SIAM J. OPTIM.
Vol. 9, No. 3, pp. 551–577
© 1999 Society for Industrial and Applied Mathematics

POLYNOMIAL CONVERGENCE OF A NEW FAMILY OF PRIMAL-DUAL ALGORITHMS FOR SEMIDEFINITE PROGRAMMING*

RENATO D. C. MONTEIRO† AND TAKASHI TSUCHIYA‡

Abstract. This paper establishes the polynomial convergence of a new class of primal-dual interior-point path-following feasible algorithms for semidefinite programming (SDP) whose search directions are obtained by applying Newton's method to the symmetric central path equation

$$(PXP^T)^{1/2}(P^{-T}SP^{-1})(PXP^T)^{1/2} - \mu I = 0,$$

where P is a nonsingular matrix. Specifically, we show that the short-step path-following algorithm based on the Frobenius norm neighborhood and the semilong-step path-following algorithm based on the operator 2-norm neighborhood have $O(\sqrt{n}L)$ and $O(nL)$ iteration-complexity bounds, respectively. When $P = I$, this yields the first polynomially convergent semilong-step algorithm based on a pure Newton direction. Restricting the scaling matrix P at each iteration to a certain subset of nonsingular matrices, we are able to establish an $O(n^{3/2}L)$ iteration complexity for the long-step path-following method. The resulting subclass of search directions contains both the Nesterov–Todd direction and the Helmberg–Rendl–Vanderbei–Wolkowicz/Kojima–Shindoh–Hara/Monteiro direction.

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

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

PII. S1052623496312836

1. Introduction.

Several authors have discussed generalizations of interior-point algorithms for linear programming (LP) to the context of semidefinite programming (SDP). The landmark work in this direction is due to Nesterov and Nemirovskii [22, 23], where a general approach for using interior-point methods for solving convex programs is proposed, 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, Nesterov and Nemirovskii 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 [37] 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], Freund [3], Helmberg et al. [4],

*Received by the editors December 1, 1996; accepted for publication (in revised form) February 17, 1998; published electronically April 19, 1999.
http://www.siam.org/journals/siopt/9-3/31283.html

†School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332 (monteiro@isye.gatech.edu). The work of this author was based on research supported by the Office of Naval Research under grant N00014-94-1-0340 and the National Science Foundation under grants INT-9600343 and CCR-9700448.

‡The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku, Tokyo, 106, Japan (tsuchiya@sun312.ism.ac.jp). This author's research was supported in part by the Grant-in-Aid for Encouragement of Young Scientists (A) 08780227 (1996) of the Ministry of Education, Science and Culture of Japan.

Jarre [5], Kojima, Shida, and Shindoh [10], Kojima, Shindoh, and Hara [11], Lin and Saigal [12], Luo, Sturm, and Zhang [13], Monteiro [15, 16], Monteiro and Tsuchiya [19], Monteiro and Zhang [21], Nesterov and Nemirovskii [24], Nesterov and Todd [26, 27], Potra and Sheng [29], Sturm and Zhang [30], Tseng [35], Vandenberghe and Boyd [36], and Zhang [38]. Most of these more recent works are concentrated on primal-dual methods.

The first algorithms for SDP which are extensions of well-known primal-dual LP algorithms, such as the long-step path-following algorithm of Kojima, Mizuno, and Yoshise [7] and Tanabe [31, 32], the short-step path-following algorithm of Kojima, Mizuno, and Yoshise [6] and Monteiro and Adler [17, 18], and the predictor-corrector algorithm of Mizuno, Todd, and Ye [14], use one of the following three search directions: (i) the Alizadeh–Haeberly–Overton (AHO) direction proposed in [2]; (ii) a direction independently proposed by Helmberg et al. [4] and Kojima, Shindoh, and Hara [11], and later rediscovered by Monteiro [15], which we refer to as the HRVW/KSH/M direction; and (iii) the Nesterov–Todd (NT) direction introduced in [26, 27]. Application of Newton's method to the central path equation $XS = \sigma\mu I$ results in an equation of the form

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

which in general yields nonsymmetric directions. The AHO direction corresponds to the symmetric equation obtained by symmetrizing both sides of (1.1).

Another way of symmetrizing (1.1) is first to apply a similarity transformation $P(\cdot)P^{-1}$ to both sides of (1.1) and then to symmetrize it. Such an approach was first introduced by Monteiro [15] for the cases $P = X^{-1/2}$ and $P = S^{1/2}$. The resulting directions were found to be equivalent to two special directions of the KSH family of directions introduced earlier by Kojima, Shindoh, and Hara [11] using a different approach. The second direction (with $P = S^{1/2}$), which is the HRVW/KSH/M direction, was also proposed by Helmberg et al. [4] independently from [11]. (For simplicity, we refer to the first direction with $P = X^{-1/2}$ as the HRVW/KSH/M dual direction. We use the term HRVW/KSH/M directions to refer to both of them.) To unify the NT direction and the HRVW/KSH/M directions, Zhang [38] 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, usually referred to as the Monteiro and Zhang (MZ) family. Subsequently, Todd, Toh, and Tütüncü [34] and Kojima, Shida, and Shindoh [8] showed that the NT direction is a member of the MZ family and the KSH family, respectively. In contrast, it is known that the AHO direction does not belong to the KSH family.

Unified convergence analyses for the MZ family have been given by Monteiro and Zhang [21] and Monteiro [16]. In the paper [21], iteration-complexity bounds are derived for the long-step primal-dual path-following method based on a subclass of the MZ family of search directions, which contains the NT and HRVW/KSH/M directions but not the AHO direction. In particular, it is shown that the corresponding algorithms based on the NT and the HRVW/KSH/M directions perform $\mathcal{O}(nL)$ and $\mathcal{O}(n^{3/2}L)$ iterations, respectively, to reduce the duality gap by a factor of at least $2^{-\mathcal{O}(L)}$. (The $\mathcal{O}(n^{3/2}L)$ iteration-complexity bound for the HRVW/KSH/M directions was in fact obtained earlier by Monteiro [15].) More recently, Monteiro [16] proves the polynomiality of the short-step primal-dual path-following algorithm and the Mizuno–Todd–Ye predictor-corrector–type algorithm based on any member of the MZ family, thus obtaining as a by-product the important result that Frobenius-norm–type algorithms based on the AHO direction are polynomially convergent.

Unified analyses for the KSH family of directions are provided in Kojima, Shindoh, and Hara [11] and Monteiro and Tsuchiya [19]. The paper [11] introduces the KSH family and establishes: (1) the polynomiality of the short-step path-following (feasible) method based on the two KSH/HRVW/M directions (both members of the KSH family); and (2) the polynomiality of a potential reduction (feasible and infeasible) algorithm based on *any* direction of the KSH family. Using techniques developed in Monteiro [16], the paper [19] extends the result (1) above to any direction of the KSH family. It also proves polynomial convergence of a Mizuno–Todd–Ye predictor-corrector-type algorithm for semidefinite linear complementarity problems based on the whole KSH family.

This paper considers primal-dual path-following methods for SDP based on the Newton direction for the symmetric central path equation

$$X^{1/2}SX^{1/2} - \mu I = 0. \tag{1.2}$$

This pure Newton direction is quite natural in view of the fact that the neighborhoods of the central path used to develop polynomially convergent algorithms are all based on the eigenvalues of the left-hand side of (1.2). (We use the qualifier "pure Newton" for those directions that are Newton directions with respect to a central path equation of the form $\Phi(X,S) = \mu I$, where the map $\Phi(\cdot,\cdot)$ is independent of the current iterate or any parameter.) In contrast, these neighborhoods have no connection with the eigenvalues of the left-hand side of the central path equation $XS + SX - \mu I = 0$ used to derive the AHO direction. Even though it is possible to define central path neighborhoods based on the eigenvalues of $XS + SX$, primal-dual path-following methods based on these neighborhoods are not known to be polynomially convergent. The polynomial convergence result obtained in [16] for the short-step path-following method using the AHO direction is based on the Frobenius norm neighborhood defined in terms of the left-hand side of (1.2).

We consider two primal-dual SDP algorithms based on the above Newton direction: (1) a short-step path-following method based on the Frobenius norm neighborhood; and (2) a semilong-step path-following method based on the operator norm neighborhood, which in terms of the eigenvalues of $X^{1/2}SX^{1/2}$ is equivalent to the infinity norm neighborhood for LP. We establish that algorithms (1) and (2) have iteration-complexity bounds of $\mathcal{O}(\sqrt{n}L)$ and $\mathcal{O}(nL)$, respectively, to reduce the duality gap by a factor of $2^{-\mathcal{O}(L)}$. It should be noted that nothing is known regarding the polynomial convergence of the semilong-step path-following algorithm using the AHO direction.

We also introduce a family of search directions which consists of the Newton directions applied to all the central path equations of the form

$$(PXP^T)^{1/2}(P^{-T}SP^{-1})(PXP^T)^{1/2} - \mu I = 0,$$

where P is a nonsingular matrix. We argue that this new family, referred to as the MT family, is related to the above Newton direction in the same way as the MZ family is related to the AHO direction, and we show that the iteration-complexity bounds of algorithms (1) and (2) above extend to any member of the MT family. Finally, we show that the long-step path-following method based on a subclass of the MT family, called the MT* subclass, has $O(n^{3/2}L)$ iteration-complexity bound, and hence does not depend on the choice of the sequence of scaling matrices $\{P^k\}$. In contrast, the iteration-complexity bound obtained in Monteiro and Zhang [21] for the long-step path-following algorithm based on the MZ* subclass of the MZ family depends on a certain condition number determined by the choice of $\{P^k\}$. Like the MZ* subclass, the MT* subclass also contains both the NT direction and the HRVW/KSH/M directions.

This paper is organized as follows. In section 2, we introduce the SDP problem and the associated assumptions and derive the Newton direction for the central path equation (1.2). We also give some existence results for this Newton direction and state a generic primal-dual algorithm based on it. In section 3, we state and prove technical results which are used in the polynomial convergence analysis of section 4. In section 4, we establish the polynomiality of the short-step and the semilong-step path-following algorithms based on the Newton direction for (1.2). In section 5, we introduce the MT family of search directions and generalize the convergence analysis of the short-step and semilong-step algorithms of section 4 to any member of this family. In section 6, we introduce the MT* subclass of directions and give the convergence analysis of the long-step path-following algorithm based on these directions. Finally, we end the paper with some concluding remarks in section 7.

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 largest and smallest eigenvalue by $\lambda_{\max}[Q]$ and $\lambda_{\min}[Q]$, respectively. 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 $P \bullet Q \equiv \text{Tr } P^TQ$. 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.

2. The SDP problem and preliminary discussion. In this section, we describe the SDP problem considered in this paper, state our assumptions, and derive the Newton direction for the central path equation (1.2). We also give some existence results for this Newton direction and state a generic primal-dual algorithm based on it.

2.1. The SDP problem. This subsection describes the SDP problem and the corresponding assumptions. It also contains some notation and terminology that are used throughout our presentation.

We consider the SDP problem

$$(P) \qquad \min\{C \bullet X : A_i \bullet X = b_i,\ i = 1, \ldots, m,\ X \succeq 0\} \tag{2.1}$$

and its associated dual SDP problem

$$(D) \qquad \max\left\{ b^Ty : \sum_{i=1}^m y_iA_i + S = C,\ S \succeq 0 \right\}, \tag{2.2}$$

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 (2.1) and (2.2) is

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

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 (2.1) and (2.2) have optimal solutions X^* and (S^*, y^*) such that $C \bullet X^* = b^T y^*$; i.e., the optimal values of (2.1) and (2.2) 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 consists of all the solutions $(X, S, y) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+ \times \Re^m$ to the following optimality system:

$$XS = 0, \tag{2.3a}$$
$$\sum_{i=1}^m y_i A_i + S - C = 0, \tag{2.3b}$$
$$A_i \bullet X - b_i = 0, \qquad i = 1, \ldots, m, \tag{2.3c}$$

where (2.3a) is called the complementarity equation. It is well known that for every $\nu > 0$, the perturbed system

$$XS = \nu I, \tag{2.4a}$$
$$\sum_{i=1}^m y_i A_i + S - C = 0, \tag{2.4b}$$
$$A_i \bullet X - b_i = 0, \qquad i = 1, \ldots, m, \tag{2.4c}$$

has a unique solution, denoted (X_ν, S_ν, y_ν), and that the limit $\lim_{\nu \to 0}(X_\nu, S_\nu, y_\nu)$ exists and is a solution of (2.3) (e.g., see Kojima, Shindoh, and Hara [11]). The set of all solutions (X_ν, S_ν, y_ν) with $\nu > 0$ is known as the *central path.*

It is known that for each $V \in \mathcal{S}^n_+$, there exists a unique $U \in \mathcal{S}^n_+$ such that $U^2 = V$. The matrix U is called the square root of V and is denoted by $V^{1/2}$. Using the square root $X^{1/2}$, (2.4a) can be alternatively expressed in the following symmetric form:

$$X^{1/2} S X^{1/2} = \nu I, \quad (X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D). \tag{2.5}$$

The path-following algorithms studied in this paper are all based on one of the following three centrality measures of a point $(X, S) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+$:

$$d_F(X, S) \equiv \left\| X^{1/2} S X^{1/2} - \mu I \right\|_F = \left[\sum_{i=1}^n \left(\lambda_i[XS] - \mu\right)^2 \right]^{1/2},$$
$$d_\infty(X, S) \equiv \left\| X^{1/2} S X^{1/2} - \mu I \right\| = \max_{i=1,\ldots,n} \left|\lambda_i[XS] - \mu\right|,$$
$$d_{-\infty}(X, S) \equiv \left\| X^{1/2} S X^{1/2} - \mu I \right\|_{-\infty} = \max\left(0, \mu - \lambda_{\min}[XS]\right),$$

where $\mu \equiv (X \bullet S)/n = (\sum_{i=1}^n \lambda_i[XS])/n$, and $\| \cdot \|_{-\infty}$ is defined as

$$\|Q\|_{-\infty} \equiv \max\left(0, \lambda_{\max}[-Q]\right) \quad \text{for } Q \in \mathcal{S}^n.$$

Note that $\|\cdot\|_{-\infty}$ is a seminorm in the sense that it satisfies

$$\|\alpha Q\|_{-\infty} = \alpha \|Q\|_{-\infty}, \quad \|Q+R\|_{-\infty} \le \|Q\|_{-\infty} + \|R\|_{-\infty} \tag{2.6}$$

for every $Q, R \in \mathcal{S}^n$ and $\alpha > 0$. Clearly, we have

$$\|Q\|_{-\infty} \le \|Q\| \le \|Q\|_F, \tag{2.7}$$

and for $\gamma > 0$,

$$\lambda_{\min}[XS] \ge (1-\gamma)\mu \iff d_{-\infty}(X,S) \le \gamma\mu.$$

The short-step, semilong-step, and long-step path-following methods are based on the following central path neighborhoods, respectively:

$$\mathcal{N}_F(\gamma) \equiv \{(X,S,y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D) : d_F(X,S) \le \gamma\mu\}, \tag{2.8a}$$

$$\mathcal{N}_\infty(\gamma) \equiv \{(X,S,y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D) : d_\infty(X,S) \le \gamma\mu\}, \tag{2.8b}$$

$$\mathcal{N}_{-\infty}(\gamma) \equiv \{(X,S,y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D) : d_{-\infty}(X,S) \le \gamma\mu\}, \tag{2.8c}$$

where $\gamma > 0$ is a given constant.

2.2. The Newton direction and the generic algorithm. In this subsection, we derive the Newton direction for system (2.4b), (2.4c), and (2.5) and state a generic primal-dual method based on it. We end the subsection by giving some existence results for this Newton direction.

We start with the following technical result.

LEMMA 2.1. *For every $A \in \mathcal{S}^n_{++}$ and $H \in \mathcal{S}^n$, the equation*

$$AU + UA = H \tag{2.9}$$

has a unique solution $U \in \mathcal{S}^n$. Moreover, this solution satisfies

$$\|AU\|_F \le \|H\|_F/\sqrt{2}. \tag{2.10}$$

Proof. The first part of the lemma follows from the fact that the linear map $\Phi_A : \mathcal{S}^n \to \mathcal{S}^n$ defined by $\Phi_A(U) = AU + UA$ is an isomorphism. Indeed, since Φ_A has the same domain and codomain, it suffices to show that Φ_A is one-to-one, or equivalently that $AU + UA = 0$ implies $U = 0$. In turn, this last implication follows from the fact that any solution U of (2.9) satisfies (2.10) (simply set $H = 0$ in (2.10) to conclude that $U = 0$). To show the last claim, we square both sides of (2.9) to obtain

$$2\|AU\|_F^2 + 2\mathrm{Tr}\,[UAUA] = \|H\|_F^2.$$

Since $\mathrm{Tr}\,[UAUA] = \|A^{1/2}UA^{1/2}\|_F^2 \ge 0$, (2.10) follows. □

Throughout this paper, we denote the unique solution U of (2.9) by $\langle\langle H \rangle\rangle_A$.

LEMMA 2.2. *Let $\theta : \mathcal{S}^n_{++} \to \mathcal{S}^n_{++}$ denote the square root function $\theta(X) = X^{1/2}$. Then, θ is an analytic function, and*

$$\theta'(X)H = \langle\langle H \rangle\rangle_{X^{1/2}} \quad \text{for every } X \in \mathcal{S}^n_{++} \text{ and } H \in \mathcal{S}^n,$$

where $\theta'(X)$ is the derivative of θ at X and $\theta'(X)H$ is the linear map $\theta'(X)$ evaluated at H.

Proof. Observe that the inverse function of θ is the analytic function given by $\theta^{-1}(A) = A^2$ for $A \in \mathcal{S}^n_{++}$. Clearly, the derivative $(\theta^{-1})'(A)$ of θ^{-1} is equal to the function Φ_A defined in the proof of Lemma 2.1. Since Φ_A is an isomorphism for every $A \in \mathcal{S}^n_{++}$, it follows from the inverse function theorem that θ is analytic and

$$\theta'(X) = \left[(\theta^{-1})'(X^{1/2})\right]^{-1} = \Phi^{-1}_{X^{1/2}}.$$

Hence, $\theta'(X)H = \Phi^{-1}_{X^{1/2}}(H) = \langle\langle H\rangle\rangle_{X^{1/2}}$. □

Using Lemma 2.2, it is now easy to see that the Newton direction $(\Delta X, \Delta S, \Delta y)$ for system (2.5) is the solution of the following system of linear equations:

(2.11a)
$$\langle\langle \Delta X\rangle\rangle_{X^{1/2}} S X^{1/2} + X^{1/2} S \langle\langle \Delta X\rangle\rangle_{X^{1/2}} + X^{1/2}\Delta S X^{1/2} = H,$$

(2.11b) $$\sum_{i=1}^m \Delta y_i A_i + \Delta S = R,$$

(2.11c) $$A_i \bullet \Delta X = r_i, \quad i = 1, \ldots, m,$$

where

(2.12a) $$H \equiv \nu I - X^{1/2} S X^{1/2},$$

(2.12b) $$R \equiv C - \sum_{i=1}^m y_i A_i - S,$$

(2.12c) $$r_i \equiv b_i - A_i \bullet X, \quad i = 1, \ldots, m.$$

Let $U \equiv \langle\langle \Delta X\rangle\rangle_{X^{1/2}}$. Then, in terms of U, we can write (2.11a) as two equivalent equations:

(2.13) $$USX^{1/2} + X^{1/2}SU + X^{1/2}\Delta S X^{1/2} = \nu I - X^{1/2}SX^{1/2},$$

(2.14) $$UX^{1/2} + X^{1/2}U = \Delta X.$$

We next state the generic primal-dual feasible algorithm that will be studied in this paper.

ALGORITHM I.

Let $(X^0, S^0, y^0) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$, $\mu_0 \equiv (X^0 \bullet S^0)/n$ and set $k = 0$.

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

(1) Let $(X, S, y) = (X^k, S^k, y^k)$ and $\mu \equiv (X \bullet S)/n$;
(2) Choose a centrality parameter $\sigma = \sigma_k \in [0, 1]$;
(3) Compute the solution $(\Delta X^k, \Delta S^k, \Delta y^k)$ of system (2.11) with $H \equiv \sigma\mu I - X^{1/2}SX^{1/2}$ and $(R, r) = (0, 0)$;
(4) Choose a stepsize $\alpha_k > 0$ such that $(X^{k+1}, S^{k+1}, y^{k+1}) = (X^k, S^k, y^k) + \alpha_k(\Delta X^k, \Delta X^k, \Delta y^k) \in \mathcal{S}^n_{++}$;
(5) Set $\mu_{k+1} \equiv (X^{k+1} \bullet S^{k+1})/n$ and increment k by 1.

End

The complete specification of Algorithm I depends on the choices of the initial point (X^0, S^0, y^0) and the sequences $\{\sigma_k\}$ and $\{\alpha_k\}$. These elements will be specified later when we discuss specific instances of the above algorithm. In general, the initial iterate (X^0, S^0, y^0) is chosen within one of the neighborhoods (2.8a)–(2.8b), and the sequences $\{\sigma_k\}$ and $\{\alpha_k\}$ are chosen so that the subsequent iterates lie in the same neighborhood and converge to an optimal solution of (2.1) and (2.2).

The following lemma establishes some important bounds on the Newton direction (2.11) and yields as a consequence Theorem 2.4, which establishes the nonsingularity of system (2.11) for any $(X,S,y) \in \mathcal{N}_\infty(\gamma)$ for $\gamma \in (0, 1/\sqrt{2})$.

LEMMA 2.3. *Suppose that* $\gamma \in [0, 1/\sqrt{2})$ *and that* $(X,S,y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ *is such that* $d_\infty(X,S) \le \gamma\mu$. *If* $(\Delta X, \Delta S, \Delta y)$ *is a solution of* (2.11) *with* $(R,r) = (0,0)$ *and* $H \in \mathcal{S}^n$, *then*

$$\max\left\{\mu \left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\right\} \le \frac{\|H\|_F}{(1-\sqrt{2}\gamma)}. \tag{2.15}$$

Proof. Multiplying (2.14) on the left and on the right by $X^{-1/2}$ and using inequality (2.10) of Lemma 2.1, we conclude that

$$\left\|UX^{-1/2}\right\|_F \le \frac{\|X^{-1/2}\Delta X X^{-1/2}\|_F}{\sqrt{2}}. \tag{2.16}$$

Since $(R,r) = (0,0)$, it follows from (2.11b) and (2.11c) that

$$\Delta X \bullet \Delta S = 0. \tag{2.17}$$

By (2.11a) and (2.14), we have

$$\begin{aligned} \mu X^{-1/2}\Delta X X^{-1/2} + X^{1/2}\Delta S X^{1/2} = H - UX^{-1/2}(X^{1/2}SX^{1/2} - \mu I) \\ - (X^{1/2}SX^{1/2} - \mu I)X^{-1/2}U. \end{aligned}$$

Taking the Frobenius norm of both sides of this equality and using (2.16) and (2.17), we obtain

$$\begin{aligned} &\max\left\{\mu \left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\right\} \\ &\le \left(\mu^2 \left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F^2 + \left\|X^{1/2}\Delta S X^{1/2}\right\|_F^2\right)^{1/2} \\ &= \left\|H - UX^{-1/2}(X^{1/2}SX^{1/2} - \mu I) - (X^{1/2}SX^{1/2} - \mu I)X^{-1/2}U\right\|_F \\ &\le \|H\|_F + 2\left\|UX^{-1/2}\right\|_F \left\|X^{1/2}SX^{1/2} - \mu I\right\| \\ &\le \|H\|_F + \sqrt{2}\gamma\mu \left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \end{aligned} \tag{2.18}$$

which clearly implies that

$$\mu\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F \le \frac{\|H\|_F}{(1-\sqrt{2}\gamma)}.$$

Using this last inequality to bound the right-hand side of (2.18), we obtain (2.15). □

THEOREM 2.4. *If* $(X,S,y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ *is such that* $d_\infty(X,S) < \mu/\sqrt{2}$ *then, for every* $(H,R,r) \in \mathcal{S}^n \times \mathcal{S}^n \times \Re^m$, *system* (2.11) *has a unique solution.*

Proof. In terms of $(\Delta X, \Delta S, \Delta y)$, the left-hand side of system (2.11) is a linear function from the space $\mathcal{S}^n \times \mathcal{S}^n \times \Re^m$ into itself. The lemma easily follows from the fact that this linear map is an isomorphism. To prove this fact, it is sufficient to show that this map is one-to-one, or equivalently that $(\Delta X, \Delta S, \Delta y) = (0,0,0)$ is the only solution of system (2.11) with $(H, R, r) = (0,0,0)$. Indeed, it follows from Lemma 2.3 that $(\Delta X, \Delta S) = (0,0)$. Using the linear independence of the matrices A_i, $i = 1, \ldots, m$, we conclude that $\Delta y = 0$. ◻

Note that the above result holds for both feasible and infeasible points. In particular, it implies the well-definedness of the Newton direction (2.11) for any point in $\mathcal{N}_\infty(\gamma)$, where $\gamma < 1/\sqrt{2}$. By slightly modifying Lemma 2.3, it is possible to establish the nonsingularity of system (2.11) for any point $(X, S, y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ such that $\|X^{1/2}SX^{1/2} - \nu I\| \le \gamma\nu$ for some $\nu \in \Re$ and $\gamma < 1/\sqrt{2}$. This yields a larger region of points since ν is not constrained to be equal to μ.

3. Technical results. In this section, we develop technical results which will be used in section 4 to establish the polynomial convergence of two specific instances of Algorithm I, namely, the short-step and the semilong-step path-following algorithms. The main novelty of the analysis of this paper is the use of second- or third-order Taylor expansions to analyze the behavior of the centrality measure when a Newton step is taken (see Lemma 3.3).

Let $(X, S, y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ denote the current iterate and let $(\Delta X, \Delta S, \Delta y)$ denote the Newton direction for system (2.5) at the point (X, S, y), that is, the solution of (2.11) with $(R, r) = (0, 0)$ and $H = \sigma\mu I - X^{1/2}SX^{1/2}$, where $\mu \equiv (X \bullet S)/n$ and $\sigma \in [0, 1]$. Define

$$(X_\alpha, S_\alpha, y_\alpha) \equiv (X, S, y) + \alpha(\Delta X, \Delta S, \Delta y), \tag{3.1}$$

$$\mu(\alpha) \equiv \frac{X_\alpha \bullet S_\alpha}{n}, \tag{3.2}$$

$$\phi(\alpha) \equiv X_\alpha^{1/2} S_\alpha X_\alpha^{1/2} - \mu(\alpha) I. \tag{3.3}$$

LEMMA 3.1. *We have*

$$\mu(\alpha) = (1 - \alpha + \alpha\sigma)\mu.$$

Proof. By (3.1) and the fact that $\Delta X \bullet \Delta S = 0$, we have

$$X_\alpha \bullet S_\alpha = X \bullet S + \alpha\,(S \bullet \Delta X + X \bullet \Delta S). \tag{3.4}$$

Using (2.13), (2.14), and the fact that $\nu = \sigma\mu$, we obtain

$$\begin{aligned} S \bullet \Delta X + X \bullet \Delta S &= \mathrm{Tr}[S\Delta X + X\Delta S] \\ &= \mathrm{Tr}[S(UX^{1/2} + X^{1/2}U) + X\Delta S] \\ &= \mathrm{Tr}[X^{1/2}SU + USX^{1/2} + X^{1/2}\Delta S X^{1/2}] \\ &= \mathrm{Tr}[\sigma\mu I - X^{1/2}SX^{1/2}] \\ &= n\sigma\mu - X \bullet S. \end{aligned} \tag{3.5}$$

The lemma now follows by substituting this equality into (3.4) and using the relations (3.2) and $X \bullet S = n\mu$. ◻

To study how the centrality measures for the points $(X_\alpha, S_\alpha, y_\alpha)$ vary, we will use either the second- or the third-order Taylor expansions of the function $\phi(\alpha)$. The following lemma gives expressions for the derivatives of this function.

LEMMA 3.2. *For every $\alpha \in \Re$ such that $(X_\alpha, S_\alpha) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, we have*

$$\phi'(\alpha) = U^{(1)}_\alpha S_\alpha X^{1/2}_\alpha + X^{1/2}_\alpha S_\alpha U^{(1)}_\alpha + X^{1/2}_\alpha \Delta S X^{1/2}_\alpha + (1-\sigma)\mu I, \tag{3.6}$$

$$\begin{aligned}\phi''(\alpha) = {} & U^{(2)}_\alpha S_\alpha X^{1/2}_\alpha + X^{1/2}_\alpha S_\alpha U^{(2)}_\alpha + 2U^{(1)}_\alpha \Delta S X^{1/2}_\alpha \\ & + 2X^{1/2}_\alpha \Delta S U^{(1)}_\alpha + 2U^{(1)}_\alpha S_\alpha U^{(1)}_\alpha,\end{aligned} \tag{3.7}$$

$$\begin{aligned}\phi'''(\alpha) = {} & U^{(3)}_\alpha S_\alpha X^{1/2}_\alpha + X^{1/2}_\alpha S_\alpha U^{(3)}_\alpha + 3U^{(2)}_\alpha \Delta S X^{1/2}_\alpha + 3X^{1/2}_\alpha \Delta S U^{(2)}_\alpha \\ & + 3U^{(2)}_\alpha S_\alpha U^{(1)}_\alpha + 3U^{(1)}_\alpha S_\alpha U^{(2)}_\alpha + 6U^{(1)}_\alpha \Delta S U^{(1)}_\alpha,\end{aligned} \tag{3.8}$$

where

$$U^{(1)}_\alpha \equiv \frac{d}{d\alpha}[X^{1/2}_\alpha], \quad U^{(2)}_\alpha \equiv \frac{d^2}{d\alpha^2}[X^{1/2}_\alpha], \quad U^{(3)}_\alpha \equiv \frac{d^3}{d\alpha^3}[X^{1/2}_\alpha],$$

and $U^{(1)}_\alpha$, $U^{(2)}_\alpha$, and $U^{(3)}_\alpha$ satisfy

$$X^{1/2}_\alpha U^{(1)}_\alpha + U^{(1)}_\alpha X^{1/2}_\alpha = \Delta X, \tag{3.9}$$

$$X^{1/2}_\alpha U^{(2)}_\alpha + U^{(2)}_\alpha X^{1/2}_\alpha = -2U^{(1)}_\alpha U^{(1)}_\alpha, \tag{3.10}$$

$$X^{1/2}_\alpha U^{(3)}_\alpha + U^{(3)}_\alpha X^{1/2}_\alpha = -3\left(U^{(1)}_\alpha U^{(2)}_\alpha + U^{(2)}_\alpha U^{(1)}_\alpha\right). \tag{3.11}$$

Also,

$$U^{(1)}_0 = U. \tag{3.12}$$

Proof. Expressions (3.6), (3.8), and (3.8) follow immediately from (3.3) and Lemma 3.1. Observe that $X^{1/2}_\alpha = \theta(X + \alpha\Delta X)$, where θ is the function defined in Lemma 2.2. It follows from this lemma that $U^{(1)}_\alpha = \theta'(X_\alpha)\Delta X = \langle\langle \Delta X \rangle\rangle_{X^{1/2}_\alpha}$, or equivalently that (3.9) holds. Expressions (3.10) and (3.11) now follow by differentiating (3.9) once and twice, respectively. Since, by Lemma 2.1, U is uniquely determined by (2.14), it follows from (3.9) with $\alpha = 0$ that $U = U^{(1)}_0$. ◻

The analysis of this paper strongly relies on the following simple result.

LEMMA 3.3. *For every $\alpha \in [0, 1]$, we have*

$$\|\phi(\alpha)\|_. \le (1-\alpha)\,\|\phi(0)\|_. + \frac{1}{2}\alpha^2 \sup_{\xi\in[0,\alpha]} \|\phi''(\xi)\|_F\,, \tag{3.13}$$

$$\|\phi(\alpha)\|_. \le (1-\alpha)\,\|\phi(0)\|_. + \frac{1}{2}\alpha^2\,\|\phi''(0)\|_. + \frac{1}{6}\alpha^3 \sup_{\xi\in[0,\alpha]} \|\phi'''(\xi)\|_F\,, \tag{3.14}$$

where $\|\cdot\|_.$ represents one of the norms $\|\cdot\|_F$ or $\|\cdot\|$ or the seminorm $\|\cdot\|_{-\infty}$.

Proof. By (3.6) with $\alpha = 0$, (3.12), (2.13), and (3.3), we have

$$\begin{aligned}\phi'(0) &= USX^{1/2} + X^{1/2}SU + X^{1/2}\Delta S X^{1/2} + (1-\sigma)\mu I \\ &= \sigma\mu I - X^{1/2}SX^{1/2} + (1-\sigma)\mu I \;=\; -\phi(0).\end{aligned} \tag{3.15}$$

The lemma now follows from this last equality, relations (2.6) and (2.7), and the two higher-order Taylor integral formulae:

$$\phi(\alpha) = \phi(0) + \alpha\phi'(0) + \alpha^2 \int_0^1 (1-t)\phi''(t\alpha)dt,$$

$$\phi(\alpha) = \phi(0) + \alpha\phi'(0) + \frac{1}{2}\alpha^2\phi''(0) + \alpha^3 \int_0^1 \frac{(1-t)^2}{2}\phi'''(t\alpha)dt. \qquad ◻$$

The analysis of section 4 is based on the inequality (3.13). Hence, in the remaining part of the section, we derive bounds for the second derivative $\phi''(\alpha)$. The other inequality (3.14) will be used in the analysis of section 6 to establish the polynomiality of the long-step path-following method based on the new family of directions introduced in section 5.

To simplify notation, we let

$$\widehat{\Delta X} \equiv X^{-1/2} \Delta X X^{-1/2}, \quad \widehat{\Delta S} \equiv S^{-1/2} \Delta S S^{-1/2}, \tag{3.16}$$

$$D_\alpha^X \equiv X^{1/2} X_\alpha^{-1/2}, \quad D_\alpha^S \equiv S^{1/2} S_\alpha^{-1/2}. \tag{3.17}$$

Observe that D_α^X and D_α^S are only well defined when $X_\alpha \in \mathcal{S}_{++}^n$ and $S_\alpha \in \mathcal{S}_{++}^n$, respectively.

LEMMA 3.4. *Let $\tau \in (0,1)$ be given. The following statements hold:*

(a) *If $\alpha > 0$ is such that $\alpha \|\widehat{\Delta X}\| \le \tau$, then $X_\alpha \in \mathcal{S}_{++}^n$, D_α^X is well defined and*

$$\max\left\{\left\|D_\alpha^X\right\|, \left\|(D_\alpha^X)^{-1}\right\|\right\} \le \frac{1}{\sqrt{1-\tau}}. \tag{3.18}$$

(b) *If $\alpha > 0$ is such that $\alpha \|\widehat{\Delta S}\| \le \tau$, then $S_\alpha \in \mathcal{S}_{++}^n$, D_α^S is well defined and*

$$\max\left\{\left\|D_\alpha^S\right\|, \left\|(D_\alpha^S)^{-1}\right\|\right\} \le \frac{1}{\sqrt{1-\tau}}. \tag{3.19}$$

Proof. We prove only (a), since the proof of (b) is similar. Let $\alpha > 0$ satisfying $\alpha \|\widehat{\Delta X}\| \le \tau$ be given. Clearly, $I + \alpha \widehat{\Delta X} \in \mathcal{S}_{++}^n$, and hence $X_\alpha = X^{1/2}(I + \alpha \widehat{\Delta X}) X^{1/2} \in \mathcal{S}_{++}^n$ and D_α^X is well defined. By (3.17), we have

$$\left[D_\alpha^X (D_\alpha^X)^T\right]^{-1} = X^{-1/2} X_\alpha X^{-1/2} = X^{-1/2}(X + \alpha \Delta X) X^{-1/2} = I + \alpha \widehat{\Delta X}.$$

Hence,

$$\left\|(D_\alpha^X)^{-1}\right\|^2 = \lambda_{\max}\left[I + \alpha \widehat{\Delta X}\right] \le 1 + \alpha \|\widehat{\Delta X}\| \le \frac{1}{1 - \alpha \|\widehat{\Delta X}\|} \le \frac{1}{1-\tau}$$

and

$$\left\|D_\alpha^X\right\|^2 = \lambda_{\max}\left[\left(I + \alpha \widehat{\Delta X}\right)^{-1}\right] = \frac{1}{\lambda_{\min}\left[I + \alpha \widehat{\Delta X}\right]} \le \frac{1}{1 - \alpha \|\widehat{\Delta X}\|} \le \frac{1}{1-\tau};$$

that is, (3.18) holds. $\square$

LEMMA 3.5. *Let $\tau \in (0,1)$ be given. If $\alpha > 0$ is such that $\alpha \max\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\} \le \tau$, then*

$$\left\|X_\alpha^{1/2} S_\alpha^{1/2}\right\| \le \frac{\|X^{1/2} S^{1/2}\|}{1-\tau}. \tag{3.20}$$

Proof. By (3.17), (3.18), and (3.19), we have

$$\begin{aligned}\left\|X_\alpha^{1/2} S_\alpha^{1/2}\right\| = \left\|(D_\alpha^X)^{-1} X^{1/2} S^{1/2} (D_\alpha^S)^{-T}\right\| &\le \left\|(D_\alpha^X)^{-1}\right\| \, \left\|(D_\alpha^S)^{-1}\right\| \, \left\|X^{1/2} S^{1/2}\right\| \\ &\le \frac{\|X^{1/2} S^{1/2}\|}{1-\tau}. \qquad \square\end{aligned}$$

LEMMA 3.6. *Let $\tau \in (0,1)$ be given. If $\alpha > 0$ is such that $\alpha\|\widehat{\Delta X}\| \le \tau$, then*

$$\left\|U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \le \frac{\|\widehat{\Delta X}\|_F}{\sqrt{2}(1-\tau)}, \tag{3.21}$$

$$\left\|U_\alpha^{(2)} X_\alpha^{-1/2}\right\|_F \le \frac{\|\widehat{\Delta X}\|_F^2}{\sqrt{2}(1-\tau)^2}, \tag{3.22}$$

$$\left\|U_\alpha^{(3)} X_\alpha^{-1/2}\right\|_F \le \frac{3}{\sqrt{2}} \frac{\|\widehat{\Delta X}\|_F^3}{(1-\tau)^3}. \tag{3.23}$$

Proof. Multiplying (3.9) on the left and on the right by $X_\alpha^{-1/2}$ and using inequality (2.10) of Lemma 2.1 and relation (3.18), we obtain (3.21) as follows:

$$\left\|U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \le \frac{1}{\sqrt{2}} \left\|X_\alpha^{-1/2} \Delta X X_\alpha^{-1/2}\right\|_F = \frac{1}{\sqrt{2}} \left\|(D_\alpha^X)^T \widehat{\Delta X} D_\alpha^X\right\|_F \le \frac{\|\widehat{\Delta X}\|_F}{\sqrt{2}(1-\tau)}.$$

Multiplying (3.10) on the left and on the right by $X_\alpha^{-1/2}$ and using inequality (2.10) of Lemma 2.1 and relation (3.21), we obtain (3.22) as follows:

$$\left\|U_\alpha^{(2)} X_\alpha^{-1/2}\right\|_F \le \sqrt{2} \left\|X_\alpha^{-1/2} U_\alpha^{(1)} U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \le \sqrt{2} \left\|(U_\alpha^{(1)}) X_\alpha^{-1/2}\right\|_F^2 \le \frac{\|\widehat{\Delta X}\|_F^2}{\sqrt{2}(1-\tau)^2}.$$

Finally, multiplying (3.11) on the left and on the right by $X_\alpha^{-1/2}$ and using inequality (2.10) of Lemma 2.1 and relations (3.21) and (3.22), we obtain (3.23) as follows:

$$\begin{aligned}\left\|U_\alpha^{(3)} X_\alpha^{-1/2}\right\|_F &\le \frac{3}{\sqrt{2}} \left\|X_\alpha^{-1/2} U_\alpha^{(1)} U_\alpha^{(2)} X_\alpha^{-1/2} + X_\alpha^{-1/2} U_\alpha^{(2)} U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \\ &\le \frac{6}{\sqrt{2}} \left\|U_\alpha^{(2)} X_\alpha^{-1/2}\right\|_F \left\|U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \le \frac{3}{\sqrt{2}} \frac{\|\widehat{\Delta X}\|_F^3}{(1-\tau)^3}. \qquad \square\end{aligned}$$

LEMMA 3.7. *Let constants $\tau \in (0,1)$ and $\gamma \in (0, 1/\sqrt{2})$ be given. Suppose that $(X,S,y) \in \mathcal{N}_\infty(\gamma)$ and that $\alpha > 0$ satisfies*

$$\alpha \max\left\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\right\} \le \tau.$$

Then,

$$\left\|U_\alpha^{(1)} \Delta S X_\alpha^{1/2}\right\|_F \le \frac{\|H\|_F^2}{\sqrt{2}(1-\tau)^2(1-\sqrt{2}\gamma)^2\mu}, \tag{3.24}$$

$$\left\|U_\alpha^{(1)} S_\alpha U_\alpha^{(1)}\right\|_F \le \frac{(1+\gamma)\|H\|_F^2}{2(1-\tau)^4(1-\sqrt{2}\gamma)^2\mu}, \tag{3.25}$$

$$\left\|U_\alpha^{(2)} S_\alpha X_\alpha^{1/2}\right\|_F \le \frac{(1+\gamma)\|H\|_F^2}{\sqrt{2}(1-\tau)^4(1-\sqrt{2}\gamma)^2\mu}. \tag{3.26}$$

Proof. Using Lemma 2.3, (3.18), and (3.21), we obtain

$$\begin{aligned}\left\|U_\alpha^{(1)} \Delta S X_\alpha^{1/2}\right\|_F &\le \left\|U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \left\|X_\alpha^{1/2} \Delta S X_\alpha^{1/2}\right\| \\ &\le \left\|U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F \|(D_\alpha^X)^{-1}\|^2 \left\|X^{1/2} \Delta S X^{1/2}\right\| \\ &\le \frac{\|\widehat{\Delta X}\|_F \|H\|_F}{\sqrt{2}(1-\tau)^2(1-\sqrt{2}\gamma)} \le \frac{\|H\|_F^2}{\sqrt{2}(1-\tau)^2(1-\sqrt{2}\gamma)^2\mu}.\end{aligned}$$

In addition, using Lemma 2.3, relations (3.20), (3.21), and (3.22), and the fact that $\left\|X^{1/2}S^{1/2}\right\|^2 \le (1+\gamma)\mu$ whenever $(X,S,y) \in \mathcal{N}_\infty(\gamma)$, we obtain

$$\left\|U_\alpha^{(1)} S_\alpha U_\alpha^{(1)}\right\|_F \le \left\|U_\alpha^{(1)} X_\alpha^{-1/2}\right\|_F^2 \left\|X_\alpha^{1/2} S_\alpha^{1/2}\right\|^2 \le \frac{\|\widehat{\Delta X}\|_F^2 \left\|X^{1/2}S^{1/2}\right\|^2}{2(1-\tau)^4}$$
$$\le \frac{(1+\gamma)\|H\|_F^2}{2(1-\tau)^4(1-\sqrt{2}\gamma)^2\mu},$$

and

$$\left\|U_\alpha^{(2)} S_\alpha X_\alpha^{1/2}\right\|_F \le \left\|U_\alpha^{(2)} X_\alpha^{-1/2}\right\|_F \left\|X_\alpha^{1/2} S_\alpha^{1/2}\right\|^2 \le \frac{\|\widehat{\Delta X}\|_F^2 \left\|X^{1/2}S^{1/2}\right\|^2}{\sqrt{2}(1-\tau)^4}$$
$$\le \frac{(1+\gamma)\|H\|_F^2}{\sqrt{2}(1-\tau)^4(1-\sqrt{2}\gamma)^2\mu}. \qquad \square$$

The following result gives the desired bound on the second derivative $\phi''(\alpha)$.

LEMMA 3.8. *Let a constant $\gamma \in (0, 1/\sqrt{2})$ be given. Suppose that $(X,S,y) \in \mathcal{N}_\infty(\gamma)$ and $\alpha > 0$ is such that*

$$\alpha \max\left\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\right\} \le \frac{1}{2}. \tag{3.27}$$

Then, $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ and

$$\|\phi''(\alpha)\|_F \le 80 \frac{\|H\|_F^2}{(1-\sqrt{2}\gamma)^2\mu}.$$

Proof. It is easy to see that (3.27) and the fact that $(X,S,y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ imply that $(X_\alpha, S_\alpha) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$. It follows from (3.8) and Lemma 3.7 with $\tau = 1/2$ that

$$\|\phi''(\alpha)\|_F \le 2\left\|U_\alpha^{(2)} S_\alpha X_\alpha^{1/2}\right\|_F + 2\left\|U_\alpha^{(1)} S_\alpha U_\alpha^{(1)}\right\|_F + 4\left\|U_\alpha^{(1)} \Delta S X_\alpha^{1/2}\right\|_F$$
$$\le \left(16\sqrt{2}(1+\gamma) + 16(1+\gamma) + 8\sqrt{2}\right) \frac{\|H\|_F^2}{(1-\sqrt{2}\gamma)^2\mu}$$
$$\le 80 \frac{\|H\|_F^2}{(1-\sqrt{2}\gamma)^2\mu},$$

where the last inequality follows from the fact that $\gamma < 1/\sqrt{2}$. $\square$

We end this section by stating without proof the following well-known result.

LEMMA 3.9. *The following statements hold:*

(a) *if $(X,S,y) \in \mathcal{N}_F(\gamma)$, then*

$$\|H\|_F \le \left[\gamma^2 + (1-\sigma)^2 n\right]^{1/2} \mu; \tag{3.28}$$

(b) *if $(X,S,y) \in \mathcal{N}_\infty(\gamma)$, then*

$$\|H\|_F \le \left[\gamma^2 + (1-\sigma)^2\right]^{1/2} \sqrt{n}\,\mu; \tag{3.29}$$

(c) *if $(X,S,y) \in \mathcal{N}_{-\infty}(\gamma)$, then*

$$\|\widehat{H}\|_F \le \left(1 - 2\sigma + \frac{\sigma^2}{1-\gamma}\right)^{1/2} \sqrt{n\mu}, \tag{3.30}$$

where $\widehat{H} \equiv H X^{-1/2} S^{-1/2}$.

4. Path-following algorithms based on the pure Newton direction. Based on the results developed in section 3, we now prove polynomiality of the short-step and the semilong-step path-following algorithms based on the pure Newton direction (2.11).

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

$$\frac{40\,(\gamma^2+\delta^2)}{(1-\sqrt{2}\gamma)^2} \le \left(1-\frac{\delta}{\sqrt{n}}\right)\gamma. \tag{4.1}$$

Suppose that $(X,S,y) \in \mathcal{N}_F(\gamma)$ and let $(\Delta X, \Delta S, \Delta y)$ denote the solution of system (2.11) with (H,R,r) given by (2.12), $\nu \equiv \sigma\mu$, $\sigma \equiv 1-\delta/\sqrt{n}$, and $\mu \equiv (X \bullet S)/n$. Then,

(a) $(X_1, S_1, y_1) \equiv (X+\Delta X, S+\Delta S, y+\Delta y) \in \mathcal{N}_F(\gamma)$;

(b) $X_1 \bullet S_1 = (1-\delta/\sqrt{n})(X \bullet S)$.

Proof. Statement (b) is an immediate consequence of Lemma 3.1 and the definition of σ. By Lemma 3.9(a) and the definition of σ, we have

$$\|H\|_F \le (\gamma^2+\delta^2)^{1/2}\mu. \tag{4.2}$$

Using Lemma 2.3, relations (4.1) and (4.2), and the fact that $\gamma < 1/\sqrt{2}$, we obtain

$$\begin{aligned}
\max\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\} &\le \max\left\{\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \left\|X^{-1/2}S^{-1/2}\right\|^2 \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\right\} \\
&\le \frac{1}{(1-\gamma)\mu}\max\left\{\mu\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\right\} \\
&\le \frac{\|H\|_F}{(1-\gamma)(1-\sqrt{2}\gamma)\mu} \;\le\; \frac{(\gamma^2+\delta^2)^{1/2}}{(1-\gamma)(1-\sqrt{2}\gamma)} \\
&\le \frac{1}{1-\gamma}\left(\frac{\gamma}{40}\right)^{1/2} \;\le\; \frac{1}{2}.
\end{aligned}$$

Hence, it follows from Lemma 3.8 and relations (4.1) and (4.2) that $(X_1, S_1, y_1) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ and

$$\sup_{\xi\in[0,1]} \|\phi''(\xi)\|_F \le 80\frac{\gamma^2+\delta^2}{(1-\sqrt{2}\gamma)^2}\,\mu \le 2\gamma\left(1-\frac{\delta}{\sqrt{n}}\right)\mu.$$

This inequality together with relation (3.13) and Lemma 3.1 (both with $\alpha = 1$) imply that

$$\left\|X_1^{1/2}S_1X_1^{1/2} - \mu(1)I\right\|_F = \|\phi(1)\| \le \frac{1}{2}\sup_{\xi\in[0,1]}\|\phi''(\xi)\|_F \le \gamma\left(1-\frac{\delta}{\sqrt{n}}\right)\mu = \gamma\mu(1).$$

Hence, $(X_1, S_1, y_1) \in \mathcal{N}_F(\gamma)$. ☐

As an immediate consequence of Theorem 4.1, we have the following polynomial convergence result for the short-step path-following algorithm obtained from Algorithm I by letting $(X^0, S^0, y^0) \in \mathcal{N}_F(\gamma)$, $\sigma_k = 1-\delta/\sqrt{n}$, and $\alpha_k = 1$ for every $k \ge 0$.

COROLLARY 4.2 (polynomiality of short-step path-following algorithm). *Suppose that $\gamma \in (0, 1/\sqrt{2})$ and $\delta \in (0,1)$ are constants satisfying (4.1). For Algorithm* I, *assume that $(X^0, S^0, y^0) \in \mathcal{N}_F(\gamma)$, $\sigma_k = 1-\delta/\sqrt{n}$, and $\alpha_k = 1$ for every $k \ge 0$. 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.*

We now consider the semilong-step path-following algorithm based on the neighborhood $\mathcal{N}_\infty(\gamma)$. It is the special case of Algorithm I for which (X^0, S^0, y^0) is selected in $\mathcal{N}_\infty(\gamma)$, and the sequences $\{\sigma_k\}$ and $\{\alpha_k\}$ are defined as

$$\sigma_k \equiv \bar{\sigma}, \tag{4.3a}$$

$$\alpha_k \equiv \max\left\{\alpha \in [0,1] : (X^k, S^k, y^k) + \alpha(\Delta X^k, \Delta S^k, \Delta y^k) \in \mathcal{N}_\infty(\gamma)\right\} \tag{4.3b}$$

for every $k \geq 0$, where $\bar{\sigma}$ is a prespecified constant in $(0,1)$.

THEOREM 4.3. *Suppose that* $(X,S,y) \in \mathcal{N}_\infty(\gamma)$ *for some given constant* $\gamma \in (0, 1/\sqrt{2})$, *and that* $(\Delta X, \Delta S, \Delta y)$ *denote the solution of* (2.11) *with* (H,R,r) *given by* (2.12), $\nu = \sigma\mu$, $\sigma \in (0,1)$, *and* $\mu \equiv (X \bullet S)/n$. *Let*

$$\tilde{\alpha} \equiv \frac{\sigma\gamma(1-\sqrt{2}\gamma)^2}{40n\left[\gamma^2 + (1-\sigma)^2\right]}. \tag{4.4}$$

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

(a) $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{N}_\infty(\gamma)$,

(b) $X_\alpha \bullet S_\alpha = (1-\alpha+\alpha\sigma)(X \bullet S)$.

Proof. Statement (b) is an immediate consequence of Lemma 3.1. Using Lemma 2.3, relations (3.29) and (4.4), and the fact that $\gamma < 1/\sqrt{2}$, we obtain

$$\begin{aligned}
\tilde{\alpha}\,\max\left\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\right\} &\leq \tilde{\alpha}\,\max\Big\{\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \\
&\qquad\qquad \left\|X^{-1/2}S^{-1/2}\right\|^2 \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\Big\} \\
&\leq \frac{\tilde{\alpha}}{(1-\gamma)\mu}\max\left\{\mu\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F, \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\right\} \\
&\leq \frac{\tilde{\alpha}\,\|H\|_F}{(1-\gamma)(1-\sqrt{2}\gamma)\mu} \;\leq\; \frac{\tilde{\alpha}\,[\gamma^2+(1-\sigma)^2]^{1/2}\sqrt{n}}{(1-\gamma)(1-\sqrt{2}\gamma)} \\
&\leq \frac{\sigma\gamma(1-\sqrt{2}\gamma)}{40\sqrt{n}(1-\gamma)[\gamma^2+(1-\sigma)^2]^{1/2}} \;\leq\; \frac{\sigma}{40\sqrt{n}} \;\leq\; \frac{1}{2}.
\end{aligned}$$

Hence, it follows from Lemma 3.8 and Lemma 3.9(b) and relation (4.4) that $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ for any $\alpha \in [0, \tilde{\alpha}]$, and

$$\tilde{\alpha}\sup_{\xi\in[0,\tilde{\alpha}]}\|\phi''(\xi)\|_F \leq 80\tilde{\alpha}\frac{\gamma^2+(1-\sigma)^2}{(1-\sqrt{2}\gamma)^2}\,n\mu = 2\sigma\gamma\mu.$$

This inequality together with (3.14) and Lemma 3.1 imply that for every $\alpha \in [0, \tilde{\alpha}]$,

$$\begin{aligned}
\left\|X_\alpha^{1/2}S_\alpha X_\alpha^{1/2} - \mu(\alpha)I\right\| = \|\phi(\alpha)\| \;&\leq\; (1-\alpha)\|\phi(0)\| + \frac{1}{2}\alpha^2\sup_{\xi\in[0,\alpha]}\|\phi''(\xi)\|_F \\
&\leq (1-\alpha)\gamma\mu + \frac{1}{2}\alpha\tilde{\alpha}\sup_{\xi\in[0,\tilde{\alpha}]}\|\phi''(\xi)\|_F \\
&\leq (1-\alpha)\gamma\mu + \alpha\sigma\gamma\mu \;=\; \gamma\mu(\alpha).
\end{aligned}$$

Hence, $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{N}_\infty(\gamma)$ for every $\alpha \in [0, \tilde{\alpha}]$; that is, (a) holds. ∎

As an immediate consequence of Theorem 4.3, we have the following polynomial convergence result for the semilong-step path-following algorithm based on the pure Newton direction (2.11).

COROLLARY 4.4 (polynomiality of semilong-step path-following algorithm). *Let constants $\gamma \in (0, 1/\sqrt{2})$ and $\bar{\sigma} \in (0,1)$ be given. For Algorithm I, assume that $(X^0, S^0, y^0) \in \mathcal{N}_\infty(\gamma)$ and that the sequences $\{\sigma_k\}$ and $\{\alpha_k\}$ are chosen according to (4.3). Then, the sequence of iterates $\{(X^k, S^k, y^k)\} \subset \mathcal{N}_\infty(\gamma)$ generated by Algorithm I satisfies $X^k \bullet S^k \leq (1-\bar{\eta})^k (X^0 \bullet S^0)$ for all $k \geq 0$, where*

$$\bar{\eta} \equiv \frac{\bar{\sigma}(1-\bar{\sigma})\gamma(1-\sqrt{2}\gamma)^2}{40n\left[\gamma^2 + (1-\bar{\sigma})^2\right]}.$$

Moreover, if the quantity $\max\{\gamma^{-1}, (1-\sqrt{2}\gamma)^{-1}, \bar{\sigma}^{-1}, (1-\bar{\sigma})^{-1}\}$ is independent of n, then the method terminates in at most $\mathcal{O}(nL)$ iterations.

5. A family of "scaled" Newton directions. In this section we introduce a new family of search directions which arises by computing the Newton direction (2.11) with respect to a scaled problem and mapping the direction back to the original space. Each direction of the family is then associated with the scaling matrix chosen to construct the scaled problem.

For the purpose of simplifying the notation in this and the next section, we assume that the variables for the original primal and dual problems are now $\widetilde{X}$ and $(\widetilde{S}, \widetilde{y})$ and that their associated data are $\widetilde{C} \in \mathcal{S}^n$, $\widetilde{A}_i \in \mathcal{S}^n$, $i = 1, \ldots, m$, and $\widetilde{b} = (\widetilde{b}_1, \ldots, \widetilde{b}_m) \in \Re^m$; that is, we assume that these problems are

$$(\widetilde{P}) \qquad \min\{\widetilde{C} \bullet \widetilde{X} : \widetilde{A}_i \bullet \widetilde{X} = \widetilde{b}_i,\ i = 1, \ldots, m,\ \widetilde{X} \succeq 0\},$$

$$(\widetilde{D}) \qquad \max\left\{\widetilde{b}^T \widetilde{y} : \sum_{i=1}^m \widetilde{y}_i \widetilde{A}_i + \widetilde{S} = \widetilde{C},\ \widetilde{S} \succeq 0\right\}.$$

Given a nonsingular matrix $\widetilde{P}$, consider the following change of variables:

$$X \equiv \widetilde{P}\widetilde{X}\widetilde{P}^T, \qquad (S, y) \equiv (\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}, \widetilde{y}). \tag{5.1}$$

Letting

$$C \equiv \widetilde{P}^{-T}\widetilde{C}\widetilde{P}^{-1}, \qquad (A_i, b_i) \equiv (\widetilde{P}^{-T}\widetilde{A}_i\widetilde{P}^{-1}, \widetilde{b}_i) \ \text{ for } i = 1, \ldots, m,$$

problems $(\widetilde{P})$ and $(\widetilde{D})$ can be written in terms of these new variables as problems (P) and (D) of section 2. It can be easily verified that if (X, S, y) and $(\widetilde{X}, \widetilde{S}, \widetilde{y})$ in $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ are related according to (5.1), then $d_F(\widetilde{X}, \widetilde{S}) = d_F(X, S)$, $d_\infty(\widetilde{X}, \widetilde{S}) = d_\infty(X, S)$, $d_{-\infty}(\widetilde{X}, \widetilde{S}) = d_{-\infty}(X, S)$. Letting $\widetilde{\mathcal{N}}_F(\gamma)$, $\widetilde{\mathcal{N}}_\infty(\gamma)$, $\widetilde{\mathcal{N}}_{-\infty}(\gamma)$ denote the neighborhoods associated with the pair of problems $(\widetilde{P}, \widetilde{D})$, the above observation immediately implies that

$$(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{N}}_F(\gamma) \iff (X, S, y) \in \mathcal{N}_F(\gamma), \tag{5.2a}$$

$$(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{N}}_\infty(\gamma) \iff (X, S, y) \in \mathcal{N}_\infty(\gamma), \tag{5.2b}$$

$$(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{N}}_{-\infty}(\gamma) \iff (X, S, y) \in \mathcal{N}_{-\infty}(\gamma). \tag{5.2c}$$

Moreover, if $(\widetilde{X}_\nu, \widetilde{S}_\nu, \widetilde{y}_\nu)$ denote the point on the central path with parameter $\nu > 0$ for the pair $(\widetilde{P}, \widetilde{D})$, then $(X_\nu, S_\nu, y_\nu) = (\widetilde{P}\widetilde{X}_\nu\widetilde{P}^T, \widetilde{P}^{-T}\widetilde{S}_\nu\widetilde{P}^{-1}, \widetilde{y}_\nu)$.

The matrix $\widetilde{P}$ also determines a scaled Newton direction (with parameter $\sigma > 0$) as follows. An interior feasible point $(\widetilde{X}, \widetilde{S}, \widetilde{y})$ for $(\widetilde{P}, \widetilde{D})$ determines an interior feasible point (X, S, y) for (P, D) as in (5.1). At the scaled point (X, S, y), the pure Newton direction (2.11) is computed and the resulting direction $(\Delta X, \Delta S, \Delta y)$ is mapped back into the original space to yield the scaled Newton direction $(\Delta\widetilde{X}, \Delta\widetilde{S}, \Delta\widetilde{y})$ as follows:

$$(\Delta\widetilde{X}, \Delta\widetilde{S}, \Delta\widetilde{y}) \equiv (\widetilde{P}^{-1}\Delta X\widetilde{P}^{-T}, \widetilde{P}^T\Delta S\widetilde{P}, \Delta y). \tag{5.3}$$

Hence, $(\Delta\widetilde{X}, \Delta\widetilde{S}, \Delta\widetilde{y})$ is a solution of

$$\begin{aligned} \nu I - X^{1/2}SX^{1/2} &= \langle\langle \widetilde{P}\Delta\widetilde{X}\widetilde{P}^T\rangle\rangle_{X^{1/2}} SX^{1/2} + X^{1/2}S\langle\langle \widetilde{P}\Delta\widetilde{X}\widetilde{P}^T\rangle\rangle_{X^{1/2}} \\ &\quad + X^{1/2}\widetilde{P}^{-T}\Delta\widetilde{S}\widetilde{P}^{-1}X^{1/2}, \\ \widetilde{C} - \sum_{i=1}^m \widetilde{y}_i\widetilde{A}_i - \widetilde{S} &= \sum_{i=1}^m \Delta\widetilde{y}_i\widetilde{A}_i + \Delta\widetilde{S}, \\ \widetilde{b}_i - \widetilde{A}_i \bullet \widetilde{X} &= \widetilde{A}_i \bullet \Delta\widetilde{X}, \quad i = 1, \ldots, m, \end{aligned} \tag{5.4}$$

where $X \equiv \widetilde{P}\widetilde{X}\widetilde{P}^T$ and $S \equiv \widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}$.

Observe that the scaled Newton direction at the point $(\widetilde{X}, \widetilde{S}, \widetilde{y})$ depends on $\widetilde{P}$, and as $\widetilde{P}$ varies over the set of nonsingular matrices, we obtain a family of search directions, which we refer to as the MT family. Several observations are in order with respect to this family. It includes both the NT direction and the two HRVW/KSH/M directions. Indeed, if $\widetilde{P} = \widetilde{X}^{-1/2}$ then $X = I$, $S = \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2}$, $\langle\langle \widetilde{P}\Delta\widetilde{X}\widetilde{P}^T\rangle\rangle_{X^{1/2}} = (\widetilde{X}^{-1/2}\Delta\widetilde{X}\widetilde{X}^{-1/2})/2$, and the first equation of system (5.4) reduces to

$$\nu I - \widetilde{X}^{1/2}\widetilde{S}\widetilde{X}^{1/2} = \frac{1}{2}\left(\widetilde{X}^{-1/2}\Delta\widetilde{X}\widetilde{S}\widetilde{X}^{1/2} + \widetilde{X}^{1/2}\widetilde{S}\Delta\widetilde{X}\widetilde{X}^{-1/2}\right) + \widetilde{X}^{1/2}\Delta\widetilde{S}\widetilde{X}^{1/2},$$

which corresponds to the HRVW/KSH/M dual direction. If $\widetilde{P} = \widetilde{S}^{1/2}$, then $X = \widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2}$, $S = I$,

$$\begin{aligned} &\langle\langle \widetilde{P}\Delta\widetilde{X}\widetilde{P}^T\rangle\rangle_{X^{1/2}} SX^{1/2} + X^{1/2}S\langle\langle \widetilde{P}\Delta\widetilde{X}\widetilde{P}^T\rangle\rangle_{X^{1/2}} \\ &= \langle\langle \widetilde{S}^{1/2}\Delta\widetilde{X}\widetilde{S}^{1/2}\rangle\rangle_{X^{1/2}} X^{1/2} + X^{1/2}\langle\langle \widetilde{S}^{1/2}\Delta\widetilde{X}\widetilde{S}^{1/2}\rangle\rangle_{X^{1/2}} = \widetilde{S}^{1/2}\Delta\widetilde{X}\widetilde{S}^{1/2}, \end{aligned}$$

and the first equation of system (5.4) becomes

$$\nu I - \widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2} = \widetilde{S}^{1/2}\Delta\widetilde{X}\widetilde{S}^{1/2} + (\widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2})^{1/2}\widetilde{S}^{-1/2}\Delta\widetilde{S}\widetilde{S}^{-1/2}(\widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2})^{1/2},$$

which is the equation corresponding to the NT direction. After the release of the first version of this paper, Todd [33] showed that the HRVW/KSH/M direction is also in the MT family and can be obtained by taking $\widetilde{P} = (\widetilde{S}\widetilde{X}\widetilde{S})^{1/2}$ so that $SXS = I$. Needless to say, we observe that if $\widetilde{P} = I$ then system (5.4) reduces to system (2.11), and hence it corresponds to the (pure) Newton direction considered in section 2.

Another possible choice is to take $\widetilde{P}$ to be the NT scaling matrix satisfying $\widetilde{P}\widetilde{X}\widetilde{P}^T = \widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}$, so that $X = S$ holds. Like the NT and HRVW/KSH/M directions, the resulting direction can be shown to have the scaling invariance property discussed in [34]. This direction is referred to as the MTW direction in [33].

The results obtained in section 2 for the pure Newton direction (2.11)–(2.12) can be extended to the whole MT family due to the fact that any member of this family reduces to the Newton direction (2.11)–(2.12) in the scaled space and the fact that the duality gap and the centrality measures remain invariant. In what follows, we summarize these results.

COROLLARY 5.1. *If* $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ *is such that* $d_\infty(\widetilde{X}, \widetilde{S}) < \widetilde{\mu}/\sqrt{2}$ *where* $\widetilde{\mu} \equiv (\widetilde{X} \bullet \widetilde{S})/n$, *then system* (5.4) *has a unique solution.*

Proof. Due to the invariance of the duality gap and the centrality measure $d_\infty(\cdot, \cdot)$, the assumption implies that $d_\infty(X, S) < \mu/\sqrt{2}$. Since the direction $(\Delta X, \Delta S, \Delta y) \equiv (\widetilde{P} \Delta \widetilde{X} \widetilde{P}^T, \widetilde{P}^{-T} \Delta \widetilde{S} \widetilde{P}^{-1}, \Delta \widetilde{y})$ is a solution of (2.11), the corollary follows immediately from Theorem 2.4. □

The generic primal-dual feasible algorithm based on the MT family of directions is stated next.

ALGORITHM II.

Let $(\widetilde{X}^0, \widetilde{S}^0, \widetilde{y}^0) \in \mathcal{F}^0(\widetilde{P}) \times \mathcal{F}^0(\widetilde{D})$, $\widetilde{\mu}_0 \equiv (\widetilde{X}^0 \bullet \widetilde{S}^0)/n$ and set $k = 0$.

Repeat until $\widetilde{\mu}_k \leq 2^{-L} \widetilde{\mu}_0$, **do**

(1) Let $(\widetilde{X}, \widetilde{S}, \widetilde{y}) = (\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k)$ and $\widetilde{\mu} \equiv (\widetilde{X} \bullet \widetilde{S})/n$;

(2) Choose a centrality parameter $\sigma = \sigma_k \in [0, 1]$ and a nonsingular matrix $\widetilde{P} = P^k$;

(3) Compute the solution $(\Delta \widetilde{X}^k, \Delta \widetilde{S}^k, \Delta \widetilde{y}^k)$ of system (5.4) with $X \equiv \widetilde{P} \widetilde{X} \widetilde{P}^T$, $S \equiv \widetilde{P}^{-T} \widetilde{S} \widetilde{P}^{-1}$, and $\nu \equiv \sigma \widetilde{\mu}$;

(4) Choose a stepsize $\alpha_k > 0$ such that $(\widetilde{X}^{k+1}, \widetilde{S}^{k+1}, \widetilde{y}^{k+1}) = (\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k) + \alpha_k (\Delta \widetilde{X}^k, \Delta \widetilde{X}^k, \Delta \widetilde{y}^k) \in \mathcal{S}^n_{++}$;

(5) Set $\widetilde{\mu}_{k+1} \equiv (\widetilde{X}^{k+1} \bullet \widetilde{S}^{k+1})/n$ and increment k by 1.

End

The following two results follow immediately from Theorems 4.1 and 4.3, the equivalences in (5.2), and the invariance of the duality gap and the centrality measures.

COROLLARY 5.2 (polynomiality of short-step path-following algorithm for the MT family). *Suppose that* $\gamma \in (0, 1/\sqrt{2})$ *and* $\delta \in (0, 1)$ *are constants satisfying* (4.1). *For Algorithm* II, *assume that* $(\widetilde{X}^0, \widetilde{S}^0, \widetilde{y}^0) \in \widetilde{\mathcal{N}}_F(\gamma)$, $\sigma_k = 1 - \delta/\sqrt{n}$, *and* $\alpha_k = 1$ *for every* $k \geq 0$. *Then, every iterate* $(\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k)$ *generated by Algorithm* II *is in the neighborhood* $\widetilde{\mathcal{N}}_F(\gamma)$ *and satisfies* $\widetilde{X}^k \bullet \widetilde{S}^k = (1 - \delta/\sqrt{n})^k (\widetilde{X}^0 \bullet \widetilde{S}^0)$. *Moreover, Algorithm* II *terminates in at most* $\mathcal{O}(\sqrt{n} L)$ *iterations.*

COROLLARY 5.3 (polynomiality of semilong-step path-following algorithm for the MT family). *Let constants* $\gamma \in (0, 1/\sqrt{2})$ *and* $\bar{\sigma} \in (0, 1)$ *be given. For Algorithm* II, *assume that* $(\widetilde{X}^0, \widetilde{S}^0, \widetilde{y}^0) \in \mathcal{N}_\infty(\gamma)$ *and that the sequences* $\{\sigma_k\}$ *and* $\{\alpha_k\}$ *are chosen according to*

$$\sigma_k = \bar{\sigma},$$
$$\alpha_k = \max \left\{ \alpha \in [0, 1] : (\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k) + \alpha (\Delta \widetilde{X}^k, \Delta \widetilde{S}^k, \Delta \widetilde{y}^k) \in \widetilde{\mathcal{N}}_\infty(\gamma) \right\}.$$

Then, the sequence of iterates $\{(\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k)\} \subset \mathcal{N}_\infty(\gamma)$ *generated by Algorithm* II *satisfies* $\widetilde{X}^k \bullet \widetilde{S}^k \leq (1 - \bar{\eta})^k (\widetilde{X}^0 \bullet \widetilde{S}^0)$ *for all* $k \geq 0$, *where*

$$\bar{\eta} \equiv \frac{\bar{\sigma}(1 - \bar{\sigma}) \gamma (1 - \sqrt{2}\gamma)^2}{40 n \left[\gamma^2 + (1 - \bar{\sigma})^2\right]}.$$

Moreover, if the quantity $\max\{\gamma^{-1}, (1 - \sqrt{2}\gamma)^{-1}, \bar{\sigma}^{-1}, (1 - \bar{\sigma})^{-1}\}$ *is independent of* n, *then the method terminates in at most* $\mathcal{O}(nL)$ *iterations.*

6. Long-step method based on a subclass of the MT family. In this section we consider a subclass of the MT family whose members are well defined at every point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$. Moreover, we establish an $\mathcal{O}(n^{3/2}L)$ iteration-complexity bound for a long-step path-following feasible algorithm based on this subclass of the MT family. The analysis of this section is based on the third-order derivative inequality (3.14) and hence is more involved than the one presented in sections 3 and 4. It is possible to derive polynomial convergence for the long-step path-following algorithm using second-order derivative inequality (3.13), but the iteration-complexity bound obtained is worse than the $\mathcal{O}(n^{3/2}L)$ bound obtained using (3.14).

We first describe the subclass of the MT family, which we refer to as the MT* family. The members of the MT* family at a point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ consists of all the members of the MT family corresponding to those scaling matrices $\widetilde{P}$ satisfying

$$X^{1/2}S + SX^{1/2} = (\widetilde{P}\widetilde{X}\widetilde{P}^T)^{1/2}(\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}) + (\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1})(\widetilde{P}\widetilde{X}\widetilde{P}^T)^{1/2} \succ 0. \tag{6.1}$$

The next two results imply that any member of the MT* family is well defined for any point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$.

LEMMA 6.1. *Suppose that $(X, S, y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ is such that $X^{1/2}S + SX^{1/2} \succ 0$. If $(\Delta X, \Delta S, \Delta y)$ is a solution of system* (2.11) *with $(R, r) = (0, 0)$ and $H \in \mathcal{S}^n$, then*

$$\left\| S^{1/2}U \right\|_F \le \|\widehat{H}\|_F, \tag{6.2}$$

$$\left\| X^{-1/2}\Delta X X^{-1/2} \right\|_F \le \frac{2}{\sqrt{\lambda_{\min}}}\|\widehat{H}\|_F, \tag{6.3}$$

$$\left\| X^{1/2}\Delta S S^{-1/2} \right\|_F \le 3\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2}\|\widehat{H}\|_F, \tag{6.4}$$

where $\lambda_{\min} \equiv \lambda_{\min}[XS]$, $\lambda_{\max} \equiv \lambda_{\max}[XS]$, $U \equiv \langle\langle \Delta X \rangle\rangle_{X^{1/2}}$, and $\widehat{H} \equiv HX^{-1/2}S^{-1/2}$.

Proof. It follows from $(R, r) = (0, 0)$, (2.11b), and (2.11c) that $\Delta X \bullet \Delta S = \mathrm{Tr}\,(\Delta X \Delta S) = 0$, which together with (2.14) imply that

$$\mathrm{Tr}\,(UX^{1/2}\Delta S) = 0. \tag{6.5}$$

Multiplying (2.13) on the left by U and on the right by $X^{-1/2}$, taking the trace of both sides of the equality, and using (6.5), we obtain

$$\mathrm{Tr}\,(U^2S) + \mathrm{Tr}\,\left(UX^{1/2}SUX^{-1/2}\right) = \mathrm{Tr}\,\left(UHX^{-1/2}\right). \tag{6.6}$$

Since $UX^{-1/2}U \succeq 0$ and, by assumption, $X^{1/2}S + SX^{1/2} \succ 0$, we have

$$\begin{aligned} \mathrm{Tr}\,\left(UX^{1/2}SUX^{-1/2}\right) &= \mathrm{Tr}\,\left(UX^{-1/2}UX^{1/2}S\right) \\ &= \frac{1}{2}\mathrm{Tr}\,\left[UX^{-1/2}U\left(X^{1/2}S + SX^{1/2}\right)\right] \ge 0. \end{aligned}$$

Relation (6.6) together with the last inequality and the fact that $\mathrm{Tr}\,U^2S = \|S^{1/2}U\|_F^2$ imply that

$$\|S^{1/2}U\|_F^2 \le \mathrm{Tr}\,\left(UHX^{-1/2}\right) = \mathrm{Tr}\,\left(S^{1/2}U\widehat{H}\right) \le \|S^{1/2}U\|_F\|\widehat{H}\|_F,$$

from which (6.2) immediately follows. To show (6.3), observe that by (2.14) we have

$$X^{-1/2}U + UX^{-1/2} = X^{-1/2}\Delta X X^{-1/2},$$

which together with (6.2) and the fact that $\|X^{-1/2}S^{-1/2}\|^2 = 1/\lambda_{\min}[XS]$ imply

$$\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F \le 2\|X^{-1/2}U\|_F \le 2\|X^{-1/2}S^{-1/2}\|\,\|S^{1/2}U\|_F \le \frac{2}{\sqrt{\lambda_{\min}}}\|\widehat{H}\|_F;$$

that is, (6.3) holds. To show (6.4), we multiply (2.13) on the right by $X^{-1/2}S^{-1/2}$ and rearrange to obtain

$$X^{1/2}\Delta S S^{-1/2} = \widehat{H} - US^{1/2} - X^{1/2}SUX^{-1/2}S^{-1/2}.$$

Taking the Frobenius norm of both sides of the last equality and using the triangle inequality, relation (6.2) and the fact that $\|X^{1/2}S^{1/2}\|^2 = \lambda_{\max}$ and $\|X^{-1/2}S^{-1/2}\|^2 = 1/\lambda_{\min}$, we obtain

$$\begin{aligned}\left\|X^{1/2}\Delta S S^{-1/2}\right\|_F &\le \|\widehat{H}\|_F + \|S^{1/2}U\|_F + \|X^{1/2}SUX^{-1/2}S^{-1/2}\|_F \\ &\le 2\|\widehat{H}\|_F + \|X^{1/2}S^{1/2}\|\,\|S^{1/2}U\|_F\|X^{-1/2}S^{-1/2}\| \\ &\le \left[2 + \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2}\right]\|\widehat{H}\|_F \;\le\; 3\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2}\|\widehat{H}\|_F. \qquad \square\end{aligned}$$

THEOREM 6.2. *If $(X, S, y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ is such that $X^{1/2}S + SX^{1/2} \succ 0$ then, for every $(H, R, r) \in \mathcal{S}^n \times \mathcal{S}^n \times \Re^m$, system (2.11) has exactly one solution. In particular, for any $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ and any nonsingular matrix $\widetilde{P} \in \Re^{n\times n}$ satisfying (6.1), system (5.4) has exactly one solution.*

Proof. The proof of the first part is analogous to that of Theorem 2.4. The only difference is that Lemma 6.1 should be invoked in place of Lemma 2.3. The second part follows from the fact that $(\Delta\widetilde{X}, \Delta\widetilde{S}, \Delta\tilde{y})$ is a solution of (5.4) if and only if $(\Delta X, \Delta S, \Delta y) \equiv (\widetilde{P}\Delta\widetilde{X}\widetilde{P}^T, \widetilde{P}^{-T}\Delta\widetilde{S}\widetilde{P}^{-1}, \Delta\tilde{y})$ is a solution of (2.11) with (H, R, r) given by (2.12). $\square$

LEMMA 6.3. *If $(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ is such that $X^{1/2}S + SX^{1/2} \succ 0$, then*

$$\|S^{1/2}U_0^{(2)}\| \le \frac{2}{\sqrt{\lambda_{\min}}}\|\widehat{H}\|_F^2.$$

Proof. Multiplying (3.10) on the left by $X^{-1/2}$ and on the right by $S^{1/2}$, setting $\alpha = 0$ and using (3.12), we obtain

$$U_0^{(2)}S^{1/2} + X^{-1/2}U_0^{(2)}X^{1/2}S^{1/2} = -2X^{-1/2}UUS^{1/2}. \tag{6.7}$$

Using the assumption that $X^{1/2}S + SX^{1/2} \succ 0$, we have

$$\begin{aligned}\left(U_0^{(2)}S^{1/2}\right) \bullet \left(X^{-1/2}U_0^{(2)}X^{1/2}S^{1/2}\right) &= \mathrm{Tr}\left(U_0^{(2)}X^{-1/2}U_0^{(2)}X^{1/2}S\right) \\ &= \frac{1}{2}\mathrm{Tr}\left[U_0^{(2)}X^{-1/2}U_0^{(2)}\left(X^{1/2}S + SX^{1/2}\right)\right] \;\ge\; 0.\end{aligned}$$

Taking the Frobenius norm of both sides of (6.7) and using the last inequality together with (6.2) and the fact that $\|X^{-1/2}S^{-1/2}\|^2 = 1/\lambda_{\min}$, we obtain

$$\begin{aligned}\|S^{1/2}U_0^{(2)}\|_F &\le \left(\|S^{1/2}U_0^{(2)}\|_F^2 + \|X^{-1/2}U_0^{(2)}X^{1/2}S^{1/2}\|_F^2\right)^{1/2}\\ &\le \|U_0^{(2)}S^{1/2} + X^{-1/2}U_0^{(2)}X^{1/2}S^{1/2}\|_F \;=\; 2\|X^{-1/2}UUS^{1/2}\|_F\\ &\le 2\|X^{-1/2}S^{-1/2}\|\,\|S^{1/2}U\|_F^2 \;\le\; \frac{2}{\sqrt{\lambda_{\min}}}\|\widehat{H}\|_F^2. \qquad \square\end{aligned}$$

LEMMA 6.4. *Let* $\tau \in (0,1)$ *be given. If* $\alpha > 0$ *is such that*

$$\alpha \max\left\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\right\} \le \tau,$$

then

$$\left\|U_\alpha^{(2)}\Delta S X_\alpha^{1/2}\right\|_F \le \frac{6\sqrt{2}}{(1-\tau)^3}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3, \tag{6.8}$$

$$\left\|U_\alpha^{(1)}\Delta S U_\alpha^{(1)}\right\|_F \le \frac{6}{(1-\tau)^3}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3, \tag{6.9}$$

$$\left\|U_\alpha^{(2)}S_\alpha U_\alpha^{(1)}\right\|_F \le \frac{4}{(1-\tau)^5}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3, \tag{6.10}$$

$$\left\|U_\alpha^{(3)}S_\alpha X_\alpha^{1/2}\right\|_F \le \frac{12\sqrt{2}}{(1-\tau)^5}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3. \tag{6.11}$$

Proof. Using (3.18), (3.21), (3.22), Lemma 6.1, and the fact that $\|X^{1/2}S^{1/2}\| = \sqrt{\lambda_{\max}}$, we obtain

$$\begin{aligned}\left\|U_\alpha^{(2)}\Delta S X_\alpha^{1/2}\right\|_F &\le \left\|U_\alpha^{(2)}X_\alpha^{-1/2}\right\|_F \left\|X_\alpha^{1/2}\Delta S X_\alpha^{1/2}\right\|\\ &\le \frac{\|\widehat{\Delta X}\|_F^2}{\sqrt{2}(1-\tau)^2}\left\|(D_\alpha^X)^{-1}\right\|^2 \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\\ &\le \frac{\|\widehat{\Delta X}\|_F^2}{\sqrt{2}(1-\tau)^3}\left\|X^{1/2}\Delta S S^{-1/2}\right\|_F \left\|X^{1/2}S^{1/2}\right\| \;\le\; \frac{6\sqrt{2}}{(1-\tau)^3}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3\end{aligned}$$

and

$$\begin{aligned}\left\|U_\alpha^{(1)}\Delta S U_\alpha^{(1)}\right\|_F &\le \left\|U_\alpha^{(1)}X_\alpha^{-1/2}\right\|_F^2 \left\|X_\alpha^{1/2}\Delta S X_\alpha^{1/2}\right\|\\ &\le \frac{\|\widehat{\Delta X}\|_F^2}{2(1-\tau)^2}\left\|(D_\alpha^X)^{-1}\right\|^2 \left\|X^{1/2}\Delta S X^{1/2}\right\|_F\\ &\le \frac{\|\widehat{\Delta X}\|_F^2}{2(1-\tau)^3}\left\|X^{1/2}\Delta S S^{-1/2}\right\|_F \left\|X^{1/2}S^{1/2}\right\| \;\le\; \frac{6}{(1-\tau)^3}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3.\end{aligned}$$

Also, using (3.20), (3.21), (3.22), (3.23), and Lemma 6.1, we obtain

$$\begin{aligned}\left\|U_\alpha^{(2)}S_\alpha U_\alpha^{(1)}\right\|_F &\le \left\|U_\alpha^{(2)}X_\alpha^{-1/2}\right\|_F \left\|U_\alpha^{(1)}X_\alpha^{-1/2}\right\|_F \left\|X_\alpha^{1/2}S_\alpha^{1/2}\right\|^2\\ &\le \frac{\|\widehat{\Delta X}\|_F^3\left\|X^{1/2}S^{1/2}\right\|^2}{2(1-\tau)^5} \;\le\; \frac{4}{(1-\tau)^5}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3\end{aligned}$$

and

$$\begin{aligned}\left\|U_\alpha^{(3)} S_\alpha X_\alpha^{1/2}\right\|_F &\le \left\|U_\alpha^{(3)} X_\alpha^{-1/2}\right\|_F \left\|X_\alpha^{1/2} S_\alpha^{1/2}\right\|^2 \\ &\le \frac{3\|\widehat{\Delta X}\|_F^3 \, \|X^{1/2}S^{1/2}\|^2}{\sqrt{2}(1-\tau)^5} \;\le\; \frac{12\sqrt{2}}{(1-\tau)^5}\frac{\lambda_{\max}}{\lambda_{\min}^{3/2}}\|\widehat{H}\|_F^3. \qquad \square\end{aligned}$$

LEMMA 6.5. *If $(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ is such that $X^{1/2}S + SX^{1/2} \succ 0$, then*

$$\|\phi''(0)\|_F \le 18 \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} \|\widehat{H}\|_F^2.$$

In addition, if $\alpha > 0$ is such that $\alpha \max\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\} \le 1/2$, then $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ and

$$\|\phi'''(\alpha)\|_F \le 3000 \frac{\lambda_{\max}}{\lambda_{\min}^{3/2}} \|\widehat{H}\|_F^3.$$

Proof. Using (3.8) with $\alpha = 0$, (3.12), (6.2), (6.4), and Lemma 6.3, we obtain

$$\begin{aligned}\|\phi''(0)\|_F &\le 2\|X^{1/2}SU_0^{(2)}\|_F + 2\|USU\|_F + 4\|X^{1/2}\Delta SU\|_F \\ &\le 2\|X^{1/2}S^{1/2}\| \, \|S^{1/2}U_0^{(2)}\|_F + 2\|S^{1/2}U\|_F^2 + 4\|X^{1/2}\Delta S S^{-1/2}\|_F\|S^{1/2}U\|_F \\ &\le 4\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} \|\widehat{H}\|_F^2 + 2\|\widehat{H}\|_F^2 + 12\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} \|\widehat{H}\|_F^2 \\ &\le 18\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} \|\widehat{H}\|_F^2.\end{aligned}$$

Since $\alpha \max\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\} \le 1/2$ and $(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$, we have $(X_\alpha, S_\alpha) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$. It follows from (3.8) and Lemma 6.4 with $\tau = 1/2$ that

$$\begin{aligned}\|\phi'''(\alpha)\|_F &\le 2\left\|U_\alpha^{(3)} S_\alpha X_\alpha^{1/2}\right\|_F + 6\left\|U_\alpha^{(2)} S_\alpha U_\alpha^{(1)}\right\|_F \\ &\quad + 6\left\|U_\alpha^{(2)} \Delta S X_\alpha^{1/2}\right\|_F + 6\left\|U_\alpha^{(1)} \Delta S U_\alpha^{(1)}\right\|_F \\ &\le \left(768\sqrt{2} + 768 + 288\sqrt{2} + 288\right) \frac{\lambda_{\max}}{\lambda_{\min}^{3/2}} \|\widehat{H}\|_F^3 \;\le\; 3000 \frac{\lambda_{\max}}{\lambda_{\min}^{3/2}} \|\widehat{H}\|_F^3. \qquad \square\end{aligned}$$

THEOREM 6.6. *Let $\gamma, \sigma \in (0, 1)$ be given. Suppose that $(X, S, y) \in \mathcal{N}_{-\infty}(\gamma)$ satisfies $X^{1/2}S + SX^{1/2} \succ 0$ and $(\Delta X, \Delta S, \Delta y)$ is the solution of (2.11) with (H, R, r) given by (2.12) and $\mu \equiv (X \bullet S)/n$. Let*

$$\hat{\alpha} \equiv \frac{\sigma\gamma(1-\gamma)^{1/2}}{30n^{3/2}\zeta}, \tag{6.12}$$

where $\zeta \equiv 1 - 2\sigma + \sigma^2/(1-\gamma) = (1-\sigma)^2 + \gamma\sigma^2/(1-\gamma)$. Then for any $\alpha \in [0, \hat{\alpha}]$, we have:

(a) $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{N}_{-\infty}(\gamma)$;

(b) $X_\alpha \bullet S_\alpha = (1 - \alpha + \alpha\sigma)(X \bullet S)$.

Proof. Statement (b) is an immediate consequence of Lemma 3.1. Using Lemma 6.1, relations (3.30) and (6.12), and the fact that $\gamma \le 1$ and $\zeta \ge \gamma\sigma^2/(1-\gamma)$, we obtain

$$\begin{aligned}
\hat\alpha \max\left\{\|\widehat{\Delta X}\|, \|\widehat{\Delta S}\|\right\} &\le \hat\alpha \max\left\{\left\|X^{-1/2}\Delta X X^{-1/2}\right\|_F,\right.\\
&\qquad\left.\left\|X^{-1/2}S^{-1/2}\right\|\left\|X^{1/2}\Delta S S^{-1/2}\right\|_F\right\}\\
&\le \hat\alpha \max\left\{\frac{2}{\sqrt{\lambda_{\min}}}\|\widehat{H}\|_F, 3\frac{\lambda_{\max}^{1/2}}{\lambda_{\min}}\|\widehat{H}\|_F\right\}\\
&\le \hat\alpha \max\left\{\frac{2}{\sqrt{(1-\gamma)\mu}}(\zeta n\mu)^{1/2}, 3\frac{(n\mu)^{1/2}}{(1-\gamma)\mu}(\zeta n\mu)^{1/2}\right\}\\
&\le \max\left\{\frac{\sigma\gamma}{15n\sqrt{\zeta}}, \frac{\sigma\gamma}{10(\zeta(1-\gamma)n)^{1/2}}\right\} \;\le\; \frac{1}{2}.
\end{aligned}$$

Hence, it follows from Lemma 6.5, Lemma 3.9(c), and the fact that $\lambda_{\min} \ge (1-\gamma)\mu$ and $\lambda_{\max} \le n\mu$ that $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$ for any $\alpha \in [0, \hat\alpha]$, and

$$\begin{aligned}
\|\phi''(0)\|_F &\le 18\frac{n^{1/2}}{(1-\gamma)^{1/2}}\zeta n\mu,\\
\sup_{\xi\in[0,\tilde\alpha]}\|\phi'''(\xi)\|_F &\le 3000\frac{n}{(1-\gamma)^{3/2}}(\zeta n)^{3/2}\mu.
\end{aligned}$$

These two inequalities together with (3.14), (6.12), Lemma 3.1, and the fact that $\gamma \le 1$ and $\zeta \ge \gamma\sigma^2/(1-\gamma)$ imply that for every $\alpha \in [0, \tilde\alpha]$,

$$\begin{aligned}
\left\|X_\alpha^{1/2}S_\alpha X_\alpha^{1/2} - \mu(\alpha)I\right\|_{-\infty} &= \|\phi(\alpha)\|_{-\infty}\\
&\le (1-\alpha)\|\phi(0)\|_{-\infty}\\
&\quad +\alpha\left[\frac{1}{2}\hat\alpha\|\phi''(0)\|_F + \frac{1}{6}\hat\alpha^2 \sup_{\xi\in[0,\alpha]}\|\phi'''(\xi)\|_F\right]\\
&\le (1-\alpha)\gamma\mu\\
&\quad +\alpha\left[9\hat\alpha\frac{n^{1/2}}{(1-\gamma)^{1/2}}\zeta n\mu + 500\hat\alpha^2\frac{n}{(1-\gamma)^{3/2}}(\zeta n)^{3/2}\mu\right]\\
&\le (1-\alpha)\gamma\mu + \alpha\left[\frac{3}{10}\sigma\gamma\mu + \frac{5\sigma^2\gamma^2\mu}{9(1-\gamma)^{1/2}n^{1/2}\zeta^{1/2}}\right]\\
&\le (1-\alpha)\gamma\mu + \alpha\sigma\gamma\mu\left[\frac{3}{10} + \frac{5\gamma^{1/2}}{9n^{1/2}}\right]\\
&\le (1-\alpha)\gamma\mu + \alpha\sigma\gamma\mu \;=\; \gamma\mu(\alpha).
\end{aligned}$$

Hence, $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{N}_{-\infty}(\gamma)$ for every $\alpha \in [0, \tilde\alpha]$; that is, (a) holds. ◻

COROLLARY 6.7 (polynomiality of long-step path-following algorithm for the MT* family). *Let constants $\gamma, \bar\sigma \in (0,1)$ be given. For Algorithm* II, *assume that $(\widetilde{X}^0, \widetilde{S}^0, \widetilde{y}^0) \in \mathcal{N}_{-\infty}(\gamma)$ and that the sequences $\{\sigma_k\}$ and $\{\alpha_k\}$ are chosen according to*

$$\begin{aligned}
\sigma_k &= \bar\sigma,\\
\alpha_k &= \max\left\{\alpha \in [0,1] : (\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k) + \alpha(\Delta\widetilde{X}^k, \Delta\widetilde{S}^k, \Delta\widetilde{y}^k) \in \widetilde{\mathcal{N}}_{-\infty}(\gamma)\right\}.
\end{aligned}$$

Then, the sequence of iterates $\{(\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k)\} \subset \mathcal{N}_{-\infty}(\gamma)$ *generated by Algorithm* II *satisfies* $\widetilde{X}^k \bullet \widetilde{S}^k \le (1-\bar{\eta})^k (\widetilde{X}^0 \bullet \widetilde{S}^0)$ *for all* $k \ge 0$*, where*

$$\bar{\eta} \equiv \frac{\bar{\sigma}(1-\bar{\sigma})\gamma(1-\gamma)^{1/2}}{30 n^{3/2} \bar{\zeta}},$$

and $\bar{\zeta} \equiv 1 - 2\bar{\sigma} + \bar{\sigma}^2/(1-\gamma)$*. Moreover, if the quantity* $\max\{\gamma^{-1}, (1-\gamma)^{-1}, \bar{\sigma}^{-1}, (1-\bar{\sigma})^{-1}\}$ *is independent of* n*, then the method terminates in at most* $\mathcal{O}(n^{3/2}L)$ *iterations.*

Proof. One step of the algorithm in the scaled space is analyzed by Theorem 6.6. By translating the result into the terms of the original space using the invariance of μ and $d_{-\infty}$ and (5.2c), the result readily follows. □

We have thus established an $O(n^{3/2}L)$ iteration complexity for the long-step path-following feasible algorithm based on any member of the MT* family. A natural question is whether our approach yields better iteration complexities for the special cases in which $X = I$ (the HRVW/KSH/M dual direction), $SXS = I$ (the HRVW/KSH/M direction), $S = I$ (the NT direction), and $X = S$ (the MTW direction). Unfortunately, our approach does not seem to yield the $O(nL)$ iteration-complexity bound that has been obtained in Monteiro and Zhang [21] for the NT direction nor to improve the $O(n^{3/2}L)$ iteration-complexity bound for the HRVW/KSH/M dual direction obtained in Monteiro [15]. For the MTW direction, we can show that the long-step algorithm has an $O(n^{11/8}L)$ iteration-complexity bound, slightly improving the general $O(n^{3/2}L)$ bound. We omit the proof of this claim here.

7. Concluding remarks. We proposed a new family of primal-dual interior-point methods for SDP. The method is based on the application of Newton's method to the equation

$$(\widetilde{P}\widetilde{X}\widetilde{P}^T)^{1/2}(\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1})(\widetilde{P}\widetilde{X}\widetilde{P}^T)^{1/2} - \nu I = 0$$

for some $\nu > 0$ and scaling nonsingular matrix $\widetilde{P}$. We proved existence of the Newton direction for any $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{N}}_\infty(\gamma)$ with $\gamma \in (0, 1/\sqrt{2})$, and established an $O(\sqrt{n}L)$ iteration-complexity bound for the short-step path-following algorithm and an $O(nL)$ iteration-complexity bound for the semilong-step path-following algorithm. Furthermore, we showed that for any interior feasible point $(\widetilde{X}, \widetilde{S}, \widetilde{y})$, the Newton direction corresponding to those scaling matrices $\widetilde{P}$ satisfying

$$(\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1})(\widetilde{P}\widetilde{X}\widetilde{P}^T)^{1/2} + (\widetilde{P}\widetilde{X}\widetilde{P}^T)^{1/2}(\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}) \succ 0$$

always exists, and we established an $O(n^{3/2}L)$ iteration-complexity bound for the long-step path-following algorithm based on this subclass of scaling matrices. This subclass yields two well-known search directions, namely, the HRVW/KSH/M dual direction when $\widetilde{P} = \widetilde{X}^{-1/2}$, the HRVW/KSH/M direction when $\widetilde{P} = (\widetilde{S}\widetilde{X}\widetilde{S})^{1/2}$, and the NT direction when $\widetilde{P} = \widetilde{S}^{1/2}$.

It is possible to derive a symmetric MT family based on the central path equation $S^{1/2}XS^{1/2} - \nu I = 0$, obtained from the one in section 2 by interchanging the role of X and S. It is easy to see that the symmetric MT family obtained by applying Newton's method to this equation (in the scaled space) has similar properties to the one studied in this paper and that it contains the NT direction and the two HRVW/KSH/M directions.

It is interesting to compare the MZ family and the MT family in light of the motivation used in section 5 to derive the MT family. We know that search directions of the MT family correspond to the Newton direction for the central path equation $X^{1/2}SX^{1/2} - \nu I = 0$ in the scaled space for an appropriate choice of $\widetilde{P}$. In a similar vein, it is easy to see that search directions of the MZ family correspond to the Newton direction for the central path equation $XS + SX - \nu I = 0$ in the scaled space for an appropriate choice of $\widetilde{P}$. This observation indicates the existence of a natural association of the MT family with the (pure) Newton direction of section 2 and of the MZ family with the (pure) Newton AHO direction.

Based on the theoretical results obtained so far, the pure Newton direction of section 2 has clear advantages over the AHO direction in the sense that polynomial convergence of the semilong-step path-following algorithm is only known for the former direction. So far this is the only pure Newton path-following algorithm which is polynomially convergent and is based on a wide neighborhood of the central path.

The MT* family also has theoretical advantages over the MZ* family based on the results so far. While for the MZ* family, the iteration-complexity bound depends on a certain condition number associated with the sequence $\{P^k\}$ of scaling matrices, the corresponding bound for the MT* family does not depend on this sequence.

After the release of this paper, Monteiro and Zanjácomo [20] have reported promising computational results for algorithms based on the pure Newton direction (2.11) and two other pure Newton directions based on the central path equations:

$$S^{1/2}XS^{1/2} = \nu I,$$
$$L_S^T X L_S = \nu I,$$

respectively, where L_S denotes the Cholesky lower triangular factor of S, that is, $S = L_S L_S^T$ with L_S lower triangular.

Finally, we mention that an interesting topic for future study would be to develop algorithms based on the pure Newton direction (2.11) that are superlinearly or quadratically convergent. We refer the reader to [9, 13, 29, 28], where quadratically convergent SDP algorithms based on other primal-dual directions are developed under the presence of strict complementarity and/or nondegeneracy assumptions.

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: Convergence rates, stability and numerical results*, SIAM J. Optim., 8 (1998), pp. 746–768.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] M. Kojima, M. Shida, and S. Shindoh, *A Note on the Nesterov-Todd and the Kojima-Shindoh-Hara Search Directions in Semidefinite Programming*, Optim. Methods Softw., to appear.

[9] 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.
[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, 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.
[12] 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, 1995.
[13] 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.
[14] 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.
[15] R. D. C. Monteiro, *Primal-dual path-following algorithms for semidefinite programming*, SIAM J. Optim., 7 (1997), pp. 663–678.
[16] R. D. C. Monteiro, *Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions*, SIAM J. Optim., 8 (1998), pp. 59–81.
[17] R. D. C. Monteiro and I. Adler, *Interior path-following primal-dual algorithms. Part* I: *Linear programming*, Math. Programming, 44 (1989), pp. 27–41.
[18] 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.
[19] R. D. C. Monteiro and T. Tsuchiya, *Polynomiality of primal-dual algorithms for semidefinite linear complementarity problems based on the Kojima-Shindoh-Hara family of directions*, Manuscript, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, 1996; Math. Programming, to appear.
[20] R. D. C. Monteiro and P. R. Zanjácomo, *Implementation of Primal-Dual Methods for Semidefinite Programming Based on Monteiro and Tsuchiya Newton Directions and Their Variants*, Tech. Report, School of ISyE, Georgia Institute of Technology, Atlanta, 1997; Optim. Math. Softw., to appear.
[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, *A general approach to the design of optimal methods for smooth convex functions minimization*, Ekonomika i Matem. Metody, 24 (1988), pp. 509–517 (in Russian). (English transl.: 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*, Tech. 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, 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, *On the Local Convergence of a Predictor-Corrector Method for Semidefinite Programming*, Reports on Computational Mathematics 98, Dept. of Mathematics, The University of Iowa, Iowa City, 1997.
[29] 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.
[30] 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.
[31] K. Tanabe, *Complementarity-enforcing centered Newton method for mathematical program-*

ming: Global methods, in New Method for Linear Programming, K. Tone, ed., Cooperative Research Report 5, Institute of Statistical Mathematics, Tokyo, Japan, 1987, pp. 118–144.

[32] K. TANABE, *Centered Newton method for mathematical programming*, in System Modelling and Optimization, M. Iri and K. Yajima, eds., Springer-Verlag, Tokyo, Japan, 1988, pp. 197–206.

[33] M. TODD, *On the Search Directions in Interior-Point Methods for Semidefinite Programming*, Tech. Report 1205, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, October, 1997.

[34] 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.

[35] P. TSENG, *Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP*, Optim. Methods Softw., 9 (1998), pp. 245–268.

[36] L. VANDENBERGHE AND S. BOYD, *A primal-dual potential reduction method for problems involving matrix inequalities*, Math. Programming, 69 (1995), pp. 205–236.

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

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