SIAM J. OPTIM.
Vol. 20, No. 6, pp. 2853–2875

A NEW CLASS OF LARGE NEIGHBORHOOD PATH-FOLLOWING INTERIOR POINT ALGORITHMS FOR SEMIDEFINITE OPTIMIZATION WITH $O(\sqrt{n}\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$ ITERATION COMPLEXITY*

YANG LI† AND TAMÁS TERLAKY‡

Abstract. In this paper, we extend the Ai–Zhang direction to the class of semidefinite optimization problems. We define a new wide neighborhood $\mathcal{N}(\tau_1,\tau_2,\eta)$, and, as usual but with a small change, we make use of the scaled Newton equations for symmetric search directions. After defining the "positive part" and the "negative part" of a symmetric matrix, we recommend solving the Newton equation with its right-hand side replaced first by its positive part and then by its negative part, respectively. In such a way, we obtain a decomposition of the classical Newton direction and use different step lengths for each of them. Starting with a feasible point (X^0, y^0, S^0) in $\mathcal{N}(\tau_1,\tau_2,\eta)$, the algorithm terminates in at most $O(\eta\sqrt{\kappa_\infty n}\log(\mathrm{Tr}(X^0S^0)/\epsilon))$ iterations, where κ_∞ is a parameter associated with the scaling matrix P and ϵ is the required precision. To our best knowledge, when the parameter η is a constant, this is the first large neighborhood path-following interior point method (IPM) with the same complexity as small neighborhood path-following IPMs for semidefinite optimization that use the Nesterov–Todd direction. In the case where η is chosen to be in the order of $\sqrt{n}$, our result coincides with the results for classical large neighborhood IPMs. Some preliminary numerical results also confirm the efficiency of the algorithm.

Key words. interior point methods, large neighborhood, path-following algorithm, semidefinite optimization

AMS subject classifications. 90C22, 90C51

DOI. 10.1137/080729311

Notation. Throughout the paper, we use the following notations:

$\mathcal{R}^n$	The n-dimensional Euclidean space
$\mathcal{R}^{m\times n}$	The set of all $m\times n$ matrices
$\mathcal{S}^n$	The set of all $n\times n$ symmetric matrices
$\mathcal{S}^n_+$	The set of all $n\times n$ symmetric positive semidefinite matrices
$\mathcal{S}^n_{++}$	The set of all $n\times n$ symmetric positive definite matrices
$Q\succeq 0$	Q is positive semidefinite, where $Q\in\mathcal{S}^n$
$Q\succ 0$	Q is positive definite, where $Q\in\mathcal{S}^n$
Tr(Q)	The trace of a matrix $Q\in\mathcal{R}^{n\times n}$, i.e., $\mathrm{Tr}(Q) := \sum_{i=1}^n Q_{ii}$
$\lambda_i(Q)$	The eigenvalues of $Q\in\mathcal{S}^n$, $i=1,2,\dots,n$
$\lambda_{\min}(Q)$	The smallest eigenvalue of $Q\in\mathcal{S}^n$
$\lambda_{\max}(Q)$	The largest eigenvalue of $Q\in\mathcal{S}^n$
$\Lambda(Q)$	The diagonal matrix with all eigenvalues of Q as diagonal elements

*Received by the editors July 7, 2008; accepted for publication (in revised form) June 10, 2010; published electronically September 23, 2010. The authors acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chairs (CRC) program, Lehigh University, and the Advanced Optimization Laboratory (AdvOL), McMaster University.

http://www.siam.org/journals/siopt/20-6/72931.html

†Department of Mathematics and Statistics, McMaster University, Hamilton, ON, L8S 4K1, Canada (liy99@mcmaster.ca).

‡Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015-1582 (terlaky@lehigh.edu). This author's research was also partially supported by a start-up grant from Leigh University, by the Hungarian National Development Agency, and the European Union within the frame of project TAMOP 4.2.2-08/1-2008-0021 at the Széchenyi István University entitled "Simulation and Optimization—Basic Research in Numerical Mathematics."

$\text{cond}(Q)$	The condition number of Q, defined as $\text{cond}(Q) = \lambda_{\max}(Q)/\lambda_{\min}(Q)$
$\|Q\|$	The Euclidean norm for $Q \in \mathcal{R}^{n\times n}$, i.e., $\|Q\| = \max_{\|\mu\|=1} \|Qu\|$
$\|Q\|_F$	The Frobenius norm of $Q \in \mathcal{R}^{n\times n}$, i.e., $\|Q\|_F = \sqrt{\text{Tr}(Q^T Q)}$
$\mathbf{vec}(Q)$	The vector obtained by stacking Q's columns one by one

1. Introduction. Semidefinite optimization (SDO) problems yield a generalization of linear optimization (LO) problems. Since Alizadeh [3] explored various applications of SDO in combinatorial optimization, SDO has been applied in many areas, including control theory, probability, and signal processing [22].

Due to the success of interior point methods (IPMs) in solving LO, most IPM variants were extended to SDO. The first IPMs for SDO were independently developed by Alizadeh [3] and Nesterov and Nemirovskii [15]. Alizadeh [3] applied Ye's potential reduction idea to SDO and showed how variants of dual IPMs could be extended to SDO. Almost at the same time, in their milestone book [15], Nesterov and Nemirovskii proved that IPMs are able to solve general conic optimization problems, in particular SDO problems, to ϵ precision in polynomial time.

The difficulty to extend primal-dual path-following IPMs from LO to SDO lies in acquiring a symmetric search direction. The Newton method applied to the central path equation $XS = \tau\mu I$ leads to the linear system

$$X\Delta S + \Delta X S = \tau\mu I - XS, \tag{1.1}$$

which generally results in nonsymmetric search directions. Over the years, people suggested many strategies to deal with this problem. Alizadeh, Haeberly, and Overton (AHO) [4] proposed to symmetrize both sides of (1.1). Alternatively, a similarity transformation $P(\cdot)P^{-1}$ could be applied to both sides of (1.1). This strategy was first investigated by Monteiro [11] for $P = X^{-1/2}$ and $P = S^{1/2}$. It turned out that the resulting directions by this approach could be seen as two special cases of the class of directions introduced earlier by Kojima, Shindoh, and Hara [10]. At the same time, another motivation led Helmberg et al. [8] to the direction given by $P = S^{1/2}$. The search directions given by $P = X^{-1/2}$ and $P = S^{1/2}$ are usually referred to as the H..K..M directions. Nesterov and Todd [16, 17] introduced the so-called Nesterov–Todd (NT) direction in their attempt to generalize primal-dual IPMs beyond SDO. In [23], based on Monteiro's idea, Zhang generalized all these aforementioned approaches to a unified scheme parameterized by a nonsingular scaling matrix P. This family of search directions is referred to as the Monterio–Zhang (MZ) family of search directions.

As in the case of LO, there is an intriguing fact about IPMs for SDO. Although the theoretical complexity is worse, large neighborhood algorithms perform better in practice than small neighborhood algorithms. Many efforts were spent to bridge this gap. In [19], Peng, Roos, and Terlaky established a new paradigm based on the class of the so-called self-regular functions. Under their new paradigm, large neighborhood IPMs can come arbitrarily close to the best known iteration bounds of small neighborhood IPMs. Later, based on Ai's original paper [1], a result of interest was given by Ai and Zhang [2] for linear complementarity problems (LCPs). Their algorithm decomposes the classical Newton direction into two orthogonal ones and proceeds in a wide neighborhood. It is proved that their algorithm stops after at most $O(\sqrt{n}L)$ iterations, where n is the number of variables and L is the input data length. This result yields the first large neighborhood path-following algorithm having the same theoretical complexity as a small neighborhood path-following algorithm for monotone LCPs, which include LO as a special case.

In this paper, we extend the Ai–Zhang scheme to SDO. We first define a new neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$, where $0 < \tau_2 < \tau_1 < 1$ and $\eta \geq 1$ are given parameters. This new neighborhood is proved to be a wide neighborhood itself. Not surprisingly, the neighborhood defined by Ai and Zhang [2] is a simple case of our wide neighborhood for LO. Another important ingredient of our algorithm is the decomposition of the classical Newton direction into two separate ones. To make this point clear, let us consider only LO, where (1.1) becomes

$$x\Delta s + s\Delta x = \tau\mu e - xs. \tag{1.2}$$

Moving from (x, s) to $(x+\alpha\Delta x, s+\alpha\Delta s)$, at each iteration we expect the duality gap to decrease by

$$x^T s - (x+\alpha\Delta x)^T(s+\alpha\Delta s) = -\alpha(x^T\Delta s + s^T\Delta x) = -\alpha e^T(\tau\mu e - xs) = \alpha(1-\tau)x^T s > 0,$$

where the relation $\mu = \frac{x^T s}{n}$ is used. Further, we may also write

$$\begin{aligned} -\alpha(x^T\Delta s + s^T\Delta x) &= -\alpha \sum \min([\tau\mu e - xs]_i, 0) - \alpha \sum \max([\tau\mu e - xs]_i, 0) \\ &= -\alpha \sum [\tau\mu e - xs]_i^- - \alpha \sum [\tau\mu e - xs]_i^+, \end{aligned}$$

from which we see that the negative components of $\tau\mu e - xs$ are responsible for reducing the duality gap. On the other hand, the small components of $x_i s_i$, i.e., $x_i s_i < \tau\frac{x^T s}{n}$, indicate that the iterate is "close" to the boundary of the positive orthant. For these coordinates, using (1.2) we have

$$\begin{aligned} x_i(\alpha)s_i(\alpha) &= (x_i + \alpha\Delta x_i)(s_i + \alpha\Delta s_i) \\ &= x_i s_i + \alpha(x_i\Delta s_i + s_i\Delta x_i) + \alpha^2\Delta x_i\Delta s_i \\ &= x_i s_i + \alpha\left(\tau\frac{x^T s}{n} - x_i s_i\right) + \alpha^2\Delta x_i\Delta s_i. \end{aligned}$$

Because the coefficient of the first order term of α, $\tau\frac{x^T s}{n} - x_i s_i$, is bigger than zero, then $x_i(\alpha)s_i(\alpha)$ increases locally and pushes the iterate to the interior of the first orthant, i.e., keeping the centrality. Ai and Zhang [2] suggested to treat negative and positive components of $\tau\mu e - xs$ separately to obtain a better iteration complexity bound for large neighborhood IPMs. We generalize this idea to SDO and show that, given a feasible starting point (X^0, y^0, S^0) in $\mathcal{N}(\tau_1, \tau_2, \eta)$, our algorithm terminates in at most $O(\eta\sqrt{\kappa_\infty n}\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$ iterations for SDO, where n is the dimension of the problem, κ_∞ is a parameter associated with the scaling matrix P, and ϵ is the required precision. In particular, when the parameter η is a fixed constant, our large neighborhood path following algorithm has the same theoretical complexity as a small neighborhood algorithm that uses NT scaling. Likewise, when η is chosen to be in the order of $\sqrt{n}$, this complexity coincides with the known results for the classical large neighborhood algorithms.

We organize our paper as follows. In section 2, we introduce the primal-dual pair of SDO problems and briefly explain how path-following IPMs work. In section 3, we define the positive and negative parts of a symmetric matrix and prove some of their intriguing properties. By using these new definitions, we introduce a new neighborhood which is proved to be a large neighborhood. In section 4, we explain the way to decompose the classical Newton direction and present the framework of our algorithm. The convergence analysis and theoretical complexity bounds are presented in section 5. Finally, we finish the paper with some conclusions and considerations about future work.

2. SDO problem. We consider the following SDO problem

$$(\mathcal{P}) \quad \begin{array}{ll} \min & \operatorname{Tr}(CX) \\ \text{subject to (s.t.)} & \operatorname{Tr}(A_iX) = b_i,\ i = 1,\dots,m, \\ & X \succeq 0, \end{array}$$

where $C,\ X \in \mathcal{S}^n$, and $A_i \in \mathcal{S}^n$, $i = 1,\dots,m$, are linearly independent and $b = (b_1,\dots,b_m)^T \in \mathcal{R}^m$. We call problem $(\mathcal{P})$ the primal form of SDO, and X is the primal matrix variable.

Corresponding to every primal problem $(\mathcal{P})$, there exists a dual problem $(\mathcal{D})$:

$$(\mathcal{D}) \quad \begin{array}{ll} \max & b^Ty \\ \text{s.t.} & \sum_{i=1}^m y_iA_i + S = C, \\ & S \succeq 0, \end{array}$$

where $y \in \mathcal{R}^m$, $S \in \mathcal{S}^n$, and (y, S) is the dual variable.

The primal-dual feasible set is defined as

$$\mathcal{F} = \left\{ (X, y, S) \in \mathcal{S}^n_+ \times \mathcal{R}^m \times \mathcal{S}^n_+ \,\middle|\, \begin{array}{c} \operatorname{Tr}(A_iX) = b_i,\ i = 1,\dots,m \\ \sum_{i=1}^m y_iA_i + S = C \end{array} \right\},$$

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

$$\mathcal{F}^0 = \left\{ (X, y, S) \in \mathcal{S}^n_{++} \times \mathcal{R}^m \times \mathcal{S}^n_{++} \,\middle|\, \begin{array}{c} \operatorname{Tr}(A_iX) = b_i,\ i = 1,\dots,m \\ \sum_{i=1}^m y_iA_i + S = C \end{array} \right\}.$$

Under the assumptions that $\mathcal{F}^0$ is nonempty and the matrices A_i, $i = 1, 2,\dots,m$, are linearly independent, then X^* and (y^*, S^*) are optimal if and only if they satisfy the optimality conditions [6],

$$\begin{aligned} \operatorname{Tr}(A_iX) = b_i,\ X \succeq 0,\ i = 1,\dots,m, \\ \sum_{i=1}^m y_iA_i + S = C,\ S \succeq 0, \\ XS = 0. \end{aligned} \tag{2.1}$$

Path-following IPMs follow the central path that is given as the set of solutions of the perturbed optimality conditions

$$\begin{aligned} \operatorname{Tr}(A_iX) = b_i,\ X \succ 0,\ i = 1,\dots,m, \\ \sum_{i=1}^m y_iA_i + S = C,\ S \succ 0, \\ XS = \mu I, \end{aligned} \tag{2.2}$$

rather than (2.1). It is proved [10, 15] that there is a unique solution $(X(\mu), y(\mu), S(\mu))$ to the central path equations (2.2) for any barrier parameter $\mu > 0$, assuming that $\mathcal{F}^0$ is nonempty and the coefficient matrices A_i, $i = 1,\dots,m$, are linearly independent. Moreover, the limit point (X^*, y^*, S^*), as μ goes to 0, is a primal-dual optimal solution of the corresponding SDO problem.

3. Neighborhood. Although path-following interior point algorithms follow the central path while the barrier parameter μ is decreasing to 0, they do not stay on the central path exactly. All the iterates are merely required to reside in a neighborhood of the central path, while steadily approaching the optimal set.

One of the popular neighborhoods is the so-called *small neighborhood*, defined as

$$\mathcal{N}_F(\theta) := \left\{ (X, y, S) \in \mathcal{F}^0 \,\middle|\, \left\| X^{1/2} S X^{1/2} - \mu_g I \right\|_F = \left[\sum_{i=1}^{n} (\lambda_i(XS) - \mu_g)^2 \right]^{1/2} \leq \theta \mu_g \right\},$$

where $\theta \in (0,1)$ and $\mu_g := \mathrm{Tr}(XS)/n$ is associated with the actual duality gap. Another popular alternative is called the *negative infinity neighborhood* that is a *large neighborhood*, defined as

$$\mathcal{N}_\infty^-(1-\gamma) := \left\{ (X, y, S) \in \mathcal{F}^0 \mid \lambda_{\min}(XS) \geq \gamma \mu_g \right\},$$

where $\gamma \in (0,1)$.

Theoretically, IPMs based on the small neighborhood $\mathcal{N}_F(\theta)$, e.g., short step algorithms, have a better iteration complexity bound than algorithms based on the large neighborhood, e.g., large update algorithms. However, computational experience reveals that large update IPMs usually perform better in practice than short step algorithms.

In this paper we explore a variant of large neighborhood path-following IPMs and prove that its iteration complexity is $O(\eta\sqrt{\kappa_\infty n} \log \frac{\mathrm{Tr}(X^0 S^0)}{\epsilon})$, where n is the dimension of the problem, κ_∞ is a parameter associated with the scaling matrix P, and ϵ is the required precision. The new parameter η is used to define our new neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$. In particular, when η is chosen to be a constant, our new algorithm has the best complexity result $O(\sqrt{n} \log \frac{\mathrm{Tr}(X^0 S^0)}{\epsilon})$, which coincides with the complexity of short step IPMs when NT scaling is used.

In order to introduce our new algorithm, we need to investigate a new neighborhood which combines the classical small and large neighborhoods. Before doing so, we need to present some notations.

Let M be a symmetric real matrix, i.e., $M \in \mathcal{S}^n$, with the *spectral decomposition* $M = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i q_i q_i^T$, where Λ is a diagonal matrix with all the eigenvalues of M along its diagonal, Q is an orthonormal matrix, i.e., $QQ^T = I$, and each column q_i of Q is an eigenvector of M corresponding to the eigenvalue λ_i. Then we define the positive part M^+ and the negative part M^- of M as

$$M^+ := \sum_{\lambda_i \geq 0} \lambda_i q_i q_i^T \ , \quad M^- := \sum_{\lambda_i \leq 0} \lambda_i q_i q_i^T. \tag{3.1}$$

In particular, for a real number $M \in \mathcal{S}^1$, $M^+ = \max\{M, 0\}$ and $M^- = \min\{M, 0\}$. Likewise, if $M \in \mathcal{S}^n$ is a diagonal matrix, M^+ and M^- could be constructed by simply separating the nonnegative and nonpositive entries. Apparently, $M = M^+ + M^-$, where $M^+, -M^- \succeq 0$.

It turns out that the positive and negative parts of a symmetric matrix have many interesting properties. We present and verify some of them which play a crucial role in the complexity analysis.

First, we show that the triangle inequality holds for the positive part.

PROPOSITION 3.1. *Assume $U, V \in \mathcal{S}^n$. Then we have*

$$\left\| (U+V)^+ \right\|_F \leq \left\| U^+ + V^+ \right\|_F \leq \left\| U^+ \right\|_F + \left\| V^+ \right\|_F.$$

Proof. The second inequality is straightforward. We show only the first one. Note that

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

and

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

According to Theorem 8.1.5 in Golub and Van Loan [7], we obtain

$$\lambda_i(U+V) \le \lambda_i(U^+ + V^+)$$

for $i = 1, \ldots, n$. Let $\mathcal{I}$ denote the index set such that $\mathcal{I} := \{\, i \mid \lambda_i(U+V) \ge 0\}$. Then

$$\left\|(U+V)^+\right\|_F = \left[\sum_{i\in\mathcal{I}} \lambda_i^2(U+V)\right]^{1/2} \le \left[\sum_{i\in\mathcal{I}} \lambda_i^2(U^+ + V^+)\right]^{1/2} \le \left\|U^+ + V^+\right\|_F,$$

which completes the proof. ∎

The next lemma reveals that a similarity transformation preserves the Frobenius norm over the positive part of a symmetric matrix.

LEMMA 3.2. *Let $M \in \mathcal{S}^n$ and W be a nonsingular matrix. Then we have*

$$\left\|M^+\right\|_F = \left\|(WMW^{-1})^+\right\|_F.$$

Proof. The result is readily available from the similarity of M and WMW^{-1}. ∎

The next lemma exhibits that the positive part of a symmetric matrix does not exceed, in the sense of Frobenius norm, its positive part after a similarity transformation.

LEMMA 3.3. *Suppose that $W \in \mathcal{R}^{n\times n}$ is a nonsingular matrix. Then, for any $M \in \mathcal{S}^n$, we have*

$$\left\|M^+\right\|_F \le \frac{1}{2}\left\|\left[WMW^{-1} + (WMW^{-1})^T\right]^+\right\|_F. \tag{3.2}$$

To prove this result, we need to first demonstrate an interesting fact about symmetric matrices.

LEMMA 3.4. *Let $M \in \mathcal{S}^n$, and let λ_i and m_{ii} denote the ith eigenvalue and the ith diagonal element of M, respectively. Then we have*

$$\sum_{\lambda_i \ge 0} \lambda_i^2 \ge \sum_{m_{ii}\ge 0} m_{ii}^2.$$

Proof. Recall the definitions of M^+ and M^- as in (3.1), and let m_{ij}^+ and m_{ij}^- denote the (i,j) element for M^+ and M^-, respectively. Note the fact that for any i, $m_{ii}^+ \ge 0$ and $m_{ii}^- \le 0$; then we can define the set $\mathcal{I}$ as

$$\mathcal{I} = \{i \mid m_{ii} = m_{ii}^+ + m_{ii}^- \ge 0\}.$$

For any $i \in \mathcal{I}$, we have $m_{ii}^+ \ge m_{ii}^+ + m_{ii}^- \ge 0$, since $m_{ii}^- \le 0$. Further, we obtain $(m_{ii}^+)^2 \ge (m_{ii}^+ + m_{ii}^-)^2$ for all $i \in \mathcal{I}$.

The proof of the lemma follows by

$$\sum_{\lambda_i \geq 0} \lambda_i^2 = \left\| M^+ \right\|_F^2 \geq \sum_{i=1}^{n} (m_{ii}^+)^2 \geq \sum_{i \in \mathcal{I}} (m_{ii}^+)^2 \geq \sum_{i \in \mathcal{I}} (m_{ii}^+ + m_{ii}^-)^2 = \sum_{m_{ii} \geq 0} m_{ii}^2. \qquad \square$$

Now, we are ready to prove Lemma 3.3.

Proof of Lemma 3.3. We see that $\|M^+\|_F^2 = \|[\Lambda(M)]^+\|_F^2 = \sum_{\lambda_i(M) \geq 0} \lambda_i^2(M)$. Let us consider the right-hand side of (3.2). According to Schur's triangularization theorem, there exists a unitary matrix U such that $U(WMW^{-1})U^* = \Lambda(WMW^{-1}) + N = \Lambda(M) + N$, where U^* is the Hermitian adjoint of U and N is a strictly upper triangular matrix. From Lemma 3.2, we know that

$$\begin{aligned} \frac{1}{2} \left\| \left[WMW^{-1} + (WMW^{-1})^T \right]^+ \right\|_F &= \frac{1}{2} \left\| \left[U(WMW^{-1} + (WMW^{-1})^T)U^* \right]^+ \right\|_F \\ &= \frac{1}{2} \left\| \left[\Lambda(M) + N + \Lambda(M) + N^T \right]^+ \right\|_F \\ &= \left\| \left[\Lambda(M) + \frac{N + N^T}{2} \right]^+ \right\|_F . \end{aligned}$$

On the other hand, from Lemma 3.4 we conclude that

$$\left\| [\Lambda(M)]^+ \right\|_F^2 \leq \left\| \left[\Lambda(M) + \frac{N + N^T}{2} \right]^+ \right\|_F^2 ,$$

which implies $\|M^+\|_F \leq \frac{1}{2} \left\| [WMW^{-1} + (WMW^{-1})^T]^+ \right\|_F$. $\square$

Until now, we have adequate results on the positive and negative parts of symmetric matrices so far. This will assist us in the analysis of convergence and complexity. Next, we move to the definition of our new large neighborhood. Analogous to the neighborhood introduced by Ai and Zhang [2], we define our neighborhood, using the positive part in (3.1), as

(3.3)
$$\mathcal{N}(\tau_1, \tau_2, \eta) := \mathcal{N}_\infty^-(1 - \tau_2) \cap \left\{ (X, y, S) \in \mathcal{F}^0 : \left\| [\tau_1 \mu_g I - X^{1/2} S X^{1/2}]^+ \right\|_F \leq \eta(\tau_1 - \tau_2)\mu_g \right\},$$

where $\eta \geq 1$ and $0 < \tau_2 < \tau_1 < 1$.

The next proposition indicates that the neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$ is indeed a large neighborhood.

PROPOSITION 3.5. *If $\eta \geq 1$ and $0 < \tau_2 < \tau_1 < 1$, then we have*

$$\mathcal{N}_\infty^-(1 - \tau_1) \subseteq \mathcal{N}(\tau_1, \tau_2, \eta) \subseteq \mathcal{N}_\infty^-(1 - \tau_2).$$

Proof. From the definition of $\mathcal{N}(\tau_1, \tau_2, \eta)$, it is obvious that $\mathcal{N}(\tau_1, \tau_2, \eta) \subseteq \mathcal{N}_\infty^-(1 - \tau_2)$. For the first inclusion, we need to prove that

$$\mathcal{N}_\infty^-(1 - \tau_1) \subseteq \left\{ (X, y, S) \in \mathcal{F}^0 : \left\| [\tau_1 \mu_g I - X^{1/2} S X^{1/2}]^+ \right\|_F \leq \eta(\tau_1 - \tau_2)\mu_g \right\}.$$

Given that for $(X, y, S) \in \mathcal{N}_\infty^-(1 - \tau_1)$, one has

$$\tau_1 \mu_g I - X^{1/2} S X^{1/2} \preceq 0, \tag{3.4}$$

which implies $[\tau_1 \mu_g I - X^{1/2} S X^{1/2}]^+ = 0$, leading to the inclusion relation. $\square$

Moreover, if the parameter $\eta \geq \sqrt{n}$, then the neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$ is exactly the negative infinity neighborhood $\mathcal{N}_\infty^-(1 - \tau_2)$.

PROPOSITION 3.6. *If $\eta \geq \sqrt{n}$ and $0 < \tau_2 < \tau_1 < 1$, then we have*

$$\mathcal{N}(\tau_1, \tau_2, \eta) = \mathcal{N}_\infty^-(1-\tau_2).$$

Proof. To complete the proof, it is sufficient to show that for any $(X, y, S) \in \mathcal{N}_\infty^-(1-\tau_2)$, we have

$$\mathcal{N}_\infty^-(1-\tau_2) \subseteq \left\{(X,y,S) \in \mathcal{F}^0 : \left\|[\tau_1\mu_g I - X^{1/2}SX^{1/2}]^+\right\|_F \leq \eta(\tau_1-\tau_2)\mu_g\right\}. \tag{3.5}$$

Because $(X, y, S) \in \mathcal{N}_\infty^-(1-\tau_2)$, it follows that

$$\lambda_{\min}(X^{1/2}SX^{1/2}) = \lambda_{\min}(XS) \geq \tau_2\mu_g.$$

Therefore,

$$\lambda_{\max}([\tau_1\mu_g I - X^{1/2}SX^{1/2}]^+) \leq (\tau_1-\tau_2)\mu_g.$$

This implies

$$\left\|[\tau_1\mu_g I - X^{1/2}SX^{1/2}]^+\right\|_F \leq \sqrt{n}(\tau_1-\tau_2)\mu_g,$$

which shows that (3.5) holds when $\eta \geq \sqrt{n}$. ∎

4. Search direction. Given an iterate (X, y, S), path-following IPMs generate the next iterate by taking a Newton step to system (2.2). Let the target be a point on the central path corresponding to $\mu = \tau\mu_g$, where $\tau \in [0, 1]$ is called centering parameter and $\mu_g = \mathrm{Tr}(XS)/n$ corresponds to the actual duality gap. To move from the current point (X, y, S) toward the target on the central path, we wish we could obtain a symmetric search direction from the following linear system:

$$\begin{aligned} \mathrm{Tr}(A_i\Delta X) &= 0,\\ \sum_{i=1}^m \Delta y_i A_i + \Delta S &= 0,\\ \Delta XS + X\Delta S &= \tau\mu_g I - XS. \end{aligned} \tag{4.1}$$

Although the second equality guarantees a symmetric ΔS, system (4.1) does not allow a symmetric solution of ΔX. Various remedies have been proposed since the middle of 1990s. The readers who are interested are referred to the paper [20] for a comprehensive discussion. We use the approach proposed by Zhang [23], who suggested to replace the last equation in system (2.2) by

$$H_P(XS) = \mu I, \tag{4.2}$$

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

$$H_P(M) = \frac{1}{2}\left[PMP^{-1} + (PMP^{-1})^T\right]$$

for a given matrix M and a given nonsingular matrix P. In particular, if $P = I$, then for any symmetric matrix M, $H_I(M) = M$. In [23], Zhang observed that if P is nonsingular, then

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

Thus, the search direction is well defined by the following system:

$$\text{Tr}(A_i \Delta X) = 0, \tag{4.3a}$$

$$\sum_{i=1}^{m} \Delta y_i A_i + \Delta S = 0, \tag{4.3b}$$

$$H_P(\Delta X S + X \Delta S) = \tau \mu_g I - H_P(XS). \tag{4.3c}$$

We refer to the directions derived from (4.3) as the Monteiro–Zhang (MZ) family. In particular, when $P = I$, the direction obtained from (4.3) coincides with the AHO direction [4]. If $P = X^{-1/2}$ or $S^{1/2}$, then (4.3) gives the H..K..M directions [8, 10, 11, 12]. Further, we obtain the NT direction when $P = W_{NT}^{-1/2}$, where W_{NT} is the solution of the system $W_{NT}^{-1} X W_{NT}^{-1} = S$. Nesterov and Todd [16, 17] proved the existence and uniqueness of such a solution as $W_{NT} = S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}$.

In terms of Kronecker product[1], (4.3c) can be expressed as

$$E\mathbf{vec}(\Delta X) + F\mathbf{vec}(\Delta S) = \mathbf{vec}(\tau \mu_g I - H_P(XS)),$$

where

$$E = \frac{1}{2}(P^{-T} S \otimes P + P \otimes P^{-T} S), \quad F = \frac{1}{2}(PX \otimes P^{-T} + P^{-T} \otimes PX). \tag{4.4}$$

Todd, Toh, and Tütüncü [21] proved that system (4.3) has a unique solution for any $(X, y, S) \in \mathcal{S}_{++}^n \times \mathcal{R}^m \times \mathcal{S}_{++}^n$ and for the scaling matrix P satisfying $PXSP^{-1} \in \mathcal{S}^n$. Actually, this still holds under some weaker conditions, as the authors pointed out in [13]. Throughout this paper, we restrict the scaling matrix P to a specific class

$$\mathcal{P}(X, S) := \{P \in \mathcal{S}_{++}^n \mid PXSP^{-1} \in \mathcal{S}^n\}, \tag{4.5}$$

where $X, S \in \mathcal{S}_{++}^n$. Apparently, $P = X^{-1/2}$, $S^{1/2}$, and $W_{NT}^{1/2}$ belong to this specific class. Unfortunately, $P = I$ does not. We should mention that this restriction on P is common for large neighborhood path-following algorithms proposed in [14]. Moreover, this restriction on P does not lose any generality, in terms of the solution set of system (4.3), as Monteiro indicates in [12].

After obtaining the search direction, most classic large neighborhood path-following algorithms do a linear search to decide how far they move along the search direction in attempt to minimize the duality gap as much as possible within the neighborhood $\mathcal{N}_\infty^-(1 - \tau_2)$. The algorithms repeat such a process until the optimal solution is identified.

In our new algorithm, we decompose the Newton direction into two separate parts according to the positive and negative parts of $\tau \mu_g I - H_P(XS)$. Thus, we need to solve the following two systems:

$$\text{Tr}(A \Delta X_-) = 0, \tag{4.6a}$$

$$\sum_{i=1}^{m} (\Delta y_i)_- A_i + \Delta S_- = 0, \tag{4.6b}$$

$$H_P(\Delta X_- S + X \Delta S_-) = [\tau \mu_g I - H_P(XS)]^-, \tag{4.6c}$$

[1]For the definition and properties of Kronecker product, please refer to Horn and Johnson [9].

and

$$\text{Tr}(A\Delta X_+) = 0, \tag{4.7a}$$

$$\sum_{i=1}^{m}(\Delta y_i)_+ A_i + \Delta S_+ = 0, \tag{4.7b}$$

$$H_P(\Delta X_+ S + X\Delta S_+) = [\tau\mu_g I - H_P(XS)]^+, \tag{4.7c}$$

where $P \in \mathcal{P}(X,S)$ and $(\Delta X_-, \Delta y_-, \Delta S_-)$ denotes the negative part of the search direction, while $(\Delta X_+, \Delta y_+, \Delta S_+)$ analogously denotes the positive part of the search direction. Again, (4.6c) and (4.7c) could be written in Kronecker product form as

$$E\mathbf{vec}(\Delta X_-) + F\mathbf{vec}(\Delta S_-) = \mathbf{vec}([\tau\mu_g I - H_P(XS)]^-) \tag{4.8}$$

and

$$E\mathbf{vec}(\Delta X_+) + F\mathbf{vec}(\Delta S_+) = \mathbf{vec}([\tau\mu_g I - H_P(XS)]^+), \tag{4.9}$$

respectively.

It is easy to see that systems (4.6) and (4.7) are also well defined and have a unique solution because $P \in \mathcal{P}(X,S)$. To get the best step lengths for both of the directions, we expect to solve the following subproblem:

$$\begin{aligned} \min \quad & \text{Tr}(X(\alpha)S(\alpha)) \\ \text{s.t.} \quad & (X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}(\tau_1, \tau_2, \eta), \\ & 0 \le \alpha_- \le 1,\ 0 \le \alpha_+ \le 1, \end{aligned} \tag{4.10}$$

where $\alpha = (\alpha_-, \alpha_+)$ denotes the step lengths along the direction $(\Delta X_-, \Delta y_-, \Delta S_-)$ and $(\Delta X_+, \Delta y_+, \Delta S_+)$, respectively. Consequently, the new iterate is given by

$$\begin{aligned} (X(\alpha), y(\alpha), S(\alpha)) &:= (X, y, S) + (\Delta X(\alpha), \Delta y(\alpha), \Delta S(\alpha)) \\ &:= (X, y, S) + \alpha_-(\Delta X_-, \Delta y_-, \Delta S_-) + \alpha_+(\Delta X_+, \Delta y_+, \Delta S_+). \end{aligned} \tag{4.11}$$

We have already introduced the most important ingredients of the new algorithm: a newly defined neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$ given by (3.3) and new search directions based on systems (4.6) and (4.7). We present a generic framework as follows.

There are three comments we would like to address about the presented algorithm. First of all, although we suggest solving problem (4.10) to decide the best step lengths, solving this problem could be expensive. Hence, a "sufficient" duality gap decrease obtained for low computational cost is preferred against the "maximal possible" duality gap decrease for high computational cost. Moreover, it is also not necessary to solve problem (4.10). Even if we do not use the optimal solution of problem (4.10) as the step lengths, we are still able to achieve polynomial convergence as it is discussed in section 5. Second, in spite of the fact that two linear systems (4.6) and (4.7) have to be solved, however, the additional cost is very marginal, since both (4.6) and (4.7) have the same coefficient matrix. At each iteration, the algorithm needs only to form and factorize the Schur matrix once, which together yield the majority of the total running time. Then backsolving needs to be executed once, simultaneously for two right-hand sides specified in (4.6) and (4.7). Third, it might

ALGORITHM 1 PATH-FOLLOWING IPM BASED ON THE $\mathcal{N}(\tau_1, \tau_2, \eta)$ NEIGHBORHOOD.

Input:
required precision $\epsilon > 0$;
neighborhood parameters $\eta \geq 1$, $0 < \tau_2 < \tau_1 < 1$;
reference parameter $0 \leq \tau \leq 1$;
an initial point $(X^0, y^0, S^0) \in \mathcal{N}(\tau_1, \tau_2, \eta)$ with $\mu_g^0 = \mathrm{Tr}(X^0 S^0)/n$;
while $\mu_g^k > \epsilon$ **do**
(1) Compute the scaling matrix $P^k \in \mathcal{P}(X^k, S^k)$.
(2) Compute the directions $(\Delta X_-^k, \Delta y_-^k, \Delta S_-^k)$ by (4.6) and $(\Delta X_+^k, \Delta y_+^k, \Delta S_+^k)$ by (4.7).
(3) Find a step length vector $\alpha^k = (\alpha_-^k, \alpha_+^k) > 0$ giving a sufficient reduction of the duality gap and assuring $(X(\alpha^k), y(\alpha^k), S(\alpha^k)) \in \mathcal{N}(\tau_1, \tau_2, \eta)$.
(4) Set $(X^{k+1}, y^{k+1}, S^{k+1}) = (X(\alpha^k), y(\alpha^k), S(\alpha^k))$.
(5) Set $\mu_g^{k+1} := \mathrm{Tr}(X^{k+1} S^{k+1})/n$ and $k := k+1$.
end while

appear computationally expensive to obtain the negative and positive parts in (4.6) and (4.7). However, if NT scaling is used, the strategy proposed by Todd, Toh, and Tütüncü [21] for computing the NT scaling matrix and the NT direction can help us. The basic idea is to simultaneously scale X and S to diagonal matrices. In this case, after scaling, the right-hand sides of (4.6c) and (4.7c) also become diagonal. Then it is straightforward to decide the negative and positive parts. Our preliminary numerical tests, which are available on request, also confirm that (4.6) and (4.7) can be simultaneously solved without a significant increase of computational time.

5. Complexity analysis. In this part, we present the convergence and complexity proofs for Algorithm 1. Recall that our algorithm is based on the MZ family; we scale problems $(\mathcal{P})$ and $(\mathcal{D})$ as Monteiro and Todd proposed in [13] in order to analyze the algorithm in a unified way for the class of matrices $P \in \mathcal{P}(X, S)$. In addition, this scaling procedure simplifies the proofs of the main results. After demonstrating several technical lemmas, we present the most important result of polynomial convergence at the end of this section.

5.1. Scaling procedure. We scale the primal and dual variables of problems $(\mathcal{P})$ and $(\mathcal{D})$ in the form of

$$\widetilde{X} := PXP, (\widetilde{y}, \widetilde{S}) := (y, P^{-1}SP^{-1}). \tag{5.1}$$

To keep consistency, we also have to apply the same scaling to the other data in $(\mathcal{P})$ and $(\mathcal{D})$ as well; i.e.,

$$\widetilde{C} := P^{-1}CP^{-1}, \quad (\widetilde{A_i}, \widetilde{b_i}) := (P^{-1}A_iP^{-1}, b_i) \text{ for } i = 1, \ldots, m.$$

As mentioned, we restrict the scaling matrix $P \in \mathcal{P}(X, S)$ as defined by (4.5). It is easy to see that for $X, S \in \mathcal{S}_{++}^n$ one has

$$\mathcal{P}(X, S) := \{P \in \mathcal{S}_{++}^n \mid PXSP^{-1} \in \mathcal{S}^n\} = \{P \in \mathcal{S}_{++}^n : \widetilde{X}\widetilde{S} = \widetilde{S}\widetilde{X}\}; \tag{5.2}$$

i.e., $\widetilde{X}$ and $\widetilde{S}$ become commutable after scaling under P. The commutativity of $\widetilde{X}\widetilde{S}$ also implies that $\widetilde{X}\widetilde{S}$ is symmetric if $\widetilde{X}$ and $\widetilde{S}$ are both symmetric. Further, the requirement on P also guarantees that $\widetilde{X}$ and $\widetilde{S}$ can be simultaneously diagonalized (i.e., they have spectral decompositions with the same Q) according to Proposition 4.2 in Monteiro and Zhang [14].

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 $i = 1, \ldots, n$ are the eigenvalues of $\widetilde{X}\widetilde{S}$ with increasing order $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. We should emphasize that the matrices $\widetilde{X}\widetilde{S}$, $\widetilde{S}\widetilde{X}$, 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.

In the scaled space problems $(\mathcal{P})$ and $(\mathcal{D})$ are equivalent to the following pair of problems:

$$(\widetilde{\mathcal{P}}) \quad \begin{array}{ll} \min & \mathrm{Tr}(\widetilde{C}\widetilde{X}) \\ \text{s.t.} & \mathrm{Tr}(\widetilde{A_i}\widetilde{X}) = \widetilde{b}_i, \ i = 1, \ldots, m, \\ & \widetilde{X} \succeq 0, \end{array}$$

and

$$(\widetilde{\mathcal{D}}) \quad \begin{array}{ll} \max & \widetilde{b}^T\widetilde{y} \\ \text{s.t.} & \displaystyle\sum_{i=1}^{m} \widetilde{y}_i\widetilde{A_i} + \widetilde{S} = \widetilde{C}, \\ & \widetilde{S} \succeq 0. \end{array}$$

The search direction $(\Delta X, \Delta y, \Delta S)$ based on system (4.6) and (4.7) corresponds to the scaled direction $(\widetilde{\Delta X}, \widetilde{\Delta y}, \widetilde{\Delta S})$ defined as

$$\widetilde{\Delta X}_- = P\Delta X_- P, \quad \widetilde{\Delta y}_- = \Delta y_-, \quad \widetilde{\Delta S}_- = P^{-1}\Delta S_- P^{-1}, \tag{5.3}$$

$$\widetilde{\Delta X}_+ = P\Delta X_+ P, \quad \widetilde{\Delta y}_+ = \Delta y_+, \quad \widetilde{\Delta S}_+ = P^{-1}\Delta S_+ P^{-1}. \tag{5.4}$$

The directions $(\widetilde{\Delta X}_-, \widetilde{\Delta y}_-, \widetilde{\Delta S}_-)$ and $(\widetilde{\Delta X}_+, \widetilde{\Delta y}_+, \widetilde{\Delta S}_+)$ are readily verified to be solutions of the scaled Newton systems

$$\mathrm{Tr}(\widetilde{A_i}\widetilde{\Delta X}_-) = 0, \tag{5.5a}$$

$$\sum_{i=1}^{m} (\widetilde{\Delta y_i})_-\widetilde{A_i} + \widetilde{\Delta S}_- = 0, \tag{5.5b}$$

$$H_I(\widetilde{\Delta X}_-\widetilde{S} + \widetilde{X}\widetilde{\Delta S}_-) = [\tau\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-, \tag{5.5c}$$

and

$$\mathrm{Tr}(\widetilde{A_i}\widetilde{\Delta X}_+) = 0, \tag{5.6a}$$

$$\sum_{i=1}^{m} (\widetilde{\Delta y_i})_+\widetilde{A_i} + \widetilde{\Delta S}_+ = 0, \tag{5.6b}$$

$$H_I(\widetilde{\Delta X}_+\widetilde{S} + \widetilde{X}\widetilde{\Delta S}_+) = [\tau\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+, \tag{5.6c}$$

respectively. To simplify the notation, we use $\widetilde{X}\widetilde{S}$ rather than $H_I(\widetilde{X}\widetilde{S})$, since $\widetilde{X}\widetilde{S} = H_I(\widetilde{X}\widetilde{S})$ when the scaling matrix $P \in \mathcal{P}(X, S)$. In terms of the Kronecker product, (5.5c) and (5.6c) become

$$\widetilde{E}\mathbf{vec}(\widetilde{\Delta X}_-) + \widetilde{F}\mathbf{vec}(\widetilde{\Delta S}_-) = \mathbf{vec}([\tau\mu_g I - \widetilde{X}\widetilde{S}]^-), \tag{5.7a}$$

$$\widetilde{E}\mathbf{vec}(\widetilde{\Delta X}_+) + \widetilde{F}\mathbf{vec}(\widetilde{\Delta S}_+) = \mathbf{vec}([\tau\mu_g I - \widetilde{X}\widetilde{S}]^+), \tag{5.7b}$$

respectively, where

$$\widetilde{E} = \frac{1}{2}(\widetilde{S} \otimes I + I \otimes \widetilde{S}), \quad \widetilde{F} = \frac{1}{2}(\widetilde{X} \otimes I + I \otimes \widetilde{X}). \tag{5.8}$$

After deciding the step lengths, the iterates are updated as follows:

$$\begin{aligned}(\widetilde{X}(\alpha), \widetilde{y}(\alpha), \widetilde{S}(\alpha)) &= (\widetilde{X}, \widetilde{y}, \widetilde{S}) + (\widetilde{\Delta X}(\alpha), \widetilde{\Delta y}(\alpha), \widetilde{\Delta S}(\alpha)) \\ &= (\widetilde{X}, \widetilde{y}, \widetilde{S}) + \alpha_-(\widetilde{\Delta X}_-, \Delta y_-, \widetilde{\Delta S}_-) + \alpha_+(\widetilde{\Delta X}_+, \widetilde{\Delta y}_+, \widetilde{\Delta S}_+).\end{aligned} \tag{5.9}$$

The next proposition formalizes the equivalence between the original and scaled problems.

PROPOSITION 5.1. *If* (X, y, S) *and* $(\widetilde{X}, \widetilde{y}, \widetilde{S})$ *are related to each other as specified by* (5.1) *and* $(X(\alpha), y(\alpha), S(\alpha))$ *and* $(\widetilde{X}(\alpha), \widetilde{y}(\alpha), \widetilde{S}(\alpha))$ *are defined by* (4.11) *and* (5.9)*, respectively, then we have*

1. $(X, y, S) \in \mathcal{F}$ *if and only if* $(\widetilde{X}, \widetilde{y}, \widetilde{S})$ *is feasible for* $(\widetilde{\mathcal{P}})$ *and* $(\widetilde{\mathcal{D}})$*;*
2. $(X, y, S) \in \mathcal{N}(\tau_1, \tau_2, \eta)$ *if and only if* $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$*, where* $\widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$ *is the neighborhood corresponding to* $(\widetilde{\mathcal{P}})$ *and* $(\widetilde{\mathcal{D}})$*;*
3. $\widetilde{X}(\alpha) = PX(\alpha)P$, $\widetilde{y}(\alpha) = y(\alpha)$, $\widetilde{S}(\alpha) = P^{-1}S(\alpha)P^{-1}$, *and* $\mu(\alpha) = \widetilde{\mu}_g(\alpha)$, *where* $\widetilde{\mu}_g(\alpha) = \mathrm{Tr}(\widetilde{X}(\alpha)\widetilde{S}(\alpha))/n$.

5.2. Technical results. Before proceeding to the complexity result, we have to prove some technical lemmas. Throughout this section we fix the reference parameter to $\tau = \tau_1$ and let

A.1 $(\widetilde{\Delta X}_-, \widetilde{\Delta y}_-, \widetilde{\Delta S}_-)$ and $(\widetilde{\Delta X}_+, \widetilde{\Delta y}_+, \widetilde{\Delta S}_+)$ be the solutions of (5.5) and (5.6), respectively;

A.2 $\widetilde{\Delta X}(\alpha) := \alpha_-\widetilde{\Delta X}_- + \alpha_+\widetilde{\Delta X}_+$ and $\widetilde{\Delta S}(\alpha) := \alpha_-\widetilde{\Delta S}_- + \alpha_+\widetilde{\Delta S}_+$.

From the following lemma, we see that if the current iterate is feasible, then the search directions are orthogonal.

LEMMA 5.2. *Under* A.1 *and* A.2, *we have*

$$\mathrm{Tr}(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha)) = 0.$$

Proof. The proof is straightforward by using (5.5a), (5.5b), (5.6a), and (5.6b). ∎

LEMMA 5.3. *If* $P \in \mathcal{P}(X, S)$*, then we have*

$$\mathrm{Tr}(\widetilde{X}\widetilde{\Delta S}_-) + \mathrm{Tr}(\widetilde{\Delta X}_-\widetilde{S}) = \mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-), \tag{5.10}$$

and

$$\mathrm{Tr}(\widetilde{X}\widetilde{\Delta S}_+) + \mathrm{Tr}(\widetilde{\Delta X}_+\widetilde{S}) = \mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+). \tag{5.11}$$

Proof. Using the fact that $\mathrm{Tr}(M) = \mathrm{Tr}(H_I(M))$ for any matrix $M \in \mathcal{R}^{n\times n}$ and the systems of (5.5) and (5.6), the results are easily established. ∎

Intuitively, we wish to reduce the duality gap as much as possible in every iteration. The next result, however, shows that Algorithm 1 holds a lower bound for duality gap reduction. This bound derives from feasibility considerations, as we will see in later discussions.

LEMMA 5.4. *Let* $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \mathcal{F}^0$*. Then for every* $\alpha := (\alpha_-, \alpha_+) \in [0, 1]$*, we have*

$$\mathrm{Tr}(\widetilde{X}(\alpha)\widetilde{S}(\alpha)) = \mathrm{Tr}(\widetilde{X}\widetilde{S}) + \alpha_-\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) + \alpha_+\mathrm{Tr}([\tau\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+).$$

Furthermore,

$$\widetilde{\mu}_g(\alpha) = \widetilde{\mu}_g + \alpha_- \frac{\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-)}{n} + \alpha_+ \frac{\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)}{n} \geq (1-\alpha_-)\widetilde{\mu}_g.$$

Proof. Using Lemmas 5.2 and 5.3, the first equality is easy to see. We now apply it to prove the second one,

$$\begin{aligned}\widetilde{\mu}_g(\alpha) &= \frac{\mathrm{Tr}(\widetilde{X}(\alpha)\widetilde{S}(\alpha))}{n} \\ &= \frac{\mathrm{Tr}(\widetilde{X}\widetilde{S})}{n} + \alpha_- \frac{\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-)}{n} + \alpha_+ \frac{\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+}{n} \\ &\geq \widetilde{\mu}_g - \alpha_- \frac{\mathrm{Tr}(\widetilde{X}\widetilde{S})}{n} \\ &= (1-\alpha_-)\widetilde{\mu}_g,\end{aligned}$$

where the inequality is due to the fact that $\widetilde{X}, \widetilde{S} \in \mathcal{S}_+^n$ implies

$$\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) \geq \mathrm{Tr}(-\widetilde{X}\widetilde{S}). \qquad \square$$

The next lemma shows that the negative part of $\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}$ is also bounded in terms of the duality gap at the current iteration.

LEMMA 5.5. *Let* $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \mathcal{F}^0$. *Then*

$$\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) \leq -(1-\tau_1)\mathrm{Tr}(\widetilde{X}\widetilde{S}). \tag{5.12}$$

Proof. Recall that

$$[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^- + [\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+ = \tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}.$$

Taking the trace of both sides, we have

$$\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) = (\tau_1 - 1)\mathrm{Tr}(\widetilde{X}\widetilde{S}) - \mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+) \leq -(1-\tau_1)\mathrm{Tr}(\widetilde{X}\widetilde{S}). \qquad \square$$

The next results, Proposition 5.6 and Corollary 5.7, imply that Algorithm 1 reduces the duality gap steadily if the feasibility of the iterates can be preserved. From now on, we introduce the notation $\beta = (\tau_1 - \tau_2)/\tau_1$. Then we have $\beta \in (0,1)$ and $\tau_2 = (1-\beta)\tau_1$. Further, let us denote

$$\hat{\eta} = \max\left\{ \frac{\left\| \left[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}\right]^+ \right\|_F}{\beta\tau_1\widetilde{\mu}_g}, 1 \right\}.$$

It follows that if $(\widetilde{X}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$, then $1 \leq \hat{\eta} \leq \eta$ and $\hat{\eta} \leq \sqrt{n}$.

PROPOSITION 5.6. *Let* $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$. *Then we have*

$$\widetilde{\mu}_g(\alpha) \leq \widetilde{\mu}_g - \alpha_-(1-\tau_1)\widetilde{\mu}_g + \alpha_+ \frac{\hat{\eta}\beta\tau_1\widetilde{\mu}_g}{\sqrt{n}}.$$

Proof. Using Lemma 5.4, we see that

$$\begin{aligned}\widetilde{\mu}_g(\alpha) &= \widetilde{\mu}_g + \alpha_- \frac{\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-)}{n} + \alpha_+ \frac{\mathrm{Tr}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)}{n} \\ &\le \widetilde{\mu}_g - \alpha_-(1-\tau_1)\frac{\mathrm{Tr}(\widetilde{X}\widetilde{S})}{n} + \alpha_+\sqrt{n}\frac{\left\|[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+\right\|_F}{n} \\ &\le \widetilde{\mu}_g - \alpha_-(1-\tau_1)\widetilde{\mu}_g + \alpha_+\frac{\hat{\eta}\beta\tau_1\widetilde{\mu}_g}{\sqrt{n}},\end{aligned}$$

where the first inequality is due to Lemma 5.5 and the Cauchy–Schwarz inequality and the last inequality derives from the assumption that $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$. ∎

When the parameters τ_1 and β are chosen appropriately and all the iterates reside in the neighborhood $\widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$, we claim that the duality gap is decreasing at the rate of $O(1/\sqrt{n})$.

COROLLARY 5.7. *Let $\tau_1 \le \frac{4}{9}$, $\beta \le \frac{1}{4}$, and $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$. If $\alpha_- = \alpha_+\hat{\eta}\sqrt{\beta\tau_1/n}$, then we have*

$$\widetilde{\mu}_g(\alpha) \le \left(1 - \alpha_+\frac{2\hat{\eta}\sqrt{\beta\tau_1}}{9\sqrt{n}}\right)\widetilde{\mu}_g.$$

Proof. From Proposition 5.6, it follows that

$$\begin{aligned}\widetilde{\mu}_g(\alpha) &\le \widetilde{\mu}_g - \alpha_-(1-\tau_1)\widetilde{\mu}_g + \alpha_+\frac{\hat{\eta}\beta\tau_1\widetilde{\mu}_g}{\sqrt{n}} \\ &\le \widetilde{\mu}_g - \frac{5}{9}\alpha_-\widetilde{\mu}_g + \alpha_+\frac{\hat{\eta}\beta\tau_1\widetilde{\mu}_g}{\sqrt{n}} \\ &= \widetilde{\mu}_g - \alpha_+\left(\frac{5}{9} - \sqrt{\beta\tau_1}\right)\frac{\hat{\eta}\sqrt{\beta\tau_1}}{\sqrt{n}}\widetilde{\mu}_g \\ &\le \left(1 - \alpha_+\frac{2\hat{\eta}\sqrt{\beta\tau_1}}{9\sqrt{n}}\right)\widetilde{\mu}_g. \qquad \blacksquare\end{aligned}$$

Subsequently, we show how to ensure all the iterates in the neighborhood $\widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$. Although we wish to decrease the duality gap as much as possible, we still need to control the smallest eigenvalue of $\widetilde{X}(\alpha)\widetilde{S}(\alpha)$ in order to stay in the neighborhood $\widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$.

LEMMA 5.8. *Suppose $P \in \mathcal{P}(X, S)$ and $\chi(\alpha) = \widetilde{X}\widetilde{S} + \alpha_-[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^- + \alpha_+[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+$. If $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$, then we have*

$$\lambda_{\min}(\chi(\alpha)) \ge \tau_2\widetilde{\mu}_g + \alpha_+(\tau_1 - \tau_2)\widetilde{\mu}_g. \tag{5.13}$$

Proof. We first consider the situation when $\lambda_{\min}(\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}) \ge 0$ and note that $\lambda_{\min}(\cdot)$ is a homogeneous concave function. Then we have

$$\begin{aligned}\lambda_{\min}(\chi(\alpha)) &= \lambda_{\min}(\widetilde{X}\widetilde{S} + \alpha_+[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+) \\ &= \lambda_{\min}(\widetilde{X}\widetilde{S} + \alpha_+(\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S})) \\ &= \lambda_{\min}((1-\alpha_+)\widetilde{X}\widetilde{S} + \alpha_+\tau_1\widetilde{\mu}_g I) \\ &\ge (1-\alpha_+)\lambda_{\min}(\widetilde{X}\widetilde{S}) + \alpha_+\tau_1\widetilde{\mu}_g \\ &\ge (1-\alpha_+)\tau_2\widetilde{\mu}_g + \alpha_+\tau_1\widetilde{\mu}_g \\ &= \tau_2\widetilde{\mu}_g + \alpha_+(\tau_1 - \tau_2)\widetilde{\mu}_g.\end{aligned}$$

The second inequality holds due to $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta)$.

When $\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}$ is negative semidefinite, i.e., $[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^- = \tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}$ and $[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+ = 0$, we have

$$\begin{aligned}\lambda_{\min}(\chi(\alpha)) &= \lambda_{\min}(\widetilde{X}\widetilde{S} + \alpha_-[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-)\\ &= \lambda_{\min}(\Lambda + \alpha_-(\tau_1\widetilde{\mu}_g I - \Lambda))\\ &\geq \lambda_{\min}(\Lambda + (\tau_1\widetilde{\mu}_g I - \Lambda))\\ &= \tau_1\widetilde{\mu}_g\\ &= \tau_2\widetilde{\mu}_g + (\tau_1 - \tau_2)\widetilde{\mu}_g\\ &\geq \tau_2\widetilde{\mu}_g + \alpha_+(\tau_1 - \tau_2)\widetilde{\mu}_g.\end{aligned}$$

Now let us consider the last case where $\tau_1\mu_g I - \widetilde{X}\widetilde{S}$ is indefinite. Recall that the eigenvalues of $\widetilde{X}\widetilde{S}$ are ordered increasingly, i.e., $\lambda_1 \leq \cdots \leq \lambda_n$. Assume λ_k is the first eigenvalue of $\widetilde{X}\widetilde{S}$ such that $\tau_1\widetilde{\mu}_g - \lambda_k \leq 0$, i.e., $\tau_1\widetilde{\mu}_g - \lambda_1 \geq \cdots \geq \tau_1\widetilde{\mu}_g - \lambda_{k-1} > 0 \geq \tau_1\widetilde{\mu}_g - \lambda_k \geq \cdots \geq \tau_1\widetilde{\mu}_g - \lambda_n$. It is easy to see that

$$\begin{aligned}\lambda_{\min}(\chi(\alpha)) &= \lambda_{\min}(\widetilde{X}\widetilde{S} + \alpha_-[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) + \alpha_+[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)\\ &= \lambda_{\min}(\Lambda + \alpha_-[\tau_1\widetilde{\mu}_g I - \Lambda]^- + \alpha_+[\tau_1\widetilde{\mu}_g I - \Lambda]^+)\\ &= \min\{\lambda_1 + \alpha_+(\tau_1\widetilde{\mu}_g - \lambda_1), \lambda_k + \alpha_-(\tau_1\widetilde{\mu}_g - \lambda_k)\}\\ &\geq \min\{\tau_2\widetilde{\mu}_g + \alpha_+(\tau_1 - \tau_2)\widetilde{\mu}_g, \tau_1\widetilde{\mu}_g\}\\ &\geq \tau_2\widetilde{\mu}_g + \alpha_+(\tau_1 - \tau_2)\widetilde{\mu}_g.\end{aligned}$$

Taking all of the possible cases into account, we conclude that (5.13) is true. ☐

To follow the central path, we also need to make sure that the iterates remain in the prescribed neighborhood of the central path.

LEMMA 5.9. *Suppose* $P \in \mathcal{P}(X,S)$ *and* $\chi(\alpha) = \widetilde{X}\widetilde{S} + \alpha_-[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^- + \alpha_+[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+$. *If* $(\widetilde{X},\widetilde{y},\widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1,\tau_2,\eta)$, *then we have*

$$\left\|[\tau_1\widetilde{\mu}_g(\alpha)I - \chi(\alpha)]^+\right\|_F \leq (1-\alpha_+)\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha). \tag{5.14}$$

Proof. Assume that the eigenvalues of $\widetilde{X}\widetilde{S}$ are ordered so that

$$\tau_1\widetilde{\mu}_g - \lambda_1 \geq \tau_1\widetilde{\mu}_g - \lambda_2 \geq \cdots \geq \tau_1\widetilde{\mu}_g - \lambda_{k-1} \geq 0 \geq \tau_1\widetilde{\mu}_g - \lambda_k \geq \cdots \geq \tau_1\widetilde{\mu}_g - \lambda_n.$$

Now let us consider the diagonal elements of $\Lambda + \alpha_-[\tau_1\widetilde{\mu}_g I - \Lambda]^- + \alpha_+[\tau_1\widetilde{\mu}_g I - \Lambda]^+$. For $i = 1, \dots, k-1$, $\lambda_i + \alpha_+(\tau_1\widetilde{\mu}_g - \lambda_i) = (1-\alpha_+)\lambda_i + \alpha_+\tau_1\widetilde{\mu}_g$, then

$$\begin{aligned}\tau_1\widetilde{\mu}_g(\alpha) - (\lambda_i + \alpha_+(\tau_1\widetilde{\mu}_g - \lambda_i)) &\leq \tau_1\widetilde{\mu}_g(\alpha) - \frac{\widetilde{\mu}_g(\alpha)}{\widetilde{\mu}_g}(\lambda_i + \alpha_+(\tau_1\widetilde{\mu}_g - \lambda_i))\\ &= \frac{\widetilde{\mu}_g(\alpha)}{\widetilde{\mu}_g}(1-\alpha_+)(\tau_1\widetilde{\mu}_g - \lambda_i).\end{aligned}$$

For $i = k, \dots, n$, $\lambda_i + \alpha_-(\tau_1\widetilde{\mu}_g - \lambda_i) \geq \lambda_i + \tau_1\widetilde{\mu}_g - \lambda_i = \tau_1\widetilde{\mu}_g \geq 0$, then

$$\tau_1\widetilde{\mu}_g(\alpha) - (\lambda_i + \alpha_-(\tau_1\widetilde{\mu}_g - \lambda_i)) \leq \tau_1\widetilde{\mu}_g - \tau_1\widetilde{\mu}_g = 0.$$

For convenience, let $\varphi(\alpha) = [\tau_1\widetilde{\mu}_g(\alpha)I - (\Lambda + \alpha_-[\tau_1\widetilde{\mu}_g I - \Lambda]^- + \alpha_+[\tau_1\widetilde{\mu}_g I - \Lambda]^+)]^+$.

Therefore, together with Lemma 3.2, we have

$$\begin{aligned}\|\varphi(\alpha)\|_F &\le \frac{\widetilde{\mu}_g(\alpha)}{\widetilde{\mu}_g}(1-\alpha_+)\left\|[\tau_1\widetilde{\mu}_g I-\Lambda]^+\right\|_F\\ &= \frac{\widetilde{\mu}_g(\alpha)}{\widetilde{\mu}_g}(1-\alpha_+)\left\|[\tau_1\widetilde{\mu}_g I-\widetilde{X}\widetilde{S}]^+\right\|_F\\ &\le (1-\alpha_+)\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha).\end{aligned} \tag{5.15}$$

Note that $\|[\tau_1\widetilde{\mu}_g(\alpha)I-\chi(\alpha)]^+\|_F = \|\varphi(\alpha)\|_F$; the proof is complete. ∎

The next two lemmas together bound the distance between the current iterate and our reference point $\tau_1\widetilde{\mu}_g I$ on the central path.

LEMMA 5.10. *Let $X, S \in \mathcal{S}^n_{++}$, $P \in \mathcal{P}(X,S)$, $\widetilde{X}$ and $\widetilde{S}$ be defined by* (5.1), *and $\widetilde{E}$ and $\widetilde{F}$ be defined by* (5.8). *Then*

$$\left\|(\widetilde{F}\widetilde{E})^{-1/2}\mathbf{vec}([\tau_1\widetilde{\mu}_g I-\widetilde{X}\widetilde{S}]^-)\right\|^2 \le \mathrm{Tr}(\widetilde{X}\widetilde{S}). \tag{5.16}$$

Proof. Using (5.8) and Proposition 4.2 of Monteiro and Zhang [14], we find the spectral decompositions of $\widetilde{E}$ and $\widetilde{F}$ to be

$$\widetilde{E} = \frac{1}{2}(\widetilde{S}\otimes I + I\otimes\widetilde{S}) = \frac{1}{2}Q_K(\Lambda(\widetilde{S})\otimes I + I\otimes\Lambda(\widetilde{S}))Q_K^T,$$

$$\widetilde{F} = \frac{1}{2}(\widetilde{X}\otimes I + I\otimes\widetilde{X}) = \frac{1}{2}Q_K(\Lambda(\widetilde{X})\otimes I + I\otimes\Lambda(\widetilde{X}))Q_K^T,$$

where $Q_K = Q\otimes Q$ is an $n^2\times n^2$ orthogonal matrix. Furthermore, because $\widetilde{X}$ and $\widetilde{S}$ commute, from Proposition 4.1 of Monteiro and Zhang [14], we have $\widetilde{F}\widetilde{E}\in\mathcal{S}^{n^2}_{++}$. Then we have

$$(\widetilde{F}\widetilde{E})^{-1} = 4Q_K(\Lambda\otimes I + I\otimes\Lambda + \Lambda(\widetilde{X})\otimes\Lambda(\widetilde{S}) + \Lambda(\widetilde{S})\otimes\Lambda(\widetilde{X}))^{-1}Q_K^T,$$

where the matrix in the middle is diagonal with the properties that the $((i-1)n+i)$th component is $1/(4\lambda_i)$ and the largest component is $1/(4\lambda_1)$. On the other hand,

$$\begin{aligned}\mathbf{vec}(\tau_1\widetilde{\mu}_g I-\widetilde{X}\widetilde{S}) &= \mathbf{vec}(\tau_1\widetilde{\mu}_g I - Q\Lambda Q^T)\\ &= (Q\otimes Q)\mathbf{vec}(\tau_1\widetilde{\mu}_g I-\Lambda)\\ &= Q_K\mathbf{vec}(\tau_1\widetilde{\mu}_g I-\Lambda),\end{aligned}$$

where $\mathbf{vec}(\tau_1\widetilde{\mu}_g I-\Lambda)$ is an n^2-vector with at most n nonzeros at the $((i-1)n+i)$th entries which are equal to $\tau_1\widetilde{\mu}_g-\lambda_i$. Finally, we have

$$\begin{aligned}\left\|(\widetilde{F}\widetilde{E})^{-1/2}\mathbf{vec}([\tau_1\widetilde{\mu}_g I-\widetilde{X}\widetilde{S}]^-)\right\|^2 &= (\mathbf{vec}([\tau_1\widetilde{\mu}_g I-\widetilde{X}\widetilde{S}]^-))^T(\widetilde{E}\widetilde{F})^{-1}\mathbf{vec}([\tau_1\widetilde{\mu}_g I-\widetilde{X}\widetilde{S}]^-)\\ &= \sum_{i=1}^n([\tau_1\widetilde{\mu}_g-\lambda_i]^-)^2/\lambda_i\\ &= \sum_{i=1}^n\left(\left[\sqrt{\lambda_i}-\tau_1\widetilde{\mu}_g/\sqrt{\lambda_i}\right]^+\right)^2\\ &\le \sum_{i=1}^n\lambda_i\\ &= \mathrm{Tr}(\widetilde{X}\widetilde{S}),\end{aligned}$$

which leads to inequality (5.16). ∎

LEMMA 5.11. *Let $P \in \mathcal{P}(X,S)$, $\widetilde{X}$ and $\widetilde{S}$ be defined by* (5.1)*, and $\widetilde{E}$ and $\widetilde{F}$ be defined by* (5.8)*. If $(\widetilde{X},\widetilde{y},\widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1,\tau_2,\eta)$ and $\beta \le 1/4$, then*

$$\left\|(\widetilde{F}\widetilde{E})^{-1/2}\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)\right\|^2 \le \hat{\eta}^2\beta\tau_1\widetilde{\mu}_g/3.$$

Proof. Note the fact that $\lambda_{\min}(\widetilde{F}\widetilde{E}) = \lambda_1 \ge \tau_2\widetilde{\mu}_g$. Then one has

$$\begin{aligned}\left\|(\widetilde{F}\widetilde{E})^{-1/2}\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)\right\|^2 &\le \left\|(\widetilde{F}\widetilde{E})^{-1/2}\right\|^2\left\|\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)\right\|^2\\ &= \left\|(\widetilde{F}\widetilde{E})^{-1/2}\right\|^2\left\|[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+\right\|_F^2\\ &\le \hat{\eta}^2\beta^2\tau_1^2\widetilde{\mu}_g^2/(\tau_2\widetilde{\mu}_g)\\ &\le \hat{\eta}^2\beta\tau_1\widetilde{\mu}_g/3.\end{aligned}$$

The last inequality follows from the fact that $\beta \le 1/4$ implies $\beta\tau_1/\tau_2 \le 1/3$. ∎

We now apply Lemmas 5.10 and 5.11 to conclude the following result.

LEMMA 5.12. *Let $P \in \mathcal{P}(X,S)$ and $G = \widetilde{E}^{-1}\widetilde{F}$. If $(\widetilde{X},\widetilde{y},\widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1,\tau_2,\eta)$ and $\beta \le 1/4$, then*

$$\begin{aligned}\left\|G^{-1/2}\mathbf{vec}(\widetilde{\Delta X}(\alpha))\right\|^2 + \left\|G^{1/2}\mathbf{vec}(\widetilde{\Delta S}(\alpha))\right\|^2 &+ 2\mathrm{Tr}(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\\ &\le \alpha_-^2\,\mathrm{Tr}(\widetilde{X}\widetilde{S}) + \alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g/3.\end{aligned}$$

Proof. From (5.7), we have

$$\widetilde{E}\mathbf{vec}(\widetilde{\Delta X}(\alpha)) + \widetilde{F}\mathbf{vec}(\widetilde{\Delta S}(\alpha)) = \alpha_-\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) + \alpha_+\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+).$$

Applying Proposition 1.1 of Zhang [23] to this equality, we obtain

$$\begin{aligned}&\left\|(\widetilde{F}\widetilde{E})^{-1/2}\widetilde{E}\mathbf{vec}(\widetilde{\Delta X}(\alpha))\right\|^2 + \left\|(\widetilde{F}\widetilde{E})^{-1/2}\widetilde{F}\mathbf{vec}(\widetilde{\Delta S}(\alpha))\right\|^2 + 2\mathrm{Tr}(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\\ &\qquad = \left\|(\widetilde{F}\widetilde{E})^{-1/2}[\alpha_-\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) + \alpha_+\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)]\right\|^2.\end{aligned}$$

The commutativity of $\widetilde{E}$ and $\widetilde{F}$ implies that

$$(\widetilde{F}\widetilde{E})^{-1/2}\widetilde{E} = (\widetilde{E}^{-1}\widetilde{F})^{-1/2} = G^{-1/2}, \quad (\widetilde{F}\widetilde{E})^{-1/2}\widetilde{F} = (\widetilde{E}^{-1}\widetilde{F})^{1/2} = G^{1/2}.$$

Hence, to complete the proof, it is sufficient to show that

$$\begin{aligned}&\left\|(\widetilde{F}\widetilde{E})^{-1/2}[\alpha_-\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-) + \alpha_+\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)]\right\|^2\\ &\quad \le \alpha_-^2\left\|(\widetilde{F}\widetilde{E})^{-1/2}\mathbf{vec}([\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^-)\right\|^2 + \alpha_+^2\left\|(\widetilde{F}\widetilde{E})^{-1/2}\mathbf{vec}((\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)\right\|^2\\ &\quad \le \alpha_-^2\,\mathrm{Tr}(\widetilde{X}\widetilde{S}) + \alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g/3,\end{aligned}$$

where the last inequality can be derived from Lemmas 5.10 and 5.11. ∎

We now would like to explore a bound for the second order term $\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha)$.

LEMMA 5.13. *Let $P \in \mathcal{P}(X,S)$ and $G = \widetilde{E}^{-1}\widetilde{F}$. If $\beta \le 1/4$, $\alpha_- = \alpha_+\hat{\eta}\sqrt{\beta\tau_1/n}$ and $(\widetilde{X},\widetilde{y},\widetilde{S}) \in \widetilde{\mathcal{N}}(\tau_1,\tau_2,\eta)$, then we have*

(5.17)
$$\left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F \le \left\|\mathbf{vec}(\widetilde{\Delta X}(\alpha))\right\|\left\|\mathbf{vec}(\widetilde{\Delta S}(\alpha))\right\| \le \frac{2}{3}\sqrt{\mathrm{cond}(G)}\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g.$$

Proof. From the last inequality in Lemma 3.2 in Monteiro [11], we have

$$\begin{aligned}\left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F &\leq \left\|\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha)\right\|_F \\ &\leq \left\|\widetilde{\Delta X}(\alpha)\right\|_F \left\|\widetilde{\Delta S}(\alpha)\right\|_F \\ &= \left\|\mathbf{vec}(\widetilde{\Delta X}(\alpha))\right\| \left\|\mathbf{vec}(\widetilde{\Delta S}(\alpha))\right\|.\end{aligned}$$

Further because of Lemmas 5.2 and 4.6 in Monteiro and Zhang [14], it follows that

$$\begin{aligned}\left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F &\leq \left\|\mathbf{vec}(\widetilde{\Delta X}(\alpha))\right\| \left\|\mathbf{vec}(\widetilde{\Delta S}(\alpha))\right\| \\ &\leq \frac{\sqrt{\mathrm{cond}(G)}}{2}\left(\left\|G^{-1/2}\mathbf{vec}(\widetilde{\Delta X}(\alpha))\right\|^2 + \left\|G^{1/2}\mathbf{vec}(\widetilde{\Delta S}(\alpha))\right\|^2\right).\end{aligned}$$

Substitute α_- with $\alpha_+\hat{\eta}\sqrt{\beta\tau_1/n}$ and apply Lemma 5.12. Then we eventually obtain

$$\begin{aligned}\left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F &\leq \frac{\sqrt{\mathrm{cond}(G)}}{2}\left(\frac{\alpha_-^2\,\mathrm{Tr}(\widetilde{X}\widetilde{S}) + \alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g}{3}\right) \\ &\leq \frac{\sqrt{\mathrm{cond}(G)}}{2}\left(\frac{\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g + \alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g}{3}\right) \\ &= \frac{2}{3}\sqrt{\mathrm{cond}(G)}\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g,\end{aligned}$$

observing that $\mathrm{Tr}(\widetilde{X}\widetilde{S}) = n\widetilde{\mu}_g$. ☐

The next proposition provides a sufficient condition which guarantees that all the iterates remain in the neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$.

PROPOSITION 5.14. *Let $(X, y, S) \in \mathcal{N}(\tau_1, \tau_2, \eta)$, $\tau_1 < 4/9$, $\beta \leq 1/4$, $P \in \mathcal{P}(X, S)$, and $G = \widetilde{E}^{-1}\widetilde{F}$. If $\alpha_- = \alpha_+\hat{\eta}\sqrt{\beta\tau_1/n}$ and $\alpha_+ \leq 1/(\sqrt{\mathrm{cond}(G)}\hat{\eta}^2)$, then*

$$(X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}(\tau_1, \tau_2, \eta).$$

Proof. By Corollary 5.7 we have $\widetilde{\mu}_g(\alpha) \leq \widetilde{\mu}_g$. Further, using Lemma 5.8 and the fact that $\lambda_{\min}(\cdot)$ is a homogeneous concave function on the space of symmetric matrices, one has

$$\begin{aligned}\lambda_{\min}(H_I(\widetilde{X}(\alpha)\widetilde{S}(\alpha))) &\geq \lambda_{\min}(H_I(\widetilde{X}\widetilde{S} + \alpha_-[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^- + \alpha_+[\tau_1\widetilde{\mu}_g I - \widetilde{X}\widetilde{S}]^+)) \\ &\quad + \lambda_{\min}(H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))) \\ &\geq \lambda_{\min}(\chi(\alpha)) - \left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\| \\ &\geq \tau_2\widetilde{\mu}_g + \alpha_+(\tau_1 - \tau_2)\widetilde{\mu}_g - \left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F,\end{aligned}$$

where the second inequality follows from the fact that the absolute value of the smallest eigenvalue of a symmetric matrix is less than or equal to its Euclidean norm and the third inequality is due to the fact that the Euclidean norm of a symmetric matrix is less than or equal to its Frobenius norm. These two results follow directly from

Lemma 3.1 of Monteiro [11]. From Lemma 5.13 one can further derive that

$$\begin{aligned}\lambda_{\min}(H_I(\widetilde{X}(\alpha)\widetilde{S}(\alpha))) &\geq \tau_2\widetilde{\mu}_g + \alpha_+(\tau_1-\tau_2)\widetilde{\mu}_g - \frac{2}{3}\sqrt{\operatorname{cond}(G)}\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g\\ &\geq \tau_2\widetilde{\mu}_g + \alpha_+\beta\tau_1\widetilde{\mu}_g - \alpha_+\beta\tau_1\widetilde{\mu}_g\\ &= \tau_2\widetilde{\mu}_g\\ &\geq \tau_2\widetilde{\mu}_g(\alpha)\\ &> 0.\end{aligned}$$

This reveals that $\widetilde{X}(\alpha)\widetilde{S}(\alpha)$ is nonsingular and further implies that each of the factors $\widetilde{X}(\alpha)$ and $\widetilde{S}(\alpha)$ are nonsingular as well. By using continuity, it follows that $\widetilde{X}(\alpha)$ and $\widetilde{S}(\alpha)$ are also in $\mathcal{S}_{++}^n$, since $\widetilde{X},\ \widetilde{S} \succ 0$. Then we claim that

$$\lambda_{\min}(\widetilde{X}(\alpha)\widetilde{S}(\alpha)) \geq \lambda_{\min}(H_I(\widetilde{X}(\alpha)\widetilde{S}(\alpha))) \geq \tau_2\widetilde{\mu}_g(\alpha). \tag{5.18}$$

Since $\beta \leq 1/4$ and $\tau_1 \leq 4/9$, from Lemma 5.4, we have

$$\widetilde{\mu}_g(\alpha) \geq (1-\alpha_-)\widetilde{\mu}_g = (1-\alpha_+\hat{\eta}\sqrt{\beta\tau_1/n})\widetilde{\mu}_g \geq (1-\sqrt{\beta\tau_1})\widetilde{\mu}_g \geq (2/3)\widetilde{\mu}_g. \tag{5.19}$$

From Lemma 3.3, we have

$$\begin{aligned}\psi(\alpha) &:= \left\|[\tau_1\widetilde{\mu}_g(\alpha)I - \widetilde{X}^{1/2}(\alpha)\widetilde{S}(\alpha)\widetilde{X}^{1/2}(\alpha)]^+\right\|_F\\ &\leq \left\|[H_{\widetilde{X}^{1/2}(\alpha)}(\tau_1\widetilde{\mu}_g(\alpha)I - \widetilde{X}^{1/2}(\alpha)\widetilde{S}(\alpha)\widetilde{X}^{1/2}(\alpha))]^+\right\|_F\\ &= \left\|[H_I(\tau_1\widetilde{\mu}_g(\alpha)I - \widetilde{X}(\alpha)\widetilde{S}(\alpha))]^+\right\|_F.\end{aligned}$$

Because $\widetilde{X}(\alpha)\widetilde{S}(\alpha) = (\widetilde{X} + \alpha_-\widetilde{\Delta X}_- + \alpha_+\widetilde{\Delta X}_+)(\widetilde{S} + \alpha_-\widetilde{\Delta S}_- + \alpha_+\widetilde{\Delta S}_+)$, we have

$$\begin{aligned}\psi(\alpha) \leq &\left\|\left[H_I(\tau_1\widetilde{\mu}_g(\alpha)I - \widetilde{X}\widetilde{S} - \alpha_-[\tau\widetilde{\mu}_gI - \widetilde{X}\widetilde{S}]^- - \alpha_+[\tau\widetilde{\mu}_gI - \widetilde{X}\widetilde{S}]^+)\right]^+\right\|_F +\\ &\left\|\left[-H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right]^+\right\|_F\\ = &\left\|\left[\tau_1\widetilde{\mu}_g(\alpha)I - \widetilde{X}\widetilde{S} - \alpha_-[\tau\widetilde{\mu}_gI - \widetilde{X}\widetilde{S}]^- - \alpha_+[\tau\widetilde{\mu}_gI - \widetilde{X}\widetilde{S}]^+\right]^+\right\|_F +\\ &\left\|\left[H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right]^-\right\|_F,\end{aligned}$$

where the second inequality is from Proposition 3.1.

Using the fact that $\left\|[H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))]^-\right\|_F \leq \left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F$ and Lemma 5.9, we can prove

$$\psi(\alpha) \leq (1-\alpha_+)\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha) + \left\|H_I(\widetilde{\Delta X}(\alpha)\widetilde{\Delta S}(\alpha))\right\|_F.$$

Further, from Lemma 5.13 and inequality (5.19), one has

$$\begin{aligned}\psi(\alpha) &\leq (1-\alpha_+)\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha) + \frac{2}{3}\sqrt{\operatorname{cond}(G)}\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g\\ &\leq (1-\alpha_+)\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha) + \sqrt{\operatorname{cond}(G)}\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g(\alpha).\end{aligned}$$

Since $\alpha_+ \le 1/(\sqrt{\mathrm{cond}(G)}\hat{\eta}^2)$ and $\widetilde{\eta} \ge 1$, we have $\sqrt{\mathrm{cond}(G)}\alpha_+^2\hat{\eta}^2\beta\tau_1\widetilde{\mu}_g(\alpha) \le \alpha_+\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha)$. Thus,

$$\begin{aligned}\psi(\alpha) &\le (1-\alpha_+)\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha) + \alpha_+\hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha)\\ &= \hat{\eta}\beta\tau_1\widetilde{\mu}_g(\alpha)\\ &\le \eta\beta\tau_1\widetilde{\mu}_g(\alpha)\\ &= \eta(\tau_1-\tau_2)\widetilde{\mu}_g(\alpha).\end{aligned}$$

This, together with (5.18), implies that

$$(\widetilde{X}(\alpha), \widetilde{y}(\alpha), \widetilde{S}(\alpha)) \in \widetilde{\mathcal{N}}(\tau_1, \tau_2, \eta).$$

Consequently, according to Proposition (5.1), one has

$$(X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}(\tau_1, \tau_2, \eta). \qquad \square$$

5.3. Polynomial complexity. In this section we present our main complexity result. The next theorem gives an iteration-complexity bound for Algorithm 1 in terms of a parameter κ_∞ defined as

$$\kappa_\infty = \sup\left\{\mathrm{cond}((\widetilde{E}^k)^{-1}\widetilde{F}^k) : \ k = 0, 1, \ldots\right\}. \tag{5.20}$$

Obviously, $\kappa_\infty \ge 1$.

THEOREM 5.15. *Suppose that $\kappa_\infty < \infty$, $\eta \ge 1$, $0 < \tau_2 < \tau_1 \le 4/9$, and $\beta \le 1/4$ are fixed parameters. At each iteration, let $P^k \in \mathcal{P}(X^k, S^k)$. Then Algorithm 1 will terminate in $O(\eta\sqrt{\kappa_\infty n}\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$ iterations with a solution $\mathrm{Tr}(XS) \le \epsilon$.*

Proof. In every iteration, let $\hat{\alpha} = (\sqrt{\beta\tau_1/(\kappa_\infty n)}/\hat{\eta}, 1/(\sqrt{\kappa_\infty}\hat{\eta}^2))$. By Proposition 5.14, we have

$$(X(\hat{\alpha}), y(\hat{\alpha}), S(\hat{\alpha})) \in \mathcal{N}(\tau_1, \tau_2, \eta).$$

Furthermore, from Lemma 5.7, we also conclude that

$$\widetilde{\mu}_g(\alpha) \le \left(1 - \frac{2\sqrt{\beta\tau_1}}{9\hat{\eta}\sqrt{\mathrm{cond}(G)n}}\right)\widetilde{\mu}_g \le \left(1 - \frac{2\sqrt{\beta\tau_1}}{9\eta\sqrt{\mathrm{cond}(G)n}}\right)\widetilde{\mu}_g \le \left(1 - \frac{2\sqrt{\beta\tau_1}}{9\eta\sqrt{\kappa_\infty n}}\right)\widetilde{\mu}_g,$$

from which the statement of the theorem follows. $\square$

Theorem 5.15 allows us to derive various iteration complexities of Algorithm 1 in terms of some specific aforementioned scaling matrices P.

COROLLARY 5.16. *If the parameter η is a constant, then for Algorithm* 1, *when it is based on the NT direction, the iteration-complexity bound is $O(\sqrt{n}\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$. When the H..K..M scaling is used, then Algorithm* 1 *terminates in at most $O(n\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$ iterations.*

COROLLARY 5.17. *If the parameter η is in the order of $\sqrt{n}$, then for Algorithm* 1, *when it is based on the NT direction, the iteration-complexity bound is $O(n\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$. When the H..K..M scaling is used, then Algorithm* 1 *terminates in at most $O(n^{3/2}\log\frac{\mathrm{Tr}(X^0S^0)}{\epsilon})$ iterations.*

Corollaries 5.16 and 5.17 are readily achieved if we notice that $\kappa_\infty = 1$ for the NT scaling and $\kappa_\infty \le n/\tau_2$ for the H..K..M scaling.

As we see, when η is a constant and the NT scaling is used, the Algorithm 1 achieves its best complexity bound which coincides with the best known complexity of IPMs for SDO. When η is in the order of $\sqrt{n}$, our complexity result is the same as the one for classical large neighborhood IPMs, since we have shown in Proposition 3.6 that in that case our neighborhood $\mathcal{N}(\tau_1, \tau_2, \eta)$ is exactly the large neighborhood $\mathcal{N}(1 - \tau_2)$.

6. Conclusions and further work. As discussed previously, this paper provides a new large neighborhood path-following algorithm with the same theoretical complexity bound as the best short step path-following algorithm. Our preliminary implementation and benchmark for pure SDO problems from SDPLIB [5] and DIMACS [18] also provide us with a promising evidence that our new algorithm may also perform well in practice. With NT scaling, Algorithm 1 on average saves 1 or 2 iterations compared with SDPT3 without using the Mehrotra predictor-corrector heuristic, while not increasing the total computational time. However, there are still some unsettled issues for implementation. For example, sophisticated and efficient strategies to choose step lengths deserve more work. Another issue of interest is how to compute efficiently the positive and negative parts of the right-hand side in the Newton equation when we do not use the NT scaling. It is apparent that, due to the high computational cost, explicitly computing them with eigenvalue decomposition is not desirable. To develop computationally efficient implementation strategies for our large neighborhood algorithm remains the subject of further research.

Because we were successful in combining Ai and Zhang's work [2] for LO and our work for SDO, a natural question to ask is whether this also applies to second order cone optimization problems and further to general conic optimization. Lastly, we are also curious about the relationship between the classical IPMs, the self-regular IPMs, and our new algorithm.

Acknowledgments. We would like to thank the anonymous referees and the associate editor for their constructive comments that helped us to improve the presentation of the paper.

REFERENCES

[1] W. Ai, *Neighborhood-following algorithms for linear programming*, Sci. China Ser. A, 47 (2004), pp. 812–820.

[2] 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.

[3] F. Alizadeh, *Combinatorial Optimization with Interior-Point Methods and Semi-definite Matrices*, Ph.D. thesis, Computer Science Department, University of Minnesota, Minneapolis, MN, 1991.

[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] B. Borchers, *SDPLIB* 1.2*, library of semidefinite programming test problems*, Optim. Methods Softw., 11 (1999), pp. 683–690.

[6] E. de Klerk, *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[7] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 2nd ed., 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, 1990.

[10] 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.

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

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

[13] R.D.C. Monteiro and M.J. Todd, *Path-following methods*, in *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer Academic Publishers, Boston, 2000, pp. 267–306.

[14] R.D.C. Monteiro and Y. Zhang, *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.

[15] Y.E. Nesterov and A.S. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming: Theory and Applications*, SIAM, Philadelphia, PA, 1994.

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

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

[18] G. Pataki and S. Schmieta, *The DIMACS Library of Semidefinite-Quadratic-Linear Programs*, Technical report, Rutgers University, New Brunswick, NJ, 1999, also available online from http://dimacs.rutgers.edu/Challenges/Seventh/Instances/.

[19] J. Peng, C. Roos, and T. Terlaky, *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*, Princeton University Press, Princeton, New Jersey, 2002.

[20] M.J. Todd, *A study of search directions in primal-dual interior-point methods for semidefinite programming*, Optim. Methods Softw., 11 (1999), pp. 1–46.

[21] 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.

[22] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.

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