NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
*Numer. Linear Algebra Appl.* 2012; **19**:113–131
Published online 1 April 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.781

# Semidefinite relaxation approximation for multivariate bi-quadratic optimization with quadratic constraints

Chen Ling[1], Xinzhen Zhang[2,*,†] and Liqun Qi[3]

[1]*School of Science, Hangzhou Dianzi University, Hangzhou, 310018, People's Republic of China*
[2]*Department of Mathematics, School of Science, Tianjin University, Tianjin, 300072, China*
[3]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*

SUMMARY

In this paper, we consider the NP-hard problem of finding global minimum of quadratically constrained multivariate bi-quadratic optimization. We present some bounds of the considered problem via approximately solving the related bi-linear semidefinite programming (SDP) relaxation. Based on the bi-linear SDP relaxation, we also establish some approximation solution methods, which generalize the methods for the quadratic polynomial optimization in (*SIAM J. Optim.* 2003; **14**:268–283). Finally, we present a special form, whose bi-linear SDP relaxation can be approximately solved in polynomial time. Copyright © 2011 John Wiley & Sons, Ltd.

Received 30 March 2010; Revised 5 December 2010; Accepted 12 February 2011

KEY WORDS: bi-quadratic optimization; NP-hard; semidefinite programming relaxation; approximation solution

## 1. INTRODUCTION

We consider the optimization of a bi-quadratic polynomial with quadratic constraints

$$\begin{aligned} \min_{x\in\Re^n, y\in\Re^m} \quad & Q(x,y) := \sum_{i,s=1}^{n}\sum_{j,t=1}^{m} a_{ijst}x_iy_jx_sy_t + \sum_{i=1}^{n}\sum_{j=1}^{m} h_{ij}x_iy_j \\ \text{s.t.} \quad & \Phi_k(x) := x^\top M^k x + (b^k)^\top x + \alpha_k \leqslant 0, \quad k\in\mathscr{P}, \\ & \Psi_l(y) := y^\top N^l y + (d^l)^\top y + \beta_l \leqslant 0, \quad l\in\mathscr{Q}, \end{aligned} \tag{1}$$

where $M^k\in\Re^{n\times n}$ symmetric, $b^k\in\Re^n$, $\alpha_k\in\Re$ for $k\in\mathscr{P}:=\{1,\ldots,p\}$, and $N^l\in\Re^{m\times m}$ symmetric, $d^l\in\Re^m$, $\beta_l\in\Re$ for $l\in\mathscr{Q}:=\{1,\ldots,q\}$. Without loss of generality, we assume that the coefficients $a_{ijst}$ satisfy the partially symmetric property, that is: $a_{ijst}=a_{sjit}=a_{itsj}$ for $i,s=1,\ldots,n$ and $j,t=1,\ldots,m$. Furthermore, throughout this paper, we assume that the feasible set of (1) is nonempty.

It is easy to see that, for fixed $x$ or $y$, the problem (1) reduces to a quadratic optimization problem with quadratic constraints, which was studied in [1]. This motivates us to call (1) a *general bi-quadratic optimization problem*, or a *general bi-quadratic program*. Furthermore, we can assert that (1) is NP-hard since the reduced quadratic optimization problem is NP-hard.

*Correspondence to: Xinzhen Zhang, Department of Mathematics, School of Science, Tianjin University, Tianjin, 300072, China.
†E-mail: xzzhang@yahoo.cn

It is well known that the polynomial optimization is a fundamental problem in optimization. As such, it is widely used in many applications such as signal processing, biomedical engineering, investment science, quantum mechanics, and statistics. Therefore, it has been a priority for many mathematical programmers to establish efficient algorithms for polynomial optimization. In particular, higher order polynomials over quadratic constraints have been studied recently, e.g. see [2–6] for details. As a special case of higher order polynomial optimization, the problem (1) includes the nonhomogeneous bi-quadratic and quadratic functions in its objective and constraints, respectively. In fact, the problem (1) is also a generalization of the bi-quadratic optimization over unit spheres studied in [4], the bi-quadratic optimization with quadratic constraints in [6] and the quadratically constrained quadratic optimization in [1]. The bi-quadratic optimization over unit spheres arises from many applications such as solid mechanics, quantum physics, rank-one approximation to the fourth-order partially symmetric tensor, signal and image processing, wireless communication systems, data analysis, higher order statistics, and independent component analysis, e.g. see [7–20] for details. Many real problems such as the maximum cut problem in combinatorial optimization are the special cases of the model in [1]. On the other hand, the problem (1) also arises directly from portfolio selection, which will be presented in Section 2.

For the bi-quadratic optimization over unit spheres in [4], Ling *et al*. proved that there is no polynomial time algorithm returning a positive relative quality bound and presented various approximation methods based on its semidefinite programming (SDP) relaxations. Recently, for the homogeneous bi-quadratic optimization with quadratic constraints, Zhang *et al*. [6] proved 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 for the original problem, where the constant in $\mathcal{O}(r)$ does not involve the dimension of variables and the data of problems. Notice that SDP relaxation methods are important for approximately solving quadratic optimization problems and have received much attention recently, e.g. [1, 21–27]. Motivated by these, our study for (1) is also based on its SDP relaxations.

Denote $\mathscr{A} := (a_{ijst})$, then $\mathscr{A}$ is a real, fourth-order $(n\times m\times n\times m)$-dimensional partially symmetric tensor. In terms of $\mathscr{A}$, the objective function in (1) can be written as $Q(x,y) = (\mathscr{A}xx^\top)\cdot(yy^\top) + (H^\top x)^\top y$, where $\mathscr{A}xx^\top = (\sum_{i,s=1}^n a_{ijst}x_i x_s)_{1\leqslant j,t\leqslant m}$ is an $m\times m$ symmetric matrix, and $X\cdot Y$ stands for usual matrix inner product, i.e. $X\cdot Y = \mathrm{Tr}(X^\top Y)$. Denote

$$B^k := \begin{bmatrix} M^k & b^k/2 \\ (b^k)^\top/2 & \alpha_k \end{bmatrix} \ (k\in\mathscr{P}) \quad \text{and} \quad C^l := \begin{bmatrix} N^l & d^l/2 \\ (d^l)^\top/2 & \beta_l \end{bmatrix} \ (l\in\mathscr{Q}).$$

It is readily to know that (1) can be written as:

$$\begin{aligned} \min_{x\in\Re^n,\, y\in\Re^m} \quad & \begin{bmatrix} \mathscr{A}xx^\top & \frac{1}{2}H^\top x \\ \frac{1}{2}x^\top H & 0 \end{bmatrix} \cdot \begin{bmatrix} yy^\top & y \\ y^\top & 1 \end{bmatrix} \\ \text{s.t.} \quad & B^k \cdot \begin{bmatrix} xx^\top & x \\ x^\top & 1 \end{bmatrix} \leqslant 0, \quad k\in\mathscr{P}, \\ & C^l \cdot \begin{bmatrix} yy^\top & y \\ y^\top & 1 \end{bmatrix} \leqslant 0, \quad l\in\mathscr{Q}. \end{aligned} \tag{2}$$

Notice that for any symmetric matrix $X$,

$$X = \begin{bmatrix} xx^\top & x \\ x^\top & 1 \end{bmatrix}$$

if and only if

$$X = \begin{bmatrix} X_1 & x \\ x^\top & 1 \end{bmatrix}$$



DOI: 10.1002/nla

and $\mathrm{rank}(X)=1$. Therefore, by relaxing the rank-1 constraints concealed in (2), we obtain the bi-linear SDP relaxation of (1) as follows:

$$\begin{aligned}
\min_{X\in\mathscr{S}^{n+1}, Y\in\mathscr{S}^{m+1}} \quad & \phi(X,Y):=\begin{bmatrix} \mathscr{A}X_1 & \frac{1}{2}H^\top x \\ \frac{1}{2}x^\top H & 0 \end{bmatrix}\cdot Y \\
\text{s.t.} \quad & B^k\cdot X\leqslant 0, \quad k\in\mathscr{P}, \\
& C^l\cdot Y\leqslant 0, \quad l\in\mathscr{Q}, \\
& B^{p+1}\cdot X=1, \quad C^{q+1}\cdot Y=1, \\
& X=\begin{bmatrix} X_1 & x \\ x^\top & u \end{bmatrix}\succeq 0, \quad Y=\begin{bmatrix} Y_1 & y \\ y^\top & v \end{bmatrix}\succeq 0,
\end{aligned} \tag{3}$$

where $B^{p+1}=\mathrm{diag}(\underbrace{0,\ldots,0}_{n},1)$ and $C^{q+1}=\mathrm{diag}(\underbrace{0,\ldots,0}_{m},1)$. Here, $\mathscr{A}X_1$ is the $m\times m$ matrix with

$$(\mathscr{A}X_1)_{jt}=\sum_{i,s=1}^{n} a_{ijst}X_{is}, \quad j,t=1,2,\ldots,m.$$

We denote by $v_{bqp}$ and $\phi_{\min}$ the optimal values of (1) and (3), respectively. It is clear that $\phi_{\min}\leqslant v_{bqp}$, which implies that a lower bound of (1) is obtained, provided that the optimal value $\phi_{\min}$ of (3) has been found. However, the bi-linear SDP relaxation (3) by itself is also NP-hard. The reason for this is that, for the bi-quadratic optimization over unit spheres, a special form of (1), the corresponding bi-linear SDP relaxation and the original problem are equivalent, i.e. they have the same optimal value and one optimal solution pair of the original problem can be obtained from the optimal solution pair of its bi-linear SDP relaxation, see [4]. In this paper, we extend the existing methods in [1] for quadratic optimization problems to general bi-quadratic optimization problem with nonhomogeneous objective and constraint functions. Some obtained results in this paper generalize the corresponding conclusions in [4, 6].

This paper mainly focus on the theoretical analysis of the approximation algorithm and is organized as follows. After an application example in portfolio selection is presented in Section 2, we analyze the relation between (1) with ellipsoid constraints and its bi-linear SDP relaxation in Section 3. In Section 4, we first discuss the approximation solution of (1) in the sense of expectation, then present an approximation method for the partially ellipsoid constraint case in the sense of high probability. In Section 5, we present a special form of (1), whose SDP relaxation problem can be approximately solved in polynomial time.

Some words about the notation. $\Re^n$ denotes the space of real $n$-dimensional column vectors. For $x\in\Re^n$, $x_j$ denotes the $j$th component of $x$. $U_n$ stands for the unit sphere in $\Re^n$, i.e. $U_n:=\{x\in\Re^n \mid \|x\|=1\}$. $\Re^{m\times n}$ denotes the space of real $m\times n$ matrices. For $A\in\Re^{m\times n}$, $A_{ij}$ denotes the $(i,j)$th entry of $A$ and $\|A\|_F$ denotes the Frobenius norm of $A$, i.e. $\|A\|_F=(\mathrm{Tr}(A^\top A))^{1/2}$, where $\mathrm{Tr}(\cdot)$ means the trace of a matrix. $\mathscr{S}^n$ denotes the space of real symmetric $n\times n$ matrices. For $A\in\mathscr{S}^n$, $A\succeq 0$(resp.$A\succ 0$) means that $A$ is positive semidefinite (resp. positive definite). $\mathscr{S}^n_+$ denotes the cone of positive semidefinite matrices in $\mathscr{S}^n$. For $A\in\mathscr{S}^n$ with $|A_{ij}|\leqslant 1$ for all $i$ and $j$, $\arcsin(A)$ denotes the matrix in $\mathscr{S}^n$ with $(i,j)$th entry $\arcsin(A_{ij})$. $I_n$ stands for the identity matrix with $n$ dimension, and $e^k$ stands for the $k$th coordinate vector. In addition, for a given finite set $D$, $\mathrm{Card}(D)$ stands for the cardinality of $D$.

## 2. MOTIVATION: APPLICATION IN PORTFOLIO SELECTION

According to Markowitz's well-known mean-variance model [28], the general single-period portfolio selection problem can be formulated as a parametric convex quadratic program. In this section,



DOI: 10.1002/nla

we present a slightly more involved mean-variance model in portfolio selection problems, which can be reformulated as a general bi-quadratic optimization problem (1).

We consider the portfolio selection problem in two groups of securities, where investment decisions have an influence on each other. Assume that the groups consist of $N$ and $M$ securities, respectively. We assume that the investment on the $i$th security of the first group of securities is further reallocated to two different types of industries $A$ and $B$, according to the special proportion $\sigma_i$. The discounted returns of the industries $A$ and $B$ are denoted by $R_{iA}^{(1)}$ and $R_{iB}^{(1)}$, respectively. Assume that $R_{iA}^{(1)}$ is independent of the relative amount $x_i$ invested in the $i$th security, but dependent on the amount $y_j$ invested in the $j$th security of the second group of security, whereas $R_{iB}^{(1)}$ dependent only on the amount $x_i$. Let $R_{iA}^{(1)}=\xi_{i1}^{(0)}+\xi_{i1}^{(1)}y_1+\cdots+\xi_{iM}^{(1)}y_M$ $(i=1,\ldots,N)$, where $\xi_{i1}^{(0)}$ is a random variable with mean $\mu_i$, and $\xi_{ij}^{(1)}$ $(i=1,\ldots,N, j=1,\ldots,M)$ are random variables with mean zero. Let $R_{iB}^{(1)}=\xi_{i2}^{(0)}+\xi_{i1}^{(2)}x_1+\cdots+\xi_{iN}^{(2)}x_N$ $(i=1,\ldots,N)$, where $\xi_{i2}^{(0)}$ $(i=1,\ldots,N)$ are random variables with mean $\alpha_i$, and $\xi_{ik}^{(2)}$ $(i,k=1,\ldots,N)$ are random variables with mean $g_{ik}$. Then, the return of a portfolio on the industry $A$ in the first group of securities is a random variable defined by

$$R_A^{(1)}=\sum_{i=1}^{N}\sigma_i R_{iA}^{(1)}x_i=\sum_{i=1}^{N}\sigma_i\xi_{i1}^{(0)}x_i+\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_i\xi_{ij}^{(1)}x_iy_j$$

and its expected value is $E(R_A^{(1)})=\mu^\top x$, where $\mu=[\sigma_1\mu_1,\ldots,\sigma_N\mu_N]^\top$ and $x=[x_1,\ldots,x_N]^\top$. The return of a portfolio on the industry $B$ in the first group of securities is a random variable defined by

$$R_B^{(1)}=\sum_{i=1}^{N}(1-\sigma_i)R_{iB}^{(1)}x_i=\sum_{i=1}^{N}(1-\sigma_i)\xi_{i2}^{(0)}x_i+\sum_{i,k=1}^{N}(1-\sigma_i)\xi_{ik}^{(2)}x_ix_k$$

and its expected value is $E(R_B^{(1)})=\alpha^\top x+x^\top Gx$, where $\alpha=[(1-\sigma_1)\alpha_1,\ldots,(1-\sigma_N)\alpha_N]^\top$ and $G=((1-\sigma_i)g_{ik})_{1\leqslant i,k\leqslant N}$. Similarly, we assume that the investment on the $j$th security of the second group of securities is reallocated to two different types of industries $C$ and $D$, according to the special proportion $\rho_j$. By similar reasoning, we obtain that the returns of a portfolio on the industries $C$ and $D$ of the second group of securities are

$$R_C^{(2)}=\sum_{j=1}^{M}\rho_j\gamma_{j1}^{(0)}y_j+\sum_{j=1}^{M}\sum_{i=1}^{N}\rho_j\gamma_{ji}^{(1)}x_iy_j \quad\text{and}\quad R_D^{(2)}=\sum_{j=1}^{M}(1-\rho_j)\gamma_{j2}^{(0)}y_j+\sum_{j,l=1}^{M}(1-\rho_j)\gamma_{jl}^{(2)}y_jy_l,$$

respectively. Here, $\gamma_{j1}^{(0)}$ $(j=1,\ldots,M)$ are random variables with mean $v_j$, $\gamma_{ji}^{(1)}$ $(i=1,\ldots,N, j=1,\ldots,M)$ are the random variables with mean zero, $\gamma_{j2}^{(0)}$ $(j=1,\ldots,M)$ are random variables with mean $\beta_j$, and $\gamma_{jl}^{(2)}$ $(j,l=1,\ldots,N)$ are the random variables with mean $q_{jl}$. It is easy to see that $E(R_C^{(2)})=v^\top y$ and $E(R_D^{(2)})=\beta^\top y+y^\top Qy$, where $v=[\rho_1v_1,\ldots,\rho_Mv_M]^\top$, $\beta=[(1-\rho_M)\beta_1,\ldots,(1-\rho_M)\beta_M]^\top$, $Q=((1-\rho_j)q_{jl})_{1\leqslant j,l\leqslant M}$ and $y=[y_1,\ldots,y_M]^\top$. It is clear that the total return of the portfolio on the industries $A$ and $C$ is $R_{AC}=R_A^{(1)}+R_C^{(2)}$. We assume that the random variables $\xi_{i1}^{(0)}$, $\xi_{ij}^{(1)}$, $\gamma_{j1}^{(0)}$, and $\gamma_{ji}^{(1)}$ are independent of each other for $i=1,\ldots,N$ and $j=1,\ldots,M$. Under this assumption, we know that the variance of $R_{AC}$ is $\mathrm{Var}(R_{AC})=\mathrm{Var}(R_A^{(1)})+\mathrm{Var}(R_C^{(2)})$.

Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be the variance tensors of the random matrices $\Xi=(\sigma_i\xi_{ij}^{(1)})$ and $\Gamma=(\rho_j\gamma_{ji}^{(1)})$, respectively, and $P_1$ and $P_2$ be the variance matrices of the random vectors $\xi^{(0)}=[\sigma_1\xi_{11}^{(0)},\ldots,\sigma_N\xi_{N1}^{(0)}]^\top$ and $\gamma^{(0)}=[\rho_1\gamma_{11}^{(0)},\ldots,\rho_M\gamma_{M1}^{(0)}]^\top$, respectively. If we consider the portfolio selection problem associated with the industries $A$ and $C$ under the condition that the returns on the industries $B$ and $D$ reach at least the given acceptable values $g$ and $h$, respectively, then, given



DOI: 10.1002/nla

a set of values for the parameter $\tau$ as well as $\mathscr{B}_1$, $\mathscr{B}_2$, $P_1$, and $P_2$, a generalized mean-variance model can be expressed by

$$\begin{aligned}\min_{x\in\Re^N,\, y\in\Re^M}\quad & (\mathscr{B}_1 xx^\top)\cdot yy^\top + (\mathscr{B}_2 xx^\top)\cdot yy^\top + x^\top P_1 x + y^\top P_2 y - \tau(\mu^\top x + \nu^\top y)\\ \text{s.t.}\quad & \alpha^\top x + x^\top G x \geqslant g, \quad \beta^\top y + y^\top Q y \geqslant h,\\ & \sum_{i=1}^{N} x_i = a, \quad \sum_{j=1}^{M} y_j = b,\end{aligned}$$

where $a$ and $b$ stand for the total amount invested in the first and the second group of securities, respectively. It is evident that the above model can be rewritten equivalently as the form of (1).

## 3. SDP RELAXATION OF (1) WITH ELLIPSOID CONSTRAINTS

In this section, we consider the following special form of (1):

$$\begin{aligned}\min_{x\in\Re^n,\, y\in\Re^m}\quad & Q(x,y) := \sum_{i,s=1}^{n}\sum_{j,t=1}^{m} a_{ijst} x_i y_j x_s y_t + \sum_{i=1}^{n}\sum_{j=1}^{m} h_{ij} x_i y_j\\ \text{s.t.}\quad & \|F^k x + f^k\|^2 \leqslant \mu_k, \quad k\in\mathscr{P},\\ & \|G^l y + g^l\|^2 \leqslant \eta_l, \quad l\in\mathscr{Q},\end{aligned} \tag{4}$$

where $F^k\in\Re^{n\times n}$, $f^k\in\Re^n$, $\mu_k\in\{0,1\}$ for $k\in\mathscr{P}$, and $G^l\in\Re^{m\times m}$, $g^l\in\Re^m$, $\eta_l\in\{0,1\}$ for $l\in\mathscr{Q}$. In this case, $M^k=(F^k)^\top F^k$, $b^k=2(F^k)^\top f^k$ and $\alpha_k=\|f^k\|^2-\mu_k$ for $k\in\mathscr{P}$, and $N^l=(G^l)^\top G^l$, $d^l=2(G^l)^\top g^l$ and $\beta_l=\|g^l\|^2-\eta_l$ for $l\in\mathscr{Q}$.

Let $(\bar{X},\bar{Y})$ be a feasible solution pair of the SDP relaxation of (4). In this section our main task is to generate a feasible solution $(\bar{x},\bar{y})$ of the problem (4) from $(\bar{X},\bar{Y})$, such that

$$Q(\bar{x},\bar{y}) \leqslant \frac{(1-\gamma_f)^2(1-\gamma_g)^2}{(\sqrt{\tau_\mu}+\gamma_f)^2(\sqrt{\tau_\eta}+\gamma_g)^2}\,\phi(\bar{X},\bar{Y}), \tag{5}$$

where $\tau_\mu := \mathrm{Card}\{k\in\mathscr{P} \mid \mu_k=1\}$, $\tau_\eta := \mathrm{Card}\{l\in\mathscr{Q} \mid \eta_l=1\}$, $\gamma_f := \max_{k:\mu_k=1}\|f^k\|$, and $\gamma_g := \max_{l:\eta_l=1}\|g^l\|$. To this end, we need the following assumption, which is a simple generalization of the corresponding assumption in [1].

*Assumption 1*
The origin $(0,0)\in\Re^n\times\Re^m$ is a feasible solution of (4). Furthermore, for any $k\in\mathscr{P}$ and $l\in\mathscr{Q}$, there hold $\|f^k\|^2<\mu_k$ whenever $\mu_k=1$ and $\|g^l\|^2<\eta_l$ whenever $\eta_l=1$.

The following two lemmas are well-known. The first lemma is due to Sturm and Zhang [29] and the second lemma was proved by Tseng [1].

*Lemma 3.1*
Let $X\in\mathscr{S}^n$ be a positive semidefinite matrix of rank $r$. Let $B\in\mathscr{S}^n$. Then, $B\cdot X\leqslant 0$ if and only if there exist $w^j\in\Re^n$, $j=1,\dots,r$, such that

$$X=\sum_{j=1}^{r} w^j (w^j)^\top \quad \text{and} \quad (w^j)^\top B w^j \leqslant 0, \quad j=1,\dots,r.$$

*Lemma 3.2*
For any scalars $\kappa\geqslant 0$, $\alpha_j\geqslant 0$ and $\beta_j\geqslant 0$, $j=1,\dots,r$ $(r\geqslant 1)$, such that $\sum_{j=1}^{r}\alpha_j\leqslant\kappa$ and $\sum_{j=1}^{r}\beta_j=1$, there exists $\bar{j}\in\{1,\dots,r\}$ such that $\beta_{\bar{j}}>0$ and $\alpha_{\bar{j}}/\beta_{\bar{j}}\leqslant\kappa$.



DOI: 10.1002/nla

We now state our main result in this section, whose proof is similar to that of Theorem 1 in [4]. However, for the sake of completeness, we still present its proof.

*Theorem 3.1*
Let $(\bar{X}, \bar{Y})$ be a feasible solution pair of the bi-linear SDP relaxation (3) of (4). Then there exists a feasible solution $(\bar{x}, \bar{y})$ of (4) satisfying (5).

*Proof*
Let

$$\bar{X} = \begin{bmatrix} \hat{X}_1 & \hat{x} \\ \hat{x}^\top & \hat{u} \end{bmatrix}, \quad \bar{Y} = \begin{bmatrix} \hat{Y}_1 & \hat{y} \\ \hat{y}^\top & \hat{v} \end{bmatrix}, \quad \bar{\phi} = \phi(\bar{X}, \bar{Y}),$$

and $\bar{A} \in \mathscr{S}^{m+1}$ defined by:

$$\bar{A} := \begin{bmatrix} \mathscr{A}\hat{X}_1 & \frac{1}{2}H^\top \hat{x} \\ \frac{1}{2}\hat{x}^\top H & 0 \end{bmatrix} - \bar{\phi} C^{q+1}.$$

Then it follows that $\bar{A} \cdot \bar{Y} = 0$ from $C^{q+1} \cdot \bar{Y} = 1$. By Lemma 3.1, there exist $z^j = ((y^j)^\top, v_j)^\top$ $(j = 1, \ldots, m+1) \in \Re^m \times \Re$, such that

$$\bar{Y} = \sum_{j=1}^{m+1} z^j (z^j)^\top \quad \text{and} \quad (z^j)^\top \bar{A} z^j \leqslant 0 \quad \text{for } j = 1, \ldots, m+1,$$

which implies that

$$(z^j)^\top \begin{bmatrix} \mathscr{A}\hat{X}_1 & \frac{1}{2}H^\top \hat{x} \\ \frac{1}{2}\hat{x}^\top H & 0 \end{bmatrix} z^j \leqslant \bar{\phi} v_j^2 \quad \text{for } j = 1, \ldots, m+1. \tag{6}$$

Since $C^l \cdot \bar{Y} \leqslant 0$ for $l \in \mathscr{Q}$ and $C^{q+1} \cdot \bar{Y} = 1$, we have that $\sum_{j=1}^{m+1} v_j^2 = 1$ and

$$\sum_{j=1}^{m+1} ((y^j)^\top (G^l)^\top G^l y^j + 2(g^l)^\top G^l y^j v_j + (\|g^l\|^2 - \eta_l) v_j^2) = C^l \cdot \bar{Y} \leqslant 0 \quad \text{for } l \in \mathscr{Q}.$$

Consequently, it follows that

$$\sum_{j=1}^{m+1} \|G^l y^j + v_j g^l\|^2 \leqslant \eta_l, \quad \text{for } l \in \mathscr{Q}, \tag{7}$$

which implies that

$$\sum_{j=1}^{m+1} \sum_{l:\eta_l = 1} \|G^l y^j + v_j g^l\|^2 \leqslant \tau_\eta.$$

By Lemma 3.2, there exists an index $\bar{j} \in \{1, \ldots, m+1\}$, such that

$$v_{\bar{j}}^2 > 0 \quad \text{and} \quad \sum_{l:\eta_l = 1} \|G^l y^{\bar{j}} + v_{\bar{j}} g^l\|^2 / v_{\bar{j}}^2 \leqslant \tau_\eta.$$

Moreover, we can choose $\bar{j}$ to minimize the ration $\sum_{l:\eta_l=1} \|G^l y^j + v_j g^l\|^2 / v_j^2$ over all $j$ with $v_j^2 > 0$. Thus, $\|G^l y^{\bar{j}} / v_{\bar{j}} + g^l\| \leqslant \sqrt{\tau_\eta}$ whenever $\eta_l = 1$.

Define

$$\tilde{y} := \begin{cases} y^{\bar{j}} / v_{\bar{j}} & \text{if } \hat{x}^\top H y^{\bar{j}} / v_{\bar{j}} \leqslant 0, \\ -y^{\bar{j}} / v_{\bar{j}} & \text{if } \hat{x}^\top H y^{\bar{j}} / v_{\bar{j}} > 0, \end{cases}$$



DOI: 10.1002/nla

and $\tilde{\sigma} := \max\{\sigma \in [0,1] | \|G^l(\sigma\tilde{y}) + g^l\|^2 \leqslant \eta_l,\ l \in \mathscr{Q}\}$. In the case where $\eta_l = 0$, from (7) and Assumption 1, we know that $\|G^l y^{\bar{j}}\| = 0$ which implies that $\|G^l(\sigma\tilde{y}) + g^l\|^2 \leqslant \eta_l$ for all $\sigma \in [0,1]$. In the case where $\eta_l = 1$, we know that if $\hat{x}^\top H y^{\bar{j}}/v_{\bar{j}} \leqslant 0$, then $\|G^l\tilde{y} + g^l\| \leqslant \sqrt{\tau_\eta}$, and otherwise $\|G^l\tilde{y} + g^l\| \leqslant \|G^l y^{\bar{j}}/v_{\bar{j}} + g^l\| + 2\|g^l\| \leqslant \sqrt{\tau_\eta} + 2\|g^l\|$. Hence, for any $\sigma \in [0,1]$

$$\|G^l(\sigma\tilde{y}) + g^l\| = \|\sigma(G^l\tilde{y} + g^l) + (1-\sigma)g^l)\| \leqslant \sigma(\sqrt{\tau_\eta} + 2\|g^l\|) + (1-\sigma)\|g^l\|.$$

Since $\|g^l\| < 1$, it holds that $\|G^l(\sigma\tilde{y}) + g^l\| \leqslant 1$ whenever $\sigma \leqslant (1 - \|g^l\|)/(\sqrt{\tau_\eta} + \|g^l\|)$. Therefore,

$$\tilde{\sigma} \geqslant \min_{l:\eta_l = 1} \frac{1 - \|g^l\|}{\sqrt{\tau_\eta} + \|g^l\|} = \frac{1 - \max_{l:\eta_l=1}\|g^l\|}{\sqrt{\tau_\eta} + \max_{l:\eta_l=1}\|g^l\|},$$

where the equality follows from $(1-\lambda)/(\sqrt{\tau_\eta} + \lambda)$ being a decreasing function on $\lambda \in [0,1)$. Let

$$\bar{\sigma} = \frac{1 - \max_{l:\eta_l=1}\|g^l\|}{\sqrt{\tau_\eta} + \max_{l:\eta_l=1}\|g^l\|}$$

and $\bar{y} = \bar{\sigma}\tilde{y}$. It is clear that

$$\|G^l\bar{y} + g^l\|^2 \leqslant \eta_l \quad \text{for } l \in \mathscr{Q}. \tag{8}$$

Moreover, by the choice of $\tilde{y}$, we know that $\hat{x}^\top H\tilde{y} \leqslant 0$ and $\hat{x}^\top H\tilde{y} \leqslant \hat{x}^\top H y^{\bar{j}}/v_{\bar{j}}$. Consequently, it holds that

$$\begin{aligned} \bar{y}^\top(\mathscr{A}\hat{X}_1)\bar{y} + \hat{x}^\top H\bar{y} &= \bar{\sigma}^2\tilde{y}^\top(\mathscr{A}\hat{X}_1)\tilde{y} + \bar{\sigma}\hat{x}^\top H\tilde{y} \\ &\leqslant \bar{\sigma}^2(\tilde{y}^\top(\mathscr{A}\hat{X}_1)\tilde{y} + \hat{x}^\top H\tilde{y}) \\ &\leqslant \bar{\sigma}^2((y^{\bar{j}})^\top(\mathscr{A}\hat{X}_1)y^{\bar{j}} + \hat{x}^\top H y^{\bar{j}} v_{\bar{j}})/v_{\bar{j}}^2 \\ &\leqslant \bar{\sigma}^2\bar{\phi}, \end{aligned}$$

which implies

$$\begin{bmatrix} \bar{y}\bar{y}^\top\mathscr{A} & \frac{1}{2}H\bar{y} \\ \frac{1}{2}(H\bar{y})^\top & 0 \end{bmatrix} \cdot \bar{X} \leqslant \bar{\sigma}^2\bar{\phi}, \tag{9}$$

where $\bar{y}\bar{y}^\top\mathscr{A} = (\sum_{j,t=1}^m a_{ijst}\bar{y}_j\bar{y}_t)_{1\leqslant i,s\leqslant n}$ is an $n \times n$ symmetric matrix. Let $A_{\bar{y}} \in \mathscr{S}^{n+1}$ defined by:

$$A_{\bar{y}} = \begin{bmatrix} \bar{y}\bar{y}^\top\mathscr{A} & \frac{1}{2}H\bar{y} \\ \frac{1}{2}(H\bar{y})^\top & 0 \end{bmatrix} - \bar{\sigma}^2\bar{\phi}B^{p+1}.$$

Then, by (9) and the fact that $B^{p+1} \cdot \bar{X} = 1$, it follows that $A_{\bar{y}} \cdot \bar{X} \leqslant 0$. Applying Lemma 3.1 to $\bar{X}$ and $A_{\bar{y}}$ again, we can find $w^i = ((x^i)^\top, u_i)^\top \in \Re^n \times \Re$, $i = 1, \ldots, n+1$, such that

$$\bar{X} = \sum_{i=1}^{n+1} w^i(w^i)^\top \quad \text{and} \quad (w^i)^\top A_{\bar{y}} w^i \leqslant 0 \quad \text{for } i = 1, \ldots, n+1,$$

which implies

$$(w^i)^\top \begin{bmatrix} \bar{y}\bar{y}^\top\mathscr{A} & \frac{1}{2}H\bar{y} \\ \frac{1}{2}(H\bar{y})^\top & 0 \end{bmatrix} w^i \leqslant \bar{\sigma}^2\bar{\phi}u_i^2 \quad \text{for } i = 1, \ldots, n+1. \tag{10}$$



DOI: 10.1002/nla

Similarly, there exists an index $\bar{i}\in\{1,\ldots,n+1\}$, such that

$$u_{\bar{i}}^2>0 \quad \text{and} \quad \sum_{k:\mu_k=1}\|F^k x^{\bar{i}}+u_{\bar{i}}f^k\|^2/u_{\bar{i}}^2\leqslant\tau_\mu.$$

Moreover, denote

$$\tilde{x}:=\begin{cases} x^{\bar{i}}/u_{\bar{i}} & \text{if } \bar{y}^\top H^\top x^{\bar{i}}/u_{\bar{i}}\leqslant 0,\\ -x^{\bar{i}}/u_{\bar{i}} & \text{if } \bar{y}^\top H^\top x^{\bar{i}}/u_{\bar{i}}>0 \end{cases}$$

and $\bar{x}:=\tilde{\rho}\tilde{x}$, where

$$\tilde{\rho}=\frac{1-\max_{k:\mu_k=1}\|f^k\|}{\sqrt{\tau_\mu}+\max_{k:\mu_k=1}\|f^k\|}.$$

Thus, arguing identically as in the proof of (8), we have that $\|F^k\bar{x}+f^k\|^2\leqslant\mu_k$ for $k\in\mathscr{P}$, which implies, together with (8), that $(\bar{x},\bar{y})$ is a feasible solution of (4). Moreover, from the choice of $\bar{x}$, arguing similarly as in the proof of (9), we can obtain

$$[\bar{x}^\top,1]\begin{bmatrix}\bar{y}\bar{y}^\top\mathscr{A} & \frac{1}{2}H\bar{y}\\ \frac{1}{2}(H\bar{y})^\top & 0\end{bmatrix}\begin{bmatrix}\bar{x}\\1\end{bmatrix}\leqslant\tilde{\sigma}^2\tilde{\rho}^2\bar{\phi}.$$

That is, $Q(\bar{x},\bar{y})=\bar{x}^\top(\bar{y}\bar{y}^\top\mathscr{A})\bar{x}+\bar{x}^\top H\bar{y}\leqslant(\tilde{\sigma}\tilde{\rho})^2\phi(\bar{X},\bar{Y})$. We obtain (5) and complete the proof. □

From Theorem 3.1, we have

*Corollary 3.1*
Let $(X^*,Y^*)$ be an optimal solution pair of the bi-linear SDP relaxation (3) of (4). Then there exists a feasible solution $(\bar{x},\bar{y})$ of (4) satisfying

$$Q(\bar{x},\bar{y})\leqslant\frac{(1-\gamma_f)^2(1-\gamma_g)^2}{(\sqrt{\tau_\mu}+\gamma_f)^2(\sqrt{\tau_\eta}+\gamma_g)^2}\phi_{\min}.$$

Specially, when $f^k=0$, $\mu_k=1$ for all $k\in\mathscr{P}$ and $g^l=0$, $\eta_l=1$ for all $l\in\mathscr{Q}$, the problem (4) is a special case of the maximization model in [6]. For this case, it holds that $\gamma_f=\gamma_g=0, \tau_\eta=p$ and $\tau_\mu=q$, and a feasible solution $(\bar{x},\bar{y})$ such that $Q(\bar{x},\bar{y})\leqslant(1/pq)\phi_{\min}$ can be generated by a deterministic way, which is more practicable than the way presented in [6].

## 4. APPROXIMATION SOLUTION OF (1)

In this section, we study approximation solutions of (1) in the sense of expectation, based upon its bi-linear SDP relaxation. Denote

$$\mathscr{P}_0=\{k\in\mathscr{P}|M^k \text{ is diagonal and } b^k=0\},\quad \mathscr{Q}_0=\{l\in\mathscr{Q}|N^l \text{ is diagonal and } d^l=0\}$$

and make the following assumption.

*Assumption 2*
$\{x\in\Re^n|\Phi_k(x)\leqslant 0,\ k\in\mathscr{P}_0\}$ and $\{y\in\Re^m|\Psi_l(x)\leqslant 0,\ l\in\mathscr{Q}_0\}$ are both bounded.

Let $(\bar{X},\bar{Y})$ be a feasible solution of (3) with objective value $\bar{\phi}=\phi(\bar{X},\bar{Y})$. Since $\bar{X}\succeq 0$ and $\bar{Y}\succeq 0$, there exist two factorization matrices $U=[u^1,\ldots,u^{n+1}]\in\Re^{(n+1)\times(n+1)}$ and $V=[v^1,\ldots,v^{m+1}]\in$



DOI: 10.1002/nla

$\Re^{(m+1)\times(m+1)}$, such that $\bar{X}=U^\top U$ and $\bar{Y}=V^\top V$. It is clear that $u^{n+1}\in U_{n+1}$ and $v^{m+1}\in U_{m+1}$, since $\|u^{n+1}\|^2=B^{p+1}\cdot\bar{X}=1$ and $\|v^{m+1}\|^2=C^{q+1}\cdot\bar{Y}=1$. Motivated by the generation method of single random vector in [1, 30, 31], we choose independently two random vectors $u$ and $v$ uniformly distributed on $U_{n+1}$ and $U_{m+1}$, respectively. For the obtained $u$, if $u^\top u^{n+1}\leqslant 0$, then we set $\hat{x}=(\hat{x}_1,\ldots,\hat{x}_{n+1})^\top$ with

$$\hat{x}_i=\begin{cases}\sqrt{\bar{X}_{ii}} & \text{if } u^\top u^i\leqslant 0,\\ -\sqrt{\bar{X}_{ii}} & \text{otherwise,}\end{cases}\qquad i=1,\ldots,n+1.$$

If $u^\top u^{n+1}>0$, then set $\hat{x}=(\hat{x}_1,\ldots,\hat{x}_{n+1})^\top$ with

$$\hat{x}_i=\begin{cases}-\sqrt{\bar{X}_{ii}} & \text{if } u^\top u^i\leqslant 0,\\ \sqrt{\bar{X}_{ii}} & \text{otherwise,}\end{cases}\qquad i=1,\ldots,n+1.$$

Hence, we denote the $(n+1)$ dimension random vector $\hat{x}$ with $\hat{x}_{n+1}=1$. Similarly, for the obtained $v$, we can generate an $(m+1)$ dimension random variable $\hat{y}$ with $\hat{y}_{m+1}=1$. Notice that $\hat{x}$ and $\hat{y}$ are independent from the independence of $u$ and $v$.

We now consider the following standard SDP problems:

$$\begin{aligned}\bar{\phi}_{sdp}:=\max_X\ & \begin{bmatrix}\bar{Y}_1\mathscr{A} & \frac{1}{2}H\bar{y}\\ \frac{1}{2}(H\bar{y})^\top & 0\end{bmatrix}\cdot X\\ \text{s.t.}\ & B^k\cdot X=\bar{b}_k,\quad k\in\mathscr{P}_0,\\ & B^{p+1}\cdot X=1,\quad X\succeq 0\end{aligned}\tag{11}$$

and

$$\begin{aligned}\bar{\varphi}_{sdp}:=\max_Y\ & \begin{bmatrix}\mathscr{A}E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}}\\ \frac{1}{2}E_{\hat{x}}^\top H & 0\end{bmatrix}\cdot Y\\ \text{s.t.}\ & C^l\cdot Y=\bar{c}_l,\quad l\in\mathscr{Q}_0,\\ & C^{q+1}\cdot Y=1,\quad Y\succeq 0,\end{aligned}\tag{12}$$

where $\bar{Y}_1\mathscr{A}$ is the $n\times n$ matrix with $(\bar{Y}_1\mathscr{A})_{is}=\sum_{j,t=1}^m b_{ijst}\bar{Y}_{jt}$ $(i,s=1,2,\ldots,n)$, $\bar{y}\in\Re^m$ consist of the first $m$ components in the last column of $\bar{Y}$, $E_{\hat{x}\hat{x}}\in\mathscr{S}^n$ consist of the first $n$ columns and the first $n$ rows of $E[\hat{x}\hat{x}^\top]$, $E_{\hat{x}}\in\Re^n$ consist of the first $n$ components in $E[\hat{x}]$, $\bar{b}_k=B^k\cdot\bar{X}\leqslant 0$ and $\bar{c}_l=C^l\cdot\bar{Y}\leqslant 0$. Here, $E[\hat{x}\hat{x}^\top]$ and $E[\hat{x}]$ are the expectations of $\hat{x}\hat{x}^\top$ and $\hat{x}$, respectively.

Now we are ready to state and prove the following theorem, which implies that, if we obtain an approximation solution $(\bar{X},\bar{Y})$ of (3), then by the selection process described above, an random vector pair $(\hat{x},\hat{y})$ can be found, which is an approximation solution of (1) in expectation. Before proceeding, we need the following technical lemma which was proved by Nesterov [32].

*Lemma 4.1*
Let $X\succcurlyeq 0$ and $X_{ii}\leqslant 1$ for every $i=1,\ldots,n$. Then $\arcsin(X)\succcurlyeq X$.



DOI: 10.1002/nla

*Theorem 4.1*
Let $(\bar{X}, \bar{Y})$ be a feasible solution of (3) with objective value $\bar{\phi} = \phi(\bar{X}, \bar{Y})$. If Assumption 2 holds, then the random vector $(\hat{x}, \hat{y})$ generated by the pair selection process described above, satisfies

$$\begin{aligned} &E[\Phi_k(\hat{x})] \leqslant 0 \quad \text{if } k \in \mathscr{P}_0, \\ &E[\Phi_k(\hat{x})] \leqslant \left(1 - \frac{2}{\pi}\right) \phi^k_{sdp} \quad \text{if } k \in \mathscr{P} \backslash \mathscr{P}_0, \\ &E[\Psi_l(\hat{y})] \leqslant 0 \quad \text{if } l \in \mathscr{Q}_0, \\ &E[\Psi_l(\hat{y})] \leqslant \left(1 - \frac{2}{\pi}\right) \varphi^l_{sdp} \quad \text{if } l \in \mathscr{Q} \backslash \mathscr{Q}_0 \end{aligned} \tag{13}$$

and

$$E[Q(\hat{x}, \hat{y})] \leqslant \frac{4}{\pi^2} \phi(\bar{X}, \bar{Y}) + \frac{2}{\pi}\left(1 - \frac{2}{\pi}\right) \bar{\phi}_{sdp} + \left(1 - \frac{2}{\pi}\right) \bar{\varphi}_{sdp}. \tag{14}$$

Here, $\phi^i_{sdp}$ $(i \in \mathscr{P} \backslash \mathscr{P}_0)$ and $\varphi^j_{sdp}$ $(j \in \mathscr{Q} \backslash \mathscr{Q}_0)$ are the optimal values of the following SDP problems:

$$\begin{aligned} \phi^i_{sdp} := \max_X \quad & B^i \cdot X \\ \text{s.t.} \quad & B^k \cdot X = \bar{b}_k, \quad k \in \mathscr{P}_0, \\ & B^{p+1} \cdot X = 1, \quad X \succeq 0 \end{aligned} \tag{15}$$

and

$$\begin{aligned} \varphi^j_{sdp} := \max_Y \quad & C^j \cdot Y \\ \text{s.t.} \quad & C^l \cdot Y = \bar{c}_l, \quad l \in \mathscr{Q}_0, \\ & C^{q+1} \cdot Y = 1, \quad Y \succeq 0, \end{aligned} \tag{16}$$

respectively.

*Proof*
The proof of (13) is the same as that in [1] and is omitted here. Now we prove (14). Since $|\hat{x}_i \hat{x}_s| = \sqrt{\bar{X}_{ii} \bar{X}_{ss}}$ for $i, s = 1, \ldots, n+1$, it is easy to see that, if $\bar{X}_{ii} \bar{X}_{ss} \neq 0$, then $\hat{x}_i \hat{x}_s = \sqrt{\bar{X}_{ii} \bar{X}_{ss}}$ if and only if $u^\top u^i$ and $u^\top u^s$ have same sign. By Lemma 3.2 of Goemans and Williamson [30], the probability that this event occurs is

$$p = 1 - \frac{1}{\pi} \arccos\left(\frac{(u^i)^\top u^s}{\|u^i\| \|u^s\|}\right) = 1 - \frac{1}{\pi} \arccos\left(\frac{\bar{X}_{is}}{\sqrt{\bar{X}_{ii} \bar{X}_{ss}}}\right).$$

Consequently, it follows that

$$E[\hat{x}_i \hat{x}_s] = \sqrt{\bar{X}_{ii} \bar{X}_{ss}}\, p + \left(-\sqrt{\bar{X}_{ii} \bar{X}_{ss}}\right)(1 - p) = \frac{2}{\pi} \sqrt{\bar{X}_{ii} \bar{X}_{ss}} \arcsin\left(\frac{\bar{X}_{is}}{\sqrt{\bar{X}_{ii} \bar{X}_{ss}}}\right).$$

This indicates that the expectation of $\hat{x}\hat{x}^\top$ is

$$E[\hat{x}\hat{x}^\top] = \frac{2}{\pi} D_{\bar{X}} \arcsin\left(D^{-1}_{\bar{X}} \bar{X} D^{-1}_{\bar{X}}\right) D_{\bar{X}}, \tag{17}$$

where $D_{\bar{X}} = \mathrm{diag}(\sqrt{\bar{X}_{11}}, \ldots, \sqrt{\bar{X}_{nn}}, 1)$. Similarly, it can be proved that

$$E[\hat{y}\hat{y}^\top] = \frac{2}{\pi} D_{\bar{Y}} \arcsin(D^{-1}_{\bar{Y}} \bar{Y} D^{-1}_{\bar{Y}}) D_{\bar{Y}}, \tag{18}$$



DOI: 10.1002/nla

where $D_{\bar{Y}} = \mathrm{diag}(\sqrt{\bar{Y}_{11}}, \ldots, \sqrt{\bar{Y}_{mm}}, 1)$. From (18) and the independence of $\hat{x}$ and $\hat{y}$, we have

$$E[Q(\hat{x},\hat{y})] = \begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot E[\hat{y}\hat{y}^\top]$$
$$= \frac{2}{\pi}\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot (D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1}\bar{Y}D_{\bar{Y}}^{-1})D_{\bar{Y}}). \tag{19}$$

On the other hand, it is easy to know that the dual of (12) is

$$\begin{aligned} \bar{\varphi}_{sdp} = \inf_z \quad & \sum_{l\in\mathscr{Q}_0} \bar{c}_l z_l + z_{q+1} \\ \text{s.t.} \quad & -\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \dfrac{1}{2}H^\top E_{\hat{x}} \\ \dfrac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} + \sum_{l\in\mathscr{Q}_0\cup\{q+1\}} C^l z_l \succeq 0. \end{aligned} \tag{20}$$

For every $\varepsilon > 0$, there exists a feasible solution $\{z_l\}_{l\in\mathscr{Q}_0\cup\{q+1\}}$ of (20) such that

$$\sum_{l\in\mathscr{Q}_0} \bar{c}_l z_l + z_{q+1} \leqslant \bar{\varphi}_{sdp} + \varepsilon. \tag{21}$$

From Lemma 4.1, it follows that $D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1}\bar{Y}D_{\bar{Y}}^{-1})D_{\bar{Y}} \succeq \bar{Y}$. Together with the fact that

$$\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} - \sum_{l\in\mathscr{Q}_0\cup\{q+1\}} C^l z_l \preceq 0,$$

there holds

$$\left(\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} - \sum_{l\in\mathscr{Q}_0\cup\{q+1\}} C^l z_l\right) \cdot (D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1}\bar{Y}D_{\bar{Y}}^{-1})D_{\bar{Y}})$$
$$\leqslant \left(\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} - \sum_{l\in\mathscr{Q}_0\cup\{q+1\}} C^l z_l\right) \cdot \bar{Y}.$$

By this, it holds that

$$\begin{aligned} & \begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot (D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1}\bar{Y}D_{\bar{Y}}^{-1})D_{\bar{Y}}) \\ & \leqslant \left(\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} - \sum_{l\in\mathscr{Q}_0\cup\{q+1\}} C^l z_l\right) \cdot \bar{Y} \\ & \quad + \left(\sum_{l\in\mathscr{Q}_0\cup\{q+1\}} C^l z_l\right) \cdot (D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1}\bar{Y}D_{\bar{Y}}^{-1})D_{\bar{Y}}) \\ & = \begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot \bar{Y} + \left(\frac{\pi}{2} - 1\right) \sum_{l\in\mathscr{Q}_0\cup\{q+1\}} z_l C^l \cdot \bar{Y} \\ & \leqslant \begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2}H^\top E_{\hat{x}} \\ \frac{1}{2}E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot \bar{Y} + \left(\frac{\pi}{2} - 1\right)(\bar{\varphi}_{sdp} + \varepsilon), \end{aligned} \tag{22}$$



DOI: 10.1002/nla

where the first equality comes from the observations that $C^l$ is diagonal for $l \in \mathcal{Q}_0 \cup \{q+1\}$ and that $D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1} \bar{Y} D_{\bar{Y}}^{-1}) D_{\bar{Y}}$ has diagonal entries $(\pi/2)\bar{Y}_{ll}$ for all $l$, and the last inequality comes from the fact that $\bar{c}_l = C^l \cdot \bar{Y}$ for $l \in \mathcal{Q}_0$, $C^{q+1} \cdot \bar{Y} = 1$ and (21). By taking $\varepsilon \to 0$, one obtain

$$\begin{aligned}
&\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2} H^\top E_{\hat{x}} \\ \frac{1}{2} E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot (D_{\bar{Y}} \arcsin(D_{\bar{Y}}^{-1} \bar{Y} D_{\bar{Y}}^{-1}) D_{\bar{Y}}) \\
&\quad \leqslant \begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2} H^\top E_{\hat{x}} \\ \frac{1}{2} E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot \bar{Y} + \left(\frac{\pi}{2} - 1\right) \bar{\varphi}_{sdp}.
\end{aligned} \tag{23}$$

Moreover, it is easy to see that

$$\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2} H^\top E_{\hat{x}} \\ \frac{1}{2} E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot \bar{Y} = \begin{bmatrix} \bar{Y}_1 \mathscr{A} & \frac{1}{2} H \bar{y} \\ \frac{1}{2} (H \bar{y})^\top & 0 \end{bmatrix} \cdot E[\hat{x}\hat{x}^\top].$$

From this and (17), we can similarly prove that

$$\begin{aligned}
\begin{bmatrix} \mathscr{A} E_{\hat{x}\hat{x}} & \frac{1}{2} H^\top E_{\hat{x}} \\ \frac{1}{2} E_{\hat{x}}^\top H & 0 \end{bmatrix} \cdot \bar{Y} &= \frac{2}{\pi} \begin{bmatrix} \bar{Y}_1 \mathscr{A} & \dfrac{1}{2} H \bar{y} \\ \dfrac{1}{2} (H \bar{y})^\top & 0 \end{bmatrix} \cdot D_{\bar{X}} \arcsin(D_{\bar{X}}^{-1} \bar{X} D_{\bar{X}}^{-1}) D_{\bar{X}} \\
&\leqslant \frac{2}{\pi} \begin{bmatrix} \bar{Y}_1 \mathscr{A} & \dfrac{1}{2} H \bar{y} \\ \dfrac{1}{2} (H \bar{y})^\top & 0 \end{bmatrix} \cdot \bar{X} + \left(1 - \frac{2}{\pi}\right) \bar{\phi}_{sdp} \\
&= \frac{2}{\pi} \phi(\bar{X}, \bar{Y}) + \left(1 - \frac{2}{\pi}\right) \bar{\phi}_{sdp}.
\end{aligned} \tag{24}$$

By combining (19), (23) and (24), we have

$$E[Q(\hat{x}, \hat{y})] \leqslant \frac{4}{\pi^2} \phi(\bar{X}, \bar{Y}) + \frac{2}{\pi}\left(1 - \frac{2}{\pi}\right) \bar{\phi}_{sdp} + \left(1 - \frac{2}{\pi}\right) \bar{\varphi}_{sdp}.$$

We obtain the desired result and complete the proof. □

The vector pair $(\bar{x}, \bar{y})$ described in Theorem 4.1 is an approximation solution of (1), but maybe infeasible. To overcome this drawback, for some special forms of (1), we will present a method for finding feasible solution with high probability, based on the method presented above.

In what follows, we consider the following case where the constraints not indexed by $\mathscr{P}_0$ and $\mathcal{Q}_0$ are ellipsoid constraints, i.e.

$$\begin{aligned}
\Phi_k(x) &= \|F^k x + f^k\| - 1 \quad \text{for } k \in \mathscr{P} \backslash \mathscr{P}_0, \\
\Psi_l(y) &= \|G^l y + g^l\| - 1 \quad \text{for } l \in \mathcal{Q} \backslash \mathcal{Q}_0.
\end{aligned} \tag{25}$$

To obtain our desired result, we assume that the origin $(0, 0)$ is a feasible solution of (1), which satisfies strictly those constraints not indexed by $\mathscr{P}_0$ and $\mathcal{Q}_0$, i.e.

$$\alpha_k \leqslant 0 \ (k \in \mathscr{P}_0), \ \|f^k\| < 1 \ (k \in \mathscr{P} \backslash \mathscr{P}_0) \quad \text{and} \quad \beta_l \leqslant 0 \ (l \in \mathcal{Q}_0), \ \|g^l\| < 1 \ (l \in \mathcal{Q} \backslash \mathcal{Q}_0). \tag{26}$$

The following lemma [33] refines the Chebychev inequality for bounded random variables.

*Lemma 4.2*
Let $\xi$ be a random variable with standard deviation $\sigma$. Suppose $\sigma \leqslant C$ and $|\xi - E[\xi]| \leqslant K$ always for some constants $C$ and $K$. Then, for any $t \in (0, C/K]$,

$$\text{Prob}[\xi - E[\xi] \geqslant \tfrac{3}{2} t C] \leqslant e^{-t^2/2}.$$



DOI: 10.1002/nla

Let $(\bar{X}, \bar{Y})$ be a feasible solution of (3) and $(\hat{x}, \hat{y})$ be a random vector pair generated by the method described above. Let $\sigma_0$ denote the standard deviation of $Q(\hat{x}, \hat{y})$. For every $k \in \mathscr{P}$ and $l \in \mathscr{Q}$, let $\sigma_x^k$ and $\sigma_y^l$ denote the standard deviations of $\Phi_k(\hat{x})$ and $\Psi_l(\hat{y})$, respectively. From the generation process of $(\hat{x}, \hat{y})$, it is clear that for $k \in \mathscr{P}_0$ and $l \in \mathscr{Q}_0$, we have $\Phi_k(\hat{x}) \leqslant 0$ and $\Psi_l(\hat{y}) \leqslant 0$ with probability 1. And for each $k \in \mathscr{P} \backslash \mathscr{P}_0$ and $l \in \mathscr{Q} \backslash \mathscr{Q}_0$, by the well-known Chebychev inequality, it holds that

$$\text{Prob}\{|\Phi_k(\hat{x}) - E[\Phi_k(\hat{x})]| \geqslant \varepsilon_k \sigma_x^k\} \leqslant \varepsilon_k^{-2}$$

and

$$\text{Prob}\{|\Psi_l(\hat{y}) - E[\Psi_l(\hat{y})]| \geqslant \delta_l \sigma_y^l\} \leqslant \delta_l^{-2},$$

provided that $\varepsilon_k > 1$ and $\delta_l > 1$. Moreover, it is easy to prove that

$$|Q(\hat{x}, \hat{y}) - E[Q(\hat{x}, \hat{y})]| \leqslant 2 \left[ \sum_{i,s=1}^{n} \sum_{j,t=1}^{m} |a_{ijst}| \sqrt{\bar{X}_{ii} \bar{X}_{ss} \bar{Y}_{jj} \bar{Y}_{tt}} + \sum_{i=1}^{n} \sum_{j=1}^{m} |h_{ij}| \sqrt{\bar{X}_{ii} \bar{Y}_{jj}} \right],$$

since $|\hat{x}_i \hat{x}_s| = \sqrt{\bar{X}_{ii} \bar{X}_{ss}}$ for $i, s = 1, \ldots, n$ and $|\hat{y}_j \hat{y}_t| = \sqrt{\bar{Y}_{jj} \bar{Y}_{tt}}$ for $j, t = 1, \ldots, m$. Consequently, applying Lemma 4.2 with $\xi = Q(\hat{x}, \hat{y})$ and $t = \frac{2}{3}\varepsilon_0$, we have

$$\text{Prob}[Q(\hat{x}, \hat{y}) - E[Q(\hat{x}, \hat{y})] \geqslant \varepsilon^0 \sigma_0] \leqslant \mathrm{e}^{-\frac{2}{9}\varepsilon_0^2},$$

when $0 < \varepsilon_0 < \frac{3}{2}\sigma_0 / K$ with

$$K = 2 \left[ \sum_{i,s=1}^{n} \sum_{j,t=1}^{m} |a_{ijst}| \sqrt{\bar{X}_{ii} \bar{X}_{ss} \bar{Y}_{jj} \bar{Y}_{tt}} + \sum_{i=1}^{n} \sum_{j=1}^{m} |h_{ij}| \sqrt{\bar{X}_{ii} \bar{Y}_{jj}} \right].$$

Therefore, if we generate $\hat{x}$ and $\hat{y}$ randomly and independently $L$ times, then the probability that one of these $L$ samples satisfies

$$\begin{aligned} Q(\hat{x}, \hat{y}) &\leqslant E[Q(\hat{x}, \hat{y})] + \varepsilon_0 \sigma_0, \\ \Phi_k(\hat{x}) &\leqslant E[\Phi_k(\hat{x})] + \varepsilon_k \sigma_x^k \quad \text{for } k \in \mathscr{P}, \\ \Psi_l(\hat{y}) &\leqslant E[\Psi_l(\hat{y})] + \delta_l \sigma_y^l \quad \text{for } l \in \mathscr{Q}, \end{aligned} \tag{27}$$

is at least $1 - \theta^L$, where

$$\theta := e^{-\frac{2}{9}\varepsilon_0^2} + \sum_{k \in \mathscr{P} \backslash \mathscr{P}_0} \varepsilon_k^{-2} + \sum_{l \in \mathscr{Q} \backslash \mathscr{Q}_0} \delta_l^{-2}. \tag{28}$$

We now construct feasible solutions with certainty, by moving $(\hat{x}, \hat{y})$ sufficiently close toward the origin. For each randomly generated $(\hat{x}, \hat{y})$, let $\bar{x} = \hat{x}$ and

$$\bar{y} = \begin{cases} \hat{y} & \text{if } \bar{x}^\top H \hat{y} \leqslant 0, \\ -\hat{y} & \text{otherwise.} \end{cases}$$

Denote

$$\bar{\tau}_x := \max\{\tau_x \in [0, 1] | \Phi_k(\tau_x \bar{x}) \leqslant 0, \ \forall\, k \in \mathscr{P}\}, \quad \bar{\tau}_y := \max\{\tau_y \in [0, 1] | \Psi_l(\tau_y \bar{y}) \leqslant 0, \ \forall\, l \in \mathscr{Q}\}$$

and

$$(\check{\tau}_x, \check{\tau}_y) := \text{argmin}\{Q(\tau_x \bar{x}, \tau_y \bar{y}) | (\tau_x, \tau_y) \in [0, \bar{\tau}_x] \times [0, \bar{\tau}_y]\}. \tag{29}$$

*Remark*

It is clear that $(\bar{\tau}_x, \bar{\tau}_y)$ is well defined and can be easily computed. Moreover, $(\check{\tau}_x, \check{\tau}_y)$ is an optimal solution of a 1-dimension bi-quadratic program with special structure and can be easily computed. In addition, from (29), we know that $(\check{\tau}_x, \check{\tau}_y)$ is a pair of random variable.



DOI: 10.1002/nla

For the considered problem (1) with constraints (25) satisfying (26), we have the following result.

*Theorem 4.2*
Let $L\geqslant 1$ be a given integer. If we generate $(\hat{x},\hat{y})$ randomly and independently $L$ times and construct $(x,y)=(\check{\tau}_x\bar{x},\check{\tau}_y\bar{y})$ as described above, then under Assumption 2, each $(x,y)$ is a feasible solution of (1) with probability 1. Moreover, for any $0<\varepsilon_0\leqslant\frac{3}{2}\sigma_0/K$, $\varepsilon_k>1$, $\eta_x^k\geqslant E[\Phi_k(\hat{x})]$ $(k\in\mathscr{P}\backslash\mathscr{P}_0)$, $\delta_l>1$ and $\eta_y^l\geqslant E[\Psi_l(\hat{y})]$ $(l\in\mathscr{Q}\backslash\mathscr{Q}_0)$, the probability that one of these $L$ samples satisfies

$$Q(x,y)\leqslant\min_{k\notin\mathscr{P}_0}\left(\frac{1-\|f^k\|}{\sqrt{1+\eta_x^k+\varepsilon_k\sigma_x^k}}\right)^2\min_{l\notin\mathscr{Q}_0}\left(\frac{1-\|g^l\|}{\sqrt{1+\eta_y^l+\delta_l\sigma_y^l}+\|g^l\|}\right)^2(E[Q(\hat{x},\hat{y})]+\varepsilon_0\sigma_0)\qquad(30)$$

is at least $1-\theta^L$, where $\theta$ is given by (28), $\sigma_x^k$ and $\sigma_y^l$ denote the standard deviations of $\Phi_k(\hat{x})$ and $\Psi_l(\hat{y})$ respectively, and $K$ is any constant satisfying $|Q(\hat{x},\hat{y})-E[Q(\hat{x},\hat{y})]|\leqslant K$.

*Proof*
It is easy to see that $(x,y)=(\check{\tau}_x\bar{x},\check{\tau}_y\bar{y})$ is a feasible solution of (1) with probability 1, since $(\check{\tau}_x,\check{\tau}_y)\in[0,\bar{\tau}_x]\times[0,\bar{\tau}_y]$.

We now prove (30). From (26) and the definitions of $\mathscr{P}_0$ and $\mathscr{Q}_0$, it follows that for each $k\in\mathscr{P}_0$ and each $l\in\mathscr{Q}_0$

$$\Phi_k(\tau_x\bar{x})=\tau_x^2\Phi_k(\hat{x})+(1-\tau_x^2)\alpha_k\leqslant 0\quad\text{for any }\tau_x\in[0,1]$$

and

$$\Psi_l(\tau_y\bar{y})=\tau_y^2\Psi_l(\hat{y})+(1-\tau_y^2)\beta_l\leqslant 0\quad\text{for any }\tau_y\in[0,1].$$

From (25) and (27), we see that $\|F^k\bar{x}+f^k\|\leqslant\sqrt{\varsigma_x^k}$ for each $k\in\mathscr{P}\backslash\mathscr{P}_0$, where $\varsigma_x^k:=1+E[\Phi_k(\hat{x})]+\varepsilon_k\sigma_x^k$. Moreover, for each $l\in\mathscr{Q}\backslash\mathscr{Q}_0$, if $\bar{x}^\top H\hat{y}\leqslant 0$, then $\|G^l\bar{y}+g^l\|\leqslant\sqrt{\varsigma_y^l}$; otherwise

$$\|G^l\bar{y}+g^l\|=\|-(G^l\hat{y}+g^l)+2g^l\|\leqslant\|G^l\hat{y}+g^l\|+2\|g^l\|\leqslant\sqrt{\varsigma_y^l}+2\|g^l\|,$$

where $\varsigma_y^l:=1+E[\Psi_l(\hat{y})]+\delta_l\sigma_y^l$. Thus, arguing identically as in the proof of Theorem 3.1, we obtain that

$$\bar{\tau}_x\geqslant\min_{k\notin\mathscr{P}_0}\frac{1-\|f^k\|}{\sqrt{\varsigma_x^k}}\quad\text{and}\quad\bar{\tau}_y\geqslant\min_{l\notin\mathscr{Q}_0}\frac{1-\|g^l\|}{\sqrt{\varsigma_y^l}+\|g^l\|}.\qquad(31)$$

Since $\eta_x^k\geqslant E[\Phi_k(\hat{x})]$ for $k\in\mathscr{P}\backslash\mathscr{P}_0$ and $\eta_y^l\geqslant E[\Psi_l(\hat{y})]$ for $l\in\mathscr{Q}\backslash\mathscr{Q}_0$, it follows from (31) that

$$\bar{\tau}_x\geqslant\hat{\tau}_x:=\min_{k\notin\mathscr{P}_0}\frac{1-\|f^k\|}{\sqrt{1+\eta_x^k+\varepsilon_k\sigma_x^k}}>0\quad\text{and}\quad\bar{\tau}_y\geqslant\hat{\tau}_y:=\min_{k\notin\mathscr{P}_0}\frac{1-\|g^l\|}{\sqrt{1+\eta_y^l+\delta_l\sigma_y^l}+\|g^l\|}>0,$$

which implies that $(\hat{\tau}_x,\hat{\tau}_y)\in(0,\bar{\tau}_x]\times(0,\bar{\tau}_y]$. Finally, our choice of $(\bar{x},\bar{y})$ implies that $\bar{x}^\top H\bar{y}\leqslant 0$ and $\bar{x}^\top H\bar{y}\leqslant\hat{x}^\top H\hat{y}$. Then, arguing similarly as in the proof of Theorem 3.1, we obtain

$$\begin{aligned}Q(\check{\tau}_x\bar{x},\check{\tau}_y\bar{y})&\leqslant Q(\hat{\tau}_x\bar{x},\hat{\tau}_y\bar{y})\\&\leqslant\hat{\tau}_x^2\hat{\tau}_y^2((\mathscr{A}\hat{x}\hat{x}^\top)\cdot\hat{y}\hat{y}^\top+\hat{x}^\top H\hat{y})\\&=\hat{\tau}_x^2\hat{\tau}_y^2Q(\hat{x},\hat{y})\\&\leqslant\hat{\tau}_x^2\hat{\tau}_y^2(E[Q(\hat{x},\hat{y})]+\varepsilon_0\sigma_0),\end{aligned}$$

where the last inequality comes from (27). We obtain the desired result and complete the proof. □



DOI: 10.1002/nla

## 5. APPROXIMATION SOLUTION OF THE RELAXED PROBLEM

In this paper, we mainly focus on the theoretical analysis of the polynomial time approximation algorithms for (1), which depends strongly on our ability to approximately solve the corresponding relaxed problem (3). In this section, we discuss the following special form of (1) whose related bi-linear SDP relaxation can be approximately solved in polynomial time.

$$\begin{array}{ll} \min & \sum_{i,s=1}^{n} \sum_{j,t=1}^{m} a_{ijst} x_i y_j x_s y_t \\ \text{s.t.} & \|x\|=1, \quad \|y\|=1, \\ & x^\top M^k x \leqslant 1, \quad k \in \mathscr{P}, \\ & y^\top N^l y \leqslant 1, \quad l \in \mathscr{Q}, \end{array} \tag{32}$$

where $M^k$ ($k \in \mathscr{P}$) and $N^l$ ($l \in \mathscr{Q}$) are positive semidefinite matrices. The model (32) is also a generalization of the problem studied in [4] and is slightly different from the problem in [6]. In this case, the related bi-linear SDP relaxation of (32) is

$$\begin{array}{rl} \phi_{\min} := \min & (\mathscr{A}X) \cdot Y \\ \text{s.t.} & I_n \cdot X = 1, \quad I_m \cdot Y = 1, \\ & M^k \cdot X \leqslant 1, \quad k \in \mathscr{P}, \\ & N^l \cdot Y \leqslant 1, \quad l \in \mathscr{Q}, \\ & X \succeq 0, \ Y \succeq 0. \end{array} \tag{33}$$

In order to study the approximation solution of (33), we need the following assumption and lemma, where the lemma generalized the result in [4] and was proved in [6].

*Assumption 3*
$I_n \cdot M^k < n$ for every $k \in \mathscr{P}$, and $I_m \cdot N^l < m$ for every $l \in \mathscr{Q}$.

*Lemma 5.1*
For any $X \in \mathscr{S}^n$, the following statements hold:

(1) If $\|X\|_F \leqslant 1/n$, then $\bar{X} := X + (1/n) I_n \succeq 0$.
(2) Suppose $n \geqslant 2$. If $I_n \cdot X \leqslant 0$ and $X \succeq -(1/n) I_n$, then $\|X\|_F \leqslant \sqrt{1 - 1/n}$.

Since $M^k (k \in \mathscr{P})$ and $N^l (l \in \mathscr{Q})$ are positive semidefinite, we have that, after some linear transformations $X := X - (1/n) I_n$ and $Y := Y - (1/m) I_m$, by Lemma 5.1, a restriction and a relaxation of (33) can be written as

$$\begin{array}{rl} \psi(\lambda) := \min & \Phi(X, Y) = (\mathscr{A}X) \cdot Y + \frac{1}{m} (\mathscr{A}X) \cdot I_m + \frac{1}{n} (\mathscr{A} I_n) \cdot Y + \frac{1}{mn} (\mathscr{A} I_n) \cdot I_m \\ \text{s.t.} & I_n \cdot X = 0, \quad I_m \cdot Y = 0, \\ & \left( M^k \cdot X + \frac{1}{n} I_n \cdot M^k \right)^2 \leqslant 1, \quad k \in \mathscr{P}, \\ & \left( N^l \cdot Y + \frac{1}{m} I_m \cdot N^l \right)^2 \leqslant 1, \quad l \in \mathscr{Q}, \\ & \|X\|_F \leqslant \lambda, \quad \|Y\|_F \leqslant \lambda, \end{array} \tag{34}$$

where $\lambda = 1/\max\{n, m\}$ and $\lambda = \sqrt{1 - 1/\max\{n, m\}}$ correspond to a restriction and a relaxation, respectively. Obviously,

$$\psi\left( \sqrt{1 - \frac{1}{\max\{n, m\}}} \right) \leqslant \phi_{\min} \leqslant \psi\left( \frac{1}{\max\{n, m\}} \right).$$



DOI: 10.1002/nla

For any $X \in \mathscr{S}^n$, we stack up the entries of $X$ (ignoring the symmetric part) into a vector, denoted by $vec_S(X)$, i.e.

$$vec_S(X)=(X_{11}, \sqrt{2}X_{12}, \ldots, \sqrt{2}X_{1n}, X_{22}, \sqrt{2}X_{23}, \ldots, \sqrt{2}X_{(n-1)n}, X_{nn})^\top.$$

Then there exists a suitable quadratic function $q_0(u, v)$, such that (34) can be rewritten as:

$$\begin{aligned} \min \quad & q_0(\mathrm{vec}(X), \mathrm{vec}(Y)) \\ \text{s.t.} \quad & \mathrm{vec}_S(I_n)^\top \mathrm{vec}_S(X)=0, \quad \mathrm{vec}_S(I_m)^\top \mathrm{vec}_S(Y)=0, \\ & \left(\mathrm{vec}_S(M^k)^\top \mathrm{vec}_S(X)+\frac{1}{n}I_n \cdot M^k\right)^2 \leqslant 1, \quad k \in \mathscr{P}, \\ & \left(\mathrm{vec}_S(N^l)^\top \mathrm{vec}_S(Y)+\frac{1}{m}I_m \cdot N^l\right)^2 \leqslant 1, \quad l \in \mathscr{Q}, \\ & \|\mathrm{vec}_S(X)\| \leqslant \lambda, \quad \|\mathrm{vec}_S(Y)\| \leqslant \lambda. \end{aligned} \tag{35}$$

The first two equation constraints can be used to eliminate two variables, denoted by $X_{11}$ and $Y_{11}$, by their linear relation with the other variables. Let $u=\mathrm{vec}_S(X)\setminus X_{11}$ and $v=\mathrm{vec}_S(Y)\setminus Y_{11}$. Then (35) can be equivalently formulated as

$$\begin{aligned} \min \quad & \bar{q}_0(u, v) \\ \text{s.t.} \quad & \left((a^k)^\top u+\frac{1}{m}I_n \cdot M^k\right)^2 \leqslant 1, \quad k \in \mathscr{P}, \\ & \left((b^l)^\top v+\frac{1}{m}I_m \cdot N^l\right)^2 \leqslant 1, \quad l \in \mathscr{Q}, \\ & \|A_0 u\| \leqslant \lambda, \quad \|B_0 v\| \leqslant \lambda, \end{aligned} \tag{36}$$

where $\bar{q}_0(u, v)$ is a quadratic function, $A_0$ and $B_0$ are two suitable matrices, and $a^k$ $(k \in \mathscr{P})$ and $b^l$ $(l \in \mathscr{Q})$ are some suitable vectors.

It is well known that quadratic function $q(x)=x^\top Ax+2b^\top x+c$ can be represented by the matrix denoted by:

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

Consequently, a standard SDP relaxation for the homogenized version of (36) is

$$\begin{aligned} z(\lambda^2):=\min \quad & \bar{Q}_0 \cdot Z \\ \text{s.t.} \quad & \bar{M}^k \cdot Z \leqslant 1, \quad k \in \mathscr{P}, \\ & \bar{N}^l \cdot Z \leqslant 1, \quad l \in \mathscr{Q}, \\ & \bar{I}_n \cdot Z \leqslant \lambda^2, \quad \bar{I}_m \cdot Z \leqslant \lambda^2, \\ & Z=\begin{bmatrix} 1 & u^\top & v^\top \\ u & W & U^\top \\ v & U & V \end{bmatrix} \succeq 0, \end{aligned} \tag{37}$$

where $\bar{Q}_0$, $\bar{M}^k$ $(k \in \mathscr{P})$, $\bar{N}^l$ $(l \in \mathscr{Q})$, $\bar{I}_n$, and $\bar{I}_m$ correspond to the matrix representation of the quadratic constraint functions in (35), respectively.

Notice that (37) can be solved in polynomial time. Based on the analysis above, we have the following main conclusion in this section, whose proof is similar to that of Theorem 3 in [6].



DOI: 10.1002/nla

*Theorem 5.1*
Under Assumption 3, a $(1-\gamma)^2/(\sqrt{p+q+2}+\gamma)^2\rho(\rho-1)$-approximation solution of (33) can be found in polynomial time, where $\rho=\max\{m,n\}$ and

$$\gamma=\max\left\{\frac{1}{n}I_n\cdot M^k, i\in\mathscr{P}, \frac{1}{m}I_m\cdot N^l, l\in\mathscr{Q}\right\}.$$

*Proof*
Let $\bar{Z}$ be an optimal solution of (37) with $\lambda=1/\rho$. By Theorem 1 in [1] and Assumption 3, a feasible solution pair $(\bar{u},\bar{v})$ of (36) can be found in polynomial time, such that

$$q_0(\bar{u},\bar{v})\leqslant\frac{(1-\gamma)^2}{(\sqrt{p+q+2}+\gamma)^2}z\left(\frac{1}{\rho^2}\right).$$

Based on the obtained $(\bar{u},\bar{v})$ and the stack relation between the vector and the matrix, we can find a feasible solution pair $(\bar{X},\bar{Y})$ for (34) with $\lambda=1/\rho$, such that $\Phi(\bar{X},\bar{Y})=q_0(\bar{u},\bar{v})$. Denote $X^*=\bar{X}+(1/n)I_n$ and $Y^*=\bar{Y}+(1/m)I_m$. By Lemma 5.1(1), it holds that $(X^*,Y^*)$ is a feasible solution of (33) satisfying

$$(\mathscr{A}X^*)\cdot Y^*\leqslant\frac{(1-\gamma)^2}{(\sqrt{p+q+2}+\gamma)^2}z\left(\frac{1}{\rho^2}\right). \tag{38}$$

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

$$\begin{aligned}z\left(\frac{1}{\rho^2}\right)&\leqslant\left(1-\frac{1}{\rho(\rho-1)}\right)z(0)+\frac{1}{\rho(\rho-1)}z\left(1-\frac{1}{\rho}\right)\\&\leqslant\frac{1}{\rho(\rho-1)}z\left(1-\frac{1}{\rho}\right)\\&\leqslant\frac{1}{\rho(\rho-1)}\phi_{\min}\end{aligned} \tag{39}$$

where the second inequality holds from the fact that $z(0)\leqslant 0$ and the last inequality holds since that $z(1-1/\rho)\leqslant\psi(\sqrt{1-1/\rho})\leqslant\phi_{\min}$. Combining (38) and (39), one obtain

$$(\mathscr{A}X^*)\cdot Y^*\leqslant\frac{(1-\gamma)^2}{(\sqrt{p+q+2}+\gamma)^2\rho(\rho-1)}\phi_{\min},$$

which shows that $(X^*,Y^*)$ is a $(1-\gamma)^2/(\sqrt{p+q+2}+\gamma)^2\rho(\rho-1)$-approximation solution of (33), since $0\leqslant\gamma<1$ from Assumption 3. We complete the proof. □

*Remark 5.1*
(a) For two special forms of maximization problems with quadratic constraints, some methods for finding approximation solution in polynomial time were presented by Zhang *et al.* under some mild conditions, see Theorems 4–6 in [6] for details. Theorems 5.1 in this paper are slightly different from these conclusions. (b) It is interesting to find some other forms of (1) whose bi-linear SDP relaxation can be approximately solved in polynomial time.

ACKNOWLEDGEMENTS

The authors thank the anonymous referees and an anonymous associate editor whose helpful comments and suggestions led to an improved presentation of this paper.

Chen Ling's work is supported by the National Natural Science Foundation of China (Grant No.10871168 and 10971187), and the Zhejiang Provincial Natural Science Foundation of China (Grant

No.Y6100366). Xinzhen Zhang's work was supported by the National Natural Science Foundation of China (Grant No.10771120). Liqun Qi's work is supported by the Hong Kong Research Grant Council.

## REFERENCES


1. Tseng P. Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation. *SIAM Journal on Optimization* 2003; **14**:268–283.
2. He S, Li Z, Zhang S. Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. *Mathematical Programming* 2010; **125**:353–383.
3. He S, Li Z, Zhang S. General constrained polynomial optimization: an approximation approach. *Technical Report SEEM2009-06*, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, 2009.
4. Ling C, Nie J, Qi L, Ye Y. Bi-quadratic optimization over unit spheres and semidefinite programming relaxations. *SIAM Journal on Optimization* 2009; **20**:1286–1310.
5. Luo C, Zhang S, A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints. *SIAM Journal on Optimization* 2010; **20**:1716–1736.
6. Zhang X, Ling C, Qi L. Semidefnite relaxation bounds for bi-quadratic optimization problems with quadratic constraints. *Journal of Global Optimization* 2011; **49**:293–311.
7. Cardoso JF. High-order contrasts for independent component analysis. *Neural Computation* 1999; **11**:157–192.
8. Comon P. Independent component analysis, a new concept? *Signal Processing* 1994; **36**:287–314.
9. De Lathauwer L, Comon P, De Moor B, Vandewalle J. Higher-order power method—application in independent component analysis. *Proceedings of the International Symposium on Nonlinear Theory and its Applications* (*NOLTA'95*), Las Vegas, NV, 1995; 91–96.
10. 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 Journal on Matrix Analysis and Applications* 2000; **21**:1324–1342.
11. Grigorascu VS, Regalia PA. Tensor displacement structures and polyspectral matching. In *Fast Reliable Algorithms for Structured Matrices*, Chapter 9, Kailath T, Sayed AH (eds). SIAM Publications: Philadeliphia, 1999.
12. Knowles JK, Sternberg E. On the ellipticity of the equations of non-linear elastostatics for a special material. *Journal of Elasticity* 1975; **5**:341–361.
13. Knowles JK, Sternberg E. On the failure of ellipticity of the equations for finite elastostatic plane strain. *Archive for Rational Mechanics and Analysis* 1977; **63**:321–336.
14. Kofidis E, Regalia PA. On the best rank-1 approximation of higher-order supersymmetric tensors. *SIAM Journal on Matrix Analysis and Applications* 2002; **23**:863–884.
15. Nikias CL, Petropulu AP. Higher-order spectra analysis. *A Nonlinear Signal Processing Framework*. Prentice-Hall: Englewood Cliffs, NJ, 1993.
16. Qi L, Wang F, Wang Y. Z-Eigenvalue methods for a global optimization polynomial optimization problem. *Mathematical Programming* 2009; **118**:301–316.
17. Rosakis P. Ellipticity and deformations with discontinuous deformation gradients in finite elastostatics. *Archive for Rational Mechanics and Analysis* 1990; **109**:1–37.
18. Simpson HC, Spector SJ. On copositive matrices and strong ellipticity for isotropic elastic materials. *Archive for Rational Mechanics and Analysis* 1983; **84**:55–68.
19. Wang Y, Aron M. A reformulation of the strong ellipticity conditions for unconstrained hyperelastic media. *Journal of Elasticity* 1996; **44**:89–96.
20. Zhang T, Golub GH. Rank-1 approximation of higher-order tensors. *SIAM Journal on Matrix Analysis and Applications* 2001; **23**:534–550.
21. Faybusovich L. Global optimization of homogeneous polynomials on the simplex and on the sphere. In *Frontiers in Global Optimization*, Floudas C, Pardalos P (eds). Kluwer Academic Publishers: Dordrecht, 2003.
22. Fu M, Luo Z, Ye Y. Approximation algorithms for quadratic programming. *Journal of Combinatorial Optimization* 1998; **2**:29–50.
23. He S, Luo Z, Nie J, Zhang S. Semidefinite relaxation bounds for indefinite homogeneous quadratic optimization. *SIAM Journal on Optimization* 2008; **19**:503–523.
24. Luo Z, Sidiropoulos ND, Tseng P, Zhang S. Approximation bounds for quadratic optimization with homogeneous quadratic constraints. *SIAM Journal on Optimization* 2007; **18**:1–28.
25. Pardalos PM, Wolkowicz H. Topics in semidefinite and interior-point methods. *Fields Institute Communications*, vol. 18. American Mathematical Society: Providence, RI, 1998.
26. Ramana M, Pardalos PM. Semidefinite programming. In *Interior Point Methods of Mathematical Programming*, Terlaky T (ed.). Kluwer Academic Publishers: Dordrecht, 1996; 369–398.
27. Ye Y, Zhang S. New results on quadratic minimization. *SIAM Journal on Optimization* 2003; **14**:245–267.
28. Markowitz HM. Portfolio selection. *The Journal of Finance* 1952; **7**:77–91.
29. Sturm JF, Zhang S. On cones of nonnegative quadratic functions. *Mathematics of Operations Research* 2003; **28**:246–267.

30. Goemans MX, Williamson DP. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM* 1995; **42**:1115–1145.
31. Ye Y. Approximating quadratic programming with bound and quadratic constraints. *Mathematical Programming* 1999; **84**:219–226.
32. Nesterov Y. Quality of semidefinite relaxation for nonconvex quadratic optimization. *CORE Discussion Paper*, *9719*, Belgium, March 1997.
33. Rényi A. *Probability Theory*. North-Holland: Amsterdam, 1970.



DOI: 10.1002/nla