International Journal of Computer Mathematics, 2014
Vol. 91, No. 5, 1082–1096, http://dx.doi.org/10.1080/00207160.2013.827784

Taylor & Francis
Taylor & Francis Group

A second-order Mehrotra-type predictor–corrector algorithm with a new wide neighbourhood for semi-definite programming

Ximei Yang[a*], Hongwei Liu[a] and Yinkui Zhang[b]

[a]*Department of Mathematics, Xidian University, Xi'an, China;* [b]*School of Mathematics and Computer Engineering, Xihua University, Chengdu, China*

(*Received 6 May 2013; revised version received 7 July 2013; accepted 16 July 2013*)

In this paper, we present a new second-order Mehrotra-type predictor–corrector algorithm for semi-definite programming (SDP). The proposed algorithm is based on a new wide neighbourhood. We are particularly concerned with an important inequality. Based on the inequality, the convergence is shown for a specific class of search directions. In particular, the complexity bound is $\mathcal{O}(\sqrt{n}\log\varepsilon^{-1})$ for the Nesterov–Todd search direction and $\mathcal{O}(n\log\varepsilon^{-1})$ for the Helmberg-Kojima-Monteiro search direction. The derived complexity bounds coincide with the currently best known theoretical complexity bounds obtained so far for SDP. We provide some preliminary numerical results as well.

Keywords: semi-definite programming; Mehrotra-type predictor–corrector algorithm; Schatten 1-norm; wide neighbourhood; complexity bound

2000 AMS Subject Classifications: 90C22; 90C51

1. Introduction

Semi-definite programming (SDP) is a generalization of linear programming (LP), which arises in many scientific and engineering fields. SDP has various applications in combinatorial optimization [3] and system and control theory [5]. Moreover, many problems in SDP come from eigenvalue optimization [23,24]. Therefore, SDP has received considerable attention and has been one of the most active research areas in mathematical programming.

With the success of interior-point methods (IPMs) in solving LP, the most IPMs were extended to SDP. The landmark work in this direction is introduced independently by Alizadeh [2] and Nesterov and Nemirovsky [20]. Alizadeh [3] extended Ye's projective potential reduction algorithm [29] from LP to SDP and argued that many known interior-point algorithms for LP could be transformed into algorithms for SDP. In addition, Nesterov and Nemirovsky [20] and Nesterov and Todd [21] presented a deep and unified theory of IPMs for solving the more general conic optimization problems using the notation of self-concordant barriers. There are other IPMs of solving SDP [4,6,8,12,17,25,27,28]. Most of these works are concentrated on primal–dual methods. The first proposed Mehrotra(M)-type predictor–corrector algorithm by Mehrotra [16] is typical representative of primal–dual IPMs and has a number of remarkable properties. Zhang

*Corresponding author. Email: yangximeiluoyang@126.com

and Zhang [31] first established convergence theory and complexity bounds for two M-type's second-order algorithms. Recently, Koulaei and Terlaky [11], Liu *et al.* [13,14] and Feng and Fang [6] present an extension of the variant of M-type predictor–corrector algorithm for SDP. Motivated by their work, we also present a new M-type predictor–corrector algorithm for SDP. Moreover, Ai and Zhang [1] introduced a new wide neighbourhood for linear complementarity problem. Later, Li and Terlaky [12] extend it to SDP and Liu *et al.* [15] extend it to symmetric cone programming. After learning those proposed neighbourhoods carefully, we find that these proposed neighbourhoods are defined by using the Frobenius norm. However, there has been no literature to define a neighbourhood using the Schatten 1-norm. This motivated us to propose a new wide neighbourhood using the Schatten 1-norm.

In this paper, based on the proposed wide neighbourhood, we propose a new second-order M-type predictor–corrector algorithm for SDP. In order to establish the iteration complexity for a specific class of search directions, which are called Monteiro-Zhang-search direction in [30], we prove an important relationship $\|UV + VU\|_1 \leq 2\|U\|_F\|V\|_F$. In particular, the complexity bound is $\mathcal{O}(\sqrt{n}\log\varepsilon^{-1})$ for the Nesterov–Todd (NT) search direction and $\mathcal{O}(n\log\varepsilon^{-1})$ for the Helmberg-Kojima-Monteiro (HKM) search direction. To our knowledge, we establish the complexity bound of the wide neighbourhood algorithm that coincides with the currently best complexity result for the small neighbourhood M-type predictor–corrector algorithm. Moreover, we present some numerical experiments illustrating the efficient properties of our algorithm.

This paper is organized as follows. In Section 2, we provide some properties of Schatten 1-norm. In Section 3, we give the definition of 1-norm neighbourhood and some properties. In Section 4, we introduce the SDP and give algorithmic framework. In Section 5, we establish the iteration complexity for the proposed algorithm. In Section 6, we present some numerical experiments illustrating the efficient properties of our algorithm. Finally, we close the paper by some conclusions.

1.1 *Notations*

$\mathbb{R}^n$	the n-dimensional Euclidean space
$\mathbb{R}^{m\times n}$	the set of all $m \times n$ matrices
S^n	the set of all $n \times n$ symmetric matrices
S^n_+	the set of all $n \times n$ symmetric positive semi-definite matrices
S^n_{++}	the set of all $n \times n$ symmetric positive definite matrices
$Q \succeq 0$	Q is positive semi-definite, where $Q \in S^n$
$Q \succ 0$	Q is positive definite, where $Q \in S^n$
$\mathrm{Tr}(Q)$	the trace of a matrix, i.e. $\mathrm{Tr}(Q) = \sum_{i=1}^n Q_{ii}$. Moreover, $G \bullet H = \mathrm{Tr}(G^{\mathrm{T}}H)$
$\lambda_i(Q)$	the eigenvalues of $Q \in S^n, i = 1, 2, \ldots, n$
$\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$	the smallest and the largest eigenvalue of $Q \in S^n$
$\Lambda(Q)$	the diagonal matrix with all the eigenvalues of Q as diagonal elements
cond(Q)	the condition number of Q, defined as $\mathrm{cond(Q)} = \lambda_{\max}(Q)/\lambda_{\min}(Q)$
$\|\|Q\|\|_F$	the Frobenius norm (F-norm) of $Q \in R^{n\times n}$, i.e. $\|\|Q\|\|_F = \sqrt{\mathrm{Tr}(Q^{\mathrm{T}}Q)}$

2. Some properties of the Schatten 1-norm

In order to introduce a new neighbourhood and the search direction, we need to present some notations. For any $x \in \mathbb{R}$, we define $x^+ = \max\{x, 0\}$ and $x^- = \min\{x, 0\}$. Let $M = Q\Lambda(M)Q^{\mathrm{T}} = \sum_{i=1}^n \lambda_i q_i q_i^{\mathrm{T}}$ be the eigenvalue decomposition of $M \in \mathcal{S}^n$, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all the eigenvalues of M and Q is an orthonormal matrix. Then, we define the positive part M^+ and the negative part M^- of M as

$$M^- = \sum \lambda_i^- q_i q_i^{\mathrm{T}} \quad \text{and} \quad M^+ = \sum \lambda_i^+ q_i q_i^{\mathrm{T}}.$$

Similarly, if $M \in S^n$ is a diagonal matrix, M^+ and M^- could be constructed by simply separating the non-negative and non-positive entries. Apparently, $M = M^+ + M^-$, where $M^+, -M^- \succeq 0$.

Let $\sigma_i(M), i = 1, \ldots, n$, be the singular values of $M \in \mathcal{S}^n$. The Schatten 1-norm (1-norm) is defined by $\|M\|_1 = \sum_{i=1}^n \sigma_i(M) = \sum_{i=1}^n |\lambda_i|$. We give some useful properties of the 1-norm. Similar to Proposition 3.1 in [12], we obtain the following lemma.

LEMMA 2.1 *Let $U, V \in \mathcal{S}^n$. Then $\|(U+V)^+\|_1 \leq \|U^+\|_1 + \|V^+\|_1$.*

Proof For $U, V \in \mathcal{S}^n$, we have

$$U = U^+ + U^- = U^+ + \sum_{\lambda_i(U) \leq 0} \lambda_i(U) q_i(U) q_i(U)^{\mathrm{T}}$$

and

$$V = V^+ + V^- = V^+ + \sum_{\lambda_i(V) \leq 0} \lambda_i(V) q_i(V) q_i(V)^{\mathrm{T}}.$$

Using Theorem 8.1.5 in Gulub and Van Loan [7], we have

$$\lambda_i(U+V) \leq \lambda_i(U^+ + V^+) \quad \text{for } i = 1, \ldots, n. \tag{1}$$

Hence using Equation (1), one has

$$\begin{aligned} \|(U+V)^+\|_1 &= \sum_{\lambda_i(U+V) \geq 0} \lambda_i(U+V) \leq \sum_{\lambda_i(U+V) \geq 0} \lambda_i(U^+ + V^+) \\ &= \|U^+ + V^+\|_1 \leq \|U^+\|_1 + \|V^+\|_1, \end{aligned}$$

which completes the proof. ∎

In the following, we prove the important inequality in this paper.

LEMMA 2.2 *Let $U, V \in S^n$. Then $\|UV + VU\|_1 \leq 2\|U\|_F \|V\|_F$.*

Proof Using the definition of F-norm, we have

$$\|U\|_F = \left[\sum_{i=1}^n \sigma_i^2(U)\right]^{1/2} \quad \text{and} \quad \|V\|_F = \left[\sum_{i=1}^n \sigma_i^2(V)\right]^{1/2},$$

where $\sigma_i(U)$ and $\sigma_i(V)$ are singular values of U and V.

Using the Corollary 3.4.3 and Theorem 3.3.14 in [9], we have

$$\begin{aligned} \frac{1}{2}\|UV + VU\|_1 &= \frac{1}{2}\sum_{i=1}^n \sigma_i(UV + VU) \leq \frac{1}{2}\left[\sum_{i=1}^n \sigma_i(UV) + \sum_{i=1}^n \sigma_i(VU)\right] \\ &= \sum_{i=1}^n \sigma_i(U)\sigma_i(V) \leq \left[\sum_{i=1}^n \sigma_i^2(U) \sum_{i=1}^n \sigma_i^2(V)\right]^{1/2} = \|U\|_F \|V\|_F, \end{aligned}$$

which completes the proof. ∎

Using the proof techniques of Lemma 3.3 in [12], it is easy to obtain the next lemma.

LEMMA 2.3 *Let $W \in \mathbb{R}^{n\times n}$ be a nonsingular matrix. Then for any $M \in S^n$*

$$\|M^+\|_1 \le \tfrac{1}{2}\|[WMW^{-1} + (WMW^{-1})^{\mathrm{T}}]^+\|_1.$$

3. 1-Norm neighbourhood

The so-called negative infinity neighbourhood, i.e. a wide neighbourhood, is defined as

$$\mathcal{N}_\infty^-(1-\tau) := \{(X,y,S) \in \mathcal{F}^0 : \lambda_{\min}(XS) \ge \tau\mu\}, \tag{2}$$

where $\tau \in (0,1)$ and $\mu = X \bullet S/n$.

The neighbourhood introduced by Li and Terlaky [12] is as follows:

$$\mathcal{N}(\tau_1,\tau_2) := \{(X,y,S) \in \mathcal{F}^0 : \|[\tau_1\mu I - X^{1/2}SX^{1/2}]^+\|_F \le (\tau_1-\tau_2)\mu\},$$

where $0 < \tau_2 < \tau_1 < 1$ and $\mu = X \bullet S/n$. $\mathcal{N}(\tau_1,\tau_2)$ is a wide neighbourhood and the following relationship holds [6,12]:

$$\mathcal{N}_\infty^-(1-\tau_1) \subseteq \mathcal{N}(\tau_1,\tau_2) \subseteq \mathcal{N}_\infty^-(1-\tau_2), \quad \forall\, 0 < \tau_2 < \tau_1 < 1.$$

In this paper, we define a new wide neighbourhood as follows:

$$\mathcal{N}_1(\tau,\beta) := \{(X,y,S) \in \mathcal{F}^0 : \|(\tau\mu I - X^{1/2}SX^{1/2})^+\|_1 \le \beta\tau\mu\}, \tag{3}$$

where $\beta,\tau \in (0,1)$ and $\mu = X \bullet S/n$.

For $\mathcal{N}_1(\tau,\beta)$, we give the following propositions.

PROPOSITION 3.1 *Let $\beta,\tau \in (0,1)$. Then $\mathcal{N}_\infty^-(1-\tau) \subseteq \mathcal{N}_1(\tau,\beta)$.*

Proof For $(X,S) \in \mathcal{N}_\infty^-(1-\tau)$, we have

$$\tau\mu - \lambda_i(X^{1/2}SX^{1/2}) \le 0, \quad i = 1,\ldots,n,$$

which is equivalent to $\tau\mu I - X^{1/2}SX^{1/2} \preceq 0$, which implies

$$(\tau\mu I - X^{1/2}SX^{1/2})^+ = 0,$$

leading to the claimed relationship. ■

PROPOSITION 3.2 *Let $(X,y,S) \in \mathcal{N}_1(\tau,\beta)$. Then*

(i) $\|(\tau\mu I - X^{1/2}SX^{1/2})^+\|_1 \le \beta\tau\mu$ *implies* $\lambda_{\min}(X^{1/2}SX^{1/2}) \ge (1-\beta)\tau\mu$.
(ii) *The matrices XS, SX, $X^{1/2}SX^{1/2}$ and $S^{1/2}XS^{1/2}$ have the same eigenvalues, since they are all similar to each other.*
(iii) *$\mathcal{N}_1(\tau,\beta)$ are symmetric with respect to X and S.*

4. SDP problem and algorithm

4.1 *SDP problem*

We consider the SDP given the following standard form:

$$(P) \quad \min C \bullet X, \quad \text{s.t. } A_i \bullet X = b_i, \ \ i = 1, 2, \ldots, m, \ X \succeq 0, \tag{4}$$

where $C, X \in \mathcal{S}^n, b \in \mathbb{R}^m$ and $A_i \in S^n$ is linearly independent. The dual problem associated with Equation (4) is

$$(D) \quad \max b^{\mathrm{T}} y, \quad \text{s.t. } \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0, \tag{5}$$

where $y \in \mathbb{R}^m$ and $S \in S^n$.

For convenience of reference, we define the following two sets:

$$\mathcal{F} := \left\{ (X, y, S) : A_i \bullet X = b_i, \ \sum_{i=1}^{m} y_i A_i + S = C, X, S \succeq 0 \right\},$$
$$\mathcal{F}^0 := \{(X, y, S) \in \mathcal{F} : X, S \succ 0\}.$$

We call $\mathcal{F}$ and $\mathcal{F}^0$, respectively, the (primal–dual) feasibility set and strictly feasibility set of (P) and (D). (X, y, S) is said to be feasible if $(X, y, S) \in \mathcal{F}$ and strictly feasible if $(X, y, S) \in \mathcal{F}^0$. In this paper, we assume that A_i is linearly independent and $\mathcal{F}^0 \neq \emptyset$.

It is well known that under the assumptions that $\mathcal{F}^0$ is nonempty and A_i is linearly independent, X^* and (y^*, S^*) are Equations (4) and (5) optimal solutions if and only if they satisfy the following system (see [10], P_{33}):

$$A_i \bullet X = b_i, X \succeq 0, \quad \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0, \quad XS = 0. \tag{6}$$

Applying Newton method for the perturbed system of equations (6), it leads to the following Newton equations:

$$A_i \bullet \triangle X = 0, \quad \sum_{i=1}^{m} A_i \triangle y_i + \triangle S = 0, \quad X \triangle S + \triangle X S = \tau \mu I - XS. \tag{7}$$

Although the second equality guarantees us a symmetric $\triangle S$, system (7) does not allow a symmetric solution matrix $\triangle X$. To overcome this difficulty, we use the approach proposed by Zhang [30], who suggested to replace the equation $XS = \tau \mu I$ by

$$H_P(XS) = \tau \mu I, \tag{8}$$

where $H_P(\cdot)$ is a symmetrization transformation defined as

$$H_P(M) = \tfrac{1}{2}[PMP^{-1} + (PMP^{-1})^{\mathrm{T}}], \quad \forall\, M \in \mathbb{R}^{n \times n},$$

for a given matrix M and a given nonsingular matrix P. In particular, if $P = I$ then for any symmetric matrix M, $H_I(M) = H(M) = M$. In [30], Zhang observed that if P is nonsingular, then

$$H_P(M) = \tau \mu I \Longleftrightarrow M = \tau \mu I.$$

In this paper, the scaling matrix P is selected from the specific class

$$\mathfrak{C}(X,S) := \{P \in \mathcal{S}^n_{++} : PXSP^{-1} \in \mathcal{S}^n\} = \{P \in \mathcal{S}^n_{++} : PXSP^{-1} = P^{-1}SXP\}.$$

In particular, choosing $P = S^{1/2}$ and $P = X^{-1/2}$ we get the HKM search directions, respectively. For the choice of $P = W^{1/2}$, we obtain the NT search direction [19,21,22,26], where

$$W = S^{1/2}(S^{1/2}XS^{1/2})^{-1/2}S^{1/2} = X^{-1/2}(X^{1/2}SX^{1/2})^{1/2}X^{-1/2}.$$

We compute the predictor direction and the corrector direction by solving the systems:

$$\begin{aligned} A_i \bullet \triangle X^a &= 0, \\ \sum_{i=1}^{m} A_i \triangle y_i^a + \triangle S^a &= 0, \\ H_P(X\triangle S^a + S\triangle X^a) &= R_c^- + \sqrt{n}R_c^+, \end{aligned} \tag{9}$$

and

$$\begin{aligned} A_i \bullet \triangle X^c &= 0, \\ \sum_{i=1}^{m} A_i \triangle y_i^c + \triangle S^c &= 0, \\ H_P(X\triangle S^c + S\triangle X^c) &= -H_P(\triangle X^a \triangle S^a), \end{aligned} \tag{10}$$

where $R_c = \tau\mu I - H_P(XS)$.

Let α be the step sizes taken along the predictor direction $(\triangle X^a, \triangle y^a, \triangle S^a)$ and the corrector direction $(\triangle X^c, \triangle y^c, \triangle S^c)$, then the new iterate is given by

$$(X(\alpha), y(\alpha), S(\alpha)) := (X, y, S) + \alpha(\triangle X^a, \triangle y^a, \triangle S^a) + \alpha^2(\triangle X^c, \triangle y^c, \triangle S^c).$$

We require that the largest step size $\bar{\alpha}$ satisfy the following two conditions:

A.1 For all $\alpha \in [0, \bar{\alpha}]$ such that $\mu(\bar{\alpha}) \le \mu(\alpha)$.
A.2 For all $\alpha \in [0, \bar{\alpha}]$ satisfies $(X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}_1(\tau, \beta)$.

4.2 *Algorithm*

In the following, we state the generic framework of our algorithm.

Algorithm 1 M-type predictor–corrector algorithm

Input: A threshold parameter $\tau \le 1/4, \beta \le 1/2$. An accuracy parameter $\varepsilon > 0$.
An initial point (X^0, y^0, S^0) such that $(X^0, y^0, S^0) \in \mathcal{N}_1(\tau, \beta)$. Set $\mu^0 = X^0 \bullet S^0/n, k := 0$.

Step 1. If $\mu^k \le \varepsilon\mu^0$, then stop.
Step 2. Choose a scaling matrix $P \in \mathfrak{C}(X^k, S^k)$.
Step 3. (Predictor step) Compute the predictor direction by solving Equation (9).
Step 4. (Corrector step) Compute the corrector direction by solving Equation (10) and the largest step size $\bar{\alpha}^k \in (0, 1]$ such that A.1 and A.2.
Step 5. Let $(X^{k+1}, y^{k+1}, S^{k+1}) = (X(\alpha^k), y(\alpha^k), S(\alpha^k))$, $\mu^{k+1} = X^{k+1} \bullet S^{k+1}/n$. Set $k := k + 1$ and go to Step 1.

---

5. Complexity analysis

In this section, we are going to establish the iteration complexity for Algorithm 4.2. In order to achieve the purpose, we have to make preparation work in advance.

5.1 *Scaling procedure*

In order to analyse the proposed algorithm in a unified way for the scaling matrix $P \in \mathfrak{C}(X,S)$, we scale the primal and dual variables as follows:

$$(\tilde{X},\tilde{y},\tilde{S}) = (PXP, y, P^{-1}SP^{-1}) \quad \text{and} \quad (\tilde{C},\tilde{A}_i,\tilde{b}_i) = (P^{-1}CP^{-1}, P^{-1}A_iP^{-1}, b_i).$$

We scale Newton systems (9) and (10) and obtain the following systems:

$$\begin{aligned} \tilde{A}_i \bullet \triangle\tilde{X}^a &= 0, \\ \sum_{i=1}^{m} \tilde{A}_i \triangle\tilde{y}_i^a + \triangle\tilde{S}^a &= 0, \\ H(\tilde{X}\triangle\tilde{S}^a + \tilde{S}\triangle\tilde{X}^a) &= \tilde{R}_c^- + \sqrt{n}\tilde{R}_c^+, \end{aligned} \tag{11}$$

and

$$\begin{aligned} \tilde{A}_i \bullet \triangle\tilde{X}^c &= 0, \\ \sum_{i=1}^{m} \tilde{A}_i \triangle\tilde{y}_i^c + \triangle\tilde{S}^c &= 0, \\ H(\tilde{X}\triangle\tilde{S}^c + \tilde{S}\triangle\tilde{X}^c) &= -H(\triangle\tilde{X}^a\triangle\tilde{S}^a), \end{aligned} \tag{12}$$

where $\tilde{R}_c = \mu I - \tilde{X}\tilde{S}$, $H(\tilde{X}\tilde{S}) = \tilde{X}\tilde{S}$, $\triangle\tilde{X}^a = P\triangle X^a P, \triangle\tilde{S}^a = P^{-1}\triangle S^a P^{-1}$, $\triangle\tilde{y}^a = \triangle y^a$ and $\triangle\tilde{X}^c = P\triangle X^c P$, $\triangle\tilde{S}^c = P^{-1}\triangle S^c P^{-1}$, $\triangle\tilde{y}^c = \triangle y^c$.

After scaled direction, the iterates are rewritten as

$$(\tilde{X}(\alpha),\tilde{y}(\alpha),\tilde{S}(\alpha)) = (\tilde{X},\tilde{y},\tilde{S}) + \alpha(\triangle\tilde{X}^a,\triangle\tilde{y}^a,\triangle\tilde{S}^a) + \alpha^2(\triangle\tilde{X}^c,\triangle\tilde{y}^c,\triangle\tilde{S}^c). \tag{13}$$

In order to illustrate the relationship between the original and the scaled problems, we give the next results.

(i) $(X,y,S) \in \mathcal{F}$ if and only if $(\tilde{X},\tilde{y},\tilde{S})$ is feasible for scaled primal and dual problems.
(ii) $\tilde{X}(\alpha) = PX(\alpha)P, \tilde{y}(\alpha) = y(\alpha), \tilde{S}(\alpha) = P^{-1}S(\alpha)P^{-1}$ and $\mu(\alpha) = \tilde{\mu}(\alpha)$, where $\tilde{\mu}(\alpha) = \tilde{X}(\alpha) \bullet \tilde{S}(\alpha)/n$.

Due to $\tilde{X}^{1/2}\tilde{S}\tilde{X}^{1/2}$ and $X^{1/2}SX^{1/2}$ having the same eigenvalues, one has the following lemma:

LEMMA 5.1 *The neighbourhood $\mathcal{N}(\tau,\beta)$ is a scaling invariant, that is (X,y,S) is in the neighbourhood if and only if $(\tilde{X},y,\tilde{S})$ is.*

5.2 *Lyapunov-like operator*

Let $A \in R^{n\times n}$ be given, and define a linear operator $L_A : S^n \to S^n$ as

$$L_A(X) = AX + XA^{\mathrm{T}}, \tag{14}$$

which is called the Lyapunov operator [10, Theorem E.2]. For convenience, we define a Lyapunov-like operator $\tilde{L}_A : S^n \to S^n$ as follows:

$$\tilde{L}_A(X) = \tfrac{1}{2}(AX + XA^{\mathrm{T}}). \tag{15}$$

It is clear that the operator $\tilde{L}_A$ has all the properties of L_A. In what follows, we list some properties of $\tilde{L}_A$, which play a key role in the following analysis.

PROPOSITION 5.2 *Let $\tilde{L}_A$ be defined in Equation (15). Then*

(i) *Let $A \in S^n_{++}$, then the $\tilde{L}_A$ is guaranteed to be invertible.*
(ii) *For $A \in S^n$, we have $\tilde{L}_A^{-1}(A^2) = A, \tilde{L}_A^{-1}(A) = I$ and $\tilde{L}_A^{-1}(I) = A^{-1}$.*
(iii) *$\tilde{L}_A$ is symmetric with respect to $\langle \cdot, \cdot \rangle$, that is $\langle \tilde{L}_A(X), S\rangle = \langle X, \tilde{L}_A(S)\rangle$.*
(iv) *Let $A \succ 0$, $B \succeq 0$, then $\tilde{L}_A^{-1}(B) \succeq 0$.*

Moreover, using the Lyapunov-like operator, the third equality of Equations (11) and (12) can be rewritten as

$$\tilde{L}_{\tilde{X}}(\triangle\tilde{S}^a) + \tilde{L}_{\tilde{S}}(\triangle\tilde{X}^a) = \tilde{R}_c^- + \sqrt{n}\tilde{R}_c^+, \tag{16}$$

$$\tilde{L}_{\tilde{X}}(\triangle\tilde{S}^c) + \tilde{L}_{\tilde{S}}(\triangle\tilde{X}^c) = -H(\triangle\tilde{X}^a\triangle\tilde{S}^a). \tag{17}$$

5.3 *Step size calculation*

In this section, we discuss how to calculate the largest step size $\bar{\alpha}$. For simplicity, we will often write $\tilde{X}, \tilde{y}, \tilde{S}$ and $\bar{\alpha}$ for $\tilde{X}^k, \tilde{y}^k, \tilde{S}^k$ and $\bar{\alpha}^k$, respectively.

By using Lemma 2.3 and the fact $-\mathrm{Tr}[\tau\mu I - \tilde{X}\tilde{S}]^- \geq (1-\tau)\mu n$, we immediately obtain the following result.

LEMMA 5.3 *Let $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{N}_1(\tau, \beta)$ and $\tau \leq \frac{1}{4}, \beta \leq \frac{1}{2}$. Then*

$$(\tau - 1)\mu n \leq \mathrm{Tr}([\tau\mu I - \tilde{X}\tilde{S}]^- + \sqrt{n}[\tau\mu I - \tilde{X}\tilde{S}]^+) \leq \left(\tau + \frac{\beta\tau}{\sqrt{n}} - 1\right)\mu n < 0.$$

In the next lemma, we deduce immediately that $\alpha = 1$ satisfies A.1.

LEMMA 5.4 *Let $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{N}_1(\tau, \beta)$. Then $\mu(\alpha)$ is strictly monotonically decreasing in $\alpha \in [0, 1]$.*

Proof By computing $\mu(\alpha)$ directly, we have

$$\mu(\alpha) = \tilde{\mu}(\alpha) = \frac{1}{n}\tilde{X}(\alpha) \bullet \tilde{S}(\alpha) = \mu + \alpha\left[(\tau - 1)\mu + \frac{\sqrt{n}-1}{n}\mathrm{Tr}(\tilde{R}_c^+)\right]. \tag{18}$$

Taking the derivative with respect to α, we have

$$\begin{aligned}\mu(\alpha)' &= (\tau - 1)\mu + \frac{\sqrt{n}-1}{n}\operatorname{Tr}(\tilde{R}_c^+)\\ &\le (\tau - 1)\mu + \frac{\sqrt{n}-1}{n}\beta\tau\mu\\ &\le (\tau + \beta\tau - 1)\mu < 0,\end{aligned}$$

which implies $\mu(\alpha)$ is strictly monotonically decreasing in $\alpha \in [0, 1]$. ■

Thus, the largest step size $\bar{\alpha}$ will be calculated as follows:

$$\bar{\alpha} = \max\{\alpha : (\tilde{X}(\alpha), \tilde{y}(\alpha), \tilde{S}(\alpha)) \in \mathcal{N}_1(\tau, \beta), \forall\alpha \in [0, 1]\}. \tag{19}$$

Indeed, we calculate $\bar{\alpha}$ by replacing Equation (19) with the following method. Let

$$\bar{\alpha} = \max\{\alpha : g(\alpha) \le 0, \alpha \in [0, 1]\}, \tag{20}$$

where $D_a = \triangle\tilde{X}^a\triangle\tilde{S}^c + \triangle\tilde{S}^a\triangle\tilde{X}^c$, $D_c = \triangle\tilde{X}^c\triangle\tilde{S}^c$ and

$$g(\alpha) = \begin{cases}\alpha^3\|[D_a]^-\|_1 + \alpha^4\|[D_c]^-\|_1 - \beta\tau\mu(\alpha) & \text{if } \alpha \ge \dfrac{1}{\sqrt{n}},\\ \dfrac{\alpha^2}{\sqrt{n}}\|[D_a]^-\|_1 + \dfrac{\alpha^3}{\sqrt{n}}\|[D_c]^-\|_1 - \beta\tau\mu(\alpha) & \text{if } \alpha < \dfrac{1}{\sqrt{n}}.\end{cases} \tag{21}$$

In the following, we explain the reasons why we define $g(\alpha)$ in Equation (21). First, let us define $\Gamma(\alpha)$ as follows:

$$\begin{aligned}\Gamma(\alpha) :&= \tilde{X}\tilde{S} + \alpha[[\tau\mu I - \tilde{X}\tilde{S}]^- + \sqrt{n}[\tau\mu I - \tilde{X}\tilde{S}]^+]\\ &= (1-\alpha)\tilde{X}\tilde{S} + \alpha\tau\mu I + \alpha(\sqrt{n}-1)[\tau\mu I - \tilde{X}\tilde{S}]^+,\end{aligned}$$

which implies $\Gamma(\alpha) \succeq 0$.

LEMMA 5.5 *Let* $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{N}_1(\tau, \beta)$. *Then*
If $1/\sqrt{n} \le \alpha$, *we have*

$$\|[\tau\mu(\alpha)I - \Gamma(\alpha)]^+\|_F = 0. \tag{22}$$

If $\alpha < 1/\sqrt{n}$, *we have*

$$\|[\tau\mu(\alpha)I - \Gamma(\alpha)]^+\|_F \le (1 - \alpha\sqrt{n})\beta\tau\mu(\alpha). \tag{23}$$

Proof Due to $\mu(\alpha) \le \mu$ and $\Gamma(\alpha) \succeq 0$ for all $\alpha \in [0, 1]$, we have

$$\begin{aligned}[\tau\mu(\alpha)I - \Gamma(\alpha)]^+ &\preceq \left[\tau\mu(\alpha)I - \frac{\mu(\alpha)}{\mu}\Gamma(\alpha)\right]^+\\ &= \frac{\mu(\alpha)}{\mu}[\tau\mu I - \Gamma(\alpha)]^+\end{aligned}$$

$$= \frac{\mu(\alpha)}{\mu}[(1-\alpha)[\tau\mu I - \tilde{X}\tilde{S}]^- + (1-\alpha\sqrt{n})[\tau\mu I - \tilde{X}\tilde{S}]^+]^+$$
$$= \frac{\mu(\alpha)}{\mu}(1-\alpha\sqrt{n})^+[\tau\mu I - \tilde{X}\tilde{S}]^+. \tag{24}$$

For $\alpha < 1/\sqrt{n}$, using Equation (24) and $1-\alpha\sqrt{n} > 0$, one has

$$\|[\tau\mu(\alpha)I - \Gamma(\alpha)]^+\|_1 \le \frac{\mu(\alpha)}{\mu}(1-\alpha\sqrt{n})\|[\tau\mu I - \tilde{X}\tilde{S}]^+\|_1 \le (1-\alpha\sqrt{n})\beta\tau\mu(\alpha),$$

where the last inequality follows from $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{N}_1(\tau, \beta)$.

For $1/\sqrt{n} \le \alpha$, using Equation (24) and $1-\alpha\sqrt{n} \le 0$, one has

$$\|[\tau\mu(\alpha)I - \Gamma(\alpha)]^+\|_1 = 0.$$

The proof is completed. ■

$\|[\tau\mu(\alpha)I - \tilde{X}(\alpha)\tilde{S}(\alpha)]^+\|_1 \le \beta\tau\mu(\alpha)$ holds if

$$\|[\tau\mu(\alpha)I - \Gamma(\alpha)]^+\|_1 + \alpha^3\|[D_a]^-\|_1 + \alpha^4\|[D_c]^-\|_1 \le \beta\tau\mu(\alpha).$$

Therefore, using Lemma 5.5, we well define $g(\alpha)$ as Equation (21).

The next lemma shows that $(\tilde{X}(\alpha), \tilde{y}(\alpha), \tilde{S}(\alpha)) \in \mathcal{N}_1(\tau, \beta)$, for $\alpha \in [0, \bar{\alpha}]$.

LEMMA 5.6 *Let $\bar{\alpha}$ be defined in Equation (20). Then for all $\alpha \in [0, \bar{\alpha}]$*

$$(\tilde{X}(\alpha), \tilde{y}(\alpha), \tilde{S}(\alpha)) \in \mathcal{N}_1(\tau, \beta).$$

Proof From Lemma 5.5, we have

$$\|[\tau\mu(\alpha)I - X^{1/2}(\alpha)S(\alpha)X^{1/2}(\alpha)]^+\|_1 \le \|[H_{X^{1/2}(\alpha)}(\tau\mu(\alpha)I - X^{1/2}(\alpha)S(\alpha)X^{1/2}(\alpha))]^+\|_1$$
$$= \|[\tau\mu(\alpha)I - \tilde{X}(\alpha)\tilde{S}(\alpha)]^+\|_1 \le \beta\tau\mu(\alpha),$$

where the first inequality follows from Lemma 2.3.

Moreover, for all $\alpha \in [0, \bar{\alpha}]$, $\|[\tau\mu(\alpha)I - \tilde{X}(\alpha)\tilde{S}(\alpha)]^+\|_1 \le \beta\tau\mu(\alpha)$ implies $\tilde{X}(\alpha)\tilde{S}(\alpha) \in S^n_{++}$. Thus, we have $\det(\tilde{X}(\alpha)) \ne 0, \det(\tilde{S}(\alpha)) \ne 0$ for all $\alpha \in [0, \bar{\alpha}]$. Then, since $\tilde{X} \in S^n_{++}$, $\tilde{S} \in S^n_{++}$, by the continuity, it follows that in this interval both $\tilde{X}(\alpha) \in S^n_{++}, \tilde{S}(\alpha) \in S^n_{++}$.

From Lemma 2.3 and the definition of $\mathcal{N}_1(\tau, \beta)$ in (13), the proof is completed. ■

5.4 *Technical results*

We begin with several lemmas that are frequently used in this section.

LEMMA 5.7 [18, Lemma 4.6] *Let $U, V \in S^n$ and $G \in S^n_{++}$. Then*

$$\|U\|_F\|V\|_F \le \sqrt{\text{cond }(G)}\|G^{-1/2}U\|_F\|G^{1/2}V\|_F$$
$$\le \frac{\sqrt{\text{cond}(G)}}{2}(\|G^{-1/2}U\|^2 + \|G^{1/2}V\|^2).$$

LEMMA 5.8 *Let* $G=\tilde{L}_{\tilde{S}}^{-1}\tilde{L}_{\tilde{X}}$, $(\tilde{X},\tilde{S})\in\mathcal{N}_1(\tau,\beta)$ *and* $\beta\le\frac{1}{2},\tau\le\frac{1}{4}$. *Then*

$$\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]^{+}\|_F^2\le\beta\tau\mu \quad\text{and}\quad \|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]^{-}\|_F^2\le\mu n.$$

Proof Some straightforward computations yield

$$\begin{aligned}\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]^{+}\|_F^2&\le\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}\|^2\|[\tau_1\mu I-\tilde{X}\tilde{S}]^{+}\|_F^2\\&\le\frac{1}{\lambda_{\min}(\tilde{X}\tilde{S})}\|[\tau\mu I-\tilde{X}\tilde{S}]^{+}\|_1^2\le\frac{(\beta\tau\mu)^2}{(1-\beta)\tau\mu}\le\beta\tau\mu,\end{aligned}$$

where the third inequality follows from $\lambda_{\min}(\tilde{X}\tilde{S})\ge(1-\beta)\tau\mu$, $\|[\tau\mu I-\tilde{X}\tilde{S}]^{+}\|_1\le\beta\tau\mu$ and the last inequality follows from $\beta\le\frac{1}{2}$.

In what follows, we prove the second inequality.

$$\begin{aligned}\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]^{-}\|_F^2&\le\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]\|_F^2\\&=\langle\tau\mu I-\tilde{X}\tilde{S},(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1}(\tau\mu I-\tilde{X}\tilde{S})\rangle\\&=\langle\tau\mu I,(L_{\tilde{X}}L_{\tilde{S}})^{-1}\tau\mu I\rangle-2\langle\tau\mu I,I\rangle+\langle\tilde{X}\tilde{S},I\rangle\\&\le\|(L_{\tilde{X}}L_{\tilde{S}})^{-1}\|\langle\tau\mu I,\tau\mu I\rangle-2\langle\tau\mu I,I\rangle+\langle\tilde{X}\tilde{S},I\rangle\\&\le\tau^2\mu^2n/\lambda_{\min}(\tilde{X}\tilde{S})-2\tau\mu n+\mu n\\&\le\mu n,\end{aligned}$$

which completes the proof. ∎

LEMMA 5.9 *Let* $G=\tilde{L}_{\tilde{S}}^{-1}\tilde{L}_{\tilde{X}}$. *Then* $\|\triangle\tilde{X}^a\|_F\|\triangle\tilde{S}^a\|_F\le 9\sqrt{\mathrm{cond}(G)}\mu n/16$.

Proof Multiplying the equation of (16) by $(\tilde{L}_{\tilde{S}}\tilde{L}_{\tilde{X}})^{-1/2}$ and taking norm-squared on both sides, we have

$$\begin{aligned}\|G^{-1/2}\triangle\tilde{X}^a+G^{1/2}\triangle\tilde{S}^a\|_F^2&=\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[[\tau\mu I-\tilde{X}\tilde{S}]^{-}+\sqrt{n}[\tau\mu I-\tilde{X}\tilde{S}]^{+}]\|_F^2\\&\le\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]^{-}\|_F^2+n\|(\tilde{L}_{\tilde{X}}\tilde{L}_{\tilde{S}})^{-1/2}[\tau\mu I-\tilde{X}\tilde{S}]^{+}\|_F^2\\&\le\mu n+\beta\tau\mu n=(1+\beta\tau)\mu n.\end{aligned}\tag{25}$$

Using the second inequality in Lemma 5.7, we have

$$\begin{aligned}\|\triangle\tilde{X}^a\|_F\|\triangle\tilde{S}^a\|_F&\le\frac{\sqrt{\mathrm{cond}(G)}}{2}(\|G^{-1/2}\triangle\tilde{X}^a\|_F^2+\|G^{1/2}\triangle\tilde{S}^a\|_F^2)\\&=\frac{\sqrt{\mathrm{cond}(G)}}{2}(\|G^{-1/2}\triangle\tilde{X}^a+G^{1/2}\triangle\tilde{S}^a\|_F^2)\\&\le\frac{\sqrt{\mathrm{cond}(G)}}{2}(1+\beta\tau)\mu n\le\frac{9}{16}\sqrt{\mathrm{cond}(G)}\mu n.\end{aligned}$$

The proof is completed. ∎

The following lemma will play a key role in our analysis.

LEMMA 5.10 *Let* $G=\tilde{L}_{\tilde{S}}^{-1}\tilde{L}_{\tilde{X}}$. *Then* $\|\triangle\tilde{X}^c\|_F\|\triangle\tilde{S}^c\|_F\le 21\mathrm{cond(G)}^{3/2}n^2\mu/(64\tau)$.

Proof Multiplying the last equation of (17) by $(\tilde{L}_{\tilde{S}}\tilde{L}_{\tilde{X}})^{-1/2}$ and taking norm-squared on both sides, we have

$$\begin{aligned}\|G^{-1/2}\triangle\tilde{X}^c + G^{1/2}\triangle\tilde{S}^c\|_F^2 &= \|(L_{\tilde{X}}L_{\tilde{S}})^{-1/2}(-\triangle\tilde{X}^a\triangle\tilde{S}^a)\|_F^2\\ &\leq \frac{1}{\lambda_{\min}(\tilde{X}\tilde{S})}\|\triangle\tilde{X}^a\triangle\tilde{S}^a\|_F^2 \leq \frac{1}{(1-\beta)\tau\mu}(\|\triangle\tilde{X}^a\|_F\|\triangle\tilde{S}^a\|_F)^2\\ &\leq \frac{1}{(1-\beta)\tau\mu}\left(\frac{9}{16}\sqrt{\mathrm{cond(G)}}n\mu\right)^2 \leq \frac{21}{32\tau}\mathrm{cond(G)}n^2\mu. \qquad (26)\end{aligned}$$

where the second inequality follows from $\lambda_{\min}(\tilde{\omega}) \geq (1-\beta)\tau\mu$ and the third inequality follows from Lemma 5.9.

Using the second inequality in Lemma 5.7 and (26), we have

$$\begin{aligned}\|\triangle\tilde{X}^c\|_F\|\triangle\tilde{S}^c\|_F &\leq \frac{\sqrt{\mathrm{cond(G)}}}{2}(\|G^{-1/2}\triangle\tilde{X}^c\|^2 + \|G^{1/2}\triangle\tilde{S}^c\|_F^2)\\ &= \frac{\sqrt{\mathrm{cond(G)}}}{2}(\|G^{-1/2}\triangle\tilde{X}^c + G^{1/2}\triangle\tilde{S}^c\|_F^2)\\ &\leq \frac{\sqrt{\mathrm{cond(G)}}}{2}\cdot\frac{21}{32\tau}\mathrm{cond(G)}n^2\mu\\ &= \frac{21}{64\tau}\mathrm{cond(G)}^{3/2}n^2\mu,\end{aligned}$$

which shows that the proof of the lemma is completed. ■

Using the fact $G^{-1/2}\triangle\tilde{X}^a \bullet G^{-1/2}\triangle\tilde{S}^a = 0$, $G^{-1/2}\triangle\tilde{X}^c \bullet G^{-1/2}\triangle\tilde{S}^c = 0$ and (25), (26) and the first inequality in Lemma 5.7, we easily obtain the following corollary.

COROLLARY 5.11 *Let $G = \tilde{L}_{\tilde{S}}^{-1}\tilde{L}_{\tilde{X}}$ and $\beta \leq \frac{1}{2}$. Then*

(i) $\|\triangle\tilde{X}^a\|_F\|\triangle\tilde{S}^c\|_F \leq 15\mathrm{cond}(G)n^{3/2}\mu/(16\sqrt{\tau})$.
(ii) $\|\triangle\tilde{S}^a\|_F\|\triangle\tilde{X}^c\|_F \leq 15\mathrm{cond}(G)n^{3/2}\mu/(16\sqrt{\tau})$.

LEMMA 5.12 *Let $\bar{\alpha}$ be defined in (20). Then $\bar{\alpha} \geq 4\beta\tau^2/(5\sqrt{\mathrm{cond}(G)}\sqrt{n})$.*

Proof From Equation (21), we obtain $1/\sqrt{n} \leq \bar{\alpha}$ or $\bar{\alpha} < 1/\sqrt{n}$. If $1/\sqrt{n} \leq \bar{\alpha}$, we immediately obtain the lower bound on $\bar{\alpha}$. Thus, we only mainly consider $\bar{\alpha} < 1/\sqrt{n}$. Let $\alpha = 4\beta\tau^2/(5\sqrt{\mathrm{cond}(G)}\sqrt{n})$, we have

$$\begin{aligned}&\frac{\alpha^2}{\sqrt{n}}\|[\triangle\tilde{X}^a\triangle\tilde{S}^c + \triangle\tilde{S}^a\triangle\tilde{X}^c]^-\|_1 + \frac{\alpha^3}{\sqrt{n}}\|[\triangle\tilde{X}^c\triangle\tilde{S}^c]^-\|_1 - \beta\tau\mu(\alpha)\\ &\leq \alpha^2\frac{15}{8\sqrt{\tau}}\mathrm{cond(G)}\mu n + \alpha^3\frac{21}{64\tau}\mathrm{cond(G)}^{3/2}\mu n^{3/2} - \beta\tau^2\mu\\ &= \frac{6\beta^2\tau^{7/2}}{5}\mu + \frac{21\beta^3\tau^5}{125}\mu - \beta\tau^2\mu = \beta\tau^2\mu\left[\frac{6\beta\tau^{3/2}}{5} + \frac{21\beta^2\tau^3}{125} - 1\right]\\ &\leq \beta\tau^2\mu\left[\frac{1}{10} + \frac{1}{1000} - 1\right] \leq 0,\end{aligned}$$

where the first inequality follows from Corollary 5.11, Lemma 5.10 and the fact $\mu(\alpha) = \mu + \alpha[(\tau-1)\mu + (\sqrt{n}-1)/n\,\mathrm{Tr}(\tilde{R}_c^+)] \geq \mu + \alpha(\tau-1)\mu \geq \mu + \tau\mu - \mu = \tau\mu$.

Thus, from the definition of $\bar{\alpha}$ in (20), we have $\bar{\alpha} \geq 4\beta\tau^2/(\sqrt{\mathrm{cond}(G)}\sqrt{n})$. ■

The following theorem gives an upper bound for the number of iterations in which Algorithm 4.2 stops with an ε-approximate solution.

THEOREM 5.13 *Let $\sqrt{\mathrm{cond}(G)}$ be bounded from above by $\kappa < \infty$ for all iterations. Then Algorithm 4.2 will terminate with (X^k, y^k, S^k) such that $\langle X^k, S^k\rangle \leq \varepsilon\langle X^0, S^0\rangle$ in $\mathcal{O}(\kappa\sqrt{n}\log\varepsilon^{-1})$ iterations.*

Proof Let $\bar{\alpha}_0 = 4\beta\tau^2/(5\sqrt{\mathrm{cond}(G)}\sqrt{n})$. By the definition of $\bar{\alpha}$ in (20), we have

$$\begin{aligned}\mu(\bar{\alpha}) \leq \mu(\bar{\alpha}_0) &= \mu + \frac{\bar{\alpha}_0}{n}\mathrm{Tr}[[\tau\mu I - \tilde{X}\tilde{S}]^- + \sqrt{n}[\tau\mu I - \tilde{X}\tilde{S}]^+]\\ &\leq \mu + \bar{\alpha}^0(\tau\mu - \mu + \beta\tau\mu) = [1 - (1 - \tau - \beta\tau)\bar{\alpha}^0]\mu\\ &= (1 - \xi\bar{\alpha}^0)\mu,\end{aligned}$$

where $\xi = 1 - \tau - \beta\tau$, the inequality follows from Lemma 5.3 and Corollary 5.11.

Because we need to have $\mu(\bar{\alpha}) \leq \varepsilon\mu_0$, it suffices to have

$$\left[1 - \frac{4\beta\tau^2\xi}{5\sqrt{\mathrm{cond}(G)}\sqrt{n}}\right]^k \mu_0 \leq \left[1 - \frac{4\beta\tau^2\xi}{5\kappa\sqrt{n}}\right]^k \mu_0 = \left[1 - \frac{\xi_0}{\kappa\sqrt{n}}\right]^k \mu_0 \leq \varepsilon\mu_0,$$

where $\xi_0 = 4\beta\tau^2\xi/5$.

Substitution gives $k \geq (\kappa\sqrt{n}\log\varepsilon^{-1})/\xi_0$. ■

In order to obtain polynomial complexity for our Algorithm 4.2, we give the next lemma, which is the bound on $\mathrm{cond}(G)$.

LEMMA 5.14 [19, Lemma 3.1] *If the NT search direction is used, then* $\mathrm{cond}(G) = 1$. *If the HKM search directions are used, then* $\mathrm{cond}(G) \leq n/(1 - \beta)\tau$.

By using Lemma 5.14, we have the following iteration complexities.

COROLLARY 5.15 *If the NT search direction is used, then the iteration complexity of Algorithm 4.2 is $\mathcal{O}(\sqrt{n}\log\varepsilon^{-1})$. If the HKM search directions are used, then the iteration complexities of Algorithm 4.2 are $\mathcal{O}(n\log\varepsilon^{-1})$.*

6. Computational results

We compare the proposed Algorithm 4.2 with the Algorithm 2 in [12] for some semi-definite problems which are generated randomly. These test problems are random SDP (Random), Max-Cut problem (Max-Cut), educational testing problem (ETP), norm minimization problem (Norm Minim), see [26] for details.

We only implemented the NT scaling for the test problems. All of our tests are run on an Intel Core i5 PC (3.10 GHz) under Windows 7 and MATLAB R2011(b). We select the optimization parameter $\tau = 0.05, \beta = 0.01$ for the Algorithm 4.2 and $\tau_1 = 0.1, \tau_2 = 0.05, \eta = 100$ for the Algorithm 2. The proposed algorithm terminates after the normalized duality gap satisfies $\langle X^k, S^k\rangle/n \leq 10^{-10} \times \langle X^0, S^0\rangle/n$. We list the names of the test problems, the number of iterations

Table 1. Computational results.

			Algorithm 1			Algorithm 2		
Problem	m	n	Iter	Gap	Time	Iter	Gap	Time
Random	100	100	21.5	6.2517e−007	4.50	23.5	1.3363e−006	4.33
	200	100	23.7	7.0592e−007	8.69	26.2	7.6678e−007	9.02
	200	300	25.7	5.6228e−006	120.96	27.0	9.8993e−006	121.03
Max-Cut	50	50	11.4	4.5436e−010	0.28	23.9	7.4568e−010	0.51
	100	100	11.7	1.0879e−009	1.81	26.2	2.1322e−009	3.74
	200	200	13.4	1.4075e−009	13.85	27.6	3.6584e−009	26.81
ETP	25	50	25.3	1.6004e−009	0.50	33.2	2.5040e−009	0.52
	50	100	35.3	2.6207e−009	3.60	39.5	3.2204e−009	3.42
	100	200	46.9	5.1276e−009	29.04	50.8	7.4809e−009	28.44
Norm Minim	100	100	12.1	2.0761e−011	2.69	25.9	4.6584e−011	5.42
	100	200	12.3	3.2084e−011	10.82	28.2	5.2465e−011	23.34
	200	200	12.4	1.9209e−011	19.53	27.3	5.3120e−011	41.40

(iter), the duality gap (gap) when the algorithms terminate and the CPU time (time) in seconds. Moreover, m represents the number of the constraint equations and n stands for the dimension of the block $A_i \in S^{n\times n}$. We run 10 times for the same m and n and show the results below. From the results in Table 1 we find that, on the average, the number of iterations is about 32.36%, less than the Algorithm 2. Although our implementations are very coarse, the proposed algorithm is comparable to Algorithm 2 as a whole.

7. Conclusion

In this paper, we proved the relationship $\|UV + VU\|_1 \leq 2\|U\|_F\|V\|_F$. Using the relationship, we showed the convergence of our algorithm for a specific class of search directions. Moreover, our numerical experiments also provide us an encouraging evidence that our new algorithm may also perform well in practice.

Acknowledgements

We are grateful to the anonymous referees and editor for their useful comments that help us improve the presentation of this paper. We would also like to thank the supports of National Natural Science Foundation of China (NNSFC) under grant nos. 61072144 and 61179040.

References

[1] W. Ai and S. Zhang, *An $O(\sqrt{n}L)$ iteration primal–dual path-following method, based on wide neighborhoods and large updates, for monotone LCP*, SIAM J. Optim. 16 (2005), pp. 400–417.
[2] F. Alizadeh, *Combinatorial optimization with interior-point methods and semi-definite matrices*, Ph.D. thesis, Computer Science Department, University of Minnesota, Minneapolis, MN, 1991.
[3] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim. 5 (1995), pp. 13–51.
[4] F. Alizadeh, J.A. Haeberly, and M.L. Overton, *Primal–dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results*, SIAM J. Optim. 8 (1998), pp. 746–768.
[5] S. Boyd, L.E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, PA, 1994.
[6] Z. Feng and L. Fang, *A wide neighbourhood interior-point method with iteration-complexity bound for semidefinite programming*, Optimization 59 (2010), pp. 1235–1246.
[7] G.H. Gulub and C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1989.

[8] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM J. Optim. 6 (1996), pp. 342–361.

[9] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

[10] E.D. Klerk, *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, Kluwer Academic Publishers, Dordrecht, 2002.

[11] M.H. Koulaei and T. Terlaky, *On the complexity analysis of a Mehrotra-type primal–dual feasible algorithm for semidefinite optimization*, Optim. Methods Softw. 25 (2010), pp. 467–485.

[12] Y. Li and T. Terlaky, *A new class of large neighborhood path-following interior point algorithms for semidefinite optimization with* $O(\sqrt{n}\log\frac{\mathrm{Tr}(X^0S^0)}{\varepsilon})$ *iteration complexity*, SIAM J. Optim. 20 (2010), pp. 2853–2875.

[13] C. Liu and H. Liu, *A new second-order corrector interior-point algorithm for semidefinite programming*, Math. Methods Oper. Res. 75 (2012), pp. 165–183.

[14] H. Liu, C. Liu, and X. Yang, *New complexity analysis of a Mehrotra-type predictor–corrector algorithm for semidefinite programming*, Optim. Methods Softw. 1 (2012), pp. 1–16.

[15] H. Liu, X. Yang, and C. Liu, *A new wide neighborhood primal–dual infeasible-interior-point method for symmetric cone programming*, J. Optim. Theory Appl. (2013). DOI:10.1007/s10957-013-0303-y.

[16] S. Mehrotra, *On the implementation of a primal–dual interior point method*, SIAM J. Optim. 2 (1992), pp. 575–601.

[17] R.D.C. Monteiro, *Primal–dual path-following algorithms for semidefinite programming*, SIAM J. Optim. 7 (1997), pp. 663–678.

[18] R.D.C. Monteiro, *Polynomial convergence of primal–dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions*, SIAM J. Optim. 8 (1998), pp. 797–812.

[19] R.D.C. Monteiro and Z. Yin, *A unified analysis for a class of long-step primal–dual path-following interior-point algorithms for semidefinite programming*, Math. Program. 81 (1998), pp. 281–299.

[20] Y.E. Nesterov and A. Nemirovsky, *Interior-Point Methods in Convex Programming: Theory and Applications*, SIAM, Philadelphia, PA, 1994.

[21] Y. Nesterov and M. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Math. Oper. Res. 22 (1997), pp. 1–42.

[22] Y. Nesterov and M. Todd, *Primal–dual interior-point methods for self-scaled cones*, SIAM J. Optim. 8 (1998), pp. 324–364.

[23] M. Overton, *On minimizing the maximum eigenvalue of a symmetric matrix*, SIAM. J. Matrix Anal. Appl. 9 (1988), pp. 256–268.

[24] M. Overton, *Large-scale optimization of eigenvalues*, SIAM J. Optim. 2 (1992), pp. 88–120.

[25] J. Peng, C. Roos, and T. Terlaky, *Self-regular functions and new search directions for linear and semidefinite optimization*, Math. Program. 93 (2002), pp. 129–171.

[26] M.J. Todd, K.C. Toh, and R.H. Tütüncü, *On the Nesterov–Todd direction in semidefinite programming*, SIAM J. Optim. 8 (1998), pp. 769–796.

[27] G. Wang and Y. Bai, *A new primal–dual path-following interior-point algorithm for semidefinite optimization*, J. Math. Anal. Appl. 353 (2009), pp. 339–343.

[28] G. Wang and Y. Bai, *Primal–dual interior-point algorithms for convex quadratic semidefinite optimization*, Nonlinear Anal. 71 (2009), pp. 3389–3402.

[29] Y. Ye, *A class of projective transformations for linear programming*, SIAM J. Comput. 19 (1990), pp. 457–466.

[30] Y. Zhang, *On extending some primal–dual interior-point algorithms from linear programming to semidefinite programming*, SIAM J. Optim. 8 (1998), pp. 365–386.

[31] Y. Zhang and D. Zhang, *On polynomial of the Mehrotra-type predictor–corrector interior-point algorithms*, Math. Program. 68 (1995), pp. 303–318.