds. This shows that $\bar{X} \succeq 0$.

(2) Since $\bar{X} = X + \frac{1}{m}I_m \succeq 0$, it follows that

$$-1 \le \mathrm{tr}(X) \le 0, \tag{28}$$

from the given condition that $\mathrm{tr}(X) \le 0$. Moreover, it holds that

$$\begin{aligned}\|\bar{X}\|_F^2 &\le (\mathrm{tr}(\bar{X}))^2 \\ &= (\mathrm{tr}(X))^2 + 2\mathrm{tr}(X) + 1,\end{aligned}$$

where the inequality is due to the positive semidefiniteness of $\bar{X}$. This implies, together with (27), that

$$\|X\|_F^2 \le (\mathrm{tr}(X))^2 + 2\left(1 - \frac{1}{m}\right)\mathrm{tr}(X) + 1 - \frac{1}{m}. \tag{29}$$

Consider the optimization problem as follows

$$\begin{aligned}p_{\max} := \max\ & p(t) = t^2 + 2bt + c \\ \text{s.t. } & l \le t \le u.\end{aligned}$$

It is easy to verify that $p_{\max} = \max\{p(l), p(u)\}$. Consequently, by this, (28) and (29), we know that $\|X\|_F^2 \le 1 - \frac{1}{m}$ and complete the proof. □

Considering linear transformations $X := X - \frac{1}{m}I_m$, $Y := Y - \frac{1}{n}I_n$, we know that under Assumptions (A1)–(A3), a restriction and a relaxation for (6) can be written in a unified form as

$$\begin{aligned}p_\lambda := \max\ & \Phi(X,Y) = (\mathcal{B}X)\bullet Y + \tfrac{1}{m}(\mathcal{B}I_m)\bullet Y + \tfrac{1}{n}(\mathcal{B}X)\bullet I_n + \tfrac{1}{mn}(\mathcal{B}I_m)\bullet I_n \\ \text{s.t. } & \left(A_p \bullet X + \tfrac{1}{m}\mathrm{tr}(A_p)\right)^2 \le 1,\ p = 0, 1, \ldots, m_1, \\ & \left(B_q \bullet Y + \tfrac{1}{n}\mathrm{tr}(B_q)\right)^2 \le 1,\ q = 1, \ldots, n_1, \\ & \|X\|_F \le \lambda, \\ & \|Y\|_F \le \lambda,\end{aligned} \tag{30}$$

where $\lambda = \frac{1}{\max\{m,n\}}$ and $\lambda = \sqrt{1 - \frac{1}{\max\{m,n\}}}$ correspond to the restriction and the relaxation, respectively. It is easy to see that matrix pair $(0, 0) \in \mathcal{S}^m \times \mathcal{S}^n$ is a feasible solution of (30) for any $\lambda \ge 0$. Furthermore $p_0 = \frac{1}{mn}(\mathcal{B}I_m)\bullet I_n$.

By stacking up the entries of a symmetric matrix (ignoring the symmetric part) into a vector, denoted by $vec_S(\cdot)$, there exists a suitable quadratic function $q_0(u, v)$ such that (30) can be rewritten into the following form

$$\begin{aligned} p_\lambda := \max\ & q_0(u, v) \\ \text{s.t.}\ & \left(vec_S(A_p)^\top u + \tfrac{1}{m}\mathrm{tr}(A_p)\right)^2 \le 1,\ p = 0, 1, \ldots, m_1, \\ & \left(vec_S(B_q)^\top v + \tfrac{1}{n}\mathrm{tr}(B_q)\right)^2 \le 1,\ q = 1, \ldots, n_1, \\ & \|u\| \le \lambda,\ \|v\| \le \lambda, \end{aligned} \tag{31}$$

where $u = vec_S(X)$, $v = vec_S(Y)$. It is well-known that for a quadratic function $q(x) = c+2b^\top x+x^\top Ax$, the homogenized version of $q(x)$ can be represented by the matrix denoted by

$$M(q(\cdot)) = \begin{pmatrix} c & b^\top \\ b & A \end{pmatrix}.$$

Hence, a standard SDP relaxation for the homogenized version of (31) is

$$\begin{aligned} z(\lambda^2) := \max\ & \bar{Q}_0 \bullet Z \\ \text{s.t.}\ & \bar{C}_p \bullet Z \le 1,\ p = 0, 1, \ldots, m_1 + n_1, \\ & \bar{C} \bullet Z \le \lambda^2, \\ & \bar{D} \bullet Z \le \lambda^2, \\ & Z = \begin{pmatrix} 1 & u^\top & v^\top \\ u & W & U^\top \\ v & U & V \end{pmatrix} \succeq 0, \end{aligned} \tag{32}$$

where $\bar{Q}_0, \bar{C}_p,\ (p = 0, 1, \ldots, m_1 + n_1)$, $\bar{C}$ and $\bar{D}$ are some suitable matrices, which correspond to the matrix representations of the homogenized version of the quadratic functions with respect to (u, v) in problem (31), respectively. Note that (32) can be solved in polynomial time.

Based upon the analysis above, we arrive at the following conclusion.

Theorem 3 *Suppose that Assumptions (A1)–(A3) hold and* $(\mathcal{B}I_m) \bullet I_n \ge 0$. *Then a* $\frac{(1-\gamma)^2}{(\sqrt{m_1+n_1+3}+\gamma)^2\rho(\rho-1)}$*-bound approximation solution of* (6) *can be found in polynomial time, where* $\rho = \max\{m, n\}$ *and*

$$\gamma = \max\left\{\frac{1}{m}|\mathrm{tr}(A_0)|,\ \frac{1}{m}\mathrm{tr}(A_p),\ p = 1, \ldots, m_1,\ \frac{1}{n}\mathrm{tr}(B_q), q = 1, 2, \ldots, n_1\right\}.$$

Proof We consider the problem (31) with $\lambda = \frac{1}{\rho}$. By Theorem 1 in Tseng [29], there exists a feasible solution (u, v) of problem (31) satisfying

$$q_0(u, v) \ge \frac{(1-\gamma)^2}{(\sqrt{m_1+n_1+3}+\gamma)^2} z\left(\frac{1}{\rho^2}\right).$$

On the other hand, it is easy to see that $z(\lambda)$ is concave on $\lambda \ge 0$, and hence

$$\begin{aligned} z\left(\frac{1}{\rho^2}\right) &\ge \left(1 - \frac{1}{\rho(\rho-1)}\right) z(0) + \frac{1}{\rho(\rho-1)} z\left(1 - \frac{1}{\rho}\right) \\ &\ge \frac{1}{\rho(\rho-1)} z\left(1 - \frac{1}{\rho}\right) \\ &\ge \frac{1}{\rho(\rho-1)} g_{\max}^{\mathrm{sdp}}, \end{aligned}$$

where the second inequality is due to $z(0) = p_0 = \frac{1}{mn}(\mathcal{B}I_m) \bullet I_n \geq 0$, and the last inequality comes from the fact that $z\left(1 - \frac{1}{\rho}\right) \geq p_{\sqrt{1-\frac{1}{\rho}}} \geq g_{\max}^{\text{sdp}}$. Therefore,

$$q_0(u, v) \geq \frac{(1-\gamma)^2}{(\sqrt{m_1 + n_1 + 3} + \gamma)^2 \rho(\rho - 1)} g_{\max}^{\text{sdp}}. \tag{33}$$

By the obtained (u, v) and the stack relation between the vector and the matrix, we can find a feasible matrix pair $(\bar{X}, \bar{Y})$ for (30) with $\lambda = \frac{1}{\rho}$ such that $\Phi(\bar{X}, \bar{Y}) = q_0(u, v)$. Denote $X^* = \bar{X} + \frac{1}{m} I_m$ and $Y^* = \bar{Y} + \frac{1}{n} I_n$. By Lemma 4 (1), it holds that (X^*, Y^*) is a feasible solution of (6), satisfying

$$(\mathcal{B}X^*) \bullet Y^* \geq \frac{(1-\gamma)^2}{(\sqrt{m_1 + n_1 + 3} + \gamma)^2 \rho(\rho - 1)} g_{\max}^{\text{sdp}}.$$

Therefore, we can assert that (X^*, Y^*) is a $\frac{(1-\gamma)^2}{(\sqrt{m_1+n_1+3}+\gamma)^2\rho(\rho-1)}$-bound approximation solution of (6). Combining with the fact that $0 \leq \gamma < 1$, the desired result follows. □

3.2 The bi-quadratic optimization problem with two constraints

In this subsection, we first consider the following problem

$$\begin{aligned} f_{\max} := \max\ & \mathcal{B}xxyy \\ \text{s.t. } & x^\top A x \leq 1, \\ & y^\top B y \leq 1, \end{aligned} \tag{34}$$

where $A \in \mathcal{S}^m$ and $B \in \mathcal{S}^n$ are positive definite. We assume that $(\mathcal{B}I_m) \bullet I_n \geq 0$. Without loss of generality, we further assume that $A = I_m$ and $B = I_n$.

Notice that the optimal solution must satisfy the constraints with equality. Therefore, the bi-linear SDP relaxation of (34) can be written equivalently as follows

$$\begin{aligned} g_{\max}^{\text{sdp}} := \max\ & (\mathcal{B}X) \bullet Y \\ \text{s.t. } & \text{tr}(X) = 1, \\ & \text{tr}(Y) = 1, \\ & X \succeq 0,\ Y \succeq 0. \end{aligned} \tag{35}$$

By a similar procedure used in Subsect. 3.1, a restriction and a relaxation of (35) can be written in a unified form as

$$\begin{aligned} p_\lambda := \max\ & (\mathcal{B}X) \bullet Y + \tfrac{1}{m}(\mathcal{B}I_m) \bullet Y + \tfrac{1}{n}(\mathcal{B}X) \bullet I_n + \tfrac{1}{mn}(\mathcal{B}I_m) \bullet I_n \\ \text{s.t. } & \text{tr}(X) = 0, \\ & \text{tr}(Y) = 0, \\ & \|X\|_F \leq \lambda, \\ & \|Y\|_F \leq \lambda, \end{aligned} \tag{36}$$

where $\lambda = \frac{1}{\max\{m,n\}}$ and $\lambda = \sqrt{1 - \frac{1}{\max\{m,n\}}}$ correspond to the restriction and the relaxation, respectively. Hence, it follows that $p_{\sqrt{1-\frac{1}{\max\{m,n\}}}} \geq g_{\max}^{\text{sdp}} \geq p_{\frac{1}{\max\{m,n\}}} \geq p_0 = \frac{1}{mn}(\mathcal{B}I_m) \bullet I_n \geq 0$. Furthermore, for $vec_S(X)$ and $vec_S(Y)$, we can eliminate two variables, say X_{11} and Y_{11}, by their linear relation with the other variables. For convenience, let

$$u = vec_S(X) \backslash X_{11} \ \text{ and } \ v = vec_S(Y) \backslash Y_{11}.$$

Then, there exist $Q_0 \in \Re^{L_m \times L_n}$, $Q_1 \in \mathcal{S}^{L_m}$, $Q_2 \in \mathcal{S}^{L_n}$, $b_0 \in \Re^{L_m}$, $c_0 \in \Re^{L_n}$, and $d_0 = \frac{1}{mn}(\mathcal{B}I_m) \bullet I_n \in \Re$ such that the above problem is equivalent to

$$\begin{aligned} p_\lambda := \max\ & q(u,v) = u^\top Q_0 v + 2b_0^\top u + 2c_0^\top v + d_0 \\ \text{s.t. } & q_1(u,v) = u^\top Q_1 u \le \lambda^2, \\ & q_2(u,v) = v^\top Q_2 v \le \lambda^2, \end{aligned} \tag{37}$$

where $L_m = m(m+1)/2 - 1$, $L_n = n(n+1)/2 - 1$ and Q_1, Q_2 are positive definite. Furthermore, it is easy to see that the SDP relaxation of the homogenized version of (37) is

$$\begin{aligned} z(\lambda^2) := \max\ & \bar{Q}_0 \bullet Z \\ \text{s.t. } & \bar{Q}_1 \bullet Z \le \lambda^2 \\ & \bar{Q}_2 \bullet Z \le \lambda^2, \\ & Z = \begin{pmatrix} 1 & u^\top & v^\top \\ u & W & U^\top \\ v & U & V \end{pmatrix} \succeq 0, \end{aligned} \tag{38}$$

where $\bar{Q}_0, \bar{Q}_1, \bar{Q}_2$ are three matrices which correspond to the homogenized version of the quadratic functions $q(u,v)$, $q_1(u,v)$, and $q_2(u,v)$, respectively.

Consider the problem (38) with $\lambda_0 = \frac{1}{\sqrt{2}\rho}$ and $\rho = \max\{m,n\}$. Since this SDP has three constraints, so that an optimal solution Z^* can be computed in polynomial time such that its rank equals 2 (e.g., see [32]). Let us denote by I_{11} the $(L_m + L_n + 1) \times (L_m + L_n + 1)$ symmetric matrix with 1 at its (1, 1)th position and 0 elsewhere. It is clear that $I_{11} \bullet Z^* = 1$. Hence, by Corollary 4 in Sturm and Zhang [28], one can always find two vectors $z^i = (t_i, (u^i)^\top, (v^i)^\top)^\top (i = 1, 2) \in \Re^{1+L_m+L_n}$ such that $Z^* = z^1(z^1)^\top + z^2(z^2)^\top$ and

$$I_{11} \bullet z^i (z^i)^\top = I_{11} \bullet Z^*/2 = 1/2, \quad \text{for } i = 1, 2,$$

which implies that $t_1^2 = t_2^2 = 1/2$. From the structure of the constraints of (37), it is ready to know that both $\bar{Q}_1$ and $\bar{Q}_2$ are positive semidefinite. Consequently, since Z^* is feasible for (38), it holds that

$$(z^i)^\top \bar{Q}_1 z^i \le \lambda_0^2 \ \text{ and } \ (z^i)^\top \bar{Q}_2 z^i \le \lambda_0^2, \quad \text{for } i = 1, 2,$$

which implies that $(\bar{u}^i, \bar{v}^i) = (u^i/t_i, v^i/t_i)$, $i = 1, 2$, are feasible solution of (37) with $\lambda = \frac{1}{\rho}$. Furthermore, we have

$$q(\bar{u}^1, \bar{v}^1) + q(\bar{u}^2, \bar{v}^2) = \left(\bar{Q}_0 \bullet z^1(z^1)^\top + \bar{Q}_0 \bullet z^2(z^2)^\top\right)/t_1^2 = 2\bar{Q}_0 \bullet Z^* = 2z(\lambda_0^2),$$

which implies that either $(\bar{u}^1, \bar{v}^1)$ or $(\bar{u}^2, \bar{v}^2)$, denoted by $(\bar{u}, \bar{v})$, satisfies

$$q(\bar{u}, \bar{v}) \ge z(\lambda_0^2). \tag{39}$$

On the other hand, it is easy to see that $z(\cdot)$ is concave, and hence

$$z(\lambda_0^2) \ge \left(1 - \frac{1}{2\rho(\rho-1)}\right) z(0) + \frac{1}{2\rho(\rho-1)} z(1 - 1/\rho) \ge \frac{1}{2\rho(\rho-1)} z(1 - 1/\rho),$$

where the last inequality due to the assumption that $z(0) \ge d_0 \ge 0$. Therefore,

$$q(\bar{u}, \bar{v}) \ge z(\lambda_0^2) \ge \frac{1}{2\rho(\rho-1)} z(1 - 1/\rho) \ge \frac{1}{2\rho(\rho-1)} f_{\max}, \tag{40}$$

where the last inequality comes from the fact that $z(1 - 1/\rho) \ge p_{\sqrt{1-1/\rho}} \ge g_{\max}^{\text{sdp}} \ge f_{\max}$. Similar to the process of the proof of Theorem 3, from the obtained $(\bar{u}, \bar{v})$, we can find a

feasible matrix pair $(\bar{X}, \bar{Y})$ of (35) such that $(\mathcal{B}\bar{X}) \bullet \bar{Y} = q(\bar{u}, \bar{v})$. Consequently, by using a similar procedure to that used in Theorem 2.4 in Ling et al. [17], we can get a vector pair $(\bar{x}, \bar{y})$ such that $\|\bar{x}\| = \|\bar{y}\| = 1$ and $\bar{y}^\top(\mathcal{B}\bar{x}\bar{x}^\top)\bar{y} \geq q(\bar{u}, \bar{v})$. This shows that $(\bar{x}, \bar{y})$ is a feasible solution of (34), and hence $f_{\max} \geq \bar{y}^\top(\mathcal{B}\bar{x}\bar{x}^\top)\bar{y} \geq q(\bar{u}, \bar{v})$. Together with (40), we can assert that $(\bar{x}, \bar{y})$ is a $\frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}$-bound approximation solution of (34). Therefore, the following assertion is established.

Theorem 4 *If $(\mathcal{B}I_m) \bullet I_n \geq 0$, then a $\frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}$-bound approximation solution of (34) can be found in polynomial time.*

In fact, from above procedure, we can see that assumption $(\mathcal{B}I_m) \bullet I_n \geq 0$ is used to guarantee that $z(0) \geq 0$. Therefore, if we replace $\mathcal{B}$ by $\mathcal{B} - cI_m \otimes I_n$ with constant $c \leq \frac{1}{mn}(\mathcal{B}I_m) \bullet I_n$, then $z(0) \geq 0$ is guaranteed. By Theorem 4, there exist a feasible solution pair $(\bar{x}, \bar{y})$ such that

$$\mathcal{B}\bar{x}\bar{x}\bar{y}\bar{y} - c \geq \frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}(f_{max} - c).$$

Let $c = \bar{g}_{\min}$, where $\bar{g}_{\min}$ is the minimum value of the objective in (35), then $c \leq \frac{1}{mn}(\mathcal{B}I_m) \bullet I_n$. This lead to the following result.

Theorem 5 *There exists a $\left(1 - \frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}\right)$-relative approximation solution for (34) in polynomial time.*

We conclude this subsection by considering the following minimization problem

$$\begin{array}{ll} \min & \mathcal{B}xxyy \\ \text{s.t.} & x^\top x \geq 1, \\ & y^\top y \geq 1. \end{array} \tag{41}$$

It is easy to see that the optimal solution must satisfy the constraints with equality if $f_{\min}$ is attainable. Thus, the bi-linear SDP relaxation can be written as (35) with tensor $-\mathcal{B}$, which leads to the following result.

Theorem 6 *Suppose that the optimal value is attainable. If $(\mathcal{B}I_m) \bullet I_n \leq 0$, then a $\frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}$-bound approximation solution of (41) can be found in polynomial time. Otherwise, there exist a $\left(1 - \frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}\right)$-relative approximation solution for (41) in polynomial time.*

4 Extensions and discussions

Motivated by the aforementioned work on complex SDP in Luo et al. [18], our analysis can be extended to the so-called complex bi-quadratic optimization problem. In this section, we further consider the minimization model

$$\begin{array}{ll} \min & f(x, y) := \mathcal{B}xxyy \\ \text{s.t.} & x^H A_p x \geq 1, \ p = 1, \ldots, m_1, \\ & y^H B_q y \geq 1, \ q = 1, \ldots, n_1, \\ & x \in \mathcal{C}^m, \ y \in \mathcal{C}^n \end{array} \tag{42}$$

and the maximization model

$$\begin{aligned} \max\ & f(x,y) = \mathcal{B}xxyy \\ \text{s.t. } & x^H A_p x \le 1,\ p = 0, 1, \ldots, m_1, \\ & y^H B_q y \le 1,\ q = 1, \ldots, n_1, \\ & x \in \mathcal{C}^m,\ y \in \mathcal{C}^n, \end{aligned} \tag{43}$$

where $\mathcal{C}$ is the field of complex numbers, and H represents Hermitian transpose, $A_p \in \mathcal{H}^m$ $(p = 1, \ldots, m_1)$ and $B_q \in \mathcal{H}^n$ $(q = 1, \ldots, n_1)$ are positive semidefinite, whereas $A_0 \in \mathcal{H}^m$ is indefinite.

A similar procedure to that in Sect. 2 can be applied to yield the approximation bounds for the complex bi-quadratic optimization problems above. To this end, we need the following probability estimation results, which comes from He et al. [12] and Luo et al. [18], respectively.

Lemma 5 *Let A, Z be two Hermitian matrices satisfying $Z \succeq 0$ and $\mathrm{tr}(AZ) \ge 0$. Let $\xi \sim N_c(0, Z)$ be a complex normal random vector. Then,*

(a) *For any $0 \le \gamma \le 1$, it holds that*

$$\mathrm{Prob}\left\{\xi^H A\xi < \gamma E(\xi^H A\xi)\right\} < 1 - \frac{1}{20}.$$

(b) *For any $\beta \ge 1$, it holds that*

$$\mathrm{Prob}\left\{\xi^H A\xi > \beta E(\xi^H A\xi)\right\} < 1 - \frac{1}{20}.$$

Lemma 6 *Let A, Z be two Hermitian positive semidefinite matrices. Suppose that ξ is a random vector generated from the complex-valued normal distribution $N_c(0, Z)$. Then for any $\gamma > 0$, the following probability estimation hold.*

(a) $\mathrm{Prob}\left\{\xi^H A\xi < \gamma E(\xi^H A\xi)\right\} \le \max\left\{\frac{4}{3}\gamma, 16(r-1)^2\gamma^2\right\}$,

(b) $\mathrm{Prob}\left\{\xi^H A\xi > \gamma E(\xi^H A\xi)\right\} \le re^{-\gamma}$,

where $r := \min\{\mathrm{rank}(A), \mathrm{rank}(Z)\}$.

The following main result in this section can be proved in the similar ways to that used in the proofs of Theorems 1 and 2.

Theorem 7 *Let $(\bar{X}, \bar{Y})$ be an r-bound approximation solution of the bi-linear SDP relaxation of (42). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (42) and the probability that*

$$\frac{r}{1600 m_1 n_1} f(\bar{x}, \bar{y}) \le f_{\min} \le f(\bar{x}, \bar{y})$$

is at least $\frac{1}{3600}$.

Suppose that $(\bar{X}, \bar{Y})$ be an r-bound approximation solution of the bi-linear SDP relaxation for (43). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (43) and the probability that

$$\frac{r}{\left(1 + 2\ln 100 m_1^2\right)\ln\left(40\sqrt{2}n_1^{\frac{3}{2}}\right)} f_{\max} \le f(\bar{x}, \bar{y}) \le f_{\max}$$

is at least $\frac{1}{4000}$.

It is well-known that if the numbers of constraints in the considered complex SDP problem is at most 3, then its rank-one optimal solution can be found, see Theorem 2.1, Proposition 5.1 in Huang and Zhang [14]. As a consequence, we get the following proposition which can be proved by the similar ways to that used in the proofs of Propositions 1 and 2.

Proposition 3 *Suppose that the numbers of constraints on x and y are less than* 4*, respectively. Then, the bi-quadratic optimization problems* (42)*,* (43) *and their relaxations are equivalent, respectively.*

References

1. Ben-Tal, A., Nemirovski, A., Roos, C.: Robust solutions of uncertain quadratic and conic quadratic problems. SIAM J. Optim. **13**, 535–560 (2002)
2. Berkelaar, A.B., Sturm, J.F., Zhang, S.Z.: Polynomial primal-dual cone affine scaling for semidefinite programming. Appl. Numer. Math. **29**, 317–333 (1999)
3. Cardoso, J.F.: High-order contrasts for independent component analysis. Neural Comput. **11**, 157–192 (1999)
4. Comon, P.: Independent component analysis, a new concept? Signal Process. **36**, 287–314 (1994)
5. Dahl, G., Leinaas, J.M., Myrheim, J., Ovrum, E.: A tensor product matrix aproximation problem in quantum physics. Linear Algebra Appl. **420**, 711–725 (2007)
6. De Lathauwer, L., Comon, P., De Moor, B., Vandewalle, J.: Higher-order power method—application in independent component analysis. In: Proceedings of the International Symposium on Nonlinear Theory and its Applications (NOLTA'95), Las Vegas, NV, pp. 91–96 (1995)
7. De Lathauwer, L., De Moor, B., Vandewalle, J.: On the best rank-1 and rank-$(R_1, R_2, \ldots, R_N)$ approximation of higher-order tensor. SIAM J. Matrix Anal. Appl. **21**, 1324–1342 (2000)
8. Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777–780 (1935)
9. Fujisawa, K., Futakata, Y., Kojima, M., Matsuyama, S., Nakamura, S., Nakata, K., Yamashita, M.: SDPA-M (SemiDefinite Programming Algorithm in MATLAB). http://homepage.mac.com/klabtitech/sdpa-homepage/download.html
10. Grigorascu, V.S., Regalia, P.A .: Tensor displacement structures and polyspectral matching. In: Kailath, T., Sayed, A.H. Chapter 9 of Fast Reliable Algorithms for Structured Matrices, SIAM Publications, Philadeliphia (1999)
11. Han, D., Dai, H.H., Qi, L.: Conditions for strong ellipticity of anisotropic elastic materials. J. Elast. **97**, 1–13 (2009)
12. He, S.M., Luo, Z.Q., Nie, J., Zhang, S.Z.: Semidefinite relaxation bounds for indefinite homogeneous quadratic optimization. SIAM J. Optim. **19**, 503–523 (2008)
13. Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization. Kluwer, Dordrecht (2000)
14. Huang, Y.M., Zhang, S.Z.: Complex matrix decomposition and quadratic programming. Math. Oper. Res. **32**, 758–768 (2007)
15. Knowles, J.K., Sternberg, E.: On the ellipticity of the equations of the equations for finite elastostatics for a special material. J. Elast. **5**, 341–361 (1975)
16. Kofidis, E., Regalia, P.A.: On the best rank-1 approximation of higher-order supersymmetric tensors. SIAM J. Matrix Anal. Appl. **23**, 863–884 (2002)
17. Ling, C., Nie, J., Qi, L., Ye, Y.: Bi-quadratic optimization over unit spheres and semidefinite programming relaxations. SIAM J. Optim. **20**, 1286–1310 (2009)
18. Luo, Z.Q., Sidiropoulos, N., Tseng, P., Zhang, S.Z.: Approximation bounds for quadratic optimization with homogeneous quadratic constraints. SIAM J. Optim. **18**, 1–28 (2007)
19. Luo, Z.Q., Zhang, S.Z.: A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints, Technical Report Seem 2008-06, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong (2009)
20. Nemirovski, A., Roos, C., Terlaky, T.: On maximization of quadratic form over intersection of ellipsoids with common center. Math. Program. **86**, 463–473 (1999)
21. Nikias, C.L., Petropulu, A.P.: Higher-Order Spectra Analysis, A Nonlinear Signal Processing Framework. Prentice-Hall, Englewood Cliffs, NJ (1993)
22. Pardalos, P.M., Wolkowicz, H.: Topics in Semidefinite and Interior-Point Methods, Fields Institute Communications, Vol. 18. AMS, Providence, Rhode Island (1998)

23. Pataki, G.: On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. Math. Oper. Res. **23**, 339–358 (1998)
24. Qi, L., Dai, H.H., Han, D.: Conditions for strong ellipticity and M-eigenvalues. Front. Math. China **4**, 349–364 (2009)
25. Ramana, M., Pardalos, P.M.: Semidefinite programming. In: Terlaky, T. Interior Point Methods of Mathematical Programming, pp. 369–398. Kluwer, Dordrecht, The Netherlands (1996)
26. Rosakis, P.: Ellipticity and deformations with discontinuous deformation gradients in finite elastostatics. Arch. Ration. Mech. Anal. **109**, 1–37 (1990)
27. So, A.M.-C., Ye, Y., Zhang, J.: A unified theorem on SDP rank reduction. Math. Oper. Res. **33**, 910–920 (2008)
28. Sturm, J.F., Zhang, S.Z.: On cones of nonnegative quadratic functions. Math. Oper. Res. **28**, 246–267 (2003)
29. Tseng, P.: Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation. SIAM J. Optim. **14**, 263–283 (2003)
30. Wang, Y., Aron, M.: A reformulation of the strong ellipticity conditions for unconstrained hyperelastic media. J. Elast. **44**, 89–96 (1996)
31. Wang, Y., Qi, L., Zhang, X.: A practical method for computing the largest M-eigenvalue of a fourth-order partially symmetric tensor. Numer. Linear Algebra Appl. **16**, 589–601 (2009)
32. Ye, Y., Zhang, S.Z.: New results on quadratic minimization. SIAM J. Optim. **14**, 245–267 (2003)
33. Zhang, T., Golub, G.H.: Rank-1 approximation of higher-order tensors. SIAM J. Matrix Anal. Appl. **23**, 534–550 (2001)