METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 6, No. 4, pp. 549–570, December 1999

ON A GENERAL CLASS OF INTERIOR-POINT ALGORITHMS FOR SEMIDEFINITE PROGRAMMING WITH POLYNOMIAL COMPLEXITY AND SUPERLINEAR CONVERGENCE*

JUN JI†, FLORIAN A. POTRA‡, AND RONGQIN SHENG§

Abstract. We propose a unified analysis for a class of infeasible-start predictor-corrector algorithms for semidefinite programming problems, using the Monteiro-Zhang unified direction. The algorithms are direct generalizations of the Mizuno-Todd-Ye predictor-corrector algorithm for linear programming. We show that the algorithms belonging to this class are globally convergent, provided the problem has a solution, and have the best known complexity. We also give simple sufficient conditions for superlinear convergence. Our results generalize the results obtained by Potra and Sheng for the infeasible-interior-point algorithm proposed by Kojima, Shida and Shindoh and Potra and Sheng.

1. Introduction. In this paper, we consider the semidefinite programming (SDP) problem:

$$(1.1) \qquad (P) \qquad \min\{C \bullet X : A_i \bullet X = b_i, \quad i = 1, \dots, m, \ X \succeq 0\},$$

and its associated dual problem:

$$(1.2) \qquad (D) \qquad \max\left\{b^T y : \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0\right\}$$

where C, X, and A_i are symmetric matrices in $\mathbb{R}^{n\times n}$, $b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$, $G \bullet H = \mathrm{Tr}(G^T H)$, and $X \succeq 0$ indicates that X is positive semidefinite.

SDP arises in many scientific and engineering fields, including system and control theory, combinatorial optimization, and eigenvalue optimization. Over the last couple of years SDP has attracted very active research, focusing on the extension of the existing methods in LP to the context of SDP. Several generalizations of the Mizuno-Todd-Ye predictor-corrector method for SDP have been recently analyzed by Lin and Saigal [6], Luo, et al. [7], Kojima, et al. [3, 4], Potra and Sheng [12]–[14], and Zhang [18]. The algorithm proposed by Kojima, et al. [3] and Potra and Sheng [13] uses the HRVW/KSH/M direction and has polynomial complexity. Also, Potra and Sheng [13] proposed a sufficient condition for the superlinear convergence of the algorithm while Kojima, et al. [3] established the superlinear convergence under the following three assumptions:

- A1 SDP has a strictly complementary solution;
- A2 SDP is nondegenerate in the sense that the Jacobian matrix of its KKT system is nonsingular;
- A3 the iterates converge tangentially to the central path in the sense that the size of the neighborhood in which the iterates reside must approach zero, namely,

$$\lim_{k\to\infty} \left\|(X^k)^{\frac{1}{2}} S^k (X^k)^{\frac{1}{2}} - (X^k \bullet S^k/n) I\right\|_F / (X^k \bullet S^k/n) = 0.$$

*Received January 5, 1998; revised March 26, 1999.

†Department of Mathematics and Computer Science, Valdosta State University, Valdosta, GA 31698, USA (junji@grits.valdosta.peachnet.edu).

‡Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA (potra@math.umbc.edu).

§Lucent Technologies, 2443 Warrenville Road, Lisle, IL 60532, USA (rsheng@lucent.com).

These were the first two papers investigating the local convergence properties of interior-point algorithms for semidefinite programming.

More recently, Kojima, et al. [4] proposed a predictor-corrector algorithm using the Alizadeh-Haeberly-Overton search direction, and proved the quadratic convergence of the algorithm under Assumptions A1 and A2, but the algorithm does not seem to be polynomial. Using the Nesterov-Todd search direction, Luo, et al. [7] investigated a symmetric primal-dual path following algorithm, which was proposed originally by Nesterov and Todd [11] and derived differently in [15]. They proved the superlinear convergence under Assumption A1, while the Assumption A3 is enforced by the algorithm. In a recent paper, Potra and Sheng [14] proved the superlinear convergence of the infeasible-interior-point algorithm of Kojima, et al. [3] and Potra and Sheng [13], under Assumption A1 and the following assumption

A4 $\lim_{k\to\infty} X^k S^k/\sqrt{X^k \bullet S^k} = 0$.

Among these four assumptions, Assumption A1 is standard. Assumptions A3 and A4 are similar, but Assumption A4 is a little bit weaker than Assumption A3. As shown by the example in [3], Assumption A3 or A4 is also needed to ensure the duality gap being reduced superlinearly. However, Assumption A3 or A4 could be enforced by proper parameter selection in the algorithm (cf. [7]).

In a very recent paper, Monteiro and Zhang [10] proposed a unified analysis for a class of long-step interior-point algorithms for SDP. In what follows, we will call the unified direction the Monteiro-Zhang direction. Using this direction, we propose a unified analysis for a class of infeasible-start predictor-corrector algorithms for SDP, which generalize the Mizuno-Todd-Ye predictor-corrector algorithm for linear programming. By extending the analysis of Potra and Sheng [13, 14], we show that this class of predictor-corrector algorithms shares similar global and local convergence properties with one of its members — the infeasible-interior-point algorithm proposed earlier by Kojima, et al. and Potra and Sheng. In particular we prove polynomial complexity for general problems and superlinear convergence for problems satisfying Assumptions A1 and A4.

Notation. The following notation is used throughout the paper. $\mathbb{R}^p$, $\mathbb{R}^p_+$, and $\mathbb{R}^p_{++}$ denote the p-dimensional Euclidean space, the nonnegative orthant of $\mathbb{R}^p$, and the positive orthant of $\mathbb{R}^p$, respectively. The set of all $p \times q$ matrices with real entries is denoted by $\mathbb{R}^{p\times q}$. The set of all $p \times p$ symmetric matrices is denoted by $\mathcal{S}^p$. For $Q \in \mathcal{S}^p$, $Q \succeq 0$ means Q is positive semidefinite and $Q \succ 0$ means Q is positive definite. The trace of a $p \times p$ matrix Q is denoted by $\mathrm{Tr}(Q) \equiv \sum_{i=1}^p [Q]_{ii}$. The eigenvalues of $Q \in \mathcal{S}^p$ are denoted by $\lambda_i(Q)$, $i = 1, \dots, p$, and its smallest and largest eigenvalues are denoted by $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$, respectively. Given P, $Q \in \mathbb{R}^{p\times q}$, the inner product between them is defined as $P \bullet Q \equiv \mathrm{Tr}(P^T Q)$. The Euclidean norm of a vector and the corresponding norm of a matrix are both denoted by $\|\cdot\|$; hence, $\|Q\| \equiv \max\{\|Qy\| : \|y\| = 1\}$. The Frobenius norm of a matrix Q is $\|Q\|_F \equiv \sqrt{\sum_{i=1}^p \sum_{j=1}^q [Q]_{ij}^2}$. $Q^k = O(1)$ means $\|Q\|^k$ is bounded, while $M^k = O(\nu_k)$ means $M^k/\nu_k = O(1)$.

2. The unified direction. Throughout this paper we assume that both (1.1) and (1.2) have finite solutions and their optimal values are equal. Under this assumption, X^* and (y^*, S^*) are solutions of (1.1) and (1.2) if and only if they are solutions of the following nonlinear system:

$$A_i \bullet X = b_i, \quad i = 1, \dots, m, \tag{2.1a}$$

(2.1b) $$\sum_{i=1}^{m} y_i A_i + S = C,$$

(2.1c) $$XS = 0, \quad X \succeq 0, \quad S \succeq 0.$$

We denote the feasible set of the problem (2.1) by

$$\mathcal{F} = \{(X, y, S) \in \mathcal{S}_+^n \times \mathbb{R}^m \times \mathcal{S}_+^n : (X, y, S) \text{ satisfies (2.1a) and (2.1b)}\}$$

and its solution set by $\mathcal{F}^*$, i.e.,

$$\mathcal{F}^* = \{(X, y, S) \in \mathcal{F} : X \bullet S = 0\}.$$

The residues of (2.1a) and (2.1b) are denoted by:

(2.2a) $$R_i = b_i - A_i \bullet X, \quad i = 1, \ldots, m,$$

(2.2b) $$R_d = C - \sum_{i=1}^{m} y_i A_i - S.$$

For any given $\epsilon > 0$ we define the set of *ϵ-approximate solutions* of (2.1) as

$$\mathcal{F}_\epsilon = \{Z = (X, y, S) \in \mathcal{S}_+^n \times \mathbb{R}^m \times \mathcal{S}_+^n : \\ X \bullet S \le \epsilon, \quad |R_i| \le \epsilon, \quad i = 1, \ldots, m, \quad \|R_d\| \le \epsilon\}.$$

In [18], Zhang defines the linear transformation

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

for a given invertible matrix P and observes that if P is invertible and M has a real spectrum, then

$$H_P(M) = \tau I \quad \text{iff} \quad M = \tau I.$$

Thus, the system $XS = 0$ is equivalent to the system $H_P(XS) = 0$, since XS is similar to the positive definite matrix $S^{\frac{1}{2}} X S^{\frac{1}{2}}$. Therefore, (2.1) is equivalent to

(2.3a) $$A_i \bullet X = b_i, \quad i = 1, \ldots, m,$$

(2.3b) $$\sum_{i=1}^{m} y_i A_i + S = C,$$

(2.3c) $$H_P(XS) = 0, \quad X \succeq 0, \quad S \succeq 0.$$

A perturbed Newton method applied to the system (2.3) leads to the following linear system:

(2.4a) $$H_P(US + XV) = \xi \tau I - H_P(XS),$$

(2.4b) $$A_i \bullet U = (1 - \xi) R_i, \quad i = 1, \ldots, m,$$

(2.4c) $$\sum_{i=1}^{m} w_i A_i + V = (1 - \xi) R_d$$

where $\tau > 0$ is a parameter, and $\xi \in [0, 1]$. Monteiro and Zhang [10] established the polynomiality of a long-step path following method based on the search direction (U, w, V) obtained from (2.4) when the scaling matrix P belongs to the class

$$\{W^{\frac{1}{2}} : W \in \mathcal{S}^n_{++} \text{ such that } WXS = SXW\}. \tag{2.5}$$

Being motivated by Monteiro and Zhang's work, in this paper we consider the following set of permissible matrices associated with $(X, S) \in \mathcal{S}^n_{++}$

$$\mathbb{P}(X, S) = \{P : P \in \mathbb{R}^{n\times n} \text{ is invertible and } PXSP^{-1} \in \mathcal{S}^n\}.$$

It is interesting to note that the members of $\mathbb{P}(X, S)$ could be nonsymmetric while those in (2.5) are only symmetric. We also note that Theorem 3.2 of [16] indicates that there is a unique symmetric solution $(U, w, V) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ to the linear system (2.4) for every $P \in \mathbb{P}(X, S)$. The following characterization of $\mathbb{P}(X, S)$ will be frequently used in our analysis.

LEMMA 2.1. *Let X, $S \in \mathcal{S}^n_{++}$. Then,*

$$\mathbb{P}(X, S) = \{P : X^{\frac{1}{2}}P^TPX^{\frac{1}{2}} \in \mathcal{S}^n_{++} \textit{ commutes with } X^{\frac{1}{2}}SX^{\frac{1}{2}}\} \tag{2.6}$$

$$= \{P : S^{-\frac{1}{2}}P^TPS^{-\frac{1}{2}} \in \mathcal{S}^n_{++} \textit{ commutes with } S^{\frac{1}{2}}XS^{\frac{1}{2}}\}. \tag{2.7}$$

Proof. The lemma follows from the fact that

$$PXSP^{-1} = (PXSP^{-1})^T \tag{2.8}$$

is equivalent to both

$$[X^{\frac{1}{2}}P^TPX^{\frac{1}{2}}][X^{\frac{1}{2}}SX^{\frac{1}{2}}] = [X^{\frac{1}{2}}SX^{\frac{1}{2}}][X^{\frac{1}{2}}P^TPX^{\frac{1}{2}}]$$

and

$$[S^{-\frac{1}{2}}P^TPS^{-\frac{1}{2}}][S^{\frac{1}{2}}XS^{\frac{1}{2}}] = [S^{\frac{1}{2}}XS^{\frac{1}{2}}][S^{-\frac{1}{2}}P^TPS^{-\frac{1}{2}}] .$$

□

We mention that the above lemma can also be derived from Proposition 3.4 and Theorem 3.1 of Monteiro and Zhang [10]. Given a matrix $P \in \mathbb{P}(X, S)$, let us define

$$J_x = X^{\frac{1}{2}}P^TPX^{\frac{1}{2}} \quad \text{and} \quad J_s = S^{-\frac{1}{2}}P^TPS^{-\frac{1}{2}}. \tag{2.9}$$

It follows from Lemma 2.1 that J_x, $J_s \in \mathcal{S}^n_{++}$ commute with $X^{\frac{1}{2}}SX^{\frac{1}{2}}$ and $S^{\frac{1}{2}}XS^{\frac{1}{2}}$, respectively.

LEMMA 2.2. *For any $P \in \mathbb{P}(X, S)$, there exist orthogonal matrices Q_x and Q_s such that*

$$P = Q_xJ_x^{\frac{1}{2}}X^{-\frac{1}{2}} = Q_sJ_s^{\frac{1}{2}}S^{\frac{1}{2}} .$$

Proof. In view of (2.9), we obtain

$$[J_x^{\frac{1}{2}}X^{-\frac{1}{2}}P^{-1}]^T[J_x^{\frac{1}{2}}X^{-\frac{1}{2}}P^{-1}] = I$$

and

$$[J_s^{\frac{1}{2}} S^{\frac{1}{2}} P^{-1}]^T [J_s^{\frac{1}{2}} S^{\frac{1}{2}} P^{-1}] = I \ .$$

The lemma is proved by taking $Q_x^T = J_x^{\frac{1}{2}} X^{-\frac{1}{2}} P^{-1}$ and $Q_s^T = J_s^{\frac{1}{2}} S^{\frac{1}{2}} P^{-1}$. □

Note that $J_x = J_s = I$ and $P = X^{-\frac{1}{2}}$ or $S^{\frac{1}{2}}$ define the directions formulated by Monteiro [9] which are particular cases of the direction originally proposed by Kojima, et al. [5]. The direction defined by $J_x = J_s = I$ and $P = S^{\frac{1}{2}}$ was derived independently by Helmberg, et al. [2]. Finally, the case $J_x = [X^{\frac{1}{2}} S X^{\frac{1}{2}}]^{\frac{1}{2}}$, or $J_s = [S^{\frac{1}{2}} X S^{\frac{1}{2}}]^{-\frac{1}{2}}$ corresponds to the Nesterov-Todd direction [11] (see [16] and [15]).

3. A class of predictor-corrector algorithms. In this section, we propose an infeasible-interior-point predictor-corrector algorithm for solving (2.1), which generalizes the interior-point method for linear programming proposed by Mizuno, et al. [8]. The algorithm performs in a neighborhood of the infeasible central path:

$$\begin{aligned} \mathcal{C}(\tau) = \{ Z = (X, y, S) \in \mathcal{S}_{++}^n \times \mathbb{R}^m \times \mathcal{S}_{++}^n : \\ XS = \tau I, \quad R_i = (\tau/\tau^0) R_i^0, \quad i = 1, \ldots, m, \quad R_d = (\tau/\tau^0) R_d^0 \}. \end{aligned}$$

The positive parameter τ is driven to zero and therefore the residues are also driven to zero at the same rate as τ. The iterates reside in the following neighborhood of the above central path:

$$\begin{aligned} \mathcal{N}(\gamma, \tau) &= \{ (X, y, S) \in \mathcal{S}_{++}^n \times \mathbb{R}^m \times \mathcal{S}_{++}^n : \| H_P(XS) - \tau I \|_F \le \gamma \tau \} \\ &= \{ (X, y, S) \in \mathcal{S}_{++}^n \times \mathbb{R}^m \times \mathcal{S}_{++}^n : \| P X S P^{-1} - \tau I \|_F \le \gamma \tau \} \\ &= \{ (X, y, S) \in \mathcal{S}_{++}^n \times \mathbb{R}^m \times \mathcal{S}_{++}^n : \left(\sum_{i=1}^m (\lambda_i(XS) - \tau)^2 \right)^{\frac{1}{2}} \le \gamma \tau \} \\ &= \{ (X, y, S) \in \mathcal{S}_{++}^n \times \mathbb{R}^m \times \mathcal{S}_{++}^n : \| X^{\frac{1}{2}} S X^{\frac{1}{2}} - \tau I \|_F \le \gamma \tau \} \end{aligned}$$

where γ is a constant such that $0 < \gamma < 1$. It is interesting to note that the neighborhood $\mathcal{N}(\gamma, \tau)$ is independent on the scaling matrix P. Throughout the paper we use the notation:

$$\kappa \equiv \sup \min\{\kappa_x, \ \kappa_s\} \tag{3.1}$$

where $\kappa_x = \|J_x\| \|J_x^{-1}\|$, $\kappa_s = \|J_s\| \|J_s^{-1}\|$, and the supremum is taken over all matrices P that are used in our algorithm. The κ in (3.1) is known for several important examples. If $P = X^{-\frac{1}{2}}$, then $J_x = I$ and $\kappa = 1$. If $P = S^{\frac{1}{2}}$, then $J_s = I$ and $\kappa = 1$. For the Nesterov-Todd direction, we have $J_x = (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{\frac{1}{2}}$ or $J_s = (S^{\frac{1}{2}} X S^{\frac{1}{2}})^{-\frac{1}{2}}$, and thus $\kappa \le \sqrt{(1+\beta)/(1-\beta)} < \sqrt{3}$ for $\alpha = 0.19$ and $\beta = 0.31$. More generally, if $J_x = (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{\sigma}$ or $J_s = (S^{\frac{1}{2}} X S^{\frac{1}{2}})^{-\sigma}$ for $\sigma \in \mathbb{R}$, then

$$\kappa \le ((1+\beta)(1-\beta))^{|\sigma|} \le 3^{|\sigma|} \quad \text{for} \quad \alpha = 0.25 * 3^{-|\sigma|/2}, \quad \beta = 0.41 * 3^{-|\sigma|/2}.$$

Throughout the paper, we assume that the spectral condition number of J_x or J_s is bounded and some upper bound on κ is known. The neighborhood size in our algorithm will depend on the quantity κ. First, let us choose two positive parameters α, β satisfying the inequalities

$$\frac{\sqrt{\kappa}\beta^2}{2(1-\sqrt{\kappa}\beta)^2} \le \alpha < \beta < \frac{\beta}{1-\sqrt{\kappa}\beta} < 1. \tag{3.2a}$$

(3.2b) $$\beta - \alpha \geq \Omega(1/\sqrt{\kappa}).$$

For example, $\alpha = 0.25/\sqrt{\kappa}$, $\beta = 0.41/\sqrt{\kappa}$ verify (3.2). We note that the parameter α defined in (3.2) always satisfies $\alpha\sqrt{\kappa} < 0.5$ and hence $\alpha < 0.5$, which will be frequently used in our analysis.

At a typical step of our algorithm we are given $(X, y, S) \in \mathcal{N}(\alpha, \tau)$ and obtain a predictor direction $(U, w, V) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ by solving the linear system (2.4) with $\xi = 0$. By taking a steplength θ along this direction, we obtain the points

$$X(\theta) = X + \theta U, \quad y(\theta) = y + \theta w, \quad S(\theta) = S + \theta V.$$

Theoretically we would like to compute the step length

(3.3)

$$\check{\theta} = \max\left\{\tilde{\theta} \in [0,1] : \left(\sum_{i=1}^{n}(\lambda_i(X(\theta)S(\theta)) - (1-\theta)\tau)^2\right)^{\frac{1}{2}} \leq \beta(1-\theta)\tau, \ \forall \theta \in [0, \tilde{\theta}]\right\}.$$

However this involves computing the root of a complicated nonlinear equation. In Lemma 3.5, we will show that

(3.4) $$\check{\theta} \geq \widehat{\theta}$$

where

(3.5) $$\widehat{\theta} \equiv \frac{2}{\sqrt{1 + 4\delta/(\beta - \alpha)} + 1},$$

and

(3.6) $$\delta \equiv \frac{1}{\tau}\|PUVP^{-1}\|_F.$$

Actually, $\widehat{\theta}$ is the positive root of $\delta\tau\theta^2 + (\beta - \alpha)\theta - (\beta - \alpha) = 0$. In what follows we assume that a steplength $\bar{\theta}$ satisfying

(3.7) $$\check{\theta} \geq \bar{\theta} \geq \widehat{\theta}$$

is computed, and we define

(3.8) $$\overline{X} = X + \bar{\theta}U, \qquad \bar{y} = y + \bar{\theta}w, \qquad \bar{S} = S + \bar{\theta}V, \qquad \tau^+ = (1 - \bar{\theta})\tau.$$

In case $\bar{\theta} = 1$ (which is very unlikely), it is easily seen that $(\overline{X}, \bar{y}, \bar{S}) \in \mathcal{F}^*$ and therefore the algorithm terminates with an exact solution. Now suppose that $\bar{\theta} < 1$. Then $\overline{X}$ and $\bar{S}$ are symmetric positive definite matrices since $\lambda_i(X(\theta)S(\theta)) \geq (1 - \beta)(1-\theta)\tau > 0$, $i = 1, \ldots, n$, $\forall \theta \in [0, \bar{\theta}]$. Therefore we can define the corrector direction $(\bar{U}, \bar{w}, \bar{V})$ as the unique symmetric solution of the following linear system (2.4) with $(X, y, S) = (\overline{X}, \bar{y}, \bar{S})$, $P = \bar{P} \in \mathbb{P}(\overline{X}, \ \bar{S})$, and $\tau = \tau^+$. By taking a unit steplength along this direction we obtain a new point

(3.9) $$X^+ = \bar{X} + \bar{U}, \quad y^+ = \bar{y} + \bar{w}, \quad S^+ = \bar{S} + \bar{V}.$$

Clearly

(3.10) $$R_i^+ = (1 - \bar{\theta})R_i, \ i = 1, \ldots, m, \quad R_d^+ = (1 - \bar{\theta})R_d.$$

Summarizing, we can formally define our algorithm as follows:

ALGORITHM 3.1. *Choose* $(X^0, y^0, S^0) \in \mathcal{N}(\alpha, \tau^0)$ *with* $\tau^0 = \mu^0 \equiv (X^0 \bullet S^0)/n$ *and set* $\psi^0 = 1$. *For* $k = 0, 1, \ldots,$ *do A1 through A5:*

A1 *Set* $X = X^k$, $y = y^k$, $S = S^k$ *and define* $R_d = C - \sum_{i=1}^m y_i A_i - S$, $R_i = b_i - A_i \bullet X$, $i = 1, \ldots, m$.

A2 *If* $\max\{X \bullet S, \|R_d\|, |R_i|, i = 1, \ldots, m\} \leq \epsilon$, *then report* $(X, y, S) \in \mathcal{F}_\epsilon$ *and terminate.*

A3 *Find the unique symmetric solution* U, w, V *of the linear system (2.4) with* $\xi = 0$, *define* $\bar{X}, \bar{y}, \bar{S}, \tau^+$ *as in (3.8), and set* $\psi^+ = (1-\bar{\theta})\psi$, *for a* $\bar{\theta}$ *satisfying (3.7). If* $\bar{\theta} = 1$, *then report* $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{F}^*$ *and terminate.*

A4 *Find the unique symmetric solution* $\bar{U}, \bar{w}, \bar{V}$ *of the linear system (2.4) with* $\xi = 1$, $(X, y, S) = (\overline{X}, \bar{y}, \bar{S})$, $P = \bar{P}$, $\tau = \tau^+$, *and define* X^+, y^+, S^+ *as in (3.9).*

A5 *Set* $X^{k+1} = X^+$, $S^{k+1} = s^+$, $\tau^{k+1} = \tau^+$, $\bar{\theta}^k = \bar{\theta}$, $\psi^{k+1} = \psi^+$, $R_d^k = R_d$, $R_i^k = R_i$, $i = 1, \ldots, m$.

The following results of Monteiro [9, Lemma 3.3] will be used in the analysis of Algorithm 3.1.

LEMMA 3.2. *Suppose that* $M \in \mathbb{R}^{p \times p}$ *is invertible. Then, for any* $E \in \mathcal{S}^p$, *we have*

$$\lambda_{\max}(E) \leq \frac{1}{2}\lambda_{\max}(MEM^{-1} + (MEM^{-1})^T), \tag{3.11}$$

$$\lambda_{\min}(E) \geq \frac{1}{2}\lambda_{\min}(MEM^{-1} + (MEM^{-1})^T), \tag{3.12}$$

$$\|E\|_F \leq \frac{1}{2}\|MEM^{-1} + (MEM^{-1})^T\|_F\ . \tag{3.13}$$

LEMMA 3.3. *Let* $(X, y, S) \in \mathcal{N}(\gamma, \tau)$ *for some* $\gamma \in [0, 1/\sqrt{\kappa})$ *and* $\tau > 0$. *Suppose that* $(D_x, \Delta y, D_s) \in \mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n}$ *is a solution of the linear system:*

$$H_P(D_x S + X D_s) = H, \tag{3.14a}$$

$$A_i \bullet D_x = 0, \quad i = 1, \ldots, m, \tag{3.14b}$$

$$\sum_{i=1}^m \Delta y_i A_i + D_s = 0, \tag{3.14c}$$

for some $H \in \mathbb{R}^{n \times n}$. *Then, the following three statements hold:*
(a) if $\kappa_x \leq \kappa_s$, *then*

$$\|X^{\frac{1}{2}} D_s X^{\frac{1}{2}}\|_F^2 + \tau^2 \|X^{-\frac{1}{2}} D_x X^{-\frac{1}{2}}\|_F^2 \leq \frac{\|H\|_F^2}{(1-\sqrt{\kappa}\gamma)^2};$$

(b) if $\kappa_s \leq \kappa_x$, *then*

$$\|S^{\frac{1}{2}} D_x S^{\frac{1}{2}}\|_F^2 + \tau^2 \|S^{-\frac{1}{2}} D_s S^{-\frac{1}{2}}\|_F^2 \leq \frac{\|H\|_F^2}{(1-\sqrt{\kappa}\gamma)^2};$$

(c) $\|PD_x D_s P^{-1}\|_F \leq \frac{\sqrt{\kappa}\|H\|_F^2}{2(1-\sqrt{\kappa}\gamma)^2\tau}$.

Proof. Let us prove (a). Suppose that $\kappa_x \leq \kappa_s$, then $\kappa_x \leq \kappa$. Writing

$$\begin{aligned} 2H &= 2H_P(XD_s + D_x S) \\ &= PXD_sP^{-1} + PD_xSP^{-1} + [PXD_sP^{-1} + PD_xSP^{-1}]^T \end{aligned}$$

$$\begin{aligned}
&= PXD_sP^{-1} + \tau PD_xX^{-1}P^{-1} + PD_xX^{-1}P^{-1}(PXSP^{-1} - \tau I) \\
&\quad + [PXD_sP^{-1} + \tau PD_xX^{-1}P^{-1} + PD_xX^{-1}P^{-1}(PXSP^{-1} - \tau I)]^T \\
&= B + B^T + PD_xX^{-1}P^{-1}(PXSP^{-1} - \tau I) \\
&\quad + [PD_xX^{-1}P^{-1}(PXSP^{-1} - \tau I)]^T
\end{aligned}$$

where

$$B = PXD_sP^{-1} + \tau PD_xX^{-1}P^{-1},$$

we have

$$\begin{aligned}
2\|H\|_F &\geq \|B + B^T\|_F - 2\|PD_xX^{-1}P^{-1}(PXSP^{-1} - \tau I)\|_F \\
&\geq \|B + B^T\|_F - 2\|PD_xX^{-1}P^{-1}\|_F\|PXSP^{-1} - \tau I\|_F \\
&\geq \|B + B^T\|_F - 2\gamma\tau\|PD_xX^{-1}P^{-1}\|_F.
\end{aligned}$$

According to Lemma 2.2, $P = Q_xJ_x^{\frac{1}{2}}X^{-\frac{1}{2}}$, so that we have

$$\begin{aligned}
\|PD_xX^{-1}P^{-1}\|_F &= \|Q_xJ_x^{\frac{1}{2}}X^{-\frac{1}{2}}D_xX^{-1}X^{\frac{1}{2}}J_x^{-\frac{1}{2}}Q_x^T\|_F \\
&= \|J_x^{\frac{1}{2}}X^{-\frac{1}{2}}D_xX^{-1}X^{\frac{1}{2}}J_x^{-\frac{1}{2}}\|_F \\
&\leq \|J_x^{\frac{1}{2}}\|\,\|J_x^{-\frac{1}{2}}\|\,\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F \\
&= \sqrt{\kappa_x}\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F \leq \sqrt{\kappa}\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F.
\end{aligned}$$

Using Lemma 3.2 with $M = PX^{\frac{1}{2}}$ and $E = X^{\frac{1}{2}}D_sX^{\frac{1}{2}} + \tau X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}$, we obtain

$$\begin{aligned}
\|B + B^T\|_F &\geq 2\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}} + \tau X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F \\
&= 2(\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F^2 + \tau^2\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F^2)^{\frac{1}{2}} \quad (\text{since } D_x \bullet D_s = 0).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|H\|_F &\geq \frac{1}{2}\|B + B^T\|_F - \gamma\tau\|PD_xX^{-1}P^{-1}\|_F \\
&\geq (\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F^2 + \tau^2\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F^2)^{\frac{1}{2}} - \sqrt{\kappa}\gamma\tau\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F \\
&\geq (\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F^2 + \tau^2\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F^2)^{\frac{1}{2}}(1 - \gamma\sqrt{\kappa}),
\end{aligned}$$

which proves (a). We can prove (b) similarly. To prove (c), we may assume $\kappa_x \leq \kappa_s$ without loss of generality, and deduce

$$\begin{aligned}
\|PD_xD_sP^{-1}\|_F &= \|Q_xJ_x^{\frac{1}{2}}X^{-\frac{1}{2}}D_xD_sX^{\frac{1}{2}}J_x^{-\frac{1}{2}}Q_x^T\|_F \\
&= \|J_x^{\frac{1}{2}}X^{-\frac{1}{2}}D_xD_sX^{\frac{1}{2}}J_x^{-\frac{1}{2}}\|_F \\
&\leq [\|J_x^{\frac{1}{2}}\|\,\|J_x^{-\frac{1}{2}}\|]\|X^{-\frac{1}{2}}D_xD_sX^{\frac{1}{2}}\|_F \\
&\leq \sqrt{\kappa}[\tau\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F][\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F]/\tau \\
&\leq \sqrt{\kappa}(\frac{1}{2})[\tau^2\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F^2 + \|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F^2]/\tau \\
&\leq \frac{\sqrt{\kappa}\|H\|_F^2}{2(1 - \gamma\sqrt{\kappa})^2\tau} \quad (\text{from (a)}).
\end{aligned}$$

□

The next corollary will be essentially used in the proof of global and local convergence properties of Algorithm 3.1.

COROLLARY 3.4. *Under the hypothesis of Lemma 3.2,*

$$\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F \le \frac{(1+\gamma)\|H\|_F}{1-\sqrt{\kappa}\gamma},$$
$$\tau\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F \le \frac{\|H\|_F}{(1-\sqrt{\kappa}\gamma)(1-\gamma)}.$$

Proof. If $\kappa_x \le \kappa_s$, then the results follow immediately from (a) of Lemma 3.2. Suppose $\kappa_x \ge \kappa_s$. Then from (b) of Lemma 3.2, we obtain

$$\|S^{\frac{1}{2}}D_xS^{\frac{1}{2}}\|_F \le \frac{\|H\|_F}{1-\sqrt{\kappa}\gamma}, \quad \tau\|S^{-\frac{1}{2}}D_sS^{-\frac{1}{2}}\|_F \le \frac{\|H\|_F}{1-\sqrt{\kappa}\gamma}.$$

Hence,

$$\begin{aligned}
\|X^{\frac{1}{2}}D_sX^{\frac{1}{2}}\|_F &= \|[X^{\frac{1}{2}}S^{\frac{1}{2}}]S^{-\frac{1}{2}}D_sS^{-\frac{1}{2}}[S^{\frac{1}{2}}X^{\frac{1}{2}}]\|_F \\
&\le \|X^{\frac{1}{2}}S^{\frac{1}{2}}\|^2\|S^{-\frac{1}{2}}D_sS^{-\frac{1}{2}}\|_F \\
&= \lambda_{\max}(X^{\frac{1}{2}}SX^{\frac{1}{2}})\|S^{-\frac{1}{2}}D_sS^{-\frac{1}{2}}\|_F \\
&\le \frac{(1+\gamma)\|H\|_F}{1-\sqrt{\kappa}\gamma}, \\
\tau\|X^{-\frac{1}{2}}D_xX^{-\frac{1}{2}}\|_F &= \tau\|[X^{-\frac{1}{2}}S^{-\frac{1}{2}}]S^{\frac{1}{2}}D_xS^{\frac{1}{2}}[S^{-\frac{1}{2}}X^{-\frac{1}{2}}]\|_F \\
&\le \tau\|X^{-\frac{1}{2}}S^{-\frac{1}{2}}\|^2\|S^{\frac{1}{2}}D_xS^{\frac{1}{2}}\|_F \\
&= \tau\|S^{\frac{1}{2}}D_xS^{\frac{1}{2}}\|_F/\lambda_{\min}(X^{\frac{1}{2}}SX^{\frac{1}{2}}) \\
&\le \frac{\|H\|_F}{(1-\sqrt{\kappa}\gamma)(1-\gamma)}\ .
\end{aligned}$$

□

The next lemma justifies our definition of steplength $\bar{\theta}$ in the algorithm.

LEMMA 3.5. *If* $(X, y, S) \in \mathcal{N}(\alpha, \tau)$, *then* $\breve{\theta} \in [0, 1]$ *defined by* (3.3) *satisfies* $\breve{\theta} \ge \hat{\theta}$ *where* $\widehat{\theta}$ *is given by* (3.5) *and* (3.6).

Proof. By definition, we have

$$\begin{aligned}
X(\theta)S(\theta) - (1-\theta)\tau I &= (X+\theta U)(S+\theta V) - (1-\theta)\tau I \\
&= (1-\theta)(XS-\tau I) + \theta XS + \theta(XV+US) + \theta^2 UV\ . \qquad (3.15)
\end{aligned}$$

If we set

$$R(\theta) \equiv P[X(\theta)S(\theta) - (1-\theta)\tau I]P^{-1},$$

then, in view of (3.15) and (2.4a) with $\xi = 0$, we obtain

$$\begin{aligned}
R(\theta) + R(\theta)^T = {}&2(1-\theta)(H_P(XS) - \tau I) + \theta[2H_P(XV+US)PX^{\frac{1}{2}} \\
&+ 2H_P(XS)] + \theta^2[PUVP^{-1} + (PUVP^{-1})^T]
\end{aligned}$$

$$= 2(1-\theta)(X^{\frac{1}{2}}SX^{\frac{1}{2}} - \tau I) + \theta^2[PUVP^{-1} + (PUVP^{-1})^T].$$

Therefore,

$$\begin{aligned}\frac{1}{2}\|R(\theta) + R(\theta)^T\|_F &\le (1-\theta)\|X^{\frac{1}{2}}SX^{\frac{1}{2}} - \tau I\|_F + \theta^2\|PUVP^{-1}\|_F \\ &\le \alpha\tau(1-\theta) + \theta^2\delta\tau. \end{aligned} \tag{3.16}$$

Hence, for any given parameter $\nu \in [0,1)$, we must have $X(\theta) \succ 0$, $S(\theta) \succ 0$ for all $\theta \in [0, \min(\widehat{\theta}, \nu))$. Otherwise, there must exist $0 \le \theta' \le \min(\widehat{\theta}, \nu) \le \nu < 1$ such that $X(\theta')S(\theta')$ is singular, which means

$$\lambda_{\min}(X(\theta')S(\theta') - (1-\theta')\tau) \le -(1-\theta')\tau\ . \tag{3.17}$$

However, using (3.12) with $M = P$ and $E = X(\theta')S(\theta') - (1-\theta')\tau$, we have

$$\begin{aligned}\lambda_{\min}(X(\theta')S(\theta') - (1-\theta')\tau) &\ge \frac{1}{2}\lambda_{\min}(R(\theta') + R(\theta')^T) \\ &\ge -\frac{1}{2}\|R(\theta') + R(\theta')^T\|_F \\ &\ge -[\alpha\tau(1-\theta') + (\theta')^2\delta\tau] \quad \text{(from (3.16))} \\ &\ge -\beta(1-\theta')\tau, \end{aligned}$$

which contradicts (3.17). Since $X(\theta) \succ 0$, its square root $X(\theta)^{\frac{1}{2}}$ exists and is uniquely defined. Applying (3.13) of Lemma 3.2 with $E = X(\theta)^{\frac{1}{2}}S(\theta)X(\theta)^{\frac{1}{2}} - (1-\theta)\tau I$, $M = PX(\theta)^{\frac{1}{2}}$, and noting that $R(\theta) = MEM^{-1}$, we obtain

$$\begin{aligned}\left(\sum_{i=1}^{n}(\lambda_i(X(\theta)S(\theta)) - (1-\theta)\tau)^2\right)^{\frac{1}{2}} &= \|X(\theta)^{\frac{1}{2}}S(\theta)X(\theta)^{\frac{1}{2}} - (1-\theta)\tau I\|_F \\ &\le \frac{1}{2}\|R(\theta) + R(\theta)^T\|_F \\ &\le \alpha\tau(1-\theta) + \theta^2\delta\tau \quad \text{(from (3.16))} \\ &\le \beta(1-\theta)\tau, \quad \text{for all } \theta \le \widehat{\theta}. \end{aligned}$$

Therefore, $(X(\theta), y(\theta), S(\theta)) \in \mathcal{N}(\beta, (1-\theta)\tau)$, for all $\theta \in [0, \min(\widehat{\theta}, \nu)]$. If $\widehat{\theta} < 1$ we can choose $\nu = \widehat{\theta}$, which gives $\bar{\theta} \ge \widehat{\theta}$. Finally, if $\widehat{\theta} = 1$, then $(X(\theta), y(\theta), S(\theta)) \in \mathcal{N}(\beta, (1-\theta)\tau)$, for all $\theta \in [0,1)$, which implies $X(1) \succeq 0$, $S(1) \succeq 0$ and $X(1)S(1) = 0$, and therefore $\bar{\theta} = 1 = \widehat{\theta}$. □

Before stating our main result let us note that the standard choice of starting points

$$X^0 = \rho_p I, \quad y^0 = 0, \quad S^0 = \rho_d I$$

is perfectly centered and satisfies $(X^0, y^0, S^0) \in \mathcal{N}(\alpha, \tau^0)$, as required in the algorithm.

We will see that if the problem has a solution, then for any $\epsilon > 0$ Algorithm 3.1 terminates in a finite number (say K_ϵ) of iterations. If $\epsilon = 0$ then the algorithm is likely to generate an infinite sequence. However it may happen that at a certain iteration (let us say at iteration K_0), we have $\bar{\theta} = 1$, which implies that an exact solution is obtained, and therefore the algorithm terminates at iteration K_0. If this (unlikely) phenomenon does not happen, we set $K_0 = \infty$.

THEOREM 3.6. *For any integer* $0 \le k < K_0$, *Algorithm 3.1 defines a triple*

$$(X^k, y^k, S^k) \in \mathcal{N}(\alpha, \tau^k) \tag{3.18}$$

and the corresponding residuals satisfy

$$R_d^k = \psi^k R_d^0, \quad R_i^k = \psi^k R_i^0, \quad i = 1, \dots, m, \tag{3.19}$$

$$\tau^k = \psi^k \tau^0, \tag{3.20}$$

$$(1-\alpha)\tau^k \le \mu^k \equiv (X^k \bullet S^k)/n \le (1+\alpha)\tau^k \tag{3.21}$$

where

$$\psi^0 = 1, \quad \psi^k = \prod_{j=0}^{k-1}(1-\bar{\theta}^j),$$

and $\bar{\theta}^j$ *is defined by (3.7).*

Proof. The proof is by induction. For $k = 0$, (3.18)–(3.21) are clearly satisfied. Suppose they are satisfied for some $k \ge 0$. As in Algorithm 3.1 we will omit the index k. Therefore we can write

$$\begin{aligned} &(X, y, S) \in \mathcal{N}(\alpha, \tau), \\ R_d = \psi R_d^0, &\quad R_i = \psi R_i^0, i = 1, \dots, m, \\ \tau = \psi\tau^0, &\quad (1-\alpha)\tau \le \mu \le (1+\alpha)\tau. \end{aligned}$$

The results in (3.19) and (3.20) follow immediately from (3.8) and (3.10). From the fact that $(\overline{U}, \bar{w}, \overline{V})$ is the solution to the system (2.4) with $\xi = 1$, $(X, y, S) = (\overline{X}, \bar{y}, \overline{S})$, $P = \bar{P}$, $\tau = \tau^+$, together with (3.9), we obtain

$$\begin{aligned} X^+S^+ - \tau^+ I &= (\overline{X} + \overline{U})(\overline{S} + \bar{V}) - \tau^+ I \\ &= \overline{XS} - (1-\bar{\theta})\tau I + \overline{XU} + \overline{US} + \overline{UV}. \end{aligned} \tag{3.22}$$

Define

$$B = \overline{P}(X^+S^+ - \tau^+ I)\overline{P}^{-1}.$$

Then, (3.22) implies

$$\begin{aligned} B + B^T &= 2[H_{\overline{P}}(\overline{XS}) - (1-\bar{\theta})\tau I] \\ &\quad + [2H_{\overline{P}}(\overline{US} + \overline{XV})] + [\overline{PUVP}^{-1} + (\overline{PUVP}^{-1})^T] \\ &= [\overline{PUVP}^{-1} + (\overline{PUVP}^{-1})^T]. \end{aligned} \tag{3.23}$$

Since $k < K_0$, we see that $\bar{\theta} < 1$ and that $\overline{X} \succ 0$, $\overline{S} \succ 0$. Applying (c) of Lemma 3.3, we deduce

$$\begin{aligned} \|\overline{PUVP}^{-1}\|_F &\le \frac{\sqrt{\kappa}\|H_P(\overline{XS}) - (1-\bar{\theta})\tau\|_F^2}{2(1-\sqrt{\kappa}\beta)^2(1-\bar{\theta})\tau} \\ &\le \frac{\sqrt{\kappa}\beta^2(1-\bar{\theta})\tau}{2(1-\sqrt{\kappa}\beta)^2} \end{aligned}$$

$$\leq \alpha(1-\bar{\theta})\tau = \alpha\tau^+ \qquad \text{(from (3.2))}. \tag{3.24}$$

Without loss of generality, we may assume $\kappa_{\overline{x}} \leq \kappa_{\overline{s}}$. Hence by applying (a) of Lemma 3.3 with $H = (1-\bar{\theta})\tau I - H_P(\overline{XS})$, we have

$$\|\overline{X}^{-\frac{1}{2}}\overline{UX}^{-\frac{1}{2}}\|_F \leq \frac{\|H_P(\overline{XS}) - (1-\bar{\theta})\tau I\|_F}{(1-\sqrt{\kappa}\beta)(1-\bar{\theta})\tau} \leq \frac{\beta}{1-\sqrt{\kappa}\beta} < 1,$$

which implies that $I + \overline{X}^{-\frac{1}{2}}\overline{UX}^{-\frac{1}{2}} \succ 0$, and therefore, $X^+ = \overline{X} + \overline{U} \succ 0$. Thus $(X^+)^{\frac{1}{2}}$ exists. Using (3.22), (3.23), applying Lemma 3.2 with $E = (X^+)^{\frac{1}{2}}S^+(X^+)^{\frac{1}{2}} - \tau^+ I$, $M = \bar{P}(X^+)^{\frac{1}{2}}$, and noting that $B = MEM^{-1}$, we have

$$\begin{aligned}
\|(X^+)^{\frac{1}{2}}S^+(X^+)^{\frac{1}{2}} - \tau^+ I\|_F &\leq \frac{1}{2}\|B + B^T\|_F \\
&= \frac{1}{2}\|\overline{PUVP}^{-1} + [\overline{PUVP}^{-1}]^T\|_F \qquad \text{(from (3.23))} \\
&\leq \|\overline{PUVP}^{-1}\|_F \\
&\leq \alpha\tau^+ \qquad \text{(from (3.24))}.
\end{aligned} \tag{3.25}$$

The above inequality implies that

$$\lambda_{\min}((X^+)^{\frac{1}{2}}S^+(X^+)^{\frac{1}{2}}) \geq (1-\alpha)\tau^+ > 0.$$

Hence $(X^+)^{\frac{1}{2}}S^+(X^+)^{\frac{1}{2}} \succ 0$, which gives $S^+ \succ 0$. In view of (3.25), this shows that (3.18) holds for $k+1$. Finally, (3.21) is an immediate consequence of (3.18). □

4. Global convergence and iteration complexity. In this section we assume that $\mathcal{F}^*$ is nonempty. Under this assumption we will prove that Algorithm 3.1, with $\epsilon = 0$, is globally convergent in the sense that

$$\lim_{k\to\infty} \mu^k = 0, \quad \lim_{k\to\infty} R_d^k = 0, \quad \lim_{k\to\infty} R_i^k = 0, \ i = 1, \ldots, m.$$

In the sequel, we will frequently use the following well-known inequality:

$$\|M_1 M_2\|_F \leq \min\{\|M_1\|\,\|M_2\|_F\,,\ \|M_1\|_F\|M_2\|\}, \quad \text{for any } M_1,\, M_2 \in \mathbb{R}^{n\times n}. \tag{4.1}$$

LEMMA 4.1 (Potra-Sheng[13], Lemma 3.2). *Assume that $\mathcal{F}^*$ is nonempty. Then for any $(X^*, y^*, S^*) \in \mathcal{F}^*$ and $(X, y, S) \in \mathcal{N}(\alpha, \tau)$ we have*

$$X \bullet S^0 + X^0 \bullet S \leq (2 + \alpha + \zeta)n\tau^0\,, \tag{4.2a}$$

$$X \bullet S^* + X^* \bullet S \leq ((1+\alpha+\psi)/(1-\psi) + \zeta)n\tau \tag{4.2b}$$

where

$$\zeta = (X^0 \bullet S^* + X^* \bullet S^0)/(X^0 \bullet S^0). \tag{4.3}$$

Lemma 4.1 shows that the pair (X^k, S^k) generated by Algorithm 3.1 is bounded. More precisely, we have the following corollary, which can easily be deduced from Lemma 4.1 and Theorem 3.6.

COROLLARY 4.2 (Potra-Sheng [13], Corollary 3.3). *Under the hypothesis of Lemma 4.1 we have*

$$\|X^{\frac{1}{2}}(S^0)^{\frac{1}{2}}\|_F \leq \sqrt{(2+\alpha+\zeta)n\tau^0}, \tag{4.4}$$

$$\|S^{\frac{1}{2}}(X^0)^{\frac{1}{2}}\|_F \le \sqrt{(2+\alpha+\zeta)n\tau^0}, \tag{4.5}$$

$$\|X^{\frac{1}{2}}\|_F \le \|(S^0)^{-\frac{1}{2}}\|\sqrt{(2+\alpha+\zeta)n\tau^0}, \tag{4.6}$$

$$\|S^{\frac{1}{2}}\|_F \le \|(X^0)^{-\frac{1}{2}}\|\sqrt{(2+\alpha+\zeta)n\tau^0}, \tag{4.7}$$

$$\|X^{\frac{1}{2}}S^{\frac{1}{2}}\|^2 = \|X^{\frac{1}{2}}SX^{\frac{1}{2}}\| \le (1+\alpha)\tau, \tag{4.8}$$

$$\|X^{-\frac{1}{2}}S^{-\frac{1}{2}}\|^2 = \|X^{-\frac{1}{2}}S^{-1}X^{-\frac{1}{2}}\| \le \frac{1}{(1-\alpha)\tau}. \tag{4.9}$$

LEMMA 4.3. *Let* $(X, y, S) \in \mathcal{N}(\alpha, \tau)$, $P \in \mathbb{P}(X, S)$. *Then,*
(i) $\max\{\kappa_x, \kappa_s\} \le 3\kappa$;
(ii) *for any* $M \in \mathbb{R}^{n\times n}$, $\|H_P(M)\|_F \le \|PMP^{-1}\|_F \le \sqrt{3\kappa}\|X^{-\frac{1}{2}}MX^{\frac{1}{2}}\|_F$.

Proof. In view of (2.9), we have

$$J_x = [X^{\frac{1}{2}}S^{\frac{1}{2}}]J_s[X^{\frac{1}{2}}S^{\frac{1}{2}}]^T \text{ and } J_x^{-1} = [(X^{\frac{1}{2}}S^{\frac{1}{2}})^{-1}]^T J_s^{-1}[X^{\frac{1}{2}}S^{\frac{1}{2}}]^{-1}.$$

Therefore we obtain,

$$\|J_x\| \le \|J_s\|\,\|X^{\frac{1}{2}}S^{\frac{1}{2}}\|^2 = \|J_s\|\lambda_{\max}(X^{\frac{1}{2}}SX^{\frac{1}{2}}) \le (1+\alpha)\tau\|J_s\|,$$
$$\|J_x^{-1}\| \le \|J_s^{-1}\|\,\|(X^{\frac{1}{2}}S^{\frac{1}{2}})^{-1}\|^2 = \|J_s^{-1}\|/\lambda_{\min}(X^{\frac{1}{2}}SX^{\frac{1}{2}}) \le \|J_s^{-1}\|/((1-\alpha)\tau).$$

Hence,

$$k_x = \|J_x\|\,\|J_x^{-1}\| \le \frac{1+\alpha}{1-\alpha}\|J_s\|\,\|J_s^{-1}\| \le 3k_s. \tag{4.10}$$

Similarly we can prove $k_s \le 3k_x$. Thus, (i) follows immediately. (ii) can be proved by noting that

$$\begin{aligned}\|H_P(M)\|_F &\le \|PMP^{-1}\|_F \\ &= \|Q_x J_x^{\frac{1}{2}} X^{-\frac{1}{2}} M X^{\frac{1}{2}} J_x^{-\frac{1}{2}} Q_x^T\|_F \\ &= \|J_x^{\frac{1}{2}} X^{-\frac{1}{2}} M X^{\frac{1}{2}} J_x^{-\frac{1}{2}}\|_F \\ &\le \sqrt{\kappa_x}\|X^{-\frac{1}{2}}MX^{\frac{1}{2}}\|_F \\ &\le \sqrt{3\kappa}\|X^{-\frac{1}{2}}MX^{\frac{1}{2}}\|_F.\end{aligned}$$

□

LEMMA 4.4. *Suppose* $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \mathcal{F}$, *and denote*

$$T = \psi[X^{\frac{1}{2}}(S^0 - \widetilde{S})X^{\frac{1}{2}} + X^{-\frac{1}{2}}(X^0 - \widetilde{X})SX^{\frac{1}{2}}] - X^{\frac{1}{2}}SX^{\frac{1}{2}},$$

$$T_x = \psi X^{-\frac{1}{2}}(X^0 - \widetilde{X})X^{-\frac{1}{2}},$$

$$T_s = \psi X^{\frac{1}{2}}(S^0 - \widetilde{S})X^{\frac{1}{2}}.$$

Then the quantity δ *defined by (3.6) satisfies the inequality:*

$$\delta \le \frac{\sqrt{3\kappa}}{\tau}\left(\tau\|T_x\|_F + 4\sqrt{3\kappa}\|T\|_F\right)\left(\|T_s\|_F + 3\sqrt{3\kappa}\|T\|_F\right).$$

Proof. It is easily seen that $(U + \psi(X^0 - \widetilde{X}),\ w + \psi(y^0 - \bar{y}),\ V + \psi(S^0 - \widetilde{S}))$ satisfies (3.14) with $H = H_P(M)$ where

$$M = \psi(X(S^0 - \widetilde{S}) + (X^0 - \widetilde{X})S) - XS.$$

Hence, according to Corollary 3.4, we have

$$\tau||X^{-\frac{1}{2}}(U + \psi(X^0 - \widetilde{X}))X^{-\frac{1}{2}}||_F \leq \frac{||H||_F}{(1-\sqrt{\kappa}\alpha)(1-\alpha)} \leq 4||H||_F,$$

$$||X^{\frac{1}{2}}(V + \psi(S^0 - \widetilde{S}))X^{\frac{1}{2}}||_F \leq \frac{(1+\alpha)||H||_F}{1-\sqrt{\kappa}\alpha} \leq 3||H||_F.$$

By (ii) of Lemma 4.3, we have

$$||H||_F = ||H_P(M)||_F \leq \sqrt{3\kappa}||X^{-\frac{1}{2}}MX^{\frac{1}{2}}||_F = \sqrt{3\kappa}||T||_F.$$

Therefore

$$\tau||X^{-\frac{1}{2}}UX^{-\frac{1}{2}}||_F \leq \tau||T_x||_F + 4||H||_F \leq \tau||T_x||_F + 4\sqrt{3\kappa}||T||_F,$$

$$||X^{\frac{1}{2}}VX^{\frac{1}{2}}||_F \leq ||T_s||_F + 3||H||_F \leq ||T_s||_F + 3\sqrt{3\kappa}||T||_F.$$

Again, using (ii) of Lemma 4.3, we deduce

$$\begin{aligned}\delta &= ||PUVP^{-1}||/\tau \leq \sqrt{3\kappa}||X^{-\frac{1}{2}}UVX^{\frac{1}{2}}||_F/\tau \\ &\leq \sqrt{3\kappa}[\tau||X^{-\frac{1}{2}}UX^{-\frac{1}{2}}||_F][||X^{\frac{1}{2}}VX^{\frac{1}{2}}||_F]/\tau^2 \\ &\leq \frac{\sqrt{3\kappa}}{\tau^2}\left(\tau||T_x||_F + 4\sqrt{3\kappa}||T||_F\right)\left(||T_s||_F + 3\sqrt{3\kappa}||T||_F\right).\end{aligned}$$

□

LEMMA 4.5. *Under the hypothesis of Lemma 4.4 we have*

$$\delta < 3.5\kappa^{1.5}\left(42.6(2.5+\zeta)nd_0 + 7.8\sqrt{n}\right)^2, \tag{4.11}$$

where

$$d_0 = \max(||(X^0)^{-\frac{1}{2}}(X^0 - \widetilde{X})(X^0)^{-\frac{1}{2}}||_F,\ ||(S^0)^{-\frac{1}{2}}(S^0 - \widetilde{S})(S^0)^{-\frac{1}{2}}||_F).$$

Proof. Using the notation of Lemma 4.4, and Corollary 4.2, we can write

$$\begin{aligned}||T||_F &\leq \psi||X^{\frac{1}{2}}(S^0 - \widetilde{S})X^{\frac{1}{2}}||_F + \psi||X^{-\frac{1}{2}}(X^0 - \widetilde{X})SX^{\frac{1}{2}}||_F + ||X^{\frac{1}{2}}SX^{\frac{1}{2}}||_F \\ &= \psi||X^{\frac{1}{2}}(S^0)^{\frac{1}{2}}(S^0)^{-\frac{1}{2}}(S^0 - \widetilde{S})(S^0)^{-\frac{1}{2}}(S^0)^{\frac{1}{2}}X^{\frac{1}{2}}||_F \\ &\quad + \psi||X^{-\frac{1}{2}}S^{-\frac{1}{2}}S^{\frac{1}{2}}(X^0)^{\frac{1}{2}}(X^0)^{-\frac{1}{2}}(X^0 - \widetilde{X})(X^0)^{-\frac{1}{2}}(X^0)^{\frac{1}{2}}S^{\frac{1}{2}}S^{\frac{1}{2}}X^{\frac{1}{2}}||_F \\ &\quad + ||X^{\frac{1}{2}}SX^{\frac{1}{2}}||_F \\ &\leq \psi||X^{\frac{1}{2}}(S^0)^{\frac{1}{2}}||^2||(S^0)^{-\frac{1}{2}}(S^0 - \widetilde{S})(S^0)^{-\frac{1}{2}}||_F \\ &\quad + \psi||X^{-\frac{1}{2}}S^{-\frac{1}{2}}||\ ||S^{\frac{1}{2}}(X^0)^{\frac{1}{2}}||^2||X^{\frac{1}{2}}S^{\frac{1}{2}}||\ ||(X^0)^{-\frac{1}{2}}(X^0 - \widetilde{X})(X^0)^{-\frac{1}{2}}||_F\end{aligned}$$

$$\begin{aligned}
&+\sqrt{n}\|X^{\frac{1}{2}}SX^{\frac{1}{2}}\| \\
&\leq \psi(2+\alpha+\zeta)n\tau^0 d_0(1+\frac{1+\alpha}{1-\alpha}) + \sqrt{n}(1+\alpha)\tau \\
&= \tau[2(2+\alpha+\zeta)nd_0/(1-\alpha) + (1+\alpha)\sqrt{n}] \\
&< \tau[2(5+2\zeta)nd_0 + 1.5\sqrt{n}].
\end{aligned}$$

Also,

$$\begin{aligned}
\|T_x\|_F &\leq \psi\|X^{-\frac{1}{2}}S^{-\frac{1}{2}}S^{\frac{1}{2}}(X^0)^{\frac{1}{2}}(X^0)^{-\frac{1}{2}}(X^0-\widetilde{X})(X^0)^{-\frac{1}{2}}(X^0)^{\frac{1}{2}}S^{\frac{1}{2}}S^{-\frac{1}{2}}X^{-\frac{1}{2}}\|_F \\
&\leq \psi\|X^{-\frac{1}{2}}S^{-\frac{1}{2}}\|^2\|S^{\frac{1}{2}}(X^0)^{\frac{1}{2}}\|^2 d_0 \\
&\leq \frac{\psi(2+\alpha+\zeta)n\tau^0}{\tau(1-\alpha)}d_0 \\
&= \frac{(2+\alpha+\zeta)n}{1-\alpha}d_0 \leq (5+2\zeta)nd_0,
\end{aligned}$$

and

$$\begin{aligned}
\|T_s\|_F &\leq \psi\|X^{\frac{1}{2}}(S^0)^{\frac{1}{2}}(S^0)^{-\frac{1}{2}}(S^0-\widetilde{S})(S^0)^{-\frac{1}{2}}(S^0)^{\frac{1}{2}}X^{\frac{1}{2}}\| \\
&\leq \psi\|X^{\frac{1}{2}}(S^0)^{\frac{1}{2}}\|^2 d_0 \\
&\leq \psi(2+\alpha+\zeta)n\tau^0 d_0 = (2+\alpha+\zeta)n\tau d_0 \\
&\leq (2.5+\zeta)n\tau d_0.
\end{aligned}$$

Then (4.11) follows from Lemma 4.4. □

According to Lemma 3.5 and Lemma 4.5, it follows that if $\mathcal{F}^*$ is not empty, then the step length $\bar{\theta}^k$ defined by (3.7) is bounded away from 0. This implies global convergence as shown in the following theorem.

THEOREM 4.6. *If $\mathcal{F}^*$ is not empty, then Algorithm 3.1 is globally convergent at a linear rate. Moreover, the iteration sequence (X^k, y^k, S^k) is bounded and every accumulation point of (X^k, y^k, S^k) belongs to $\mathcal{F}^*$ (i.e., is a primal dual optimal solution of the SDP problem).*

Using Lemma 4.5, we can easily deduce the following result.

THEOREM 4.7. *Suppose that $\mathcal{F}^*$ is nonempty and that the starting point is chosen such that there is a constant γ^* independent of n satisfying the inequality*

$$\begin{aligned}
&(2.5+\zeta)\max(\|(X^0)^{-\frac{1}{2}}(X^0-\widetilde{X})(X^0)^{-\frac{1}{2}}\|_F,\ \|(S^0)^{-\frac{1}{2}}(S^0-\widetilde{S})(S^0)^{-\frac{1}{2}}\|_F) \\
&\leq n^{-\frac{1}{2}}\gamma^*,
\end{aligned}$$

for some $(\widetilde{X}, \widetilde{y}, \widetilde{S}) \in \mathcal{F}$. Then Algorithm 3.1 terminates in at most $O(\kappa\sqrt{n}\ln(\epsilon_0/\epsilon))$ iterations, where

$$\epsilon_0 = \max\{X^0 \bullet S^0, \|R_d^0\|, |R_i^0|, i = 1, \ldots, m\}. \tag{4.12}$$

COROLLARY 4.8. *Suppose that $\mathcal{F}^*$ is nonempty and that the starting point is feasible, i.e., $(X^0, y^0, S^0) \in \mathcal{F}$. Then Algorithm 3.1 terminates in at most $O(\kappa\sqrt{n}\ln(\epsilon_0/\epsilon))$ iterations, where ϵ_0 is defined by (4.12).*

THEOREM 4.9. *Suppose $X^0 = S^0 = \rho I$, where $\rho > 0$ is a constant such that $||X^*|| \le \rho$, $||S^*|| \le \rho$ for some $(X^*, y^*, S^*) \in \mathcal{F}^*$. Then the step length $\bar{\theta}_k$ defined by (3.7) satisfies the inequality*

$$\bar{\theta}_k > \frac{1}{95\kappa n/\sqrt{\omega}+1}, \tag{4.13}$$

where ω is a constant such that $\beta - \alpha \ge \omega/\sqrt{\kappa}$.

Proof. According to Lemma 4.1, we have

$$\rho(\mathrm{Tr}(X) + \mathrm{Tr}(S)) \le (2+\alpha+\zeta)n\tau^0 = (2+\alpha+\zeta)n\rho^2,$$

i.e.,

$$\sum_{i=1}^{n}(\lambda_i(X) + \lambda_i(S)) \le (2+\alpha+\zeta)\rho n.$$

Since $X^* \bullet S^* = 0$ we get the relation

$$\zeta = (S^* \bullet X^0 + X^* \bullet S^0)/(X^0 \bullet S^0) = (\mathrm{Tr}(X^*) + \mathrm{Tr}(S^*))/(n\rho) \le 1,$$

which implies

$$||X^{\frac{1}{2}}||_F^2 + ||S^{\frac{1}{2}}||_F^2 = \sum_{i=1}^{n}(\lambda_i(X) + \lambda_i(S)) \le (3+\alpha)\rho n. \tag{4.14}$$

It is easily seen that

$$||X^0 - X^*|| \le \rho \quad \text{and} \quad ||S^0 - S^*|| \le \rho. \tag{4.15}$$

Applying (4.14)-(4.15), Corollary 4.2, and Lemma 4.4 with $(\widetilde{X}, \widetilde{y}, \widetilde{S}) = (X^*, y^*, S^*)$, we have

$$||X^{\frac{1}{2}}(S^0 - S^*)X^{\frac{1}{2}}||_F \le ||X^{\frac{1}{2}}||_F^2 ||S^0 - S^*|| \le 3.5\rho^2 n,$$

$$\begin{aligned} ||X^{-\frac{1}{2}}(X^0 - X^*)SX^{\frac{1}{2}}||_F &= ||(X^{-\frac{1}{2}}S^{-\frac{1}{2}})S^{\frac{1}{2}}(X^0 - X^*)S^{\frac{1}{2}}(S^{\frac{1}{2}}X^{\frac{1}{2}})||_F \\ &\le ||X^{-\frac{1}{2}}S^{-\frac{1}{2}}||\,||S^{\frac{1}{2}}X^{\frac{1}{2}}||\,||S^{\frac{1}{2}}||_F^2||X^0 - X^*|| \\ &= ||X^{-\frac{1}{2}}S^{-1}X^{-\frac{1}{2}}||^{\frac{1}{2}}||X^{\frac{1}{2}}SX^{\frac{1}{2}}||^{\frac{1}{2}}||S^{\frac{1}{2}}||_F^2||X^0 - X^*|| \\ &\le 6.1\rho^2 n. \end{aligned} \tag{4.16}$$

In view of (4.16)-(4.17) and Corollary 4.2, we obtain

$$\begin{aligned} ||T||_F &\le \psi||X^{\frac{1}{2}}(S^0-S^*)X^{\frac{1}{2}}||_F + \psi||X^{-\frac{1}{2}}(X^0-X^*)SX^{\frac{1}{2}}||_F + ||X^{\frac{1}{2}}SX^{\frac{1}{2}}||_F \\ &< \tau[9.6n + 1.5\sqrt{n}] \le 11.1\tau n, \end{aligned} \tag{4.17}$$

$$\begin{aligned} \tau||T_x||_F &= \tau\psi||X^{-\frac{1}{2}}(X^0 - X^*)X^{-\frac{1}{2}}||_F \\ &= \tau\psi||(X^{-\frac{1}{2}}S^{-\frac{1}{2}})S^{\frac{1}{2}}(X^0 - X^*)S^{\frac{1}{2}}(S^{-\frac{1}{2}}X^{-\frac{1}{2}})||_F \\ &\le \tau\psi||S^{\frac{1}{2}}||_F^2||X^{-\frac{1}{2}}S^{-\frac{1}{2}}||^2||X^0 - X^*|| \le 7\tau n, \end{aligned} \tag{4.18}$$

$$||T_s||_F = \psi||X^{\frac{1}{2}}(S^0 - S^*)X^{\frac{1}{2}}||_F \le \psi||X^{\frac{1}{2}}||_F^2||S^0 - S^*|| \le 3.5\tau n. \tag{4.19}$$

Therefore,

$$\delta \leq \frac{\sqrt{3\kappa}}{\tau^2}\left(\tau\|T_x\|_F + 4\sqrt{3\kappa}\|T\|_F\right)\left(\|T_s\|_F + 3\sqrt{3\kappa}\|T\|_F\right) < 8890.6\kappa^{1.5}n^2.$$

Consequently,

$$\bar{\theta} \geq \hat{\theta} = \frac{2}{\sqrt{1+4\delta/(\beta-\alpha)}+1} > \frac{1}{\sqrt{\delta/(\beta-\alpha)}+1} > \frac{1}{95\kappa n/\sqrt{\omega}+1}.$$

□

In the following corollary we summarize the complexity results for standard starting point of the form $X^0 = S^0 = \rho I$.

COROLLARY 4.10. *Assume that in Algorithm 3.1 we choose a starting point of the form* $X^0 = S^0 = \rho I$, *where* $\rho > 0$ *is a constant. Let* ϵ_0 *be given by (4.12) and let* $\epsilon > 0$ *be arbitrary. Then the following statements hold:*

(i) If $\mathcal{F}^* \neq \emptyset$, *then the algorithm terminates with an* ϵ*-approximate solution* $(X^k, y^k, S^k) \in \mathcal{F}_\epsilon$ *in a finite number of steps* $k = K_\epsilon < \infty$.

(ii) If $\rho \geq \max\{\|X^*\|, \|S^*\|\}$, *for some* $(X^*, y^*, S^*) \in \mathcal{F}^*$ *then* $K_\epsilon = O(\kappa n \ln(\epsilon_0/\epsilon))$.

(iii) For any choice of $\rho > 0$ *there is an index* $k = \widehat{K}_\epsilon = O(\kappa n \ln(\epsilon_0/\epsilon))$ *such that either*

(iiia) $(X^k, y^k, S^k) \in \mathcal{F}_\epsilon$,

or,

(iiib) $\bar{\theta} \leq 1/(95\kappa n/\sqrt{\omega}+1)$, *and in the latter case there is no solution* $(X^*, y^*, S^*) \in \mathcal{F}^*$ *with* $\rho \geq \max\{\|X^*\|, \|S^*\|\}$.

5. Local convergence. In this section we will investigate the asymptotic behavior of Algorithm 3.1. Throughout the paper we assume that the SDP problem has a strictly complementary solution (X^*, y^*, S^*) of (2.1), i.e., $X^* + S^* \succ 0$. For a strict complementarity solution (X^*, S^*), there exists an orthogonal matrix $Q = (q_1, \ldots, q_n)$ whose columns $q_1, \ldots, q_n$ are common eigenvectors of X^* and S^*, and define

$$\mathbb{B} = \{i : q_i^T X^* q_i > 0\}, \qquad \mathbb{N} = \{i : q_i^T S^* q_i > 0\}.$$

It is easily seen that $\mathbb{B} \cup \mathbb{N} = \{1, 2, \ldots, n\}$. For simplicity, let us assume that

$$Q^T X^* Q = \begin{pmatrix} \Lambda_B & 0 \\ 0 & 0 \end{pmatrix}, \quad Q^T S^* Q = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_N \end{pmatrix},$$

where Λ_B and Λ_N are diagonal matrices. Here and in the sequel, if we write a matrix M in the block form

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$

then we assume that the dimensions of M_{11} and M_{22} are $|\mathbb{B}| \times |\mathbb{B}|$ and $|\mathbb{N}| \times |\mathbb{N}|$, respectively.

Lemma 4.4 of Potra and Sheng [13] indicates that we can write

$$(5.1)\quad Q^T (X^k)^{\frac{1}{2}} Q = \begin{pmatrix} O(1) & O(\sqrt{\tau^k}) \\ O(\sqrt{\tau^k}) & O(\sqrt{\tau^k}) \end{pmatrix}, \quad Q^T (X^k)^{-\frac{1}{2}} Q = \begin{pmatrix} O(1) & O(1) \\ O(1) & O(1/\sqrt{\tau^k}) \end{pmatrix},$$

$$(5.2)\quad Q^T(S^k)^{\frac{1}{2}}Q = \begin{pmatrix} O(\sqrt{\tau^k}) & O(\sqrt{\tau^k}) \\ O(\sqrt{\tau^k}) & O(1) \end{pmatrix}, \quad Q^T(S^k)^{-\frac{1}{2}}Q = \begin{pmatrix} O(1/\sqrt{\tau^k}) & O(1) \\ O(1) & O(1) \end{pmatrix}.$$

As in [13], we define a linear manifold:

$$(5.3)\quad \begin{aligned} \mathcal{M} \equiv \{(X', y', S') \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n : A_i \bullet X' &= b_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^m y_i' A_i + S' &= C, \\ q_i^T X' q_j &= 0 \text{ if } i \text{ or } j \in \mathbb{N}, \\ q_i^T S' q_j &= 0 \text{ if } i \text{ or } j \in \mathbb{B}\}. \end{aligned}$$

It is easily seen that if $(X', y', S') \in \mathcal{M}$, then

$$Q^T X' Q = \begin{pmatrix} M_B & 0 \\ 0 & 0 \end{pmatrix}, \quad Q^T S' Q = \begin{pmatrix} 0 & 0 \\ 0 & M_N \end{pmatrix}.$$

In the next theorem, we provide a sufficient condition for the superlinear convergence of Algorithm 3.1. This sufficient condition will be characterized by the asymptotic behavior of the following quantity η_k

$$(5.4)\quad \eta_k = \eta_k(\Gamma) = \frac{1}{\tau^k} ||(X^k)^{-\frac{1}{2}}(X^k - \breve{X}^k)(S^k - \breve{S}^k)(X^k)^{\frac{1}{2}}||_F,$$

where $(\breve{X}^k, \breve{y}^k, \breve{S}^k)$ is the solution of the following minimization problem:

$$(5.5)\quad \min\{||(X^k)^{-\frac{1}{2}}(X^k - X')(S^k - S')(X^k)^{\frac{1}{2}}||_F : (X', y', S') \in \mathcal{M},\ ||(X', S')||_F \le \Gamma\},$$

and Γ is a constant such that $||(X^k, S^k)||_F \le \Gamma$, $\forall k$. Note that every accumulation point of (X^k, y^k, S^k) belongs to the feasible set of the above minimization problem and the feasible set is bounded. Therefore $(\breve{X}^k, \breve{S}^k)$ exists for each k.

THEOREM 5.1. *Under the strict complementarity assumption, if $\eta_k \to 0$ as $k \to \infty$, then Algorithm 3.1 is superlinearly convergent. Moreover, if there exists a constant $\sigma > 0$ such that $\eta_k = O((\tau^k)^\sigma)$, then the convergence has Q-order at least $1 + \sigma$ in the sense that $\mu^{k+1} = O((\mu^k)^{1+\sigma})$.*

Proof. By Lemma 3.5, it remains to prove that $\delta^k \to 0$ as $k \to \infty$. For simplicity, let us omit the index k. It is easily seen that $(U + X - \breve{X}, w + y - \breve{y}, V + S - \breve{S})$ satisfies (3.14) with

$$\begin{aligned} H &= H_P(X(S - \breve{S}) + (X - \breve{X})S - XS) \\ &= H_P((X - \breve{X})(S - \breve{S})) \ \text{(since } \breve{X}\breve{S} = 0). \end{aligned}$$

Let

$$(5.6)\quad \Delta = X^{-\frac{1}{2}}(X - \breve{X})(S - \breve{S})X^{\frac{1}{2}}.$$

Then, according to (ii) of Lemma 4.3,

$$(5.7)\quad ||H||_F \le \sqrt{3\kappa}||\Delta||_F = \sqrt{3\kappa}\eta\tau.$$

Denoting

$$\Delta_x = X^{-\frac{1}{2}}(U + X - \breve{X})X^{-\frac{1}{2}}, \quad \Delta_s = X^{\frac{1}{2}}(V + S - \breve{S})X^{\frac{1}{2}},$$

and applying Corollary 3.4, we obtain

$$\tau||\Delta_x||_F \leq \frac{||H||_F}{(1-\sqrt{\kappa}\alpha)(1-\alpha)} \leq 4\sqrt{3\kappa}\eta\tau,$$

which implies

$$||\Delta_x||_F \leq 4\sqrt{3\kappa}\eta. \tag{5.8}$$

Similarly,

$$||\Delta_s||_F \leq \frac{(1+\alpha)||H||_F}{1-\sqrt{\kappa}\alpha} \leq 3\sqrt{3\kappa}\eta\tau. \tag{5.9}$$

The fact that $(\breve{X}, \breve{y}, \breve{S}) \in \mathcal{M}$, together with (5.1)-(5.2), implies

$$\begin{aligned}
||X^{-\frac{1}{2}}(X - \breve{X})X^{-\frac{1}{2}}||_F &= ||I - X^{-\frac{1}{2}}\breve{X}X^{-\frac{1}{2}}||_F \\
&\leq ||I||_F + ||X^{-\frac{1}{2}}\breve{X}X^{-\frac{1}{2}}||_F \\
&= \sqrt{n} + ||Q^T X^{-\frac{1}{2}} Q Q^T \breve{X} Q Q^T X^{-\frac{1}{2}} Q||_F \\
&= \sqrt{n} + ||(\widehat{x_1}, \dots, \widehat{x_n}) Q^T \breve{X} Q (\widehat{x_1}, \dots, \widehat{x_n})^T||_F \\
&= \sqrt{n} + ||\sum_{i,j \in \mathbb{B}} (q_i^T \breve{X} q_j)\widehat{x_i}\widehat{x_j}^T||_F = O(1).
\end{aligned} \tag{5.10}$$

Similarly,

$$||S^{-\frac{1}{2}}(S - \breve{S})S^{-\frac{1}{2}}||_F = O(1). \tag{5.11}$$

Let us observe that

$$\begin{aligned}
X^{-\frac{1}{2}}UVX^{\frac{1}{2}} &= \left(X^{-\frac{1}{2}}UX^{-\frac{1}{2}}\right)\left(X^{\frac{1}{2}}VX^{\frac{1}{2}}\right) \\
&= \left(\Delta_x - X^{-\frac{1}{2}}(X - \breve{X})X^{-\frac{1}{2}}\right)\left(\Delta_s - X^{\frac{1}{2}}(S - \breve{S})X^{\frac{1}{2}}\right) \\
&= \Delta_x\Delta_s - \left(X^{-\frac{1}{2}}(X - \breve{X})X^{-\frac{1}{2}}\right)\Delta_s \\
&\quad -\Delta_x X^{\frac{1}{2}}S^{\frac{1}{2}}\left(S^{-\frac{1}{2}}(S - \breve{S})S^{-\frac{1}{2}}\right)S^{\frac{1}{2}}X^{\frac{1}{2}} + \Delta.
\end{aligned}$$

Then from (5.7), (5.8), (5.9), (5.10), (5.11) and Corollary 4.2, we have

$$\begin{aligned}
||X^{-\frac{1}{2}}UVX^{\frac{1}{2}}||_F &\leq ||\Delta_x||_F||\Delta_s||_F + ||X^{-\frac{1}{2}}(X - \breve{X})X^{-\frac{1}{2}}||_F||\Delta_s||_F \\
&\quad + ||X^{\frac{1}{2}}S^{\frac{1}{2}}||^2||\Delta_x||_F||S^{-\frac{1}{2}}(S - \breve{S})S^{-\frac{1}{2}}||_F + ||\Delta||_F \\
&= O(\eta\tau).
\end{aligned}$$

Hence, according to statement (ii) of Lemma 4.3 we get,

$$\delta = ||PUVP^{-1}||_F/\tau \leq \sqrt{3\kappa}||X^{-\frac{1}{2}}UVX^{\frac{1}{2}}||_F/\tau = O(\eta).$$

Therefore, $\delta^k \to 0$ if $\eta_k \to 0$. Finally, if $\eta_k = O(\tau^k \sigma)$ for some constant $\sigma > 0$, then we have $\delta^k = O((\tau^k)^\sigma)$. From Lemma 3.5,

$$\begin{aligned}1-\bar{\theta} &\le 1-\frac{2}{1+\sqrt{1+4\delta/(\beta-\alpha)}}\\ &= \frac{\sqrt{1+4\delta/(\beta-\alpha)}-1}{\sqrt{1+4\delta/(\beta-\alpha)}+1}\\ &= \frac{4\delta/(\beta-\alpha)}{(\sqrt{1+\delta/(\beta-\alpha)}+1)^2}\\ &\le \delta/(\beta-\alpha) = O(\tau^\sigma).\end{aligned}$$

Therefore, $\tau^{k+1} = (1-\bar{\theta}^k)\tau^k = O((\tau^k)^{1+\sigma})$. Recalling (3.21), we obtain $\mu^{k+1} = O((\mu^k)^{1+\sigma})$. □

Because of Theorem 5.1, the local convergence analysis established in [14] also applies to Algorithm 3.1 if Assumption A4 is satisfied. We end this section by stating the following result without proof. For its proof, we refer the reader to [14].

THEOREM 5.2. *Under the the strict complementarity assumption and Assumption A4, Algorithm 3.1 is superlinearly convergent. Moreover, if* $X^k S^k = O((\tau^k)^{0.5+\sigma})$ *for some constant* $\sigma > 0$*, then the convergence has Q-order at least* $1 + \min\{\sigma,\ 0.5\}$.

6. Further remarks. We have shown that the class of predictor-corrector algorithms defined by Algorithm 3.1 shares the same global and local convergence properties with the infeasible-interior-point algorithm of Kojima, Shida and Shindoh and Potra and Sheng. This result suggests that the practical performance of these algorithms should be similar.

The iteration complexity of Algorithm 3.1 depends on the spectral condition number of J_x^k or J_x^k. From a computational point of view, the choice of $J_s = I$ (where $P^T P = S$) seems to be preferable (cf. Zhang [18]). However more computational experiments are necessary before a definitive conclusion is reached (see also [16] for a comparison between the performance of several Mehrotra predictor-corrector algorithms for SDP).

Acknowledgment. The authors are grateful to anonymous referees for their valuable comments. The work of F. A. Potra and R. Sheng was supported in part by NSF Grant DMS 9996154.

REFERENCES

[1] F. ALIZADEH, J.-P. 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.

[2] 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.

[3] M. KOJIMA, M. SHIDA, AND S. SHINDOH, *Local convergence of predictor-corrector infeasible-interior-point method for SDPs and SDLCPs*, Math. Programming, 80(1998). pp. 129–160.

[4] M. KOJIMA, M. SHIDA, AND S. SHINDOH, *A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton search direction*, SIAM J. Optim., 9 (1999), pp. 444–465.

[5] M. KOJIMA, S. SHINDOH, AND S. HARA, *Interior-point methods for the monotone linear complementarity problem in symmetric matrices*, SIAM J. Optim., 7(1997). pp. 86–125.

[6] C-J. LIN AND R. SAIGAL, *A predictor-corrector method for semidefinite programming*, Working paper, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109–2177, USA, 1995.

[7] Z-Q. LUO, J. F. STURM, AND S. ZHANG, *Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming*, SIAM J. Optim., 8(1998). pp. 59–81.

[8] S. MIZUNO, M. J. TODD, AND Y. YE, *On adaptive-step primal-dual interior-point algorithms for linear programming*, Math. of Operations Research, 18(1993), pp. 964–981.

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

[10] R. D. C. MONTEIRO AND Y. ZHANG, *A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming*, Math. Programming, 81(1998), pp. 281–299.

[11] Y. E. NESTEROV AND M. J. TODD, *Primal–dual interior–point methods for self–scaled cones*, SIAM J. Optim., 8(1998), pp. 324–364.

[12] F. A. POTRA AND R. SHENG, *On homogeneous interior-point algorithms for semidefinite programming*, Optimization Methods and Software, 9(1998), pp. 161–184.

[13] F. A. POTRA AND R. SHENG, *A superlinearly convergent primal–dual infeasible–interior–point algorithm for semidefinite programming*, SIAM J. Optim., 8(1998), pp. 1007–1028.

[14] F. A. POTRA AND R. SHENG, *Superlinear convergence of interior-point algorithms for semidefinite programming*, Journal of Optimization Theory and Applications, 99:1 (1998), pp. 103–119.

[15] J. F. STURM AND S. ZHANG, *Symmetric primal-dual path following algorithms for semidefinite programming*, Report 9554/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1995.

[16] 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.

[17] L. VANDENBERGHE AND S. BOYD, *Semidefinite programming*, SIAM Review, 38(1996), pp. 49–95.

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