A GENERAL FRAMEWORK FOR ESTABLISHING POLYNOMIAL CONVERGENCE OF LONG-STEP METHODS FOR SEMIDEFINITE PROGRAMMING

SAMUEL BURER*

Department of Management Sciences, University of Iowa, Iowa City, Iowa 52242-1000, USA

RENATO D. C. MONTEIRO†

School of ISyE, Georgia Tech, Atlanta, Georgia 30332, USA

This paper considers feasible long-step primal-dual path-following methods for semidefinite programming based on Newton directions associated with central path equations of the form $\Phi(PXP^T, P^{-T}SP^{-1}) - \nu I = 0$, where the map Φ and the nonsingular matrix P satisfy several key properties. An iteration-complexity bound for the long-step method is derived in terms of an upper bound on a certain scaled norm of the second derivative of Φ. As a consequence of our general framework, we derive polynomial iteration-complexity bounds for long-step algorithms based on the following four maps: $\Phi(X,S) = (XS+SX)/2$, $\Phi(X,S) = L_x^T S L_x$, $\Phi(X,S) = X^{1/2}SX^{1/2}$, and $\Phi(X,S) = W^{1/2}XSW^{-1/2}$, where L_x is the lower Cholesky factor of X and W is the unique symmetric matrix satisfying $S = WXW$.

KEY WORDS: semidefinite programming, interior-point methods, path-following methods, long-step methods, Newton directions, central path

1 Introduction

Semidefinite programming (SDP) is a generalization of linear programming (LP) in which a linear function of a symmetric matrix variable X is minimized over an affine subspace of real symmetric matrices subject to the constraint that X be positive semidefinite. Semidefinite programming shares many features of linear programming, including a large number of applications, a rich duality theory, and the ability to be solved (more precisely, approximated) in polynomial time.

In the past several years, a major part of the research into SDP has focused on both the theoretical and practical solution of SDP problems using extensions of interior-point methods for LP. Many authors have proposed interior-point algorithms for solving SDP problems (see for example [1, 2, 4, 6, 7, 8, 9, 10, 11, 13, 14, 17, 18, 19, 20, 21, 22, 24, 25, 26]). Many of the recent works on interior-point algorithms for SDP are concentrated on primal-dual methods. Feasible primal-dual path-following algorithms for SDP simultaneously solve the primal and dual SDP problems by maintaining primal feasibility in X and dual feasibility in (S, y) while iteratively solving the system $XS = 0$. The key idea is to follow the central path by moving in the direction obtained by the application of Newton's method to the central path equation $XS = \nu I$. Newton's method, however, results in an equation of the form

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

which in general yields nonsymmetric directions. Many authors have investigated alternate yet equivalent equations of the central path for which Newton's method does yield symmetric directions (see for example [2, 4, 7, 10, 11, 13, 14, 17, 20, 24]).

* The work of this author was based on research supported bythe National Science Foundation under grants INT-9600343, INT-9910084, CCR-9700448, CCR-9902010 and CCR-0203426.

† The work of this author was based on research supported bythe National Science Foundation under grants INT-9600343, INT-9910084,CCR-9700448, CCR-9902010, and CCR-0203113.

This paper considers feasible long-step primal-dual path-following methods for SDP based on search directions contained in the family of all symmetric Newton directions obtained from central path equations of the form

$$\Phi(PXP^T, P^{-T}SP^{-1}) - \nu I = 0, \tag{2}$$

where Φ is a fixed map satisfying several key properties and P is an arbitrary nonsingular matrix. To reflect the role of Φ, we generalize the concepts of the duality gap of a primal-dual point and the distance of a primal-dual point from the central path, and then prove that the sequence of iterates produced by the long-step algorithm converges to the optimal primal-dual solution set of the underlying primal and dual SDP problems. As a consequence of the convergence analysis, we express the iteration-complexity of the long-step algorithm in terms of an upper bound on a certain scaled norm of the second derivative of Φ.

We demonstrate how the long-step framework applies to the following specific central-path maps:

$$\Phi(X,S) = (XS+SX)/2, \tag{3}$$
$$\Phi(X,S) = L_x^T S L_x, \tag{4}$$
$$\Phi(X,S) = X^{1/2} S X^{1/2}, \tag{5}$$
$$\Phi(X,S) = W^{1/2} X S W^{-1/2}, \tag{6}$$

where L_x is the lower Cholesky factor of X and W is the unique symmetric matrix satisfying $S = WXW$. Polynomially convergent long-step algorithms for the first and third maps have been given in Monteiro and Zhang [17] and Monteiro and Tsuchiya [13], respectively, for which respective iteration-complexities of $\mathcal{O}(n\sqrt{\kappa}L)$ and $\mathcal{O}(n^{3/2}L)$ have been established, where κ is a certain condition number determined by the sequence of scaling matrices $\{P^k\}$. To illustrate the usefulness and generality of our approach, we use our framework to rederive the polynomial convergence of these two algorithms in a unified way; more specifically, we obtain the same iteration-complexity of $\mathcal{O}(n\sqrt{\kappa}L)$ for the first map and a slightly worse iteration-complexity of $\mathcal{O}(n^2L)$ for the third map. The second and fourth maps have been studied in [15], [16] and [23], yet no polynomial convergence analysis of long-step algorithms based on these maps has been established. We show that the second and fourth maps also fit nicely into our general framework and hence obtain for the first time polynomially convergent feasible long-step algorithms for these maps, having respective iteration-complexities of $\mathcal{O}(n^2L)$ and $\mathcal{O}(n^3L)$.

This paper is organized as follows. In Section 2, we introduce the SDP problem, its central path, and the reformulation of the central path using the map Φ. We also state sufficient conditions for the Newton direction of (2) to exist and define the concepts of the duality gap and centrality measure, which are then incorporated into a long-step algorithm. In Section 3, we develop the analysis of the long-step method and, under certain assumptions, prove the convergence of the algorithm. Section 4 details how the four maps given above fit into our framework and establishes the polynomiality of the long-step algorithm for each map.

1.1 Notation and Terminology

The following notation is used throughout this paper. The superscript ^{T} denotes transpose. Let $\Re^m$ denote the m-dimensional real Euclidean space and $\Re^{m\times n}$ denote the space of $m \times n$ real matrices. Let $\mathcal{S}^n$ denote the space of $n \times n$ real symmetric matrices, and $\mathcal{S}^n_+$ and $\mathcal{S}^n_{++}$ denote the subsets of $\mathcal{S}^n$ consisting of the positive semidefinite and positive definite matrices, respectively. The symbols $\succeq$ and $\succ$ denote, respectively, the positive semidefinite and positive definite ordering over the set of symmetric matrices; that is, for $X, Y \in \mathcal{S}^n$, $X \succeq Y$ (or $Y \preceq X$) means $X - Y \in \mathcal{S}^n_+$, and $X \succ Y$ (or $Y \prec X$) means $X - Y \in \mathcal{S}^n_{++}$. For $A \in \Re^{n\times n}$, let $\mathrm{Tr}\,(A) \equiv \sum_{i=1}^n A_{ii}$ denote the trace of A. For $P, Q \in \Re^{m\times n}$, let $P \bullet Q \equiv \mathrm{Tr}\ P^TQ$ denote standard inner product in $\Re^{m\times n}$. For a matrix $A \in \Re^{n\times n}$ with all real eigenvalues, we denote its smallest and largest eigenvalues by $\lambda_{\min}[A]$ and $\lambda_{\max}[A]$, respectively. The operator norm of a matrix $P \in \Re^{m\times n}$ is $\|P\| \equiv [\lambda_{\max}(P^TP)]^{1/2}$ and its Frobenius norm is $\|P\|_F \equiv (P \bullet P)^{1/2}$. The $-\infty$ seminorm of a matrix $A \in \mathcal{S}^n$ is $\|A\|_{-\infty} \equiv \max\{0\,,\,\lambda_{\max}[-A]\}$. For any matrix $A \in \Re^{n\times n}$, we denote the symmetric matrix $A + A^T$ by $\mathrm{sim}(A)$.

Let $\mathcal{L}^n$ denote the space of lower triangular matrices in $\Re^{n\times n}$, and $\mathcal{L}^n_{++}$ the subset of $\mathcal{L}^n$ consisting of lower triangular matrices with positive diagonal elements. It is well known that for any matrix $V \in \mathcal{S}^n_{++}$, there exists a unique $L \in \mathcal{L}^n_{++}$ such that $V = LL^T$ (see Theorem 4.2.5 in [3]). The matrix L is called the Cholesky factor of V and is denoted by $\mathrm{chol}\,(V)$.

Let $\Xi : \mathcal{S}^n \to \mathcal{S}^n$ be a linear operator. Its adjoint is the unique linear operator $\Xi^* : \mathcal{S}^n \to \mathcal{S}^n$, such that $\Xi(A) \bullet B = A \bullet \Xi^*(B)$ for all $A, B \in \mathcal{S}^n$. Ξ is called symmetric if $\Xi = \Xi^*$ and positive definite if $\Xi(A) \bullet A > 0$ for all $0 \neq A \in \mathcal{S}^n$. We let $\mathcal{T}^n$ denote the set of all symmetric operators from $\mathcal{S}^n$ to $\mathcal{S}^n$, and $\mathcal{T}^n_{++}$ denote the subset of $\mathcal{T}^n$ consisting of the positive definite operators. For $\Xi \in \mathcal{T}^n$, we define $\|\Xi\| \equiv \max\{\|\Xi(A)\|_F : \|A\|_F \leq 1\}$. It can be shown that if $\Xi \in \mathcal{T}^n_{++}$ then $\|\Xi\| = \max\{A \bullet \Xi(A) : \|A\|_F \leq 1\}$. Any $\Xi \in \mathcal{T}^n_{++}$ has a square root operator, which we denote by $\Xi^{1/2}$; hence $\Xi^{1/2}(\Xi^{1/2}(A)) = \Xi(A)$ for all $A \in \mathcal{S}^n$.

2 The Long-Step Algorithm for General Central Path Equations

In this section, we describe the SDP problem studied in this paper and its associated central path, and we propose a general central path system (parametrized by a positive scalar) whose solution set yields the central path. Using this general central path system and a scaling procedure, we then define a scaled Newton direction which is used to describe a primal-dual long-step path-following algorithm for solving the SDP problem.

2.1 The SDP problem and central path

This subsection describes the SDP problem, its central path, and the general central path system. It also contains some notation and terminology which we use throughout our presentation.

We consider the primal SDP problem

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

and its dual SDP problem

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

where $\widetilde{C} \in \mathcal{S}^n$, $\widetilde{A}_i \in \mathcal{S}^n, i = 1, \ldots, m$, and $\widetilde{b} = (\widetilde{b}_1, \ldots, \widetilde{b}_m)^T \in \Re^m$ are the data and $\widetilde{X} \in \mathcal{S}^n_+$ and $(\widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_+ \times \Re^m$ are the primal and dual variables, respectively. We denote the set of interior feasible solutions of $(\widetilde{P})$ and $(\widetilde{D})$ respectively by

$$\mathcal{F}^0(\widetilde{P}) \equiv \left\{\widetilde{X} \in \mathcal{S}^n : \widetilde{A}_i \bullet \widetilde{X} = \widetilde{b}_i,\ i = 1, \ldots, m,\ \widetilde{X} \succ 0\right\},$$

$$\mathcal{F}^0(\widetilde{D}) \equiv \left\{(\widetilde{S}, \widetilde{y}) \in \mathcal{S}^n \times \Re^m : \sum_{i=1}^m \widetilde{y}_i \widetilde{A}_i + \widetilde{S} = \widetilde{C},\ \widetilde{S} \succ 0\right\}.$$

Furthermore, we denote $\mathcal{F}^0(\widetilde{P}) \times \mathcal{F}^0(\widetilde{D})$ as $\widetilde{\mathcal{F}}^0$. We assume throughout that $\widetilde{\mathcal{F}}^0 \neq \emptyset$ and that the matrices $\widetilde{A}_1, \ldots, \widetilde{A}_m$ are linearly independent. Under the first assumption, it is well-known that both $(\widetilde{P})$ and $(\widetilde{D})$ have optimal solutions $\widetilde{X}^*$ and $(\widetilde{S}^*, \widetilde{y}^*)$ such that $\widetilde{C} \bullet \widetilde{X}^* = \widetilde{b}^T \widetilde{y}^*$. This last condition, called strong duality, can alternatively be expressed as $\widetilde{X}^* \bullet \widetilde{S}^* = 0$ or $\widetilde{X}^* \widetilde{S}^* = 0$. Thus, the set of primal-dual optimal solutions consists of all the solutions $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+ \times \Re^m$ to the optimality system

$$\widetilde{X}\widetilde{S} = 0, \tag{7a}$$

$$\textstyle\sum_{i=1}^m \widetilde{y}_i \widetilde{A}_i + \widetilde{S} - \widetilde{C} = 0, \tag{7b}$$

$$\widetilde{A}_i \bullet \widetilde{X} - \widetilde{b}_i = 0,\ i = 1, \ldots, m. \tag{7c}$$

It is well-known that for every $\nu > 0$, the perturbed system formed from (7) by replacing (7a) with $\widetilde{X}\widetilde{S} = \nu I$ has a unique solution, denoted by $(\widetilde{X}_\nu, \widetilde{S}_\nu, \widetilde{y}_\nu)$, and that the limit $\lim_{\nu \to 0}(\widetilde{X}_\nu, \widetilde{S}_\nu, \widetilde{y}_\nu)$ exists and is a solution of (7). The set of all solutions $\{(\widetilde{X}_\nu, \widetilde{S}_\nu, \widetilde{y}_\nu) : \nu > 0\}$ is known as the central path.

The equation $\widetilde{X}\widetilde{S} = \nu I$ provides the canonical formulation of the central path, and in this paper, we consider alternative equations which lead to different formulations of the central path. More precisely, we consider equations of the form

$$\Phi(\widetilde{X}, \widetilde{S}) = \nu I \tag{8}$$

where $\Phi : \mathcal{D} \to \mathcal{S}^n$ is a map whose domain satisfies

$$(\mathcal{S}^n_{++} \times \mathcal{S}^n_{++}) \cup \{(X, S) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+ : X \bullet S = 0\} \subseteq \mathcal{D} \subseteq \mathcal{S}^n_+ \times \mathcal{S}^n_+,$$

and the following assumptions (in addition to others that will be presented later) hold:

Assumption 1. *The map $\Phi : \mathcal{D} \to \mathcal{S}^n$ is continuous and twice continuously differentiable on $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$.*

Assumption 2. *There exists a homeomorphism $\varphi : \Re_+ \to \Re_+$ such that, for all $\nu \geq 0$ and $(\widetilde{X}, \widetilde{S}) \in \mathcal{D}$, $\Phi(\widetilde{X}, \widetilde{S}) = \nu I$ if and only if $\widetilde{X}\widetilde{S} = \varphi(\nu) I$.*

The need for considering maps Φ whose domains are not equal to the whole set $\mathcal{S}^n_+ \times \mathcal{S}^n_+$ is illustrated by the map (4) which is not well-defined in the boundary of $\mathcal{S}^n_+ \times \mathcal{S}^n_+$. However, we can continuously extend this map Φ to the set $\{(X, S) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+ : X \bullet S = 0\}$ by defining it to be identically zero there, thereby obtaining a map which satisfies the assumptions.

The reformulation (8) is necessary in light of the fact that the system given by the equations $\widetilde{X}\widetilde{S} = \nu I$, (7b), and (7c) maps a point in $\mathcal{S}^n \times \mathcal{S}^n \times \Re^m$ into $\Re^{n \times n} \times \mathcal{S}^n \times \Re^m$, and hence a corresponding Newton direction is not defined.

Polynomial convergence of feasible primal-dual long-step path-following algorithms for solving problems $(\widetilde{P})$ and $(\widetilde{D})$ based on the maps (3) and (5) has been given in [17] and [13], respectively. In this paper, we seek to unify these two earlier approaches by providing a long-step algorithm based on the formulation of the central path given by equation (8), where $\Phi(\widetilde{X}, \widetilde{S})$ is taken to be an arbitrary map satisfying Assumptions 1 and 2 and certain other assumptions, which we detail in the remainder of this section and in Section 3.

2.2 A family of scaled Newton directions

In this subsection, we derive a family of scaled Newton directions for the perturbed optimality system given by (7b), (7c), and (8) and state sufficient conditions for the directions in this family to exist.

Given a fixed, nonsingular matrix $\widetilde{P} \in \Re^{n \times n}$ and a point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$, consider the change of variables and data

$$X \equiv \widetilde{P}\widetilde{X}\widetilde{P}^T, \quad (S, y) \equiv (\widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}, \widetilde{y}) \tag{9}$$

and

$$C \equiv \widetilde{P}^{-T}\widetilde{C}\widetilde{P}^{-1}, \quad (A_i, b_i) \equiv (\widetilde{P}^{-T}\widetilde{A}_i\widetilde{P}^{-1}, \widetilde{b}_i), \; i = 1, \ldots, m. \tag{10}$$

This change allows us to recast problems $(\widetilde{P})$ and $(\widetilde{D})$ as

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

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

We define the sets $\mathcal{F}^0(P)$, $\mathcal{F}^0(D)$, and $\mathcal{F}^0$ for problems (P) and (D) in analogy with the corresponding sets for problems $(\widetilde{P})$ and $(\widetilde{D})$.

Given a parameter $\nu > 0$, the perturbed optimality system for problems (P) and (D)

$$\Phi(X, S) = \nu I, \tag{11a}$$

$$\textstyle\sum_{i=1}^m y_i A_i + S - C = 0, \tag{11b}$$

$$A_i \bullet X - b_i = 0, \; i = 1, \ldots, m, \tag{11c}$$

gives rise to the pure Newton direction $(\Delta X, \Delta S, \Delta y)$ determined by the equations

$$\Phi_x(\Delta X) + \Phi_s(\Delta S) \quad = \quad H, \tag{12a}$$

$$\sum_{i=1}^m \Delta y_i A_i + \Delta S \quad = \quad R, \tag{12b}$$

$$A_i \bullet \Delta X \quad = \quad r_i, \quad i = 1, \ldots, m, \tag{12c}$$

where the linear operators $\Phi_x, \Phi_s : \mathcal{S}^n \to \mathcal{S}^n$ are defined as

$$\Phi_x(A) \equiv \Phi'(X,S)[A,0], \quad \forall\, A \in \mathcal{S}^n,$$
$$\Phi_s(B) \equiv \Phi'(X,S)[0,B], \quad \forall\, B \in \mathcal{S}^n$$

and where

$$H \;=\; \nu I - \Phi(X,S), \tag{13a}$$
$$R \;=\; C - \sum_{i=1}^m y_i A_i - S, \tag{13b}$$
$$r_i \;=\; b_i - A_i \bullet X, \quad i = 1,\ldots,m. \tag{13c}$$

A nonsingular matrix $\widetilde{P} \in \Re^{n\times n}$ and a parameter $\nu > 0$ determine a scaled Newton direction at a given interior feasible point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{F}}^0$ as follows. Using the change of variables and data given by (9) and (10), we obtain an interior feasible point $(X,S,y) \in \mathcal{F}^0$. We then compute the pure Newton direction $(\Delta X, \Delta S, \Delta y)$ of system (12) with (H,R,r) given by (13). Notice that $R = 0$ and $r = 0$ since (X,S,y) is a primal-dual feasible solution. Finally, we map $(\Delta X, \Delta S, \Delta y)$ back to the original space, obtaining

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

Hence, for fixed $\nu > 0$ we obtain a family of Newton directions at $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{F}}^0$ by varying the choice of scaling matrix $\widetilde{P}$. Clearly, the scaled Newton direction is well-defined only when the scaling matrix $\widetilde{P}$ is such that the corresponding system (12) is nonsingular. Theorem 2.2 below gives sufficient conditions on the scaled point (X,S,y) for system (12) to be nonsingular.

Let

$$\mathcal{C} \equiv \{(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} : \Phi_x\Phi_s^* \in \mathcal{T}^n_{++}\}.$$

It is easy to see that for any $(X,S) \in \mathcal{C}$, the operators Φ_x and Φ_s are invertible and that $\Phi_s^{-1}\Phi_x \in \mathcal{T}^n_{++}$. Hence, for $(X,S) \in \mathcal{C}$, both operators $(\Phi_x\Phi_s^*)^{-1}$ and $\Phi_s^{-1}\Phi_x$ have square roots, which we denote as

$$\mathcal{G}_{(X,S)} \equiv \left(\Phi_s^{-1}\Phi_x\right)^{1/2}, \qquad \mathcal{R}_{(X,S)} \equiv \left(\Phi_x\Phi_s^*\right)^{-1/2}. \tag{14}$$

Lemma 2.1. *Let $(X,S) \in \mathcal{C}$, $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, and $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$. Let $H \in \mathcal{S}^n$ be arbitrary, and suppose that $\Phi_x(U) + \Phi_s(V) = H$ with $U \bullet V \geq 0$ for some $U, V \in \mathcal{S}^n$. Then*

$$\|\mathcal{R}(H)\|_F^2 \geq \|\mathcal{G}(U)\|_F^2 + \|\mathcal{G}^{-1}(V)\|_F^2, \tag{15}$$

and equality holds if and only if $U \bullet V = 0$.

Proof. Using that $\Phi_x(U) + \Phi_s(V) = H$, $\mathcal{G} \in \mathcal{T}^n_{++}$, $\mathcal{R} \in \mathcal{T}^n_{++}$, and $U \bullet V \geq 0$, we obtain

$$\begin{aligned}
\|\mathcal{R}(H)\|_F^2 &= \Phi_x^{-1}(H) \bullet \Phi_s^{-1}(H) = \left(U + \mathcal{G}^{-2}(V)\right) \bullet \left(\mathcal{G}^2(U) + V\right) \\
&= U \bullet \mathcal{G}^2(U) + U \bullet V + \mathcal{G}^{-2}(V) \bullet \mathcal{G}^2(U) + \mathcal{G}^{-2}(V) \bullet V \\
&= U \bullet \mathcal{G}^2(U) + 2\, U \bullet V + \mathcal{G}^{-2}(V) \bullet V \\
&\geq U \bullet \mathcal{G}^2(U) + \mathcal{G}^{-2}(V) \bullet V = \|\mathcal{G}(U)\|_F^2 + \|\mathcal{G}^{-1}(V)\|_F^2,
\end{aligned}$$

with equality holding if and only if $U \bullet V = 0$. ∎

Theorem 2.2. *Let $(X,S,y) \in \mathcal{C} \times \Re^m$. Then, for any $(H,R,r) \in \mathcal{S}^n \times \mathcal{S}^n \times \Re^m$, system (12) has a unique solution. In particular, the Newton direction of system (12) with (H,R,r) given by (13) exists at (X,S,y).*

Proof. In terms of $(\Delta X, \Delta S, \Delta y)$, the left-hand side of system (12) is a linear map from $\mathcal{S}^n \times \mathcal{S}^n \times \Re^m$ into itself. Hence, to establish the theorem it suffices to show that $(\Delta X, \Delta S, \Delta y) = (0,0,0)$ is the only solution in $\mathcal{S}^n \times \mathcal{S}^n \times \Re^m$ of system (12) with $(H,R,r) = (0,0,0)$. Indeed, let $(\Delta X, \Delta S, \Delta y)$ be such a solution. Using (12b) and (12c), we easily see that $\Delta X \bullet \Delta S = 0$. Hence, applying Lemma 2.1 with $U = \Delta X$, $V = \Delta S$ and $H = 0$, we see that

$$0 = \|\mathcal{R}(H)\|_F^2 = \Phi_x^{-1}(H) \bullet \Phi_s^{-1}(H) = \|\mathcal{G}(\Delta X)\|_F^2 + \|\mathcal{G}^{-1}(\Delta S)\|_F^2,$$

where $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$ and $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$. This relation together with the positive definiteness of $\mathcal{G}$ implies that $\Delta X = \Delta S = 0$. It now follows from (12b) with $R = \Delta S = 0$ and the linear independence of $A_1, \ldots, A_m$ that $\Delta y = 0$. ■

In light of Theorem 2.2, to obtain a Newton direction $(\Delta\widetilde{X}, \Delta\widetilde{S}, \Delta\widetilde{y})$ at a point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{F}}^0$, we can simply choose as our scaling matrix a nonsingular $\widetilde{P} \in \Re^{n\times n}$ such that the scaled pair (X, S) given by (9) is in $\mathcal{C}$.

It is worth noting that the scaled Newton direction can also be obtained as the Newton direction for the following system of equations in terms of the original variables $(\widetilde{X}, \widetilde{S}, \widetilde{y})$:

$$\begin{gathered} \Phi(\widetilde{P}\widetilde{X}\widetilde{P}^T, \widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}) = \nu I, \\ \textstyle\sum_{i=1}^m \widetilde{y}_i \widetilde{A}_i + \widetilde{S} - \widetilde{C} = 0, \\ \widetilde{A}_i \bullet \widetilde{X} - \widetilde{b}_i = 0, \ i = 1, \ldots, m. \end{gathered}$$

2.3 A centrality measure

In this subsection, we use the seminorm $\|\cdot\|_{-\infty}$ to define a measure of centrality of a point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$ and then use it to define a neighborhood of the central path associated with $(\widetilde{P})$ and $(\widetilde{D})$ which will play an important role in the design and analysis of the long-step algorithm.

For any $(\widetilde{X}, \widetilde{S}) \in \mathcal{D}$, the quantity

$$\mu(\widetilde{X}, \widetilde{S}) \equiv \frac{\Phi(\widetilde{X}, \widetilde{S}) \bullet I}{n}$$

is referred to as the duality gap at $(\widetilde{X}, \widetilde{S})$ (with respect to the map Φ). Furthermore, given $(\widetilde{X}, \widetilde{S}), (X, S) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+$, we write $(\widetilde{X}, \widetilde{S}) \sim (X, S)$ if there exists a nonsingular $n \times n$ matrix $\widetilde{P}$ such that $X = \widetilde{P}\widetilde{X}\widetilde{P}^T$ and $S = \widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}$.

We make two additional assumptions relating the map $\Phi : \mathcal{D} \to \mathcal{S}^n$ with its associated duality gap map μ.

Assumption 3. *$\mu(\widetilde{X}, \widetilde{S}) = \mu(X, S)$ for all $(\widetilde{X}, \widetilde{S})$, $(X, S) \in \mathcal{D}$ such that $(\widetilde{X}, \widetilde{S}) \sim (X, S)$.*

Assumption 4. *If $\mu(\widetilde{X}, \widetilde{S}) = 0$ for a point $(\widetilde{X}, \widetilde{S}) \in \mathcal{D}$, then $\Phi(\widetilde{X}, \widetilde{S}) = 0$.*

Given a point $(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \Re^m$, we define the distance measure of $(\widetilde{X}, \widetilde{S}, \widetilde{y})$ from the central path associated with problems $(\widetilde{P})$ and $(\widetilde{D})$ as the quantity

$$d_{-\infty}(\widetilde{X}, \widetilde{S}) \equiv \inf\left\{ \|\Phi(X, S) - \mu(X, S)I\|_{-\infty} : (X, S) \sim (\widetilde{X}, \widetilde{S}) \right\}. \tag{16}$$

Moreover, given $\gamma \in (0, 1)$, we define

$$\begin{aligned} \mathcal{N}_{-\infty}(\gamma) &\equiv \{(\widetilde{X}, \widetilde{S}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} : d_{-\infty}(\widetilde{X}, \widetilde{S}) \le \gamma\mu(\widetilde{X}, \widetilde{S})\}, \\ \widetilde{\mathcal{F}}^0_{-\infty}(\gamma) &\equiv \{(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{F}}^0 : (\widetilde{X}, \widetilde{S}) \in \mathcal{N}_{-\infty}(\gamma)\}, \end{aligned} \tag{17}$$

and, for a fixed nonsingular $\widetilde{P} \in \Re^{n\times n}$,

$$\mathcal{F}^0_{-\infty}(\gamma) \equiv \{(X, S, y) \in \mathcal{F}^0 : (X, S) \in \mathcal{N}_{-\infty}(\gamma)\}.$$

It is easy to see that $d_{-\infty}(\widetilde{X}, \widetilde{S}) = d_{-\infty}(X, S)$ for all pairs $(\widetilde{X}, \widetilde{S}), (X, S)$ such that $(\widetilde{X}, \widetilde{S}) \sim (X, S)$ and that $d_{-\infty}(\widetilde{X}, \widetilde{S})$ equals zero if $(\widetilde{X}, \widetilde{S}, \widetilde{y})$ is a point on the central path. Furthermore, for a fixed nonsingular $\widetilde{P} \in \Re^{n\times n}$, if (X, S, y) is given by (9), then

$$(\widetilde{X}, \widetilde{S}, \widetilde{y}) \in \widetilde{\mathcal{F}}^0_{-\infty}(\gamma) \Longleftrightarrow (X, S, y) \in \mathcal{F}^0_{-\infty}(\gamma) \tag{18}$$

due to Assumption 3 and the fact that $d_{-\infty}(\widetilde{X}, \widetilde{S}) = d_{-\infty}(X, S)$.

2.4 The long-step path-following algorithm

In this subsection, we state the generic feasible primal-dual long-step path-following algorithm based on the scaled Newton direction defined in Subsection 2.2.

The full specification of the long-step algorithm requires the choice of a constant $L > 1$ which is used in the algorithm's stopping criterion, a neighborhood-opening constant $\gamma \in (0,1)$, and a centrality parameter $\sigma \in (0,1)$. Furthermore, it requires the specification of a nonempty subset $\mathcal{C}_0$ of $\mathcal{C}$ satisfying Assumptions 5 through 9 stated below.

Given $(\widetilde{X}, \widetilde{S}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, define

$$\mathcal{P}_0(\widetilde{X}, \widetilde{S}) \equiv \{\widetilde{P} \in \Re^{n\times n} : \widetilde{P} \text{ is nonsingular and } (\widetilde{P}\widetilde{X}\widetilde{P}^T, \widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}) \in \mathcal{C}_0\}. \tag{19}$$

The first two assumptions on the set $\mathcal{C}_0$ are as follows.

Assumption 5. *$\mathcal{C}_0 \subseteq \mathcal{C}$ and $\mathcal{P}_0(\widetilde{X}, \widetilde{S}) \neq \emptyset$ for all $(\widetilde{X}, \widetilde{S}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$.*

Assumption 6. *$d_{-\infty}(X,S) = \|\Phi(X,S) - \mu(X,S)I\|_{-\infty}$ for every $(X,S) \in \mathcal{C}_0$.*

From a practical viewpoint, the choice of $\mathcal{C}_0$ is made so that the scaling matrix $\widetilde{P}$ and the scaled iterates $X \equiv \widetilde{P}\widetilde{X}\widetilde{P}^T$ and $S \equiv \widetilde{P}^{-T}\widetilde{S}\widetilde{P}^{-1}$ have some desirable properties. As an example, we will see in Section 4 that the choice $\mathcal{C}_0 = \{(X,S) \in \mathcal{C} : XS = SX\}$ is suitable for establishing polynomial convergence of the long-step algorithm for two specific instances of the map Φ.

For $\sigma \in \Re$ and $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, define

$$H_\sigma(X,S) \equiv \sigma\mu(X,S)I - \Phi(X,S). \tag{20}$$

Moreover, given $(X,S,y) \in \mathcal{C} \times \Re^m$ and $\nu > 0$, let

$$(X_\alpha, S_\alpha, y_\alpha) \equiv (X,S,y) + \alpha(\Delta X, \Delta S, \Delta y), \tag{21}$$

where $(\Delta X, \Delta S, \Delta y)$ is the Newton direction for system (12) with (H,R,r) given by (13). We are now ready to state the long-step algorithm studied in this paper.

Long-Step Algorithm:

Let $L > 1$, $\gamma \in (0,1)$, $\sigma \in (0,1)$, and $(\widetilde{X}^0, \widetilde{S}^0, \widetilde{y}^0) \in \widetilde{\mathcal{F}}^0_{-\infty}(\gamma)$ be given. Set $\mu_0 = \mu(\widetilde{X}^0, \widetilde{S}^0)$ and $k = 0$. Let $\mathcal{C}_0$ be a nonempty subset of $\mathcal{C}$ satisfying Assumption 5.

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

1. Let $(\widetilde{X}, \widetilde{S}, \widetilde{y}) = (\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k)$.
2. Choose $\widetilde{P} \in \mathcal{P}_0(\widetilde{X}, \widetilde{S})$, and let (X,S,y) be given by (9).
3. Compute the solution $(\Delta X, \Delta S, \Delta y)$ of system (12) with C, b, and A_i, for $i = 1, \ldots, m$, given by (10) and $(H,R,r) = (H_\sigma(X,S), 0, 0)$.
4. Let $\widehat{\alpha} \geq 0$ be the largest scalar such that $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0_{-\infty}(\gamma)$ for all $\alpha \in [0, \widehat{\alpha}]$. Choose a stepsize $\alpha_k > 0$ such that $\mu(X_{\alpha_k}, S_{\alpha_k}) \leq \mu(X_\alpha, S_\alpha)$ for all $\alpha \in [0, \widehat{\alpha}]$.
5. Set $(\widetilde{X}^{k+1}, \widetilde{S}^{k+1}, \widetilde{y}^{k+1}) = (\widetilde{P}^{-1}X_{\alpha_k}\widetilde{P}^{-T}, \widetilde{P}^T S_{\alpha_k}\widetilde{P}, y_{\alpha_k})$ and $\mu_{k+1} = \mu(\widetilde{X}^{k+1}, \widetilde{S}^{k+1})$.
6. Increment k by 1.

End

Notice that (17) implies that each iterate $(\widetilde{X}^k, \widetilde{S}^k, \widetilde{y}^k)$ generated by the long-step algorithm has a nonnegative duality gap, i.e., $\mu(\widetilde{X}^k, \widetilde{S}^k) \geq 0$ for all $k \geq 0$. Hence, the long-step algorithm clearly produces a sequence of primal-dual feasible interior points for $(\widetilde{P})$ and $(\widetilde{D})$ such that the corresponding sequence of duality gaps converges to some nonnegative value. In Section 3, we prove that the sequence of duality gaps does in fact converge to zero under Assumptions 1, 2, 3, 5, and 6 presented in this section as well as Assumptions 7 through 9 which we introduce in Section 3. Note that Assumption 4 is not needed to prove the convergence of the duality gaps to zero. Assumption 4, however, together with Assumptions 1 and 2, is needed to ensure that any limit point of the sequence of iterates produced by the long-step algorithm is a primal-dual optimal solution.

3 Analysis of the long-step method

In this section, we state further assumptions on the map Φ and the set $\mathcal{C}_0$ introduced in Subsection 2.4 and develop the convergence analysis of the long-step algorithm given in Subsection 2.4.

Before stating the additional assumptions on the set $\mathcal{C}_0$ which appears in the Long-Step Algorithm, we need to introduce some definitions. Given a bilinear map $B : W \times W \to V$, where W and V are two normed vector spaces with norms $\|\cdot\|_w$ and $\|\cdot\|_v$, respectively, its norm is defined as

$$\|B\| = \|B\|_{w,v} \equiv \max\{\, \|B[u,u']\|_v : \|u\|_w \le 1,\ \|u'\|_w \le 1\,\}. \tag{22}$$

It can be shown that $\|B\| = \max\{\|B[u,u]\|_v : \|u\|_w \le 1\}$ (see Proposition 9.1.1 in Appendix 1 of [18]). For simplicity, we write $B[u]^{(2)} = B[u,u]$. Of particular interest to us is the bilinear map $B = \Phi''(X,S)$ for which $W = \mathcal{S}^n \times \mathcal{S}^n$ and $V = \mathcal{S}^n$. Given an invertible operator $\mathcal{V} : \mathcal{S}^n \to \mathcal{S}^n$, a norm for $\Phi''(X,S)$ can be defined as

$$\begin{aligned}\|\Phi''(X,S)\|_{\mathcal{V}} &\equiv \max\left\{\|\Phi''(X,S)[A,B]^{(2)}\|_F : \|\mathcal{V}(A)\|_F^2 + \|\mathcal{V}^{-1}(B)\|_F^2 \le 1\right\}\\ &= \max\left\{\|\Phi''(X,S)[\mathcal{V}^{-1}(\widetilde{A}),\mathcal{V}(\widetilde{B})]^{(2)}\|_F : \|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2 \le 1\right\}.\end{aligned} \tag{23}$$

Observe that this norm is the one obtained from (22) by letting $\|\cdot\|_v$ be the Frobenius norm on $V = \mathcal{S}^n$ and $\|\cdot\|_w$ be the norm on $W = \mathcal{S}^n \times \mathcal{S}^n$ defined by $\|(A,B)\|_w \equiv \left(\|\mathcal{V}(A)\|_F^2 + \|\mathcal{V}^{-1}(B)\|_F^2\right)^{1/2}$ for all $(A,B) \in \mathcal{S}^n \times \mathcal{S}^n$. Clearly, for every $(A,B) \in \mathcal{S}^n \times \mathcal{S}^n$, we have

$$\|\Phi''(X,S)[A,B]^{(2)}\|_F \le \|\Phi''(X,S)\|_{\mathcal{V}}\left(\|\mathcal{V}(A)\|_F^2 + \|\mathcal{V}^{-1}(B)\|_F^2\right). \tag{24}$$

A primal-dual ellipsoid centered at $(X,S) \in \mathcal{C}$ with radius $\theta > 0$ can be defined as

$$\mathcal{E}_\theta(X,S) \equiv \left\{(\widehat{X},\widehat{S}) \in \mathcal{S}^n \times \mathcal{S}^n : \|\mathcal{G}(\widehat{X}-X)\|_F^2 + \|\mathcal{G}^{-1}(\widehat{S}-S)\|_F^2 < \theta^2\,\|\mathcal{R}(\Phi)\|_F^2\right\}, \tag{25}$$

where $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$, and $\Phi \equiv \Phi(X,S)$. Also, for $(X,S) \in \mathcal{C}$, let

$$\Pi(X,S) \equiv \frac{\mu\,\|\mathcal{R}(I)\|_F}{\|\mathcal{R}(\Phi)\|_F}, \tag{26}$$

$$\Theta(X,S) \equiv \sup\{\theta > 0 : \mathcal{E}_\theta(X,S) \subseteq \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}\}, \tag{27}$$

$$\Omega_\theta(X,S) \equiv \sup\{\|\Phi''(\widehat{X},\widehat{S})\|_{\mathcal{G}} : (\widehat{X},\widehat{S}) \in \mathcal{E}_\theta(X,S)\}, \tag{28}$$

for all $0 < \theta \le \Theta(X,S)$, where $\mu \equiv \mu(X,S)$. Finally, given $\gamma \in (0,1)$, define

$$\Pi(\gamma) \equiv \sup\{\,\Pi(X,S) : (X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)\,\}, \tag{29}$$

$$\Theta(\gamma) \equiv \inf\{\,\Theta(X,S) : (X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)\,\} \tag{30}$$

$$\Omega_\theta(\gamma) \equiv \sup\left\{\frac{1}{\mu}\Omega_\theta(X,S)\|\mathcal{R}(\Phi)\|_F^2 : (X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)\right\}, \tag{31}$$

for all $0 < \theta \le \Theta(\gamma)$. Note that, by Assumption 2, $\Phi(X,S) \ne 0$ for every $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \supseteq \mathcal{C}$. As a result, $\Pi(X,S)$ is well-defined for all $(X,S) \in \mathcal{C}$.

With these definitions, we make the following assumptions:

Assumption 7. *For every $\gamma \in (0,1)$, $\Pi(\gamma)$ is finite.*

Assumption 8. *For every $\gamma \in (0,1)$, $\Theta(\gamma)$ is positive.*

Assumption 9. *For every $\gamma \in (0,1)$, there exists $\theta \in (0,\Theta(\gamma)]$ such that $\Omega_\theta(\gamma)$ is finite.*

In what follows we establish an iteration-complexity bound for Long-Step Algorithm which depends on the quantities (29), (30) and (31) (see Theorem 3.3 below).

Lemma 3.1. *Let $\sigma \in (0,1)$ and $(X,S,y) \in \mathcal{C} \times \Re^m$ be given. Suppose $\alpha \in [0,1]$ satisfies*

$$\alpha < \frac{\Theta}{1+\sigma\Pi}, \tag{32}$$

where $\Pi \equiv \Pi(X,S)$ *and* $\Theta \equiv \Theta(X,S)$. *Then,* $(X_\alpha, S_\alpha) \in \mathcal{E}_\Theta(X,S) \subset \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, *the quantity*

$$T_\alpha = T_\alpha(X,S) \equiv \int_0^1 (1-t)\Phi''(X_{t\alpha}, S_{t\alpha})[\Delta X, \Delta S]^{(2)}\, dt \tag{33}$$

is well-defined, and

$$\mu(X_\alpha, S_\alpha) \le (1-\alpha+\alpha\sigma)\mu + \alpha^2 \|T_\alpha\|_F. \tag{34}$$

where $\mu \equiv \mu(X,S)$. *If, in addition,* $(X,S,y) \in (\mathcal{C}_0 \times \Re^m) \cap \mathcal{F}^0_{-\infty}(\gamma)$ *for some* $\gamma \in (0,1)$ *and* α *satisfies*

$$\alpha \|T_\alpha\|_F \le \frac{\gamma\sigma\mu}{1+\gamma}, \tag{35}$$

then $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0_{-\infty}(\gamma)$.

Proof. By (26), (20) and the triangle inequality for norms, we have

$$\|\mathcal{R}(H_\sigma)\|_F = \|\mathcal{R}(\sigma\mu I - \Phi)\|_F \le \sigma\mu \|\mathcal{R}(I)\|_F + \|\mathcal{R}(\Phi)\|_F \le (\sigma\Pi + 1)\,\|\mathcal{R}(\Phi)\|_F, \tag{36}$$

where $H_\sigma \equiv H_\sigma(X,S)$. Using (21), (12), Lemma 2.1 with $(U,V) = (\Delta X, \Delta S)$, (32) and (36), we obtain

$$\begin{aligned} \|\mathcal{G}(X_\alpha - X)\|_F^2 + \|\mathcal{G}^{-1}(S_\alpha - S)\|_F^2 &= \alpha^2 \left(\|\mathcal{G}(\Delta X)\|_F^2 + \|\mathcal{G}^{-1}(\Delta S)\|_F^2 \right) \\ &= \alpha^2 \,\|\mathcal{R}(H_\sigma)\|_F^2 < \frac{\Theta^2 \|\mathcal{R}(H_\sigma)\|_F^2}{(1+\sigma\Pi)^2} \le \Theta^2\, \|\mathcal{R}(\Phi)\|_F^2. \end{aligned}$$

Hence, $(X_\alpha, S_\alpha) \in \mathcal{E}_\Theta(X,S) \subseteq \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ where the inclusion follows immediately from (25) and (27). This together with Assumption 1 clearly implies that T_α is well-defined. Moreover, using the Taylor integral formula and (12a), we obtain

$$\begin{aligned} \Phi(X_\alpha, S_\alpha) &= \Phi + \alpha\, \Phi'(X,S)[\Delta X, \Delta S] + \alpha^2 T_\alpha = \Phi + \alpha\left(\Phi_x(\Delta X) + \Phi_s(\Delta S)\right) + \alpha^2 T_\alpha \\ &= \Phi + \alpha\left(\sigma\mu I - \Phi\right) + \alpha^2 T_\alpha = (1-\alpha)\Phi + \alpha\sigma\mu I + \alpha^2 T_\alpha. \end{aligned}$$

Using this expression and the fact that $\mu(X_\alpha, S_\alpha) \equiv \Phi(X_\alpha, S_\alpha) \bullet I/n$, we obtain

$$\begin{aligned} \mu(X_\alpha, S_\alpha) &= (1-\alpha)\left(\frac{\Phi \bullet I}{n}\right) + \alpha\sigma\mu\left(\frac{I \bullet I}{n}\right) + \alpha^2 \left(\frac{T_\alpha \bullet I}{n}\right) \\ &\le (1-\alpha)\mu + \alpha\sigma\mu + \alpha^2 \|T_\alpha\|_F, \end{aligned} \tag{37}$$

and hence (34) follows. Now suppose in addition that $(X,S,y) \in (\mathcal{C}_0 \times \Re^m) \cap \mathcal{F}^0_{-\infty}(\gamma)$ and α satisfies (35). Using (12b), (12c), the fact that $(X,S,y) \in \mathcal{F}^0$ and the first part of the lemma, we easily see that $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0$. Using (16), the two last relations, some standard norm properties, Assumption 6, and the assumption that $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, we deduce that

$$\begin{aligned} d_{-\infty}(X_\alpha, S_\alpha) &\le \|\Phi(X_\alpha, S_\alpha) - \mu(X_\alpha, S_\alpha) I\|_{-\infty} = \left\| (1-\alpha)\left(\Phi - \mu I\right) + \alpha^2 \left(T_\alpha - \frac{T_\alpha \bullet I}{n} I \right) \right\|_{-\infty} \\ &\le (1-\alpha)\,\|\Phi - \mu I\|_{-\infty} + \alpha^2 \left\| T_\alpha - \frac{T_\alpha \bullet I}{n} I \right\|_{-\infty} \le (1-\alpha)\,\|\Phi - \mu I\|_{-\infty} + \alpha^2\,\|T_\alpha\|_F \\ &= (1-\alpha)\, d_{-\infty}(X,S) + \alpha^2\,\|T_\alpha\|_F \le (1-\alpha)\gamma\mu + \alpha^2 \|T_\alpha\|_F. \end{aligned}$$

Using the expression for $\mu(X_\alpha, S_\alpha)$ in (37) to substitute for $(1-\alpha)\mu$ in the previous inequality, we obtain

$$\begin{aligned} \|\Phi(X_\alpha, S_\alpha) - \mu(X_\alpha, S_\alpha) I\|_{-\infty} &\le \gamma\,\mu(X_\alpha, S_\alpha) - \alpha\gamma\sigma\mu - \alpha^2\gamma\left(\frac{T_\alpha \bullet I}{n}\right) + \alpha^2 \|T_\alpha\|_F \\ &\le \gamma\,\mu(X_\alpha, S_\alpha) - \alpha\gamma\sigma\mu + \alpha^2(1+\gamma)\|T_\alpha\|_F \le \gamma\mu(X_\alpha, S_\alpha), \end{aligned}$$

where the last inequality follows from (35). We have thus shown that $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0_{-\infty}(\gamma)$. ∎

Lemma 3.2. *Let $\gamma, \sigma \in (0,1)$, $(X,S,y) \in (\mathcal{C}_0 \times \Re^m) \cap \mathcal{F}^0_{-\infty}(\gamma)$ be given. Then, for all $\theta \in (0, \Theta(\gamma)]$ and $\alpha \in [0, \bar\alpha(\theta)]$, where*

$$\bar\alpha(\theta) \equiv \min\left\{1,\, \frac{\theta}{2\,(1+\sigma\Pi(\gamma))},\, \frac{2\gamma\sigma q}{(1+\gamma)\Omega_\theta(\gamma)(1+\sigma\Pi(\gamma))^2}\right\} \tag{38}$$

and

$$q \equiv \min\left\{1,\, \frac{(1-\sigma)(1+\gamma)}{2\gamma\sigma}\right\}, \tag{39}$$

there hold:

(a) $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0_{-\infty}(\gamma)$;
(b) $\mu(X_\alpha, S_\alpha) \le (1 - \frac{\alpha}{2}(1-\sigma))\mu(X,S)$.

Proof. By (38), (30), (29) and the assumption on θ and α, we have

$$\alpha \le \bar\alpha(\theta) < \theta/(1+\sigma\Pi(\gamma)) \le \Theta(\gamma)/(1+\sigma\Pi(\gamma)) \le \Theta/(1+\sigma\Pi),$$

where $\Pi \equiv \Pi(X,S)$ and $\Theta \equiv \Theta(X,S)$. Hence, by Lemma 3.1, it suffices to establish (35) in order to show that $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0_{-\infty}(\gamma)$. Letting $\mu \equiv \mu(X,S)$, $\Phi \equiv \Phi(X,S)$, $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, and $H_\sigma \equiv H_\sigma(X,S)$, it follows from (33), (24), Lemma 2.1 with $(U,V) = (\Delta X, \Delta S)$, (36), (31), (29), (38), and (39) that

$$\begin{aligned}
\alpha\|T_\alpha\|_F &= \alpha \left\| \int_0^1 (1-t)\Phi''(X_{t\alpha}, S_{t\alpha})[\Delta X, \Delta S]^{(2)} dt \right\|_F \\
&\le \alpha \int_0^1 (1-t)\, \left\|\Phi''(X_{t\alpha}, S_{t\alpha})[\Delta X, \Delta S]^{(2)}\right\|_F dt \\
&\le \frac{\alpha}{2} \sup_{t\in[0,1]} \left\{ \left\|\Phi''(X_{t\alpha}, S_{t\alpha})[\Delta X, \Delta S]^{(2)}\right\|_F \right\} \\
&\le \frac{\alpha}{2} \sup_{t\in[0,1]} \left\{ \|\Phi''(X_{t\alpha}, S_{t\alpha})\|_{\mathcal{G}} \right\} \left(\|\mathcal{G}(\Delta X)\|_F^2 + \|\mathcal{G}^{-1}(\Delta S)\|_F^2 \right) \\
&= \frac{\alpha}{2}\Omega_\theta(X,S)\,\|\mathcal{R}(H_\sigma)\|_F^2 \le \frac{\alpha}{2}\Omega_\theta(X,S)\,\|\mathcal{R}(\Phi)\|_F^2\,(1+\sigma\Pi)^2 \\
&\le \frac{\bar\alpha(\theta)}{2}\,\mu\,\Omega_\theta(\gamma)\,(1+\sigma\Pi(\gamma))^2 \le \frac{\gamma\sigma q\mu}{1+\gamma} \le \frac{\gamma\sigma\mu}{1+\gamma}.
\end{aligned}$$

Hence, $(X_\alpha, S_\alpha, y_\alpha) \in \mathcal{F}^0_{-\infty}(\gamma)$, and (a) follows. To prove (b), note that the previous inequality and relations (34) and (39) imply that

$$\begin{aligned}
\mu(X_\alpha, S_\alpha) &\le (1-\alpha+\alpha\sigma)\mu + \alpha^2\|T_\alpha\|_F \le (1-\alpha+\alpha\sigma)\,\mu + \frac{\gamma\sigma q\alpha}{1+\gamma}\,\mu \\
&\le (1-\alpha+\alpha\sigma)\,\mu + \frac{(1-\sigma)\alpha}{2}\,\mu = \left(1 - \frac{1-\sigma}{2}\,\alpha\right)\mu.
\end{aligned}$$

∎

Theorem 3.3. (Iteration Complexity of the Long-Step Algorithm) *Let constants $\gamma, \sigma \in (0,1)$ be given and define*

$$\Omega(\gamma) \equiv \inf_{0<\theta\le\Theta(\gamma)} \max\{\Omega_\theta(\gamma)\,,\, 4\,\theta^{-1}\}. \tag{40}$$

Then the sequence of iterates $\{(\widetilde X^k, \widetilde S^k, \widetilde y^k)\}_{k\ge 0} \subset \widetilde{\mathcal{F}}^0_{-\infty}(\gamma)$ generated by the long-step algorithm satisfies $\mu(\widetilde X^k, \widetilde S^k) \le [1 - \eta(1-\sigma)]^k \mu(\widetilde X^0, \widetilde S^0)$ for all $k \ge 0$, where

$$\eta \equiv \frac{1}{2}\min\left\{1,\, \frac{\gamma\sigma q}{(1+\gamma)\Omega(\gamma)(1+\sigma\Pi(\gamma))^2}\right\}. \tag{41}$$

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

Proof. First note that $\Omega(\gamma)$ is finite due to Assumption 9. By (40), there exists $\bar\theta \in (0, \Theta(\gamma)]$ such that

$$\max\{\Omega_{\bar\theta}(\gamma)\,,\, 4\,\bar\theta^{-1}\} \le 2\Omega(\gamma). \tag{42}$$

It follows from the definition of α_k in step 4 of the Long-Step Algorithm, Lemma 3.2(b) and the scale invariance of the duality gap that

$$\mu(\widetilde X^{k+1}, \widetilde S^{k+1}) \le \left(1 - \frac{1-\sigma}{2}\,\bar\alpha(\bar\theta)\right)\mu(\widetilde X^k, \widetilde S^k). \tag{43}$$

Now, by (38), (42) and (41), we have

$$\begin{aligned}
\bar\alpha(\bar\theta) &= \min\left\{1,\ \frac{\bar\theta}{2\,(1+\sigma\Pi(\gamma))}, \frac{2\gamma\sigma q}{(1+\gamma)\Omega_{\bar\theta}(\gamma)(1+\sigma\Pi(\gamma))^2}\right\}\\
&\ge \min\left\{1,\ \frac{1}{\Omega(\gamma)(1+\sigma\Pi(\gamma))}, \frac{\gamma\sigma q}{(1+\gamma)\Omega(\gamma)(1+\sigma\Pi(\gamma))^2}\right\},\\
&= \min\left\{1,\ \frac{\gamma\sigma q}{(1+\gamma)\Omega(\gamma)(1+\sigma\Pi(\gamma))^2}\right\} = 2\eta,
\end{aligned}$$

which together with (43) yields the first statement of the theorem. The second statement is a straightforward consequence of the first one. ∎

4 Examples

In this section we give four examples of pairs $(\Phi, \mathcal{C}_0)$ for which it is possible to establish polynomial iteration-complexity bounds for the long-step algorithm based on $(\Phi, \mathcal{C}_0)$. The maps for the first and third pairs are the $XS + SX$ map and the $X^{1/2}SX^{1/2}$ map. Long-step algorithms based on these two maps have been extensively studied in Monteiro and Zhang [17] and Monteiro and Tsuchiya [13], respectively. We study them again here using our general framework of Sections 2 and 3 to illustrate the usefulness and generality of our approach. Long-step algorithms based on the other two maps, namely the $L_x^T S L_x$ map and the $W^{1/2}XSW^{-1/2}$ map, where $L_x \equiv \mathrm{chol}\,(X)$ and W is the unique symmetric matrix such that $S = WXW$, are studied here for the first time. The two last maps have been introduced in [15], [16] and [23] and were studied there from different points of view.

We now state a technical result that will be used in the upcoming subsections.

Lemma 4.1. *Let $(X,S) \in \mathcal{C}$, $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, and $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$. Then, for all $A \in \mathcal{S}^n$,*

$$\|\mathcal{G}(A)\|_F^2 \ge \frac{\|(\Phi_s^*)^{-1}(A)\|_F^2}{\|\mathcal{R}^2\|}, \qquad \|\mathcal{G}^{-1}(A)\|_F^2 \ge \frac{\|(\Phi_x^*)^{-1}(A)\|_F^2}{\|\mathcal{R}^2\|}, \qquad \|\mathcal{G}^{-1}(A)\|_F^2 \ge \frac{\|\Phi_s(A)\|_F^2}{\|\mathcal{R}^{-2}\|}.$$

Proof. Let $A \in \mathcal{S}^n$ and define $\widetilde A \equiv (\Phi_s^*)^{-1}(A)$. The first inequality of the lemma is proved using (14) as follows:

$$\begin{aligned}
\|\mathcal{G}(A)\|_F^2 &= \mathcal{G}^2(A) \bullet A = (\Phi_s^{-1}\Phi_x)(A) \bullet A = \Phi_x(A) \bullet (\Phi_s^*)^{-1}(A) = (\Phi_x\Phi_s^*)(\widetilde A) \bullet \widetilde A\\
&= \mathcal{R}^{-2}(\widetilde A) \bullet \widetilde A \ge \|\widetilde A\|_F^2/\|\mathcal{R}^2\| = \|(\Phi_s^*)^{-1}(A)\|_F^2/\|\mathcal{R}^2\|.
\end{aligned}$$

Similar arguments prove the second and third inequalities. ∎

In addition, we point out the following standard fact that will be used implicitly in several places within this section. Its proof follows directly from the fact that two diagonalizable matrices X and S commute if and only if they share a common basis of eigenvectors.

Proposition 4.2. *Let $X, S \in \mathcal{S}_+^n$, and suppose $XS \in \mathcal{S}^n$ or, equivalently, $XS = SX$. Then $X^{1/2}S^{1/2} = S^{1/2}X^{1/2}$.*

4.1 The $XS+SX$ Map

In this subsection, we consider the map $\Phi(X,S) = (XS+SX)/2$ and show that Φ satisfies Assumptions 1 through 4 and that $\mathcal{C}_0$ can be chosen so that Assumptions 5 through 9 are satisfied. As a result, we obtain an iteration-complexity for the long-step algorithm based on this map which duplicates the complexity found by Monteiro and Zhang in [17].

It is well-known that the map Φ satisfies Assumptions 1 through 4. Moreover, for any $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$,

$$\Phi_x(A) = (SA+AS)/2, \qquad \Phi_s(B) = (XB+BX)/2, \tag{44}$$

$$\Phi''(X,S)[A,B]^{(2)} = AB+BA \tag{45}$$

for all $A,B \in \mathcal{S}^n$.

We now state our choice for the set $\mathcal{C}_0$. Let $\kappa \geq 1$ be any constant, and define

$$\mathcal{C}_0 = \mathcal{C}_0(\kappa) \equiv \{(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} : XS \in \mathcal{S}^n \text{ and } \operatorname{cond}[X^{-1}S] \leq \kappa\}. \tag{46}$$

Note that the condition that $XS \in \mathcal{S}^n$ is equivalent to saying that X and S commute. The proof of the following result is easy and can be found in section 4.3 of [12].

Lemma 4.3. *For any $U \in \mathcal{S}^n_{++}$, let $\Xi_U : \mathcal{S}^n \to \mathcal{S}^n$ be the linear operator given by $\Xi_U(A) \equiv (UA+AU)/2$. Then:*

(a) $\Xi_U \in \mathcal{T}^n_{++}$ for all $U \in \mathcal{S}^n_{++}$.

(b) If $U,V \in \mathcal{S}^n_{++}$ are such that $UV = VU$, then $\Xi_U \Xi_V = \Xi_V \Xi_U$.

Proposition 4.4. *For every $(X,S) \in \mathcal{C}_0$, the operators Φ_x and Φ_s are in $\mathcal{T}^n_{++}$ and commute. Moreover, the set $\mathcal{C}_0$ given by (46) satisfies Assumption 5.*

Proof. Since by (44) $\Phi_x = \Xi_S$ and $\Phi_s = \Xi_X$ for every $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, Lemma 4.3 clearly implies the first statement of the proposition and hence that $\mathcal{C}_0$ is a subset of $\mathcal{C}$. Let $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ be given and define $\widetilde{P} \equiv (\widetilde{S}^{1/2}(\widetilde{S}^{1/2}\widetilde{X}\widetilde{S}^{1/2})^{-1/2}\widetilde{S}^{1/2})^{1/2}$. One can easily see that (X,S) given by (9) satisfies $X = S$. Hence $(X,S) \in \mathcal{C}_0$ and $\widetilde{P} \in \mathcal{P}_0(\widetilde{X},\widetilde{S}) \neq \emptyset$. ∎

The following result shows that the pair $(\Phi, \mathcal{C}_0)$ satisfies Assumption 6.

Proposition 4.5. *For every $(X,S) \in \mathcal{C}_0$, we have $d_{-\infty}(X,S) = \|\Phi - \mu I\|_{-\infty}$, where $\Phi \equiv \Phi(X,S)$ and $\mu \equiv \mu(X,S)$.*

Proof. Since $(X,S) \in \mathcal{C}_0$, we have $\Phi = XS$. Let $(\widetilde{X},\widetilde{S}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ be such that $(X,S) \sim (\widetilde{X},\widetilde{S})$, and let $\widetilde{\mu} \equiv \mu(\widetilde{X},\widetilde{S})$. Observing that $\Phi = XS = \widetilde{P}\widetilde{X}\widetilde{S}\widetilde{P}^{-1}$ for some nonsingular $\widetilde{P}$, we have

$$\lambda_{\min}[\Phi] = \lambda_{\min}[XS] = \lambda_{\min}[\widetilde{X}\widetilde{S}] \geq \frac{1}{2}\lambda_{\min}[\widetilde{X}\widetilde{S} + \widetilde{S}\widetilde{X}] = \lambda_{\min}[\Phi(\widetilde{X},\widetilde{S})],$$

where the inequality follows from the fact that the real part of the spectrum of a real matrix is contained between the largest and the smallest eigenvalues of its Hermitian part (see p. 187 of [5], for example). It follows from this relation, the definition of $\|\cdot\|_{-\infty}$, and the fact that $\mu = \widetilde{\mu}$, due to Assumption 3, that

$$\|\Phi - \mu I\|_{-\infty} = \max\{0, \mu - \lambda_{\min}[\Phi]\} \leq \max\{0, \mu - \lambda_{\min}[\Phi(\widetilde{X},\widetilde{S})]\} = \|\Phi(\widetilde{X},\widetilde{S}) - \widetilde{\mu} I\|_{-\infty}.$$

Hence, by (16), we have $d_{-\infty}(X,S) = \|\Phi - \mu I\|_{-\infty}$. ∎

Lemma 4.6. *Let $(X,S) \in \mathcal{C}$, $\Phi \equiv \Phi(X,S)$ and $\mu \equiv \mu(X,S)$. Then*

$$\|\mathcal{R}(\Phi)\|_F^2 = X \bullet S = n\mu, \qquad \|\mathcal{R}(I)\|_F^2 = X^{-1} \bullet S^{-1}. \tag{47}$$

Proof. The proof is an immediate verification. ∎

The following result shows that the pair $(\Phi, \mathcal{C}_0)$ satisfies Assumption 7 and also establishes a valuable inequality that we will use later in this subsection.

Proposition 4.7. *Let $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, and define $\mu \equiv \mu(X,S)$. Then*

$$\|\Phi(X,S)^{-1}\| \leq \frac{1}{(1-\gamma)\mu}, \qquad \Pi(X,S)^2 \leq \frac{1}{1-\gamma}.$$

In particular, $\Pi(\gamma)^2 \leq 1/(1-\gamma)$.

Proof. Define $\Phi \equiv \Phi(X,S)$ and $\Pi \equiv \Pi(X,S)$. Using Assumption 6 (which has been verified in Proposition 4.4), the definition of the seminorm $\|\cdot\|_{-\infty}$, and the assumption that $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, it is easy to see that $\lambda_{\min}[\Phi] \geq (1-\gamma)\mu$. The first inequality of the lemma is proved by combining this inequality with the fact that $\lambda_{\min}[\Phi] = 1/\|\Phi^{-1}\|$.

We can rewrite the first inequality of the lemma as $\lambda_{\max}[X^{-1}S^{-1}] \leq 1/((1-\gamma)\mu)$, which implies that $X^{-1} \bullet S^{-1} \leq n/((1-\gamma)\mu)$ since $X^{-1} \bullet S^{-1}$ is the sum of the eigenvalues of $X^{-1}S^{-1}$. This together with Lemma 4.6 implies that

$$\Pi^2 = \mu^2 \left(\frac{\|\mathcal{R}(I)\|_F^2}{\|\mathcal{R}(\Phi)\|_F^2} \right) = \mu^2 \left(\frac{X^{-1} \bullet S^{-1}}{n\mu} \right) \leq \mu^2 \left(\frac{1}{(1-\gamma)\mu^2} \right) = \frac{1}{1-\gamma}.$$

The last statement of the proposition follows from (29). ∎

The following result shows that the pair $(\Phi, \mathcal{C}_0)$ satisfies Assumption 8.

Proposition 4.8. *Let $\gamma \in (0,1)$, and define*

$$\theta^* = \theta^*(\gamma) \equiv \frac{1}{2} \left(\frac{1-\gamma}{n} \right)^{1/2}. \tag{48}$$

Then $\mathcal{E}_{\theta^*}(X,S) \subseteq \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ for all $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$. In particular, $\Theta(\gamma) \geq \theta^*$.

Proof. It suffices to show that any $(\widehat{X}, \widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ satisfies

$$\|X^{-1/2}(\widehat{X} - X)X^{-1/2}\| < 1 \quad \text{and} \quad \|S^{-1/2}(\widehat{S} - S)S^{-1/2}\| < 1. \tag{49}$$

Let $\Phi \equiv \Phi(X,S)$, $\mu \equiv \mu(X,S)$, $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, and $U \equiv \Phi_s^{-1}(\widehat{X} - X)$. Hence, by (44) we have $(XU + UX)/2 = \widehat{X} - X$, which together with the triangle inequality for norms implies that

$$\|X^{-1/2}UX^{1/2}\|_F \geq \|X^{-1/2}(\widehat{X} - X)X^{-1/2}\|_F. \tag{50}$$

Using (14), (44), (48), (50), Proposition 4.4, Lemma 4.6, Proposition 4.7, the assumption that $(\widehat{X}, \widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ and $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, we obtain

$$\begin{aligned}
\frac{1}{4}(1-\gamma)\mu &= (\theta^*)^2 n\mu = (\theta^*)^2 \|\mathcal{R}(\Phi)\|_F^2 \geq \|\mathcal{G}(\widehat{X} - X)\|_F^2 + \|\mathcal{G}^{-1}(\widehat{S} - S)\|_F^2 \\
&\geq \|\mathcal{G}(\widehat{X} - X)\|_F^2 = \Phi_x(\widehat{X} - X) \bullet (\Phi_s^*)^{-1}(\widehat{X} - X) \\
&= \Phi_x(\widehat{X} - X) \bullet \Phi_s^{-1}(\widehat{X} - X) = \frac{1}{2}\Big(S(\widehat{X} - X) + (\widehat{X} - X)S \Big) \bullet U \\
&= \mathrm{Tr}\,[S(\widehat{X} - X)U] = \frac{1}{2}\mathrm{Tr}\,[S(XU + UX)U] \\
&= \frac{1}{2}\mathrm{Tr}\,[SXU^2 + SUXU] = \frac{1}{2}\Big(\|S^{1/2}X^{1/2}U\|_F^2 + \|S^{1/2}UX^{1/2}\|_F^2 \Big) \\
&\geq \frac{1}{2}\|S^{1/2}UX^{1/2}\|_F^2 = \frac{1}{2}\|S^{1/2}X^{1/2}X^{-1/2}UX^{1/2}\|_F^2 \\
&\geq \frac{\|X^{-1/2}UX^{1/2}\|_F^2}{2\|X^{-1}S^{-1}\|} \geq \frac{1}{2}(1-\gamma)\mu\|X^{-1/2}(\widehat{X} - X)X^{-1/2}\|_F^2,
\end{aligned}$$

which clearly yields the first inequality in (49). The second inequality in (49) can be established in a similar way. ∎

We refer the reader to lemma 2.1 of [13] for a proof of the following technical result.

Lemma 4.9. *For every $A \in \mathcal{S}^n_{++}$ and $H \in \mathcal{S}^n$, the equation $AU + UA = H$ has a unique solution $U \in \mathcal{S}^n$. Moreover, this solution satisfies $\|AU\|_F \le \|H\|_F/\sqrt{2}$.*

The following proposition shows that the pair $(\Phi, \mathcal{C}_0)$ satisfies Assumption 9.

Proposition 4.10. *For any $\gamma \in (0,1)$ and $\theta \in (0, \Theta(\gamma)]$, the following statements hold:*

(a) $\Omega_\theta(X,S) \le \sqrt{2\kappa}$ *for any* $(X,S) \in \mathcal{C}_0 = \mathcal{C}_0(\kappa)$;

(b) $\Omega_\theta(\gamma) \le n\sqrt{2\kappa}$.

Proof. To prove (a), fix $(X,S) \in \mathcal{C}_0$ and $\theta \in (0, \Theta(\gamma)]$, and define $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$. Also, let $(\widehat{X}, \widehat{S}) \in \mathcal{E}_\theta(X,S)$ and matrices $A, B \in \mathcal{S}^n$ such that $\|A\|_F^2 + \|B\|_F^2 \le 1$ be arbitrarily given and define $U \equiv \Phi_x^{-1}(A) = (\Phi_x^*)^{-1}(A)$. Then, by (44), $U \in \mathcal{S}^n$ satisfies $SU + US = 2A$. Using (14), (44) and Lemma 4.9, we obtain

$$\begin{aligned}\|\mathcal{G}^{-1}(A)\|_F^2 = A \bullet \mathcal{G}^{-2}(A) = (\Phi_x^*)^{-1}(A) \bullet \Phi_s(A) = \frac{1}{2} U \bullet (XA + AX) = \mathrm{Tr}\,[UXA] \\ = \mathrm{Tr}\,[USS^{-1}XA] \le \|S^{-1}X\| \|A\|_F \|US\|_F \le \sqrt{2}\,\|S^{-1}X\|\,\|A\|_F^2.\end{aligned}$$

Similarly, we can show that $\|\mathcal{G}(B)\|_F^2 \le \sqrt{2}\,\|X^{-1}S\|\,\|B\|_F^2$. Using these two inequalities together with (45) and the inequality $\|A\|_F^2 + \|B\|_F^2 \le 1$, we obtain

$$\begin{aligned}\|\Phi''(\widehat{X}, \widehat{S})[\mathcal{G}^{-1}(A), \mathcal{G}(B)]^{(2)}\|_F &= \|\mathcal{G}^{-1}(A)\mathcal{G}(B) + \mathcal{G}(B)\mathcal{G}^{-1}(A)\|_F \le 2\|\mathcal{G}^{-1}(A)\|_F \|\mathcal{G}(B)\|_F \\ &\le 2\sqrt{2}\, \left(\mathrm{cond}\,[X^{-1}S]\right)^{1/2} \|A\|_F \|B\|_F \le \sqrt{2}\, \left(\mathrm{cond}\,[X^{-1}S]\right)^{1/2} \left(\|A\|_F^2 + \|B\|_F^2\right) \\ &\le \sqrt{2}\left(\mathrm{cond}\,[X^{-1}S]\right)^{1/2} \le \sqrt{2\kappa},\end{aligned}$$

where the last inequality follows from (46) and the fact that $(X,S) \in \mathcal{C}_0$. Hence, we conclude from (23) that $\|\Phi''(\widehat{X}, \widehat{S})\|_{\mathcal{G}} \le \sqrt{2\kappa}$. Since this upper bound does not depend on the choice of $(\widehat{X}, \widehat{S}) \in \mathcal{E}_\theta(X,S)$, (28) implies that $\Omega_\theta(X,S) \le \sqrt{2\kappa}$.

Statement (b) follows immediately from (a), Lemma 4.6 and (31). ∎

The main iteration-complexity result for the long-step algorithm based on the AHO map (3) and the set (46) can now be derived as a consequence of Theorem 3.3 and the results of this subsection.

Theorem 4.11. *The long-step algorithm based on the map $\Phi(X,S) = (XS+SX)/2$ and the set $\mathcal{C}_0 = \mathcal{C}_0(\kappa)$ of (46) terminates in at most $\mathcal{O}(n\sqrt{\kappa}L)$ iterations.*

Proof. We have already seen that Φ and $\mathcal{C}_0$ satisfy Assumptions 1 through 9. Let θ^* be as in Proposition 4.8. By (40) and Propositions 4.8 and 4.10, we have

$$\Omega(\gamma) \le \max\{\Omega_{\theta^*}(\gamma), 4(\theta^*)^{-1}\} \le \max\left\{ n\sqrt{2\kappa}, 8\left(\frac{2n}{1-\gamma}\right)^{1/2} \right\}.$$

The conclusion of the theorem now follows from Theorem 3.3. ∎

4.2 The $L_x^T S L_x$ Map

In this subsection, we consider the map $\Phi(X,S) = L_x^T S L_x$, where $L_x \equiv \mathrm{chol}\,(X)$, and show that Φ satisfies Assumptions 1 through 4 and that $\mathcal{C}_0$ can be chosen so that Assumptions 5 through 9 are satisfied. As a consequence, we derive for the first time a polynomial iteration-complexity for the long-step algorithm based on this map.

The map Φ clearly satisfies Assumptions 1 through 4. Moreover, for every $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ and $A, B \in \mathcal{S}^n$,

$$\Phi_x(A) = (U')^T S L_x + L_x^T S U', \qquad \Phi_s(B) = L_x^T B L_x, \tag{51}$$

$$\Phi''(X,S)[A,B]^{(2)} = (U'')^T S L_x + L_x^T S U'' + 2(U')^T B L_x + 2L_x^T B U' + 2(U')^T S U', \tag{52}$$

where $U', U'' \in \mathcal{L}^n$ are the unique solutions of the equations

$$U' L_x^T + L_x (U')^T = A, \tag{53}$$

$$U'' L_x^T + L_x (U'')^T = -2U'(U')^T, \tag{54}$$

respectively.

A suitable set $\mathcal{C}_0$ that fits well with the map (4) is given by

$$\mathcal{C}_0 \equiv \{(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} : XS \in \mathcal{L}^n\}. \tag{55}$$

It is interesting to note that if $(X,S) \in \mathcal{C}_0$ then $L_x^T S L_x$ is a diagonal matrix.

Lemma 4.12. *Let $(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, and let $L_x \equiv chol(X)$. Then, Φ_s is invertible, $\Phi_s^{-1}(B) = L_x^{-T} B L_x^{-1}$, $\Phi_s^*(B) = L_x B L_x^T$ and $(\Phi_s^*)^{-1}(B) = L_x^{-1} B L_x^{-T}$ for all $B \in \mathcal{S}^n$. If, in addition, $(X,S) \in \mathcal{C}_0$, then Φ_x is invertible, $\Phi_x \Phi_s^* \in \mathcal{T}^n_{++}$, and $\Phi_x^{-1}(A) = V L_x^T + L_x V^T$ for all $A \in \mathcal{S}^n$, where $V \in \mathcal{L}^n$ is the unique solution of the equation $V^T S L_x + L_x^T S V = A$.*

Proof. The statements regarding Φ_s are immediate. Let $L \equiv L_x^T S$. In view of (51) and (53), to prove that Φ_x is invertible, it is enough to show that the equation $V^T L^T + LV = A$ has a unique solution $V \in \mathcal{L}^n$. Indeed, using the assumption that $(X,S) \in \mathcal{C}_0$, we see that $L = L_x^{-1} X S \in \mathcal{L}^n_{++}$ and hence that $LV \in \mathcal{L}^n$. This clearly implies that the unique solution of $V^T L^T + LV = A$ is $V \equiv L^{-1} L_A$, where L_A is the unique matrix in $\mathcal{L}^n$ satisfying $(L_A)_{ij} = A_{ij}$ for all $1 \le j < i \le n$, and $(L_A)_{ii} = A_{ii}/2$ for all $1 \le i \le n$. We will now show that $\Phi_x \Phi_s^* \in \mathcal{T}^n_{++}$ by establishing the equivalent statement that $\Phi_s^{-1}\Phi_x \in \mathcal{T}^n_{++}$. Indeed, let $A, B \in \mathcal{S}^n$ be given. Also, let $U'_A, U'_B \in \mathcal{L}^n$ be the (unique) matrices satisfying

$$U'_A L_x^T + L_x (U'_A)^T = A, \qquad U'_B L_x^T + L_x (U'_B)^T = B.$$

Then, $\Phi_x(A) = (U'_A)^T S L_x + L_x^T S U'_A$ and

$$\begin{aligned}\Phi_s^{-1}\Phi_x(A) \bullet B &= \left(L_x^{-T} (U'_A)^T S + S U'_A L_x^{-1}\right) \bullet \left(U'_B L_x^T + L_x (U'_B)^T\right) \\ &= 2\left(S^{1/2} U'_A\right) \bullet \left(S^{1/2} U'_B\right) + 2\,\mathrm{Tr}\,\left[U'_B (L_x^T S) U'_A L_x^{-1}\right].\end{aligned}$$

Similarly, we can see that

$$A \bullet \Phi_s^{-1}\Phi_x(B) = 2\left(S^{1/2} U'_A\right) \bullet \left(S^{1/2} U'_B\right) + 2\,\mathrm{Tr}\,\left[U'_A (L_x^T S) U'_B L_x^{-1}\right]$$

The two last expressions are identical due to the fact that the four matrices U'_A, $L_x^T S$, U'_B and L_x^{-1} are in $\mathcal{L}^n$ and the fact that the trace of a product of a finite number of matrices in $\mathcal{L}^n$ does not change by permuting the order of the matrices. Hence, $\Phi_s^{-1}\Phi_x$ is symmetric. If $0 \ne A = B$ then $0 \ne U'_A = U'_B$ and one easily sees that Tr $\left[U'_A (L_x^T S) U'_B L_x^{-1}\right] \ge 0$, and hence that $A \bullet \Phi_s^{-1}\Phi_x(A) \ge 2\|S^{1/2} U' A\|_F^2 > 0$. Hence, $\Phi_s^{-1}\Phi_x$ is positive definite. We have thus shown that $\Phi_s^{-1}\Phi_x \in \mathcal{T}^n_{++}$. ∎

Proposition 4.13. *The set $\mathcal{C}_0$ given by (55) satisfies Assumptions 5 and 6.*

Proof. Lemma 4.12 immediately implies that $\mathcal{C}_0 \subseteq \mathcal{C}$. We will now show that $\mathcal{P}_0(\widetilde{X}, \widetilde{S}) \ne \emptyset$ for all $(\widetilde{X}, \widetilde{S}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$. Indeed, let $L_{\widetilde{x}} \equiv \mathrm{chol}\,(\widetilde{X})$ and consider the orthogonal eigenvalue decomposition $L_{\widetilde{x}}^T \widetilde{S} L_{\widetilde{x}} = Q \Lambda Q^T$, where Q is an orthogonal matrix and Λ is a diagonal matrix. Since $Q = Q^{-T}$, we have

$$(L_{\widetilde{x}} Q)^{-1} (\widetilde{X}\widetilde{S}) (L_{\widetilde{x}} Q) = Q^{-1} L_{\widetilde{x}}^{-1} \widetilde{X}\widetilde{S} L_{\widetilde{x}} Q = Q^{-1} L_{\widetilde{x}}^T \widetilde{S} L_{\widetilde{x}} Q^{-T} = \Lambda \in \mathcal{L}^n.$$

Hence, $L_{\widetilde{x}} Q \in \mathcal{P}_0(\widetilde{X}, \widetilde{S})$, and so we have verified Assumption 5. The verification of Assumption 6 is straightforward. ∎

The proof of the following lemma is a simple verification and hence it is omitted.

Lemma 4.14. *Let $(X,S) \in \mathcal{C}_0$, $\Phi \equiv \Phi(X,S)$, and $\mu \equiv \mu(X,S)$. Then*

$$\|\mathcal{R}(\Phi)\|_F^2 = X \bullet S = n\mu, \qquad \|\mathcal{R}(I)\|_F^2 = X^{-1} \bullet S^{-1}.$$

The following proposition verifies Assumption 7 and also establishes two important inequalities that we will use throughout the remainder of this subsection.

Proposition 4.15. *Let $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ be given and Φ denote the map (4). Then,*

$$\|\Phi(X,S)\| \le n\mu, \qquad \|\Phi(X,S)^{-1}\| \le \frac{1}{(1-\gamma)\mu}, \qquad \Pi(X,S)^2 \le \frac{1}{1-\gamma},$$

where $\mu \equiv \mu(X,S)$*. In particular,* $\Pi(\gamma)^2 \le 1/(1-\gamma)$*.*

Proof. The first inequality of the lemma is a direct consequence of the facts that $\|\Phi(X,S)\| = \lambda_{\max}[\Phi(X,S)]$ and that $n\mu$ is the sum of the eigenvalues of $\Phi(X,S)$. The proof of the second and third inequalities and the final statement of the proposition are similar to the proof of Proposition 4.7. ∎

The proof of the following lemma can be found as lemma 7 in Monteiro and Zanjácomo [16].

Lemma 4.16. *If* $M \in \Re^{n\times n}$ *is such that Tr* $[M^2] \ge 0$*, then* $\|M\|_F \le \|M + M^T\|_F/\sqrt{2}$*. In particular, the inequality holds if* $M \in \mathcal{L}^n$*.*

Lemma 4.17. *Let* $\gamma \in (0,1)$ *and* $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ *be given and define* $\mu \equiv \mu(X,S)$ *and* $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$*. Then*

$$\|\mathcal{R}^2\| \le \frac{\sqrt{2}}{(1-\gamma)\mu}, \qquad \|\mathcal{R}^{-2}\| \le \sqrt{2}n\mu. \tag{56}$$

Proof. By standard properties of the norm $\|\cdot\|$ in $\mathcal{T}^n$ (see Subsection 1.1) and relation (14), we have

$$\|\mathcal{R}^2\| = \max_{A\in\mathcal{S}^n, \|A\|_F\le 1} \mathcal{R}^2(A)\bullet A = \max_{A\in\mathcal{S}^n, \|A\|_F\le 1} = \Phi_x^{-1}(A)\bullet\Phi_s^{-1}(A), \tag{57}$$

$$\|\mathcal{R}^{-2}\| = \max_{A\in\mathcal{S}^n, \|A\|_F\le 1} \mathcal{R}^{-2}(A)\bullet A = \max_{A\in\mathcal{S}^n, \|A\|_F\le 1} = \Phi_x(A)\bullet\Phi_s(A). \tag{58}$$

Now let $A \in \mathcal{S}^n$ such that $\|A\|_F \le 1$ be given, and let $V \in \mathcal{L}^n$ be the unique solution of the system $V^TSL_x + L_x^TSV = A$, where $L_x \equiv \text{chol}\,(X)$. The assumption that $(X,S) \in \mathcal{C}_0$ implies that $L_x^TSV = L_x^{-1}(XS)V \in \mathcal{L}^n$. Hence, by Lemma 4.16 with $M = L_x^TSV$, we conclude that $\|L_x^TSV\|_F \le \|A\|_F/\sqrt{2}$. Moreover, Lemma 4.12 implies that $\Phi_x^{-1}(A) = VL_x^T + L_xV^T$. Using these two last relations, (14) and Proposition 4.15, we obtain

$$\begin{aligned}\mathcal{R}^2(A)\bullet A &= \Phi_x^{-1}(A)\bullet\Phi_s^{-1}(A) = \text{Tr}\,[(VL_x^T + L_xV^T)L_x^{-T}AL_x^{-1}]\\ &= 2\,\text{Tr}\,[L_x^{-1}VA] \le 2\,\|L_x^{-1}V\|_F\|A\|_F \le 2\,\|L_x^{-1}S^{-1}L_x^{-T}\|\,\|L_x^TSV\|_F\|A\|_F\\ &\le \frac{\sqrt{2}}{(1-\gamma)\mu}\|A\|_F^2 \le \frac{\sqrt{2}}{(1-\gamma)\mu},\end{aligned}$$

which clearly yields the first inequality of (56). Similarly, letting U' be the unique solution of $U'L_x^T + L_x(U')^T = L_xAL_x^T$ so that, by (51) and Lemma 4.12, $(\Phi_x\Phi_s^*)(A) = (U')^TSL_x + L_x^TSU'$, and using Lemma 4.16 with $M = L_x^{-1}U' \in \mathcal{L}^n$, Proposition 4.15 and the standard properties of norms, we obtain

$$\begin{aligned}\mathcal{R}^{-2}(A)\bullet A &= (\Phi_x\Phi_s^*)(A)\bullet A \le \|(\Phi_x\Phi_s^*)(A)\|_F\|A\|_F \le 2\|L_x^TSU'\|_F\\ &\le 2\|L_x^TSL_x\|\,\|L_x^{-1}U'\|_F \le \sqrt{2}n\mu\|A\|_F \le \sqrt{2}n\mu,\end{aligned}$$

which clearly yields the second inequality of (56). ∎

The following result shows that the pair $(\Phi, \mathcal{C}_0)$ satisfies Assumption 8.

Proposition 4.18. *Let* $\gamma \in (0,1)$*, and define*

$$\theta^* = \theta^*(\gamma) \equiv \frac{1}{2\sqrt{2}}\left(\frac{1-\gamma}{n}\right)^{1/2}. \tag{59}$$

Then, for all $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ *and* $(\widehat{X},\widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$*, we have*

$$\|L_x^{-1}(\widehat{X}-X)L_x^{-T}\| \le \frac{1}{2}, \qquad \|L_s^{-1}(\widehat{S}-S)L_s^{-T}\| \le \frac{1}{2}, \tag{60}$$

where $L_x \equiv chol(X)$ *and* $L_s \equiv chol(S)$*. Hence,* $\mathcal{E}_{\theta^*}(X,S) \subseteq \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ *for all* $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$*. In particular,* $\Theta(\gamma) \ge \theta^*$*.*

Proof. Let $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ and $(\widehat{X},\widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ be given. Let $L_x \equiv \text{chol}\,(X)$, $L_s \equiv \text{chol}\,(S)$, $\Phi \equiv \Phi(X,S)$, $\mu \equiv \mu(X,S)$, $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, and $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$. Using (59), Lemmas 4.12, 4.14, 4.17, and 4.1, and the fact that $(X,S)\in\mathcal{C}_0\cap\mathcal{N}_{-\infty}(\gamma)$ and $(\widehat{X},\widehat{S})\in\mathcal{E}_{\theta^*}(X,S)$, we obtain

$$\begin{aligned}\frac{1}{8}(1-\gamma)\mu=(\theta^*)^2n\mu=(\theta^*)^2\|\mathcal{R}(\Phi)\|_F^2 &\geq \|\mathcal{G}(\widehat{X}-X)\|_F^2+\|\mathcal{G}^{-1}(\widehat{S}-S)\|_F^2\geq\|\mathcal{G}(\widehat{X}-X)\|_F^2\\ &\geq \|(\Phi_s^*)^{-1}(\widehat{X}-X)\|_F/\|\mathcal{R}^2\|\geq\frac{1}{\sqrt{2}}(1-\gamma)\mu\|L_x^{-1}(\widehat{X}-X)L_x^{-T}\|_F,\end{aligned}$$

from which the first inequality in (60) follows. Now let $V\in\mathcal{L}^n$ be the unique solution of the equation $V^TSL_x+L_x^TSV=L_x^T(\widehat{S}-S)L_x$ so that $(\Phi_x^{-1}\Phi_s)(\widehat{S}-S)=VL_x^T+L_xV^T$. Clearly, we have $\|L_s^{-1}(\widehat{S}-S)L_s^{-T}\|_F\leq 2\,\|L_s^TVL_x^{-1}L_s^{-T}\|_F$. Using these identities, (14), Proposition 4.15, and arguments similar to the ones above, we obtain

$$\begin{aligned}\frac{1}{8}(1-\gamma)\mu &\geq \|\mathcal{G}^{-1}(\widehat{S}-S)\|_F^2=(\Phi_x^{-1}\Phi_s)(\widehat{S}-S)\bullet(\widehat{S}-S)\\ &=(VL_x^T+L_xV^T)\bullet(L_x^{-T}V^TS+SVL_x^{-1})\\ &=2\,\mathrm{Tr}\,[L_xV^TL_x^{-T}V^TS]+2\,\mathrm{Tr}\,[L_xV^TSVL_x^{-1}]\\ &=2\,\mathrm{Tr}\,[L_x^{-T}V^TL_x^{-T}V^TSX]+2\,\mathrm{Tr}\,[V^TSV]\geq 2\,\|L_s^TV\|_F^2\\ &\geq 2\,\frac{\|L_s^TVL_x^{-1}L_s^{-T}\|_F^2}{\|L_x^{-1}S^{-1}L_x^{-T}\|}\geq\frac{1}{2}(1-\gamma)\mu\|L_s^{-1}(\widehat{S}-S)L_s^{-T}\|_F^2,\end{aligned}$$

where in the second inequality we used the inequality $\mathrm{Tr}\,[(L_x^{-T}V^T)^2SX]\geq 0$, which trivially holds due to the fact that both $(L_x^{-T}V^T)^2$ and SX are upper triangular matrices with nonnegative diagonal elements. ∎

The following technical lemma is a direct consequence of lemmas 3.4 and 3.5 of Monteiro and Tsuchiya [13].

Lemma 4.19. *Let $X,S\in\mathcal{S}^n_{++}$, $\widehat{X},\widehat{S}\in\mathcal{S}^n$, $L_x\equiv chol(X)$, and $L_s\equiv chol(S)$. Suppose there exists some $\tau\in(0,1)$ such that*

$$\|L_x^{-1}(\widehat{X}-X)L_x^{-T}\|\leq\tau,\qquad \|L_s^{-1}(\widehat{S}-S)L_s^{-T}\|\leq\tau. \tag{61}$$

Then $\widehat{X},\widehat{S}\in\mathcal{S}^n_{++}$,

$$\max\left\{\|L_x^{-1}L_{\widehat{x}}\|,\|L_{\widehat{x}}^{-1}L_x\|,\|L_s^{-1}L_{\widehat{s}}\|,\|L_{\widehat{s}}^{-1}Ls\|\right\}\leq\frac{1}{\sqrt{1-\tau}}$$

and

$$\|L_{\widehat{x}}^TL_{\widehat{s}}\|\leq\frac{\|L_x^TL_s\|}{1-\tau},\qquad \|L_{\widehat{x}}^T\widehat{S}L_{\widehat{x}}\|\leq\frac{\|L_x^TSL_x\|}{(1-\tau)^2},$$

where $L_{\widehat{x}}\equiv chol(\widehat{X})$ and $L_{\widehat{s}}\equiv chol(\widehat{S})$.

Lemma 4.20. *Let $(X,S)\in\mathcal{S}^n_{++}\times\mathcal{S}^n_{++}$ and $L_x\equiv chol(X)$. Then, for every $A,B\in\mathcal{S}^n$, we have*

$$\|\Phi''(X,S)[A,B]^{(2)}\|_F\leq(\sqrt{2}+1)\|L_x^TSL_x\|\,\|L_x^{-1}AL_x^{-T}\|_F^2+2\sqrt{2}\,\|L_x^{-1}AL_x^{-T}\|_F\|L_x^TBL_x\|_F.$$

Proof. We know that $\Phi''(X,S)[A,B]^{(2)}$ is given by (52), where $U',U''\in\mathcal{L}^n$ are determined by (53) and (54), respectively. Pre- and post-multiplying (53) by L_x^{-1} and L_x^{-T}, respectively, and applying Lemma 4.16 with $M=L_x^{-1}U'$ gives

$$\|L_x^{-1}U'\|_F\leq\frac{\|L_x^{-1}AL_x^{-T}\|_F}{\sqrt{2}}. \tag{62}$$

Moreover, pre- and post-multiplying (54) by L_x^{-1} and L_x^{-T}, respectively, rearranging to obtain

$$L_x^{-1}U''+(U'')^TL_x^{-T}=-2L_x^{-1}U'(U')^TL_x^{-T},$$

and using Lemma 4.16 with $M=L_x^{-1}U''$, (62), and standard properties of norms, we obtain

$$\|L_x^{-1}U''\|_F\leq\sqrt{2}\|L_x^{-1}U'(U')^TL_x^{-T}\|_F\leq\sqrt{2}\|L_x^{-1}U'\|_F^2\leq\frac{\|L_x^{-1}AL_x^{-T}\|_F^2}{\sqrt{2}}. \tag{63}$$

Taking the Frobenius norm of (52), we have

$$\|\Phi''(X,S)[A,B]^{(2)}\|_F \le 2\|L_x^T S U''\|_F + 2\|(U')^T S U'\|_F + 4\|L_x^T B U'\|_F.$$

Using (62) and (63) and the standard properties of norms, we can bound each of the three terms on the right-hand side as follows:

$$\begin{aligned}
\|L_x^T S U''\|_F &= \|L_x^T S L_x L_x^{-1} U''\|_F \le \|L_x^T S L_x\| \|L_x^{-1} U''\|_F \\
&\le \frac{1}{\sqrt{2}}\, \|L_x^T S L_x\| \|L_x^{-1} A L_x^{-T}\|_F^2, \\
\|(U')^T S U'\|_F &= \|(U')^T L_x^{-T} L_x^T S L_x L_x^{-1} U'\|_F \le \|L_x^T S L_x\| \|L_x^{-1} U'\|_F^2 \\
&\le \frac{1}{2}\, \|L_x^T S L_x\| \|L_x^{-1} A L_x^{-T}\|_F^2, \\
\|L_x^T B U'\|_F &= \|L_x^T B L_x L_x^{-1} U'\|_F \le \frac{1}{\sqrt{2}}\, \|L_x^{-1} A L_x^{-T}\|_F \|L_x^T B L_x\|_F.
\end{aligned}$$

The desired inequality follows by combining the last four relations. ∎

Proposition 4.21. *Let $\gamma \in (0,1)$, and let $\theta^* = \theta^*(\gamma)$ be given by (59). The following statements hold:*

(a) $\Omega_{\theta^*}(X,S) \le (83/(1-\gamma))n$ *for any* $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$*;*

(b) $\Omega_{\theta^*}(\gamma) \le (83/(1-\gamma))n^2$.

Proof. To prove (a), fix $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, and define $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$. Also, let $(\widehat{X},\widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ and matrices $\widetilde{A}, \widetilde{B} \in \mathcal{S}^n$ such that $\|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2 \le 1$ be arbitrarily given. Note that Proposition 4.18 implies that $(\widehat{X},\widehat{S}) \in \mathcal{S}_{++}^n \times \mathcal{S}_{++}^n$ and that (X,S), $(\widehat{X},\widehat{S})$ satisfy the hypotheses of Lemma 4.19 with $\tau = 1/2$. Hence,

$$\begin{aligned}
\max\left\{\|L_x^{-1} L_{\widehat{x}}\|, \|L_{\widehat{x}}^{-1} L_x\|\right\} &\le \sqrt{2}, \\
\|L_{\widehat{x}}^T \widehat{S} L_{\widehat{x}}\| &\le 4\, \|L_x^T S L_x\|.
\end{aligned}$$

Using these inequalities, Lemma 4.20 applied to $(\widehat{X},\widehat{S})$ with $A = \mathcal{G}^{-1}(\widetilde{A})$ and $B = \mathcal{G}(\widetilde{B})$, Lemma 4.12, and the standard properties of norms, we obtain

$$\begin{aligned}
\|\Phi''(\widehat{X},\widehat{S})[\mathcal{G}^{-1}(\widetilde{A}),\mathcal{G}(\widetilde{B})]^{(2)}\|_F &\le (\sqrt{2}+1)\|L_{\widehat{x}}^T \widehat{S} L_{\widehat{x}}\|\, \|L_{\widehat{x}}^{-1}\mathcal{G}^{-1}(\widetilde{A}) L_{\widehat{x}}^{-T}\|_F^2 \\
&\quad + 2\sqrt{2}\|L_{\widehat{x}}^{-1}\mathcal{G}^{-1}(\widetilde{A}) L_{\widehat{x}}^{-T}\|_F \|L_{\widehat{x}}^T \mathcal{G}(\widetilde{B}) L_{\widehat{x}}\|_F \\
&\le 16\,(\sqrt{2}+1)\|L_x^T S L_x\|\, \|L_x^{-1}\mathcal{G}^{-1}(\widetilde{A}) L_x^{-T}\|_F^2 \\
&\quad + 8\sqrt{2}\|L_x^{-1}\mathcal{G}^{-1}(\widetilde{A}) L_x^{-T}\|_F \|L_x^T \mathcal{G}(\widetilde{B}) L_x\|_F \\
&= 16\,(\sqrt{2}+1)\|L_x^T S L_x\|\, \|((\Phi_s^*)^{-1}\mathcal{G}^{-1})(\widetilde{A})\|_F^2 \\
&\quad + 8\sqrt{2}\|((\Phi_s^*)^{-1}\mathcal{G}^{-1})(\widetilde{A})\|_F \|(\Phi_s\mathcal{G})(\widetilde{B})\|_F. \qquad (64)
\end{aligned}$$

An easy application of Lemma 4.1 shows that

$$\|((\Phi_s^*)^{-1}\mathcal{G}^{-1})(\widetilde{A})\|_F^2 \le \|\mathcal{R}^2\|\, \|\widetilde{A}\|_F^2, \qquad \|(\Phi_s\mathcal{G})(\widetilde{B})\|_F^2 \le \|\mathcal{R}^{-2}\|\, \|\widetilde{B}\|_F^2, \qquad (65)$$

where $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$. Using (64), (65), Lemma 4.17, Proposition 4.15 and the inequality $\|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2 \le 1$, we have

$$\begin{aligned}
&\|\Phi''(\widehat{X},\widehat{S})[\mathcal{G}^{-1}(\widetilde{A}),\mathcal{G}(\widetilde{B})]^{(2)}\|_F \le\ 16\,(\sqrt{2}+1)\|L_x^T S L_x\|\, \|\mathcal{R}^2\|\, \|\widetilde{A}\|_F^2 + 8\sqrt{2}\|\mathcal{R}\|\, \|\mathcal{R}^{-1}\|\, \|\widetilde{A}\|_F \|\widetilde{B}\|_F \\
&\quad \le \left(\frac{16(\sqrt{2}+1)\sqrt{2}}{1-\gamma}\right) n\, \|\widetilde{A}\|_F^2 + \left(\frac{32}{\sqrt{1-\gamma}}\right) \sqrt{n}\, \|\widetilde{A}\|_F\, \|\widetilde{B}\|_F \\
&\quad \le \left(\frac{55}{1-\gamma}\right) n \left(\|\widetilde{A}\|_F^2 + \|\widetilde{A}\|_F \|\widetilde{B}\|_F\right) \le \left(\frac{83}{1-\gamma}\right) n \left(\|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2\right) \le \left(\frac{83}{1-\gamma}\right) n.
\end{aligned}$$

Hence, we conclude from (23) that $\|\Phi''(\widehat{X},\widehat{S})\|_{\mathcal{G}} \leq (83/(1-\gamma))n$. Since this upper bound does not depend on the choice of $(\widehat{X},\widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$, (28) implies that $\Omega_{\theta^*}(X,S) \leq (83/(1-\gamma))n$.

Statement (b) follows immediately from (a), Lemma 4.14, and (31). ∎

The main iteration-complexity result for the long-step algorithm based on the map (4) and the set (55) can now be derived as a consequence of Theorem 3.3 and the results of this subsection.

Theorem 4.22. *The long-step algorithm based on the map $\Phi(X,S) = L_x^T S L_x$ and the set $\mathcal{C}_0$ of (55) terminates in at most $\mathcal{O}(n^2 L)$ iterations.*

Proof. We have already seen that Φ and $\mathcal{C}_0$ satisfy Assumption 1 through 9. Let $\theta^* = \theta^*(\gamma)$ be as in (59). Then by (40) and Lemma 4.21, we have

$$\Omega(\gamma) \leq \max\left\{\Omega_{\theta^*}(\gamma), 4(\theta^*)^{-1}\right\} \leq \max\left\{\left(\frac{83}{1-\gamma}\right)n^2, 8\sqrt{2}\left(\frac{n}{1-\gamma}\right)^{1/2}\right\}.$$

The conclusion of the theorem now follows from Theorem 3.3. ∎

4.3 The $X^{1/2}SX^{1/2}$ Map

A polynomial iteration complexity for the long-step algorithm based on the map $\Phi(X,S) = X^{1/2}SX^{1/2}$ can be established in a manner very similar to that of the previous subsection for the map $\Phi(X,S) = L_x^T S L_x$. In fact, each result for $L_x^T S L_x$ has an easily derived, analogous result for $X^{1/2}SX^{1/2}$, and the same overall iteration complexity of $\mathcal{O}(n^2 L)$ holds for the $X^{1/2}SX^{1/2}$ map.

4.4 The $W^{1/2}XSW^{-1/2}$ Map

In this section, we consider the long-step algorithm based on the map $\Phi(X,S) = W^{1/2}XSW^{-1/2}$, where

$$W \equiv W(X,S) \equiv X^{-1/2}(X^{1/2}SX^{1/2})^{1/2}X^{-1/2}$$

is the unique symmetric matrix such that $WXW = S$. We show that Φ satisfies Assumptions 1 through 4 and that $\mathcal{C}_0$ can be chosen so that Assumptions 5 through 9 are satisfied. As a result, we obtain for the first time a polynomial long-step algorithm based on this choice of the map Φ.

It is easy to see that the map Φ satisfies Assumptions 1 through 4. Moreover, letting $V = V(X,S) \equiv W^{1/2}XW^{1/2} = W^{-1/2}SW^{-1/2}$, we have $\Phi(X,S) = V^2$. This alternative expression for Φ can be used to determine the functions Φ_x and Φ_s. Indeed, for all $A, B \in \mathcal{S}^n$,

$$\Phi_x(A) = V_x V + V V_x, \qquad \Phi_s(B) = V_s V + V V_s, \tag{66}$$

where $V_x \equiv V'(X,S)[A,0]$ and $V_s \equiv V'(X,S)[0,B]$. A good choice for the C_0 that fits well with the map (6) is

$$\mathcal{C}_0 \equiv \{(X,S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} : XS \in \mathcal{S}^n\}. \tag{67}$$

The proof of the following lemma, while not difficult, is quite lengthy, and for this reason we include it in the appendix. As a consequence of the lemma and some additional straightforward arguments, we obtain Proposition 4.24.

Lemma 4.23. *Let $(X,S) \in \mathcal{C}_0$. Then $\Phi_x, \Phi_s \in \mathcal{T}^n_{++}$ and $\Phi_x\Phi_s = \Phi_s\Phi_x$.*

Proposition 4.24. *The set $\mathcal{C}_0$ given by (67) satisfies Assumptions 5 and 6.*

The following lemma and proposition together verify Assumption 7 and establish two important inequalities which we will use throughout the remainder of this subsection. The proof of the lemma is straightforward, and the proof of the proposition is similar to the proof of Proposition 4.15. Hence, both proofs have been omitted.

Lemma 4.25. *Let $(X,S) \in \mathcal{C}_0$, $\Phi \equiv \Phi(X,S)$, and $\mu \equiv \mu(X,S)$. Then*

$$\|\mathcal{R}(\Phi)\|_F^2 = X \bullet S = n\mu \qquad \|\mathcal{R}(I)\|_F^2 = X^{-1} \bullet S^{-1}.$$

Proposition 4.26. *Let $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, and define $\mu \equiv \mu(X,S)$. Then*

$$\|\Phi(X,S)\| \le n\mu, \qquad \|\Phi(X,S)^{-1}\| \le \frac{1}{(1-\gamma)\mu}, \qquad \Pi(X,S)^2 \le \frac{1}{1-\gamma}.$$

In particular, $\Pi(\gamma)^2 \le 1/(1-\gamma)$.

The next three results together verify that the pair $(\Phi, \mathcal{C}_0)$ satisfies Assumption 8. Since Lemma 4.27 is technical and its proof is long, the proof is given in the appendix.

Lemma 4.27. *Let $(X,S) \in \mathcal{C}_0$, and suppose that $E \in \mathcal{S}^n$ commutes with both $W \equiv W(X,S)$ and $V \equiv V(X,S)$. Then, for all $M, N \in \mathcal{S}^n$,*

$$\frac{\|EW^{-1/2}MW^{-1/2}E\|_F}{2\,\|V\|} \le \|E\Phi_x^{-1}(M)E\|_F \le 2\,\|V^{-1}\|\,\|EW^{-1/2}MW^{-1/2}E\|_F, \tag{68}$$

$$\frac{\|EW^{1/2}NW^{1/2}E\|_F}{2\,\|V\|} \le \|E\Phi_s^{-1}(N)E\|_F \le 2\,\|V^{-1}\|\,\|EW^{1/2}NW^{1/2}E\|_F. \tag{69}$$

Lemma 4.28. *Let $(X,S) \in \mathcal{C}_0$, and let $W \equiv W(X,S)$ and $V \equiv V(X,S)$. Then,*

$$\|\mathcal{R}^2\| \le 4\|V^{-2}\|, \qquad \|\mathcal{R}^{-2}\| \le 4\|V^2\|.$$

Moreover, for all $A, B \in \mathcal{S}^n$, there hold:

$$\|\mathcal{G}(A)\|_F^2 \ge \frac{\|W^{1/2}AW^{1/2}\|_F^2}{16\,\|V^2\|\,\|V^{-2}\|}, \qquad \|\mathcal{G}^{-1}(B)\|_F^2 \ge \frac{\|W^{-1/2}BW^{-1/2}\|_F^2}{16\,\|V^2\|\,\|V^{-2}\|}.$$

Proof. Let $M \in \mathcal{S}^n$ such that $\|M\|_F \le 1$ be given. Using (14), Lemma 4.23, (69) with $E = I$ and $N = \Phi_x^{-1}(M)$, and (68) with $E = W^{1/2}$, we obtain

$$\begin{aligned}\|\mathcal{R}^2(M)\|_F &= \|(\Phi_s^*)^{-1}\Phi_x^{-1}(M)\|_F = \|\Phi_s^{-1}\Phi_x^{-1}(M)\|_F \\ &\le 2\,\|V^{-1}\|\,\|W^{1/2}\Phi_x^{-1}(M)W^{1/2}\|_F \le 4\,\|V^{-2}\|\,\|M\|_F \le 4\,\|V^{-2}\|.\end{aligned}$$

By the definition of $\|\cdot\|$ in $\mathcal{T}_{++}^n$, we obtain the first inequality of the lemma. A similar argument yields the second inequality.

To prove the third inequality of the lemma, let $A \in \mathcal{S}^n$ be given. From Lemma 4.1, the inequality proved in the previous paragraph, and (69) with $E = I$ and $N = A$, we obtain

$$\|\mathcal{G}(A)\|_F^2 \ge \|\Phi_s^{-1}(A)\|_F^2/\|\mathcal{R}^2\| \ge \|\Phi_s^{-1}(A)\|_F^2/(4\|V^{-2}\|) \ge \|W^{1/2}AW^{1/2}\|_F^2/(16\,\|V^2\|\,\|V^{-2}\|).$$

In a similar fashion, we can establish the final inequality of the lemma. ∎

Proposition 4.29. *Let $\gamma \in (0,1)$, and define $\theta^* \equiv \theta^*(\gamma) \equiv (1-\gamma)/(8n)$. Then,*

$$\|X^{-1/2}(\widehat{X} - X)X^{-1/2}\| \le \frac{1}{2}, \qquad \|S^{-1/2}(\widehat{S} - S)S^{-1/2}\| \le \frac{1}{2}$$

for all $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ and $(\widehat{X}, \widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$. Hence, $\mathcal{E}_{\theta^*}(X,S) \subseteq \mathcal{S}_{++}^n \times \mathcal{S}_{++}^n$ for all $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$. In particular, $\Theta(\gamma) \ge \theta^*$.

Proof. Let $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ and $(\widehat{X}, \widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ be given. Let $\Phi \equiv \Phi(X,S)$, $W \equiv W(X,S)$, $V \equiv V(X,S)$, $\mu \equiv \mu(X,S)$, $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$, and $\mathcal{R} \equiv \mathcal{R}_{(X,S)}$. Notice that, for any $A \in \mathcal{S}^n$, we have

$$\begin{aligned}\|W^{1/2}AW^{1/2}\|_F &= \|W^{1/2}X^{1/2}(X^{-1/2}AX^{-1/2})X^{1/2}W^{1/2}\|_F \\ &\ge \|X^{-1/2}AX^{-1/2}\|_F/\|W^{-1/2}X^{-1}W^{-1/2}\| = \|X^{-1/2}AX^{-1/2}\|_F/\|V^{-1}\|.\end{aligned} \tag{70}$$

Hence, (70) with $A = \widehat{X} - X$, Lemma 4.28 with $A = \widehat{X} - X$, Lemma 4.25, Proposition 4.26, and the fact that $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$ and $(\widehat{X}, \widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ imply

$$\begin{aligned}\frac{(1-\gamma)^2\mu}{64n} &= (\theta^*)^2 n\mu = (\theta^*)^2 \|\mathcal{R}(\Phi)\|_F^2 \geq \|\mathcal{G}(\widehat{X} - X)\|_F^2 + \|\mathcal{G}^{-1}(\widehat{S} - S)\|_F^2 \\ &\geq \|\mathcal{G}(\widehat{X} - X)\|_F^2 \geq \|W^{1/2}(\widehat{X} - X)W^{1/2}\|_F^2 / (16\,\|V^2\|\,\|V^{-2}\|) \\ &\geq \|X^{-1/2}(\widehat{X} - X)X^{-1/2}\|_F^2 / (16\,\|V^2\|\,\|V^{-2}\|^2) \\ &\geq \left(\frac{(1-\gamma)^2\mu}{16n}\right) \|X^{-1/2}(\widehat{X} - X)X^{-1/2}\|_F^2,\end{aligned}$$

which implies the first inequality of the proposition. Similarly, we can show the second. ∎

Finally, the next three results together verify that Assumption 9 is satisfied. The proofs of Lemmas 4.30 and 4.31 are given in the appendix.

Lemma 4.30. *Let $(X,S) \in \mathcal{C}_0$, $W \equiv W(X,S)$, and $V \equiv V(X,S)$. Then, for all $A, B \in \mathcal{S}^n$,*

$$\|\Phi''(X,S)[A,B]^{(2)}\|_F \leq 18\,\|V^2\|\|V^{-2}\| \left(\|W^{1/2}AW^{1/2}\|_F^2 + \|W^{-1/2}BW^{-1/2}\|_F^2\right).$$

Lemma 4.31. *Let $X, S \in \mathcal{S}^n_{++}$ and $\widehat{X}, \widehat{S} \in \mathcal{S}^n$. Suppose there exists some $\tau \in (0,1)$ such that*

$$\|X^{-1/2}(\widehat{X} - X)X^{-1/2}\| \leq \tau, \qquad \|S^{-1/2}(\widehat{S} - S)S^{-1/2}\| \leq \tau. \tag{71}$$

Then $\widehat{X}, \widehat{S} \in \mathcal{S}^n_{++}$,

$$\|\widehat{V}\| \leq \frac{\|V\|}{1-\tau}, \qquad \|\widehat{V}^{-1}\| \leq \frac{\|V^{-1}\|}{1-\tau},$$
$$\max\left\{\|\widehat{W}^{1/2}W^{-1/2}\|^2, \|\widehat{W}^{-1/2}W^{1/2}\|^2\right\} \leq \frac{1}{1-\tau},$$

where $W \equiv W(X,S)$, $V \equiv V(X,S)$, $\widehat{W} \equiv W(\widehat{X}, \widehat{S})$, and $\widehat{V} \equiv V(\widehat{X}, \widehat{S})$.

Proposition 4.32. *Let $\gamma \in (0,1)$ be given, and let $\theta^* = \theta^*(\gamma)$ be as in Proposition 4.29. The following statements hold:*

(a) $\Omega_{\theta^*}(X,S) \leq (18432/(1-\gamma)^2)n^2$ *for any* $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$*;*
(b) $\Omega_{\theta^*}(\gamma) \leq (18432/(1-\gamma)^2)n^3$.

Proof. To prove (a), fix $(X,S) \in \mathcal{C}_0 \cap \mathcal{N}_{-\infty}(\gamma)$, and define $\mathcal{G} \equiv \mathcal{G}_{(X,S)}$. In addition, let $(\widehat{X}, \widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$ and matrices $A, B \in \mathcal{S}^n$ such that $\|A\|_F^2 + \|B\|_F^2 \leq 1$ be arbitrarily given. Note that Proposition 4.29 implies that $(\widehat{X}, \widehat{S}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ and that $(X,S), (\widehat{X}, \widehat{S})$ satisfy the hypotheses of Lemma 4.31 with $\tau = 1/2$. Hence

$$\|\widehat{W}^{1/2}M\widehat{W}^{1/2}\|_F^2 \leq \|\widehat{W}^{1/2}W^{-1/2}\|^4 \|W^{1/2}MW^{1/2}\|_F^2 \leq 4\,\|W^{1/2}MW^{1/2}\|_F^2,$$

for any $M \in \mathcal{S}^n$, where $W \equiv W(X,S)$, $V \equiv V(X,S)$, $\widehat{W} \equiv W(\widehat{X}, \widehat{S})$, and $\widehat{V} \equiv V(\widehat{X}, \widehat{S})$ and similarly $\|\widehat{W}^{-1/2}N\widehat{W}^{-1/2}\|_F^2 \leq 4\|W^{-1/2}NW^{-1/2}\|_F^2$ for any $N \in \mathcal{S}^n$. These two inequalities, Lemmas 4.30 and 4.31, and Proposition 4.26 imply

$$\begin{aligned}\|\Phi''(\widehat{X}, \widehat{S})[M,N]^{(2)}\|_F &\leq 18\,\|\widehat{V}^2\|\,\|\widehat{V}^{-2}\| \left(\|\widehat{W}^{1/2}M\widehat{W}^{1/2}\|_F^2 + \|\widehat{W}^{-1/2}N\widehat{W}^{-1/2}\|_F^2\right) \\ &\leq 1152\,\|V^2\|\,\|V^{-2}\| \left(\|W^{1/2}MW^{1/2}\|_F^2 + \|W^{-1/2}NW^{-1/2}\|_F^2\right).\end{aligned}$$

Now letting $M = \mathcal{G}^{-1}(A)$ and $N = \mathcal{G}(B)$ and applying Lemma 4.28, Proposition 4.26, and the inequality $\|A\|_F^2 + \|B\|_F^2 \leq 1$, we obtain

$$\begin{aligned}\|\Phi''(\widehat{X}, \widehat{S})[\mathcal{G}^{-1}(A), \mathcal{G}(B)]^{(2)}\|_F &\leq 1152\,\|V^2\|\,\|V^{-2}\| \left(\|W^{1/2}\mathcal{G}^{-1}(A)W^{1/2}\|_F^2 + \|W^{-1/2}\mathcal{G}(B)W^{-1/2}\|_F^2\right) \\ &\leq 18432\,\|V^2\|^2\,\|V^{-2}\|^2 \left(\|A\|_F^2 + \|B\|_F^2\right) \\ &\leq 18432 \left(\frac{n}{1-\gamma}\right)^2 \left(\|A\|_F^2 + \|B\|_F^2\right) \leq \left(\frac{18432}{(1-\gamma)^2}\right) n^2.\end{aligned}$$

Hence, we conclude from (23) that $\|\Phi''(\widehat{X},\widehat{S})\|_{\mathcal{G}} \le (18432/(1-\gamma)^2)n^2$. Since this upper bound does not depend on the choice of $(\widehat{X},\widehat{S}) \in \mathcal{E}_{\theta^*}(X,S)$, (28) implies that $\Omega_{\theta^*}(X,S) \le (18432/(1-\gamma)^2)n^2$.

Statement (b) follows immediately from (a), Lemma 4.25, and (31). ∎

The main iteration-complexity result for the long-step algorithm based on the map Φ and the set $\mathcal{C}_0$ can now be derived as a consequence of Theorem 3.3 and the results of this subsection.

Theorem 4.33. *The long-step algorithm based on the map $\Phi(X,S) = W^{1/2}XSW^{-1/2}$ and the set $\mathcal{C}_0$ terminates in at most $\mathcal{O}(n^3L)$ iterations.*

Proof. We have already seen that Φ and $\mathcal{C}_0$ satisfy Assumption 1 through 9. Let θ^* be as in Proposition 4.29. Then by (40) and Proposition 4.32, we have

$$\Omega(\gamma) \le \max\left\{\Omega_{\theta^*}(\gamma), 4(\theta^*)^{-1}\right\} \le \max\left\{\left(\frac{18432}{(1-\gamma)^2}\right)n^3, \frac{32n}{1-\gamma}\right\}.$$

The conclusion of the theorem now follows from Theorem 3.3. ∎

REFERENCES

1. F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5:13–51, 1995.
2. F. Alizadeh, J.-P.A. Haeberly, and M.L. Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM Journal on Optimization*, 8:746–768, 1998.
3. G. H. Golub and C. E. Van Loan. *Matrix Computations: Second Edition.* The John Hopkins University Press, Baltimore, MA 21211, 1989.
4. C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6:342–361, 1996.
5. R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, New York, 1991.
6. F. Jarre. An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices. *SIAM Journal on Control and Optimization*, 31:1360–1377, 1993.
7. M. Kojima, M. Shida, and S. Shindoh. Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. *Mathematical Programming*, 80:129–160, 1998.
8. M. Kojima, S. Shindoh, and S. Hara. Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM Journal on Optimization*, 7:86–125, 1997.
9. Z.-Q. Luo, J.F. Sturm, and S. Zhang. Superlinear convergence of a symmetric primal-dual path-following algorithm for semidefinite programming. *SIAM Journal on Optimization*, 8:59–81, 1998.
10. R.D.C. Monteiro. Primal-dual path following algorithms for semidefinite programming. *SIAM Journal on Optimization*, 7:663–678, 1997.
11. R.D.C. Monteiro. Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions. *SIAM Journal on Optimization*, 8:797–812, 1998.
12. R.D.C. Monteiro and M. Todd. Path-following methods for semidefinite programming. In R. Saigal, L. Vandenberghe, and H. Wolkowicz, editors, *Handbook of Semidefinite Programming.* Kluwer Academic Publishers, 2000.
13. R.D.C. Monteiro and T. Tsuchiya. Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming. *SIAM Journal on Optimization*, 9:551–577, 1999.
14. 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. *Mathematical Programming*, 84:39–53, 1999.
15. 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. *Optimization Methods and Software*, 11/12:91–140, 1999.
16. R.D.C. Monteiro and P.R. Zanjácomo. General interior-point maps and existence of weighted paths for nonlinear semidefinite complementarity problems. *Mathematics of Operations Research*, 25:381–399, 2000.
17. R.D.C. Monteiro and Y. Zhang. A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. *Mathematical Programming*, 81:281–299, 1998.
18. Y. E. Nesterov and A. S. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming.* Society for Industrial and Applied Mathematics, Philadelphia, 1994.
19. Y. E. Nesterov and M. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22:1–42, 1997.
20. Y. E. Nesterov and M. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8:324–364, 1998.
21. F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. *SIAM Journal on Optimization*, 8:1007–1028, 1998.
22. J.F. Sturm and S. Zhang. Symmetric primal-dual path-following algorithms for semidefinite programming. *Applied Numerical Mathematics*, 29:301–315, 1999.
23. M.J. Todd. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Optimization Methods and Software*, 11:1–46, 1999.

24. P. Tseng. Search directions and convergence analysis of some infeasible path-following methods for the monotone semidefinite LCP. *Optimization Methods and Software*, 9:245–268, 1998.
25. L. Vandenberghe and S. Boyd. A primal-dual potential reduction method for problems involving matrix inequalities. *Mathematical Programming*, 69:205–236, 1995.
26. Y. Zhang. On extending some primal–dual interior–point algorithms from linear programming to semidefinite programming. *SIAM Journal on Optimization*, 8:365–386, 1998.

Appendix

In the appendix, we prove several technical results that we left unproven in Section 4.4 of the main body of the paper.

First, we must further elaborate on the derivatives of the map $\Phi(X,S) = W^{1/2}XSW^{1/2}$. Letting $U_x \equiv (W^{-1/2})'(X,S)[A,0]$ and $T_x \equiv (W^{-1})'(X,S)[A,0]$ and differentiating the identities $V = W^{-1/2}SW^{-1/2}$, $W^{-1} = W^{-1/2}W^{-1/2}$, and $X = W^{-1}SW^{-1}$, we obtain the system

$$V_x = U_xSW^{-1/2} + W^{-1/2}SU_x, \tag{72a}$$

$$T_x = U_xW^{-1/2} + W^{-1/2}U_x, \tag{72b}$$

$$A = T_xSW^{-1} + W^{-1}ST_x. \tag{72c}$$

Similarly, letting $U_s \equiv (W^{1/2})'(X,S)[0,B]$ and $T_s \equiv W'(X,S)[0,B]$ and differentiating the identities $V = W^{1/2}XW^{1/2}$, $W = W^{1/2}W^{1/2}$, and $S = WXW$, we have

$$V_s = U_sXW^{1/2} + W^{1/2}XU_s, \tag{73a}$$

$$T_s = U_sW^{1/2} + W^{1/2}U_s, \tag{73b}$$

$$B = T_sXW + WXT_s. \tag{73c}$$

Define $V_{xx} \equiv \Phi''(X,S)[A,0]^{(2)}$, $V_{xs} \equiv \Phi''(X,S)[A,0][0,B] = \Phi''(X,S)[0,B][A,0]$, and $V_{ss} \equiv \Phi''(X,S)[0,B]^{(2)}$. Then $\Phi''(X,S)[A,B]^{(2)} = V''V + 2(V')^2 + VV''$, where

$$V' \equiv V'(X,S)[A,B] = V_x + V_s,$$
$$V'' \equiv V''(X,S)[A,B]^{(2)} = V_{xx} + 2\,V_{xs} + V_{ss}.$$

Differentiating (72) at the point (X,S) in the direction $(A,0)$ yields the following system which defines V_{xx}:

$$V_{xx} = U_{xx}SW^{-1/2} + 2U_xSU_x + W^{-1/2}SU_{xx}, \tag{74a}$$

$$T_{xx} = U_{xx}W^{-1/2} + 2(U_x)^2 + W^{-1/2}U_{xx}, \tag{74b}$$

$$0 = T_{xx}SW^{-1} + 2T_xST_x + W^{-1}ST_{xx}. \tag{74c}$$

In a similar manner, differentiating (73) at the point (X,S) in the direction $(0,B)$ yields the system defining V_{ss}:

$$V_{ss} = U_{ss}XW^{1/2} + 2U_sXU_s + W^{1/2}XU_{ss}, \tag{75a}$$

$$T_{ss} = U_{ss}W^{1/2} + 2(U_s)^2 + W^{1/2}U_{ss}, \tag{75b}$$

$$0 = T_{ss}XW + 2T_sXT_s + WXT_{ss}. \tag{75c}$$

The system defining V_{xs} is obtained by differentiating (72) at the point (X,S) in the direction $(0,B)$, but before giving the system, we note that the identity $W^{-1/2}W^{1/2} = I$ implies that the directional derivative $(W^{-1/2})_s \equiv (W^{-1/2})'(X,S)[0,B]$ is related to U_s by the equation $(W^{-1/2})_s = -W^{-1/2}U_sW^{-1/2}$. Similarly, we have $(W^{-1})_s = -W^{-1}T_sW^{-1}$. Hence, the system defining V_{xs} is as follows:

$$V_{xs} = \mathrm{sim}(U_{xs}SW^{-1/2} + U_xBW^{-1/2} - U_xW^{1/2}VU_sW^{-1/2}), \tag{76a}$$

$$T_{xs} = \mathrm{sim}(U_{xs}W^{-1/2} - U_xW^{-1/2}U_sW^{-1/2}), \tag{76b}$$

$$0 = \mathrm{sim}(T_{xs}SW^{-1} + T_xBW^{-1} - T_xSW^{-1}T_sW^{-1}). \tag{76c}$$

We are now ready to prove Lemma 4.23, which states that, whenever $(X,S) \in \mathcal{C}_0$, Φ_x and Φ_s are symmetric positive definite operators that commute.

Proof of Lemma 4.23. To demonstrate that Φ_x is a symmetric operator, we must show that $\Phi_x(A) \bullet B = A \bullet \Phi_x(B)$ for all $A, B \in \mathcal{S}^n$. Indeed, let $A, B \in \mathcal{S}^n$ be arbitrary, and let $W \equiv W(X,S)$ and $V \equiv V(X,S)$. Then $\Phi_x(A) = V_x^a V + V V_x^a$, where V_x^a, U_x^a, and T_x^a are related by (72a), (72b), and (72c), and $\Phi_x(B) = V_x^b V + V V_x^b$, where V_x^b, U_x^b, and T_x^b are related by (72a), (72b), and (72c) with A replaced by B. It follows that

$$\begin{aligned}
\Phi_x(A) \bullet B &= (V_x^a V + V V_x^a) \bullet (T_x^b V + V T_x^b) \\
&= 2(V_x^a V) \bullet (T_x^b V + V T_x^b) \\
&= 2(U_x^a S W^{-1/2} V + W^{-1/2} S U_x^a V) \bullet (U_x^b W^{-1/2} V + W^{-1/2} U_x^b V + V U_x^b W^{-1/2} + V W^{-1/2} U_x^b) \\
&= 2(U_x^a S W^{-1/2} V) \bullet (U_x^b W^{-1/2} V) + 2(U_x^a S W^{-1/2} V) \bullet (W^{-1/2} U_x^b V) + \\
&\quad\; 2(U_x^a S W^{-1/2} V) \bullet (V U_x^b W^{-1/2}) + 2(U_x^a S W^{-1/2} V) \bullet (V W^{-1/2} U_x^b) + \\
&\quad\; 2(W^{-1/2} S U_x^a V) \bullet (U_x^b W^{-1/2} V) + 2(W^{-1/2} S U_x^a V) \bullet (W^{-1/2} U_x^b V) + \\
&\quad\; 2(W^{-1/2} S U_x^a V) \bullet (V U_x^b W^{-1/2}) + 2(W^{-1/2} S U_x^a V) \bullet (V W^{-1/2} U_x^b) \\
&= 2(U_x^a S^{1/2} W^{-1/2} V) \bullet (U_x^b S^{1/2} W^{-1/2} V) + \\
&\quad\; 2(W^{-1/4} U_x^a S^{1/2} W^{-1/4} V) \bullet (W^{-1/4} U_x^b S^{1/2} W^{-1/4} V) + \\
&\quad\; 2(V^{1/2} U_x^a S^{1/2} W^{-1/2} V^{1/2}) \bullet (V^{1/2} U_x^b S^{1/2} W^{-1/2} V^{1/2}) + \\
&\quad\; 2(V^{1/2} W^{-1/4} U_x^a S^{1/2} W^{-1/4} V^{1/2}) \bullet (V^{1/2} W^{-1/4} U_x^b S^{1/2} W^{-1/4} V^{1/2}) + \\
&\quad\; 2(W^{-1/4} S^{1/2} U_x^a W^{-1/4} V) \bullet (W^{-1/4} S^{1/2} U_x^b W^{-1/4} V) + \\
&\quad\; 2(W^{-1/2} S^{1/2} U_x^a V) \bullet (W^{-1/2} S^{1/2} U_x^b V) + \\
&\quad\; 2(V^{1/2} W^{-1/4} S^{1/2} U_x^a W^{-1/4} V^{1/2}) \bullet (V^{1/2} W^{-1/4} S^{1/2} U_x^b W^{-1/4} V^{1/2}) + \\
&\quad\; 2(V^{1/2} W^{-1/2} S^{1/2} U_x^a V^{1/2}) \bullet (V^{1/2} W^{-1/2} S^{1/2} U_x^b V^{1/2}) \qquad (77)
\end{aligned}$$

This final expression (77) for $\Phi_x(A) \bullet B$ is symmetric with respect to U_x^a and U_x^b and hence is symmetric with respect to A and B, i.e.,

$$\Phi_x(A) \bullet B = \Phi_x(B) \bullet A = A \bullet \Phi_x(B).$$

So Φ_x is a symmetric operator. In addition, taking $B = A$ in (77), we see that $\Phi_x(A) \bullet A$ is the sum of squares of norms of matrices involving U_x^a. Hence, using this and the fact that $U_x^a \neq 0$ whenever $A \neq 0$, we conclude that $\Phi_x(A) \bullet A > 0$ for all $A \neq 0$. Thus, $\Phi_x \in \mathcal{T}_{++}^n$. A similar argument shows that $\Phi_s \in \mathcal{T}_{++}^n$.

We now prove that Φ_x and Φ_s commute. Let $A \in \mathcal{S}^n$ be arbitrary, and consider $\Phi_x(A)$ and $\Phi_s(A)$. We have $\Phi_x(A) = V_x^a V + V V_x^a$, where V_x^a, U_x^a, and T_x^a are related by (72a), (72b), and (72c), and similarly, $\Phi_s(A) = V_s^a V + V V_s^a$, where V_s^a, U_s^a, and T_s^a are related by (73a), (73b), and (73c) with B replaced by A. Using the positive definiteness of V and the assumption that $(X,S) \in \mathcal{C}_0$, it is easy to see that the equation defining $\Phi_x(A)$ and $\Phi_s(A)$ can be simplified to

$$\begin{aligned}
V_x^a &= W^{1/2} U^a S + S U^a W^{1/2}, && (78) \\
V_s^a &= U^a X W^{1/2} + W^{1/2} X U^a, && (79) \\
T^a &= U^a W^{1/2} + W^{1/2} U^a, \\
A &= T^a V + V T^a.
\end{aligned}$$

Define $B \equiv \Phi_x(A)$ and $C \equiv \Phi_s(A)$. We have $\Phi_x(C) = V_x^c V + V V_x^c$, where V_x^c, U_x^c, and T_x^c satisfy (72a), (72b), and (72c) with A replaced by C, and $\Phi_s(B) = V_s^b V + V V_s^b$, where V_s^b, U_s^b, and T_s^b satisfy (73a), (73b), and (73c). Using the definitions of B and C, the assumption that $(X,S) \in \mathcal{C}_0$, and the positive definiteness of V, we conclude that $T_x^c = V_s^a$ and that $T_s^b = V_x^a$. Hence, $\Phi_x(C)$ and $\Phi_s(B)$ are fully specified by the sets of equations

$$V_x^c = U_x^c S W^{-1/2} + W^{-1/2} S U_x^c, \qquad V_s^b = U_s^b X W^{1/2} + W^{1/2} X U_s^b, \tag{80}$$
$$V_s^a = U_x^c W^{-1/2} + W^{-1/2} U_x^c, \qquad V_x^a = U_s^b W^{1/2} + W^{1/2} U_s^b. \tag{81}$$

We wish to show that $\Phi_x(\Phi_s(A)) = \Phi_s(\Phi_x(A))$ or equivalently that $\Phi_x(C) = \Phi_s(B)$. The definitions of $\Phi_x(C)$ and $\Phi_s(B)$ and the positive definiteness of V imply that $\Phi_x(C) = \Phi_s(B)$ if and only if $V_x^c = V_s^b$, and the positive definiteness of $W^{1/2}$ in turn implies that $V_x^c = V_s^b$ if and only if

$$W^{1/2} V_x^c + V_x^c W^{1/2} = W^{1/2} V_s^b + V_s^b W^{1/2}. \tag{82}$$

Thus, to show the commutativity of Φ_x and Φ_s, we prove (82) as follows:

$$\begin{aligned}
W^{1/2} V_x^c + V_x^c W^{1/2} &= W^{1/2} U_x^c S W^{-1/2} + U_x^c S + S U_x^c + W^{-1/2} S U_x^c W^{1/2} \\
&= W^{1/2} (U_x^c W^{-1/2} + W^{-1/2} U_x^c) S + S (U_x^c W^{-1/2} + W^{-1/2} U_x^c) W^{1/2} \\
&= W^{1/2} (U^a X W^{1/2} + W^{1/2} X U^a) S + S (U^a X W^{1/2} + W^{1/2} X U^a) W^{1/2} \\
&= (W^{1/2} U^a S + S U^a W^{1/2}) X W^{1/2} + W^{1/2} X (W^{1/2} U^a S + S U^a W^{1/2}) \\
&= (U_s^b W^{1/2} + W^{1/2} U_s^b) X W^{1/2} + W^{1/2} X (U_s^b W^{1/2} + W^{1/2} U_s^b) \\
&= W^{1/2} U_s^b X W^{1/2} + V U_s^b + U_s^b V + W^{1/2} X U_s^b W^{1/2} \\
&= W^{1/2} V_s^b + V_s^b W^{1/2},
\end{aligned}$$

where the first equality follows from the left-hand equation of (80), the third equality follows from the left-hand equation of (81) and from (79), the fifth equality follows from (78) and from the right-hand equation of (81), and the last equality follows from the right-hand equation of (80). ∎

We now provide the proof of Lemma 4.27, which is most easily seen as an application of Lemma A1 below.

Lemma A1 Let $(X, S) \in \mathcal{C}_0$, and suppose that $D \in \mathcal{S}^n$ commutes with both $W \equiv W(X, S)$ and $V \equiv V(X, S)$. Then, for all $A, B \in \mathcal{S}^n$,

$$\frac{\|D W^{1/2} A W^{1/2} D\|_F}{2\,\|V^{-1}\|} \le \|D \Phi_x(A) D\|_F \le 2\,\|V\|\,\|D W^{1/2} A W^{1/2} D\|_F, \tag{83}$$
$$\frac{\|D W^{-1/2} B W^{-1/2} D\|_F}{2\,\|V^{-1}\|} \le \|D \Phi_s(B) D\|_F \le 2\,\|V\|\,\|D W^{-1/2} B W^{-1/2} D\|_F. \tag{84}$$

Proof. We prove only (83) because the proof of (84) is similar. Letting $\widetilde{A} \equiv W^{1/2} A W^{1/2}$ and $\widetilde{T}_x \equiv W^{1/2} T_x W^{1/2}$ and using the identity $S W^{-1/2} = W^{1/2} V$, (72) can be rewritten as

$$V_x = U_x W^{1/2} V + V W^{1/2} U_x, \tag{85a}$$
$$\widetilde{T}_x = U_x W^{1/2} + W^{1/2} U_x, \tag{85b}$$
$$\widetilde{A} = \widetilde{T}_x V + V \widetilde{T}_x. \tag{85c}$$

Using (85b) and the commutativity of W, V, and D, we have

$$\begin{aligned}
\|D \widetilde{T}_x V D\|_F^2 &= \|D U_x W^{1/2} V D\|_F^2 + 2\,\mathrm{Tr}\,[D V W^{1/2} U_x D^2 W^{1/2} U_x V D] + \|D W^{1/2} U_x V D\|_F^2 \\
&= \|D U_x W^{1/2} V D\|_F^2 + 2\,\|W^{1/4} D U_x W^{1/4} V D\|_F^2 + \|D W^{1/2} U_x V D\|_F^2 \\
&\ge \|D U_x W^{1/2} V D\|_F^2 + \|D W^{1/2} U_x V D\|_F^2,
\end{aligned}$$

which in turn implies that

$$\|D U_x W^{1/2} V D\|_F + \|D W^{1/2} U_x V D\|_F \le \sqrt{2}\,\|D \widetilde{T}_x V D\|_F. \tag{86}$$

In addition, the assumption that D commutes with V implies

$$\text{Tr}\;[(D\widetilde{T}_xVD)^2] = \text{Tr}\;[V^{1/2}D\widetilde{T}_xDVD\widetilde{T}_xDV^{1/2}] = \|V^{1/2}D\widetilde{T}_xDV^{1/2}\|_F^2 \geq 0.$$

Hence, pre- and post-multiplying (85c) by D and applying Lemma 4.16 to the resulting equation $D\widetilde{A}D = D\widetilde{T}_xVD + DV\widetilde{T}_xD$ yields

$$\|D\widetilde{T}_xVD\|_F \leq \|D\widetilde{A}D\|_F/\sqrt{2}. \tag{87}$$

Now using (66), (85a), the commutativity of D and V, (86), (87), and the standard properties of norms, we obtain

$$\begin{aligned}\|D\Phi_x(A)D\|_F &\leq 2\,\|DV_xVD\|_F \leq 2\,\|DU_xW^{1/2}V^2D\|_F + 2\,\|DVW^{1/2}U_xVD\|_F \\ &\leq 2\,\|V\|\left(\|DU_xW^{1/2}VD\|_F + \|DW^{1/2}U_xVD\|_F\right) \leq 2\sqrt{2}\,\|V\|\,\|D\widetilde{T}_xVD\|_F \\ &\leq 2\,\|V\|\,\|D\widetilde{A}D\|_F,\end{aligned}$$

which proves the right-hand inequality of (83).

We now prove the left-hand inequality of (83). Using (85a) and the commutativity of W, V, and D, we have

$$\begin{aligned}\|DV_xVD\|_F^2 &= \|DU_xW^{1/2}V^2D\|_F^2 + 2\,\text{Tr}\;[DV^2W^{1/2}U_xD^2VW^{1/2}U_xVD] + \|DVW^{1/2}U_xVD\|_F^2 \\ &= \|DU_xW^{1/2}V^2D\|_F^2 + 2\,\|DV^{1/2}W^{1/4}U_xW^{1/4}V^{3/2}D\|_F^2 + \|DVW^{1/2}U_xVD\|_F^2 \\ &\geq \|DU_xW^{1/2}V^2D\|_F^2 + \|DVW^{1/2}U_xVD\|_F^2.\end{aligned} \tag{88}$$

In addition, the assumption that V commutes with D implies

$$\text{Tr}\;[(DV_xVD)^2] = \text{Tr}\;[V^{1/2}DV_xDVDV_xDV^{1/2}] = \|V^{1/2}DV_xDV^{1/2}\|_F^2 \geq 0,$$

and hence, pre- and post-multiplying $\Phi_x(A)$ by D and applying Lemma 4.16 on the resulting equation $D\Phi_x(A)D = DV_xVD + DVV_xD$ yields

$$\|D\Phi_x(A)D\|_F \geq \sqrt{2}\,\|DV_xVD\|_F. \tag{89}$$

Using (89), (88), the commutativity of V and D, (85a), (85b), and the standard properties of norms, we obtain the left-hand inequality of (83) as follows:

$$\begin{aligned}\|D\Phi_x(A)D\|_F &\geq \sqrt{2}\,\|DV_xVD\|_F \geq \sqrt{2}\left(\|DU_xW^{1/2}V^2D\|_F^2 + \|DVW^{1/2}U_xVD\|_F^2\right)^{1/2} \\ &\geq \left(\|DU_xW^{1/2}VD\|_F + \|DW^{1/2}U_xVD\|_F\right)/\|V^{-1}\| \geq \|D\widetilde{T}_xVD\|_F/\|V^{-1}\| \\ &\geq \|D\widetilde{A}D\|_F/\left(2\,\|V^{-1}\|\right).\end{aligned}$$

■

Proof of Lemma 4.27. Applying (83) with $D = EW^{-1/2}$ and $A = \Phi_x^{-1}(M)$, we obtain (68). Similarly, (69) follows from (84) with $D = EW^{1/2}$ and $B = \Phi_s^{-1}(N)$. ■

The proof of Lemma 4.30, which involves a bound on the second derivative of Φ, is given next.

Proof of Lemma 4.30. Consider (72) through (76), and let $\widetilde{A} \equiv W^{1/2}AW^{1/2}$, $\widetilde{B} \equiv W^{-1/2}BW^{-1/2}$, $\widetilde{T}_x \equiv W^{1/2}T_xW^{1/2}$, $\widetilde{T}_s \equiv W^{-1/2}T_sW^{-1/2}$, $\widetilde{T}_{xx} \equiv W^{1/2}T_{xx}W^{1/2}$, $\widetilde{T}_{ss} \equiv W^{-1/2}T_{ss}W^{-1/2}$, and $\widetilde{T}_{xs} \equiv W^{1/2}T_{xs}W^{1/2}$. It is easy to see that the systems of equations (72), (74), and (76) can be rewritten as the systems (85),

$$V_{xx} = U_{xx}W^{1/2}V + 2\,U_xW^{1/2}VW^{1/2}U_x + VW^{1/2}U_{xx}, \tag{90a}$$

$$\widetilde{T}_{xx} - 2\,W^{1/2}(U_x)^2W^{1/2} = U_{xx}W^{1/2} + W^{1/2}U_{xx}, \tag{90b}$$

$$-2\,\widetilde{T}_xV\widetilde{T}_x = \widetilde{T}_{xx}V + V\widetilde{T}_{xx}, \tag{90c}$$

and

$$
\begin{aligned}
V_{xs} &= \mathrm{sim}(U_{xs}W^{1/2}V + U_xW^{1/2}\widetilde{B} - U_xW^{1/2}VU_sW^{-1/2}), && \text{(91a)}\\
\widetilde{T}_{xs} + \mathrm{sim}(W^{1/2}U_xW^{-1/2}U_s) &= U_{xs}W^{1/2} + W^{1/2}U_{xs}, && \text{(91b)}\\
\mathrm{sim}(\widetilde{T}_xV\widetilde{T}_s - \widetilde{T}_x\widetilde{B}) &= \widetilde{T}_{xs}V + V\widetilde{T}_{xs}, && \text{(91c)}
\end{aligned}
$$

respectively. Then (85b), (85c), Lemma 4.9, and the standard properties of norms imply

$$
\|U_xW^{1/2}\|_F \le \frac{1}{\sqrt{2}}\|\widetilde{T}_x\|_F \le \frac{1}{\sqrt{2}}\|V^{-1}\|\,\|\widetilde{T}_xV\|_F \le \frac{1}{2}\|V^{-1}\|\,\|\widetilde{A}\|_F. \tag{92}
$$

In addition, (90b), (90c), (92), and Lemma 4.9 imply

$$
\begin{aligned}
\|U_{xx}W^{1/2}\|_F &\le \frac{1}{\sqrt{2}}\|\widetilde{T}_{xx}\|_F + \sqrt{2}\,\|U_xW^{1/2}\|_F^2 \le \frac{1}{\sqrt{2}}\|V^{-1}\|\,\|\widetilde{T}_{xx}V\|_F + \frac{1}{2\sqrt{2}}\,\|V^{-2}\|\,\|\widetilde{A}\|_F^2\\
&\le \|V^{-1}\|\,\|\widetilde{T}_xV\widetilde{T}_x\|_F + \frac{1}{2\sqrt{2}}\,\|V^{-2}\|\,\|\widetilde{A}\|_F^2\\
&\le \|V^{-2}\|\,\|\widetilde{T}_xV\|_F^2 + \frac{1}{2\sqrt{2}}\,\|V^{-2}\|\,\|\widetilde{A}\|_F^2 \le \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)\|V^{-2}\|\,\|\widetilde{A}\|_F^2. && \text{(93)}
\end{aligned}
$$

Similar arguments with (73b), (73c), (75b), and (75c) show

$$
\begin{aligned}
\|U_sW^{-1/2}\|_F &\le \frac{1}{\sqrt{2}}\|V^{-1}\|\,\|\widetilde{T}_sV\|_F \le \frac{1}{2}\|V^{-1}\|\,\|\widetilde{B}\|_F, && \text{(94)}\\
\|U_{ss}W^{-1/2}\|_F &\le \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)\|V^{-2}\|\,\|\widetilde{B}\|_F^2. && \text{(95)}
\end{aligned}
$$

Using (91b), (91c), (92), (94), and Lemma 4.9, we also obtain

$$
\begin{aligned}
\|U_{xs}W^{1/2}\|_F &\le \frac{1}{\sqrt{2}}\|\widetilde{T}_{xs}\|_F + \sqrt{2}\,\|U_xW^{1/2}\|_F\|U_sW^{-1/2}\|_F\\
&\le \frac{1}{\sqrt{2}}\|V^{-1}\|\,\|\widetilde{T}_{xs}V\|_F + \frac{1}{2\sqrt{2}}\|V^{-2}\|\,\|\widetilde{A}\|_F\|\widetilde{B}\|_F\\
&\le \|V^{-1}\|\left(\|\widetilde{T}_x\|_F\|\widetilde{T}_sV\|_F + \|\widetilde{T}_x\|_F\|\widetilde{B}\|_F\right) + \frac{1}{2\sqrt{2}}\|V^{-2}\|\,\|\widetilde{A}\|_F\|\widetilde{B}\|_F\\
&\le \|V^{-2}\|\left(\|\widetilde{T}_xV\|_F\|\widetilde{T}_sV\|_F + \|\widetilde{T}_xV\|_F\|\widetilde{B}\|_F\right) + \frac{1}{2\sqrt{2}}\|V^{-2}\|\,\|\widetilde{A}\|_F\|\widetilde{B}\|_F\\
&\le \left(\frac{1}{2} + \frac{1}{\sqrt{2}} + \frac{1}{2\sqrt{2}}\right)\|V^{-2}\|\,\|\widetilde{A}\|_F\|\widetilde{B}\|_F. && \text{(96)}
\end{aligned}
$$

We next use (85a), (92), and the standard properties of norms to establish

$$
\|V_x\|_F \le 2\,\|U_xW^{1/2}V\|_F \le \|V\|\,\|V^{-1}\|\,\|\widetilde{A}\|_F,
$$

and using (90a), (92), and (93), we see that

$$
\|V_{xx}\|_F \le 2\,\|V\|\left(\|U_{xx}W^{1/2}\|_F + \|U_xW^{1/2}\|_F^2\right) \le \left(\frac{3}{2} + \frac{1}{\sqrt{2}}\right)\|V\|\,\|V^{-2}\|\,\|\widetilde{A}\|_F^2.
$$

Similar arguments with (73a), (75a), (94), and (95) show that

$$
\|V_s\|_F \le \|V\|\,\|V^{-1}\|\,\|\widetilde{B}\|_F, \qquad \|V_{ss}\|_F \le \left(\frac{3}{2} + \frac{1}{\sqrt{2}}\right)\|V\|\,\|V^{-2}\|\,\|\widetilde{B}\|_F^2.
$$

Moreover, (91a), (92), (94), (96), and the inequality $\|V\|\,\|V^{-1}\| = \mathrm{cond}\,[V] \geq 1$ imply

$$\begin{aligned}\|V_{xs}\|_F &\leq 2\,\|V\|\,\|U_{xs}W^{1/2}\|_F + 2\,\|U_xW^{1/2}\|_F\|\widetilde{B}\|_F + 2\,\|V\|\,\|U_xW^{1/2}\|_F\|U_sW^{-1/2}\|_F \\ &\leq \left(\frac{5}{2} + \sqrt{2} + \frac{1}{\sqrt{2}}\right)\|V\|\,\|V^{-2}\|\,\|\widetilde{A}\|_F\|\widetilde{B}\|_F.\end{aligned}$$

Combining the last five inequalities, we obtain

$$\begin{aligned}\|\Phi''(X,S)[A,B]^{(2)}\|_F &\leq 2\,\|V\|\,\|V''\|_F + 2\|V'\|_F^2 \\ &\leq 2\,\|V\|\,(\|V_{xx}\|_F + 2\,\|V_{xs}\|_F + \|V_{ss}\|_F) + 2\,(\|V_x\|_F + \|V_s\|_F)^2 \\ &\leq \|V^2\|\,\|V^{-2}\|\left((5+\sqrt{2})\|\widetilde{A}\|_F^2 + (14+6\sqrt{2})\|\widetilde{A}\|_F\|\widetilde{B}\|_F + (5+\sqrt{2})\|\widetilde{B}\|_F^2\right) \\ &\leq 18\,\|V^2\|\,\|V^{-2}\|\left(\|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2\right).\end{aligned}$$

■

The following technical lemma is a direct consequence of lemmas 3.4 and 3.5 of Monteiro and Tsuchiya [13] and is used in the proof of Lemma 4.31 below.

Lemma A2 Let $X, S \in \mathcal{S}^n_{++}$ and $\widehat{X}, \widehat{S} \in \mathcal{S}^n$. Suppose there exists some $\tau \in (0,1)$ such that

$$\|X^{-1/2}(\widehat{X} - X)X^{-1/2}\| \leq \tau, \qquad \|S^{-1/2}(\widehat{S} - S)S^{-1/2}\| \leq \tau. \tag{97}$$

Then $\widehat{X}, \widehat{S} \in \mathcal{S}^n_{++}$,

$$\max\left\{\|\widehat{X}^{1/2}X^{-1/2}\|, \|X^{1/2}\widehat{X}^{-1/2}\|, \|\widehat{S}^{1/2}S^{-1/2}\|, \|S^{1/2}\widehat{S}^{-1/2}\|\right\} \leq \frac{1}{\sqrt{1-\tau}},$$

and

$$\|\widehat{X}^{1/2}\widehat{S}^{1/2}\| \leq \frac{\|X^{1/2}S^{1/2}\|}{1-\tau}, \qquad \|\widehat{X}^{-1/2}\widehat{S}^{-1/2}\| \leq \frac{\|X^{-1/2}S^{-1/2}\|}{1-\tau},$$

$$\|\widehat{X}^{1/2}\widehat{S}\widehat{X}^{1/2}\| \leq \frac{\|X^{1/2}SX^{1/2}\|}{(1-\tau)^2}, \qquad \|\widehat{X}^{-1/2}\widehat{S}^{-1}\widehat{X}^{-1/2}\| \leq \frac{\|X^{-1/2}S^{-1}X^{-1/2}\|}{(1-\tau)^2}.$$

Proof of Lemma 4.31. Lemma A2 implies that $\widehat{X}, \widehat{S} \in \mathcal{S}^n_{++}$, and the first inequality of the lemma is proved as follows:

$$\begin{aligned}\|\widehat{V}^2\| &= \lambda_{\max}[\widehat{V}^2] = \lambda_{\max}[\widehat{X}^{1/2}\widehat{S}\widehat{X}^{1/2}] = \|\widehat{X}^{1/2}\widehat{S}^{1/2}\|^2 \\ &\leq \frac{\|X^{1/2}S^{1/2}\|^2}{(1-\tau)^2} = \frac{\lambda_{\max}[X^{1/2}SX^{1/2}]}{(1-\tau)^2} = \frac{\lambda_{\max}[V^2]}{(1-\tau)^2} = \frac{\|V^2\|}{(1-\tau)^2},\end{aligned}$$

where the inequality follows from Lemma A2. The second inequality of the lemma follows in a similar manner using Lemma A2.

To prove the third inequality of the lemma, we bound both $\|\widehat{W}^{1/2}W^{-1/2}\|^4$ and $\|\widehat{W}^{-1/2}W^{1/2}\|^4$ by $1/(1-\tau)^2$. Using the definition of the matrix operator $\|\cdot\|$, the identities $S = WXW$ and $\widehat{S} = \widehat{W}\widehat{X}\widehat{W}$, Lemma A2, and the fact that, for all $A \in \mathcal{S}^n$ and for all nonsingular $P \in \Re^{n\times n}$, $\|A\| \leq \|PAP^{-1}\|$, we have

$$\begin{aligned}\|\widehat{W}^{1/2}W^{-1/2}\|^4 &= \|W^{-1/2}\widehat{W}W^{-1/2}\|^2 \leq \|X^{1/2}\widehat{W}W^{-1}X^{-1/2}\|^2 \\ &\leq \|X^{1/2}\widehat{X}^{-1/2}\|^2\|\widehat{X}^{1/2}\widehat{W}W^{-1}X^{-1/2}\| \\ &= \|X^{1/2}\widehat{X}^{-1/2}\|^2\,\lambda_{\max}[X^{-1/2}W^{-1}\widehat{W}\widehat{X}\widehat{W}W^{-1}X^{-1/2}] \\ &= \|X^{1/2}\widehat{X}^{-1/2}\|^2\,\lambda_{\max}[\widehat{W}\widehat{X}\widehat{W}W^{-1}X^{-1}W^{-1}] \\ &= \|X^{1/2}\widehat{X}^{-1/2}\|^2\,\lambda_{\max}[\widehat{S}S^{-1}] = \|X^{1/2}\widehat{X}^{-1/2}\|^2\,\lambda_{\max}[S^{-1/2}\widehat{S}S^{-1/2}] \\ &= \|X^{1/2}\widehat{X}^{-1/2}\|^2\|\widehat{S}^{1/2}S^{-1/2}\|^2 \leq \frac{1}{(1-\tau)^2}.\end{aligned}$$

A similar argument shows that $\|\widehat{W}^{-1/2}W^{1/2}\|^4 \leq 1/(1-\tau)^2$, and hence, the third inequality of the lemma follows. ∎