SIAM J. OPTIM.
Vol. 7, No. 3, pp. 663–678, August 1997

005

PRIMAL–DUAL PATH-FOLLOWING ALGORITHMS FOR SEMIDEFINITE PROGRAMMING*

RENATO D. C. MONTEIRO†

Abstract. This paper deals with a class of primal–dual interior-point algorithms for semidefinite programming (SDP) which was recently introduced by Kojima, Shindoh, and Hara [*SIAM J. Optim.*, 7 (1997), pp. 86–125]. These authors proposed a family of primal-dual search directions that generalizes the one used in algorithms for linear programming based on the scaling matrix $X^{1/2}S^{-1/2}$. They study three primal–dual algorithms based on this family of search directions: a short-step path-following method, a feasible potential-reduction method, and an infeasible potential-reduction method. However, they were not able to provide an algorithm which generalizes the long-step path-following algorithm introduced by Kojima, Mizuno, and Yoshise [*Progress in Mathematical Programming: Interior Point and Related Methods*, N. Megiddor, ed., Springer-Verlag, Berlin, New York, 1989, pp. 29–47]. In this paper, we characterize two search directions within their family as being (unique) solutions of systems of linear equations in symmetric variables. Based on this characterization, we present a simplified polynomial convergence proof for one of their short-step path-following algorithms and, for the first time, a polynomially convergent long-step path-following algorithm for SDP which requires an extra $\sqrt{n}$ factor in its iteration-complexity order as compared to its linear programming counterpart, where n is the number of rows (or columns) of the matrices involved.

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

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

PII. S1052623495293056

1. Introduction.

This paper studies primal–dual path-following algorithms for semidefinite programming (SDP) based on a search direction that has been proposed by Kojima, Shindoh, and Hara [11] as a natural extension of the one used in algorithms for linear programming based on the scaling matrix $X^{1/2}S^{-1/2}$. The first primal–dual algorithm for linear programming (LP) to use this scaling matrix was presented by Kojima, Mizuno, and Yoshise [10] and is referred in here to as the *long-step path-following method.* Another variant independently developed by Kojima, Mizuno, and Yoshise [9] and Monteiro and Adler [12, 13], referred to here as the *short-step path-following method*, improves the worst-case iteration complexity of the algorithm of [10] by a factor of $\sqrt{n}$ by generating iterates in a narrower neighborhood of the central path.

Several authors have discussed generalizations of interior-point algorithms for linear programming to the context of SDP. The landmark work in this direction is due to Nesterov and Nemirovskii [14, 15], 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 [17] for a comprehensive treatment of this subject.) They show that the problem of minimizing a linear function over a convex set K can be solved in "polynomial time" as long as a self-concordant barrier function for K is known. In particular, Nesterov and Nemirovskii show that linear programs,

* Received by the editors October 11, 1995; accepted for publication (in revised form) April 2, 1996.

http://www.siam.org/journals/siopt/7-3/29305.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 National Science Foundation under grant INT-9600343 and the Office of Naval Research under grants N00014-93-1-0234 and N00014-94-1-0340.

convex quadratic programs with convex quadratic constraints, and semidefinite programs all have explicit and easily computable self-concordant functions and hence can be solved in "polynomial time." Subsequently, Alizadeh [2] extends in a direct way Ye's projective potential-reduction algorithm (see [21]) for LP to the context of SDP and argues that many known interior-point LP algorithms can also be transformed into an algorithm for SDP in a mechanical way. Since then, several authors have proposed interior-point algorithms for solving SDP problems, including Helmberg et al. [5], Jarre [8], Kojima, Shindoh, and Hara [11], Nesterov and Nemirovskii [16], Nesterov and Todd [19, 18] and Vandenberghe and Boyd [20].

Among the above works, Kojima, Shindoh, and Hara [11] and Nesterov and Todd [18] present some algorithms which extend the primal–dual methods for linear programming based on the scaling $X^{1/2}S^{-1/2}$. In particular, they both provide short-step path-following methods for SDP which generalize the algorithm in [9, 12, 13]; however, no extensions of the long-step path-following algorithm in [10] are provided. In fact, Kojima, Shindoh, and Hara mention in section 9 of [11] that they encountered difficulty in providing such an extension.

In this paper, by characterizing two of the search directions introduced in [11] as solutions of systems of linear equations in symmetric variables, we present a simplified polynomial convergence proof for a short-step path-following algorithm in [11] and for the first time, a polynomially convergent long-step path-following algorithm for SDP. We show that the long-step method requires $\mathcal{O}(n^{3/2}\log(t^0\epsilon^{-1}))$ iterations to generate a feasible solution with objective function within ϵ of the optimal value when initialized at an interior feasible point whose duality gap is t^0. Hence, the algorithm of [10] when extended to SDP has its iteration-complexity increased by a factor of $\sqrt{n}$.

This paper is organized as follows. In section 2, we describe the generic primal–dual algorithm for SDP which will be the subject of our study in this paper. Section 3 contains some matrix results that are frequently used in our presentation. Section 4 discusses the short-step path-following method for SDP while section 5 discusses its long-step counterpart.

1.1. Notation and terminology. The following notation is used throughout the paper. The superscript ^{T} denotes transpose. $\Re^p$, $\Re^p_+$, and $\Re^p_{++}$ denote the p-dimensional Euclidean space, the nonnegative orthant of $\Re^p$, and the positive orthant of $\Re^p$, respectively. The ith component of a vector $u \in \Re^p$ is denoted by u_i. The set of all $p \times q$ matrices with real entries is denoted by $\Re^{p\times q}$. The (i,j)th entry of a matrix $Q \in \Re^{p\times q}$ is denoted by Q_{ij}. The set of all symmetric $p \times p$ matrices is denoted by $\mathcal{S}^{(p)}$ or, simply, by $\mathcal{S}$ when the dimension p is clear from the context. For $Q \in \mathcal{S}$, $Q \succeq 0$ ($Q \preceq 0$) means Q is positive (negative) semidefinite and $Q \succ 0$ ($Q \prec 0$) means Q is positive (negative) definite. The trace of a matrix $Q \in \Re^{p\times p}$ is denoted by $\text{Tr } Q \equiv \sum_{i=1}^n Q_{ii}$. The eigenvalues of $Q \in \mathcal{S}^{(p)}$ are denoted by $\lambda_i(Q)$, $i = 1,\ldots,p$, and its smallest and largest eigenvalues are denoted by $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$, respectively. Given P and Q in $\Re^{p\times q}$, the inner product between them is defined as $P \bullet Q \equiv \text{Tr } P^TQ = \sum_{i=1,j=1}^n P_{ij}Q_{ij}$. Given u and v in $\Re^p$, $u \le v$ means $u_i \le v_i$ for every $i = 1,\ldots,p$. 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 $\|Q\|_F \equiv (Q \bullet Q)^{1/2}$. $\mathcal{S}_+$ and $\mathcal{S}_{++}$ denote the set of all matrices in $\mathcal{S}$ which are positive semidefinite and positive definite, respectively. Finally, $\mathcal{S}_\perp^{(p)}$, or simply $\mathcal{S}_\perp$ when p is understood from the context, denote the set of all skew-symmetric matrices in $\Re^{p\times p}$. Since $\mathcal{S}^{(p)} + \mathcal{S}_\perp^{(p)} = \Re^{p\times p}$ and $U \bullet V = 0$ for every $U \in \mathcal{S}^{(p)}$ and $V \in \mathcal{S}_{\perp}^{(p)}$, it follows that $\mathcal{S}_{\perp}^{(p)}$ is the orthogonal complement of $\mathcal{S}^{(p)}$ with respect to the inner product $\bullet$.

2. The primal–dual algorithm and some technical results. In this section we describe the generic primal–dual algorithm which will be the subject of our study in this paper. We then show that the search direction used by the generic algorithm is a particular one from the family of the search directions introduced in the revised version of [11]. We end the section by giving some basic results about the generic algorithm.

This paper studies primal–dual path-following algorithms for solving the *semidefinite programming problem* (SDP)

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

and its associated dual SDP

$$(2) \qquad (D) \qquad \max\left\{b^T y : \sum_{i=1}^{m} y_i A_i + S = C,\ S \succeq 0\right\},$$

where $C \in \Re^{n\times n}$, $A_i \in \Re^{n\times n}$, $i = 1, \ldots, m$, 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. We assume without loss of generality that the matrices C and A_i, $i = 1, \ldots, m$, are symmetric (otherwise, replace C by $(C + C^T)/2$ and A_i by $(A_i + A_i^T)/2$).

The set of *interior feasible solutions* of (1) and (2) are

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

respectively. Throughout this paper, we assume that $F^0(P) \times F^0(D) \neq \emptyset$. Under this assumption, it is well known that both (1) and (2) have optimal solutions X^* and (S^*, y^*) such that $C \bullet X^* = b^T y^*$ (that is, the optimal values of (1) and (2) are equal). This last condition alternatively can be expressed as $X^* \bullet S^* = 0$, since for feasible solutions X and (S, y) for (1) and (2) there hold $C \bullet X - b^T y = (\sum_{i=1}^{n} y_i A_i + S) \bullet X - b^T y = X \bullet S + \sum_{i=1}^{n} y_i (A_i \bullet X) - b^T y = X \bullet S + \sum_{i=1}^{n} y_i b_i - b^T y = X \bullet S$. For simplicity, we will also assume that the matrices A_i, $i = 1, \ldots, m$ are linearly independent.

We next outline a generic interior-point primal–dual algorithm for solving the pair of SDPs (1) and (2) which was introduced in [11]. The system of linear equations defining the search direction in the following algorithm is actually different from the one used in [11], but the resulting search direction is the same as will be shown in Lemma 2.1.

GENERIC PRIMAL–DUAL ALGORITHM.

Step 0. Let $X^0 \in F^0(P)$ and $(S^0, y^0) \in F^0(D)$ be given and set $k = 0$.

Step 1. Let $X = X^k$, $(S, y) = (S^k, y^k)$ and $\mu = (X \bullet S)/n$;

Step 2. Choose a centrality parameter $\sigma = \sigma_k \in [0, 1]$ and set $H \equiv (\sigma\mu I - X^{1/2} S X^{1/2})$;

Step 3. Compute the search direction $(\Delta X, \Delta S, \Delta y) \in \mathcal{S} \times \mathcal{S} \times \Re^m$ by solving the following system of linear equations:

$$X^{-1/2}(X\Delta S + \Delta X S)X^{1/2} + X^{1/2}(\Delta S X + S\Delta X)X^{-1/2} = 2H, \tag{3}$$

$$A_i \bullet \Delta X = 0, \quad \text{for all } i = 1, \ldots, m, \tag{4}$$

$$\sum_{i=1}^{m} \Delta y_i A_i + \Delta S = 0; \tag{5}$$

Step 4. Choose a step-size $\alpha = \alpha_k \geq 0$ such that $\hat{X} \equiv X + \alpha \Delta X \in \mathcal{S}_{++}$, and $(\hat{S}, \hat{y}) \equiv (S, y) + \alpha(\Delta S, \Delta y) \in \mathcal{S}_{++} \times \Re^m$;

Step 5. Let $X^{k+1} = \hat{X}$, $(S^{k+1}, y^{k+1}) = (\hat{S}, \hat{y})$, replace k by $k+1$, and go to step (1).

End

In what follows, we show that the search direction used by the generic algorithm is a particular one from the family of the search directions introduced in the revised version of [11]. We first describe this family of search directions. Given a fixed $t \in [0, 1]$, Kojima et al. show that the system of linear equations consisting of (4), (5), and the equation

$$X(\Delta S + tW) + (\Delta X + (1-t)W)S = \sigma\mu I - XS, \tag{6}$$

has a unique solution $(\Delta X(t), \Delta S(t), \Delta y(t), W(t)) \in \mathcal{S} \times \mathcal{S} \times \Re^m \times \mathcal{S}_{\perp}$ (see Theorem 4.2 of [11]). The search direction for their algorithm is $(\Delta X(t), \Delta S(t), \Delta y(t))$ for some fixed $t \in [0, 1]$. (They have in fact introduced a larger family of search directions but this one suffices for the purpose of our discussion.) The following result shows that system (4), (5), and (6) with $t = 1$ determines exactly the same direction as system (3)–(5) does; that is, $(\Delta X(1), \Delta S(1), \Delta y(1)) = (\Delta X, \Delta S, \Delta y)$.

LEMMA 2.1. *$(\Delta X(1), \Delta S(1), \Delta y(1))$ is the unique solution of the system* (3)–(5).

Proof. Let $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y, \hat{W}) \equiv (\Delta X(1), \Delta S(1), \Delta y(1), W(1))$. We first show that $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y)$ is a solution of (3)–(5). It suffices to show that $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y)$ satisfies (3). Indeed, by definition, $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y)$ satisfies (6) with $t = 1$. After multiplying this relation on the left by $X^{-1/2}$ and on right by $X^{1/2}$, we obtain

$$X^{1/2}(\hat{\Delta} S + \hat{W})X^{1/2} + X^{-1/2}\hat{\Delta} X S X^{1/2} = \sigma\mu I - X^{1/2} S X^{1/2}.$$

Hence, the sum of the symmetric parts of the two terms on the left-hand side is equal to the right-hand side. This fact together with the fact that $\hat{W} + \hat{W}^T = 0$ imply

$$\begin{aligned}
&2\left(\sigma\mu I - X^{1/2} S X^{1/2}\right) \\
&= X^{1/2}(2\hat{\Delta} S + \hat{W} + \hat{W}^T)X^{1/2} + X^{-1/2}\hat{\Delta} X S X^{1/2} + X^{1/2} S \hat{\Delta} X X^{-1/2} \\
&= 2X^{1/2}\hat{\Delta} S X^{1/2} + X^{-1/2}\hat{\Delta} X S X^{1/2} + X^{1/2} S \hat{\Delta} X X^{-1/2} \\
&= X^{-1/2}(X\hat{\Delta} S + \hat{\Delta} X S)X^{1/2} + X^{1/2}(\hat{\Delta} S X + S\hat{\Delta} X)X^{-1/2}.
\end{aligned}$$

That is, $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y)$ satisfies (3). To show that $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y)$ is the only solution of (3)–(5), assume that $(\Delta X, \Delta S, \Delta y)$ is an arbitrary solution of (3)–(5) and let $E \equiv X^{-1/2}(X\Delta S + \Delta X S)X^{1/2}$. Then, by (3) we have $E + E^T = 2H$, and hence $W \equiv X^{-1/2}(H - E)X^{-1/2} = X^{-1/2}(E^T - E)X^{-1/2}/2$ is skew symmetric. A simple algebraic manipulation shows that $(\Delta X, \Delta S, \Delta y, W)$ satisfies (6) with $t = 1$ and, hence, that it is a solution of the system defined by (4), (5), and (6) with $t = 1$. Since $(\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y, \hat{W})$ is the unique solution of this system in $\mathcal{S} \times \mathcal{S} \times \Re^m \times \mathcal{S}_{\perp}$, we conclude that $(\Delta X, \Delta S, \Delta y, W) = (\hat{\Delta} X, \hat{\Delta} S, \hat{\Delta} y, \hat{W})$. ∎

In a similar vein, it is possible to characterize $(\Delta X(0), \Delta S(0), \Delta y(0))$ as the unique solution in $\mathcal{S} \times \mathcal{S} \times \Re^m$ of the system of linear equations consisting of (4), (5), and the equation

$$S^{1/2}(X\Delta S + \Delta X S)S^{-1/2} + S^{-1/2}(\Delta S X + S\Delta X)S^{1/2} = 2(\sigma\mu I - S^{1/2} X S^{1/2}). \tag{7}$$

Results analogous to the ones proved in this paper easily can be obtained with respect to path-following algorithms based on this search direction.

It should be noted that the two systems of linear equations (3)–(5) and (3), (4), (7) were introduced for the first time in a preliminary version of this paper. The result stated in Lemma 2.1 was subsequently pointed out by Masakazu Kojima to the author in a personal communication. The present version of this paper is essentially a modification of the previous version which takes into account this important observation.

From the discussion above, we see that both directions $(\Delta X(0), \Delta S(0), \Delta y(0))$ and $(\Delta X(1), \Delta S(1), \Delta y(1))$ are solutions of systems of linear equations in symmetric matrices, a property which is also shared by the NT-direction introduced by Nesterov and Todd [18], namely, the unique solution $(\Delta X, \Delta S, \Delta y)$ of (4), (5), and the equation

$$\begin{aligned}&(X^{1/2}SX^{1/2})^{1/2}X^{-1/2}\Delta X X^{-1/2}(X^{1/2}SX^{1/2})^{1/2}\\ &\quad + X^{1/2}\Delta S X^{1/2} = \sigma\mu I - X^{1/2}SX^{1/2}.\end{aligned} \tag{8}$$

But unlike the NT-direction, computing the directions $(\Delta X(t), \Delta S(t), \Delta y(t))$ do not require computation of matrix square roots, which is certainly an advantage from the computational point of view.

Another primal–dual search direction which has been considered by a few authors (see, for example, Adler and Alizadeh [1] and Alizadeh, Haeberly, and Overton [3]) is the one that is the solution of the linear system consisting of (4), (5), and the equation

$$X\Delta S + \Delta S X + S\Delta X + \Delta X S = 2\sigma\mu I - XS - SX. \tag{9}$$

At the time of this writing, no polynomial convergence has been proven for an algorithm based on this direction.

We end this section by stating the following straightforward result regarding the generic algorithm.

LEMMA 2.2. *Let $X \in F^0(P)$ and $(S, y) \in F^0(D)$ be given and suppose that $(\Delta X, \Delta S, \Delta y)$ is a solution of (3)–(5) for some $H \in \Re^{n\times n}$. Then, the following statements hold:*

(a) $\Delta S \bullet \Delta X = 0$,

(b) $X \bullet \Delta S + S \bullet \Delta X = \mathrm{Tr}\, H$,

(c) *if $H = \sigma\mu I - X^{1/2}SX^{1/2}$ where $\sigma \in \Re$ and $\mu \equiv (X \bullet S)/n$, then*

$$(X + \alpha\Delta X) \bullet (S + \alpha\Delta S) = (1 - \alpha + \alpha\sigma)(X \bullet S) \quad \forall \alpha \in \Re.$$

Proof. Using (4) and (5), we obtain

$$\Delta S \bullet \Delta X = -\left(\sum_{i=1}^{n} \Delta y_i A_i\right) \bullet \Delta X = -\sum_{i=1}^{n} \Delta y_i (A_i \bullet \Delta X) = 0,$$

and hence (a) follows. In view of (3), we have

$$\begin{aligned}2\,\mathrm{Tr}\, H &= \mathrm{Tr}\ X^{-1/2}(X\Delta S + \Delta X S)X^{1/2} + \mathrm{Tr}\ X^{1/2}(\Delta S X + S\Delta X)X^{-1/2}\\ &= \mathrm{Tr}\ (X\Delta S + \Delta X S) + \mathrm{Tr}\ (\Delta S X + S\Delta X)\\ &= 2\,\mathrm{Tr}\ (X\Delta S + S\Delta X) \ = \ 2(X \bullet \Delta S + S \bullet \Delta X),\end{aligned}$$

and hence (b) follows. Using statements (a) and (b) and the fact that $H = \sigma\mu I - X^{1/2}SX^{1/2}$ and Tr $(X^{1/2}SX^{1/2}) = X \bullet S = n\mu$, we obtain

$$\begin{aligned}(X + \alpha\Delta X) \bullet (S + \alpha\Delta S) &= X \bullet S + \alpha(X \bullet \Delta S + S \bullet \Delta X) + \alpha^2(\Delta X \bullet \Delta S) \\ &= X \bullet S + \alpha \mathrm{Tr}\ \left(\sigma\mu I - X^{1/2}SX^{1/2}\right) \\ &= X \bullet S + \alpha\,(\sigma n\mu - X \bullet S) \\ &= (1 - \alpha + \alpha\sigma)(X \bullet S)\end{aligned}$$

for every $\alpha \in \Re$. Hence, (c) holds. □

3. Some technical results about matrices. This section states some inequalities about matrices which play an important role in the convergence analysis of the algorithms presented in sections 4 and 5.

In the next result, we collect some useful facts about symmetric matrices. For its proof, we refer the reader to Golub and Van Loan [4] or Horn and Johnson [6].

LEMMA 3.1. *For any $E \in \mathcal{S}^{(p)}$, we have*

$$\lambda_{\max}(E) = \max_{\|u\|=1} u^T E u, \tag{10}$$

$$\lambda_{\min}(E) = \min_{\|u\|=1} u^T E u, \tag{11}$$

$$\|E\| = \max_{i=1,\dots,p} |\lambda_i(E)|, \tag{12}$$

$$\|E\|_F^2 = \sum_{i=1}^{p} [\lambda_i(E)]^2. \tag{13}$$

The following result about general matrices is also useful.

LEMMA 3.2. *For any $W \in \Re^{p\times p}$, the following relations hold:*

$$\max_{i=1,\dots,n} Re[\lambda_i(W)] \le \frac{1}{2}\lambda_{\max}(W + W^T), \tag{14}$$

$$\min_{i=1,\dots,n} Re[\lambda_i(W)] \ge \frac{1}{2}\lambda_{\min}(W + W^T), \tag{15}$$

$$\sum_{i=1}^{p} |\lambda_i(W)|^2 \le \|W\|_F^2 = \|W^T\|_F^2, \tag{16}$$

$$\lambda_{\max}(W^T W) = \|W^T W\| \ = \ \|W\|^2 \ = \ \|W^T\|^2. \tag{17}$$

Proof. Inequality (14) is stated as an exercise in Horn and Johnson; see [7], page 187, exercise 20. Inequality (15) follows from (14) applied to the matrix $-W$. For a proof of (16) and (17), see Golub and Van Loan [4], pages 58 and 336. □

As a consequence of Lemma 3.2, we obtain the following result.

LEMMA 3.3. *Suppose that $W \in \Re^{p\times p}$ is a nonsingular matrix. Then, for any $E \in \mathcal{S}^{(p)}$, we have*

$$\lambda_{\max}(E) \le \frac{1}{2}\lambda_{\max}\left(WEW^{-1} + (WEW^{-1})^T\right), \tag{18}$$

$$\lambda_{\min}(E) \ge \frac{1}{2}\lambda_{\min}\left(WEW^{-1} + (WEW^{-1})^T\right), \tag{19}$$

$$\|E\| \le \frac{1}{2}\,\|WEW^{-1} + (WEW^{-1})^T\|, \tag{20}$$

(21) $$\|E\|_F \leq \frac{1}{2} \|WEW^{-1} + (WEW^{-1})^T\|_F.$$

Proof. Using (14), we obtain

$$\lambda_{\max}(E) = \lambda_{\max}(WEW^{-1}) \leq \frac{1}{2}\lambda_{\max}\left(WEW^{-1} + (WEW^{-1})^T\right)$$

for every $E \in \Re^{p\times p}$, and hence (18) follows. Inequality (19) is proved in a similar way by using (15). Inequality (20) follows from (18), (19), and (12). To prove (21), we use (13) and (16) to get

$$\|E\|_F^2 = \sum_{i=1}^{p}[\lambda_i(E)]^2 = \sum_{i=1}^{p}[\lambda_i(WEW^{-1})]^2 \leq \|WEW^{-1}\|_F^2.$$

Hence, we obtain

$$\begin{aligned} 4\,\|E\|_F^2 &\leq 2\,\|WEW^{-1}\|_F^2 + 2\|E\|_F^2 \;=\; 2\,\|WEW^{-1}\|_F^2 + 2\,\mathrm{Tr}\;E^2 \\ &= 2\,\|WEW^{-1}\|_F^2 + 2\,\mathrm{Tr}\;WE^2W^{-1} \;=\; 2\,\|WEW^{-1}\|_F^2 + 2\,\mathrm{Tr}\;(WEW^{-1})^2 \\ &= \|WEW^{-1} + (WEW^{-1})^T\|_F^2, \end{aligned}$$

which clearly implies (21). □

We observe that (20) is not needed in our presentation, but it could be useful in proving polynomial convergence of other primal–dual variants not studied in this paper. The other inequalities in Lemma 3.3 play a crucial role in the analysis of the short-step and the long-step path-following methods of sections 4 and 5, respectively.

4. Short-step path-following primal–dual algorithm. As mentioned previously, Kojima, Shindoh, and Hara [11] have studied a short-step path-following algorithm based on the search direction $(\Delta X(t), \Delta S(t), \Delta y(t))$ for any $t \in [0,1]$ (see (6)). In this section, we give a simplified polynomial convergence proof of their short-step path-following algorithm based on the search direction $(\Delta X(1), \Delta S(1), \Delta y(1))$ or, equivalently, the one determined by (3)–(5). It is a straightforward task to carry out a similar analysis with respect to the search direction $(\Delta X(0), \Delta S(0), \Delta y(0))$.

The short-step path-following algorithm generates iterates in the following (narrow) neighborhood of the central path:

$$\begin{aligned} \mathcal{N}_F(\gamma) &\equiv \{(X,S,y) \in F^0(P) \times F^0(D) : \|X^{1/2}SX^{1/2} - \mu I\|_F \leq \gamma\mu\} \\ &= \left\{(X,S,y) \in F^0(P) \times F^0(D) : \left(\sum_{i=1}^{n}(\lambda_i(XS) - \mu)^2\right)^{1/2} \leq \gamma\mu\right\}, \end{aligned}$$

where $\mu \equiv (X \bullet S)/n$ and γ is a constant such that $0 < \gamma < 1$. This neighborhood is a natural extension of the one used by the short-step path-following algorithm studied in [9, 12, 13]. The algorithm, which is a special case of the generic algorithm discussed in section 2, selects the sequence of step-sizes $\{\alpha_k\}$ and centrality parameters $\{\sigma_k\}$ according to the following rule.

SHORT-STEP METHOD. *For all $k \geq 0$, let $\alpha_k = 1$ and $\sigma_k \equiv 1 - \delta/\sqrt{n}$, where $\delta > 0$ is a constant which is specified in Theorem* 4.1 *below.*

The following result analyzes the behavior of one iteration of the short-step path-following method. Its proof will be given at the end of this section.

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

$$\frac{\gamma^2+\delta^2}{2(1-\gamma)^2(1-\delta/\sqrt{n})} \le \gamma, \quad \gamma \le \frac{1}{2}. \tag{22}$$

Suppose that $(X,S,y) \in \mathcal{N}_F(\gamma)$ and let $(\Delta X, \Delta S, \Delta y)$ denote the solution of (3)–(5) with $H = \sigma\mu I - X^{1/2}SX^{1/2}$ and $\sigma = 1-\delta/\sqrt{n}$. Then,

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

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

An example of constants γ and δ satisfying the conditions stated in Theorem 4.1 is $\gamma = \delta = 0.3$. As an immediate consequence, we obtain the following result for the short-step path-following method.

COROLLARY 4.2. *Let γ and δ be as in Theorem 4.1 and let $(X^0, S^0, y^0) \in \mathcal{N}_F(\gamma)$ be given. Then the short-step path-following method generates a sequence of points $\{(X^k, S^k, y^k)\} \subset \mathcal{N}_F(\gamma)$ such that $X^k \bullet S^k \le (1-\delta/\sqrt{n})^k (X^0 \bullet S^0)$ for all $k \ge 0$. Moreover, given a tolerance $\epsilon > 0$, the short-step path-following method computes an iterate (X^k, S^k, y^k) satisfying $X^k \bullet S^k \le \epsilon$ in at most $\sqrt{n}\delta^{-1}\log[\epsilon^{-1}(X^0 \bullet S^0)] = \mathcal{O}(\sqrt{n}\log[\epsilon^{-1}(X^0 \bullet S^0)])$ iterations.*

We now turn our efforts towards proving Theorem 4.1.

LEMMA 4.3. *Suppose that $X \in F^0(P)$, $(S,y) \in F^0(D)$, and let $(\Delta X, \Delta S, \Delta y)$ denote the solution of (3)–(5) with $H \equiv \sigma\mu I - X^{1/2}SX^{1/2}$. For any $\alpha \in \Re$, let*

$$(X(\alpha), S(\alpha), y(\alpha)) \equiv (X,S,y) + \alpha(\Delta X, \Delta S, \Delta y), \tag{23}$$

$$\mu(\alpha) \equiv (X(\alpha) \bullet S(\alpha))/n, \tag{24}$$

$$Q(\alpha) \equiv X^{-1/2}[X(\alpha)S(\alpha) - \mu(\alpha)I]X^{1/2}. \tag{25}$$

Then,

$$\begin{aligned} Q(\alpha) + Q(\alpha)^T &= 2(1-\alpha)(X^{1/2}SX^{1/2} - \mu I) \\ &+ \alpha^2\left[X^{-1/2}\Delta X \Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2}\right]. \end{aligned} \tag{26}$$

Proof. Let $\alpha \in \Re$ be given. By Lemma 2.2(c), we have $\mu(\alpha) = (1-\alpha+\sigma\alpha)\mu$. Hence, we obtain

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

This relation, together with (3), implies

$$\begin{aligned} Q(\alpha) + Q(\alpha)^T &= 2(1-\alpha)(X^{1/2}SX^{1/2} - \mu I) + 2\alpha(X^{1/2}SX^{1/2} - \sigma\mu I) \\ &\quad + \alpha\left[X^{-1/2}(X\Delta S + \Delta X S)X^{1/2} + X^{1/2}(S\Delta X + \Delta S X)X^{-1/2}\right] \\ &\quad + \alpha^2(X^{-1/2}\Delta X \Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2}) \\ &= 2(1-\alpha)(X^{1/2}SX^{1/2} - \mu I) \\ &\quad + \alpha^2(X^{-1/2}\Delta X \Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2}). \qquad \square \end{aligned}$$

The following lemma bounds the size of the scaled directions $X^{-1/2}\Delta X X^{-1/2}$ and $X^{1/2}\Delta S X^{1/2}$ for points $(X,S,y) \in F^0(P) \times F^0(D)$, which are "well centered."

Alternative bounds on the size of these quantities which are valid for any $(X,S,y) \in F^0(P) \times F^0(D)$ are given in Lemma 5.6, but the proof of the result below is considerably simpler than that of Lemma 5.6. The following inequality involving norms is used in the proof of the lemma below and in other places in our presentation: for any $A_1, A_2 \in \Re^{n\times n}$, we have $\|A_1A_2\|_F \le \|A_1\|\,\|A_2\|_F$ and $\|A_1A_2\|_F \le \|A_1\|_F\,\|A_2\|$ (see exercise 20 of section 5.6 of [6]).

LEMMA 4.4. *Let $X \in F^0(P)$ and $(S,y) \in F^0(D)$ be such that $\|X^{1/2}SX^{1/2} - \nu I\| \le \nu\gamma$ for some $\gamma \in [0,1)$ and $\nu > 0$. Suppose that $(\Delta X, \Delta S, \Delta y) \in \Re^{n\times n} \times \Re^{n\times n} \times \Re^m$ is a solution of* (3)–(5) *for some $H \in \Re^{n\times n}$ and let $\delta_x \equiv \nu\|X^{-1/2}\Delta X X^{-1/2}\|_F$ and $\delta_s \equiv \|X^{1/2}\Delta S X^{1/2}\|_F$. Then,*

$$\delta_x\delta_s \le \frac{1}{2}\left(\delta_x^2 + \delta_s^2\right) \le \frac{\|H\|_F^2}{2(1-\gamma)^2}.$$

Proof. Using (3) and simple algebraic manipulation, we obtain

$$\begin{aligned} H = {} & X^{1/2}\Delta S X^{1/2} + \nu X^{-1/2}\Delta X X^{-1/2} + \frac{1}{2}X^{-1/2}\Delta X X^{-1/2}(X^{1/2}SX^{1/2} - \nu I) \\ & + \frac{1}{2}(X^{1/2}SX^{1/2} - \nu I)X^{-1/2}\Delta X X^{-1/2}, \end{aligned}$$

from which it follows that

$$\begin{aligned} \|H\|_F &\ge \|X^{1/2}\Delta S X^{1/2} + \nu X^{-1/2}\Delta X X^{-1/2}\|_F \\ &\quad - \|X^{1/2}SX^{1/2} - \nu I\|\,\|X^{-1/2}\Delta X X^{-1/2}\|_F \\ &\ge \left(\|X^{1/2}\Delta S X^{1/2}\|_F^2 + \nu^2\|X^{-1/2}\Delta X X^{-1/2}\|_F^2\right)^{1/2} - (\gamma\nu)(\delta_x/\nu) \\ &= \sqrt{\delta_x^2 + \delta_s^2} - \gamma\delta_x \;\ge\; (1-\gamma)\sqrt{\delta_x^2 + \delta_s^2}, \end{aligned}$$

where the second inequality follows from the assumption that $\|X^{1/2}SX^{1/2} - \nu I\| \le \nu\gamma$ and the fact that $(X^{-1/2}\Delta X X^{-1/2}) \bullet (X^{1/2}\Delta S X^{1/2}) = \Delta X \bullet \Delta S = 0$, due to Lemma 2.2(a). The result now follows trivially from the last inequality. □

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. Statement (b) is an immediate consequence of Lemma 2.2(c) with $\alpha = 1$ and the fact that $\sigma = (1 - \delta/\sqrt{n})$. Hence,

$$\hat{\mu} \equiv (\hat{X} \bullet \hat{S})/n = (1 - \delta/\sqrt{n})\mu. \tag{27}$$

Using the fact that $(X^{1/2}SX^{1/2} - \mu I) \bullet I = 0$, $(X,S,y) \in \mathcal{N}_F(\gamma)$, and $\sigma = (1 - \delta/\sqrt{n})$, we obtain

$$\begin{aligned} \|\sigma\mu I - X^{1/2}SX^{1/2}\|_F^2 &= \|(\sigma-1)\mu I\|_F^2 + \|\mu I - X^{1/2}SX^{1/2}\|_F^2 \\ &\le \{(1-\sigma)^2 n + \gamma^2\}\mu^2 \;=\; (\delta^2 + \gamma^2)\mu^2. \end{aligned} \tag{28}$$

Since $\|X^{1/2}SX^{1/2} - \mu I\| \le \gamma\mu$, it follows from Lemma 4.4 with $\nu = \mu$ and $H = \sigma\mu I - X^{1/2}SX^{1/2}$ that

$$\|X^{-1/2}\Delta X X^{-1/2}\|_F \le \frac{\|\sigma\mu I - X^{1/2}SX^{1/2}\|_F}{(1-\gamma)\mu} \tag{29}$$

and

$$\|X^{-1/2}\Delta X X^{-1/2}\|_F\,\|X^{1/2}\Delta S X^{1/2}\|_F \le \frac{\|\sigma\mu I - X^{1/2}SX^{1/2}\|_F^2}{2(1-\gamma)^2\mu}. \tag{30}$$

Let $\hat{Q} \equiv Q(1) = X^{-1/2}(\hat{X}\hat{S} - \hat{\mu}I)X^{1/2}$. Using (26) with $\alpha = 1$, (30), (28), (22), and (27), we obtain

$$\begin{aligned}
(31)\qquad \frac{1}{2}\|\hat{Q} + \hat{Q}^T\|_F &= \frac{1}{2}\|X^{-1/2}\Delta X \Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2}\|_F \\
&\le \|X^{-1/2}\Delta X \Delta S X^{1/2}\|_F \\
(32)\qquad &\le \|X^{-1/2}\Delta X X^{-1/2}\|_F \, \|X^{1/2}\Delta S X^{1/2}\|_F \\
(33)\qquad &\le \frac{\|\sigma\mu I - X^{1/2}SX^{1/2}\|_F^2}{2(1-\gamma)^2\mu} \;\le\; \frac{(\gamma^2+\delta^2)\mu}{2(1-\gamma)^2} \\
(34)\qquad &\le \gamma(1-\delta/\sqrt{n})\mu \;=\; \gamma\hat{\mu}.
\end{aligned}$$

Using (29), (28), and (22), we obtain

$$\begin{aligned}
\|X^{-1/2}\Delta X X^{-1/2}\|_F &\le \frac{\|\sigma\mu I - X^{1/2}SX^{1/2}\|_F}{\mu(1-\gamma)} \;\le\; \frac{(\delta^2+\gamma^2)^{1/2}}{1-\gamma} \\
&\le \left[2\gamma(1-\delta/\sqrt{n})\right]^{1/2} < 1.
\end{aligned}$$

It is easy to see that the last relation implies that $I + X^{-1/2}\Delta X X^{-1/2} \succ 0$ and, hence, $\hat{X} \equiv X + \Delta X = X^{1/2}(I + X^{-1/2}\Delta X X^{-1/2})X^{1/2} \succ 0$. In particular, $\hat{X}^{1/2}$ exists. Applying Lemma 3.3 with $E = \hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \hat{\mu}I$ and $W = X^{-1/2}\hat{X}^{1/2}$ and noting that $\hat{Q} = WEW^{-1}$, we conclude that

$$(35)\qquad \|\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \hat{\mu}I\|_F \le \frac{1}{2}\,\|\hat{Q} + \hat{Q}^T\|_F \le \;\gamma\hat{\mu},$$

where the last inequality is due to (34). This implies that $\lambda_{\min}(\hat{X}^{1/2}\hat{S}\hat{X}^{1/2}) \ge (1-\gamma)\hat{\mu} > 0$, and hence $\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} \succ 0$. Thus, $\hat{S} \succ 0$. Using (4), (5), and the fact that $(X, S, y) \in F^0(P) \times F^0(D)$, it is now easy to see that $(\hat{X}, \hat{S}, \hat{y}) \in F^0(P) \times F^0(D)$. In view of (35), we conclude that $(\hat{X}, \hat{S}, \hat{y}) \in \mathcal{N}_F(\gamma)$. □

5. Long-step path-following algorithm. In this section, we present a long-step path-following algorithm whose iterates lie within a larger conical neighborhood of the central path. The algorithm extends the long-step primal–dual path-following method of Kojima, Mizuno, and Yoshise [10] for solving linear programming problems. We show that the algorithm finds an approximate strictly feasible point (X^k, S^k, y^k) satisfying $X^k \bullet S^k \le \epsilon$ within $\mathcal{O}(n^{3/2}\log(\epsilon^{-1}(X^0 \bullet S^0)))$ iterations, therefore requiring an extra $\sqrt{n}$ factor compared to the complexity of the algorithm in [10].

To describe the algorithm, we need to introduce the following neighborhood of the central path: for $\gamma \in [0, 1)$ and $\Gamma \ge 0$, let

$$\mathcal{N}(\gamma, \Gamma) \equiv \left\{(X, S, y) \in F^0(P) \times F^0(D) : \begin{array}{l} (1-\gamma)\mu \le \lambda_i(XS) \le (1+\Gamma)\mu \\ \text{for all } i = 1, \ldots, n \end{array} \right\}$$

and

$$\mathcal{N}(\gamma, \infty) \equiv \left\{(X, S, y) \in F^0(P) \times F^0(D) : \lambda_{\min}(XS) \ge (1-\gamma)\mu\right\},$$

where $\mu \equiv (X \bullet S)/n$. Clearly, $\mathcal{N}(\gamma, \Gamma) \subset \mathcal{N}(\gamma, \infty)$. We will describe the long-step path-following algorithm in terms of the neighborhood $\mathcal{N}(\gamma, \Gamma)$ with $0 \le \Gamma < \infty$. The following straightforward result shows that the corresponding algorithm based on the neighborhood $\mathcal{N}(\gamma,\infty)$ is a special case of the algorithm described in terms of $\mathcal{N}(\gamma,\Gamma)$ for specific values of Γ.

LEMMA 5.1. *For any $\Gamma \ge (n-1)\gamma$ and $\gamma \in [0,1)$, we have $\mathcal{N}(\gamma,\infty)=\mathcal{N}(\gamma,\Gamma)$.*

Proof. Let $(X,S,y)\in\mathcal{N}(\gamma,\infty)$ be given and $\lambda_1 \le \cdots \le \lambda_n$ denote the eigenvalues of XS. We know that $\lambda_1+\cdots+\lambda_n = X\bullet S = n\mu$. Hence, we have

$$\begin{aligned}\lambda_{\max}(XS) = \lambda_n &= n\mu - (\lambda_1+\cdots+\lambda_{n-1}) \le n\mu-(n-1)(1-\gamma)\mu\\ &= [1+(n-1)\gamma]\mu \le (1+\Gamma)\mu,\end{aligned}$$

and hence, $(X,S,y)\in\mathcal{N}(\gamma,\Gamma)$. □

We next describe the path-following algorithm studied in this section. Since the algorithm is a special case of the generic algorithm of section 2, it is enough to specify the choices of the sequence of step-sizes $\{\alpha_k\}$ and centrality parameters $\{\sigma_k\}$. Fix $\gamma\in(0,1)$, $\Gamma\ge\gamma$, $\bar\sigma\in(0,1)$, and, for all $k\ge 0$, let $(\Delta X^k,\Delta S^k,\Delta y^k)$ denote the solution of (3)–(5) with $(X,S)=(X^k,S^k)$ and $H=\sigma_k\mu_k I-(X^k)^{1/2}S^k(X^k)^{1/2}$, where $\mu_k \equiv (X^k\bullet S^k)/n$.

LONG-STEP PATH-FOLLOWING METHOD. *For all $k\ge 0$, let $\sigma_k=\bar\sigma$ and*

$$\alpha_k = \max\left\{\alpha\in[0,1]: \begin{array}{l}(X^k,S^k,y^k)+\alpha'(\Delta X^k,\Delta S^k,\Delta y^k)\in\mathcal{N}(\gamma,\Gamma)\\ \text{for all } \alpha'\in[0,\alpha]\end{array}\right\}. \tag{36}$$

The following result describes the behavior of one iteration of the long-step path-following method. Its proof will be given at the end of the section after we have stated and proved several preliminary lemmas.

THEOREM 5.2. *Suppose that $(X,S,y)\in\mathcal{N}(\gamma,\Gamma)$ for some constants $\gamma\in[0,1)$ and $\Gamma\ge\gamma$, and let $(\Delta X,\Delta S,\Delta y)$ denote the solution of (3)–(5) with $H=\sigma\mu I - X^{1/2}SX^{1/2}$ and $\sigma\in[0,1]$. Let*

$$\tilde\alpha \equiv \frac{\sigma\gamma(1-\gamma)^{1/2}}{n(1+\Gamma)^{1/2}}\left((1-\sigma)^2+\frac{\gamma\sigma^2}{1-\gamma}\right)^{-1}. \tag{37}$$

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

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

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

As an immediate consequence of Theorem 5.2, we obtain the following convergence result for the long-step path-following method.

COROLLARY 5.3. *The sequence of iterates $\{(X^k,S^k,y^k)\}\subset\mathcal{N}(\gamma,\Gamma)$ generated by the long-step path-following algorithm satisfies $X^k\bullet S^k\le(1-\bar\tau)^k(X^0\bullet S^0)$ for all $k\ge 0$, where*

$$\bar\tau \equiv \frac{\bar\sigma(1-\bar\sigma)\gamma(1-\gamma)^{1/2}}{n(1+\Gamma)^{1/2}}\left((1-\bar\sigma)^2+\frac{\gamma\bar\sigma^2}{1-\gamma}\right)^{-1}.$$

Moreover, given a tolerance $\epsilon>0$, the long-step path-following method computes an iterate satisfying $X^k\bullet S^k\le\epsilon$ in at most $\bar\tau^{-1}\log[\epsilon^{-1}(X^0\bullet S^0)] = \mathcal{O}(n\Gamma^{1/2}\log[\epsilon^{-1}(X^0\bullet S^0)])$ iterations.

Proof. It follows from Theorem 5.2, relation (36), and the fact that $\sigma_k=\bar\sigma$ that $\alpha_k\ge\bar\tau/(1-\bar\sigma)$ for all $k\ge 0$. In view of Theorem 5.2(b), we conclude that $X^{k+1}\bullet S^{k+1}=[1-(1-\bar\sigma)\alpha_k](X^k\bullet S^k)\le(1-\bar\tau)(X^k\bullet S^k)$ for all $k\ge 0$. Hence, the first part of the corollary follows. The second part of the result follows from the first part and some standard arguments. □

It follows from Corollary 5.3 that if the size of the quantity

$$\max\{\gamma^{-1}, (1-\gamma)^{-1}, \sigma^{-1}, (1-\sigma)^{-1}\}$$

is independent of n then the long-step path-following algorithm finds an ϵ-approximate solution in $\mathcal{O}(n\Gamma^{1/2}\log[\epsilon^{-1}(X^0 \bullet S^0)])$ iterations. In view of Lemma 5.1, we conclude that this number of iterations is equal to $\mathcal{O}(n^{3/2}\log[\epsilon^{-1}(X^0 \bullet S^0)])$ when the algorithm uses the neighborhood $\mathcal{N}(\gamma,\infty) = \mathcal{N}(\gamma,(n-1)\gamma)$.

We now turn our efforts towards proving Theorem 5.2.

LEMMA 5.4. *Suppose that* $(X,S,y) \in \mathcal{N}(\gamma,\Gamma)$ *for some* $\gamma \geq 0$ *and* $\Gamma \geq 0$ *and let* $(\Delta X, \Delta S, \Delta y)$ *denote the solution of* (3)–(5) *with* $H = \sigma\mu I - X^{1/2}SX^{1/2}$ *and* $\sigma \in [0,1]$. *Let* $\mu(\alpha)$ *and* $Q(\alpha)$ *be defined as in* (24) *and* (25) *for any* $\alpha \in \Re$. *Then,*

$$-\gamma\,\mu(\alpha) \leq \frac{1}{2}\lambda_{\min}\left(Q(\alpha)+Q(\alpha)^T\right) \leq \frac{1}{2}\lambda_{\max}\left(Q(\alpha)+Q(\alpha)^T\right) \leq \Gamma\,\mu(\alpha) \tag{38}$$

for any $\alpha \in [0,\bar{\alpha}]$, *where*

$$\bar{\alpha} \equiv \min\left\{1,\ \frac{\sigma\mu\min\{\gamma,\Gamma\}}{\|X^{-1/2}\Delta X \Delta S X^{1/2}\|}\right\}. \tag{39}$$

Proof. Let $\alpha \in [0,\bar{\alpha}]$ be given. By Lemma 2.2(c), we have $\mu(\alpha) = (1-\alpha+\sigma\alpha)\mu$. This relation, (12), (26), and the fact that $\lambda_{\max}(X^{1/2}SX^{1/2}-\mu I) \leq \Gamma\mu$, $0 \leq \alpha \leq \bar{\alpha} \leq 1$ and $\lambda_{\max}(\cdot)$ is a homogeneous convex function on the space of symmetric matrices imply that

$$\begin{aligned}
&\frac{1}{2}\lambda_{\max}\left(Q(\alpha)+Q(\alpha)^T\right)\\
&\leq (1-\alpha)\lambda_{\max}(X^{1/2}SX^{1/2}-\mu I)\\
&\quad + \frac{1}{2}\alpha^2\lambda_{\max}(X^{-1/2}\Delta X\Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2})\\
&\leq (1-\alpha)\Gamma\mu + \frac{1}{2}\alpha^2\|X^{-1/2}\Delta X\Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2}\|\\
&\leq \Gamma\mu(\alpha) - \alpha\sigma\Gamma\mu + \alpha^2\|X^{-1/2}\Delta X\Delta S X^{1/2}\|\\
&\leq \Gamma\mu(\alpha) - \alpha\left(\sigma\Gamma\mu - \bar{\alpha}\|X^{-1/2}\Delta X\Delta S X^{1/2}\|\right) \ \leq\ \Gamma\mu(\alpha),
\end{aligned}$$

where the last inequality is due to (39). Working with the function $\lambda_{\min}(\cdot)$, which is homogeneous and concave over the space of symmetric matrices, and using (12), (26), (39), and the fact that $\lambda_{\min}(X^{1/2}SX^{1/2}-\mu I) \geq -\gamma\mu$, $0 \leq \alpha \leq \bar{\alpha} \leq 1$ and $\mu(\alpha) = (1-\alpha+\sigma\alpha)\mu$, we obtain

$$\begin{aligned}
&\frac{1}{2}\lambda_{\min}\left(Q(\alpha)+Q(\alpha)^T\right)\\
&\geq (1-\alpha)\lambda_{\min}(X^{1/2}SX^{1/2}-\mu I)\\
&\quad + \frac{1}{2}\alpha^2\lambda_{\min}(X^{-1/2}\Delta X\Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2})\\
&\geq -(1-\alpha)\gamma\mu - \frac{1}{2}\alpha^2\|X^{-1/2}\Delta X\Delta S X^{1/2} + X^{1/2}\Delta S \Delta X X^{-1/2}\|\\
&\geq -\gamma\mu(\alpha) + \alpha\sigma\gamma\mu - \alpha^2\|X^{-1/2}\Delta X\Delta S X^{1/2}\|\\
&\geq -\gamma\mu(\alpha) + \alpha\left(\sigma\gamma\mu - \bar{\alpha}\|X^{-1/2}\Delta X\Delta S X^{1/2}\|\right) \ \geq\ -\gamma\mu(\alpha).
\end{aligned}$$

We have thus shown that (38) holds. □

LEMMA 5.5. *Suppose that* $(X,S,y) \in \mathcal{N}(\gamma,\Gamma)$ *for some* $\gamma \in [0,1)$ *and* $\Gamma \geq 0$ *and let* $(\Delta X, \Delta S, \Delta y)$ *denote the solution of* (3)–(5) *with* $H = \sigma\mu I - X^{1/2}SX^{1/2}$ *and* $\sigma \in [0,1]$. *Let* $(X(\alpha), S(\alpha), y(\alpha))$ *be defined as in* (23) *for any* $\alpha \in \Re$. *Then,* $(X(\alpha), S(\alpha), y(\alpha)) \in \mathcal{N}(\gamma,\Gamma)$ *for any* $\alpha \in [0,\hat{\alpha})$, *where*

$$\hat{\alpha} \equiv \min\left\{1, \frac{1}{\|X^{-1/2}\Delta X X^{-1/2}\|}, \frac{\sigma\mu\min\{\gamma,\Gamma\}}{\|X^{-1/2}\Delta X \Delta S X^{1/2}\|}\right\}. \tag{40}$$

Proof. Fix some $\alpha \in [0,\hat{\alpha})$. We first show that $X(\alpha) \in F^0(P)$. Indeed, using (4) and the fact that X is strictly feasible, we easily see that $A_i \bullet X(\alpha) = b_i$ for every $i = 1, \ldots, m$. By (40) and the fact that $\alpha < \hat{\alpha}$, we have $\alpha\|X^{-1/2}\Delta X X^{-1/2}\| < 1$, which in turn implies that $I + \alpha X^{-1/2}\Delta X X^{-1/2} \succ 0$. Thus, $X(\alpha) \equiv X + \alpha\Delta X = X^{1/2}(I + \alpha X^{-1/2}\Delta X X^{-1/2})X^{1/2} \succ 0$. Hence, $X(\alpha) \in F^0(P)$.

Let $\mu(\alpha)$ and $Q(\alpha)$ be defined as in (24) and (25) and let $W(\alpha) \equiv X^{-1/2}[X(\alpha)]^{1/2}$ and $E(\alpha) \equiv [X(\alpha)]^{1/2}S(\alpha)[X(\alpha)]^{1/2} - \mu(\alpha)I$. Clearly, $W(\alpha)$ is nonsingular and $W(\alpha)E(\alpha)W(\alpha)^{-1} = Q(\alpha)$. In view of Lemma 3.3, we conclude that

$$\frac{1}{2}\lambda_{\min}\left(Q(\alpha) + Q(\alpha)^T\right) \leq \lambda_{\min}(E(\alpha)) \leq \lambda_{\max}(E(\alpha)) \leq \frac{1}{2}\lambda_{\max}\left(Q(\alpha) + Q(\alpha)^T\right).$$

Using this relation, Lemma 5.4, and the fact that $\hat{\alpha} \leq \bar{\alpha}$, we conclude that

$$-\gamma\mu(\alpha) \leq \lambda_{\min}(E(\alpha)) \leq \lambda_{\max}(E(\alpha)) \leq \Gamma\mu(\alpha).$$

To conclude the proof, it remains to show that $(S(\alpha), y(\alpha)) \in F^0(D)$. Indeed, the first inequality of the last relation, the definition of $E(\alpha)$, and the assumption that $\gamma < 1$ imply that

$$\lambda_{\min}\left([X(\alpha)]^{1/2}S(\alpha)[X(\alpha)]^{1/2}\right) \geq (1-\gamma)\mu(\alpha) > 0.$$

Hence, $[X(\alpha)]^{1/2}S(\alpha)[X(\alpha)]^{1/2} \succ 0$, which in turn implies that $S(\alpha) \succ 0$. Using both (5) and the fact that (S,y) is strictly feasible, we easily see that $\sum_{i=1}^n y_i(\alpha)A_i + S(\alpha) = C$. Hence, $(S(\alpha), y(\alpha)) \in F^0(D)$. We have thus shown that $(X(\alpha), S(\alpha), y(\alpha)) \in \mathcal{N}(\gamma,\Gamma)$. □

We now state the following result due to Kojima, Shindoh, and Hara [11].

LEMMA 5.6. *Suppose that* $X \in F^0(P)$, $(S,y) \in F^0(D)$ *and let* $(\Delta X, \Delta S, \Delta y)$ *denote the solution of* (3)–(5) *with* $H \equiv \sigma\mu I - X^{1/2}SX^{1/2}$. *Then,*

$$\|X^{-1/2}\Delta X S^{1/2}\|_F \leq \sqrt{\mu}\,\|\sigma R^{-T} - R\|_F, \tag{41}$$

$$\|X^{-1/2}\Delta X X^{-1/2}\|_F \leq \|R^{-1}\|\,\|\sigma R^{-T} - R\|_F, \tag{42}$$

$$\|S^{-1/2}\Delta S S^{-1/2}\|_F \leq \|R^{-1}\|\,\|\sigma R^{-T} - R\|_F, \tag{43}$$

$$\|S^{-1/2}\Delta S X^{1/2}\|_F \leq \sqrt{\mu}\,\|R\|\,\|R^{-1}\|\,\|\sigma R^{-T} - R\|_F, \tag{44}$$

where $R \equiv \mu^{-1/2}X^{1/2}S^{1/2}$.

Proof. Using the definition of R and standard norm inequalities, it is easy to see that (41) implies (42) and that (43) implies (44). In view of Lemma 2.1, there exists $W \in \mathcal{S}_\perp$ such that $(\Delta X, \Delta S, \Delta y, W)$ is a solution of the system consisting of (4), (5), and the equation $X(\Delta S + W) + \Delta X S = \sigma\mu I - XS$. In view of Corollary 7.7 of [11], we conclude that

$$\|X^{1/2}(\Delta S + W)S^{-1/2}\|_F \leq \|\sigma\mu X^{-1/2}S^{-1/2} - X^{1/2}S^{1/2}\|_F = \sqrt{\mu}\|\sigma R^{-T} - R\|_F \tag{45}$$

and

$$\|X^{-1/2}\Delta X S^{1/2}\|_F \le \|\sigma\mu X^{-1/2}S^{-1/2} - X^{1/2}S^{1/2}\|_F = \sqrt{\mu}\|\sigma R^{-T} - R\|_F,$$

which shows (41). It remains to show (43). Indeed, relation (45) and the definition of R imply that

$$\begin{aligned}\|S^{-1/2}(\Delta S + W)S^{-1/2}\|_F &\le \|S^{-1/2}X^{-1/2}\|\,\|X^{1/2}(\Delta S + W)S^{-1/2}\|_F \\ &\le \|R^{-1}\|\,\|\sigma R^{-T} - R\|_F.\end{aligned}$$

Let $E \equiv S^{-1/2}(\Delta S + W)S^{-1/2}$. Using the fact that $(E + E^T)/2 = S^{-1/2}\Delta S S^{-1/2}$ and $\|(E^T + E)/2\|_F \le \|E\|_F$ with the above inequality, we obtain (43). □

LEMMA 5.7. *Let R be a nonsingular matrix such that $\|R\|_F = \sqrt{n}$. Then, for any $\sigma \in \Re$, we have*

$$\|\sigma R^{-T} - R\|^2 \le n(1 - 2\sigma + \sigma^2\|R^{-1}\|^2).$$

Proof. Using (17), we obtain

$$\text{Tr } (R^T R)^{-1} = \sum_{i=1}^{n} \lambda_i\left((R^T R)^{-1}\right) \le n\lambda_{\max}\left(R^{-1}R^{-T}\right) = n\|R^{-1}\|^2. \tag{46}$$

This relation together with the assumption that $\|R\|_F^2 = n$ imply

$$\begin{aligned}\|R - \sigma R^{-T}\|_F^2 &= \text{Tr }\left(R^T - \sigma R^{-1}\right)\left(R - \sigma R^{-T}\right) \\ &= \text{Tr }\left(R^T R - 2\sigma I + \sigma^2 (R^T R)^{-1}\right) \\ &= \|R\|_F^2 - 2\sigma n + \sigma^2 \text{Tr } (R^T R)^{-1} \\ &\le n(1 - 2\sigma + \sigma^2\|R^{-1}\|^2). \quad \square\end{aligned}$$

We are now ready to prove Theorem 5.2.

Proof of Theorem 5.2. In view of Lemma 5.5, it is sufficient to show that

$$\tilde{\alpha} \le \min\left\{1,\ \frac{1}{\|X^{-1/2}\Delta X X^{-1/2}\|},\ \frac{\sigma\mu\gamma}{\|X^{-1/2}\Delta X \Delta S X^{1/2}\|}\right\}, \tag{47}$$

where $\tilde{\alpha}$ is defined in (37). Indeed, using (17), the definition of R, and the fact that $(1-\gamma)\mu \le \lambda_{\min}(XS) \le \lambda_{\max}(XS) \le (1+\Gamma)\mu$, we have

$$\|R\|^2 = \lambda_{\max}(R^T R) = \frac{\lambda_{\max}(S^{1/2}XS^{1/2})}{\mu} = \frac{\lambda_{\max}(XS)}{\mu} \le (1+\Gamma) \tag{48}$$

and

$$\|R^{-1}\|^2 = \frac{1}{\lambda_{\min}(R^T R)} = \frac{\mu}{\lambda_{\min}(S^{1/2}XS^{1/2})} = \frac{\mu}{\lambda_{\min}(XS)} \le \frac{1}{(1-\gamma)}. \tag{49}$$

By (49) and Lemma 5.7(a), we have

$$\|\sigma R^{-T} - R\|_F^2 \le \left(1 - 2\sigma + \frac{\sigma^2}{1-\gamma}\right) n = \left((1-\sigma)^2 + \frac{\gamma\sigma^2}{1-\gamma}\right) n. \tag{50}$$

Using (41), (44), (48), (49), (50), and (37), we obtain

$$\begin{aligned}\tilde{\alpha}\|X^{-1/2}\Delta X\Delta S X^{1/2}\| &\leq \tilde{\alpha}\|X^{-1/2}\Delta X S^{1/2}\|_F\,\|S^{-1/2}\Delta S X^{1/2}\|_F\\ &\leq \tilde{\alpha}\left[\sqrt{\mu}\,\|\sigma R^{-T}-R\|_F\right]\left[\sqrt{\mu}\,\|R\|\,\|R^{-1}\|\,\|\sigma R^{-T}-R\|_F\right]\\ &\leq \tilde{\alpha}\mu\|R\|\,\|R^{-1}\|\,\|\sigma R^{-T}-R\|_F^2\\ &\leq \tilde{\alpha}\mu\frac{(1+\Gamma)^{1/2}}{(1-\gamma)^{1/2}}\left((1-\sigma)^2+\frac{\gamma\sigma^2}{1-\gamma}\right)n \;=\; \sigma\gamma\mu.\end{aligned}$$

Moreover, using (42), (49), (50), and (37), we obtain

$$\begin{aligned}\tilde{\alpha}\|X^{-1/2}\Delta X X^{-1/2}\| &\leq \tilde{\alpha}\,\|R^{-1}\|\,\|\sigma R^{-T}-R\|_F\\ &\leq \tilde{\alpha}\,\frac{1}{(1-\gamma)^{1/2}}\left((1-\sigma)^2+\frac{\gamma\sigma^2}{1-\gamma}\right)^{1/2} n^{1/2}\\ &= \frac{\sigma\gamma}{(1+\Gamma)^{1/2}n^{1/2}}\left((1-\sigma)^2+\frac{\gamma\sigma^2}{1-\gamma}\right)^{-1/2}\\ &\leq \frac{\sigma\gamma}{(1+\Gamma)^{1/2}n^{1/2}}\,\frac{(1-\gamma)^{1/2}}{\gamma^{1/2}\sigma}\\ &= \frac{\gamma^{1/2}(1-\gamma)^{1/2}}{(1+\Gamma)^{1/2}n^{1/2}} \;<\; 1.\end{aligned}$$

It is also easy to see that $\tilde{\alpha} \leq 1$. We have thus shown that (47) holds. □

6. Concluding remarks. In this paper, we have provided results which make the task of extending polynomially convergent primal–dual path-following algorithms to SDP a routine exercise. We have illustrated these results for two well-known feasible interior-point path-following algorithms: a short-step and a long-step method. The author believes that similar techniques can be used to extend other polynomially convergent *feasible or infeasible* interior-point path-following methods to the context of SDP.

Acknowledgment. The author has mistakenly claimed in a previous version of this manuscript that the search direction determined by systems (3), (4), and (5) (respectively, (7)) was new. The author thanks Prof. Masakazu Kojima for kindly pointing out to the author that this direction was just a special one from the family of search directions introduced in [11]. The author also thanks Prof. Yin Zhang for pointing out an error in the statement of Lemma 3.2 in a previous version of this manuscript.

REFERENCES

[1] I. Adler and F. Alizadeh, *Primal–Ddual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems*, manuscript.

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

[3] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, *Primal–Dual Interior-Point Methods for Semidefinite Programming*, manuscript. Presented at the Math Programming Symposium, Ann Arbor, MI, 1994.

[4] G. H. Golub and C. E. Van Loan, *Matrix Computations: Second Edition*, The John Hopkins University Press, Baltimore, MD, 1989.

[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] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

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

[9] 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.

[10] 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, Berlin, New York, 1989, pp. 29–47.

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

[13] 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.

[14] 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).

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

[16] 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.

[17] Y. E. Nesterov and A. S. Nemirovskii, *Interior Point Methods in Convex Programming: Theory and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[18] Y. E. Nesterov and M. Todd, *Primal–Dual Interior-Point Methods for Self-Scaled Cones*, Technical report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1995.

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

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

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