Optimization Methods & Software, 2013
Vol. 28, No. 6, 1179–1194, http://dx.doi.org/10.1080/10556788.2012.679270

New complexity analysis of a Mehrotra-type predictor–corrector algorithm for semidefinite programming

Hongwei Liu[a], Changhe Liu[a,b]* and Ximei Yang[a]

[a]*Department of Mathematics, Xidian University, Xi'an, People's Republic of China;*
[b]*Department of Applied Mathematics, Henan University of Science and Technology, Luoyang, People's Republic of China*

(*Received 7 December 2010; final version received 20 March 2012*)

In this paper, we propose a new Mehrotra-type predictor–corrector interior-point algorithm for semidefinite programming. This algorithm is an extension of the variant of Mehrotra-type algorithm that was proposed by Salahi *et al.* [*On Mehrotra-type predictor–corrector algorithms*, SIAM J. Optim. 18 (2007), pp. 1377–1397] for linear programming problems. We modify the step sizes lightly in the predictor step of Koulaei and Terlaky [*On the complexity analysis of a Mehrotra-type primal–dual feasible algorithm for semidefinite optimization*, Optim. Methods Softw. 25 (2010), pp. 467–485]. In such a way, we obtain $O(n\log(\mathrm{Tr}(X^0S^0)/\varepsilon))$ iteration complexity of the algorithm, where (X^0, y^0, S^0) is the initial feasible point and ε is the required precision.

Keywords: semidefinite programming; interior-point methods; predictor–corrector algorithm; Mehrotra-type algorithm; polynomial complexity

AMS Subject Classification: 65K05; 90C22; 90C51

1. Introduction

Semidefinite programming (SDP) is a generalization of linear programming (LP). It has received considerable attention and has been one of the most active research areas in mathematical programming. SDP has been applied in many areas, such as combinatorial optimization [2,3] and system and control theory [5]. Due to the success of interior-point methods (IPMs) in solving LP, most IPM variants were extended to SDP. The first IPMs for SDP were independently developed by Alizadeh [2] and Nesterov and Nemirovskii [19]. Alizadeh [2] extended Ye's [29] projective potential reduction algorithm from LP to SDP and argued that many known interior-point algorithms for LP could be transformed into algorithms for SDP. On the other hand, Nesterov and Nemirovskii [19] and Nesterov and Todd [20] presented a deep and unified theory of IPMs for solving the more general conic optimization problems using the notation of self-concordant barriers. Other IPMs for solving SDP can be found, for example, in [4,6,8,12,14,17,18,21,22].

In LP, the most computationally successful IPMs have been primal–dual methods using Mehrotra's [16] predictor–corrector (MPC) steps. MPC algorithms for SDP have been implemented in the softwares SeDuMi by Sturm [24], SDPT3 by Toh *et al.* [27] and SDPA

*Corresponding author. Email: changheliu@163.com, hwliu@mail.xidian.edu.cn

by Fujisawa *et al.* [7]. In spite of the extensive use of this variant in IPM-based optimization packages, not much was known about its complexity before the recent paper by Salahi *et al.* [23], which presents a new variant of the Mehrotra-type predictor–corrector algorithm for LP. This variant incorporates a safeguard in the algorithm that keeps the iterates in the prescribed neighbourhood and allows us to get a reasonably large step size. This safeguard strategy is used also when the affine scaling step performs poorly, which effectively forces the algorithm to take pure centring steps. They proved that the modified algorithm, in the worst case, will terminate after at most $O(n^2L)$ iterations, where n is the number of variables and L is the input data length. By slightly modifying the Newton system in the corrector step, the iteration complexity was reduced to $O(nL)$. Their numerical results also show that the safeguard-based algorithm has a superior computational performance in real applications. Based on Ai and Zhang's work [1], Liu *et al.* [15] proposed a Mehrotra-type primal–dual second-order corrector algorithm with a fixed centre parameter. They proved that the algorithm stops after at most $O(\sqrt{n}L)$.

Recently, Koulaei and Terlaky [13] extended the Mehrotra-type algorithm of Salahi *et al.* [23] for SDP, based on the Nesterov–Todd (NT) direction [20,21], and showed that the iteration complexity bound of the algorithm is of the same order as that of the corresponding algorithm for LP. However, it needs to be pointed out that the proof of the $O(nL)$ iteration complexity in [13] is incorrect. In fact, one cannot get Lemma 4.2, which is the key lemma in the proof of the $O(nL)$ iteration complexity, by using Theorem 4.1 in [13]. In this paper, we propose a new Mehrotra-type predictor–corrector algorithm by slightly modifying the maximum step size in the predictor step of Koulaei and Terlaky [13]. However, in the case of LP, the new predictor step size is still identical with that used in [23]. We prove that the iteration complexity of the new algorithm is $O(nL)$ based on the NT direction, which is analogous to the case of LP.

We organize our paper as follows. In Section 2, we review the so-called similar symmetrization operator introduced by Zhang [30], which leads to a class of search directions. Among these directions, we would like to highlight the one originally proposed by Nesterov and Todd [20,21] which is a member of this class. Based on the NT direction, we proposed a modified Mehrotra-type predictor–corrector algorithm by slightly modifying the maximum step size in the predictor step of [13]. In Section 3, by using the machinery of the Lyapunov operator, we first demonstrate several technical lemmas, and then give the $O(n\log(\mathrm{Tr}(X^0S^0)/\varepsilon))$ iteration complexity of this new algorithm. Finally, some conclusions are given in Section 4.

The following notations are used throughout the paper. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. The set of all $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m\times n}$. A^{T} denotes the transpose of $A \in \mathbb{R}^{m\times n}$. The set of all symmetric $n \times n$ real matrices is denoted by $\mathcal{S}^n$. For $M \in \mathcal{S}^n$, we write $M \succ 0$ ($M \succeq 0$) if M is positive definite (positive semidefinite). $\mathcal{S}^n_{++}$ ($\mathcal{S}^n_{+}$) denotes the set of all matrices in $\mathcal{S}^n$ which are positive definite (positive semidefinite). For a matrix M with all real eigenvalues, we denote its eigenvalues by $\lambda_i(M), i = 1, 2, \dots, n$, and its smallest and largest eigenvalues by $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$, respectively. The Hermitian part of M is denoted by $H(M) := \frac{1}{2}(M + M^*)$, where $M^* = \bar{M}^{\mathrm{T}}$ denotes the Hermitian adjoint. Given $G, H \in \mathbb{R}^{m\times n}$, the inner product between them is defined as $\langle G, H\rangle := \mathrm{Tr}(G^{\mathrm{T}}H)$, the trace of the matrix $G^{\mathrm{T}}H$. The Frobenius norm of $M \in \mathbb{R}^{n\times n}$ is $\|M\|_{\mathrm{F}} := \langle M, M\rangle^{1/2}$.

2. SDP problem and preliminary discussions

We consider the following SDP problem:

$$\text{(P)} \quad \min\langle C, X\rangle \quad \text{s.t. } \langle A_i, X\rangle = b_i, \quad i = 1, 2, \dots, m, \;\; X \succeq 0, \tag{1}$$

where $C, X \in \mathcal{S}^n, b \in \mathbb{R}^m$ and $A_i \in \mathcal{S}^n, i = 1, 2, \ldots, m$, are linearly independent. We call problem (P) the primal form of SDP problem, and X is the primal variable.

Corresponding to every primal problem (P), there exists a dual problem (D)

$$\text{(D)} \quad \max b^{\mathrm{T}} y \quad \text{s.t.} \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0, \tag{2}$$

where $y \in \mathbb{R}^m$ and $S \in \mathcal{S}^n$ are the dual variables.

The set of primal dual feasible solutions is denoted by

$$\mathcal{F} := \left\{ (X, y, S) \in \mathcal{S}^n_+ \times \mathbb{R}^m \times \mathcal{S}^n_+ : \begin{array}{c} \langle A_i, X \rangle = b_i, \quad i = 1, 2, \ldots, m \\ \sum_{i=1}^{m} y_i A_i + S = C \end{array} \right\}$$

and the relative interior of the primal–dual feasible set is

$$\mathcal{F}^0 := \left\{ (X, y, S) \in \mathcal{S}^n_{++} \times \mathbb{R}^m \times \mathcal{S}^n_{++} : \begin{array}{c} \langle A_i, X \rangle = b_i, \quad i = 1, 2, \ldots, m \\ \sum_{i=1}^{m} y_i A_i + S = C \end{array} \right\}.$$

It is well known [11] that under the assumptions that $\mathcal{F}^0$ is non-empty and the matrices A_i, $i = 1, 2, \ldots, m$, are linearly independent, then X^* and (y^*, S^*) are optimal if and only if they satisfy the optimality conditions

$$\begin{aligned} \langle A_i, X \rangle = b_i, \quad i = 1, 2, \ldots, m, \;\; X \succeq 0, \\ \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0, \\ XS = 0, \end{aligned} \tag{3}$$

where the last equality is called the complementarity equation. The central path consists of points (X^μ, y^μ, S^μ) satisfying the perturbed system

$$\begin{aligned} \langle A_i, X \rangle = b_i, \quad i = 1, 2, \ldots, m, \;\; X \succeq 0, \\ \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0, \\ XS = \mu I, \end{aligned} \tag{4}$$

where $\mu \in \mathbb{R}$, $\mu > 0$. It is proved in [12,19] that there is a unique solution (X^μ, y^μ, S^μ) to the central path equations (4) for any barrier parameter $\mu > 0$, assuming that the $\mathcal{F}^0$ is non-empty and the coefficient matrices A_i, $i = 1, 2, \ldots, m$, are linearly independent. Moreover, the limit point (X^*, y^*, S^*) as μ goes to zero is a primal–dual optimal solution of the corresponding SDP problem.

Since for $X, S \in \mathcal{S}^n$, the product XS is generally not in $\mathcal{S}^n$, the left-hand side of (4) is a map from $\mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ to $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^n$. Thus, the system (4) is not a square system when X and S are restricted to $\mathcal{S}^n$, which it is needed for applying Newton-like methods. A remedy for this is to make the perturbed optimality system (4) square by modifying the left-hand side to a map from $\mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ to itself. To this end, we use the so-called similar symmetrization operator $H_P : \mathbb{R}^{n \times n} \to \mathcal{S}^n$ introduced by Zhang [30] defined as

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

where $P \in \mathbb{R}^{n\times n}$ is some non-singular matrix. In particular, when $P = I$,

$$H(M) := H_I(M)$$

is the Hermitian part of M. Zhang [30] also observed that

$$H_P(M) = \mu I \Leftrightarrow M = \mu I,$$

for any non-singular matrix P and any matrix M with real spectrum. Thus, for any given non-singular matrix P, system (4) is equivalent to

$$\begin{aligned} \langle A_i, X\rangle &= b_i, \quad i = 1,2,\dots,m, \;\; X \succeq 0,\\ \sum_{i=1}^{m} y_i A_i + S &= C, \quad S \succeq 0,\\ H_P(XS) &= \mu I. \end{aligned} \tag{5}$$

A Newton-like method applied to system (5) leads to the following linear system:

$$\begin{aligned} \langle A_i, \triangle X\rangle &= 0, \quad i = 1,2,\dots,m,\\ \sum_{i=1}^{m} A_i \triangle y_i + \triangle S &= 0,\\ H_P(X\triangle S + \triangle XS) &= \mu I - H_P(XS), \end{aligned} \tag{6}$$

where $(\triangle X, \triangle y, \triangle S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ is the search direction. Todd *et al.* [26] proved that system (6) has a unique solution for any $(X, y, S) \in \mathcal{S}^n_{++} \times \mathbb{R}^m \times \mathcal{S}^n_{++}$ and for the scaling matrix P which satisfies $PXSP^{-1} \in \mathcal{S}^n$. In particular, as was shown in [20,21,26], the choice of $P = W^{1/2}$ in (6), where

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

leads to the NT direction. It is easier to show that $PXSP^{-1} \in \mathcal{S}^n$ whenever X and S are positive definite and P is NT scaling matrix. In fact, for NT scaling, we have $PXP = P^{-1}SP^{-1}$ [18]. We mention that the NT direction can also be interpreted by the v-space notion, as was analysed by Sturm and Zhang [25], which provides a possibility of deriving many more search directions other than the NT direction.

In what follows, we describe a Mehrotra-type predictor–corrector interior-point algorithm. Most efficient IPM solvers work in the so-called negative infinity neighbourhood that is a wide neighbourhood, defined by

$$\mathcal{N}^-_\infty(\gamma) := \{(X, y, S) \in \mathcal{F}^0 : \lambda_{\min}(XS) \geq \gamma\mu_g\},$$

where $\gamma \in (0, 1)$ is a constant independent of n and $\mu_g = \langle X, S\rangle/n$ is the normalized duality gap corresponding to (X, y, S). In this paper, we consider algorithms that are working in $\mathcal{N}^-_\infty(\gamma)$. The Mehrotra-type algorithm, in the predictor step, computes the affine scaling search direction, that is,

$$\langle A_i, \triangle X^{\mathrm{a}}\rangle = 0, \quad i = 1,2,\dots,m,$$

$$\sum_{i=1}^{m} A_i \triangle y_i^{\mathrm{a}} + \triangle S^{\mathrm{a}} = 0,$$

$$H_P(X\triangle S^{\mathrm{a}} + \triangle X^{\mathrm{a}} S) = -H_P(XS), \tag{7}$$

then it computes the maximum step size $\alpha_{\mathrm{a}} \in (0, 1]$ that ensures

$$H_P((X + \alpha_{\mathrm{a}}\triangle X^{\mathrm{a}})(S + \alpha_{\mathrm{a}}\triangle S^{\mathrm{a}})) \succeq 0. \tag{8}$$

However, the algorithm does not take such a step right away. We use the information from the predictor step to compute the corrector direction that is defined as follows:

$$\langle A_i, \triangle X\rangle = 0, \quad i = 1, 2, \ldots, m,$$

$$\sum_{i=1}^{m} A_i \triangle y_i + \triangle S = 0,$$

$$H_P(X\triangle S + \triangle XS) = \mu I - H_P(XS) - \alpha_{\mathrm{a}} H_P(\triangle X^{\mathrm{a}} \triangle S^{\mathrm{a}}), \tag{9}$$

where the centring parameter μ is defined adaptively by

$$\mu = \left(\frac{\langle X + \alpha_{\mathrm{a}}\triangle X^{\mathrm{a}}, S + \alpha_{\mathrm{a}}\triangle S^{\mathrm{a}}\rangle}{\langle X, S\rangle}\right)^3 \mu_g.$$

Since $\langle \triangle X^{\mathrm{a}}, \triangle S^{\mathrm{a}}\rangle = 0$, the previous relation implies

$$\mu = (1 - \alpha_{\mathrm{a}})^3 \mu_g. \tag{10}$$

From (10) it is obvious that if only a small step in the affine scaling direction can be made, then we only improve the centrality of the iterate. Finally, the new iterate is given by

$$(X(\alpha_c), y(\alpha_c), S(\alpha_c)) := (X, y, S) + \alpha_c(\triangle X, \triangle y, \triangle S), \tag{11}$$

where $\alpha_c \in (0, 1]$ is the maximum step size that keeps $(X(\alpha_c), y(\alpha_c), S(\alpha_c))$ in $\mathcal{N}_\infty^-(\gamma)$, where $\mu_g(\alpha_c) = \langle X(\alpha_c), S(\alpha_c)\rangle/n$.

In the case of LP, it has been shown in [23] by an example that Mehrotra's heuristic may force the algorithm to make very small steps to keep the iterates in a certain neighbourhood of the central path, which may cause too many iterations to converge. This may also happen when the affine scaling step size is very small, in which case the algorithm might make pure centring steps that only marginally reduce the duality gap. Therefore, Mehrotra's adaptive updating scheme of the centring parameter has to be combined with certain safeguards to get a warranted step size at each iteration. This variant of the Mehrotra's algorithm for SDP can be stated as follows.

ALGORITHM 1 (Mehrotra-type predictor–corrector algorithm) *Input an accuracy parameter* $\varepsilon > 0$, *a neighbourhood parameters* $\gamma \in (0, \frac{1}{2})$, *and an initial point* $(X^0, y^0, S^0) \in \mathcal{N}_\infty^-(\gamma)$. *Set* $\mu_g^0 = \langle X^0, S^0\rangle/n$ *and* $k := 0$.

while $\langle X^k, S^k\rangle > \varepsilon$ **do:**

(1) *Compute the NT scaling matrix* P^k.
(2) (*Predictor step*) *Solve* (7) *and compute the maximum step size* α_{a}^k *by* (8).
(3) (*Corrector step*) *If* $\alpha_{\mathrm{a}}^k \geq 0.1$, *then solve* (9) *with* $\mu^k = (1 - \alpha_{\mathrm{a}}^k)^3 \mu_g^k$ *and compute the maximum step size* α_c^k *such that* $(X(\alpha_c^k), y(\alpha_c^k), S(\alpha_c^k)) \in \mathcal{N}_\infty^-(\gamma)$; *If* $\alpha_c^k < 3\gamma/(5n)$ *or* $\alpha_{\mathrm{a}}^k < 0.1$,

then solve (9) *with* $\mu^k = \gamma\mu_g^k/(1-\gamma)$ *and compute the maximum step size* α_c^k *such that* $(X(\alpha_c^k), y(\alpha_c^k), S(\alpha_c^k)) \in \mathcal{N}_\infty^-(\gamma)$.

(4) *Let* $(X^{k+1}, y^{k+1}, S^{k+1}) = (X(\alpha_c^k), y(\alpha_c^k), S(\alpha_c^k))$, $\mu_g^{k+1} = \langle X^{k+1}, S^{k+1}\rangle/n$, *and set* $k := k+1$.

end while

Remark 1 Our new algorithm is different from the modified algorithm of Koulaei and Terlaky [13] only in the computation of the step size $\alpha_{\rm a}$ in the predictor step. The motivation for this variant is based on the following lemma.

LEMMA 2.1 *Suppose that* $(X, y, S) \in \mathcal{F}^0$ *and non-singular matrix* P *satisfies* $PXSP^{-1} \in \mathcal{S}^n$. *Let* $(\triangle X^{\rm a}, \triangle y^{\rm a}, \triangle S^{\rm a})$ *be the solution of* (7) *and denote*

$$\bar{X}(\alpha) := X + \alpha\triangle X^{\rm a}, \quad \bar{S}(\alpha) := S + \alpha\triangle S^{\rm a}.$$

If $\alpha_{\rm a} \in (0, 1]$ *is the step size such that* $H_P(\bar{X}(\alpha_{\rm a})\bar{S}(\alpha_{\rm a})) \succeq 0$, *then we have*

$$H_P(\bar{X}(\alpha)\bar{S}(\alpha)) \succ 0 \quad \forall\alpha \in [0, \alpha_{\rm a}). \tag{12}$$

Moreover, we have

$$\bar{X}(\alpha) \succ 0, \quad \bar{S}(\alpha) \succ 0 \quad \forall\alpha \in [0, \alpha_{\rm a}) \quad \textit{and} \quad \bar{X}(\alpha_{\rm a}) \succeq 0, \quad \bar{S}(\alpha_{\rm a}) \succeq 0. \tag{13}$$

Proof We first note

$$H_P(\bar{X}(0)\bar{S}(0)) = H_P(XS) \succ 0.$$

In fact, since $PXSP^{-1} \in \mathcal{S}^n$, we have $H_P(XS) = PXSP^{-1}$. By similarity,

$$\lambda_{\min}(H_P(XS)) = \lambda_{\min}(XS) = \lambda_{\min}(X^{1/2}SX^{1/2}) > 0.$$

By the third equation of system (7), we have

$$H_P(\bar{X}(\alpha)\bar{S}(\alpha)) = (1-\alpha)H_P(XS) + \alpha^2 H_P(\triangle X^{\rm a}\triangle S^{\rm a}).$$

Then, for all $\alpha \in (0, \alpha_{\rm a})$, we have

$$\begin{aligned} H_P(\bar{X}(\alpha)\bar{S}(\alpha)) &= \alpha^2\left(\frac{1-\alpha}{\alpha^2}H_P(XS) + H_P(\triangle X^{\rm a}\triangle S^{\rm a})\right) \\ &= \alpha^2\left(\frac{1-\alpha}{\alpha^2} - \frac{1-\alpha_{\rm a}}{\alpha_{\rm a}^2}\right)H_P(XS) + \frac{\alpha^2}{\alpha_{\rm a}^2}H_P(\bar{X}(\alpha_{\rm a})\bar{S}(\alpha_{\rm a})). \end{aligned}$$

Let $f(t) = (1-t)/t^2,\ t \in (0, 1]$, then $f(t)$ is a strict monotone decreasing function in $(0, 1]$. For all $\alpha \in (0, \alpha_{\rm a})$, denote $\delta(\alpha) = f(\alpha) - f(\alpha_{\rm a})$, then we have $\delta(\alpha) > 0$ and

$$H_P(\bar{X}(\alpha)\bar{S}(\alpha)) = \alpha^2\delta(\alpha)H_P(XS) + \frac{\alpha^2}{\alpha_{\rm a}^2}H_P(\bar{X}(\alpha_{\rm a})\bar{S}(\alpha_{\rm a})).$$

Using the fact that $\lambda_{\min}(\cdot)$ is a homogeneous concave function on the space of symmetric matrices, one has

$$\lambda_{\min}(H_P(\bar{X}(\alpha)\bar{S}(\alpha))) \ge \alpha^2\delta(\alpha)\lambda_{\min}(H_P(XS)) + \frac{\alpha^2}{\alpha_{\rm a}^2}\lambda_{\min}(H_P(\bar{X}(\alpha_{\rm a})\bar{S}(\alpha_{\rm a}))) > 0,$$

which gives the required result of (12).

By Lemma A.1, that

$$\lambda_{\min}(H(P\bar{X}(\alpha)\bar{S}(\alpha)P^{-1})) = \lambda_{\min}(H_P(\bar{X}(\alpha)\bar{S}(\alpha))) > 0$$

reveals $P\bar{X}(\alpha)\bar{S}(\alpha)P^{-1}$ is non-singular, and further implies that $\bar{X}(\alpha)$ and $\bar{S}(\alpha)$ are non-singular as well. By using continuity of the eigenvalues of a symmetric matrix (see [11, p. 231], Theorem A.5), it follows that $\bar{X}(\alpha), \bar{S}(\alpha) \succ 0$ for all $\alpha \in [0, \alpha_a)$, and $\bar{X}(\alpha_a), \bar{S}(\alpha_a) \succeq 0$, since $X, S \succ 0$. ■

We note that, the maximum α_a that satisfies the condition (13) is the maximum feasible step size used in [13]. In general, for two symmetric matrixes A and B, the condition $A \succ 0$ and $B \succ 0$ does not implicate $H(AB) \succ 0$. Therefore, the step size α_a computed by (8) may be smaller than that used in [13]. However, it is obvious that, in the case of LP, the step size α_a computed by (8) is identical with that used in [13,23].

3. Complexity analysis of the algorithm

In this part, we present the convergence and complexity proofs for Algorithm 1. First, we scale problems (P) and (D) as Li and Terlaky proposed in [14] in order to simplify the proofs of the main results. At the end of this section, after demonstrating several technical lemmas, we prove that Algorithm 1 has an iteration-complexity bound of $O(n \log(\mathrm{Tr}(X^0S^0)/\varepsilon))$ based on the NT direction.

3.1 *Scaling procedure*

For the scaling matrix P satisfying $PXSP^{-1} \in \mathcal{S}^n$, we scale the primal and dual variables of problems (P) and (D) in the form of

$$\hat{X} := PXP, \quad (\hat{y}, \hat{S}) := (y, P^{-1}SP^{-1}). \tag{14}$$

Hence, one has $\hat{X}\hat{S} = \hat{S}\hat{X}$, that is, $\hat{X}$ and $\hat{S}$ become commutable after scaling. To keep consistency, we also have to apply the same scaling to the other data in (P) and (D) as well, that is,

$$\hat{C} := P^{-1}CP^{-1}, \quad (\hat{A}_i, \hat{b}_i) := (P^{-1}A_iP^{-1}, b_i) \quad \text{for } i = 1, 2, \ldots, m.$$

From now on, we use Λ to denote the diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, where λ_i for $i = 1, 2, \ldots, n$ are the eigenvalues of $\hat{X}\hat{S}$ with increasing order, that is,

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n.$$

We should emphasize that the matrices $\hat{X}\hat{S}, \hat{S}\hat{X}, XS, SX, X^{1/2}SX^{1/2}$ and $S^{1/2}XS^{1/2}$ have the same eigenvalues, since they are all similar to each other.

The search directions based on system (7) and (9) correspond to the scaled directions defined as

$$\triangle\hat{X}^a = P\triangle X^a P, \quad \triangle\hat{y}^a = \triangle y^a, \quad \triangle\hat{S}^a = P^{-1}\triangle S^a P^{-1},$$
$$\triangle\hat{X} = P\triangle XP, \quad \triangle\hat{y} = \triangle y, \quad \triangle\hat{S} = P^{-1}\triangle SP^{-1}.$$

The directions $(\triangle\hat{X}^{\mathrm{a}}, \triangle\hat{y}^{\mathrm{a}}, \triangle\hat{S}^{\mathrm{a}})$ and $(\triangle\hat{X}, \triangle\hat{y}, \triangle\hat{S})$ are readily verified to be solutions of the scaled Newton systems

$$\begin{aligned} \langle \hat{A}_i, \triangle\hat{X}^{\mathrm{a}} \rangle = 0, \quad i = 1, 2, \ldots, m, \\ \sum_{i=1}^{m} \hat{A}_i \triangle\hat{y}_i^{\mathrm{a}} + \triangle\hat{S}^{\mathrm{a}} = 0, \\ H(\hat{X}\triangle\hat{S}^{\mathrm{a}} + \triangle\hat{X}^{\mathrm{a}}\hat{S}) = -\hat{X}\hat{S}, \end{aligned} \tag{15}$$

and

$$\begin{aligned} \langle \hat{A}_i, \triangle\hat{X} \rangle = 0, \quad i = 1, 2, \ldots, m, \\ \sum_{i=1}^{m} \hat{A}_i \triangle\hat{y}_i + \triangle\hat{S} = 0, \\ H(\hat{X}\triangle\hat{S} + \triangle\hat{X}\hat{S}) = \hat{\mu}I - \hat{X}\hat{S} - \alpha_{\mathrm{a}} H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}). \end{aligned} \tag{16}$$

The iterates are updated as follows:

$$(\hat{X}(\alpha), \hat{y}(\alpha), \hat{S}(\alpha)) := (\hat{X}, \hat{y}, \hat{S}) + \alpha(\triangle\hat{X}, \triangle\hat{y}, \triangle\hat{S}). \tag{17}$$

Since $H_P((X + \alpha_{\mathrm{a}}\triangle X^{\mathrm{a}})(S + \alpha_{\mathrm{a}}\triangle S^{\mathrm{a}})) = H((\hat{X} + \alpha_{\mathrm{a}}\triangle\hat{X}^{\mathrm{a}})(\hat{S} + \alpha_{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))$, condition (8) becomes

$$H((\hat{X} + \alpha_{\mathrm{a}}\triangle\hat{X}^{\mathrm{a}})(\hat{S} + \alpha_{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})) \succeq 0. \tag{18}$$

By applying similarity, we have

$$\hat{\mu}_g(\alpha) = \frac{\langle \hat{X}(\alpha), \hat{S}(\alpha) \rangle}{n} = \mu_g(\alpha), \tag{19}$$

and moreover,

$$(X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}_\infty^-(\gamma) \text{ if and only if } (\hat{X}(\alpha), \hat{y}(\alpha), \hat{S}(\alpha)) \in \mathcal{N}_\infty^-(\gamma).$$

3.2 *Lyapunov operator*

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

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

which is called the Lyapunov operator. We note the following well-known property (see [11, p. 250], Theorem E.2).

LEMMA 3.1 *Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathcal{S}^n$. The Lyapunov equation*

$$L_A(X) = B$$

has a unique symmetric solution if A and $-A$ have no eigenvalues in common.

Let $A \in \mathcal{S}^n_{++}$, then the Lyapunov operator L_A is guaranteed to be invertible. Use $L_A^{-1}(B)$ to denote the unique symmetric solution X to the equation $L_A(X) = B$. It is easy to show that, for $A \in \mathcal{S}^n_{++}$, we have

$$L_A^{-1}(A^2) = \tfrac{1}{2}A, \quad L_A^{-1}(A) = \tfrac{1}{2}I, \quad L_A^{-1}(I) = \tfrac{1}{2}A^{-1}.$$

Note that since $\mathrm{Tr}((AX + XA)Y) = \mathrm{Tr}(X(AY + YA))$, it follows that L_A is symmetric with respect to $\langle \cdot, \cdot \rangle$, that is,

$$\langle L_A(X), Y \rangle = \langle X, L_A(Y) \rangle.$$

We show that L_A^{-1} is also symmetric with respect to $\langle \cdot, \cdot \rangle$. In fact, for $L_A^{-1}(B_1) = X, L_A^{-1}(B_2) = Y$, we have

$$\langle L_A^{-1}(B_1), B_2 \rangle = \langle X, L_A(Y) \rangle = \langle L_A(X), Y \rangle = \langle B_1, L_A^{-1}(B_2) \rangle.$$

LEMMA 3.2 *Let* $A \in \mathcal{S}^n_{++}$. *If* $B \succeq 0$, *then* $L_A^{-1}(B) \succeq 0$.

Proof Suppose $L_A^{-1}(B) = X$, that is, $L_A(X) = AX + XA = B$. Then, by similarity and Lemma A.2, we have

$$\lambda_{\min}(A^{1/2}XA^{1/2}) = \lambda_{\min}(AX) \geq \lambda_{\min}(H(AX)) = \tfrac{1}{2}\lambda_{\min}(B) \geq 0.$$

Let $A^{1/2}XA^{1/2} = Y$, and therefore $Y \succeq 0$. Thus,

$$X = A^{-1/2}YA^{-1/2} \succeq 0,$$

which is the required result. ■

We observed that for NT scaling, we have $\hat{X} = \hat{S} := V$ [18]. Hence, in terms of the Lyapunov operator, the third equations of systems (15) and (16) become

$$L_V(\triangle\hat{X}^{\mathrm{a}} + \triangle\hat{S}^{\mathrm{a}}) = -2V^2,$$
$$L_V(\triangle\hat{X} + \triangle\hat{S}) = 2(\hat{\mu}I - V^2 - \alpha_{\mathrm{a}}H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})).$$

Therefore, by $V \in \mathcal{S}^n_{++}$, we have

$$\triangle\hat{X}^{\mathrm{a}} + \triangle\hat{S}^{\mathrm{a}} = -L_V^{-1}(2V^2) = -V, \tag{20}$$
$$\triangle\hat{X} + \triangle\hat{S} = \hat{\mu}V^{-1} - V - 2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})). \tag{21}$$

3.3 *Technical results*

Before proceeding to the complexity result, we have to prove some technical lemmas. Throughout this section, we consider P as the NT scaling matrix.

The difficulty in analysing Mehrotra-type algorithms arises from the second-order term $H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})$ in the corrector step of the algorithm, which is one of the main differences of this algorithm and other interior-point algorithms. To overcome this difficulty, we present some important technical lemmas that give the relationship between this term and the matrix $\hat{X}\hat{S}$. Before doing so, we need to present some notations.

Let $M := H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})) \in \mathcal{S}^n$ with the eigenvalue decomposition $M = QDQ^{\mathrm{T}}$, where $D = \mathrm{diag}\{d_1, d_2, \ldots, d_n\}$ is a diagonal matrix with all the eigenvalues of M along its diagonal, Q is an orthonormal matrix, that is, $QQ^{\mathrm{T}} = I$. Let the eigenvalues be arranged in increasing order, that is,

$$d_1 \leq d_2 \leq \cdots \leq d_{k-1} \leq 0 \leq d_k \leq \cdots \leq d_n.$$

Then, we define the negative part and the positive part as

$$M^- = QD^-Q^{\mathrm{T}}, \quad M^+ = QD^+Q^{\mathrm{T}},$$

where $D^- = \mathrm{diag}\{d_1, d_2, \ldots, d_{k-1}, 0, \ldots, 0\}$, $D^+ = \mathrm{diag}\{0, \ldots, 0, d_k, \ldots, d_n\}$. Apparently, $M = M^- + M^+$, where $-M^-, M^+ \succeq 0$.

Lemma 3.3 *Let* $(\hat{X}, \hat{y}, \hat{S}) \in \mathcal{F}^0$, $(\triangle\hat{X}^{\mathrm{a}}, \triangle\hat{y}^{\mathrm{a}}, \triangle\hat{S}^{\mathrm{a}})$ *be the solution of* (15), *and* α_{a} *be the maximum step size defined by* (18). *Then*

$$H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) \preceq \tfrac{1}{4}V^2$$

and

$$-\alpha_{\mathrm{a}}^2 H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) \preceq V^2.$$

Proof It is trivial to verify that

$$H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) = \tfrac{1}{4}((\triangle\hat{X}^{\mathrm{a}} + \triangle\hat{S}^{\mathrm{a}})^2 - (\triangle\hat{X}^{\mathrm{a}} - \triangle\hat{S}^{\mathrm{a}})^2).$$

Hence, by (20), we have

$$H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) = \tfrac{1}{4}(V^2 - (\triangle\hat{X}^{\mathrm{a}} - \triangle\hat{S}^{\mathrm{a}})^2)$$

from which the former result follows.

By (18) and (15), one has

$$\hat{X}\hat{S} + \alpha_{\mathrm{a}}(-\hat{X}\hat{S}) + \alpha_{\mathrm{a}}^2 H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) \succeq 0.$$

This is equivalent to

$$(1 - \alpha_{\mathrm{a}})V^2 + \alpha_{\mathrm{a}}^2 H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) \succeq 0$$

from which the letter statement follows. ∎

Lemma 3.4 *Let* $(\hat{X}, \hat{y}, \hat{S}) \in \mathcal{F}^0$ *and* $(\triangle\hat{X}^{\mathrm{a}}, \triangle\hat{y}^{\mathrm{a}}, \triangle\hat{S}^{\mathrm{a}})$ *be the solution of* (15), *then*

$$\sum_{i=1}^{k-1}(-d_i) = \sum_{i=k}^{n} d_i \leq \frac{1}{4}n\mu_g.$$

Proof By Lemma 3.3, we have

$$H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) - \frac{1}{4}V^2 \preceq 0.$$

Thus, by Lemma A.3, we have

$$d_i \leq \tfrac{1}{4}\lambda_i, \quad i = k, \ldots, n,$$

and hence,

$$\sum_{i=k}^{n} d_i \leq \frac{1}{4}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) = \frac{1}{4}n\mu_g.$$

By using $\mathrm{Tr}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})) = \mathrm{Tr}(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) = 0$, we complete the proof. ∎

LEMMA 3.5 *Let* $(\hat{X},\hat{y},\hat{S}) \in \mathcal{F}^0$ *and* $(\triangle\hat{X}^{\mathrm{a}},\triangle\hat{y}^{\mathrm{a}},\triangle\hat{S}^{\mathrm{a}})$ *be the solution of* (15), *then*

$$(L_V^{-1})^2(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})) \preceq \tfrac{1}{16}I$$

and

$$-(L_V^{-1})^2(\alpha_{\mathrm{a}}^2 H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})) \preceq \tfrac{1}{4}I.$$

Proof Since $(L_V^{-1})^2(V^2) = \frac{1}{4}I$, by Lemma 3.2, the proof is a direct consequence of Lemma 3.3. ■

The following lemma gives an upper bound for the second-order term on the right-hand side of (21).

LEMMA 3.6 *Let* $(\hat{X},\hat{y},\hat{S}) \in \mathcal{F}^0$ *and* $(\triangle\hat{X}^{\mathrm{a}},\triangle\hat{y}^{\mathrm{a}},\triangle\hat{S}^{\mathrm{a}})$ *be the solution of* (15), *then*

$$\|2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))\|_{\mathrm{F}}^2 \leq \tfrac{5}{16}n\hat{\mu}_g.$$

Proof Since L_A^{-1} is symmetric with respect to $\langle\cdot,\cdot\rangle$, it follows that

$$\begin{aligned}\|2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))\|_{\mathrm{F}}^2 &= \langle 2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})), 2\alpha_{\mathrm{a}}L_V^{-1}H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})\rangle \\ &= \langle 4\alpha_{\mathrm{a}}^2(L_V^{-1})^2(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})), H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})\rangle.\end{aligned}$$

By using

$$M := H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) = QDQ^{\mathrm{T}},$$

we have

$$\|2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))\|_{\mathrm{F}}^2 = \langle 4\alpha_{\mathrm{a}}^2 Q^{\mathrm{T}}(L_V^{-1})^2(M)Q, D\rangle.$$

Let

$$H := 4\alpha_{\mathrm{a}}^2 Q^{\mathrm{T}}(L_V^{-1})^2(M)Q,$$

and from Lemma 3.5, we obtain

$$-I \preceq H \preceq \tfrac{1}{4}I.$$

Hence, the diagonal elements of H satisfy $-1 \leq H_{ii} \leq \frac{1}{4}$, and consequently,

$$\begin{aligned}\|2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))\|_{\mathrm{F}}^2 &= \langle H, D^-\rangle + \langle H, D^+\rangle \\ &= \sum_{i=1}^{k-1} H_{ii}d_i + \sum_{i=k}^{n} H_{ii}d_i \\ &\leq \sum_{i=1}^{k-1}(-d_i) + \sum_{i=k}^{n}\left(\frac{d_i}{4}\right) \\ &\leq \frac{5}{16}n\hat{\mu}_g,\end{aligned}$$

where the last inequality follows from Lemma 3.4. ■

The following technical lemma will be used in the next theorem, which estimates the maximum step size in the corrector step.

LEMMA 3.7 *Let* $(\hat{X}, \hat{y}, \hat{S}) \in \mathcal{N}_{\infty}^{-}(\gamma)$. *Then, we have*

$$\|H(\triangle\hat{X}\triangle\hat{S})\|_{\mathrm{F}} \leq \frac{1}{2}\left[\left(\frac{1}{\gamma}\left(\frac{\hat{\mu}}{\hat{\mu}_g}\right)^2 + 1 - \frac{2\hat{\mu}}{\hat{\mu}_g}\right)^{1/2} + \frac{\sqrt{5}}{4}\right]^2 n\hat{\mu}_g.$$

Proof By using (21) and $\langle\triangle\hat{X}, \triangle\hat{S}\rangle = 0$, we have

$$\begin{aligned}\|\triangle\hat{X}\|_{\mathrm{F}}^2 + \|\triangle\hat{S}\|_{\mathrm{F}}^2 &= \|\hat{\mu}V^{-1} - V - 2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))\|_{\mathrm{F}}^2 \\ &\leq [\|\hat{\mu}V^{-1} - V\|_{\mathrm{F}} + \|2\alpha_{\mathrm{a}}L_V^{-1}(H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}))\|_{\mathrm{F}}]^2.\end{aligned}$$

By $(\hat{X}, \hat{y}, \hat{S}) \in \mathcal{N}_{\infty}^{-}(\gamma)$, it has

$$\lambda_i(V^{-1}) = \frac{1}{\sqrt{\lambda_i}} \leq \frac{1}{\sqrt{\gamma\hat{\mu}_g}}.$$

Hence,

$$\begin{aligned}\|\hat{\mu}V^{-1} - V\|_{\mathrm{F}}^2 &= \|\hat{\mu}V^{-1}\|_{\mathrm{F}}^2 + \|V\|_{\mathrm{F}}^2 - 2n\hat{\mu} \\ &\leq \frac{n\hat{\mu}^2}{(\gamma\hat{\mu}_g)} + n\hat{\mu}_g - 2n\hat{\mu} \\ &= \left(\frac{1}{\gamma}\left(\frac{\hat{\mu}}{\hat{\mu}_g}\right)^2 + 1 - \frac{2\hat{\mu}}{\hat{\mu}_g}\right) n\hat{\mu}_g.\end{aligned}$$

Therefore, by using Lemma 3.6, we obtain

$$\begin{aligned}\|H(\triangle\hat{X}\triangle\hat{S})\|_{\mathrm{F}} &\leq \|\triangle\hat{X}\triangle\hat{S}\|_{\mathrm{F}} \\ &\leq \|\triangle\hat{X}\|_{\mathrm{F}}\|\triangle\hat{S}\|_{\mathrm{F}} \\ &\leq \frac{1}{2}(\|\triangle\hat{X}\|_{\mathrm{F}}^2 + \|\triangle\hat{S}\|_{\mathrm{F}}^2) \\ &\leq \frac{1}{2}\left[\left(\frac{1}{\gamma}\left(\frac{\hat{\mu}}{\hat{\mu}_g}\right)^2 + 1 - \frac{2\hat{\mu}}{\hat{\mu}_g}\right)^{1/2} + \frac{\sqrt{5}}{4}\right]^2 n\hat{\mu}_g,\end{aligned}$$

which completes the proof. ■

The following corollary, which follows from Lemma 3.7, gives an explicit upper bound for a specific value of $\hat{\mu}$.

COROLLARY 3.8 *If* $\mu = \gamma\mu_g/(1-\gamma)\hat{\mu}_g$, *where* $0 \leq \gamma \leq \frac{1}{2}$, *then*

$$\|H(\triangle\hat{X}\triangle\hat{S})\|_{\mathrm{F}} \leq \tfrac{5}{4}n\hat{\mu}_g.$$

THEOREM 3.9 *Suppose that the current iterate* $(\hat{X}, \hat{y}, \hat{S}) \in \mathcal{N}_{\infty}^{-}(\gamma)$, *where* $\gamma \in (0, \frac{1}{2})$, *and let* $(\triangle\hat{X}, \triangle\hat{y}, \triangle\hat{S})$ *be the solution of* (16) *with* $\mu = \gamma\mu_g/(1-\gamma)\hat{\mu}_g$. *Then, the maximum step size* α_c, *that keeps* $(\hat{X}(\alpha_c), \hat{y}(\alpha_c), \hat{S}(\alpha_c)) \in \mathcal{N}_{\infty}^{-}(\gamma)$, *satisfies*

$$\alpha_c \geq \frac{3\gamma}{5n}.$$

Proof We need to estimate the maximum non-negative α that keeps the next iterate in $\mathcal{N}_\infty^-(\gamma)$, that is, $\hat{X}(\alpha), \hat{S}(\alpha) \in \mathcal{S}_{++}^n$ and

$$\lambda_{\min}(\hat{X}(\alpha)\hat{S}(\alpha)) \geq \gamma \hat{\mu}_g(\alpha). \tag{22}$$

By Equation (16), we have

$$\begin{aligned} H(\hat{X}(\alpha)\hat{S}(\alpha)) &= \hat{X}\hat{S} + \alpha H(\hat{X}\triangle\hat{S} + \triangle\hat{X}\hat{S}) + \alpha^2 H(\triangle\hat{X}\triangle\hat{S}) \\ &= \hat{X}\hat{S} + \alpha(\hat{\mu}I - \hat{X}\hat{S} - \alpha_{\mathrm{a}} H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}})) + \alpha^2 H(\triangle\hat{X}\triangle\hat{S}) \\ &= \alpha\hat{\mu}I + (1-\alpha)\hat{X}\hat{S} - \alpha\alpha_{\mathrm{a}} H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}) + \alpha^2 H(\triangle\hat{X}\triangle\hat{S}). \end{aligned} \tag{23}$$

Hence, by using $\mathrm{Tr}(H(M)) = \mathrm{Tr}(M)$, we obtain

$$\hat{\mu}_g(\alpha) = \frac{\mathrm{Tr}(H(\hat{X}(\alpha)\hat{S}(\alpha)))}{n} = \left(1 - \alpha + \alpha\frac{\hat{\mu}}{\hat{\mu}_g}\right)\hat{\mu}_g. \tag{24}$$

Let

$$G(\alpha) := (1-\alpha)\hat{X}\hat{S} - \alpha\alpha_{\mathrm{a}} H(\triangle\hat{X}^{\mathrm{a}}\triangle\hat{S}^{\mathrm{a}}),$$

then by Lemma 3.3, we have

$$G(\alpha) \succeq \left(1 - \alpha - \tfrac{1}{4}\alpha\alpha_{\mathrm{a}}\right)\hat{X}\hat{S} \succeq \left(1 - \tfrac{5}{4}\alpha\right)\hat{X}\hat{S}.$$

If $\alpha \geq \frac{4}{5}$, we are done. Otherwise, by Lemma A.3, we have

$$\lambda_{\min}(G(\alpha)) \geq \left(1 - \tfrac{5}{4}\alpha\right)\lambda_1 \geq \left(1 - \tfrac{5}{4}\alpha\right)\gamma\hat{\mu}_g.$$

It follows from Equation (23) and from the fact that $\lambda_{\min}(\cdot)$ is a homogeneous concave function on the space of symmetric matrices, that

$$\begin{aligned} \lambda_{\min}(H(\hat{X}(\alpha)\hat{S}(\alpha))) &\geq \alpha\hat{\mu} + \lambda_{\min}(G(\alpha)) + \alpha^2\lambda_{\min}(H(\triangle\hat{X}\triangle\hat{S})) \\ &\geq \alpha\hat{\mu} + \left(1 - \frac{5}{4}\alpha\right)\gamma\hat{\mu}_g - \alpha^2\|H(\triangle\hat{X}\triangle\hat{S})\|_{\mathrm{F}} \\ &\geq \left(\alpha\frac{\hat{\mu}}{\hat{\mu}_g} + \left(1 - \frac{5}{4}\alpha\right)\gamma - \frac{5}{4}\alpha^2 n\right)\hat{\mu}_g, \end{aligned}$$

where the last inequality follows from Corollary 3.8. Using the specific value of $\hat{\mu}$, we have

$$\begin{aligned} \lambda_{\min}(H(\hat{X}(\alpha)\hat{S}(\alpha))) &\geq \left(\frac{\alpha\gamma}{1-\gamma} + \left(1 - \frac{5}{4}\alpha\right)\gamma - \frac{5}{4}\alpha^2 n\right)\hat{\mu}_g \\ &= \left(\frac{\alpha\gamma^2}{1-\gamma} + (1-\alpha)\gamma + \frac{3}{4}\alpha\gamma - \frac{5}{4}\alpha^2 n\right)\hat{\mu}_g \\ &= \hat{\mu}_g(\alpha) + \left(\frac{3}{4}\alpha\gamma - \frac{5}{4}\alpha^2 n\right)\hat{\mu}_g. \end{aligned}$$

Thus, for all $0 \leq \alpha \leq 3\gamma/(5n)$, it holds that

$$\lambda_{\min}(H(\hat{X}(\alpha)\hat{S}(\alpha))) \geq \hat{\mu}_g(\alpha) > 0,$$

which reveals $\hat{X}(\alpha)\hat{S}(\alpha)$ is non-singular by Lemma A.1, and further implies that $\hat{X}(\alpha)$ and $\hat{S}(\alpha)$ are non-singular as well. By using continuity of the eigenvalues of a symmetric matrix (see [11] Theorem A.5), it follows that $\hat{X}(\alpha) \succ 0$ and $\hat{S}(\alpha) \succ 0$ for all $\alpha \in [0, 3\gamma/(5n)]$, since $\hat{X}, \hat{S} \succ 0$. Moreover, by Lemma A.2, we have

$$\lambda_{\min}(\hat{X}(\alpha)\hat{S}(\alpha)) \geq \lambda_{\min}(H(\hat{X}(\alpha)\hat{S}(\alpha))) \geq \hat{\mu}_g(\alpha) \quad \forall \alpha \in \left[0, \frac{3\gamma}{5n}\right].$$

Hence, we can conclude that $\alpha_c \geq 3\gamma/(5n)$. ∎

3.4 *Polynomial complexity*

In this section, we present our main complexity result.

THEOREM 3.10 *Algorithm* 1 *terminates in at most* $O(n \log(\mathrm{Tr}(X^0S^0)/\varepsilon))$ *iterations with a solution for which* $\mathrm{Tr}(XS) \leq \varepsilon$.

Proof If $\alpha_{\mathrm{a}} < 0.1$ or $\alpha_c < 3\gamma/(5n)$, then the algorithm uses the safeguard strategy, and by Theorem 3.9 and relation (24) one has

$$\hat{\mu}_g(\alpha) = \left(1 - \alpha\frac{1-2\gamma}{1-\gamma}\right)\hat{\mu}_g \leq \left(1 - \frac{3\gamma(1-2\gamma)}{5(1-\gamma)n}\right)\hat{\mu}_g.$$

If $\alpha_{\mathrm{a}} \geq 0.1$ and $\alpha_c \geq 3\gamma/(5n)$, then the algorithm uses Mehrotra's updating strategy, and thus one has

$$\hat{\mu}_g(\alpha) = (1 - \alpha(1 - (1-\alpha_{\mathrm{a}})^3))\hat{\mu}_g \leq \left(1 - \frac{3\gamma}{20n}\right)\hat{\mu}_g.$$

Therefore, by (19), there have a positive constant $\delta < 1$ such that

$$\mu_g(\alpha) \leq \left(1 - \frac{\delta}{n}\right)\mu_g,$$

which completes the proof conforming to Theorem 3.2 of [28]. ∎

4. Conclusions

In this paper, we have discussed the polynomiality of Mehrotra-type predictor–corrector algorithm for SDP. In fact, this algorithm is an extension of the recent variant of Mehrotra's predictor–corrector algorithm of Salahi *et al.* [23] for LP problems. Based on the NT directions [20,21], we show that the iteration complexity of the algorithm is $O(n \log(\mathrm{Tr}(X^0S^0)/\varepsilon))$, which is analogous to the case of LP.

Acknowledgements

We are very grateful to the editor and the anonymous referees for their constructive suggestions which helped to improve the paper. Authors 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. 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. El 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 $O(\sqrt{n}L)$ iteration-complexity bound for semidefinite programming*, Optimization 59 (2010), pp. 1235–1246.
[7] K. Fujisawa, M. Fukuda, M. Kojima, and K. Nakata, *Numerical evaluation of the SDPA (SemiDefinite Programming Algorithm)*, in *High Performance Optimization*, H. Frenk, K. Roos, T. Terlaky, and S. Zhang eds., Kluwer Academic Publishers, Netherlands, 2000, pp. 267–301.
[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, *Matrix Analysis*, Cambridge University Press, New York, 1990.
[10] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1994.
[11] E. de Klerk, *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
[12] M. Kojima, S. Shindoh, and S. Hara, *Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*, SIAM J. Optim. 7 (1997), pp. 86–125.
[13] 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.
[14] Y. Li and T. Terlaky, *A new class of large neighborhood path-following interior point algorithms for semidefinite optimization with $O(\sqrt{n}\log(\mathrm{Tr}(X^0S^0)/\varepsilon))$ iteration complexity*, SIAM J. Optim. 20 (2010), pp. 2853–2875.
[15] C. Liu, H. Liu, and W. Cong, *An $O(\sqrt{n}L)$ iteration primal–dual second-order corrector algorithm for linear programming*, Optim. Lett. 5 (2011), pp. 729–743.
[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 semidefenite programming*, SIAM J. Optim. 7 (1997), pp. 663–678.
[18] R.D.C. Monteiro and Y. Zhang, *A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinie programming*, Math. Program. 81 (1998), pp. 281–299.
[19] Y. Nesterov and A.S. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, PA, 1994.
[20] Y. Nesterov and M.J. Todd, *Self-scaled barriers and interior-point methods for convex program-ming*, Math. Oper. Res. 22 (1997), pp. 1–42.
[21] Y. Nesterov and M.J. Todd, *Primal–dual interior-point methods for self-scaled cones*, SIAM J. Optim. 8 (1998), pp. 324–364.
[22] J.M. Peng, C. Roos, and T. Terlaky, *Self-regular functions and new search directions for linear and semidefinite optimization*, Math. Program. Ser. A 93 (2002), pp. 129–171.
[23] M. Salahi, J. Peng, and T. Terlaky, *On Mehrotra-type predictor–corrector algorithms*, SIAM J. Optim. 18 (2007), pp. 1377–1397.
[24] J.F. Sturm, *Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones*, Optim. Methods Softw. 11/12 (1999), pp. 625–653.
[25] J.F. Sturm and S. Zhang, *Symmetric primal-dual path following algorithms for semidefinite programming*, Appl. Numer. Math. 29 (1999), pp. 301–315.
[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] K.C. Toh, M.J. Todd, and R.H. Tütüncü, *SDPT3 – A Matlab software package for semidefinite programming*, Optim. Methods Softw. 11 (1999), pp. 545–581.
[28] S.J. Wright, *Primal–Dual Interior-Point Methods*, SIAM, Philadelphia, PA, 1997.
[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.

Appendix 1

In this section, we prove two technical lemmas that have been used frequently during the analysis.

The following lemma follows from the fact that the real part of the spectrum of a real matrix is contained between the largest and the smallest eigenvalues of its Hermitian part (see [10, p. 187], Mirsky's theorem).

LEMMA A.1 *For any $M \in \mathbb{R}^{n\times n}$, the following relations hold*:

$$\max_{i=1,\dots,n} \mathrm{Re}[\lambda_i(M)] \le \lambda_{\max}(H(M)), \quad \min_{i=1,\dots,n} \mathrm{Re}[\lambda_i(M)] \ge \lambda_{\min}(H(M)).$$

LEMMA A.2 *Suppose that $X \in \mathcal{S}^n_{++}$ and $S \in \mathcal{S}^n$, then $\lambda_{\min}(XS) \ge \lambda_{\min}(H(XS))$.*

Proof By similarity, we have $\lambda_i(XS) = \lambda_i(X^{1/2}SX^{1/2}) \in \mathbb{R}$. Hence, the result follows from Lemma A.1. ∎

We show the following Weyl theorem (see [9, p. 181], Theorem 4.3.1).

LEMMA A.3 *Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric and let the eigenvalues $\lambda_i(A), \lambda_i(B)$ and $\lambda_i(A+B)$ be arranged in increasing order. For each $k = 1, 2, \dots, n$, we have*

$$\lambda_k(A) + \lambda_1(B) \le \lambda_k(A+B) \le \lambda_k(A) + \lambda_n(B).$$