SIAM J. OPTIM.
Vol. 8, No. 3, pp. 797–812, August 1998

009

POLYNOMIAL CONVERGENCE OF PRIMAL-DUAL ALGORITHMS FOR SEMIDEFINITE PROGRAMMING BASED ON THE MONTEIRO AND ZHANG FAMILY OF DIRECTIONS*

RENATO D. C. MONTEIRO†

Abstract. This paper establishes the polynomial convergence of the class of primal-dual feasible interior-point algorithms for semidefinite programming (SDP) based on the Monteiro and Zhang family of search directions. In contrast to Monteiro and Zhang's work [*Math. Programming*, 81 (1998), pp. 281–299], here no condition is imposed on the scaling matrix that determines the search direction. We show that the polynomial iteration-complexity bounds of two well-known algorithms for linear programming, namely the short-step path-following algorithm of Kojima, Mizuno, and Yoshise [*Math. Programming*, 44 (1989), pp. 1–26] and Monteiro and Adler [*Math. Programming*, 44 (1989), pp. 27–41 and pp. 43–66] and the predictor-corrector algorithm of Mizuno, Todd, and Ye [*Math. Oper. Res.*, 18 (1993), pp. 945–981] carry over into the context of SDP. Since the Monteiro and Zhang family of directions includes the Alizadeh, Haeberly, and Overton direction, we establish for the first time the polynomial convergence of algorithms based on this search direction.

Key words. semidefinite programming, interior-point methods, polynomial complexity, path-following methods, primal-dual methods

AMS subject classifications. 65K05, 90C25, 90C30

PII. S1052623496308618

1. Introduction.

Several authors have discussed generalizations of interior-point algorithms for linear programming (LP) to the context of semidefinite programming (SDP). Nesterov and Nemirovskii propose in their landmark works [22, 23] a general approach for using interior-point methods to solve convex programs based on the notion of self-concordant functions. (See their book [25] for a comprehensive treatment of this subject.) They show that the problem of minimizing a linear function over a convex set can be solved in "polynomial time" as long as a self-concordant barrier function for the convex set is known. In particular, they show that linear programs, convex quadratic programs with convex quadratic constraints, and semidefinite programs all have explicit and easily computable self-concordant barrier functions, and hence can be solved in "polynomial time." On the other hand, Alizadeh [1] extends Ye's projective potential reduction algorithm [33] for LP to SDP and argues that many known interior point algorithms for LP can also be transformed into algorithms for SDP in a mechanical way. Since then, many authors have proposed interior-point algorithms for solving SDP problems, including Alizadeh, Haeberly, and Overton [2, 3]; Freund [4]; Helmberg et al. [5]; Jarre [7]; Kojima, Shida, and Shindoh [10, 12]; Kojima, Shindoh, and Hara [13]; Lin and Saigal [14]; Luo, Sturm, and Zhang [15]; Monteiro [17]; Monteiro and Zhang [21]; Nesterov and Nemirovskii [24]; Nesterov and Todd [26, 27]; Potra and Sheng [28]; Sturm and Zhang [29]; Tseng [31]; Vandenberghe and Boyd [32]; and Zhang [34]. Most of these more recent works concentrate on primal-dual methods.

*Received by the editors August 28, 1996; accepted for publication (in revised form) July 24, 1997; published electronically July 20, 1998. This research was partially supported by National Science Foundation grants INT-9600343 and CCR-9700448 and Office of Naval Research grant N00014-94-1-0340.

http://www.siam.org/journals/siopt/8-3/30861.html

†School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332 (monteiro@isye.gatech.edu).

The first SDP algorithms that are extensions of primal-dual LP algorithms, such as the long-step path-following algorithm of Kojima, Mizuno, and Yoshise [9], the short-step path-following algorithm of Kojima, Mizuno, and Yoshise [8] and Monteiro and Adler [18, 19], and the predictor-corrector algorithm of Mizuno, Todd, and Ye [16], use one of the following three search directions: (i) the Alizadeh, Haeberly, and Overton (AHO) direction proposed in [2]; (ii) a direction independently proposed by Kojima, Shindoh, and Hara [13] and Helmberg et al. [5], and later rediscovered by Monteiro [17], which we refer to as the KSH/HRVW/M-direction; and (iii) the Nesterov and Todd (NT) direction introduced in [26, 27]. Application of the Newton method to the central path equation $XS = \sigma\mu I$ results in an equation of the form

$$X\Delta S + \Delta X S = \sigma\mu I - XS, \tag{1}$$

which in general yields nonsymmetric directions. The AHO-direction corresponds to the symmetric equation obtained by symmetrizing both sides of the previous equation.

Another way of symmetrizing (1) is to first apply a similarity transformation $P(\cdot)P^{-1}$ to both sides and then symmetrize it. Such an approach was first introduced by Monteiro [17] for the cases of $P = X^{-1/2}$ and $P = S^{1/2}$. The resulting directions were found to be equivalent to two special directions of the class of directions introduced earlier by Kojima, Shindoh, and Hara [13] using a different motivation. The second direction (with $P = S^{1/2}$), which is the KSH/HRVW/M-direction, was also proposed by Helmberg et al. [5] independently from [13]. To unify the above directions, including the one by Alizadeh, Haeberly, and Overton [2], Zhang [34] formally introduced the above scaling and symmetrization scheme for a general nonsingular scaling matrix P, which leads to a class of search directions parametrized by P. As suggested in Monteiro and Zhang [21], we refer to this set of search directions as the Monteiro–Zhang (MZ) family of search directions (or simply, the MZ-unified direction). In a recent paper, Todd, Toh, and Tütüncü [30] study conditions for the existence and uniqueness of search directions in the MZ-family and show that the NT-direction [26] is a member of this family corresponding to any scaling matrix P such that $P^TP = S^{1/2}(S^{1/2}XS^{1/2})^{-1/2}S^{1/2}$. More recently, Kojima, Shida, and Shindoh [11] show that the NT-direction also belongs to the class of directions introduced by Kojima, Shindoh, and Hara [13]. In contrast, it should be noted that the AHO-direction does not belong to the latter family.

This paper introduces new techniques for establishing the polynomiality of interior-point primal-dual SDP algorithms whose iterates belong to a narrow (or Frobenius norm) neighborhood of the central path. We illustrate our techniques by showing the polynomiality of two primal-dual feasible algorithms based on the MZ-unified direction: a short-step path-following method and a predictor-corrector method which are extensions of the LP algorithms studied in Kojima, Mizuno, and Yoshise [8], Monteiro and Adler [18, 19], and Mizuno, Todd, and Ye [16]. The iteration complexity bounds for both algorithms are shown to be exactly the same as their corresponding LP algorithms, namely, $\mathcal{O}(\sqrt{n}L)$ iterations to reduce the duality gap by a factor of at least $2^{-\mathcal{O}(L)}$. As opposed to Monteiro and Zhang [21], who establish an iteration-complexity bound for a long-step path-following method based on a subclass of the MZ-family of directions, namely, that corresponding to scaling matrices P such that $PXSP^{-1}$ is symmetric, our analysis is valid for the whole MZ-family of directions. Since this family includes the AHO-direction, we establish for the first time, as a special case of our unified analysis, the polynomial convergence of algorithms based on the AHO-direction. Finally, we mention that Kojima, Shida, and Shindoh [12] establish the global (but not polynomial) convergence of a primal-dual predictor-corrector algorithm based on the AHO-direction whose iterates are restricted to a central path neighborhood narrower than the one used in this paper.

This paper is organized as follows. In section 2, we introduce the SDP problem, review the scaling and symmetrization scheme discussed above, and give some useful preliminary results. In section 3, we state and prove the technical results used in the polynomial convergence analysis of the algorithms of section 4. In section 4, we establish the polynomiality of two primal-dual feasible algorithms: the short-step path-following algorithm in subsection 4.1 and the predictor-corrector algorithm in subsection 4.2.

1.1. Notation and terminology. The following notation is used throughout the paper. The superscript ^{T} denotes transpose. $\Re^p$ denotes the p-dimensional Euclidean space. The set of all $p \times q$ matrices with real entries is denoted by $\Re^{p\times q}$. The set of all symmetric $p \times p$ 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 matrix $Q \in \Re^{p\times p}$ is denoted by $\text{Tr } Q \equiv \sum_{i=1}^n Q_{ii}$. For a matrix $Q \in \Re^{p\times p}$ with all real eigenvalues, we denote its eigenvalues by $\lambda_i[Q]$, $i = 1, \ldots, p$, and its smallest eigenvalue by $\lambda_{\min}[Q]$. Given P and Q in $\Re^{p\times q}$, the inner product between them in the vector space $\Re^{p\times q}$ is defined as $P \bullet Q \equiv \text{Tr } P^T Q$. The Euclidean norm and its associated operator norm are both denoted by $\|\cdot\|$; hence $\|Q\| \equiv \max_{\|u\|=1} \|Qu\|$ for any $Q \in \Re^{p\times p}$. The Frobenius norm of $Q \in \Re^{p\times p}$ is $\|Q\|_F \equiv (Q \bullet Q)^{1/2}$. $\mathcal{S}^p_+$ and $\mathcal{S}^p_{++}$ denote the set of all matrices in $\mathcal{S}^p$ which are positive semidefinite and positive definite, respectively. $\mathcal{S}^p_\perp$ denotes the set of all skew-symmetric matrices in $\Re^{p\times p}$. Since $\mathcal{S}^p + \mathcal{S}^p_\perp = \Re^{p\times p}$ and $U \bullet V = 0$ for every $U \in \mathcal{S}^p$ and $V \in \mathcal{S}^p_\perp$, it follows that $\mathcal{S}^p_\perp$ is the orthogonal complement of $\mathcal{S}^p$ with respect to the inner product $\bullet$.

2. The SDP problem and preliminary discussion. In this section, we describe the SDP problem considered in this paper and review the similar symmetrization operator introduced by Zhang [34] and the MZ-family of search directions.

This paper studies primal-dual path-following algorithms for solving the SDP problem

$$\text{(2)} \qquad \text{(P)} \qquad \min\{C \bullet X : A_i \bullet X = b_i,\, i = 1, \ldots, m,\, X \succeq 0\},$$

and its associated dual SDP problem

$$\text{(3)} \qquad \text{(D)} \qquad \max\left\{ b^T y : \sum_{i=1}^m y_i A_i + S = C,\, S \succeq 0 \right\},$$

where $C \in \mathcal{S}^n$, $A_i \in \mathcal{S}^n$, $i = 1, \ldots, m$, and $b = (b_1, \ldots, b_m)^T \in \Re^m$ are the data, and $X \in \mathcal{S}^n_+$ and $(S, y) \in \mathcal{S}^n_+ \times \Re^m$ are the primal and dual variables, respectively.

The set of *interior feasible solutions* of (2) and (3) are

$$F^0(\text{P}) \equiv \{X \in \mathcal{S} : A_i \bullet X = b_i,\, i = 1, \ldots, m,\, X \succ 0\},$$
$$F^0(\text{D}) \equiv \left\{ (S, y) \in \mathcal{S} \times \Re^m : \sum_{i=1}^m y_i A_i + S = C,\, S \succ 0 \right\},$$

respectively. Throughout this paper, we assume that $F^0(P) \times F^0(D) \neq \emptyset$ and that the matrices A_i, $i = 1, \ldots, m$, are linearly independent. Under the first assumption, it is well known that both (2) and (3) have optimal solutions X^* and (y^*, S^*) such that $C \bullet X^* = b^T y^*$, i.e., the optimal values of (2) and (3) coincide. This last condition, called the strong duality, can be alternatively expressed as $X^* \bullet S^* = 0$ or $X^* S^* = 0$. Hence, the set of primal and dual optimal solutions consist of all the solutions $(X, S, y) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+ \times \Re^m$ to the following optimality system:

$$XS = 0, \tag{4a}$$

$$\sum_{i=1}^m y_i A_i + S - C = 0, \tag{4b}$$

$$A_i \bullet X - b_i = 0, \ i = 1, \ldots, m, \tag{4c}$$

where (4a) is called the complementarity equation.

The left-hand side of the above system, viewed as a function of (X, S, y), maps $\mathcal{S}^n \times \mathcal{S}^n \times \Re^m$ into $\Re^{n\times n} \times \mathcal{S}^n \times \Re^m$, and hence it is a function between spaces of different dimensions. To apply Newton-type algorithms to solve (4), it is first necessary to symmetrize (4) so that the left-hand side of the resulting equivalent system becomes a function from $\mathcal{S}^n \times \mathcal{S}^n \times \Re^m$ into itself.

Motivated by the works of Alizadeh, Haeberly, and Overton [2] and Monteiro [17], Zhang [34] introduced a general symmetrization scheme based on the so-called *similar symmetrization* operator $H_P : \Re^{n\times n} \to \mathcal{S}^n$ defined as

$$H_P(M) \equiv \frac{1}{2}\left[PMP^{-1} + (PMP^{-1})^T\right] \quad \forall M \in \Re^{n\times n},$$

where $P \in \Re^{n\times n}$ is some nonsingular matrix. It has been shown by Zhang [34] that

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

for any nonsingular matrix P, any matrix M with real spectrum (e.g., $M = XS$ with $X, S \in \mathcal{S}^n_+$), and any $\tau \in \Re$. Consequently, for any given nonsingular matrix P, (4) is equivalent to the symmetric system

$$H_P(XS) = 0, \tag{5a}$$

$$\sum_{i=1}^m y_i A_i + S - C = 0, \tag{5b}$$

$$A_i \bullet X - b_i = 0, \ i = 1, \ldots, m, \tag{5c}$$

to which Newton-type methods can be applied. A perturbed Newton method applied to this system leads to the following linear system:

$$H_P(\Delta X S + X \Delta S) = \sigma\mu I - H_P(XS), \tag{6a}$$

$$\sum_{i=1}^m \Delta y_i A_i + \Delta S = C - S - \sum_{i=1}^m y_i A_i, \tag{6b}$$

$$A_i \bullet \Delta X = b_i - A_i \bullet X, \ i = 1, \ldots, m, \tag{6c}$$

where $(\Delta X, \Delta S, \Delta y) \in \mathcal{S}^n \times \mathcal{S}^n \times \Re^m$, $\sigma \in [0, 1]$ is the centering parameter, and $\mu = \mu(X, S) \equiv (X \bullet S)/n$ is the normalized duality gap corresponding to (X, S, y). Note that the solution $(\Delta X, \Delta S, \Delta y)$ of this linear system is the Newton direction at the point (X, S, y) with respect to a system of equations defining the (unique) point on the central path with duality gap $\sigma\mu$, namely, the system consisting of (5b), (5c) and $H_P(XS) = \sigma\mu I$.

The algorithms studied in this paper are all based on the centrality measure at a point $(X, S) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+$ defined as

$$d(X,S) \equiv \left\| X^{1/2} S X^{1/2} - \mu I \right\|_F = \left[\sum_{i=1}^{n} \left(\lambda_i[XS] - \mu\right)^2 \right]^{1/2}. \tag{7}$$

Given a constant $\gamma \in (0,1)$, we let $\mathcal{N}_F(\gamma)$ denote the following narrow (or Frobenius) neighborhood of the central path:

$$\mathcal{N}_F(\gamma) \equiv \{(X,S,y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D) : d(X,S) \le \gamma\mu\}.$$

The following simple result plays an important role in the polynomial convergence analysis of the algorithms of section 4.

LEMMA 2.1. *Suppose that* $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, $Q \in \Re^{n \times n}$ *is a nonsingular matrix, and* $\mu \equiv (X \bullet S)/n$. *Then*

(a) $d(\widetilde{X}, \widetilde{S}) = d(X,S)$, *where* $\widetilde{X} \equiv QXQ^T$ *and* $\widetilde{S} \equiv Q^{-T} S Q^{-1}$;

(b) $d(X,S) \le \|H_Q(XS - \mu I)\|_F$ *with equality holding if* $QXSQ^{-1} \in \mathcal{S}^n$.

Proof. (a) Since $\widetilde{X}\widetilde{S} = QXSQ^{-1}$, that is, XS and $\widetilde{X}\widetilde{S}$ are similar matrices, we conclude that they have the same spectrum. This implies that $d(\widetilde{X}, \widetilde{S}) = d(X,S)$, since $d(X,S)$ can be expressed only in terms of the eigenvalues of XS.

(b) It has been shown in Lemma 3.3 of Monteiro [17] that $\|E\|_F \le \|H_T(E)\|_F$ for any $E \in \mathcal{S}^n$ and any nonsingular $T \in \Re^{n \times n}$. Using this inequality, with $E \equiv X^{1/2} S X^{1/2} - \mu I$ and $T \equiv Q X^{1/2}$, and noting that $H_T(E) = H_Q(XS - \mu I)$, we obtain the inequality in (b). If $QXSQ^{-1} \in \mathcal{S}^n$, then $H_Q(XS - \mu I) = QXSQ^{-1} - \mu I$, and hence

$$\begin{aligned} \|H_Q(XS - \mu I)\|_F = \left\| QXSQ^{-1} - \mu I \right\|_F &= \left[\sum_{i=1} \left(\lambda_i[QXSQ^{-1}] - \mu\right)^2 \right]^{1/2} \\ = \left[\sum_{i=1} \left(\lambda_i[XS] - \mu\right)^2 \right]^{1/2} &= d(X,S). \qquad \square \end{aligned}$$

3. Technical results. In this section we provide some technical results which will be used to establish the polynomial convergence of the algorithms presented in section 4.

We assume throughout this section that $(X,S,y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$, $P \in \Re^{n \times n}$ is a nonsingular matrix, and $(\Delta X, \Delta S, \Delta y)$ is a solution of system (6) for some $\sigma \in [0,1]$. Moreover, for $\alpha \in \Re$, we let

$$\widetilde{X} \equiv PXP^T, \qquad \widetilde{S} \equiv P^{-T} S P^{-1}, \tag{8}$$

$$\widetilde{\Delta X} \equiv P \Delta X P^T, \qquad \widetilde{\Delta S} \equiv P^{-T} \Delta S P^{-1}, \tag{9}$$

$$X(\alpha) \equiv X + \alpha \Delta X, \qquad S(\alpha) \equiv S + \alpha \Delta S, \qquad y(\alpha) \equiv y + \alpha \Delta y, \tag{10}$$

$$\widetilde{X}(\alpha) \equiv P X(\alpha) P^T = \widetilde{X} + \alpha \widetilde{\Delta X}, \tag{11}$$

$$\widetilde{S}(\alpha) \equiv P^{-T} S(\alpha) P^{-1} = \widetilde{S} + \alpha \widetilde{\Delta S}, \tag{12}$$

$$W_X \equiv \widetilde{X}^{-1/2} \left[\widetilde{\Delta X}\widetilde{S} + \widetilde{X}\widetilde{\Delta S} + \widetilde{X}\widetilde{S} - \sigma\mu I \right] \widetilde{X}^{1/2}, \tag{13}$$

$$\mu \equiv \frac{X \bullet S}{n} = \frac{\widetilde{X} \bullet \widetilde{S}}{n}, \qquad \mu(\alpha) \equiv \frac{X(\alpha) \bullet S(\alpha)}{n} = \frac{\widetilde{X}(\alpha) \bullet \widetilde{S}(\alpha)}{n}. \tag{14}$$

LEMMA 3.1. *$(\Delta X, \Delta S, \Delta y)$ satisfies system (6) if and only if $(\widetilde{\Delta X}, \widetilde{\Delta S}, \Delta y)$ is a solution of the system*

$$H_I(\widetilde{\Delta X}\widetilde{S} + \widetilde{X}\widetilde{\Delta S}) = \sigma\mu I - H_I(\widetilde{X}\widetilde{S}), \tag{15a}$$

$$\sum_{i=1}^{m} \Delta y_i \widetilde{A}_i + \widetilde{\Delta S} = \widetilde{C} - \widetilde{S} - \sum_{i=1}^{m} y_i \widetilde{A}_i, \tag{15b}$$

$$\widetilde{A}_i \bullet \widetilde{\Delta X} = b_i - \widetilde{A}_i \bullet \widetilde{X}, \ i = 1, \dots, m, \tag{15c}$$

where $\widetilde{C} \equiv P^{-T}CP^{-1}$ and $\widetilde{A}_i \equiv P^{-T}A_iP^{-1}$ for $i = 1, \dots, m$.

Proof. The proof follows immediately from (6), (8), and (9). □

The result above says that $(\widetilde{\Delta X}, \widetilde{\Delta S}, \Delta y)$ is exactly the AHO-direction at the point $(\widetilde{X}, \widetilde{S}, y)$ for the pair of dual semidefinite programs with data $\widetilde{C}$ and $(\widetilde{A}_i, b_i)$, $i = 1, \dots, m$. As an immediate consequence of this lemma and Theorem 1 of Monteiro and Zanjácomo [20], we obtain the following result which ensures the existence and uniqueness of the solution of system (6) when (X, S, y) is well centered.

PROPOSITION 3.2. *If $(X, S, y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ is such that*

$$\|S^{1/2}XS^{1/2} - \nu I\| \le \frac{\nu}{2}$$

for some scalar $\nu > 0$, then system (6) has a unique solution.

Proof. It was shown in Theorem 1 of Monteiro and Zanjácomo [20] that system (15) has a unique solution whenever there exists $\nu > 0$ such that

$$\|\widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2} - \nu I\| \le \frac{\nu}{2}.$$

Since $\|S^{1/2}XS^{1/2} - \nu I\| = \|\widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2} - \nu I\|$, the result follows from Lemma 3.1. □

It should be observed that the above result can be strengthened if further conditions are imposed on the scaling matrix P. For example, it has been shown by Todd, Toh, and Tütüncü [30] (see also Monteiro and Zhang [21]) that system (6) has a unique solution for any $(X, S, y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ and scaling matrix P such that $PXSP^{-1} \in \mathcal{S}^n$.

In the remaining part of this section, we impose the following condition (X, S, y).

Assumption. $(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$.

LEMMA 3.3. *The following relations hold:*

$$H_{\widetilde{X}^{1/2}}(W_X) = 0, \tag{16}$$

$$\widetilde{\Delta X} \bullet \widetilde{\Delta S} = \Delta X \bullet \Delta S \ = \ 0. \tag{17}$$

Proof. Relation (16) follows immediately from (15a). The two identities in (17) follow from relations (6b), (6c), and (9) and the assumption that $(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$. □

LEMMA 3.4. *For every $\alpha \in \Re$, we have*

$$\mu(\alpha) = (1 - \alpha + \sigma\alpha)\mu, \tag{18}$$

$$\widetilde{X}^{-1/2}\left[\widetilde{X}(\alpha)\widetilde{S}(\alpha) - \mu(\alpha)I\right]\widetilde{X}^{1/2} = (1-\alpha)\left(\widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} - \mu I\right) + \alpha\, W_X \tag{19}$$

$$+ \alpha^2 \widetilde{X}^{-1/2}\widetilde{\Delta X}\widetilde{\Delta S}\widetilde{X}^{1/2}. \tag{20}$$

Proof. By (10), we have

$$X(\alpha)S(\alpha) = (X+\alpha\Delta X)(S+\alpha\Delta S) = XS + \alpha(X\Delta S + \Delta XS) + \alpha^2 \Delta X\Delta S,$$

which, together with the linearity of $H_P(\cdot)$ and (6a), implies that

$$\begin{aligned} H_P(X(\alpha)S(\alpha)) &= H_P(XS) + \alpha H_P(X\Delta S + \Delta XS) + \alpha^2 H_P(\Delta X\Delta S) \\ &= H_P(XS) + \alpha H_P(\sigma\mu I - XS) + \alpha^2 H_P(\Delta X\Delta S) \\ &= (1-\alpha)H_P(XS) + \alpha\sigma\mu I + \alpha^2 H_P(\Delta X\Delta S). \end{aligned}$$

Using this expression and the identity Tr $H_P(M) =$ Tr M for $M \in \Re^{n\times n}$, we obtain

$$\begin{aligned} X(\alpha)\bullet S(\alpha) &= \mathrm{Tr}\ [X(\alpha)S(\alpha)] \ = \ \mathrm{Tr}\ [H_P(X(\alpha)S(\alpha))] \\ &= \mathrm{Tr}\ [(1-\alpha)H_P(XS) + \alpha\sigma\mu I + \alpha^2 H_P(\Delta X\Delta S)] \\ &= (1-\alpha)\mathrm{Tr}\ H_P(XS) + \alpha\sigma\mu n + \alpha^2 \mathrm{Tr}\ H_P(\Delta X\Delta S) \\ &= (1-\alpha)X\bullet S + \alpha\sigma\mu n + \alpha^2\Delta X\bullet\Delta S. \end{aligned}$$

Dividing this expression by n and noting (14) and (17), we obtain (18). Using (11), (12), (18), and (13), we conclude that, for every $\alpha \in \Re$,

$$\begin{aligned} &\widetilde{X}(\alpha)\widetilde{S}(\alpha) - \mu(\alpha)I \\ &= (\widetilde{X} + \alpha\widetilde{\Delta X})(\widetilde{S} + \alpha\widetilde{\Delta S}) - \mu(\alpha)I \\ &= \widetilde{X}\widetilde{S} + \alpha(\widetilde{\Delta X}\widetilde{S} + \widetilde{X}\widetilde{\Delta S}) + \alpha^2\widetilde{\Delta X}\widetilde{\Delta S} - (1-\alpha+\sigma\alpha)\mu I \\ &= (1-\alpha)(\widetilde{X}\widetilde{S} - \mu I) + \alpha(\widetilde{\Delta X}\widetilde{S} + \widetilde{X}\widetilde{\Delta S} + \widetilde{X}\widetilde{S} - \sigma\mu I) + \alpha^2\widetilde{\Delta X}\widetilde{\Delta S} \\ &= (1-\alpha)(\widetilde{X}\widetilde{S} - \mu I) + \alpha\widetilde{X}^{1/2}W_X\widetilde{X}^{-1/2} + \alpha^2\widetilde{\Delta X}\widetilde{\Delta S}. \end{aligned}$$

Multiplying this identity on the left by $\widetilde{X}^{-1/2}$ and on the right by $\widetilde{X}^{1/2}$, we obtain (20). ∎

The following technical result plays a major role in our analysis.

LEMMA 3.5. *Let $W \in \Re^{n\times n}$ be such that $H_Q(W) = 0$ for some nonsingular $Q \in \Re^{n\times n}$. Then,*

$$\|H_I(W)\|_F \le \frac{1}{2}\left\|W - W^T\right\|_F, \tag{21}$$

$$\|W\|_F \le \frac{\sqrt{2}}{2}\left\|W - W^T\right\|_F. \tag{22}$$

In particular, if $W = U_1 + U_2$ for some $U_1 \in \mathcal{S}^n$ and $U_2 \in \Re^{n\times n}$, then

$$\|W\|_F \le \sqrt{2}\,\|U_2\|_F.$$

Proof. Define $U \equiv (W + W^T)/2 \in \mathcal{S}^n$ and $\widehat{U} \equiv (W - W^T)/2 \in \mathcal{S}^n_\perp$. Clearly, $W = U + \widehat{U}$ and $U \bullet \widehat{U} = 0$. Using these relations, the assumption that $H_Q(W) = 0$, and the identity $\|B + B^T\|_F^2 = 2\left(\|B\|_F^2 + \mathrm{Tr}\ B^2\right)$ for $B \in \Re^{n\times n}$, we obtain

$$\begin{aligned} 0 = 2\|H_Q(W)\|_F^2 \ &= \ \frac{1}{2}\left\|QWQ^{-1} + (QWQ^{-1})^T\right\|_F^2 \\ &= \left\|QWQ^{-1}\right\|_F^2 + \mathrm{Tr}\ \left[\left(QWQ^{-1}\right)^2\right] \ = \ \left\|QWQ^{-1}\right\|_F^2 + \mathrm{Tr}\ \left[W^2\right] \\ &\ge \mathrm{Tr}\ \left[(U+\widehat{U})^2\right] \ = \ \mathrm{Tr}\ \left[U^2 + U\widehat{U} + \widehat{U}U + \widehat{U}^2\right] \\ &= \|U\|_F^2 + 2U\bullet\widehat{U} + \mathrm{Tr}\ \widehat{U}^2 \ = \ \|U\|_F^2 - \|\widehat{U}\|_F^2, \end{aligned}$$

and hence $\|U\|_F \le \|\widehat{U}\|_F$; that is, (21) holds. Using this inequality and the fact that $U \bullet \widehat{U} = 0$, we obtain

$$\|W\|_F^2 = \|U + \widehat{U}\|_F^2 = \|U\|_F^2 + \|\widehat{U}\|_F^2 \le 2\|\widehat{U}\|_F^2;$$

that is, (22) holds. To show the last part of the lemma, observe that $\|W - W^T\|_F = \|U_2 - U_2^T\|_F \le 2\|U_2\|_F$, and hence, by (22), we obtain $\|W\|_F \le \sqrt{2}\|U_2\|_F$. □

Using the above result, we can now prove the following lemma.

LEMMA 3.6. *For every $\theta \in \Re$, we have*

$$\|W_X\|_F \;\le\; \sqrt{2}\,\delta_x \left\|\widetilde{S}^{1/2}\widetilde{X}^{1/2} - \theta\mu\widetilde{S}^{-1/2}\widetilde{X}^{-1/2}\right\| \tag{23}$$

and

$$\begin{aligned}\left\|\widetilde{X}^{-1/2}[\widetilde{X}(\alpha)\widetilde{S}(\alpha) - \mu(\alpha)I]\widetilde{X}^{1/2}\right\|_F &\le (1-\alpha)\,d(\widetilde{X},\widetilde{S}) + \alpha^2\,\delta_x\,\delta_s \\ &\quad + \alpha\sqrt{2}\,\delta_x \left\|\widetilde{X}^{1/2}\widetilde{S}^{1/2} - \theta\mu\widetilde{X}^{-1/2}\widetilde{S}^{-1/2}\right\| \end{aligned}\tag{24}$$

for all $\alpha \in [0,1]$, where

$$\delta_x \equiv \left\|\widetilde{X}^{-1/2}\widetilde{\Delta X}\widetilde{S}^{1/2}\right\|_F, \qquad \delta_s \equiv \left\|\widetilde{S}^{-1/2}\widetilde{\Delta S}\widetilde{X}^{1/2}\right\|_F. \tag{25}$$

Proof. Let $\theta \in \Re$ be given. By (13), we have $W_X = U_1 + U_2$, where

$$\begin{aligned} U_1 &\equiv \widetilde{X}^{1/2}\widetilde{\Delta S}\widetilde{X}^{1/2} + \theta\mu\widetilde{X}^{-1/2}\widetilde{\Delta X}\widetilde{X}^{-1/2} + \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} - \sigma\mu I \;\in\; \mathcal{S}^n, \\ U_2 &\equiv \widetilde{X}^{-1/2}\widetilde{\Delta X}\widetilde{S}^{1/2}\left(\widetilde{S}^{1/2}\widetilde{X}^{1/2} - \theta\mu\widetilde{S}^{-1/2}\widetilde{X}^{-1/2}\right). \end{aligned}$$

By (16), we have $H_Q(W_X) = 0$ for $Q \equiv \widetilde{X}^{1/2}$. These observations, together with Lemma 3.5, the definition of U_2 and δ_x, and the fact that $\|AB\|_F \le \|A\|_F\|B\|$ for every $A, B \in \Re^{n\times n}$ (see exercise 20 of section 5.6 of [6]), imply that

$$\|W_X\|_F \le \sqrt{2}\,\|U_2\|_F \le \sqrt{2}\,\delta_x \left\|\widetilde{S}^{1/2}\widetilde{X}^{1/2} - \theta\mu\widetilde{S}^{-1/2}\widetilde{X}^{-1/2}\right\|;$$

i.e., (23) holds. Moreover, relations (7) and (20), the fact that $\alpha \in [0,1]$, and the definitions of δ_x and δ_s, yield

$$\begin{aligned} &\left\|\widetilde{X}^{-1/2}\left[\widetilde{X}(\alpha)\widetilde{S}(\alpha) - \mu(\alpha)\right]\widetilde{X}^{1/2}\right\|_F \\ &\le (1-\alpha)\left\|\widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} - \mu I\right\|_F + \alpha\,\|W_X\|_F + \alpha^2\,\delta_x\,\delta_s \\ &= (1-\alpha)\,d(\widetilde{X},\widetilde{S}) + \alpha\,\|W_X\|_F + \alpha^2\,\delta_x\,\delta_s, \end{aligned}$$

which, together with (23), imply (24). □

To shorten the notation used in the results below, we let

$$\Phi_\theta(A,B) \equiv \left\|A^{1/2}B^{1/2} - \theta\frac{A \bullet B}{n}A^{-1/2}B^{-1/2}\right\|_F, \tag{26}$$

for every $(A, B) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ and $\theta \in \Re$.

LEMMA 3.7. *If $d(X,S) \le \gamma\mu$ for some $\gamma \in (0,1)$, then*

$$\left\| \widetilde{X}^{-1/2}\widetilde{S}^{-1/2} \right\|^2 \le \frac{1}{(1-\gamma)\mu}, \tag{27}$$

$$\left[\Phi_\theta(\widetilde{X},\widetilde{S})\right]^2 \le \frac{\gamma^2 + (1-\theta)^2 n}{1-\gamma}\,\mu. \tag{28}$$

Proof. By relation (7) and the assumption that $d(X,S) \le \gamma\mu$, we have $\lambda_{\min}[XS] \ge (1-\gamma)\mu$. This, together with (8), implies

$$\begin{aligned}\left\| \widetilde{X}^{-1/2}\widetilde{S}^{-1/2} \right\|^2 = \left\| \widetilde{X}^{-1/2}\widetilde{S}^{-1}\widetilde{X}^{-1/2} \right\| &= \frac{1}{\lambda_{\min}\left[\widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2}\right]} = \frac{1}{\lambda_{\min}\left[\widetilde{X}\widetilde{S}\right]} \\ &= \frac{1}{\lambda_{\min}\left[PXSP^{-1}\right]} = \frac{1}{\lambda_{\min}\left[XS\right]} \le \frac{1}{(1-\gamma)\mu};\end{aligned}$$

that is, (27) holds. Using relations (26), (14), (27), (7), the fact that the matrices $\widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} - \mu I$ and I are orthogonal, the assumption that $d(X,S) \le \gamma\mu$, and Lemma 2.1(a), we obtain

$$\begin{aligned}\left[\Phi_\theta(\widetilde{X},\widetilde{S})\right]^2 &= \left\| \widetilde{X}^{1/2}\widetilde{S}^{1/2} - \theta\mu\widetilde{X}^{-1/2}\widetilde{S}^{-1/2} \right\|_F^2 \\ &\le \left\| \widetilde{X}^{-1/2}\widetilde{S}^{-1/2} \right\|^2 \left\| \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} - \theta\mu I \right\|_F^2 \\ &\le \frac{1}{(1-\gamma)\mu}\left(\left\| \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} - \mu I \right\|_F^2 + \|\mu I - \theta\mu I\|_F^2 \right) \\ &= \frac{1}{(1-\gamma)\mu}\left(d(X,S)^2 + (1-\theta)^2\mu^2 n \right) \le \frac{\gamma^2 + (1-\theta)^2 n}{1-\gamma}\,\mu;\end{aligned}$$

that is, (28) holds. ∎

LEMMA 3.8. *If $(X,S,y) \in \mathcal{N}_F(\gamma)$ for some $\gamma > 0$ satisfying*

$$2\sqrt{2}\frac{\gamma}{1-\gamma} \le 1, \tag{29}$$

then

$$\max\{\delta_x, \delta_s\} \le 2\,\Phi_\sigma(\widetilde{X},\widetilde{S}),$$

where δ_x, δ_x, and $\Phi_\sigma(\cdot,\cdot)$ are defined in (25) *and* (26).

Proof. Using (13), it is easy to see that

$$\widetilde{X}^{1/2}\widetilde{\Delta S}\widetilde{S}^{-1/2} + \widetilde{X}^{-1/2}\widetilde{\Delta X}\widetilde{S}^{1/2} = W_X\widetilde{X}^{-1/2}\widetilde{S}^{-1/2} + \sigma\mu\widetilde{X}^{-1/2}\widetilde{S}^{-1/2} - \widetilde{X}^{1/2}\widetilde{S}^{1/2}. \tag{30}$$

By (17), it is easy to see that the two terms on the left-hand side of (30) are orthogonal. This observation, together with (25), (30), (26), (14), (23) with $\theta = 1$, (27), (28) with $\theta = 1$, and (29), implies that

$$
\begin{aligned}
\max\{\delta_x, \delta_s\} \le \left(\delta_x^2 + \delta_s^2\right)^{1/2} &= \left(\left\| \widetilde{X}^{-1/2} \widetilde{\Delta X} \widetilde{S}^{1/2} \right\|_F^2 + \left\| \widetilde{S}^{-1/2} \widetilde{\Delta S} \widetilde{X}^{1/2} \right\|_F^2 \right)^{1/2} \\
&= \left\| \widetilde{X}^{-1/2} \widetilde{\Delta X} \widetilde{S}^{1/2} + \widetilde{X}^{1/2} \widetilde{\Delta S} \widetilde{S}^{-1/2} \right\|_F \\
&\le \|W_X\|_F \left\| \widetilde{X}^{-1/2} \widetilde{S}^{-1/2} \right\| + \left\| \widetilde{X}^{1/2} \widetilde{S}^{1/2} - \sigma\mu \widetilde{X}^{-1/2} \widetilde{S}^{-1/2} \right\|_F \\
&\le \sqrt{2}\, \delta_x \left\| \widetilde{X}^{1/2} \widetilde{S}^{1/2} - \mu \widetilde{X}^{-1/2} \widetilde{S}^{-1/2} \right\|_F \left\| \widetilde{X}^{-1/2} \widetilde{S}^{-1/2} \right\| + \Phi_\sigma(\widetilde{X}, \widetilde{S}) \\
&\le \sqrt{2}\, \delta_x \frac{\gamma \mu^{1/2}}{(1-\gamma)^{1/2}} \frac{1}{(1-\gamma)^{1/2} \mu^{1/2}} + \Phi_\sigma(\widetilde{X}, \widetilde{S}) \\
&\le \frac{\delta_x}{2} + \Phi_\sigma(\widetilde{X}, \widetilde{S}) \ \le \ \frac{1}{2} \max\{\delta_x, \delta_s\} + \Phi_\sigma(\widetilde{X}, \widetilde{S}),
\end{aligned}
$$

from which the lemma follows. $\square$

As an immediate consequence of the above lemmas, we obtain the following result.

LEMMA 3.9. *If $(X, S, y) \in \mathcal{N}_F(\gamma)$ for some $\gamma > 0$ satisfying (29), then, for every $\alpha \in [0, 1]$, we have*

$$
\begin{aligned}
&\left\| \widetilde{X}^{-1/2} [\widetilde{X}(\alpha) \widetilde{S}(\alpha) - \mu(\alpha)] \widetilde{X}^{1/2} \right\|_F \\
&\le \left\{ (1-\alpha)\,\gamma + 2\sqrt{2}\,\alpha\, \frac{\gamma \left[\gamma^2 + (1-\sigma)^2 n\right]^{1/2}}{1-\gamma} + 4\,\alpha^2\, \frac{\gamma^2 + (1-\sigma)^2 n}{1-\gamma} \right\} \mu.
\end{aligned}
$$

Proof. Using (24) with $\theta = 1$, Lemma 2.1(a), (8), the assumption that $d(X, S) \le \gamma\mu$, Lemma 3.8, and (28) with $\theta = 1$ and $\theta = \sigma$, we obtain

$$
\begin{aligned}
&\left\| \widetilde{X}^{-1/2} [\widetilde{X}(\alpha) \widetilde{S}(\alpha) - \mu(\alpha)] \widetilde{X}^{1/2} \right\|_F \\
&\le (1-\alpha)\, d(X, S) + \sqrt{2}\,\alpha\, \delta_x\, \Phi_1(\widetilde{X}, \widetilde{S}) + \alpha^2\, \delta_x\, \delta_s \\
&\le (1-\alpha)\gamma\mu + 2\sqrt{2}\,\alpha\, \Phi_\sigma(\widetilde{X}, \widetilde{S})\, \Phi_1(\widetilde{X}, \widetilde{S}) + 4\,\alpha^2 \left[\Phi_\sigma(\widetilde{X}, \widetilde{S})\right]^2 \\
&\le \left\{ (1-\alpha)\gamma + 2\sqrt{2}\,\alpha\, \frac{\gamma \left[\gamma^2 + (1-\sigma)^2 n\right]^{1/2}}{1-\gamma} + 4\,\alpha^2\, \frac{\gamma^2 + (1-\sigma)^2 n}{1-\gamma} \right\} \mu. \qquad \square
\end{aligned}
$$

4. Algorithms and polynomial convergence.

In this section, we establish polynomial iteration-complexity bounds for two primal-dual feasible interior-point algorithms based on the MZ-family of search directions given by (6). Both algorithms are extensions of well-known algorithms for linear programming: the first one is a short-step path-following method which generalizes the algorithms presented in Kojima, Mizuno, and Yoshise [8] and Monteiro and Adler [18, 19]; the second one is a predictor-corrector algorithm similar to the predictor-corrector LP method of Mizuno, Todd, and Ye [16].

4.1. Short-step path-following algorithm.

In this subsection, we analyze the polynomial convergence of a short-step path-following algorithm based on the MZ-unified search direction (6).

We start by stating the algorithm that will be considered in this subsection.

ALGORITHM I.

Choose constants γ and δ in $(0,1)$ satisfying the conditions of Corollary 4.2 below and let $\sigma \equiv 1-\delta/\sqrt{n}$. Let $L>1$ and $(X^0,S^0,y^0)\in\mathcal{F}^0(P)\times\mathcal{F}^0(D)$ be such that $d(X^0,S^0)\le\gamma\mu_0$, where $\mu_0\equiv(X^0\bullet S^0)/n$.

Repeat until $\mu_k\le 2^{-L}\mu_0$, **do**

(1) Choose a nonsingular matrix $P^k\in\Re^{n\times n}$;

(2) Compute the solution $(\Delta X^k,\Delta S^k,\Delta y^k)$ of system (6) with $P=P^k$, $\mu=\mu_k$, and $(X,S,y)=(X^k,S^k,y^k)$;

(3) Set $(X^{k+1},y^{k+1},S^{k+1})\equiv(X^k,y^k,S^k)+(\Delta X^k,\Delta S^k,\Delta y^k)$;

(4) Set $\mu_{k+1}\equiv(X^{k+1}\bullet S^{k+1})/n$ and increment k by 1.

End

Setting $\Gamma=\gamma$ in the following result, we obtain the analysis of one iteration of Algorithm I for suitable choices of the constants γ and δ.

THEOREM 4.1. *Let $\gamma,\Gamma\in(0,1)$ and $\delta\in[0,n^{1/2})$ be constants satisfying*

$$\frac{7\,(\gamma^2+\delta^2)}{1-\gamma}\le\left(1-\frac{\delta}{\sqrt{n}}\right)\Gamma,\qquad \frac{2\sqrt{2}\,\gamma}{1-\gamma}\le 1. \tag{31}$$

Suppose that $(X,S,y)\in\mathcal{N}_F(\gamma)$, and let $(\Delta X,\Delta S,\Delta y)$ denote the solution of system (6), where $\sigma\equiv 1-\delta/\sqrt{n}$, $\mu\equiv(X\bullet S)/n$, and $P\in\Re^{n\times n}$ is an invertible matrix. Then,

(a) $(\widehat{X},\widehat{S},\widehat{y})\equiv(X+\Delta X,S+\Delta S,y+\Delta y)\in\mathcal{N}_F(\Gamma)$;

(b) $\widehat{X}\bullet\widehat{S}=(1-\delta/\sqrt{n})(X\bullet S)$.

Proof. It follows from Lemma 3.9, the definition of σ, and (31) that for every $\alpha\in[0,1]$,

$$\begin{aligned}\left\|\widetilde{X}^{-1/2}[\widetilde{X}(\alpha)\widetilde{S}(\alpha)-\mu(\alpha)]\widetilde{X}^{1/2}\right\|_F &\le\left\{(1-\alpha)\gamma+7\,\alpha\frac{\gamma^2+(1-\sigma)^2n}{1-\gamma}\right\}\mu\\ &=\left\{(1-\alpha)\gamma+\alpha\frac{7\,(\gamma^2+\delta^2)}{1-\gamma}\right\}\mu\\ &\le\left\{(1-\alpha)\gamma+\alpha\left(1-\frac{\delta}{\sqrt{n}}\right)\Gamma\right\}\mu\\ &=\{(1-\alpha)\gamma+\alpha\Gamma\sigma\}\mu,\end{aligned}\tag{32}$$

and hence, in view of (18),

$$\left\|\widetilde{X}^{-1/2}[\widetilde{X}(\alpha)\widetilde{S}(\alpha)-\mu(\alpha)]\widetilde{X}^{1/2}\right\|_F\ \le\ \max\{\gamma,\Gamma\}\,(1-\alpha+\alpha\sigma)\,\mu\ =\ \max\{\gamma,\Gamma\}\,\mu(\alpha).$$

This relation implies that the matrix

$$G(\alpha)\equiv\left(\widetilde{X}^{-1/2}\widetilde{X}(\alpha)\widetilde{S}(\alpha)\widetilde{X}^{1/2}\right)/\mu(\alpha)$$

satisfies $\|G(\alpha)-I\|\le\max\{\gamma,\Gamma\}<1$, and thus that $G(\alpha)$ is nonsingular for every $\alpha\in[0,1]$. Hence, we conclude that $\widetilde{X}(\alpha)$ and $\widetilde{S}(\alpha)$ are also nonsingular for every $\alpha\in[0,1]$. A trivial continuity argument together with (11) and (12) shows that $(\widetilde{X}(\alpha),\widetilde{S}(\alpha))$ and $(X(\alpha),S(\alpha))$ are in $\mathcal{S}^n_{++}\times\mathcal{S}^n_{++}$, for every $\alpha\in[0,1]$. Moreover, using (6b) and (6c) and the fact that $(X,S,y)\in\mathcal{F}^0(P)\times\mathcal{F}^0(D)$, we easily see that $(X(\alpha),S(\alpha),y(\alpha))$ satisfies (5b) and (5c), and hence that $(X(\alpha),S(\alpha),y(\alpha))\in\mathcal{F}^0(P)\times\mathcal{F}^0(D)$, for every $\alpha\in[0,1]$. Using Lemma 2.1(a) with $(X,S)=(X(1),S(1))$ and $Q=P$, Lemma 2.1(b) with $(X,S) = (\widetilde{X}(1), \widetilde{S}(1))$ and $Q = \widetilde{X}^{-1/2}$, and relations (11), (12), (18) with $\alpha = 1$, and (32) with $\alpha = 1$, we obtain

$$d(X(1),S(1)) = d(\widetilde{X}(1),\widetilde{S}(1)) \le \left\| H_{\widetilde{X}^{-1/2}}[\widetilde{X}(1)\widetilde{S}(1) - \mu(1)I] \right\|_F$$
$$\le \left\| \widetilde{X}^{-1/2}[\widetilde{X}(1)\widetilde{S}(1) - \mu(1)]\widetilde{X}^{1/2} \right\|_F \le \Gamma\mu(1).$$

We have thus shown that $(\widehat{X}, \widehat{S}, \widehat{y}) = (X(1), S(1), y(1)) \in \mathcal{N}_F(\Gamma)$; that is, (a) holds. Statement (b) is an immediate consequence of (18) with $\alpha = 1$ and the definition of σ. ∎

As an immediate consequence of Theorem 4.1, we have the following convergence result for Algorithm I.

COROLLARY 4.2. *Assume that* γ *and* δ *are constants in* $(0,1)$ *satisfying*

$$\frac{7\,(\gamma^2 + \delta^2)}{1-\gamma} \le \left(1 - \frac{\delta}{\sqrt{n}}\right)\gamma, \qquad \frac{2\sqrt{2}\,\gamma}{1-\gamma} \le 1.$$

Then, every iterate (X^k, S^k, y^k) *generated by Algorithm* I *is in the neighborhood* $\mathcal{N}_F(\gamma)$ *and satisfies* $X^k \bullet S^k = (1 - \delta/\sqrt{n})^k (X^0 \bullet S^0)$. *Moreover, Algorithm* I *terminates in at most* $\mathcal{O}(\sqrt{n}L)$ *iterations.*

Examples of constants γ and δ satisfying the conditions of Corollary 4.2 are $\gamma = \delta = 1/20$.

4.2. Predictor-corrector algorithm. In this subsection, we give the polynomial convergence analysis of a predictor-corrector algorithm which is a direct extension of the LP predictor-corrector algorithm studied by Mizuno, Todd, and Ye [16].

The algorithm considered in this subsection is as follows.

ALGORITHM II.

Choose a constant $0 < \tau < 1$ satisfying the conditions of Theorem 4.3 below. Let $L > 1$ and $(X^0, S^0, y^0) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ be such that $d(X^0, S^0) \le \tau\mu_0$, where $\mu_0 \equiv (X^0 \bullet S^0)/n$.

Repeat until $\mu_k \le 2^{-L}\mu_0$, **do**

(1) Choose a nonsingular matrix $P^k \in \Re^{n\times n}$;
(2) Compute the solution $(\Delta X^k, \Delta S^k, \Delta y^k)$ of system (6) with $P = P^k$, $\sigma = 0$, and $(X, S, y) = (X^k, S^k, y^k)$;
(3) Let $\alpha_k \equiv \max\{\alpha \in [0,1] : (X^k(\alpha'), S^k(\alpha'), y^k(\alpha')) \in \mathcal{N}_F(2\tau), \forall \alpha' \in [0, \alpha]\}$, where $(X^k(\alpha), S^k(\alpha), y^k(\alpha)) \equiv (X^k + \alpha\Delta X^k, S^k + \alpha\Delta S^k, y^k + \alpha\Delta y^k)$;
(4) Let $(\widehat{X}^k, \widehat{S}^k, \widehat{y}^k) \equiv (X^k, S^k, y^k) + \alpha_k(\Delta X^k, \Delta S^k, \Delta y^k)$;
(5) Choose a nonsingular matrix $\hat{P}^k \in \Re^{n\times n}$;
(6) Compute the solution $(\widehat{\Delta X}^k, \widehat{\Delta S}^k, \widehat{\Delta y}^k)$ of system (6) with $P = P^k$, $\sigma = 1$, and $(X, S, y) = (\widehat{X}^k, \widehat{S}^k, \widehat{y}^k)$;
(7) Set $(X^{k+1}, S^{k+1}, y^{k+1}) \equiv (\widehat{X}^k, \widehat{S}^k, \widehat{y}^k) + (\widehat{\Delta X}^k, \widehat{\Delta S}^k, \widehat{\Delta y}^k)$;
(8) Set $\mu_{k+1} \equiv (X^{k+1} \bullet S^{k+1})/n$ and increment k by 1.

End

The following result provides the polynomial convergence analysis of the above algorithm.

THEOREM 4.3. *Assume that* $\tau \in (0, 1/30]$. *Then, Algorithm* II *satisfies the following statements:*

(a) *for every* $k \geq 0$, $(X^k, S^k, y^k) \in \mathcal{N}_F(\tau)$ *and* $(\widehat{X}^k, \widehat{S}^k, \widehat{y}^k) \in \mathcal{N}_F(2\tau)$;

(b) *for every* $k \geq 0$, $X^{k+1} \bullet S^{k+1} = \widehat{X}^k \bullet \widehat{S}^k = (1-\alpha_k)(X^k \bullet S^k)$ *and*

$$\alpha_k = \frac{1}{\mathcal{O}(\sqrt{n})};$$

(c) *the algorithm terminates in at most* $\mathcal{O}(\sqrt{n}L)$ *iterations.*

Proof. Statement (c) and the well-definedness of Algorithm II follow directly from (a) and (b). In turn, these two statements follow by a simple induction argument and the two lemmas below. $\square$

The following lemma analyzes the predictor step of Algorithm II, namely, the step described in items (1)–(4) of Algorithm II.

LEMMA 4.4. *Suppose that* $(X, S, y) \in \mathcal{N}_F(\tau)$ *for some* $\tau \in (0, 1/2)$. *Let the triple* $(\Delta X, \Delta S, \Delta y)$ *denote the solution of* (6) *with* $\sigma = 0$ *and for some invertible matrix* $P \in \Re^{n\times n}$. *Let* $\bar{\alpha}$ *denote the unique positive root of the second-order polynomial* $p(\alpha)$ *defined as*

$$p(\alpha) \equiv 4\frac{(\tau^2+n)}{1-\tau}\alpha^2 + \left(2\sqrt{2}\tau\frac{(\tau^2+n)^{1/2}}{1-\tau} + \tau\right)\alpha - \tau. \tag{33}$$

Then, for any $\alpha \in [0, \bar{\alpha}]$, *we have*

(a) $(X(\alpha), S(\alpha), y(\alpha)) \in \mathcal{N}_F(2\tau)$;

(b) $X(\alpha) \bullet S(\alpha) = (1-\alpha)(X \bullet S)$.

Moreover, $\bar{\alpha} = 1/\mathcal{O}(n^{1/2})$.

Proof. Using Lemma 3.9 with $\gamma = \tau$ and $\sigma = 0$, the fact that $p(\alpha) \leq 0$ for $\alpha \in [0, \bar{\alpha}]$, and relation (18) with $\sigma = 0$ and (33), we obtain

$$\begin{aligned}
&\left\|\widetilde{X}^{-1/2}[\widetilde{X}(\alpha)\widetilde{S}(\alpha) - \mu(\alpha)]\widetilde{X}^{1/2}\right\|_F \\
&\leq \left\{(1-\alpha)\tau + 2\sqrt{2}\alpha\frac{\tau(\tau^2+n)^{1/2}}{1-\tau} + 4\,\alpha^2\frac{\tau^2+n}{1-\tau}\right\}\mu \\
&= 2\tau\mu(\alpha) + p(\alpha)\mu \;\leq\; 2\tau\mu(\alpha).
\end{aligned}$$

An argument similar to the one used in Theorem 4.1, together with the fact that $2\tau < 1$, can be used to show that (a) holds. Statement (b) follows from (18) with $\sigma = 0$. The last assertion of the lemma follows by a straightforward verification. $\square$

The following lemma analyzes the corrector step of Algorithm II, namely, the step described in items (5)–(7) of Algorithm II.

LEMMA 4.5. *Suppose* $(\widehat{X}, \widehat{S}, \widehat{y})$ *is in* $\mathcal{N}_F(2\tau)$ *for some* $\tau \in (0, 1/30]$. *Let the triple* $(\widehat{\Delta X}, \widehat{\Delta S}, \widehat{\Delta y})$ *denote the solution of* (6) *with* $(X, S, y) = (\widehat{X}, \widehat{S}, \widehat{y})$, $\sigma = 1$, *and* $P = \hat{P}$, *for some nonsingular matrix* $\hat{P} \in \Re^{n\times n}$. *Then,*

$$\begin{aligned}
(\widehat{X}, \widehat{S}, \widehat{y}) + (\widehat{\Delta X}, \widehat{\Delta S}, \widehat{\Delta y}) &\in \mathcal{N}_F(\tau), \\
(\widehat{X} + \widehat{\Delta X}) \bullet (\widehat{S} + \widehat{\Delta S}) &= \widehat{X} \bullet \widehat{S}.
\end{aligned}$$

Proof. The result follows immediately from Theorem 4.1 with $\delta = 0$, $(X, S, y) = (\widehat{X}, \widehat{S}, \widehat{y})$, $\gamma = 2\tau$ and $\Gamma = \tau$, and the fact that (31) holds when $0 < \tau \leq 1/30$. $\square$

5. Remarks. To keep the presentation simple and the analysis more transparent, we have focused our attention only on feasible methods. It would be interesting to check whether the techniques developed in this paper could also be used to establish the polynomial convergence of infeasible primal-dual path-following algorithms based on the MZ-unified direction.

The analysis of this paper makes strong use of the fact that the algorithms are based on the narrow neighborhood of the central path. The question of whether long-step path-following methods based on the MZ-unified direction can be proved to be "polynomially convergent" under weaker conditions than those imposed in Monteiro and Zhang [21] is an interesting topic for future research.

We end this section by stating a technical result whose proof is similar to that of Lemma 3.5 and which may eventually have application elsewhere. To state this result, define for every nonsingular matrix $P \in \Re^{n\times n}$ the operator $A_P : \Re^{n\times n} \to \mathcal{S}^n_\perp$ as

$$A_P(M) \equiv \frac{1}{2}\left[PMP^{-1} - (PMP^{-1})^T\right] \quad \forall M \in \Re^{n\times n}.$$

LEMMA 5.1. *Let $W \in \Re^{n\times n}$ be such that $A_Q(W) = 0$ for some nonsingular $Q \in \Re^{n\times n}$. Then,*

$$\|A_I(W)\|_F \le \|H_I(W)\|_F,$$
$$\|W\|_F \le \sqrt{2}\,\|H_I(W)\|_F.$$

In particular, if $W = V_1 + V_2$ for some $V_1 \in \mathcal{S}^n_\perp$ and $V_2 \in \Re^{n\times n}$, then

$$\|W\|_F \le \sqrt{2}\,\|V_2\|_F.$$

Proof. The proof is the same as the proof of Lemma 3.5 but instead of using the identity $\|B + B^T\|_F^2 = 2\left(\|B\|_F^2 + \mathrm{Tr}\ B^2\right)$ uses the alternative identity $\|B - B^T\|_F^2 = 2\left(\|B\|_F^2 - \mathrm{Tr}\ B^2\right)$. ∎

Note that, in terms of the operator A_P, the inequalities (21) and (22) can be simply written as

$$\|H_I(W)\|_F \le \|A_I(W)\|_F,$$
$$\|W\|_F \le \sqrt{2}\,\|A_I(W)\|_F,$$

respectively.

Acknowledgments. This work was done while the author was visiting the Institute of Statistical Mathematics and the Tokyo Institute of Technology in Japan. He is grateful to both institutes for providing a congenial scientific atmosphere during his stay. He would like to thank his host Dr. Takashi Tsuchiya for making his visit to Japan a very enjoyable experience and for having many useful discussions which have been invaluable toward the development of this work. Finally, thanks also are due to Prof. Masakazu Kojima for making possible the author's living at the International House in the Tokyo Institute of Technology. Finally, the financial support of the National Science Foundation and Office of Naval Research is gratefully acknowledged.

REFERENCES

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[2] F. Alizadeh, J.-P. Haeberly, and M. Overton, *Primal-dual Interior-point Methods for Semidefinite Programming*, Technical Report 659, Computer Science Department, Courant Institute of Mathematical Sciences, New York University, New York, NY, 1994.

[3] F. Alizadeh, J.-P. Haeberly, and M. Overton, *Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results*, SIAM J. Optim., 8 (1998), pp. 746–768.

[4] R. M. Freund, *Complexity of an Algorithm for Finding an Approximate Solution of a Semidefinite Program with No Regularity Condition*, Working paper OR 302-94, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, December 1994.

[5] 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.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[7] F. Jarre, *An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices*, SIAM J. Control Optim., 31 (1993), pp. 1360–1377.

[8] M. Kojima, S. Mizuno, and A. Yoshise, *A polynomial-time algorithm for a class of linear complementarity problems*, Math. Programming, 44 (1989), pp. 1–26.

[9] M. Kojima, S. Mizuno, and A. Yoshise, *A primal–dual interior point algorithm for linear programming*, in Progress in Mathematical Programming: Interior Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989, pp. 29–47.

[10] M. Kojima, M. Shida, and S. Shindoh, *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs*, Math. Programming, 80 (1998), pp. 129–160.

[11] M. Kojima, M. Shida, and S. Shindoh, *A Note on the Nesterov–Todd and the Kojima–Shindoh–Hara Search Directions in Semidefinite Programming*, Research Report B-313, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo, Japan, April 1996.

[12] 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., to appear.

[13] 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.

[14] C.-J. Lin and R. Saigal, *A Predictor-corrector Method for Semi-definite Programming*, Working paper, Dept. of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI, 1995.

[15] 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.

[16] S. Mizuno, M. J. Todd, and Y. Ye, *On adaptive step primal-dual interior-point algorithms for linear programming*, Math. Oper. Res., 18 (1993), pp. 945–981.

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

[18] R. D. C. Monteiro and I. Adler, *Interior path-following primal-dual algorithms. Part* I: *Linear programming*, Math. Programming, 44 (1989), pp. 27–41.

[19] R. D. C. Monteiro and I. Adler, *Interior path-following primal-dual algorithms. Part* II: *Convex quadratic programming*, Math. Programming, 44 (1989), pp. 43–66.

[20] R. D. C. Monteiro and P. Zanjácomo, *A note on the existence of the Alizadeh–Haeberly–Overton direction for semidefinite programming*, Math. Programming, 78 (1997), pp. 393–396.

[21] 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.

[22] Y. E. Nesterov and A. S. Nemirovskii, *Polynomial barrier methods in convex programming*, Ekonomika i Matem. Metody, 24 (1988), pp. 1084–1091 (in Russian; English translation in Matekon: Translations of Russian and East European Math. Economics).

[23] Y. E. Nesterov and A. S. Nemirovskii, *Self-concordant Functions and Polynomial Time Methods in Convex Programming*, preprint, Central Economic & Mathematical Institute, USSR Acad. Sci. Moscow, USSR, 1989.

[24] Y. E. Nesterov and A. S. Nemirovskii, *Optimization Over Positive Semidefinite Matrices: Mathematical Background and User's Manual*, Technical report, Central Economic & Mathematical Institute, USSR Acad. Sci. Moscow, USSR, 1990.

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

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

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

[28] F. A. Potra and R. Sheng, *A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming*, SIAM J. Optim., to appear.

[29] 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, November 1995.

[30] 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.

[31] P. Tseng, *Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP*, Optim. Meth. and Software, to appear.

[32] L. Vandenberghe and S. Boyd, *A primal-dual potential reduction method for problems involving matrix inequalities*, Math. Programming, 69 (1995), pp. 205–236.

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

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