SIAM J. OPTIM.
Vol. 7, No. 1, pp. 86–125, February 1997

006

INTERIOR-POINT METHODS FOR THE MONOTONE SEMIDEFINITE LINEAR COMPLEMENTARITY PROBLEM IN SYMMETRIC MATRICES*

MASAKAZU KOJIMA†, SUSUMU SHINDOH‡, AND SHINJI HARA§

Abstract. The SDLCP (semidefinite linear complementarity problem) in symmetric matrices introduced in this paper provides a unified mathematical model for various problems arising from systems and control theory and combinatorial optimization. It is defined as the problem of finding a pair (X, Y) of $n \times n$ symmetric positive semidefinite matrices which lies in a given $n(n+1)/2$ dimensional affine subspace $\mathcal{F}$ of $\mathcal{S}^2$ and satisfies the complementarity condition $X \bullet Y = 0$, where $\mathcal{S}$ denotes the $n(n+1)/2$-dimensional linear space of symmetric matrices and $X \bullet Y$ the inner product of X and Y. The problem enjoys a close analogy with the LCP in the Euclidean space. In particular, the central trajectory leading to a solution of the problem exists under the nonemptiness of the interior of the feasible region and a monotonicity assumption on the affine subspace $\mathcal{F}$. The aim of this paper is to establish a theoretical basis of interior-point methods with the use of Newton directions toward the central trajectory for the monotone SDLCP.

Key words. interior-point method, linear complementarity problem, linear matrix inequality, semidefinite program, linear program

AMS subject classifications. 90C33, 90C05, 90C25

PII. S1052623494269035

1. Introduction.

Let $\hat{\mathcal{S}}$ denote the set of all $n \times n$ real matrices and $\mathcal{S}$ the set of all $n \times n$ symmetric real matrices. We identify $\hat{\mathcal{S}}$ with the n^2-dimensional Euclidean space $R^{n\times n}$ and $\mathcal{S}$ with an $n(n+1)/2$-dimensional linear subspace of $\hat{\mathcal{S}} = R^{n\times n}$. The *inner product* $\boldsymbol{X} \bullet \boldsymbol{Y}$ of $\boldsymbol{X}$ and $\boldsymbol{Y}$ in the linear space $\hat{\mathcal{S}}$ is Tr $\boldsymbol{X}^T\boldsymbol{Y}$, i.e., the trace of $\boldsymbol{X}^T\boldsymbol{Y}$. We write $\boldsymbol{X} \succ \boldsymbol{O}$ if $\boldsymbol{X} \in \hat{\mathcal{S}}$ is positive definite, i.e., $\boldsymbol{u}^T\boldsymbol{X}\boldsymbol{u} > 0$ for every nonzero $\boldsymbol{u} \in R^n$, and $\boldsymbol{X} \succeq \boldsymbol{O}$ if $\boldsymbol{X}$ is positive semidefinite, i.e., $\boldsymbol{u}^T\boldsymbol{X}\boldsymbol{u} \geq 0$ for every $\boldsymbol{u} \in R^n$. Here $\boldsymbol{O}$ stands for the $n \times n$ zero matrix. We use the symbol $\mathcal{S}_+$ for the set of symmetric positive semidefinite matrices and $\mathcal{S}_{++}$ for the set of symmetric positive definite matrices,

$$\mathcal{S}_+ = \{\boldsymbol{X} \in \mathcal{S} : \boldsymbol{X} \succeq \boldsymbol{O}\},\ \mathcal{S}_{++} = \{\boldsymbol{X} \in \mathcal{S} : \boldsymbol{X} \succ \boldsymbol{O}\}.$$

This paper introduces the SDLCP (the semidefinite linear complementarity problem) in symmetric matrices: find an $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2$ such that

$$(1) \qquad (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F},\ \boldsymbol{X} \succeq \boldsymbol{O},\ \boldsymbol{Y} \succeq \boldsymbol{O} \ \text{ and } \boldsymbol{X} \bullet \boldsymbol{Y} = 0.$$

Here $\mathcal{F}$ is an $n(n+1)/2$-dimensional affine subspace of $\mathcal{S}^2$. We call $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}$ with $\boldsymbol{X} \succeq \boldsymbol{O}$ and $\boldsymbol{Y} \succeq \boldsymbol{O}$ a *feasible solution* of the SDLCP (1) and $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}$ with

* Received by the editors June 3, 1994; accepted for publication (in revised form) September 1, 1995. This work was partially supported by the Okawa Institute of Information and Telecommunication.

http://www.siam.org/journals/siopt/7-1/26903.html

† Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152, Japan (kojima@is.titech.ac.jp).

‡ Department of Mathematics and Physics, The National Defense Academy, Hashirimizu 1-10-20, Yokosuka, Kanagawa, 239, Japan (shindoh@cc.nda.ac.jp).

§ Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Nagatsuta-cho, Midori-ku, Yokohama 227, Japan (hara@ssie.titech.ac.jp).

$\boldsymbol{X} \succ \boldsymbol{O}$ and $\boldsymbol{Y} \succ \boldsymbol{O}$ an *interior feasible solution* of the SDLCP (1). We impose a certain monotonicity condition (Condition 1.2 below) of the affine subspace $\mathcal{F}$. *The purpose of this paper is to establish a general theoretical framework of interior-point methods for the monotone SDLCP* (1).

The SDLCP is a generalization of SDPs (semidefinite programs) which have various applications in systems and control theory and combinatorial optimization. See [1, 2, 3, 5, 12, 15, 35, 44, 45], etc. Given $\boldsymbol{C} \in \mathcal{S}$, $\boldsymbol{A}_i \in \mathcal{S}$ $(i = 1, 2, \dots, m)$, and $b_i \in R$ $(i = 1, 2, \dots, m)$, a primal-dual pair of SDPs is defined as

$$
(2) \qquad \left\{ \begin{array}{lll}
\mathcal{P}: & \text{minimize} & \boldsymbol{C} \bullet \boldsymbol{X} \\
& \text{subject to} & \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \ (i = 1, 2, \dots, m), \\
& & \boldsymbol{X} \succeq \boldsymbol{O} \ (\text{ or } \boldsymbol{X} \in \mathcal{S}_+), \\
\mathcal{D}: & \text{maximize} & \sum_{i=1}^m b_i z_i \\
& \text{subject to} & \sum_{i=1}^m \boldsymbol{A}_i z_i + \boldsymbol{Y} = \boldsymbol{C}, \\
& & \boldsymbol{Y} \succeq \boldsymbol{O} \ (\text{ or } \boldsymbol{Y} \in \mathcal{S}_+).
\end{array} \right.
$$

We call $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{z}) \in \mathcal{S}^2 \times R^m$ a *feasible solution* of the primal-dual pair (2) of SDPs if it satisfies all the constraints in (2) and an *interior feasible solution* of (2) if it satisfies $\boldsymbol{X} \succ \boldsymbol{O}$ and $\boldsymbol{Y} \succ \boldsymbol{O}$ in addition to the constraints. Let

$$
(3) \qquad \mathcal{F} = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2 \ : \ \begin{array}{l} \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \ (i = 1, 2, \dots, m), \\ \sum_{i=1}^m \boldsymbol{A}_i z_i + \boldsymbol{Y} = \boldsymbol{C} \ \text{ for some } \boldsymbol{z} \in R^m \end{array} \right\}.
$$

Then $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{z}) \in \mathcal{S}^2 \times R^m$ is a feasible solution (or an interior feasible solution) of the primal-dual pair (2) of SDPs if and only if $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2$ is a feasible solution (or an interior feasible solution) of the SDLCP (1). Furthermore, if we assume that there is an interior feasible solution of (2), we can state a common necessary and sufficient optimality condition for $\boldsymbol{X}$ to be a minimum solution of the primal problem $\mathcal{P}$ and $(\boldsymbol{Y}, \boldsymbol{z})$ to be a maximum solution of the dual problem $\mathcal{D}$ in terms of the SDLCP (1). In this case, the $n(n+1)/2$-dimensional affine subspace $\mathcal{F}$ enjoys the self-orthogonality, i.e.,

$$
(\boldsymbol{X}' - \boldsymbol{X}) \bullet (\boldsymbol{Y}' - \boldsymbol{Y}) = 0 \ \text{ for every } (\boldsymbol{X}', \boldsymbol{Y}'), \ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F},
$$

which is a special case of the monotonicity (see Condition 1.2 below). Therefore the monotone SDLCP (1) is at least as general as the primal-dual pair (2) of SDPs, and we can specialize and/or modify both theoretical results and interior-point methods presented in this paper to adapt them to the primal-dual pair (2) of SDPs. However, our primary concern is not an extension of the primal-dual pair of SDPs but rather a basic idea for designing a wide class of interior-point methods for mathematical programs in the space of symmetric matrices.

A distinctive and important feature of our theoretical framework of interior-point methods for the monotone SDLCP (1) is the use of “a Newton direction for approximating a point on the central trajectory” at each iteration. This feature enables us to transfer many useful technologies developed in the class of primal-dual interior-point methods for LPs (linear programs) ([9, 20, 25, 26, 27, 30, 33, 34, 40, 41], etc.) and their extensions to LCPs (linear complementarity problems) ([19, 21, 22], etc.) and horizontal LCPs ([47, 48, 49], etc.). Indeed the Generic Interior-Point Method presented in section 5 opens up the possibilities of extensions of a great variety of primal-dual interior-point methods developed so far — central trajectory following methods, potential-reduction methods, predictor-corrector methods, infeasible interior-point methods, etc.—to the monotone SDLCP (1).

In recent years, many studies ([1, 5, 45, 15, 35, 44], etc.) have been done on extensions of interior-point methods developed for LPs to SDPs. Our primal-dual interior-point methods for the SDLCP (1) are built on the same materials, the logarithmic barrier function, the central trajectory, the potential function, etc., as those used in the existing ones. In particular, we follow "a recipe" proposed by Alizadeh [2] to extend known interior-point algorithms for LPs into similar algorithms for SDPs using those materials (see Figures 2.1 and 3.2 of [2]). Alizadeh gave "a direct extension of Ye's projective potential-reduction method" [46] based on his recipe. It should be emphasized, however, that our extension of primal-dual interior-point methods to the SDLCP (1) is not so direct as in the case of Ye's projective potential-reduction method. There is a brief discussion of the difficulty below.

We use the notation $\boldsymbol{a}\bullet\boldsymbol{b}$ to denote the inner product $\sum_{j=1}^n a_j b_j$ of every $\boldsymbol{a},\ \boldsymbol{b} \in R^n$ and the notation diag $\boldsymbol{a}$ to denote the $n\times n$ diagonal matrix with the diagonal elements $a_1, a_2, \ldots, a_n$ for every $\boldsymbol{a} = (a_1, a_2, \ldots, a_n) \in R^n$. Let $\boldsymbol{c} \in R^n$, $\boldsymbol{a}_i \in R^n$ $(i = 1, 2, \ldots, m)$. Consider the primal-dual pair of LPs:

$$
\left\{
\begin{array}{lll}
\mathcal{P}: & \text{minimize} & \boldsymbol{c}\bullet\boldsymbol{x} \\
 & \quad\text{subject to} & \boldsymbol{a}_i\bullet\boldsymbol{x} = b_i\ (i = 1, 2, \ldots, m), \\
 & & \boldsymbol{x} \geq \boldsymbol{0}, \\
\mathcal{D}: & \text{maximize} & \sum_{i=1}^m b_i z_i \\
 & \quad\text{subject to} & \sum_{i=1}^m \boldsymbol{a}_i z_i + \boldsymbol{y} = \boldsymbol{c}, \\
 & & \boldsymbol{y} \geq \boldsymbol{0}.
\end{array}
\right. \tag{4}
$$

By taking

$$
\begin{array}{l}
\boldsymbol{C} = \text{diag}\ \boldsymbol{c},\ \ \boldsymbol{A}_i = \text{diag}\ \boldsymbol{a}_i\ (i = 1, 2, \ldots, m),\ \ \boldsymbol{O} = \ \text{diag}\ \boldsymbol{0}, \\
\boldsymbol{X} = \text{diag}\ \boldsymbol{x},\ \ \text{and}\ \boldsymbol{Y} = \ \text{diag}\ \boldsymbol{y},
\end{array}
$$

we embed the primal-dual LPs (4) into the primal-dual SDPs (2). This convenience makes it possible for us to simultaneously present one iteration of primal-dual interior-point methods for LPs and SDPs. Suppose that we know an interior feasible solution $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{z}) \in \mathcal{S}^2 \times R^m$ of the primal-dual pair (2) of SDPs. Alizadeh's recipe leads us to the Newton equation

$$
\left\{
\begin{array}{l}
d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{X}d\boldsymbol{Y} = \mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}, \\
\boldsymbol{A}_i\bullet d\boldsymbol{X} = 0\ (i = 1, 2, \ldots, m), \\
\sum_{i=1}^m \boldsymbol{A}_i dz_i + d\boldsymbol{Y} = \boldsymbol{O}
\end{array}
\right. \tag{5}
$$

for a search direction $(d\boldsymbol{X}, d\boldsymbol{Y}, d\boldsymbol{z}) \in \mathcal{S}^2 \times R^m$, where $\mu > 0$ denotes a search direction parameter. Then we generate a new point $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}, \bar{\boldsymbol{z}}) \in \mathcal{S}^2 \times R^m$ with appropriate step lengths $\alpha_p > 0$ and $\alpha_d > 0$ such that

$$
\begin{array}{c}
\bar{\boldsymbol{X}} = \boldsymbol{X} + \alpha_p\ d\boldsymbol{X}, \\
(\bar{\boldsymbol{Y}}, \bar{\boldsymbol{z}}) = (\boldsymbol{Y}, \boldsymbol{z}) + \alpha_d\ (d\boldsymbol{Y}, d\boldsymbol{z}).
\end{array}
$$

When we are concerned with the primal-dual pair (4) of LPs, all the matrices $\boldsymbol{A}_i$, $\boldsymbol{X}$, $\boldsymbol{Y}$, $d\boldsymbol{X}$, $d\boldsymbol{Y}$ appearing in the Newton equation (5) are diagonal; hence they are commutative. In this case, we can transform (5) into the system of equations

$$
\left\{
\begin{array}{l}
\boldsymbol{D}^{-1}d\boldsymbol{X} + \boldsymbol{D}d\boldsymbol{Y} = \mu(\boldsymbol{X}\boldsymbol{Y})^{-1/2} - (\boldsymbol{X}\boldsymbol{Y})^{1/2}, \\
\boldsymbol{D}\boldsymbol{A}_i\bullet\boldsymbol{D}^{-1}d\boldsymbol{X} = 0\ (i = 1, 2, \ldots, m), \\
\sum_{i=1}^m \boldsymbol{D}\boldsymbol{A}_i dz_i + \boldsymbol{D}d\boldsymbol{Y} = \boldsymbol{O}
\end{array}
\right. \tag{6}
$$

in a "scaled Newton direction" $(\boldsymbol{D}^{-1}d\boldsymbol{X}, \boldsymbol{D}d\boldsymbol{Y}, d\boldsymbol{z})$, where $\boldsymbol{D} = \boldsymbol{X}^{1/2}\boldsymbol{Y}^{-1/2}$. We can easily verify that the scaled Newton direction $(\boldsymbol{D}^{-1}d\boldsymbol{X}, \boldsymbol{D}d\boldsymbol{Y}, d\boldsymbol{z})$ satisfies

$$(7)\quad (\boldsymbol{D}^{-1}d\boldsymbol{X}) + (\boldsymbol{D}d\boldsymbol{Y}) = \mu(\boldsymbol{X}\boldsymbol{Y})^{-1/2} - (\boldsymbol{X}\boldsymbol{Y})^{1/2} \quad \text{and } (\boldsymbol{D}^{-1}d\boldsymbol{X}) \bullet (\boldsymbol{D}d\boldsymbol{Y}) = 0.$$

This relation has been playing a crucial role in the development of primal-dual interior-point algorithms for LPs. See [20, Section 2].

In the case of SDPs, we can derive from the Newton equation (5) a similar system of equations,

$$(6)' \qquad \begin{cases} \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{Y}} + \sqrt{\boldsymbol{X}} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1} = \mu \sqrt{\boldsymbol{X}}^{-1} \sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}, \\ \sqrt{\boldsymbol{X}}\boldsymbol{A}_i \sqrt{\boldsymbol{Y}}^{-1} \bullet \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{Y}} = 0 \ (i = 1, 2, \ldots, m), \\ \sum_{i=1}^{m} \sqrt{\boldsymbol{X}}\boldsymbol{A}_i \sqrt{\boldsymbol{Y}}^{-1} dz_i + \sqrt{\boldsymbol{X}} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1} = \boldsymbol{O}, \end{cases}$$

in a "scaled Newton direction" $(\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{Y}}, \sqrt{\boldsymbol{X}} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1}, d\boldsymbol{z})$ and a similar relation,

$$(7)' \qquad \begin{cases} \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{Y}} + \sqrt{\boldsymbol{X}} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1} = \mu \sqrt{\boldsymbol{X}}^{-1} \sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}, \\ \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{Y}} \bullet \sqrt{\boldsymbol{X}} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1} = 0. \end{cases}$$

But (6) and (7) do not hold any more because the matrices $\boldsymbol{X}$, $\boldsymbol{Y}$, $d\boldsymbol{X}$, and $d\boldsymbol{Y}$ appearing in the Newton equation (5) are $n \times n$ general symmetric matrices whose multiplication is not necessarily commutative, and the derivation of the scaled Newton equation (6) essentially relies on the commutativity of these matrices. This noncommutativity of general symmetric matrices certainly causes some difficulty in straightforward extensions of primal-dual interior-point methods to the SDP (2). What is worse and more substantial in the case of SDPs, however, is that the Newton equation (5) does not necessarily have a symmetric solution (i.e., a solution $(d\boldsymbol{X}, d\boldsymbol{Y}, d\boldsymbol{z})$ with symmetric $d\boldsymbol{X}$ and $d\boldsymbol{Y}$). (This fact was also pointed out in the recent paper [3] by Alizadeh, Haeberly, and Overton. They proposed some variants of the Newton equation which are different from (5) to get a symmetric search direction.) Therefore, following Alizadeh's recipe is not enough to generalize primal-dual interior-point methods from LPs to SDPs. In this paper, we will devise "a new system of equations" in a modified Newton direction towards the central trajectory and establish some fundamental results (including a system of equations similar to (6)' and a relation similar to (7)'; see Corollary 4.3) which are necessary to analyze the convergence of primal-dual interior-point methods using the modified Newton direction.

The SDLCP (1) presents an extraordinary similarity to the LCP in the Euclidean space: find an $(\boldsymbol{x}, \boldsymbol{y}) \in R^{2n}$ such that

$$(8) \qquad \boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q},\ \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0}, \quad \text{and } \boldsymbol{x} \bullet \boldsymbol{y} = 0,$$

where $\boldsymbol{M} \in \hat{\mathcal{S}}$ is a given constant matrix and $\boldsymbol{q} \in R^n$ a given constant vector. Letting F be the n-dimensional affine subspace $\{(\boldsymbol{x}, \boldsymbol{y}) \in R^{2n} : \boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q}\}$, we can rewrite (8) as

$$(9) \qquad (\boldsymbol{x}, \boldsymbol{y}) \in F,\ \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0}, \quad \text{and } \boldsymbol{x} \bullet \boldsymbol{y} = 0.$$

If we allow F to be a general affine subspace of R^{2n}, the LCP (9) is equivalent to the so-called horizontal linear complementarity problem (see, for example, [4, 6, 31, 43, 48]).

Thus we have the clear correspondence between the SDLCP (1) and the horizontal LCP (9) in the Euclidean space:

$$(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} \subset \mathcal{S}^2 \Longleftrightarrow (\boldsymbol{x}, \boldsymbol{y}) \in F \subset R^{2n},$$
$$\text{“}\succ,\ \succeq\text{”} \Longleftrightarrow \text{“}>,\ \geq\text{”},$$
$$\boldsymbol{X} \bullet \boldsymbol{Y} \Longleftrightarrow \boldsymbol{x} \bullet \boldsymbol{y}.$$

(See also Figures 2.1 and 3.2 of [2].)

It is interesting to note that the sets $\mathcal{S}_+ = \{\boldsymbol{X} \in \mathcal{S} : \boldsymbol{X} \succeq \boldsymbol{O}\}$ and $\mathcal{S}_{++} = \{\boldsymbol{X} \in \mathcal{S} : \boldsymbol{X} \succ \boldsymbol{O}\}$ play the roles of the nonnegative orthant R^n_+ and the positive orthant R^n_{++}, respectively, in the space $\mathcal{S}$ of symmetric matrices. We have the following properties.

LEMMA 1.1.

1. *$\mathcal{S}_+$ is a closed convex cone in $\mathcal{S}$ and its interior coincides with $\mathcal{S}_{++}$.*
2. *$\{\boldsymbol{Y} \in \mathcal{S} : \boldsymbol{X} \bullet \boldsymbol{Y} \geq 0$ for every $\boldsymbol{X} \in \mathcal{S}_+\} = \mathcal{S}_+$ (self-polarity).*
3. *Let $\boldsymbol{X},\ \boldsymbol{Y} \in \mathcal{S}_+$. Then $\boldsymbol{X} \bullet \boldsymbol{Y} \geq 0$, and $\boldsymbol{X} \bullet \boldsymbol{Y} = 0$ if and only if $\boldsymbol{XY} = \boldsymbol{O}$ (Lemma 2.3 of [1]).*
4. *Suppose that $\boldsymbol{A} \in \mathcal{S}_{++}$ and $\alpha > 0$. Let λ_{min} be the minimum eigenvalue of $\boldsymbol{A}$. If $\boldsymbol{X} \in \mathcal{S}_+$ and $\boldsymbol{A} \bullet \boldsymbol{X} \leq \alpha$ then the sum of all eigenvalues of $\boldsymbol{X}$ is not greater than α/λ_{min}; hence the set $\{\boldsymbol{X} \in \mathcal{S}_+ : \boldsymbol{A} \bullet \boldsymbol{X} \leq \alpha\}$ is bounded.*

The properties 1, 2, and 4 are easily verified, and their proofs are omitted here. See [13] for further properties of $\mathcal{S}_+$. In view of property 3 of Lemma 1.1, we can rewrite the SDLCP (1) as $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}, (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_+$, and $\boldsymbol{XY} = \boldsymbol{O}$.

Among many kinds of assumptions on the LCP (8) and the horizontal LCP (9), the monotonicity assumption

$$(\boldsymbol{x}' - \boldsymbol{x}) \bullet (\boldsymbol{y}' - \boldsymbol{y}) \geq 0 \quad \text{for every } (\boldsymbol{x}', \boldsymbol{y}') \text{ and } (\boldsymbol{x}, \boldsymbol{y}) \in F$$

is the most popular one. Indeed, the monotone LCP and the monotone horizontal LCP have important applications to LPs and convex quadratic programs. We impose a similar assumption on the SDLCP (1) throughout the paper.

Condition 1.2. The $n(n+1)/2$-dimensional affine subspace $\mathcal{F}$ associated with the SDLCP (1) is monotone, i.e., $(\boldsymbol{X}' - \boldsymbol{X}) \bullet (\boldsymbol{Y}' - \boldsymbol{Y}) \geq 0$ for every $(\boldsymbol{X}', \boldsymbol{Y}')$ and $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}$.

Suppose that F is an n-dimensional monotone affine subspace of R^{2n} and that the horizontal LCP (9) has an interior feasible solution, i.e., $(\boldsymbol{x}^0, \boldsymbol{y}^0) \in F$ such that $(\boldsymbol{x}^0, \boldsymbol{y}^0) > \boldsymbol{0}$. It is well known that for every $\mu > 0$ there exists a unique interior feasible solution $(\boldsymbol{x}(\mu), \boldsymbol{y}(\mu))$ satisfying $x_j(\mu)y_j(\mu) = \mu$ $(j = 1, 2, \ldots, n)$ and that the set $C = \{(\boldsymbol{x}(\mu), \boldsymbol{y}(\mu)) : \mu > 0\}$ forms a smooth trajectory converging to a solution of the horizontal LCP (9) as $\mu \to 0$. The set C is called the central trajectory or the path of centers.

Remark 1.3. Megiddo [30] presented the result above in connection with interior-point methods for the monotone LCP (8) with $F = \{(\boldsymbol{x}, \boldsymbol{y}) \in R^{2n} : \boldsymbol{y} = \boldsymbol{Mx} + \boldsymbol{q}\}$. It was shown recently that $\dim F \leq n$ for any monotone affine subspace F of R^{2n} and that any monotone horizontal LCP with F of dimension n is reducible to a positive semidefinite LCP (8) (see [4, 11, 39, 43]). Hence the result above is equivalent to the one by Megiddo [30] on the monotone LCP (8) with $F = \{(\boldsymbol{x}, \boldsymbol{y}) \in R^{2n} : \boldsymbol{y} = \boldsymbol{Mx} + \boldsymbol{q}\}$. See also [19, 28] for the existence of the central trajectory for more general complementarity problems.

The central trajectory has provided us with a theoretical basis for a wide class of interior-point methods which originated from a primal-dual interior-point method

([20, 30, 34, 40], etc.) for LPs and later extended to the monotone LCP (8) ([19, 21, 22], etc.) and the monotone horizontal LCP ([48, 49], etc.). A common feature of methods in this class is to move in "a Newton direction for approximating a point on the central trajectory" at each iteration.

It is well known that convex quadratic programs in the Euclidean space can be transformed into the LCP (8) or the horizontal LCP (9) via the Karush–Kuhn–Tucker optimality condition. As an extension of the primal SDP $\mathcal{P}$ stated above, consider a convex quadratic program of the form

$$\text{QP} : \text{minimize } \boldsymbol{C} \bullet \boldsymbol{X} + \boldsymbol{X} \bullet (\boldsymbol{Q}\boldsymbol{X}) \text{ subject to } \boldsymbol{X} \in \mathcal{S}_+ \cap (\mathcal{L} + \boldsymbol{D}).$$

Here $\boldsymbol{Q} \in \mathcal{S}_+$ is a given matrix and $\mathcal{L}$ a given linear subspace of $\mathcal{S}$. Then it is easily verified that $\boldsymbol{X} \in \mathcal{S}$ is a minimum solution of the quadratic program if the conditions

$$\begin{aligned} &\boldsymbol{X} \in \mathcal{S}_+ \cap (\mathcal{L} + \boldsymbol{D}),\ \boldsymbol{Y} - (\boldsymbol{Q}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{Q}) \in \left(\mathcal{L}^\perp + \boldsymbol{C}\right), \\ &\boldsymbol{Y} \in \mathcal{S}_+, \quad \text{and } \boldsymbol{X} \bullet \boldsymbol{Y} = 0 \end{aligned}$$

hold for some $\boldsymbol{Y} \in \mathcal{S}_+$, where $\mathcal{L}^\perp$ is the orthogonal complement of $\mathcal{L}$. We can rewrite these conditions as the SDLCP (1) with the affine subspace

$$\mathcal{F} = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2 : \boldsymbol{Y} - (\boldsymbol{Q}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{Q}) \in \left(\mathcal{L}^\perp + \boldsymbol{C}\right),\ \boldsymbol{X} \in (\mathcal{L} + \boldsymbol{D})\}.$$

It is easily verified that $\mathcal{F}$ is an $n(n+1)/2$-dimensional monotone affine subspace. Thus we can apply the generic IP method described in section 5 to the convex QP. It should be noted, however, that the convex QP is equivalent to an SDP of the form

$$\begin{array}{lll} \text{SDP:} & \text{minimize} & \boldsymbol{C} \bullet \boldsymbol{X} + \boldsymbol{I} \bullet \boldsymbol{Z} \\ & \text{subject to} & \boldsymbol{X} \in \mathcal{S}_+ \cap (\mathcal{L} + \boldsymbol{D}), \\ & & \begin{pmatrix} \boldsymbol{I} & \boldsymbol{L}^T\boldsymbol{X} \\ \boldsymbol{X}\boldsymbol{L} & \boldsymbol{Z} \end{pmatrix} \succeq \boldsymbol{O}. \end{array}$$

Here $\boldsymbol{L}$ denotes an $n \times n$ matrix such that $\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{L}^T$. This fact itself never denies the significance of the monotone SDLCP because the direct SDLCP formulation is of a smaller size than the SDP formulation but raises questions like how general the monotone SDLCP is and whether it is essentially different from the SDP. In their recent paper [24], Kojima, Shida, and Shindoh showed that the monotone SDLCP (1) is reducible to an SDP involving an additional m-dimensional variable vector and an $(m+1) \times (m+1)$ variable symmetric matrix, where $m = n(n+1)/2$.

In section 2, we list notation and symbols that are used throughout the paper. Sections 3 through 8 are devoted to our main results:

- The existence of the central trajectory (section 3).
- The existence of modified Newton directions towards the central trajectory (section 4).
- A generic interior-point method (section 5).
- Some properties of the solution set of the monotone SDLCP (1) (section 6).
- Basic lemmas necessary to analyze the computational complexity of interior-point methods for the monotone SDLCP (1) (section 7).
- A central trajectory following method which is an extension of the algorithm given by Kojima–Mizuno–Yoshise [21] for the monotone LCP (8) in the Euclidean space to the monotone SDLCP (1) (section 8.1).

- A potential-reduction method based on the algorithm given by Kojima–Mizuno–Yoshise [22] for the monotone LCP (8) to the monotone SDLCP (1) (section 8.2).
- An infeasible interior-point potential-reduction method based on the constrained potential reduction algorithm given by Mizuno–Kojima–Todd [32, Algorithm I] for LPs to the monotone SDLCP (1) (section 8.3).

2. Notation and symbols.

R^m : the m-dimensional Euclidean space.

$\hat{\mathcal{S}} = R^{n\times n}$, the set of all $n \times n$ matrices.

$\mathcal{S}$: the $n(n+1)/2$-dimensional linear subspace of $\hat{\mathcal{S}}$ consisting of all $n \times n$ symmetric matrices.

$\tilde{\mathcal{S}}$: the $n(n-1)/2$-dimensional linear subspace of $\hat{\mathcal{S}}$ consisting of all $n \times n$ skew-symmetric matrices.

$\mathcal{S}_+ = \{\boldsymbol{X} \in \mathcal{S} : \boldsymbol{X} \succeq \boldsymbol{O}\}$.

$\mathcal{S}_{++} = \{\boldsymbol{X} \in \mathcal{S} : \boldsymbol{X} \succ \boldsymbol{O}\}$.

$\hat{\mathcal{S}}_{++} = \{\boldsymbol{X} \in \hat{\mathcal{S}} : \boldsymbol{X} \succ \boldsymbol{O}\}$.

$\boldsymbol{I}$, $\boldsymbol{O}$: the $n \times n$ identity matrix, the $n \times n$ zero matrix, respectively.

Tr $\boldsymbol{X}$: the trace of $\boldsymbol{X} \in \hat{\mathcal{S}}$.

$\boldsymbol{X} \bullet \boldsymbol{Y} = \mathrm{Tr}\ \boldsymbol{X}^T\boldsymbol{Y}$ for $\boldsymbol{X}, \boldsymbol{Y} \in \hat{\mathcal{S}}$ (the inner product of $\boldsymbol{X}$ and $\boldsymbol{Y}$).

$\|\boldsymbol{X}\|_F = (\boldsymbol{X} \bullet \boldsymbol{X})^{1/2}$ (the Frobenius norm of $\boldsymbol{X} \in \hat{\mathcal{S}}$).

$$\mathcal{F}^0 = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2 : \begin{array}{l} (\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^{n(n+1)/2} c_i(\boldsymbol{M}^i, \boldsymbol{N}^i) \\ \text{for some } c_i \in R\ (i = 1, 2, \ldots, n(n+1)/2) \end{array} \right\},$$

where $(\boldsymbol{M}^i, \boldsymbol{N}^i) \in \mathcal{S}^2$ $(i = 1, 2, \ldots, n(n+1)/2)$ are linearly independent.

$\mathcal{F} = (\boldsymbol{X}^0, \boldsymbol{Y}^0) + \mathcal{F}^0$ for some $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}$ (an $n(n+1)/2$-dimensional affine subspace associated with the SDLCP (1)).

$\mathcal{F}_+ = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} : \boldsymbol{X} \succeq \boldsymbol{O},\ \boldsymbol{Y} \succeq \boldsymbol{O}\}$ (the set of feasible solutions of the SDLCP (1)).

$\mathcal{F}_{++} = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} : \boldsymbol{X} \succ \boldsymbol{O},\ \boldsymbol{Y} \succ \boldsymbol{O}\}$ (the set of interior feasible solutions of the SDLCP (1)).

$\mathcal{F}^* = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_+ : \boldsymbol{X} \bullet \boldsymbol{Y} = 0\}$ (the set of solutions of the SDLCP (1)).

$$\tilde{\mathcal{F}}^0 = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \hat{\mathcal{S}}^2 : \begin{array}{l} (\boldsymbol{X}, \boldsymbol{Y}) = \sum_{j=1}^{n(n-1)/2} \tilde{c}_j(\tilde{\boldsymbol{M}}^j, \tilde{\boldsymbol{N}}^j) \\ \text{for some } \tilde{c}_j \in R\ (j = 1, 2, \ldots, n(n-1)/2) \end{array} \right\},$$

where $(\tilde{\boldsymbol{M}}^j, \tilde{\boldsymbol{N}}^j) \in \tilde{\mathcal{S}}^2$ $(j = 1, 2, \ldots, n(n-1)/2)$ are linearly independent (an $n(n-1)/2$-dimensional linear subspace of $\tilde{\mathcal{S}}^2$).

$\phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{X} \bullet \boldsymbol{Y} - \mu \log\det \boldsymbol{X}\boldsymbol{Y}$ for every $(\mu, \boldsymbol{X}, \boldsymbol{Y}) \in R_{++} \times \mathcal{S}^2_{++}$ (the logarithmic barrier function).

$\lambda_1, \lambda_2, \ldots, \lambda_n$: the eigenvalues of $\boldsymbol{X}\boldsymbol{Y}$, where $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$. (Note that all λ_i's are positive. See below.)

$$
\begin{aligned}
\boldsymbol{\Lambda} &= \text{ diag } (\lambda_1, \lambda_2, \ldots, \lambda_n). \\
\lambda_{min} &= \min\{\lambda_1, \lambda_2, \ldots, \lambda_n\}. \\
\lambda_{max} &= \max\{\lambda_1, \lambda_2, \ldots, \lambda_n\}. \\
\boldsymbol{H}(\beta) &= \beta\mu\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{Y}} \in \hat{\mathcal{S}}, \quad \text{where } \beta \geq 0 \text{ and } \mu > 0. \\
\mathcal{N}(\gamma) &= \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : \quad \left(\sum_{j=1}^{n} (\lambda_j - \mu)^2 \right)^{1/2} \leq \gamma\mu, \quad \text{where } \mu = \frac{\boldsymbol{X} \bullet \boldsymbol{Y}}{n} \right\}, \\
& \qquad \text{where } \gamma > 0 \text{ (a horn neighborhood of the central trajectory).} \\
f(\boldsymbol{X}, \boldsymbol{Y}) &= (n + \nu) \log \boldsymbol{X} \bullet \boldsymbol{Y} - \log \det \boldsymbol{XY} - n \log n \quad \text{for every } (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}, \\
& \qquad \text{where } \nu \geq 0 \text{ is a parameter (the potential function).}
\end{aligned}
$$

Let $\boldsymbol{X} \in \mathcal{S}_+$. Then we can find a symmetric matrix $\boldsymbol{B}$ such that $\boldsymbol{X} = \boldsymbol{BB}$. Note that such a matrix $\boldsymbol{B}$ is uniquely determined and is positive semidefinite. We denote such a matrix $\boldsymbol{B}$ by $\sqrt{\boldsymbol{X}}$ throughout the paper:

$$\sqrt{\boldsymbol{X}} : \text{ the matrix in } \mathcal{S}_+ \text{ uniquely determined by } \boldsymbol{X} = \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{X}} \text{ for } \boldsymbol{X} \in \mathcal{S}_+.$$

By the definition, we see that

$$
\begin{aligned}
&\text{Tr } \boldsymbol{A} = \sum_{j=1}^{n} \alpha_j, \quad \text{where } \alpha_1, \alpha_2, \ldots, \alpha_n \text{ denote the eigenvalues of a matrix } \boldsymbol{A} \in \hat{\mathcal{S}}, \\
&\text{Tr } \boldsymbol{A} = \text{Tr } \boldsymbol{B}^{-1}\boldsymbol{AB} \quad \text{for every } \boldsymbol{A} \in \hat{\mathcal{S}} \quad \text{and every nonsingular } \boldsymbol{B} \in \hat{\mathcal{S}}, \\
&\boldsymbol{M} \bullet \boldsymbol{X} = \boldsymbol{M}^T \bullet \boldsymbol{X} \quad \text{if } \boldsymbol{M} \in \mathcal{S} \quad \text{or } \boldsymbol{X} \in \mathcal{S}.
\end{aligned}
$$

We will often use these relations throughout the paper.

The following fact is also utilized often: If $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$, all the eigenvalues of $\boldsymbol{XY}$ are real and positive. This is because $\boldsymbol{XY}$ has the same eigenvalues as the symmetric positive definite matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}$.

3. The central trajectory. Let $\mathcal{F}_+$, $\mathcal{F}_{++}$, and $\mathcal{F}^*$ denote the set of feasible solutions, the set of interior feasible solutions, and the set of solutions, of the SDLCP (1), respectively:

$$
\begin{aligned}
\mathcal{F}_+ &= \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} : \boldsymbol{X} \succeq \boldsymbol{O}, \ \boldsymbol{Y} \succeq \boldsymbol{O}\}, \\
\mathcal{F}_{++} &= \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} : \boldsymbol{X} \succ \boldsymbol{O}, \ \boldsymbol{Y} \succ \boldsymbol{O}\}, \\
\mathcal{F}^* &= \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_+ : \boldsymbol{X} \bullet \boldsymbol{Y} = 0\}.
\end{aligned}
$$

THEOREM 3.1. *Suppose that the SDLCP* (1) *has an interior feasible solution, i.e.,* $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}_{++}$.

1. *For every* $\mu > 0$*, there exists a unique* $(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu)) \in \mathcal{F}_{++}$ *such that* $\boldsymbol{X}(\mu)\boldsymbol{Y}(\mu) = \mu\boldsymbol{I}$*, where* $\boldsymbol{I}$ *denotes the* $n \times n$ *identity matrix.*
2. $(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu))$ *is the unique minimizer of the logarithmic barrier function*

$$\phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{X} \bullet \boldsymbol{Y} - \mu \log \det \boldsymbol{XY} \ \text{ over } \mathcal{F}_{++}.$$

3. *The set* $\mathcal{C} = \{(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu)) : \mu > 0\}$ *forms a smooth trajectory. (We call* $\mathcal{C}$ *the central trajectory.)*
4. $(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu))$ *converges to a solution of the SDLCP* (1), $(\boldsymbol{X}^*, \boldsymbol{Y}^*) \in \mathcal{F}^*$ *as* $\mu > 0$ *tends to zero.*

The existence of the central trajectory is known if we restrict ourselves to the primal-dual pair (2) of SDPs $\mathcal{P}$ and $\mathcal{D}$, where $\mathcal{F}$ is given as in (3); see [5, 44, 45].

Besides item 2 of Theorem 3.1, there are some other characterizations of the central trajectory $\mathcal{C}$. One is the following: an $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++}$ lies on the central trajectory $\mathcal{C}$ if and only if all eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $\boldsymbol{XY}$ have a common value $\mu > 0$.

We give some remarks on relations of the SDLCP (1) in symmetric matrices with the horizontal LCP (9) (or the LCP (8) with $F = \{(\boldsymbol{x}, \boldsymbol{y}) \in R^{2n} : \boldsymbol{y} = \boldsymbol{Mx} + \boldsymbol{q}\}$) in the Euclidean space. Each eigenvalue λ_j of the product $\boldsymbol{XY}$ of a pair of matrices $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$ plays the role of the product $x_j y_j$ of a complementary pair of variables x_j and y_j in $(\boldsymbol{x}, \boldsymbol{y}) \in R^{2n}_{++}$. We have seen above that the central trajectory $\mathcal{C}$ can be rewritten as

$$\mathcal{C} = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : \begin{array}{c} \lambda_j = \mu > 0 \ (j = 1, 2, \ldots, n) \ \text{ for some } \mu > 0, \\ \text{where } \lambda_1, \lambda_2, \ldots, \lambda_n \ \text{ are the eigenvalues of } \boldsymbol{XY} \end{array} \right\}.$$

We also see that the logarithmic barrier function ϕ and the potential function f,

$$\begin{aligned} \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) &= \boldsymbol{X} \bullet \boldsymbol{Y} - \mu \log\det \boldsymbol{XY}, \\ f(\boldsymbol{X}, \boldsymbol{Y}) &= (n + \nu) \log \boldsymbol{X} \bullet \boldsymbol{Y} - \log\det \boldsymbol{XY} - n \log n, \end{aligned}$$

which we will utilize in section 8, can be rewritten in terms of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $\boldsymbol{XY}$ as follows:

$$\begin{aligned} \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) &= \sum_{j=1}^{n} \lambda_j - \mu \sum_{j=1}^{n} \log \lambda_j, \\ f(\boldsymbol{X}, \boldsymbol{Y}) &= (n + \nu) \log \left(\sum_{j=1}^{n} \lambda_j \right) - \sum_{j=1}^{n} \log \lambda_j - n \log n. \end{aligned}$$

If we replace λ_j by $x_j y_j$, we have the central trajectory, the logarithmic barrier function, and the potential function that have been used widely in interior-point methods for the LCP (8) in the Euclidean space. See also [7] and Figures 2.1 and 3.2 of [2].

We give another characterization of the central trajectory $\mathcal{C}$ in terms of the potential function in section 8.2 where we present a potential-reduction method. See (60).

The remainder of this section is devoted to a proof of Theorem 3.1. Let $\|\boldsymbol{X}\|_F$ denote the Frobenius norm of a matrix $\boldsymbol{X} \in \hat{\mathcal{S}}$; $\|\boldsymbol{X}\|_F^2 = \boldsymbol{X} \bullet \boldsymbol{X}$. Let

$$\mathcal{F}^0 = \{(\boldsymbol{X}' - \boldsymbol{X}, \boldsymbol{Y}' - \boldsymbol{Y}) : (\boldsymbol{X}, \boldsymbol{Y}), (\boldsymbol{X}', \boldsymbol{Y}') \in \mathcal{F}\}.$$

Then $\mathcal{F}^0$ forms an $n(n+1)/2$-dimensional linear subspace of $\mathcal{S}^2$. Let $p = n(n+1)/2$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}$ and let $(\boldsymbol{M}^i, \boldsymbol{N}^i) (i = 1, 2, \ldots, p)$ be a basis of $\mathcal{F}^0$. Then

$$\mathcal{F} = (\boldsymbol{X}^0, \boldsymbol{Y}^0) + \mathcal{F}^0, \tag{10}$$

$$\mathcal{F}^0 = \left\{ (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{S}^2 : \begin{array}{l} (d\boldsymbol{X}, d\boldsymbol{Y}) = \sum_{i=1}^{p} c_i (\boldsymbol{M}^i, \boldsymbol{N}^i) \\ \text{for some } c_i \in R \ (i = 1, 2, \ldots, p) \end{array} \right\}. \tag{11}$$

We also note that $\mathcal{F}^0$ is monotone, i.e., $d\boldsymbol{X} \bullet d\boldsymbol{Y} \geq 0$ for every $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0$.

We need a series of lemmas.

LEMMA 3.2.

1. *Suppose that* $(\mu, \boldsymbol{X}, \boldsymbol{Y}) \in R_{++} \times \mathcal{F}_{++}$. *Then*

$$
\begin{aligned}
&\phi(\mu, \boldsymbol{X}+d\boldsymbol{X}, \boldsymbol{Y}+d\boldsymbol{Y}) - \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) \\
&= \Phi_1(d\boldsymbol{X}, d\boldsymbol{Y}) \; + \; \Phi_2(d\boldsymbol{X}, d\boldsymbol{Y}) \\
&\quad + o(\|\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}\|_F^2) + \; o(\|\sqrt{\boldsymbol{Y}}^{-1} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1}\|_F^2) \\
&\qquad\qquad\qquad\qquad \textit{for all } (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0,
\end{aligned} \tag{12}
$$

where

$$
\left\{
\begin{array}{rcl}
\Phi_1(d\boldsymbol{X}, d\boldsymbol{Y}) &=& (\boldsymbol{Y} - \mu\boldsymbol{X}^{-1}) \bullet d\boldsymbol{X} + (\boldsymbol{X} - \mu\boldsymbol{Y}^{-1}) \bullet d\boldsymbol{Y}, \\
\Phi_2(d\boldsymbol{X}, d\boldsymbol{Y}) &=& d\boldsymbol{X} \bullet d\boldsymbol{Y} + \dfrac{\mu}{2}\|\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}\|_F^2 \\
&& \qquad + \dfrac{\mu}{2}\|\sqrt{\boldsymbol{Y}}^{-1} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1}\|_F^2.
\end{array}
\right. \tag{13}
$$

2. $\phi(\mu, \cdot)$ *is strictly convex on* $\mathcal{F}_{++}$.

Proof. 1. Let $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0$, $\xi_1, \xi_2, \ldots, \xi_n$ be the eigenvalues of the symmetric matrix $\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}$, and $\eta_1, \eta_2, \ldots, \eta_n$ be the eigenvalues of the symmetric matrix $\sqrt{\boldsymbol{Y}}^{-1} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1}$. It suffices to derive (12) under the assumption that $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0$ is so small that the absolute values of all eigenvalues $\xi_1, \xi_2, \ldots, \xi_n$, $\eta_1, \eta_2, \ldots, \eta_n$ are less than one. The assumption ensures that $(\boldsymbol{X}+d\boldsymbol{X}, \boldsymbol{Y}+d\boldsymbol{Y}) \in \mathcal{F}_{++}$. We then see that

$$
\begin{aligned}
&\phi(\mu, \boldsymbol{X}+d\boldsymbol{X}, \boldsymbol{Y}+d\boldsymbol{Y}) \\
&= (\boldsymbol{X}+d\boldsymbol{X}) \bullet (\boldsymbol{Y}+d\boldsymbol{Y}) - \mu \log\det(\boldsymbol{X}+d\boldsymbol{X}) - \mu \log\det(\boldsymbol{Y}+d\boldsymbol{Y}) \\
&= \boldsymbol{X} \bullet \boldsymbol{Y} + d\boldsymbol{X} \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet d\boldsymbol{Y} + d\boldsymbol{X} \bullet d\boldsymbol{Y} \\
&\quad -\mu \left(\log\det \boldsymbol{X} + \log \prod_{j=1}^n (1+\xi_j) \right) - \mu \left(\log\det \boldsymbol{Y} + \log \prod_{j=1}^n (1+\eta_j) \right) \\
&= \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) + d\boldsymbol{X} \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet d\boldsymbol{Y} + d\boldsymbol{X} \bullet d\boldsymbol{Y} \\
&\quad -\mu \left(\sum_{j=1}^n \xi_j - \frac{1}{2} \sum_{j=1}^n \xi_j^2 + o\left(\sum_{j=1}^n \xi_j^2 \right) \right) - \mu \left(\sum_{j=1}^n \eta_j - \frac{1}{2} \sum_{j=1}^n \eta_j^2 + o\left(\sum_{j=1}^n \eta_j^2 \right) \right) \\
&= \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) + \Phi_1(d\boldsymbol{X}, d\boldsymbol{Y}) + \Phi_2(d\boldsymbol{X}, d\boldsymbol{Y}) \\
&\quad + o\left(\|\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}\|_F^2 \right) + o\left(\|\sqrt{\boldsymbol{Y}}^{-1} d\boldsymbol{Y} \sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \right).
\end{aligned}
$$

Thus we have shown assertion 1.

2. Note that the quadratic form $\Phi_2(d\boldsymbol{X}, d\boldsymbol{Y})$ is positive for any nonzero $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0$. This ensures that $\phi(\mu, \cdot)$ is strictly convex on $\mathcal{F}_{++}$. ∎

LEMMA 3.3. *Let* $\mu \in R_{++}$.

1. *If* $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++}$ *satisfies the condition* $\boldsymbol{XY} = \mu\boldsymbol{I}$, *then* $(\boldsymbol{X}, \boldsymbol{Y})$ *is a global minimizer of* $\phi(\mu, \cdot)$ *over* $\mathcal{F}_{++}$.

2. *Assume that there is an* $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}_{++}$. *Then there is a unique global minimizer of* $\phi(\mu, \cdot)$ *over* $\mathcal{F}_{++}$.

Proof. Let λ_j $(j = 1, 2, \cdots, n)$ denote the n positive eigenvalues of the symmetric positive definite matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}$. Then

$$
\phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{X} \bullet \boldsymbol{Y} - \mu \log\det \boldsymbol{XY} = \sum_{j=1}^n (\lambda_j - \mu \log \lambda_j).
$$

Hence we have shown that

$$\phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) = \sum_{j=1}^{n} (\lambda_j - \mu \log \lambda_j). \tag{14}$$

1. Note that each term $\lambda_j - \mu \log \lambda_j$ in the parentheses $(\ \cdot\ )$ attains the minimum under the condition $\lambda_j > 0$ if and only if $\lambda_j = \mu$. On the other hand, we can rewrite the condition $\boldsymbol{XY} = \mu \boldsymbol{I}$ as $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}} = \mu \boldsymbol{I}$ or equivalently $\lambda_j = \mu$ $(j = 1, 2, \ldots, n)$. Thus assertion 1 follows.

2. Take a real number θ such that $\phi(\mu, \boldsymbol{X}^0, \boldsymbol{Y}^0) \leq \theta$. We show that the level set

$$\Gamma = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) \leq \theta\}$$

of the function $\phi(\mu, \cdot)$ is a bounded and closed subset of $\mathcal{S}^2$. Then assertion 2 follows from the continuity and the strict convexity of the function $\phi(\mu, \cdot)$ over the level set Γ. We first show that the level set is contained in the bounded set

$$\Gamma^* = \{\boldsymbol{X} \in \mathcal{S}_+ : \boldsymbol{Y}^0 \bullet \boldsymbol{X} \leq \gamma\} \times \{\boldsymbol{Y} \in \mathcal{S}_+ : \boldsymbol{X}^0 \bullet \boldsymbol{Y} \leq \gamma\},$$

where

$$\gamma = 2n\left(\theta - n(\mu - \mu \log \mu) + \mu \log 2\right) + \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0.$$

Assume that $(\boldsymbol{X}, \boldsymbol{Y}) \in \Gamma$. Let $j \in \{1, 2, \ldots, n\}$ be fixed. We see from $(\boldsymbol{X}, \boldsymbol{Y}) \in \Gamma$ and (14) that

$$\begin{aligned}
\theta &\geq \phi(\mu, \boldsymbol{X}, \boldsymbol{Y}) \\
&= \sum_{j=1}^{n} (\lambda_j - \mu \log \lambda_j) \\
&\geq (n-1)(\mu - \mu \log \mu) + \lambda_j - \mu \log \lambda_j \\
&\geq n(\mu - \mu \log \mu) - \mu \log 2 + \lambda_j/2.
\end{aligned}$$

Hence all the positive eigenvalues λ_j $(j = 1, 2, \ldots, n)$ of the symmetric positive definite matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}$ are bounded from above by the number

$$\gamma' = 2\left(\theta - n(\mu - \mu \log \mu) + \mu \log 2\right).$$

This implies that $\boldsymbol{X} \bullet \boldsymbol{Y} = \mathrm{Tr}\ \sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}} = \sum_{j=1}^{n} \lambda_j \leq n\gamma'$. On the other hand, we see by Condition 1.2 that $\boldsymbol{X} \bullet \boldsymbol{Y} + \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \geq \boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{Y}^0 \bullet \boldsymbol{X}$. Hence $(\boldsymbol{X}, \boldsymbol{Y})$ satisfies that $\boldsymbol{Y}^0 \bullet \boldsymbol{X} \leq \gamma$ and $\boldsymbol{X}^0 \bullet \boldsymbol{Y} \leq \gamma$. See Lemma 1.1 for $\boldsymbol{Y}^0 \bullet \boldsymbol{X} \geq 0$, $\boldsymbol{X}^0 \bullet \boldsymbol{Y} \geq 0$ and the boundedness of the set Γ^*. Thus we have shown that the level set Γ is contained in the bounded set Γ^*.

Now we will prove that the level set Γ is closed. Let $\{(\boldsymbol{X}^k, \boldsymbol{Y}^k)\} \subset \Gamma$ be a sequence converging to some $(\boldsymbol{X}^*, \boldsymbol{Y}^*) \in \mathcal{S}^2$. It suffices to show that $(\boldsymbol{X}^*, \boldsymbol{Y}^*) \in \mathcal{S}^2_{++}$. Since the sequence is bounded, there is a positive number δ such that

$$\log\det \boldsymbol{X}^k \leq \delta \ \text{ and } \ \log\det \boldsymbol{Y}^k \leq \delta \ \text{ for every } k = 1, 2, \ldots\ .$$

It follows that for every $k = 1, 2, \ldots,$

$$\begin{aligned}
\mu \log\det \boldsymbol{X}^k &= -\phi(\mu, \boldsymbol{X}^k, \boldsymbol{Y}^k) + \boldsymbol{X}^k \bullet \boldsymbol{Y}^k - \mu \log\det \boldsymbol{Y}^k \geq -\theta - \mu\delta, \\
\mu \log\det \boldsymbol{Y}^k &= -\phi(\mu, \boldsymbol{X}^k, \boldsymbol{Y}^k) + \boldsymbol{X}^k \bullet \boldsymbol{Y}^k - \mu \log\det \boldsymbol{X}^k \geq -\theta - \mu\delta;
\end{aligned}$$

hence

$$\left.\begin{array}{ll} \boldsymbol{X}^k \in \mathcal{S}_{++}, & \det \boldsymbol{X}^k \geq \exp\left((-\theta - \mu\delta)/\mu\right) > 0 \\ \boldsymbol{Y}^k \in \mathcal{S}_{++}, & \det \boldsymbol{Y}^k \geq \exp\left((-\theta - \mu\delta)/\mu\right) > 0 \end{array}\right\} \text{ for every } k = 1, 2, \dots .$$

This ensures that $(\boldsymbol{X}^*, \boldsymbol{Y}^*) \in \mathcal{S}^2_{++}$. Thus we have shown that the level set Γ is closed. ∎

LEMMA 3.4. *Let L be an m-dimensional monotone linear subspace of R^{2m}. Then its orthogonal complement $L^\perp = \{(\boldsymbol{a}, \boldsymbol{b}) \in R^{2m} : \boldsymbol{a}^T\boldsymbol{u} + \boldsymbol{b}^T\boldsymbol{v} = 0$ for every $(\boldsymbol{u}, \boldsymbol{v}) \in L\}$ is antitone, i.e., $\boldsymbol{a} \bullet \boldsymbol{b} \leq 0$ for every $(\boldsymbol{a}, \boldsymbol{b}) \in L^\perp$.*

Proof. We represent the m-dimensional monotone linear space L as

$$L = \left\{(\boldsymbol{u}, \boldsymbol{v}) \in R^{2m} : \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}^T \\ -\boldsymbol{B}^T \end{pmatrix} \boldsymbol{z},\ \boldsymbol{z} \in R^m \right\},$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ are $m \times m$ matrices such that rank $(\boldsymbol{A}, -\boldsymbol{B}) = m$. We can easily verify that $L^\perp = \{(\boldsymbol{a}, \boldsymbol{b}) \in R^{2m} : \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b} = \boldsymbol{0}\}$. By the monotonicity, $\boldsymbol{B}\boldsymbol{A}^T$ is negative semidefinite. Consider the quadratic form

$$\boldsymbol{u}^T(\boldsymbol{A} - \boldsymbol{B})(\boldsymbol{A} - \boldsymbol{B})^T\boldsymbol{u} = \boldsymbol{u}^T\left((\boldsymbol{A}, -\boldsymbol{B})(\boldsymbol{A}, -\boldsymbol{B})^T - 2\boldsymbol{B}\boldsymbol{A}^T\right)\boldsymbol{u} \text{ for every } \boldsymbol{u} \in R^m.$$

The right-hand side is positive for any nonzero $\boldsymbol{u} \in R^m$ because the $m \times 2m$ matrix $(\boldsymbol{A},\ -\boldsymbol{B})$ is of full row rank, i.e., rank $(\boldsymbol{A}, -\boldsymbol{B}) = m$, and $-\boldsymbol{B}\boldsymbol{A}^T$ is positive semidefinite. Hence so is the left-hand side. This ensures rank $(\boldsymbol{A} - \boldsymbol{B}) = m$. By Theorem 11 of Sznajder–Gowda [39], we obtain that the m-dimensional subspace $\{(\boldsymbol{a}, \boldsymbol{b}) \in R^{2m} : \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{b} = \boldsymbol{0}\}$ is monotone, which implies that the orthogonal complement $L^\perp = \{(\boldsymbol{a}, \boldsymbol{b}) \in R^{2m} : \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b} = \boldsymbol{0}\}$ of L is antitone. ∎

LEMMA 3.5. *Let $\mu \in R_{++}$ be fixed. Suppose that $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++}$ is the global minimizer of $\phi(\mu, \cdot)$ over $\mathcal{F}_{++}$. Then it satisfies $\boldsymbol{X}\boldsymbol{Y} = \mu\boldsymbol{I}$.*

Proof. Recall the relations (12) and (13) in Lemma 3.2. At the global minimizer $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++}$, the linear term $\Phi_1(d\boldsymbol{X}, d\boldsymbol{Y})$ with respect to $(d\boldsymbol{X}, d\boldsymbol{Y})$ in (12) must satisfy the equality

$$\begin{aligned} \Phi_1(d\boldsymbol{X}, d\boldsymbol{Y}) = (\boldsymbol{Y} - \mu\boldsymbol{X}^{-1}) \bullet d\boldsymbol{X} + (\boldsymbol{X} - \mu\boldsymbol{Y}^{-1}) \bullet d\boldsymbol{Y} = 0 \\ \text{for every } (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0. \end{aligned}$$

Hence $(\boldsymbol{Y} - \mu\boldsymbol{X}^{-1}, \boldsymbol{X} - \mu\boldsymbol{Y}^{-1}) \in \mathcal{S}^2$ lies in the orthogonal complement of the $n(n+1)/2$-dimensional monotone linear subspace $\mathcal{F}^0$. By Lemma 3.4,

$$0 \geq (\boldsymbol{Y} - \mu\boldsymbol{X}^{-1}) \bullet (\boldsymbol{X} - \mu\boldsymbol{Y}^{-1}) = \mathrm{Tr}\ \left(\boldsymbol{X}\boldsymbol{Y} - 2\mu\boldsymbol{I} + \mu^2\boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right).$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ denote the eigenvalues of the symmetric positive definite matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}$. Then they are all real and positive. It follows from the inequality above that

$$0 \geq \mathrm{Tr}\ \left(\boldsymbol{X}\boldsymbol{Y} - 2\mu\boldsymbol{I} + \mu^2\boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right) = \sum_{j=1}^{n} \frac{(\lambda_j - \mu)^2}{\lambda_j}.$$

Hence all the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of the symmetric matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}$ are equal to μ. This implies that $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}} = \mu\boldsymbol{I}$. Therefore $\boldsymbol{X}\boldsymbol{Y} = \sqrt{\boldsymbol{X}}(\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}})\sqrt{\boldsymbol{X}}^{-1} = \mu\boldsymbol{I}$. ∎

Assertions 1 and 2 of Theorem 3.1 follow from Lemmas 3.3 and 3.5. To prove assertion 3 of Theorem 3.1, define a mapping $\boldsymbol{H} : R \times R^p \to \mathcal{S} = R^p$ by

$$\boldsymbol{H}(\mu, \boldsymbol{c}) = \left(\boldsymbol{X}^0 + \sum_{i=1}^{p} c_i \boldsymbol{M}^i\right)\left(\boldsymbol{Y}^0 + \sum_{i=1}^{p} c_i \boldsymbol{N}^i\right) - \mu \boldsymbol{I} \text{ for every } (\mu, \boldsymbol{c}) \in R^{1+p}.$$

Here $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}$ and $\{(\boldsymbol{M}^i, \boldsymbol{N}^i) \in \mathcal{S}^2 \ (i = 1, 2, \ldots, p)\}$ denote a basis of $\mathcal{F}^0$ (see (11)). Then each point $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu)) \in \mathcal{S}^2_{++}$ on the central trajectory is characterized as a unique solution of the system of equations

$$(\boldsymbol{X}, \boldsymbol{Y}) \ = \ \left(\boldsymbol{X}^0 + \sum_{i=1}^{p} c_i \boldsymbol{M}^i, \boldsymbol{Y}^0 + \sum_{i=1}^{p} c_i \boldsymbol{N}^i\right) \text{ and } \boldsymbol{H}(\mu, \boldsymbol{c}) \ = \ \boldsymbol{O}.$$

LEMMA 3.6. *Let $\mu \in R_{++}$ and $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^0 + \sum_{i=1}^{p} c_i^0 \boldsymbol{M}^i, \boldsymbol{Y}^0 + \sum_{i=1}^{p} c_i^0 \boldsymbol{N}^i) \in \mathcal{F}_{++}$. Then the Jacobian matrix of the mapping $\boldsymbol{H}(\mu, \cdot)$ with respect to $\boldsymbol{c} \in R^p$ at $\boldsymbol{c}^0$ is nonsingular.*

Proof. It suffices to show that the system of linear equations

$$\left.\frac{d\boldsymbol{H}(\mu, \boldsymbol{c}^0 + td\boldsymbol{c})}{dt}\right|_{t=0} = \boldsymbol{O} \tag{15}$$

has no nonzero solution $d\boldsymbol{c} \in R^p$. A simple calculation shows that

$$\left.\frac{d\boldsymbol{H}(\mu, \boldsymbol{c}^0 + td\boldsymbol{c})}{dt}\right|_{t=0} = \boldsymbol{X}\left(\sum_{i=1}^{p} dc_i \boldsymbol{N}^i\right) + \left(\sum_{i=1}^{p} dc_i \boldsymbol{M}^i\right)\boldsymbol{Y}.$$

Let $d\boldsymbol{c} \in R^p$ be a solution of (15) and $(d\boldsymbol{X}, d\boldsymbol{Y}) = \sum_{i=1}^{p} dc_i(\boldsymbol{M}^i, \boldsymbol{N}^i) \in \mathcal{F}^0$. Then $\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{O}$. By the monotonicity of the linear subspace $\mathcal{F}^0$ and $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0$ we see that $d\boldsymbol{X} \bullet d\boldsymbol{Y} \geq 0$. We also see that $\boldsymbol{O} = \boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y} = \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{X}}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y}$; hence

$$\begin{aligned}
\boldsymbol{O} &= \sqrt{\boldsymbol{X}}d\boldsymbol{Y}\sqrt{\boldsymbol{X}} + \sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\boldsymbol{Y}\sqrt{\boldsymbol{X}} \\
&= \sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}d\boldsymbol{Y}\sqrt{\boldsymbol{X}} + \sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\boldsymbol{Y}\sqrt{\boldsymbol{X}}.
\end{aligned}$$

Thus

$$\begin{aligned}
0 &= \mathrm{Tr}\ \left(\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}d\boldsymbol{Y}\sqrt{\boldsymbol{X}} + \sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\boldsymbol{Y}\sqrt{\boldsymbol{X}}\right) \\
&= d\boldsymbol{X} \bullet d\boldsymbol{Y} + \mathrm{Tr}\ d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\boldsymbol{Y} \\
&\geq \mathrm{Tr}\ \sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\boldsymbol{Y}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}.
\end{aligned}$$

Since the matrix $\boldsymbol{Y} \in \mathcal{S}$ is positive definite, the inequality above implies that every column of the matrix $d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}$ is zero. Hence $(\boldsymbol{O}, \boldsymbol{O}) = (d\boldsymbol{X}, d\boldsymbol{Y}) = \sum_{i=1}^{p} dc_i(\boldsymbol{M}^i, \boldsymbol{N}^i)$. Recall that $\{(\boldsymbol{M}^i, \boldsymbol{N}^i) \in \mathcal{S}^2 \ (i = 1, 2, \ldots, p)\}$ is a basis of $\mathcal{F}^0$. Therefore we obtain $d\boldsymbol{c} = \boldsymbol{0}$. ∎

Obviously, the mapping $\boldsymbol{H}$ is C^∞ on $R \times R^p$. Applying the implicit function theorem (see, for example, [14]), we obtain assertion 3 of Theorem 3.1.

LEMMA 3.7. *For every $\bar{\mu} > 0$, the subset $\{(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu)) : 0 < \mu \leq \bar{\mu}\}$ of the central trajectory is bounded.*

Proof. Let $0 < \mu \leq \bar{\mu}$. By Condition 1.2, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}_{++}$, and $(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu)) \in \mathcal{F}_{++}$, we see that $(\boldsymbol{X}(\mu) - \boldsymbol{X}^0) \bullet (\boldsymbol{Y}(\mu) - \boldsymbol{Y}^0) \geq 0$. It follows that

$$\begin{aligned} \boldsymbol{Y}^0 \bullet \boldsymbol{X}(\mu) + \boldsymbol{X}^0 \bullet \boldsymbol{Y}(\mu) &\leq \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + \boldsymbol{X}(\mu) \bullet \boldsymbol{Y}(\mu) \\ &\leq \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + n\mu \\ &\leq \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + n\bar{\mu}. \end{aligned}$$

Thus the subset $\{(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu)) : 0 < \mu \leq \bar{\mu}\}$ is contained in the bounded set

$$\{\boldsymbol{X} \in \mathcal{S}_+ : \boldsymbol{Y}^0 \bullet \boldsymbol{X} \leq \gamma\} \times \{\boldsymbol{Y} \in \mathcal{S}_+ : \boldsymbol{X}^0 \bullet \boldsymbol{Y} \leq \gamma\},$$

where $\gamma = \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + n\bar{\mu}$. ∎

In view of the lemma above, there exists at least one accumulation point of $(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu))$ as $\mu > 0$ tends to 0. By the continuity, every accumulation point is a solution of the SDLCP (1). The convergence of $(\boldsymbol{X}(\mu), \boldsymbol{Y}(\mu))$ to a single point as $\mu > 0$ tends to 0 follows from the fact that the central trajectory $\mathcal{C}$ is characterized as the algebraic system of equations $\boldsymbol{H}(\mu, \boldsymbol{c}) = \boldsymbol{O}$. The details are omitted here. See Theorem 4.4 of Kojima–Megiddo–Noma–Yoshise [19]. This completes the proof of Theorem 3.1.

4. Newton directions toward the central trajectory. Let $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$ and $\mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n$. Choose $\beta \geq 0$. It might seem natural to regard the system of linear equations

$$(\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) \in \mathcal{F} \ \text{ and } d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{X}d\boldsymbol{Y} = \boldsymbol{Q} \tag{16}$$

in variable matrices $d\boldsymbol{X},\ d\boldsymbol{Y} \in \mathcal{S}$ as the Newton equation at $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$ for approximating a point $(\boldsymbol{X}', \boldsymbol{Y}') = (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) \in \mathcal{S}^2_{++}$ on the central trajectory that satisfies

$$(\boldsymbol{X}', \boldsymbol{Y}') \in \mathcal{F} \ \text{ and } \boldsymbol{X}'\boldsymbol{Y}' = \beta\mu\boldsymbol{I}. \tag{17}$$

Here $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$. However the system (16) does not necessarily have a solution. Hence we need a suitable modification in the system (16) to consistently define Newton directions toward the central trajectory. For this purpose, we introduce an $n(n-1)/2$-dimensional linear subspace $\tilde{\mathcal{F}}^0$ of $\tilde{\mathcal{S}}^2$, where $\tilde{\mathcal{S}} \subset \hat{\mathcal{S}}$ is the $n(n-1)/2$-dimensional linear subspace consisting of all $n \times n$ skew-symmetric matrices. It should be noted that $\mathcal{S}$ and $\tilde{\mathcal{S}}$ are orthogonal complements to each other in the linear space $\hat{\mathcal{S}}$. Since $\mathcal{F}^0 \subset \mathcal{S}^2$ and $\tilde{\mathcal{F}}^0 \subset \tilde{\mathcal{S}}^2$,

$$\begin{aligned} d\boldsymbol{X} \bullet d\tilde{\boldsymbol{X}} \ = \ d\boldsymbol{Y} \bullet d\tilde{\boldsymbol{Y}} \ = \ d\boldsymbol{X} \bullet d\tilde{\boldsymbol{Y}} \ = \ d\boldsymbol{Y} \bullet d\tilde{\boldsymbol{X}} \ = \ 0 \\ \text{for every } (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0 \text{ and } (d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}}) \in \tilde{\mathcal{F}}^0. \end{aligned} \tag{18}$$

We impose $\tilde{\mathcal{F}}^0$ on the condition below.

Condition 4.1. $\tilde{\mathcal{F}}^0$ is monotone, i.e., $d\tilde{\boldsymbol{X}} \bullet d\tilde{\boldsymbol{Y}} \geq 0$ for every $(d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}}) \in \tilde{\mathcal{F}}^0$. For example, we can take $\tilde{\mathcal{F}}^0 = \{(t\boldsymbol{W}, (1-t)\boldsymbol{W}) : \boldsymbol{W} \in \tilde{\mathcal{S}}\}$, where $t \in [0,1]$ is an arbitrary constant.

Now we consider a (modified) Newton equation at $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$ for approximating a point $(\boldsymbol{X}', \boldsymbol{Y}') = (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y})$ on the central trajectory which satisfies (17):

$$\left\{ \begin{array}{l} (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) \in \mathcal{F},\ (d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}}) \in \tilde{\mathcal{F}}^0, \ \text{ and} \\ \boldsymbol{X}(d\boldsymbol{Y} + d\tilde{\boldsymbol{Y}}) + (d\boldsymbol{X} + d\tilde{\boldsymbol{X}})\boldsymbol{Y} = \boldsymbol{Q} \end{array} \right. \tag{19}$$

in variable matrices $d\boldsymbol{X}, d\boldsymbol{Y} \in \mathcal{S}$, and $d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}} \in \tilde{\mathcal{S}}$. Here $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$.

THEOREM 4.2. *Let* $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$, $\mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n$, *and* $\beta \geq 0$. *Then the Newton equation* (19) *has a unique solution* $(d\boldsymbol{X}, d\boldsymbol{Y}, d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}}) \in \mathcal{S}^2 \times \tilde{\mathcal{S}}^2$.

It should be noted that Theorem 4.2 is valid even when the feasible region $\mathcal{F}_+$ of the SDLCP (1) in symmetric matrices is empty or the central trajectory $\mathcal{C}$ does not exist.

Proof of Theorem 4.2. Let $\{(\boldsymbol{M}^i, \boldsymbol{N}^i) \in \mathcal{S}^2\ (i = 1, 2, \dots, p)\}$ be a basis of $\mathcal{F}^0$ and $\{(\tilde{\boldsymbol{M}}^j, \tilde{\boldsymbol{N}}^j) \in \tilde{\mathcal{S}}^2\ (j = 1, 2, \dots, \tilde{p})\}$ be a basis of $\tilde{\mathcal{F}}^0$, and let $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}$, where $p = n(n+1)/2$ and $\tilde{p} = n(n-1)/2$. Note that $p + \tilde{p} = n^2$. Then the first relation of the Newton equation (19) can be written as $(\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) = (\boldsymbol{X}^0, \boldsymbol{Y}^0) + \sum_{i=1}^{p} c_i(\boldsymbol{M}^i, \boldsymbol{N}^i)$, hence

$$d\boldsymbol{X} \ = \ \boldsymbol{X}^0 - \boldsymbol{X} + \sum_{i=1}^{p} c_i\boldsymbol{M}^i \ \text{ and } \ d\boldsymbol{Y} \ = \ \boldsymbol{Y}^0 - \boldsymbol{Y} + \sum_{i=1}^{p} c_i\boldsymbol{N}^i,$$

where $c_i\ (i = 1, 2, \dots, p)$ are real variables. With new variables $\tilde{c}_j\ (j = 1, 2, \dots, \tilde{p})$, we also rewrite the second relation of (19) as $(d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}}) = \sum_{j=1}^{\tilde{p}} \tilde{c}_j(\tilde{\boldsymbol{M}}^j, \tilde{\boldsymbol{N}}^j)$. Now the last equation in (19) is reduced to

$$\sum_{i=1}^{p} c_i(\boldsymbol{X}\boldsymbol{N}^i + \boldsymbol{M}^i\boldsymbol{Y}) + \sum_{j=1}^{\tilde{p}} \tilde{c}_j(\boldsymbol{X}\tilde{\boldsymbol{N}}^j + \tilde{\boldsymbol{M}}^j\boldsymbol{Y}) = \boldsymbol{Q} - \boldsymbol{X}(\boldsymbol{Y}^0 - \boldsymbol{Y}) - (\boldsymbol{X}^0 - \boldsymbol{X})\boldsymbol{Y}.$$

Hence we have only to show that the equation above in n^2 variables $c_i\ (i = 1, 2, \dots, p)$ and $\tilde{c}_j\ (j = 1, 2, \dots, \tilde{p})$ has a unique solution. It suffices to show that the set of n^2 matrices

$$(20)\quad (\boldsymbol{X}\boldsymbol{N}^i + \boldsymbol{M}^i\boldsymbol{Y})\ \ (i = 1, 2, \dots, p) \quad \text{and} \quad (\boldsymbol{X}\tilde{\boldsymbol{N}}^j + \tilde{\boldsymbol{M}}^j\boldsymbol{Y})\ \ (j = 1, 2, \dots, \tilde{p})$$

forms a basis of the n^2-dimensional linear space $\hat{\mathcal{S}}$. Assume that

$$(21)\qquad \sum_{i=1}^{p} c'_i(\boldsymbol{X}\boldsymbol{N}^i + \boldsymbol{M}^i\boldsymbol{Y}) + \sum_{j=1}^{\tilde{p}} \tilde{c}'_j(\boldsymbol{X}\tilde{\boldsymbol{N}}^j + \tilde{\boldsymbol{M}}^j\boldsymbol{Y}) = \boldsymbol{O}$$

for some $c'_i\ (i = 1, 2, \dots, p)$ and $\tilde{c}'_j\ (j = 1, 2, \dots, \tilde{p})$. Let

$$d\boldsymbol{X}' = \sum_{i=1}^{p} c'_i\boldsymbol{M}^i,\ d\boldsymbol{Y}' = \sum_{i=1}^{p} c'_i\boldsymbol{N}^i,\ d\tilde{\boldsymbol{X}}' = \sum_{j=1}^{\tilde{p}} \tilde{c}'_j\tilde{\boldsymbol{M}}^j, \text{ and } d\tilde{\boldsymbol{Y}}' = \sum_{j=1}^{\tilde{p}} \tilde{c}'_j\tilde{\boldsymbol{N}}^j.$$

Then $(d\boldsymbol{X}', d\boldsymbol{Y}') \in \mathcal{F}^0$ and $(d\tilde{\boldsymbol{X}}', d\tilde{\boldsymbol{Y}}') \in \tilde{\mathcal{F}}^0$. We also see from (21) that

$$(22)\qquad \boldsymbol{O} = \boldsymbol{X}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}') + (d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\boldsymbol{Y}.$$

Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ are positive definite, it follows from (22) that

$$\boldsymbol{O} = \sqrt{\boldsymbol{X}}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}}.$$

From the above equality, we obtain that

$$0 = \|\sqrt{\boldsymbol{X}}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \ + \ \|\sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}}\|_F^2$$

$$
\begin{aligned}
&+ \sqrt{\boldsymbol{X}}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1} \bullet \sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}} \\
&+ \sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}} \bullet \sqrt{\boldsymbol{X}}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1} \\
= {} & \|\sqrt{\boldsymbol{X}}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \ + \ \|\sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}}\|_F^2 \\
&+ 2d\boldsymbol{Y}' \bullet d\boldsymbol{X}' + 2d\tilde{\boldsymbol{X}}' \bullet d\tilde{\boldsymbol{Y}}' \\
&(\text{since } d\boldsymbol{Y}' \bullet d\tilde{\boldsymbol{X}}' = d\tilde{\boldsymbol{Y}}' \bullet d\boldsymbol{X}' = 0 \text{ from (18)}) \\
\geq {} & \|\sqrt{\boldsymbol{X}}(d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \ + \ \|\sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}}\|_F^2 \\
&(\text{since } d\boldsymbol{Y}' \bullet d\boldsymbol{X}' \geq 0 \text{ and } d\tilde{\boldsymbol{X}}' \bullet d\tilde{\boldsymbol{Y}}' \geq 0).
\end{aligned}
$$

Hence we see that $\|\sqrt{\boldsymbol{X}}(d\boldsymbol{Y}'+d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 = 0$ and $\|\sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}'+d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}}\|_F^2 = 0$. This implies that $\sqrt{\boldsymbol{X}}(d\boldsymbol{Y}'+d\tilde{\boldsymbol{Y}}')\sqrt{\boldsymbol{Y}}^{-1} = \boldsymbol{O}$ and $\sqrt{\boldsymbol{X}}^{-1}(d\boldsymbol{X}'+d\tilde{\boldsymbol{X}}')\sqrt{\boldsymbol{Y}} = \boldsymbol{O}$. By the nonsingularity of $\sqrt{\boldsymbol{X}}$ and $\sqrt{\boldsymbol{Y}}$, we obtain that $d\boldsymbol{Y}' + d\tilde{\boldsymbol{Y}}' = \boldsymbol{O}$ and $d\boldsymbol{X}' + d\tilde{\boldsymbol{X}}' = \boldsymbol{O}$. We see from (18) that $d\boldsymbol{X}' \bullet d\tilde{\boldsymbol{X}}' = d\boldsymbol{Y}' \bullet d\tilde{\boldsymbol{Y}}' = 0$. Hence the equalities above imply that

$$
(\boldsymbol{O},\boldsymbol{O}) = (d\boldsymbol{X}', d\boldsymbol{Y}') = \sum_{i=1}^{p} c'_i(\boldsymbol{M}^i, \boldsymbol{N}^i) \ \text{ and } (\boldsymbol{O},\boldsymbol{O}) = (d\tilde{\boldsymbol{X}}', d\tilde{\boldsymbol{Y}}') = \sum_{j=1}^{\tilde{p}} \tilde{c}'_j(\tilde{\boldsymbol{M}}^j, \tilde{\boldsymbol{N}}^j).
$$

Recall that $\{(\boldsymbol{M}^i, \boldsymbol{N}^i) \in \mathcal{S}^2 \ (i = 1, 2, \ldots, p)\}$ and $\{(\tilde{\boldsymbol{M}}^j, \tilde{\boldsymbol{N}}^j) \in \tilde{\mathcal{S}}^2 \ (j = 1, 2, \ldots, \tilde{p})\}$ are bases of $\mathcal{F}^0$ and $\tilde{\mathcal{F}}^0$, respectively. Hence $c'_i = 0$ $(i = 1, 2, \ldots, p)$ and $\tilde{c}'_j = 0$ $(j = 1, 2, \ldots, \tilde{p})$. This means that the set of n^2 matrices given in (20) is linearly independent and forms a basis of the n^2-dimensional linear space $\hat{\mathcal{S}}$. This completes the proof of Theorem 4.2.

We can rewrite the Newton equation (19) as

$$
(\boldsymbol{X} + d\hat{\boldsymbol{X}}, \boldsymbol{Y} + d\hat{\boldsymbol{Y}}) \in \mathcal{F} + \tilde{\mathcal{F}}^0 \ \text{ and } \boldsymbol{X} d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{X}}\boldsymbol{Y} = \boldsymbol{Q}. \tag{23}
$$

In fact,

(i) if $(d\boldsymbol{X}, d\boldsymbol{Y}, d\tilde{\boldsymbol{X}}, d\tilde{\boldsymbol{Y}}) \in \mathcal{S}^2 \times \tilde{\mathcal{S}}^2$ is a solution of (19), then $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) = (d\boldsymbol{X} + d\tilde{\boldsymbol{X}}, d\boldsymbol{Y} + d\tilde{\boldsymbol{Y}})$ is a solution of (23).

(ii) if $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$ is a solution of (23), then

$$
((d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2, (d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{Y}}^T)/2, (d\hat{\boldsymbol{X}} - d\hat{\boldsymbol{X}}^T)/2, (d\hat{\boldsymbol{Y}} - d\hat{\boldsymbol{Y}}^T)/2)
$$

is a solution of (19).

The proof of Theorem 4.2 and the argument above do not depend on the specific matrix $\boldsymbol{Q}$ of the right-hand side of the Newton equation (19) or (23). They remain valid for any $\boldsymbol{Q} \in \hat{\mathcal{S}}$. Thus we have the following corollary which will be utilized in our succeeding discussions.

Corollary 4.3. *Let* $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$ *and* $\boldsymbol{Q} \in \hat{\mathcal{S}}$.

1. *The system of equations*

$$
(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0 \ \ \text{and } \boldsymbol{X} d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{X}}\boldsymbol{Y} = \boldsymbol{Q} \tag{24}
$$

has a unique solution $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$.

2. *The solution* $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0$ *satisfies*

$$
\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}} = \sqrt{\boldsymbol{X}}^{-1} \boldsymbol{Q} \sqrt{\boldsymbol{Y}}^{-1}
$$

and

$$\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}} \bullet \sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} \geq 0.$$

Proof. To prove assertion 1, let $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}$ and $\boldsymbol{Q}' = \boldsymbol{Q} - \boldsymbol{X}(\boldsymbol{Y} - \boldsymbol{Y}^0) - (\boldsymbol{X} - \boldsymbol{X}^0)\boldsymbol{Y}$. In view of the discussion above, there exists a unique solution $(d\hat{\boldsymbol{X}}', d\hat{\boldsymbol{Y}}') \in \hat{\mathcal{S}}^2$ of the system of equations

$$(\boldsymbol{X} + d\hat{\boldsymbol{X}}', \boldsymbol{Y} + d\hat{\boldsymbol{Y}}') \ \in \ \mathcal{F} + \tilde{\mathcal{F}}^0 \ \text{ and } \ \boldsymbol{X} d\hat{\boldsymbol{Y}}' + d\hat{\boldsymbol{X}}' \boldsymbol{Y} = \boldsymbol{Q}'.$$

We can rewrite this system of equations as

$$\begin{aligned} &(\boldsymbol{X} + d\hat{\boldsymbol{X}}' - \boldsymbol{X}^0, \boldsymbol{Y} + d\hat{\boldsymbol{Y}}' - \boldsymbol{Y}^0) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0 \text{ and} \\ &\boldsymbol{X}(\boldsymbol{Y} + d\hat{\boldsymbol{Y}}' - \boldsymbol{Y}^0) + (\boldsymbol{X} + d\hat{\boldsymbol{X}}' - \boldsymbol{X}^0)\boldsymbol{Y} = \boldsymbol{Q}. \end{aligned}$$

Letting $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) = (\boldsymbol{X} + d\hat{\boldsymbol{X}}' - \boldsymbol{X}^0, \boldsymbol{Y} + d\hat{\boldsymbol{Y}}' - \boldsymbol{Y}^0)$, we see that $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}})$ is a unique solution of (24). Multiplying both sides of the last equality in the system (24) of equations by $\sqrt{\boldsymbol{X}}^{-1}$ from the left and $\sqrt{\boldsymbol{Y}}^{-1}$ from the right, we have the first relation of assertion 2. The second relation of 2 follows from Conditions 1.2 and 4.1 and the relation (18). ∎

5. A generic interior-point method.

Generic IP method.

Step 0: Choose $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$. Let $r = 0$.

Step 1: Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^r, \boldsymbol{Y}^r)$ and $\mu = (\boldsymbol{X} \bullet \boldsymbol{Y})/n$.

Step 2: Choose *a direction parameter* $\beta \geq 0$.

Step 3: Compute a solution $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$ of the system (23) of equations with $\boldsymbol{Q} = \beta\mu \boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$.

Step 4: Let $d\boldsymbol{X} = (d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2$ and $d\boldsymbol{Y} = (d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{Y}}^T)/2$.

Step 5: Choose *a step size parameter* $\alpha \geq 0$ such that

$$(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) = (\boldsymbol{X}, \boldsymbol{Y}) + \alpha(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{S}^2_{++}. \tag{25}$$

Let $(\boldsymbol{X}^{r+1}, \boldsymbol{Y}^{r+1}) = (\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}})$.

Step 6: Replace r by $r+1$ and go to Step 1.

The generic IP method involves two parameters: a search direction parameter $\beta \geq 0$ and a step size parameter $\alpha \geq 0$. If we choose an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$, which can be infeasible, and if we specify $\beta \geq 0$ and $\alpha \geq 0$ satisfying (25) in each iteration, the method consistently generates a sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ in $\mathcal{S}^2_{++}$. The lemma below is useful to determine a legitimate step size parameter α satisfying (25) and is closely related to the generalized eigenvalue problem of the matrix pencil (see, for example, [8]).

LEMMA 5.1. *Suppose that $\boldsymbol{X} \in \mathcal{S}_{++}$, $d\boldsymbol{X} \in \mathcal{S}$, and $\alpha \geq 0$. Let ξ_{min} be the minimum eigenvalue of the matrix $\boldsymbol{X}^{-1} d\boldsymbol{X}$ and let*

$$\bar{\alpha} = \sup\{\alpha : 1 + \alpha\xi_{min} \geq 0\} = \begin{cases} -1/\xi_{min} & \text{if } \xi_{min} < 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Then $\boldsymbol{X} + \alpha d\boldsymbol{X} \in \mathcal{S}_{++}$ if and only if $\alpha < \bar{\alpha}$.

If we compute the minimum ζ_{min} of all the eigenvalues of $\boldsymbol{X}^{-1}d\boldsymbol{X}$ and $\boldsymbol{Y}^{-1}d\boldsymbol{Y}$ in Step 5 of the generic IP method, then

$$\alpha_{bd} = \begin{cases} -1/\zeta_{min} & \text{if } \zeta_{min} < 0, \\ +\infty & \text{otherwise} \end{cases}$$

gives the upper bound for a step size $\alpha \geq 0$ which satisfies (25).

The generic IP method is analogous to many infeasible interior-point methods ([17, 18, 23, 26, 48], etc.) developed for the monotone LCP (8). Specifically the generic IP method shares with them the features that we can start from an infeasible initial point and that we utilize a Newton direction for approximating a point on the central trajectory. A difference lies in Step 4 of the generic IP method where we symmetrize the Newton direction $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}})$ computed at Step 3 to create a symmetric search direction $(d\boldsymbol{X}, d\boldsymbol{Y})$. This ensures that each iterate $(\boldsymbol{X}^r, \boldsymbol{Y}^r)$ runs in the set $\mathcal{S}^2_{++}$ of symmetric positive definite matrices whenever we take an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ in the set $\mathcal{S}^2_{++}$. The main reason why we use the symmetrized direction is that handling symmetric matrices is much easier than handling nonsymmetric matrices theoretically and practically. In particular,

(a) The logarithmic barrier function $\phi(\mu, \cdot)$ with a fixed $\mu > 0$ is strictly convex on $\mathcal{S}^2_{++}$ (see Lemma 3.2) but not convex on $\hat{\mathcal{S}}^2_{++}$, where $\hat{\mathcal{S}}_{++} = \{\boldsymbol{X} \in \hat{\mathcal{S}} : \boldsymbol{X} \succ \boldsymbol{O}\}$. In fact, if $(\boldsymbol{X}, \hat{\boldsymbol{Y}}) \in \mathcal{S}_{++} \times \hat{\mathcal{S}}_{++}$ but $\hat{\boldsymbol{Y}} \notin \mathcal{S}_{++}$ then

$$\det \frac{\hat{\boldsymbol{Y}} + \hat{\boldsymbol{Y}}^T}{2} < \det \hat{\boldsymbol{Y}} = \det \hat{\boldsymbol{Y}}^T \quad \text{(see [38]);}$$

hence

$$\phi\left(\mu, \boldsymbol{X}, \frac{\hat{\boldsymbol{Y}} + \hat{\boldsymbol{Y}}^T}{2}\right) > \frac{1}{2}\phi(\mu, \boldsymbol{X}, \hat{\boldsymbol{Y}}) + \frac{1}{2}\phi(\mu, \boldsymbol{X}, \hat{\boldsymbol{Y}}^T).$$

(b) If we confine the sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ within $\mathcal{F}_{++} \subset \hat{\mathcal{S}}^2$, the sequence is at least bounded and any accumulation point is a solution of the SDLCP (1) in symmetric matrices as observed in Theorem 3.1.

As special cases of the generic IP method, we present a central trajectory following method in section 8.1 and a potential-reduction method in section 8.2. Both methods may be classified into *feasible interior-point methods*; they start from a feasible interior point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}_{++}$ and generate a sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ in the interior $\mathcal{F}_{++}$ of the feasible region such that $\boldsymbol{X}^r \bullet \boldsymbol{Y}^r \to 0$ as r tends to ∞. It follows from the monotonicity that

$$\boldsymbol{Y}^0 \bullet \boldsymbol{X}^r + \boldsymbol{X}^0 \bullet \boldsymbol{Y}^r \leq \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + \boldsymbol{X}^r \bullet \boldsymbol{Y}^r \quad \text{for every } r = 0, 1, \dots .$$

Since $\boldsymbol{X}^r \bullet \boldsymbol{Y}^r \to 0$ as r tends to ∞, the right-hand side is bounded by a positive number, say ω. This implies that the sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ lies in a closed and bounded set

$$\{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_+ : \boldsymbol{Y}^0 \bullet \boldsymbol{X} + \boldsymbol{X}^0 \bullet \boldsymbol{Y} \leq \omega\}.$$

See Lemma 1.1 for the boundedness of the set. Therefore the sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ has at least one accumulation point and every accumulation point is a solution of the SDLCP (1) in symmetric matrices.

When $(\boldsymbol{X},\boldsymbol{Y}) = (\boldsymbol{X}^r,\boldsymbol{Y}^r) \in \mathcal{S}^2_{++}$ is in the interior $\mathcal{F}_{++}$ of the feasible region, the Newton equation (23) turns out to be

$$(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0 \quad \text{and } \boldsymbol{X}d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{X}}\boldsymbol{Y} = \boldsymbol{Q}, \tag{26}$$

and the search direction $(d\boldsymbol{X}, d\boldsymbol{Y}) = ((d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2, (d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{Y}}^T)/2)$ computed at Step 4 lies in $\mathcal{F}^0$. In this case, (26) coincides with (24). Hence the solution of (26) satisfies item 2 of Corollary 4.3.

6. Some properties of the solution set. It is well known (see, for example, [6]) that if the feasible region of the monotone LCP (8) in the Euclidean space is nonempty then

(i) the solution set of the LCP (8) is a nonempty convex set,

(ii) there exist subsets I, J of the index set $\{1,2,\dots,n\}$ such that

$$\begin{array}{l} I \cup J = \{1,2,\dots,n\}, \\ \left.\begin{array}{rcl} x_i &=& 0 \ (i \in I) \\ y_i &=& 0 \ (i \in J) \end{array}\right\} \text{ for every solution } (\boldsymbol{x},\boldsymbol{y}) \text{ of the LCP (8).} \end{array}$$

We can prove similar results on the monotone SDLCP (1) in symmetric matrices under a slightly stronger assumption that the interior $\mathcal{F}_{++}$ of the feasible region is nonempty.

THEOREM 6.1. *Suppose that the interior $\mathcal{F}_{++}$ of the feasible region of the SDLCP* (1) *in symmetric matrices is nonempty.*

1. *The solution set $\mathcal{F}^*$ of the monotone SDLCP* (1) *is a nonempty convex set.*
2. *There exist subsets I, J of the index set $\{1,2,\dots,n\}$ and an orthogonal matrix $\boldsymbol{P}$ such that*

$$\begin{array}{l} I \cup J = \{1,2,\dots,n\}, \\ \left.\begin{array}{rcl} \left[\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P}\right]_{ij} &=& 0 \ (i \in I \ \text{ or } j \in I) \\ \left[\boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{P}\right]_{ij} &=& 0 \ (i \in J \ \text{ or } j \in J) \end{array}\right\} \textit{for every } (\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F}^*. \end{array}$$

Proof. 1. The nonemptiness of the solution set $\mathcal{F}^*$ follows from Theorem 3.1. Suppose that $(\boldsymbol{X}^1,\boldsymbol{Y}^1),\ (\boldsymbol{X}^2,\boldsymbol{Y}^2) \in \mathcal{F}^*$. By Condition 1.2,

$$0 \ge -(\boldsymbol{X}^2 - \boldsymbol{X}^1) \bullet (\boldsymbol{Y}^2 - \boldsymbol{Y}^1) = \boldsymbol{X}^2 \bullet \boldsymbol{Y}^1 + \boldsymbol{X}^1 \bullet \boldsymbol{Y}^2.$$

Since all matrices $\boldsymbol{X}^1$, $\boldsymbol{X}^2$, $\boldsymbol{Y}^1$, and $\boldsymbol{Y}^2$ are symmetric and positive semidefinite, the inequality above implies that $\boldsymbol{X}^1 \bullet \boldsymbol{Y}^2 = 0$ and $\boldsymbol{X}^2 \bullet \boldsymbol{Y}^1 = 0$. See Lemma 1.1. Hence, for every $\lambda \in [0,1]$,

$$\begin{array}{l} \lambda(\boldsymbol{X}^1,\boldsymbol{Y}^1) + (1-\lambda)(\boldsymbol{X}^2,\boldsymbol{Y}^2) \in \mathcal{F}_+, \\ \left(\lambda\boldsymbol{X}^1 + (1-\lambda)\boldsymbol{X}^2\right) \bullet \left(\lambda\boldsymbol{Y}^1 + (1-\lambda)\boldsymbol{Y}^2\right) = 0. \end{array}$$

Thus we have shown that the solution set $\mathcal{F}^*$ is convex.

2. Let r be the dimension of the affine subspace spanned by the solution set $\mathcal{F}^*$. Then there exist $r+1$ pairs of matrices $(\boldsymbol{X}^0,\boldsymbol{Y}^0), (\boldsymbol{X}^1,\boldsymbol{Y}^1), \dots, (\boldsymbol{X}^r,\boldsymbol{Y}^r) \in \mathcal{F}^*$ such that any $(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F}^*$ can be represented as $(\boldsymbol{X},\boldsymbol{Y}) = \sum_{j=0}^r \alpha_j(\boldsymbol{X}^j,\boldsymbol{Y}^j)$, $\sum_{j=0}^r \alpha_j =$

1 for some $\alpha_0, \alpha_1, \ldots, \alpha_r$. Define

$$\begin{aligned}(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) &= \frac{1}{r+1}\sum_{j=0}^{r}(\boldsymbol{X}^j, \boldsymbol{Y}^j),\\ K &= \{\boldsymbol{p} \in R^n : \boldsymbol{p}^T\bar{\boldsymbol{X}}\boldsymbol{p} = 0\},\\ L &= \{\boldsymbol{p} \in R^n : \boldsymbol{p}^T\bar{\boldsymbol{Y}}\boldsymbol{p} = 0\}.\end{aligned}$$

(K coincides with the subspace spanned by all the eigenvectors of $\bar{\boldsymbol{X}}$ associated with its zero eigenvalue, and L coincides with the subspace spanned by all the eigenvectors of $\bar{\boldsymbol{Y}}$ associated with its zero eigenvalue.) Then for every $\boldsymbol{p} \in K$,

$$0 = \boldsymbol{p}^T\bar{\boldsymbol{X}}\boldsymbol{p} = \frac{1}{r+1}\sum_{j=0}^{r}\boldsymbol{p}^T\boldsymbol{X}^j\boldsymbol{p}.$$

Since each $\boldsymbol{X}^j$ is positive semidefinite, we have $\boldsymbol{X}^j\boldsymbol{p} = \boldsymbol{0}$ $(j = 0, 1, \ldots, r)$. Hence

$$\boldsymbol{X}\boldsymbol{p} = \boldsymbol{0} \quad \text{for every } (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^* \quad \text{and every } \boldsymbol{p} \in K. \tag{27}$$

Similarly we have that

$$\boldsymbol{Y}\boldsymbol{p} = \boldsymbol{0} \quad \text{for every } (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^* \quad \text{and every } \boldsymbol{p} \in L. \tag{28}$$

Let $\xi_1, \xi_2, \ldots, \xi_n$ denote the eigenvalues of $\bar{\boldsymbol{X}}$ and $\bar{\boldsymbol{p}}^1, \bar{\boldsymbol{p}}^2, \ldots, \bar{\boldsymbol{p}}^n$ the eigenvectors corresponding to them. We may assume that the eigenvectors form a normalized orthogonal basis of R^n. Define $I = \{j : \xi_j = 0\}$ and $J = \{j : \xi_j > 0\}$. Then $\{\bar{\boldsymbol{p}}^j : j \in I\}$ forms a basis of the subspace K. It follows from (27) that

$$\boldsymbol{X}\bar{\boldsymbol{p}}^j = \boldsymbol{0} \quad \text{for every } (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^* \quad \text{and every } j \in I. \tag{29}$$

On the other hand, $\bar{\boldsymbol{X}}\bar{\boldsymbol{Y}} = \boldsymbol{O}$ because $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) \in \mathcal{F}^*$. Hence

$$0 = (\bar{\boldsymbol{p}}^j)^T\bar{\boldsymbol{X}}\bar{\boldsymbol{Y}}\bar{\boldsymbol{p}}^j = \xi_j(\bar{\boldsymbol{p}}^j)^T\bar{\boldsymbol{Y}}\bar{\boldsymbol{p}}^j \quad \text{for every } j \in J,$$

which together with $\xi_j > 0$ $(j \in J)$ implies that $(\bar{\boldsymbol{p}}^j)^T\bar{\boldsymbol{Y}}\bar{\boldsymbol{p}}^j = 0$ for every $j \in J$, or equivalently $\bar{\boldsymbol{p}}^j \in L$ for every $j \in J$. By (28), we then see that

$$\boldsymbol{Y}\bar{\boldsymbol{p}}^j = \boldsymbol{0} \quad \text{for every } (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^* \quad \text{and every } j \in J. \tag{30}$$

Letting $\boldsymbol{P} = (\bar{\boldsymbol{p}}^1, \bar{\boldsymbol{p}}^2, \ldots, \bar{\boldsymbol{p}}^n)$, we obtain the desired result from the definition of I, J, (29), and (30). ∎

Remark 6.2.

(a) In item 1 of Theorem 6.1, the condition $\mathcal{F}_{++} \neq \emptyset$ cannot be weakened. It is well known that the nonemptiness of the feasible region gives a necessary and sufficient condition for the existence of a solution of the monotone LCP (8) in the Euclidean space (see, for example, [6]). In contrast with that case, however, the weaker condition $\mathcal{F}_{+} \neq \emptyset$ does not imply the solvability of the SDLCP(1). This is due to the fact that the cone of positive semidefinite matrices is not polyhedral. (See Gowda and Seidman [10].) There is a similar gap in the case of the monotone nonlinear complementarity problem [29].

(b) Result 2 of Theorem 6.1 is closely related to the complementary slackness theorem (Corollary 2.11 of [2]): it says that all solutions $(\boldsymbol{X}, \boldsymbol{Y})$s of the SDP(2) share a system of eigenvectors and their eigenvalues are complementary in the sense of LCP.

7. Basic lemmas. In this section we prepare basic lemmas which play important roles in what follows.

LEMMA 7.1. *Suppose that* $(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{S}^2_{++}$. *Let* λ_{min} *and* λ_{max} *denote the minimum and the maximum eigenvalues of* $\boldsymbol{XY}$, *respectively.*

1. *Let* $(\hat{d\boldsymbol{X}}, \hat{d\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$. *Then*

$$\|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{X}}^{-1}\|_F \leq \frac{\|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F}{\sqrt{\lambda_{min}}}, \tag{31}$$

$$\|\sqrt{\boldsymbol{Y}}^{-1}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F \leq \frac{\|\sqrt{\boldsymbol{X}}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F}{\sqrt{\lambda_{min}}}, \tag{32}$$

$$\|\sqrt{\boldsymbol{Y}}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F \leq \sqrt{\lambda_{max}}\|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F, \tag{33}$$

$$\|\sqrt{\boldsymbol{X}}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{X}}\|_F \leq \sqrt{\lambda_{max}}\|\sqrt{\boldsymbol{X}}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F. \tag{34}$$

2. *Let* $(\hat{d\boldsymbol{X}}, \hat{d\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$, $d\boldsymbol{X} = (\hat{d\boldsymbol{X}} + \hat{d\boldsymbol{X}}^T)/2$, $d\boldsymbol{Y} = (\hat{d\boldsymbol{Y}} + \hat{d\boldsymbol{Y}}^T)/2$, $\xi_1, \xi_2, \dots, \xi_n$ *be the eigenvalues of* $\boldsymbol{X}^{-1}d\boldsymbol{X}$, *and* $\eta_1, \eta_2, \dots, \eta_n$ *be the eigenvalues of* $\boldsymbol{Y}^{-1}d\boldsymbol{Y}$. *Then*

$$\sum_{j=1}^n \xi_j^2 \leq \|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{X}}^{-1}\|_F^2 \leq \frac{\|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F^2}{\lambda_{min}}, \tag{35}$$

$$\sum_{j=1}^n \eta_j^2 \leq \|\sqrt{\boldsymbol{Y}}^{-1}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \leq \frac{\|\sqrt{\boldsymbol{X}}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F^2}{\lambda_{min}}. \tag{36}$$

3. *(Extension of the inequalities given in Lemma* 4.20 *of* [19]*). Let* $\boldsymbol{Q} \in \hat{\mathcal{S}}$ *and let* $(\hat{d\boldsymbol{X}}, \hat{d\boldsymbol{Y}})$ *be a solution of the system* (24) *of equations. Then*

$$\begin{aligned} &\|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F^2 + \|\sqrt{\boldsymbol{X}}\hat{d\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \\ &\quad = \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}\sqrt{\boldsymbol{Y}}^{-1}\|_F^2 - 2\hat{d\boldsymbol{X}} \bullet \hat{d\boldsymbol{Y}}, \end{aligned} \tag{37}$$

$$0 \leq \hat{d\boldsymbol{X}} \bullet \hat{d\boldsymbol{Y}} \leq \frac{\|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}\sqrt{\boldsymbol{Y}}^{-1}\|_F^2}{4}. \tag{38}$$

Proof. 1. In general, the inequalities $\nu_{min}(\boldsymbol{A})\|\boldsymbol{B}\|_F^2 \leq \text{Tr } \boldsymbol{B}^T\boldsymbol{AB} \leq \nu_{max}(\boldsymbol{A})\|\boldsymbol{B}\|_F^2$ hold for every $\boldsymbol{A} \in \mathcal{S}_{++}$ and every $\boldsymbol{B} \in \hat{\mathcal{S}}$, where $\nu_{min}(\boldsymbol{A})$ and $\nu_{max}(\boldsymbol{A})$ denote the minimum and the maximum eigenvalues of $\boldsymbol{A}$, respectively. On the other hand, λ_{min} and $1/\lambda_{max}$ are the minimum eigenvalues of the matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}$ and its inverse $(\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}})^{-1}$, respectively. Hence

$$\begin{aligned} \|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F^2 &= \text{Tr } (\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{X}}^{-1})^T\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}(\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{X}}^{-1}) \\ &\geq \lambda_{min}\|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{X}}^{-1}\|_F^2, \\ \|\sqrt{\boldsymbol{X}}^{-1}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F^2 &= \text{Tr } (\sqrt{\boldsymbol{Y}}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}})^T(\sqrt{\boldsymbol{Y}}\boldsymbol{X}\sqrt{\boldsymbol{Y}})^{-1}(\sqrt{\boldsymbol{Y}}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}) \\ &\geq \frac{1}{\lambda_{max}}\|\sqrt{\boldsymbol{Y}}\hat{d\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F^2. \end{aligned}$$

Thus we have shown (31) and (33). The proof of (32) and (34) is quite similar.

2. We only show that the inequality (35) holds. The inequality (36) can be proven similarly. Since the symmetric matrix $\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}$ has the same eigenvalues

$\xi_1, \xi_2, \ldots, \xi_n$ as the matrix $\boldsymbol{X}^{-1} d\boldsymbol{X}$, we have that

$$
\begin{aligned}
\sum_{j=1}^{n} \xi_j^2 &= \|\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}\|_F^2 \\
&= \left\| \sqrt{\boldsymbol{X}}^{-1} \left(\frac{d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T}{2} \right) \sqrt{\boldsymbol{X}}^{-1} \right\|_F^2 \quad (\text{since } d\boldsymbol{X} = (d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2) \\
&\leq \left(\frac{\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{X}}^{-1}\|_F}{2} + \frac{\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{X}}^{-1}\|_F}{2} \right)^2 \\
&= \|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{X}}^{-1}\|_F^2 \\
&\leq \frac{\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F^2}{\lambda_{min}}.
\end{aligned}
$$

Here the last inequality follows from (31).

3. By Corollary 4.3, we have $\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}} = \sqrt{\boldsymbol{X}}^{-1} \boldsymbol{Q} \sqrt{\boldsymbol{Y}}^{-1}$. Hence

$$
\begin{aligned}
&\|\sqrt{\boldsymbol{X}}^{-1} \boldsymbol{Q} \sqrt{\boldsymbol{Y}}^{-1}\|_F^2 \\
&= (\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}) \bullet (\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}) \\
&= \|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F^2 + \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1}\|_F^2 + 2 d\hat{\boldsymbol{X}} \bullet d\hat{\boldsymbol{Y}}.
\end{aligned}
$$

Thus we have shown (37). Since the linear subspaces $\mathcal{F}^0$ and $\tilde{\mathcal{F}}^0$ of $\hat{\mathcal{S}}^2$ are orthogonal to each other (see (18)), the first inequality of (38) follows directly from Conditions 1.2 and 4.1. We also see that

$$
\begin{aligned}
&d\hat{\boldsymbol{X}} \bullet d\hat{\boldsymbol{Y}} \\
&= \frac{1}{4} \left\{ \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F^2 - \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F^2 \right\} \\
&\leq \frac{1}{4} \left\{ \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F^2 \right\} \\
&= \frac{\|\sqrt{\boldsymbol{X}}^{-1} \boldsymbol{Q} \sqrt{\boldsymbol{Y}}^{-1}\|_F^2}{4}.
\end{aligned}
$$

Thus we have shown (38). ∎

LEMMA 7.2. *(Extension of Lemma 4.16 of [19]). Let* $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_{++}^2$, $\mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n$, $0 \leq \beta \leq 1$ *and let* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *be the eigenvalues of the matrix* $\boldsymbol{X}\boldsymbol{Y}$. *Define the* $n \times n$ *matrix* $\boldsymbol{H}(\beta)$ *as* $\boldsymbol{H}(\beta) = \beta\mu\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}$.

1. *Define* $\boldsymbol{\Lambda} = diag\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$ *to be the* $n \times n$ *diagonal matrix with the coordinates* $\lambda_1, \lambda_2, \ldots, \lambda_n$. *Then*

$$
\|\boldsymbol{H}(\beta)\|_F^2 = \|\beta\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}\|_F^2. \tag{39}
$$

2. *Assume that* $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{N}(\gamma)$ *for some* $\gamma \in (0, \sqrt{n}]$. *Here*

$$
\mathcal{N}(\gamma) = \left\{ \begin{array}{ll} (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : & \left(\displaystyle\sum_{j=1}^{n} (\lambda_j - \mu)^2 \right)^{1/2} \leq \gamma\mu, \quad \text{where } \mu = \dfrac{\boldsymbol{X} \bullet \boldsymbol{Y}}{n}, \\ & \text{and } \lambda_1, \ldots, \lambda_n \ \text{ denote the eigenvalues of } \boldsymbol{X}\boldsymbol{Y} \end{array} \right\}
$$

$$= \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : \|\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}} - \mu \boldsymbol{I}\|_F \leq \gamma\mu, \quad \textit{where } \mu = \frac{\boldsymbol{X} \bullet \boldsymbol{Y}}{n} \right\}.$$

Then

$$(1-\gamma)\mu \leq \lambda_{min} \leq \lambda_j \leq \lambda_{max} \leq (1+\gamma)\mu \quad \textit{for every } j = 1, 2, \ldots, n, \tag{40}$$

$$\|\boldsymbol{H}(\beta)\|_F \leq \min \left\{ \frac{((1-\beta)\sqrt{n}+\gamma)\mu}{\sqrt{\lambda_{min}}}, \ \frac{\sqrt{2n}\mu}{\sqrt{\lambda_{min}}} \right\}.$$

Proof. 1. By the definition, we see that

$$\begin{aligned} \|\boldsymbol{H}(\beta)\|_F^2 &= \|\beta\mu\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F^2 \\ &= \sum_{j=1}^n \left(\frac{\beta\mu}{\sqrt{\lambda_j}} - \sqrt{\lambda_j} \right)^2 \\ &= \|\beta\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}\|_F^2. \end{aligned}$$

Thus we have shown the equality (39).

2. We see from $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{N}(\gamma)$ that

$$\left(\sum_{j=1}^n (\lambda_j - \mu)^2 \right)^{1/2} \leq \gamma\mu, \tag{41}$$

which implies (40). Hence

$$\begin{aligned} \|\boldsymbol{H}(\beta)\|_F &= \|\beta\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}\|_F \text{ (by (39))} \\ &\leq (1-\beta)\mu\|\sqrt{\boldsymbol{\Lambda}}^{-1}\|_F + \|\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}\|_F \\ &\leq (1-\beta)\mu \left(\sum_{j=1}^n \frac{1}{\lambda_{min}} \right)^{1/2} + \left(\sum_{j=1}^n \left(\frac{\mu - \lambda_j}{\sqrt{\lambda_{min}}} \right)^2 \right)^{1/2} \quad \text{(by (40))} \\ &\leq (1-\beta)\mu \cdot \frac{\sqrt{n}}{\sqrt{\lambda_{min}}} + \frac{\gamma\mu}{\sqrt{\lambda_{min}}} \quad \text{(by (41))} \\ &= \frac{((1-\beta)\sqrt{n}+\gamma)\,\mu}{\sqrt{\lambda_{min}}}. \end{aligned}$$

We also see that

$$\begin{aligned} \|\boldsymbol{H}(\beta)\|_F^2 &= \|\beta\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}\|_F^2 \text{ (by (39))} \\ &\leq \sum_{j=1}^n \left(\left(\frac{\beta\mu}{\sqrt{\lambda_j}} \right)^2 + \left(\sqrt{\lambda_j} \right)^2 \right) \\ &\leq \frac{n\mu^2}{\lambda_{min}} + n\mu \left(\text{since } 0 \leq \beta \leq 1 \text{ and } \sum_{j=1}^n \lambda_j = n\mu \right) \\ &\leq \frac{2n\mu^2}{\lambda_{min}} \text{ (since } 1 \leq \mu/\lambda_{min}) \\ &\leq \frac{2n\mu^2}{\lambda_{min}}. \end{aligned}$$

Thus we have shown assertion 2. $\square$

LEMMA 7.3. *(See* [16, 42]*, etc.)*

1. *If* $1+\xi>0$ *then* $\log(1+\xi)\leq\xi$.
2. *If* $\boldsymbol{\xi}\in R^n$ *satisfies* $\|\boldsymbol{\xi}\|_\infty\leq\tau<1$ *then* $\sum_{j=1}^n\log(1+\xi_j)\geq\sum_{j=1}^n\xi_j-\dfrac{\|\boldsymbol{\xi}\|^2}{2(1-\tau)}$.

For every $(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{F}_{++}$, define the potential function

$$f(\boldsymbol{X},\boldsymbol{Y})=(n+\nu)\log\boldsymbol{X}\bullet\boldsymbol{Y}-\log\det\boldsymbol{X}\boldsymbol{Y}-n\log n.$$

Here $\nu\geq 0$ is a parameter. This potential function is the same as the one used in the paper [44] where Vandenberghe and Boyd proposed a potential-reduction method for the primal-dual pair (2) of SDPs. (See also [2, 36].)

LEMMA 7.4. *Let* $(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{F}_{++}$, $(d\boldsymbol{X},d\boldsymbol{Y})\in\mathcal{S}^2$, $\mu=\boldsymbol{X}\bullet\boldsymbol{Y}/n$, $0<\tau<1$, *and* $\nu\geq 0$. *Let* $\xi_1,\xi_2,\dots,\xi_n$ *be the eigenvalues of the matrix* $\boldsymbol{X}^{-1}d\boldsymbol{X}$ *and* $\eta_1,\eta_2,\dots,\eta_n$ *be the eigenvalues of the matrix* $d\boldsymbol{Y}\boldsymbol{Y}^{-1}$. *Let* α *be a positive number such that* $|\alpha\xi_j|\leq\tau$ *and* $|\alpha\eta_j|\leq\tau$ *for every* $j=1,2,\dots,n$.

1. $f(\boldsymbol{X}+\alpha d\boldsymbol{X},\boldsymbol{Y}+\alpha d\boldsymbol{Y})-f(\boldsymbol{X},\boldsymbol{Y})\leq\alpha G_1(d\boldsymbol{X},d\boldsymbol{Y})+\alpha^2G_2(d\boldsymbol{X},d\boldsymbol{Y})$, *where*

$$G_1(d\boldsymbol{X},d\boldsymbol{Y})=Tr\left(\frac{n+\nu}{n\mu}\boldsymbol{I}-\boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right)(d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{X}d\boldsymbol{Y}),$$
$$G_2(d\boldsymbol{X},d\boldsymbol{Y})=\frac{(n+\nu)d\boldsymbol{X}\bullet d\boldsymbol{Y}}{n\mu}+\frac{\sum_{j=1}^n(\xi_j^2+\eta_j^2)}{2(1-\tau)}.$$

2. *(Extension of Lemma* 2.5 *of* [22]*). Let* $\lambda_1,\lambda_2,\dots,\lambda_n$ *be the eigenvalues of the matrix* $\boldsymbol{X}\boldsymbol{Y}$ *and* $\lambda_{min}=\min\{\lambda_1,\lambda_2,\dots,\lambda_n\}$. *Assume that* $\beta=n/(n+\nu)$ *and that* $(d\hat{\boldsymbol{X}},d\hat{\boldsymbol{Y}})$ *is a solution of the Newton equation* (23) *with* $\boldsymbol{Q}=\beta\mu\boldsymbol{I}-\boldsymbol{X}\boldsymbol{Y}$. *Let* $d\boldsymbol{X}=(d\hat{\boldsymbol{X}}+d\hat{\boldsymbol{X}}^T)/2$ *and* $d\boldsymbol{Y}=(d\hat{\boldsymbol{Y}}+d\hat{\boldsymbol{Y}}^T)/2$. *Then*

$$G_1(d\boldsymbol{X},d\boldsymbol{Y})=-\frac{1}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2,\tag{42}$$

$$G_2(d\boldsymbol{X},d\boldsymbol{Y})\leq\frac{\|\boldsymbol{H}(\beta)\|_F^2}{2(1-\tau)\lambda_{min}}+\frac{1}{\lambda_{min}}\left(\frac{n+\nu}{n}-\frac{1}{1-\tau}\right)d\boldsymbol{X}\bullet d\boldsymbol{Y}.\tag{43}$$

If in addition, $\nu\geq\sqrt{n}$ *then*

$$\frac{\sqrt{\lambda_{min}}}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F\geq\frac{\sqrt{3}}{2}.\tag{44}$$

Proof. 1. The desired inequality follows from the calculation below.

$$\begin{aligned}
&f(\boldsymbol{X}+\alpha d\boldsymbol{X},\boldsymbol{Y}+\alpha d\boldsymbol{Y})-f(\boldsymbol{X},\boldsymbol{Y})\\
&=(n+\nu)\log\left(1+\frac{\alpha\mathrm{Tr}\ (d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{X}d\boldsymbol{Y})}{n\mu}+\frac{\alpha^2d\boldsymbol{X}\bullet d\boldsymbol{Y}}{n\mu}\right)\\
&\quad-\sum_{j=1}^n(\log(1+\alpha\xi_j)+\log(1+\alpha\eta_j))\\
&\leq(n+\nu)\left(\frac{\alpha\mathrm{Tr}\ (d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{X}d\boldsymbol{Y})}{n\mu}+\frac{\alpha^2d\boldsymbol{X}\bullet d\boldsymbol{Y}}{n\mu}\right)\\
&\quad-\left(\alpha\sum_{j=1}^n(\xi_j+\eta_j)-\alpha^2\frac{\sum_{j=1}^n(\xi_j^2+\eta_j^2)}{2(1-\tau)}\right)\quad\text{(by Lemma 7.3)}
\end{aligned}$$

$$
\begin{aligned}
&= \alpha \mathrm{Tr}\ \left(\frac{n+\nu}{n\mu}\boldsymbol{I} - \boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right)(d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{X}d\boldsymbol{Y}) \\
&\quad + \alpha^2 \left(\frac{(n+\nu)d\boldsymbol{X}\bullet d\boldsymbol{Y}}{n\mu} + \frac{\sum_{j=1}^n(\xi_j^2+\eta_j^2)}{2(1-\tau)}\right).
\end{aligned}
$$

2. By the definition of G_1,

$$
\begin{aligned}
G_1(d\boldsymbol{X}, d\boldsymbol{Y}) &= \mathrm{Tr}\ \left(\frac{n+\nu}{n\mu}\boldsymbol{I} - \boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right)(d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{X}d\boldsymbol{Y}) \\
&= \mathrm{Tr}\ \left(\frac{n+\nu}{n\mu}\boldsymbol{I} - \boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right)(d\hat{\boldsymbol{X}}\boldsymbol{Y} + \boldsymbol{X}d\hat{\boldsymbol{Y}}) \\
&= \mathrm{Tr}\ \left(\frac{1}{\beta\mu}\boldsymbol{I} - \boldsymbol{Y}^{-1}\boldsymbol{X}^{-1}\right)(\beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}) \\
&\qquad (\text{since } \beta = n/(n+\nu) \text{ and } d\hat{\boldsymbol{X}}\boldsymbol{Y} + \boldsymbol{X}d\hat{\boldsymbol{Y}} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}) \\
&= \mathrm{Tr}\ \left(\frac{1}{\beta\mu}\boldsymbol{I} - (\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}})^{-1}\right)(\beta\mu\boldsymbol{I} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}}) \\
&= \mathrm{Tr}\ \left(\frac{1}{\beta\mu}\sqrt{\boldsymbol{\Lambda}} - \sqrt{\boldsymbol{\Lambda}}^{-1}\right)(\beta\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}) \\
&= -\frac{1}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 \ (\text{by Lemma 7.2}).
\end{aligned}
$$

Thus we have shown (42).

By the definition of G_2, $0 < \lambda_{min} \le \mu$, and Lemma 7.1, we see that

$$
\begin{aligned}
G_2(d\boldsymbol{X}, d\boldsymbol{Y}) &= \frac{(n+\nu)d\boldsymbol{X}\bullet d\boldsymbol{Y}}{n\mu} + \frac{\sum_{j=1}^n(\xi_j^2+\eta_j^2)}{2(1-\tau)} \\
&\le \frac{\|\boldsymbol{H}(\beta)\|_F^2}{2(1-\tau)\lambda_{min}} + \frac{n+\nu}{\lambda_{min}n}d\boldsymbol{X}\bullet d\boldsymbol{Y} - \frac{1}{\lambda_{min}(1-\tau)}d\hat{\boldsymbol{X}}\bullet d\hat{\boldsymbol{Y}} \\
&\le \frac{\|\boldsymbol{H}(\beta)\|_F^2}{2(1-\tau)\lambda_{min}} + \frac{1}{\lambda_{min}}\left(\frac{n+\nu}{n} - \frac{1}{1-\tau}\right)d\boldsymbol{X}\bullet d\boldsymbol{Y}.
\end{aligned}
$$

Here the last inequality is due to the fact that $d\boldsymbol{X}\bullet d\boldsymbol{Y} \le d\hat{\boldsymbol{X}}\bullet d\hat{\boldsymbol{Y}}$. Thus we obtain the inequality (43).

Finally we prove the inequality (44) under the assumption that $\nu \ge \sqrt{n}$.

$$
\begin{aligned}
&\left(\frac{\sqrt{\lambda_{min}}}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F\right)^2 \\
&= \left(\frac{\sqrt{\lambda_{min}}}{\beta\mu}\|\beta\mu\sqrt{\boldsymbol{\Lambda}}^{-1} - \sqrt{\boldsymbol{\Lambda}}\|_F\right)^2 \ (\text{by Lemma 7.2}) \\
&= \lambda_{min}\left\|\frac{n+\nu}{n\mu}\sqrt{\boldsymbol{\Lambda}} - \sqrt{\boldsymbol{\Lambda}}^{-1}\right\|_F^2 \ (\text{since } \beta = n/(n+\nu)) \\
&\ge \lambda_{min}\left\|\frac{\sqrt{\boldsymbol{\Lambda}}}{\sqrt{n}\mu}\right\|_F^2 + \lambda_{min}\left\|\frac{\sqrt{\boldsymbol{\Lambda}}}{\mu} - \sqrt{\boldsymbol{\Lambda}}^{-1}\right\|_F^2 \\
&\qquad \left(\text{since Tr } \sqrt{\boldsymbol{\Lambda}}\left(\frac{\sqrt{\boldsymbol{\Lambda}}}{\mu} - \sqrt{\boldsymbol{\Lambda}}^{-1}\right) = 0 \text{ and } \nu \ge \sqrt{n}\right)
\end{aligned}
$$

$$\begin{aligned}
&\geq \frac{\lambda_{min}}{\mu} + \lambda_{min} \left| \frac{\sqrt{\lambda_{min}}}{\mu} - \frac{1}{\sqrt{\lambda_{min}}} \right|^2 \\
&= \frac{(\mu/2 - \lambda_{min})^2 + 3\mu^2/4}{\mu^2} \\
&\geq \frac{3}{4}. \qquad \square
\end{aligned}$$

In the two lemmas below, we are concerned with the following hypothesis.

Hypothesis. (See Hypothesis 4.1 of [17].) Let $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$. There exists a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP (1) such that

$$\omega^* \boldsymbol{X}^0 \succeq \boldsymbol{X}^* \quad \text{and } \omega^* \boldsymbol{Y}^0 \succeq \boldsymbol{Y}^* \tag{45}$$

for some $\omega^* \geq 1$.

This hypothesis as well as the lemmas will be utilized in section 8.3 where we present an infeasible interior-point potential-reduction method.

LEMMA 7.5. *Let $(\boldsymbol{X}, \boldsymbol{Y})$ and $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$. Assume that the hypothesis is true. Then*

$$\begin{aligned}
\|\sqrt{\boldsymbol{X}}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\sqrt{\boldsymbol{X}}\|_F \leq \omega^* \boldsymbol{X} \bullet \boldsymbol{Y}^0, \\
\|\sqrt{\boldsymbol{Y}}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\sqrt{\boldsymbol{Y}}\|_F \leq \omega^* \boldsymbol{X}^0 \bullet \boldsymbol{Y}.
\end{aligned}$$

Proof. From the assumption,

$$\begin{aligned}
\omega^* \sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} + \left(\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}^*\sqrt{\boldsymbol{X}}\right) \in \mathcal{S}_{++}, \\
\omega^* \sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \left(\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}^*\sqrt{\boldsymbol{X}}\right) \in \mathcal{S}_{++}.
\end{aligned}$$

Hence we have

$$\begin{aligned}
0 \leq \mathrm{Tr}\ &\left(\omega^* \sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} + \left(\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}^*\sqrt{\boldsymbol{X}}\right)\right)^T \\
&\left(\omega^* \sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \left(\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}^*\sqrt{\boldsymbol{X}}\right)\right) \\
= (\omega^*)^2 &\|\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}}\|_F^2 - \|\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}^*\sqrt{\boldsymbol{X}}\|_F^2.
\end{aligned}$$

It follows that $\|\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}} - \sqrt{\boldsymbol{X}}\boldsymbol{Y}^*\sqrt{\boldsymbol{X}}\|_F \leq \omega^* \|\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}}\|_F$. Since the matrix $\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}}$ is symmetric and positive definite, we have

$$\|\sqrt{\boldsymbol{X}}\boldsymbol{Y}^0\sqrt{\boldsymbol{X}}\|_F \leq \sqrt{\boldsymbol{X}} \bullet \boldsymbol{Y}^0\sqrt{\boldsymbol{X}} = \boldsymbol{X} \bullet \boldsymbol{Y}^0.$$

Thus the first inequality in the lemma follows. We can derive the second inequality of the lemma similarly. $\square$

LEMMA 7.6. *(Lemmas 5.1 and 5.2 of [17]). Let $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$. Assume that the hypothesis is true and that $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++}$ satisfies*

$$\theta \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \leq \xi \boldsymbol{X} \bullet \boldsymbol{Y}, \tag{46}$$

$$(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^0 + \theta(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta)(\boldsymbol{X}^*, \boldsymbol{Y}^*) \tag{47}$$

for some $\xi \geq 1$ and $\theta \in [0, 1]$. Let λ_{min} be the minimum eigenvalue of the matrix $\boldsymbol{X}\boldsymbol{Y}$, $\mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n$, $\sigma = 2\omega^*\xi + 1$, $\zeta = 2 + \omega^*\sigma$. Let $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$ be a solution of the Newton equation (23) with $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$, and $d\boldsymbol{X} = (d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2$ and $d\boldsymbol{Y} = (d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{Y}}^T)/2$.

1. $\theta(\boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}^0) \leq \sigma \boldsymbol{X} \bullet \boldsymbol{Y}$.
2. $\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F,\ \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1}\|_F \leq \|\boldsymbol{H}(\beta)\|_F + \frac{\omega^* \sigma n \mu}{\sqrt{\lambda_{min}}} \leq \frac{\zeta n \mu}{\sqrt{\lambda_{min}}}$.
3. *Let* $\xi_1, \xi_2, \ldots, \xi_n$ *be the eigenvalues of* $\boldsymbol{X}^{-1} d\boldsymbol{X}$ *and* $\eta_1, \eta_2, \ldots, \eta_n$ *be the eigenvalues of* $\boldsymbol{Y}^{-1} d\boldsymbol{Y}$. *Then* $\sum_{j=1}^n \xi_j^2,\ \sum_{j=1}^n \eta_j^2 \leq \left(\frac{\zeta n \mu}{\lambda_{min}}\right)^2$.

Proof. 1. By the assumption, there is a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP (1) satisfying

$$\omega^* \boldsymbol{X}^0 \succeq \boldsymbol{X}^* \text{ and } \omega^* \boldsymbol{Y}^0 \succeq \boldsymbol{Y}^*.$$

It follows that

$$\boldsymbol{X}^0 \bullet \boldsymbol{Y}^* \leq \omega^* \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \text{ and } \boldsymbol{X}^* \bullet \boldsymbol{Y}^0 \leq \omega^* \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0. \tag{48}$$

Let $(\boldsymbol{X}', \boldsymbol{Y}') = \theta(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta)(\boldsymbol{X}^*, \boldsymbol{Y}^*)$. By the relation (47) which we have assumed, we then see that

$$(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^0 + \theta(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta)(\boldsymbol{X}^*, \boldsymbol{Y}^*) = \mathcal{F}^0 + (\boldsymbol{X}', \boldsymbol{Y}').$$

Hence $(\boldsymbol{X}' - \boldsymbol{X}, \boldsymbol{Y}' - \boldsymbol{Y}) \in \mathcal{F}^0$. By the monotonicity of the affine subspace $\mathcal{F}^0$, $\boldsymbol{X}' \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}' \leq \boldsymbol{X}' \bullet \boldsymbol{Y}' + \boldsymbol{X} \bullet \boldsymbol{Y}$. It follows that

$$\begin{aligned}
&\theta(\boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}^0) \\
&\leq \boldsymbol{X}' \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}' \text{ (since } \boldsymbol{X}' \succeq \theta \boldsymbol{X}^0 \text{ and } \boldsymbol{Y}' \succeq \theta \boldsymbol{Y}^0) \\
&\leq \boldsymbol{X}' \bullet \boldsymbol{Y}' + \boldsymbol{X} \bullet \boldsymbol{Y} \\
&\leq \theta^2 \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + \theta(1-\theta)(\boldsymbol{X}^0 \bullet \boldsymbol{Y}^* + \boldsymbol{X}^* \bullet \boldsymbol{Y}^0) + \boldsymbol{X} \bullet \boldsymbol{Y} \text{ (since } \boldsymbol{X}^* \bullet \boldsymbol{Y}^* = 0) \\
&\leq \theta^2 \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + 2\theta(1-\theta)\omega^* \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + \boldsymbol{X} \bullet \boldsymbol{Y} \text{ (by (48))} \\
&\leq \left(2\theta\omega^* \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 + \boldsymbol{X} \bullet \boldsymbol{Y}\right) \text{ (since } \omega^* \geq 1 \text{ and } \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \geq 0) \\
&\leq (2\omega^* \xi \boldsymbol{X} \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}) \text{ (by (46))} \\
&\leq (2\omega^* \xi + 1) \boldsymbol{X} \bullet \boldsymbol{Y}.
\end{aligned}$$

Thus we have shown assertion 1.

2. By assumption (47) and the Newton equation (23) which $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}})$ satisfies, we know that

$$\begin{aligned}
&-\theta\left((\boldsymbol{X}^0, \boldsymbol{Y}^0) - (\boldsymbol{X}^*, \boldsymbol{Y}^*)\right) \in \mathcal{F}^0 - (\boldsymbol{X}, \boldsymbol{Y}) + (\boldsymbol{X}^*, \boldsymbol{Y}^*), \\
&(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0 - (\boldsymbol{X}, \boldsymbol{Y}) + (\boldsymbol{X}^*, \boldsymbol{Y}^*).
\end{aligned}$$

Hence, letting

$$(d\hat{\boldsymbol{X}}', d\hat{\boldsymbol{Y}}') = (d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) + \theta\left((\boldsymbol{X}^0, \boldsymbol{Y}^0) - (\boldsymbol{X}^*, \boldsymbol{Y}^*)\right) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0,$$

we obtain the system of equations in the variable matrices $d\hat{\boldsymbol{X}}',\ d\hat{\boldsymbol{Y}}' \in \hat{\mathcal{S}}$:

$$\begin{cases} (d\hat{\boldsymbol{X}}', d\hat{\boldsymbol{Y}}') \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0, \\ \boldsymbol{X} d\hat{\boldsymbol{Y}}' + d\hat{\boldsymbol{X}}' \boldsymbol{Y} = \boldsymbol{Q}_1 + \boldsymbol{Q}_2 + \boldsymbol{Q}_3, \end{cases} \tag{49}$$

where $\boldsymbol{Q}_1 = \beta\mu \boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y},\ \boldsymbol{Q}_2 = \theta \boldsymbol{X}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*),\ \boldsymbol{Q}_3 = \theta(\boldsymbol{X}^0 - \boldsymbol{X}^*)\boldsymbol{Y}$. We note that the system (49) of equations has a unique solution in view of Corollary 4.3. Let $(d\hat{\boldsymbol{X}}_j, d\hat{\boldsymbol{Y}}_j)$ $(j = 1, 2, 3)$ denote the solution of the following system of equations:

$$(d\hat{\boldsymbol{X}}_j, d\hat{\boldsymbol{Y}}_j) \in \mathcal{F}^0 + \tilde{\mathcal{F}}^0 \text{ and } \boldsymbol{X} d\hat{\boldsymbol{Y}}_j + d\hat{\boldsymbol{X}}_j \boldsymbol{Y} = \boldsymbol{Q}_j.$$

Then the solution $(d\hat{\boldsymbol{X}}', d\hat{\boldsymbol{Y}}')$ of the system (49) of equations can be represented as

$$(d\hat{\boldsymbol{X}}', d\hat{\boldsymbol{Y}}') = (d\hat{\boldsymbol{X}}_1, d\hat{\boldsymbol{Y}}_1) + (d\hat{\boldsymbol{X}}_2, d\hat{\boldsymbol{Y}}_2) + (d\hat{\boldsymbol{X}}_3, d\hat{\boldsymbol{Y}}_3).$$

On the other hand, we see by the definition of $(d\hat{\boldsymbol{X}}', d\hat{\boldsymbol{Y}}')$ that

$$\begin{aligned}
d\hat{\boldsymbol{X}} &= d\hat{\boldsymbol{X}}' - \theta\left(\boldsymbol{X}^0 - \boldsymbol{X}^*\right) \\
&= d\hat{\boldsymbol{X}}_1 + d\hat{\boldsymbol{X}}_2 + d\hat{\boldsymbol{X}}_3 - \theta\left(\boldsymbol{X}^0 - \boldsymbol{X}^*\right) \\
&= d\hat{\boldsymbol{X}}_1 + d\hat{\boldsymbol{X}}_2 + \boldsymbol{Q}_3\boldsymbol{Y}^{-1} - \boldsymbol{X}d\hat{\boldsymbol{Y}}_3\boldsymbol{Y}^{-1} - \theta\left(\boldsymbol{X}^0 - \boldsymbol{X}^*\right) \\
&= d\hat{\boldsymbol{X}}_1 + d\hat{\boldsymbol{X}}_2 + \left(\theta(\boldsymbol{X}^0 - \boldsymbol{X}^*)\boldsymbol{Y}\right)\boldsymbol{Y}^{-1} - \boldsymbol{X}d\hat{\boldsymbol{Y}}_3\boldsymbol{Y}^{-1} - \theta\left(\boldsymbol{X}^0 - \boldsymbol{X}^*\right) \\
&= d\hat{\boldsymbol{X}}_1 + d\hat{\boldsymbol{X}}_2 - \boldsymbol{X}d\hat{\boldsymbol{Y}}_3\boldsymbol{Y}^{-1}, \\
d\hat{\boldsymbol{Y}} &= d\hat{\boldsymbol{Y}}' - \theta\left(\boldsymbol{Y}^0 - \boldsymbol{Y}^*\right) \\
&= d\hat{\boldsymbol{Y}}_1 + d\hat{\boldsymbol{Y}}_2 + d\hat{\boldsymbol{Y}}_3 - \theta\left(\boldsymbol{Y}^0 - \boldsymbol{Y}^*\right) \\
&= d\hat{\boldsymbol{Y}}_1 + \boldsymbol{X}^{-1}\boldsymbol{Q}_2 - \boldsymbol{X}^{-1}d\hat{\boldsymbol{X}}_2\boldsymbol{Y} + d\hat{\boldsymbol{Y}}_3 - \theta\left(\boldsymbol{Y}^0 - \boldsymbol{Y}^*\right) \\
&= d\hat{\boldsymbol{Y}}_1 + \boldsymbol{X}^{-1}\left(\theta\boldsymbol{X}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\right) - \boldsymbol{X}^{-1}d\hat{\boldsymbol{X}}_2\boldsymbol{Y} + d\hat{\boldsymbol{Y}}_3 - \theta\left(\boldsymbol{Y}^0 - \boldsymbol{Y}^*\right) \\
&= d\hat{\boldsymbol{Y}}_1 - \boldsymbol{X}^{-1}d\hat{\boldsymbol{X}}_2\boldsymbol{Y} + d\hat{\boldsymbol{Y}}_3.
\end{aligned}$$

It follows that

$$\begin{aligned}
\sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}\sqrt{\boldsymbol{Y}} &= \sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_1\sqrt{\boldsymbol{Y}} + \sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_2\sqrt{\boldsymbol{Y}} - \sqrt{\boldsymbol{X}}^{-1}(\boldsymbol{X}d\hat{\boldsymbol{Y}}_3\boldsymbol{Y}^{-1})\sqrt{\boldsymbol{Y}} \\
&= \sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_1\sqrt{\boldsymbol{Y}} + \sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_2\sqrt{\boldsymbol{Y}} - \sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_3\sqrt{\boldsymbol{Y}}^{-1}, \\
\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1} &= \sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_1\sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}(\boldsymbol{X}^{-1}d\hat{\boldsymbol{X}}_2\boldsymbol{Y})\sqrt{\boldsymbol{Y}}^{-1} + \sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_3\sqrt{\boldsymbol{Y}}^{-1} \\
&= \sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_1\sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_2\sqrt{\boldsymbol{Y}} + \sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_3\sqrt{\boldsymbol{Y}}^{-1}.
\end{aligned}$$

Hence, by item 3 of Lemma 7.1 and the definition of $\boldsymbol{H}(\beta)$,

$$\begin{aligned}
&\|\sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F \\
&\le \|\sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_1\sqrt{\boldsymbol{Y}}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_2\sqrt{\boldsymbol{Y}}\|_F + \|\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\le \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_1\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\le \|\boldsymbol{H}(\beta)\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F, \\
&\|\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\le \|\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_1\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}_2\sqrt{\boldsymbol{Y}}\|_F + \|\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\le \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_1\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\le \|\boldsymbol{H}(\beta)\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F.
\end{aligned}$$

Therefore we have shown

$$\begin{aligned}
&\|\sqrt{\boldsymbol{X}}^{-1}d\hat{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F,\ \|\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\quad\le \|\boldsymbol{H}(\beta)\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F.
\end{aligned} \tag{50}$$

We now evaluate $\|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F$.

$$\|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F$$

$$
\begin{aligned}
&= \left\| \sqrt{\boldsymbol{X}}^{-1} \left(\theta \boldsymbol{X}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\right) \sqrt{\boldsymbol{Y}}^{-1} \right\|_F \\
&= \theta \left(\mathrm{Tr}\ (\sqrt{\boldsymbol{X}}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\sqrt{\boldsymbol{X}})(\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}})^{-1}(\sqrt{\boldsymbol{X}}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\sqrt{\boldsymbol{X}}) \right)^{1/2} \\
&\leq \frac{\theta}{\sqrt{\lambda_{min}}} \|\sqrt{\boldsymbol{X}}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\sqrt{\boldsymbol{X}}\|_F \\
&= \frac{\theta\omega^*}{\sqrt{\lambda_{min}}} \boldsymbol{X} \bullet \boldsymbol{Y}^0 \text{ (by Lemma 7.5).}
\end{aligned}
$$

Thus we have shown that

$$
\|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F \leq \frac{\theta\omega^*}{\sqrt{\lambda_{min}}}\boldsymbol{X} \bullet \boldsymbol{Y}^0. \tag{51}
$$

Similarly we can prove that

$$
\|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F \leq \frac{\theta\omega^*}{\sqrt{\lambda_{min}}}\boldsymbol{X}^0 \bullet \boldsymbol{Y}. \tag{52}
$$

Therefore

$$
\begin{aligned}
&\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F,\ \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1}\|_F \\
&\leq \|\boldsymbol{H}(\beta)\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_2\sqrt{\boldsymbol{Y}}^{-1}\|_F + \|\sqrt{\boldsymbol{X}}^{-1}\boldsymbol{Q}_3\sqrt{\boldsymbol{Y}}^{-1}\|_F \quad \text{(by (50))} \\
&\leq \|\boldsymbol{H}(\beta)\|_F + \frac{\theta\omega^*}{\sqrt{\lambda_{min}}}(\boldsymbol{X} \bullet \boldsymbol{Y}^0 + \boldsymbol{X}^0 \bullet \boldsymbol{Y}) \quad \text{(by (51) and (52))} \\
&\leq \|\boldsymbol{H}(\beta)\|_F + \frac{\omega^*\sigma n\mu}{\sqrt{\lambda_{min}}} \quad \text{(by assertion 1)} \\
&\leq \frac{\sqrt{2n}\mu}{\sqrt{\lambda_{min}}} + \frac{\omega^*\sigma n\mu}{\sqrt{\lambda_{min}}} \quad \text{(by Lemma 7.2)} \\
&\leq \frac{\zeta n\mu}{\sqrt{\lambda_{min}}}.
\end{aligned}
$$

3. The assertion follows from assertion 2 of both Lemma 7.6 and 7.1. □

By the argument above, it is easily seen that if $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}_{++}$ then we can take $\theta = 0$, which implies that $\boldsymbol{Q}_2 = \boldsymbol{Q}_3 = \boldsymbol{O}$. Hence we can obtain the better evaluation of $\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F,\ \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1}\|_F$ in assertion 2 of Lemma 7.6. We state this result as a corollary.

COROLLARY 7.7. *Suppose that all the assumptions of Lemma* 7.6 *hold. In addition, let* $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}_{++}$. *Then* $\|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}} \sqrt{\boldsymbol{Y}}\|_F,\ \|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{Y}}^{-1}\|_F \leq \|\boldsymbol{H}(\beta)\|_F$.

8. Some interior-point methods.

In this section we present three types of interior-point methods, a central trajectory following method, a potential-reduction method, and an infeasible interior-point potential-reduction method as special cases of the generic IP method.

8.1. A central trajectory following method.

This method is based on the $O(\sqrt{n}L)$ iteration interior-point method proposed by Kojima–Mizuno–Yoshise [21] for the monotone LCP (8) in the Euclidean space. A horn neighborhood of the central trajectory is defined as

$$
\mathcal{N}(\gamma) = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : \begin{array}{l} \left(\displaystyle\sum_{j=1}^{n} (\lambda_j - \mu)^2 \right)^{1/2} \leq \gamma\mu, \quad \text{where } \mu = \dfrac{\boldsymbol{X} \bullet \boldsymbol{Y}}{n}, \\ \text{and } \lambda_1, \ldots, \lambda_n \text{ denote the eigenvalues of } \boldsymbol{X}\boldsymbol{Y} \end{array} \right\}
$$

$$= \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++} : \|\sqrt{\boldsymbol{X}}\boldsymbol{Y}\sqrt{\boldsymbol{X}} - \mu \boldsymbol{I}\|_F \le \gamma\mu, \quad \text{where } \mu = \frac{\boldsymbol{X} \bullet \boldsymbol{Y}}{n} \right\}.$$

Here $\gamma > 0$ denotes a parameter which determines the width of the neighborhood $\mathcal{N}(\gamma)$. We obtain the central trajectory following method by imposing the following additional restrictions on the generic IP method:

- Let $\tilde{\mathcal{F}}^0 = \boldsymbol{O} \times \tilde{\mathcal{S}}$.
- Let $0 < \gamma \le 0.1$. Choose an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{N}(\gamma)$ in Step 0.
- Let $\beta = 1 - \gamma/\sqrt{n}$ in Step 2.
- Let $\alpha = 1$ in Step 5.

The first restriction needs some explanation. When we compute the solution $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$ of the Newton equation (26) with $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$ in Step 3, the restriction above implies $d\hat{\boldsymbol{X}}$ is symmetric so that $d\boldsymbol{X} = (d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2 = d\hat{\boldsymbol{X}}$ in Step 4. The fact that $d\hat{\boldsymbol{X}}$ is symmetric is necessary in our proof of the lemma below. The authors tried to employ a general $n(n-1)/2$-dimensional monotone linear subspace $\tilde{\mathcal{F}}^0$ of $\tilde{\mathcal{S}}^2$, but we had some difficulty deriving the inequality (55) in the general case.

THEOREM 8.1. *Let* $\gamma \in (0,\ 0.1]$. *Suppose that* $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{N}(\gamma)$. *Let* $\beta = 1 - \gamma/\sqrt{n}$ *in Step* 2 *and* $\alpha = 1$ *in Step* 5. *Let* $\bar{\mu} = \bar{\boldsymbol{X}} \bullet \bar{\boldsymbol{Y}}/n$. *Then*

$$(53) \qquad \begin{aligned} &(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) = (\boldsymbol{X}, \boldsymbol{Y}) + (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{N}(\gamma), \\ &\beta\mu \le \bar{\mu} \le \left(1 - \frac{\gamma}{2\sqrt{n}}\right)\mu. \end{aligned}$$

Proof. First note that $d\hat{\boldsymbol{X}} = d\boldsymbol{X}$. By the definition of $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}})$,

$$\begin{aligned} n\bar{\mu} &= \bar{\boldsymbol{X}} \bullet \bar{\boldsymbol{Y}} \\ &= \boldsymbol{X} \bullet \boldsymbol{Y} + \operatorname{Tr}\,(\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y}) + d\boldsymbol{X} \bullet d\boldsymbol{Y} \\ &= \boldsymbol{X} \bullet \boldsymbol{Y} + \operatorname{Tr}\,(\boldsymbol{X}d\hat{\boldsymbol{Y}} + d\boldsymbol{X}\boldsymbol{Y}) + d\boldsymbol{X} \bullet d\boldsymbol{Y} \\ &= n\mu + \operatorname{Tr}\,(\beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}) + d\boldsymbol{X} \bullet d\boldsymbol{Y} \ \text{(by the Newton equation (26))} \\ &= n\mu + n\beta\mu - n\mu + d\boldsymbol{X} \bullet d\boldsymbol{Y} \ \text{(since } \mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n) \\ &= n\beta\mu + d\boldsymbol{X} \bullet d\boldsymbol{Y}. \end{aligned}$$

In view of item 3 of Lemma 7.1 with $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$ and item 2 of Lemma 7.2, we see that

$$0 \le d\boldsymbol{X} \bullet d\boldsymbol{Y} \le \frac{\|\boldsymbol{H}(\beta)\|_F^2}{4} \le \frac{\gamma^2\mu}{1-\gamma}.$$

Hence

$$\beta\mu \le \bar{\mu} = \beta\mu + \frac{d\boldsymbol{X} \bullet d\boldsymbol{Y}}{n} \le \left(1 - \frac{\gamma}{2\sqrt{n}}\right)\mu.$$

Thus we have shown (53).

By Lemma 7.1 with $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$ and Lemma 7.2,

$$\sum_{j=1}^{n}(\xi_j^2 + \eta_j^2) \le \frac{1}{\lambda_{min}}\|\boldsymbol{H}(\beta)\|_F^2 \le \frac{4\gamma^2}{(1-\gamma)^2} < 1.$$

Hence Lemma 5.1 together with $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}$ and $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}^0$ ensure that

$$(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) = (\boldsymbol{X}, \boldsymbol{Y}) + (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}_{++}. \tag{54}$$

To complete the proof, we only need to show the inequality

$$\|\bar{\boldsymbol{B}}^T \bar{\boldsymbol{Y}} \bar{\boldsymbol{B}} - \bar{\mu}\boldsymbol{I}\|_F \leq \gamma\bar{\mu} \tag{55}$$

for some $\bar{\boldsymbol{B}}$ such that $\bar{\boldsymbol{X}} = \bar{\boldsymbol{B}}\bar{\boldsymbol{B}}^T$, since (54) and (55) imply that $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) \in \mathcal{N}(\gamma)$. By Lemma 7.1, with $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$, Lemma 7.2, and $0 < \gamma \leq 0.1$,

$$\|\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}\|_F \leq \frac{1}{\sqrt{(1-\gamma)\mu}} \cdot \frac{2\gamma\sqrt{\mu}}{\sqrt{(1-\gamma)}} = \frac{2\gamma}{1-\gamma} < 1, \tag{56}$$

$$\|\sqrt{\boldsymbol{X}} d\hat{\boldsymbol{Y}} \sqrt{\boldsymbol{X}}\|_F \leq \sqrt{(1+\gamma)\mu} \cdot \frac{2\gamma\sqrt{\mu}}{\sqrt{(1-\gamma)}} = \frac{2\mu\gamma\sqrt{1+\gamma}}{\sqrt{1-\gamma}}. \tag{57}$$

We see from the definition of $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}})$ that

$$\bar{\boldsymbol{X}} = \boldsymbol{X} + d\boldsymbol{X} = \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{X}} + d\boldsymbol{X} = \sqrt{\boldsymbol{X}}(\boldsymbol{I} + \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1})\sqrt{\boldsymbol{X}}.$$

Since the inequality (56) implies that the absolute values of all the eigenvalues of the symmetric matrix $\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}$ are less than 1, the symmetric matrix $\boldsymbol{I} + \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1}$ is positive definite. Hence it can be represented as $\boldsymbol{I} + \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1} = \boldsymbol{P}\boldsymbol{\Xi}\boldsymbol{P}^T$ for an orthogonal matrix $\boldsymbol{P}$ and a diagonal matrix $\boldsymbol{\Xi} = \operatorname{diag}(\bar{\xi}_1, \bar{\xi}_2, \ldots, \bar{\xi}_n)$ of its eigenvalues $\bar{\xi}_1, \bar{\xi}_2, \ldots, \bar{\xi}_n$ such that

$$\frac{1-3\gamma}{1-\gamma} = 1 - \frac{2\gamma}{1-\gamma} \leq \bar{\xi}_j \leq 1 + \frac{2\gamma}{1-\gamma} = \frac{1+\gamma}{1-\gamma} \text{ for every } j = 1, 2, \ldots, n. \tag{58}$$

Letting $\bar{\boldsymbol{B}} = \sqrt{\boldsymbol{X}}\boldsymbol{P}\sqrt{\boldsymbol{\Xi}}$, we obtain

$$\bar{\boldsymbol{X}} = \sqrt{\boldsymbol{X}}(\boldsymbol{I} + \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X} \sqrt{\boldsymbol{X}}^{-1})\sqrt{\boldsymbol{X}} = \sqrt{\boldsymbol{X}}\boldsymbol{P}\boldsymbol{\Xi}\boldsymbol{P}^T\sqrt{\boldsymbol{X}} = \bar{\boldsymbol{B}}\bar{\boldsymbol{B}}^T.$$

We also see from the Newton equation (26) that

$$\bar{\boldsymbol{X}}(\boldsymbol{Y} + d\hat{\boldsymbol{Y}}) - \beta\mu\boldsymbol{I} = (\boldsymbol{X} + d\boldsymbol{X})(\boldsymbol{Y} + d\hat{\boldsymbol{Y}}) - \beta\mu\boldsymbol{I} = d\boldsymbol{X} d\hat{\boldsymbol{Y}};$$

hence $\bar{\boldsymbol{B}}^T(\boldsymbol{Y} + d\hat{\boldsymbol{Y}})\bar{\boldsymbol{B}} - \beta\mu\boldsymbol{I} = \bar{\boldsymbol{B}}^{-1} d\boldsymbol{X} d\hat{\boldsymbol{Y}} \bar{\boldsymbol{B}}$. Now we are ready to evaluate $\|\bar{\boldsymbol{B}}^T \bar{\boldsymbol{Y}} \bar{\boldsymbol{B}} - \bar{\mu}\boldsymbol{I}\|_F$ to derive the inequality (55):

$$\begin{aligned}
&\|\bar{\boldsymbol{B}}^T \bar{\boldsymbol{Y}} \bar{\boldsymbol{B}} - \bar{\mu}\boldsymbol{I}\|_F \\
&\leq \|\bar{\boldsymbol{B}}^T \bar{\boldsymbol{Y}} \bar{\boldsymbol{B}} - \beta\mu\boldsymbol{I}\|_F \ (\text{since } \|\bar{\boldsymbol{B}}^T \bar{\boldsymbol{Y}} \bar{\boldsymbol{B}} - \bar{\mu}\boldsymbol{I}\|_F = \min\nolimits_{\nu \in R} \|\bar{\boldsymbol{B}}^T \bar{\boldsymbol{Y}} \bar{\boldsymbol{B}} - \nu\boldsymbol{I}\|_F) \\
&= \left\|\bar{\boldsymbol{B}}^T \left(\frac{\boldsymbol{Y} + d\hat{\boldsymbol{Y}} + \boldsymbol{Y}^T + d\hat{\boldsymbol{Y}}^T}{2}\right) \bar{\boldsymbol{B}} - \beta\mu\boldsymbol{I}\right\|_F \\
&\leq \frac{\|\bar{\boldsymbol{B}}^T(\boldsymbol{Y} + d\hat{\boldsymbol{Y}})\bar{\boldsymbol{B}} - \beta\mu\boldsymbol{I}\|_F}{2} + \frac{\|\bar{\boldsymbol{B}}^T(\boldsymbol{Y} + d\hat{\boldsymbol{Y}})^T\bar{\boldsymbol{B}} - \beta\mu\boldsymbol{I}\|_F}{2} \\
&= \|\bar{\boldsymbol{B}}^T(\boldsymbol{Y} + d\hat{\boldsymbol{Y}})\bar{\boldsymbol{B}} - \beta\mu\boldsymbol{I}\|_F \\
&= \|\bar{\boldsymbol{B}}^{-1} d\boldsymbol{X} d\hat{\boldsymbol{Y}} \bar{\boldsymbol{B}}\|_F
\end{aligned}$$

$$
\begin{aligned}
&= \|\sqrt{\boldsymbol{\Xi}}^{-1}\boldsymbol{P}^T\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{X}}\boldsymbol{P}\sqrt{\boldsymbol{\Xi}}\|_F \text{ (since } \bar{\boldsymbol{B}} = \sqrt{\boldsymbol{X}}\boldsymbol{P}\sqrt{\boldsymbol{\Xi}}) \\
&\le \frac{\sqrt{(1+\gamma)/(1-\gamma)}}{\sqrt{(1-3\gamma)/(1-\gamma)}}\|\boldsymbol{P}^T\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{X}}\boldsymbol{P}\|_F \text{ (by (58))} \\
&\le \frac{\sqrt{1+\gamma}}{\sqrt{1-3\gamma}}\|\sqrt{\boldsymbol{X}}^{-1}d\boldsymbol{X}\sqrt{\boldsymbol{X}}^{-1}\|_F\cdot\|\sqrt{\boldsymbol{X}}d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{X}}\|_F \\
&\le \frac{\sqrt{1+\gamma}}{\sqrt{1-3\gamma}}\cdot\frac{2\gamma}{1-\gamma}\cdot\frac{2\mu\gamma\sqrt{1+\gamma}}{\sqrt{1-\gamma}} \quad \text{(by (56) and (57))} \\
&= \frac{4\gamma^2(1+\gamma)\mu}{(1-\gamma)^{3/2}\sqrt{1-3\gamma}} \\
&\le \frac{4\gamma^2(1+\gamma)\bar{\mu}}{(1-\gamma)^{3/2}\sqrt{1-3\gamma}\beta} \quad \text{(since } \beta\mu \le \bar{\mu} \text{ by (53))} \\
&\le \frac{4\gamma^2(1+\gamma)}{(1-\gamma)^{3/2}\sqrt{1-3\gamma}(1-\gamma)}\bar{\mu} \quad \text{(since } 1-\gamma \le \beta).
\end{aligned}
$$

Hence

$$\|\bar{\boldsymbol{B}}^T\bar{\boldsymbol{Y}}\bar{\boldsymbol{B}} - \bar{\mu}\boldsymbol{I}\|_F \le \frac{4\gamma(1+\gamma)}{(1-\gamma)^{5/2}\sqrt{1-3\gamma}}\gamma\bar{\mu} \le \gamma\bar{\mu}.$$

Here the last inequality follows from $\gamma \in (0, 0.1]$. Thus we have shown the inequality (55). $\square$

Let $\epsilon > 0$. In view of the theorem above, the central trajectory following method generates a sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ such that

$$(\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{N}(\gamma) \text{ and } \boldsymbol{X}^r \bullet \boldsymbol{Y}^r \le \left(1 - \frac{\gamma}{2\sqrt{n}}\right)^r \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0$$

for every $r = 0, 1, \ldots$. Hence if

$$r \ge \frac{2\sqrt{n}}{\gamma}\log\frac{\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0}{\epsilon} = O\left(\sqrt{n}\log\frac{\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0}{\epsilon}\right)$$

then $(\boldsymbol{X}^r, \boldsymbol{Y}^r)$ gives an approximate solution of the SDLCP (1) in symmetric matrices such that

$$(\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{F}_{++} \text{ and } \boldsymbol{X}^r \bullet \boldsymbol{Y}^r \le \epsilon. \tag{59}$$

8.2. A potential-reduction method. For every $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{++}$, define the potential function

$$f(\boldsymbol{X}, \boldsymbol{Y}) = (n+\nu)\log\boldsymbol{X}\bullet\boldsymbol{Y} - \log\det\boldsymbol{X}\boldsymbol{Y} - n\log n.$$

Here $\nu \ge 0$ is a parameter. This potential function is the same as the one used in the paper [44] by Vandenberghe and Boyd. Our potential-reduction method described below is different from their method in search directions. Our method may be regarded as an extension of the Kojima–Mizuno–Yoshise potential-reduction method [22] for the monotone LCP (8) in the Euclidean space (see also [19]).

The potential function f defined above enjoys similar properties as the potential function [40, 42] used for the monotone LCP (8). In particular, if we rewrite the potential function f as

$$\begin{aligned} f(\boldsymbol{X},\boldsymbol{Y}) &= \nu f_{cp}(\boldsymbol{X},\boldsymbol{Y}) + f_{cen}(\boldsymbol{X},\boldsymbol{Y}), \\ f_{cp}(\boldsymbol{X},\boldsymbol{Y}) &= \log \boldsymbol{X}\bullet\boldsymbol{Y}, \\ f_{cen}(\boldsymbol{X},\boldsymbol{Y}) &= n\log \boldsymbol{X}\bullet\boldsymbol{Y} - \log\det \boldsymbol{X}\boldsymbol{Y} - n\log n, \end{aligned}$$

we have

$$\begin{aligned} & f_{cen}(\boldsymbol{X},\boldsymbol{Y}) \geq 0 \text{ for every } (\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F}_{++}, \\ & f_{cen}(\boldsymbol{X},\boldsymbol{Y}) = 0 \text{ if and only if } (\boldsymbol{X},\boldsymbol{Y}) \text{ lies in the central trajectory } \mathcal{C}. \end{aligned} \tag{60}$$

See the paper [44].

We impose the following restrictions on the generic IP method:

- Choose an initial point $(\boldsymbol{X}^0,\boldsymbol{Y}^0) \in \mathcal{F}_{++}$ in Step 0.
- Let $\beta = n/(n+\nu)$ in Step 2.
- Let

$$\left\{ \begin{array}{rcl} 0 &<& \tau \;<\; 1, \\ \boldsymbol{H}(\beta) &=& \beta\mu\sqrt{\boldsymbol{X}}^{-1}\sqrt{\boldsymbol{Y}}^{-1} - \sqrt{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}, \\ \lambda_{min} &=& \min\{\lambda_1,\lambda_2,\ldots,\lambda_n\}, \end{array} \right. \tag{61}$$

  where $\lambda_1,\lambda_2,\ldots,\lambda_n$ denote the eigenvalues of the matrix $\boldsymbol{X}\boldsymbol{Y}$. Take a step size parameter α in Step 5 such that $\alpha = \tau\sqrt{\lambda_{min}}/\|\boldsymbol{H}(\beta)\|_F$.

We remark here that if $\boldsymbol{X} = \boldsymbol{L}\boldsymbol{L}^T$, $\boldsymbol{L} \in \hat{\mathcal{S}}$, $\boldsymbol{Y} = \boldsymbol{M}\boldsymbol{M}^T$, and $\boldsymbol{M} \in \hat{\mathcal{S}}$, then

$$\|\boldsymbol{H}(\beta)\|_F = \|\beta\mu\boldsymbol{L}^{-1}\boldsymbol{M}^{-T} - \boldsymbol{L}^T\boldsymbol{M}\|_F.$$

This makes the computation of the step length $\alpha = \tau\sqrt{\lambda_{min}}/\|\boldsymbol{H}(\beta)\|_F$ more flexible and efficient.

THEOREM 8.2. *Let $n \geq 3$, $\nu = \sqrt{n}$, $\tau = 0.4$, and $\delta = 0.2$. Suppose that $(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F}_{++}$. Let $\beta = n/(n+\nu)$ in Step 2 and $\alpha = \tau\sqrt{\lambda_{min}}/\|\boldsymbol{H}(\beta)\|_F$ in Step 5, where τ, λ_{min}, and $\boldsymbol{H}(\beta)$ are given in (61). Then*

$$(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}}) = (\boldsymbol{X},\boldsymbol{Y}) + \alpha(d\boldsymbol{X},d\boldsymbol{Y}) \in \mathcal{F}_{++} \text{ and } f(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}}) \leq f(\boldsymbol{X},\boldsymbol{Y}) - \delta.$$

Proof. Let $\xi_1,\xi_2,\ldots,\xi_n$ be the eigenvalues of $\boldsymbol{X}^{-1}d\boldsymbol{X}$ and $\eta_1,\eta_2,\ldots,\eta_n$ be the eigenvalues of $\boldsymbol{Y}^{-1}d\boldsymbol{Y}$. By Lemma 7.1 and $\alpha = \tau\sqrt{\lambda_{min}}/\|\boldsymbol{H}(\beta)\|_F$,

$$\sum_{j=1}^{n}\left((\alpha\xi_j)^2 + (\alpha\eta_j)^2\right) \leq \alpha^2 \cdot \frac{1}{\lambda_{min}}\|\boldsymbol{H}(\beta)\|_F^2 = \tau^2.$$

Hence $|\alpha\xi_j| \leq \tau = 0.4$ and $|\alpha\eta_j| \leq \tau = 0.4$ for every $j = 1,2,\ldots,n$. By Lemma 5.1, $(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F}$ and $(d\boldsymbol{X},d\boldsymbol{Y}) \in \mathcal{F}^0$, we obtain that

$$(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}}) = (\boldsymbol{X},\boldsymbol{Y}) + \alpha(d\boldsymbol{X},d\boldsymbol{Y}) \in \mathcal{F}_{++}.$$

It follows from $\tau = 0.4$, $\nu = \sqrt{n}$, and $n \geq 3$ that $(n+\sqrt{n})/n < 1/(1-\tau)$. By Lemma 7.4,

$$G_2(d\boldsymbol{X},d\boldsymbol{Y}) \leq \frac{\|\boldsymbol{H}(\beta)\|_F^2}{2(1-\tau)\lambda_{min}}. \tag{62}$$

Thus we consequently obtain

$$\begin{aligned}
&f(\boldsymbol{X}+\alpha d\boldsymbol{X}, \boldsymbol{Y}+\alpha d\boldsymbol{Y}) - f(\boldsymbol{X}, \boldsymbol{Y}) \\
&\leq \alpha G_1(d\boldsymbol{X}, d\boldsymbol{Y}) + \alpha^2 G_2(d\boldsymbol{X}, d\boldsymbol{Y}) \quad \text{(by 1 of Lemma 7.4)} \\
&\leq -\frac{1}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 \alpha + \frac{\|\boldsymbol{H}(\beta)\|^2}{2(1-\tau)\lambda_{min}}\alpha^2 \quad \text{(by (42) and (62))} \\
&= -\frac{\tau\sqrt{\lambda_{min}}}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F + \frac{\tau^2}{2(1-\tau)} \quad (\text{since } \alpha = \tau\sqrt{\lambda_{min}}/\|\boldsymbol{H}(\beta)\|_F) \\
&\leq -\frac{\sqrt{3}\tau}{2} + \frac{\tau^2}{2(1-\tau)} \quad \text{(by 2 of Lemma 7.4)} \\
&\leq -0.2. \ (\text{since } \tau = 0.4).
\end{aligned}$$

This completes the proof of Theorem 8.2. ∎

Let $\epsilon > 0$. By Theorem 8.2, the potential-reduction method generates a sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r)\}$ such that $(\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{F}_{++}$ and $f(\boldsymbol{X}^r, \boldsymbol{Y}^r) \leq f(\boldsymbol{X}^0, \boldsymbol{Y}^0) - r\delta$ for every $r = 1, 2, \ldots$. Hence if $r \geq (f(\boldsymbol{X}^0, \boldsymbol{Y}^0) - \sqrt{n}\log\epsilon)/\delta$, then $(\boldsymbol{X}^r, \boldsymbol{Y}^r)$ gives an approximate solution of the SDLCP (1) satisfying (59). If in addition $f_{cen}(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ is bounded by a constant independent of n, the right-hand side of the inequality above is of $O(\sqrt{n}\log(\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0/\epsilon))$, the same order as the one in the case of the central trajectory following method described in section 8.1.

8.3. An infeasible interior-point potential-reduction method. The IIP potential-reduction method presented below is based on the $O(n^{2.5}L)$ iteration constrained potential-reduction algorithm (Algorithm I of [32]) for linear programs and its modification [17]. We can start the IIP potential-reduction method from any $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$. Before running the method, we assume that the hypothesis given in section 7 holds.

As we will see below, the IIP potential-reduction method either detects in a finite number of iterations that the hypothesis is false (i.e., there is no solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP (1) satisfying (45)) or reduces the potential function by at least a given constant δ at every iteration.

We impose some additional requirements on Steps 0, 2, and 5 of the generic IP method to describe the IIP potential-reduction method:

IIP potential-reduction method.

Step 0_{iip}: Choose an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}^2_{++}$ and two parameters $\nu \geq \sqrt{n}$, $\xi \geq 1$. Let $\sigma = 2\omega^*\xi + 1, \zeta = 2 + \omega^*\sigma$, and $\delta = 1/(10\zeta^2(n+\nu)^2)$. Let $\theta^0 = 1$ and $r = 0$.

Step 1: Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^r, \boldsymbol{Y}^r)$ and

$$\mu = \frac{\boldsymbol{X} \bullet \boldsymbol{Y}}{n}.$$

Step 2_{iip}: If

$$\theta^r(\boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}^0) \leq \sigma \boldsymbol{X} \bullet \boldsymbol{Y} \tag{63}$$

does not hold then stop; in this case there is no solution of the SDLCP (1) satisfying (45) (see item 2 of Theorem 8.3). Otherwise let $\beta = n/(n+\nu)$.

Step 3: Compute a solution $(d\hat{\boldsymbol{X}}, d\hat{\boldsymbol{Y}}) \in \hat{\mathcal{S}}^2$ of the system (23) of equations with $\boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}$.

Step 4: Let $d\boldsymbol{X} = (d\hat{\boldsymbol{X}} + d\hat{\boldsymbol{X}}^T)/2$ and $d\boldsymbol{Y} = (d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{Y}}^T)/2$.

Step 5_{iip}: Choose a step size parameter $\alpha \geq 0$ such that

$$(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) = (\boldsymbol{X}, \boldsymbol{Y}) + \alpha(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{S}^2_{++}, \tag{64}$$

$$(1-\alpha)\theta^r \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \leq \xi \bar{\boldsymbol{X}} \bullet \bar{\boldsymbol{Y}}, \tag{65}$$

$$f(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) \leq f(\boldsymbol{X}, \boldsymbol{Y}) - \delta. \tag{66}$$

Let $(\boldsymbol{X}^{r+1}, \boldsymbol{Y}^{r+1}) = (\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}})$ and $\theta^{r+1} = (1-\alpha)\theta^r$.

Step 6: Replace r by $r+1$ and go to Step 1.

THEOREM 8.3. *Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{S}^2_{++}$ be the rth iterate generated by the IIP potential-reduction method. Let $\mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n$ and let λ_{min} denote the minimum eigenvalue of the matrix $\boldsymbol{X}\boldsymbol{Y}$.*

1. *Let $(\boldsymbol{X}', \boldsymbol{Y}')$ be a pair of matrices in $\mathcal{F}$; for example, we can take an orthogonal projection of $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ onto $\mathcal{F}$, or a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP* (1) *satisfying* (45) *under the hypothesis in section 7. Then*

$$\theta^r \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \leq \xi \boldsymbol{X} \bullet \boldsymbol{Y}, \tag{67}$$

$$(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} + \theta^r \left((\boldsymbol{X}^0, \boldsymbol{Y}^0) - (\boldsymbol{X}', \boldsymbol{Y}') \right). \tag{68}$$

2. *Assume that inequality* (63) *holds at Step 2_{iip}. Then $\alpha = \lambda_{min}/(5\zeta^2(n+\nu)n\mu)$ fulfills all the requirements* (64), (65), *and* (66) *in Step 5_{iip} for a legitimate step size parameter.*

3. *If the inequality* (63) *does not hold at Step 2_{iip} then there is no solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP* (1) *satisfying* (45).

In order to prove the assertion 1, we need the following lemma.

LEMMA 8.4. *Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{S}^2_{++}$ be the rth iterate generated by the IIP potential-reduction method, and let $\theta = \theta^r$. Let $(\boldsymbol{X}', \boldsymbol{Y}')$ be a pair of matrices in $\mathcal{F}$. Then*

$$\theta \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \leq \xi \boldsymbol{X} \bullet \boldsymbol{Y}, \tag{69}$$

$$(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}^0 + \theta(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta)(\boldsymbol{X}', \boldsymbol{Y}'). \tag{70}$$

Proof. By the construction, the inequality (69) holds with $\theta = \theta^r \in [0,1]$. Hence it suffices to show by induction that $(\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{F}^0 + \theta^r(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta^r)(\boldsymbol{X}', \boldsymbol{Y}')$. When $r = 0$, the relation above obviously holds because $\theta^0 = 1$. Assume that the relation holds for $r = k$. Then

$$\begin{aligned}
&(\boldsymbol{X}^{k+1}, \boldsymbol{Y}^{k+1}) \\
&= (1-\alpha)(\boldsymbol{X}^k, \boldsymbol{Y}^k) + \alpha(\boldsymbol{X}^k + d\boldsymbol{X}, \boldsymbol{Y}^k + d\boldsymbol{Y}) \\
&\in (1-\alpha)\left(\mathcal{F}^0 + \theta^k(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta^k)(\boldsymbol{X}', \boldsymbol{Y}')\right) + \alpha\left(\mathcal{F}^0 + (\boldsymbol{X}', \boldsymbol{Y}')\right) \\
&= \mathcal{F}^0 + (1-\alpha)\theta^k(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-(1-\alpha)\theta^k)(\boldsymbol{X}', \boldsymbol{Y}') \\
&= \mathcal{F}^0 + \theta^{k+1}(\boldsymbol{X}^0, \boldsymbol{Y}^0) + (1-\theta^{k+1})(\boldsymbol{X}', \boldsymbol{Y}').
\end{aligned}$$

Thus we have shown the desired relation for $r = k+1$. □

Proof of Theorem 8.3. Assertion 1 of the theorem follows from Lemma 8.4 since we can rewrite the relation (70) as the relation (68).

To prove assertion 2 of the theorem, let $\alpha = \lambda_{min}/(5\zeta^2 n(n+\nu)\mu)$. By the definition, $\zeta = 2 + \omega^*\sigma \geq 4$, so that we see $\alpha = \lambda_{min}/(5\zeta^2 n(n+\nu)\mu) \leq 1/80$. By Lemma 7.6, we then see that for every $j = 1, 2, \ldots, n$

$$|\alpha\xi_j|,\ |\alpha\eta_j| \leq \alpha \cdot \frac{\zeta n\mu}{\lambda_{min}} = \frac{\lambda_{min}}{5\zeta^2 n(n+\nu)\mu} \cdot \frac{\zeta n\mu}{\lambda_{min}} \leq \frac{1}{20}.$$

Hence, letting $\tau = 1/20$, we obtain that $|\alpha\xi_j|$, $|\alpha\eta_j| \le \tau$ $(j = 1, 2, \ldots, n)$. By Lemma 5.1, $\boldsymbol{X} + \alpha d\boldsymbol{X} \in \mathcal{S}_{++}$ and $\boldsymbol{Y} + \alpha d\boldsymbol{Y} \in \mathcal{S}_{++}$. Thus we have shown the relation (64).

To derive the inequality (65) we observe that

$$
\begin{aligned}
&\xi(\boldsymbol{X} + \alpha d\boldsymbol{X}) \bullet (\boldsymbol{Y} + \alpha d\boldsymbol{Y}) \\
&= \xi\left(\boldsymbol{X} \bullet \boldsymbol{Y} + \alpha \mathrm{Tr}\ (\boldsymbol{X} d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y}) + \alpha^2 d\boldsymbol{X} \bullet d\boldsymbol{Y}\right) \\
&= \xi\left(\boldsymbol{X} \bullet \boldsymbol{Y} + \alpha \mathrm{Tr}\ (\boldsymbol{X} d\hat{\boldsymbol{Y}} + d\hat{\boldsymbol{X}}\boldsymbol{Y}) + \alpha^2 d\boldsymbol{X} \bullet d\boldsymbol{Y}\right) \\
&\quad \text{(since } \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are symmetric)} \\
&= \xi\left(\boldsymbol{X} \bullet \boldsymbol{Y} + \alpha \mathrm{Tr}\ (\beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}) + \alpha^2 \mathrm{Tr}\ \sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X}\sqrt{\boldsymbol{Y}}\sqrt{\boldsymbol{Y}}^{-1} d\boldsymbol{Y}\sqrt{\boldsymbol{X}}\right) \\
&\quad \text{(by the Newton equation (23))} \\
&\ge \xi\left((1 - (1-\beta)\alpha)\boldsymbol{X} \bullet \boldsymbol{Y} - \alpha^2 \|\sqrt{\boldsymbol{X}}^{-1} d\boldsymbol{X}\sqrt{\boldsymbol{Y}}\|_F \cdot \|\sqrt{\boldsymbol{Y}}^{-1} d\boldsymbol{Y}\sqrt{\boldsymbol{X}}\|_F\right) \\
&\quad \text{(since } n\mu = \boldsymbol{X} \bullet \boldsymbol{Y}) \\
&\ge \xi\left((1 - (1-\beta)\alpha)\boldsymbol{X} \bullet \boldsymbol{Y} - \alpha^2 \|\sqrt{\boldsymbol{X}}^{-1} d\hat{\boldsymbol{X}}\sqrt{\boldsymbol{Y}}\|_F \cdot \|\sqrt{\boldsymbol{Y}}^{-1} d\hat{\boldsymbol{Y}}\sqrt{\boldsymbol{X}}\|_F\right) \\
&\ge \xi\left((1 - (1-\beta)\alpha)\boldsymbol{X} \bullet \boldsymbol{Y} - \alpha^2 \left(\frac{\zeta n\mu}{\sqrt{\lambda_{min}}}\right)^2\right) \quad \text{(by 2 of Lemma 7.6)} \\
&= (1-\alpha)\xi\boldsymbol{X} \bullet \boldsymbol{Y} + \alpha\xi\left(\beta n\mu - \alpha \cdot \frac{\zeta^2 n^2\mu^2}{\lambda_{min}}\right) \\
&= (1-\alpha)\xi\boldsymbol{X} \bullet \boldsymbol{Y} + \alpha\xi\left(\frac{n^2\mu}{n+\nu} - \frac{\lambda_{min}}{5\zeta^2 n(n+\nu)\mu} \cdot \frac{\zeta^2 n^2\mu^2}{\lambda_{min}}\right) \\
&\quad \left(\text{since } \beta = \frac{n}{n+\nu} \text{ and } \alpha = \frac{\lambda_{min}}{5\zeta^2 n(n+\nu)\mu}\right) \\
&= (1-\alpha)\xi\boldsymbol{X} \bullet \boldsymbol{Y} + \alpha\xi\left(\frac{n^2\mu}{n+\nu} - \frac{n\mu}{5(n+\nu)}\right) \\
&\ge (1-\alpha)\xi\boldsymbol{X} \bullet \boldsymbol{Y} \\
&\ge (1-\alpha)\theta^r\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \quad \text{(by assertion 1 of the theorem).}
\end{aligned}
$$

Thus we have shown the inequality (65).

Now we derive the inequality (66):

$$
\begin{aligned}
&f(\boldsymbol{X} + \alpha d\boldsymbol{X}, \boldsymbol{Y} + \alpha d\boldsymbol{Y}) - f(\boldsymbol{X}, \boldsymbol{Y}) \\
&\le \alpha G_1(d\boldsymbol{X}, d\boldsymbol{Y}) + \alpha^2 G_2(d\boldsymbol{X}, d\boldsymbol{Y}) \text{ (by 1 of Lemma 7.4)} \\
&= -\frac{\alpha}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 + \frac{\alpha^2(n+\nu)d\boldsymbol{X} \bullet d\boldsymbol{Y}}{n\mu} + \frac{\alpha^2\sum_{j=1}^n(\xi_j^2 + \eta_j^2)}{2(1-\tau)} \\
&\quad \text{(by 1 of Lemma 7.4 with } \tau = 1/20) \\
&\le -\frac{\alpha}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 + \frac{\alpha^2(n+\nu)d\hat{\boldsymbol{X}} \bullet d\hat{\boldsymbol{Y}}}{n\mu} + \frac{\alpha^2\sum_{j=1}^n(\xi_j^2 + \eta_j^2)}{2(1-\tau)} \\
&\le -\frac{\alpha}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 + \frac{\alpha^2}{4\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 + \frac{2\alpha^2(\zeta n\mu)^2)}{2(1-1/20)\lambda_{min}^2} \\
&\quad \text{(by 3 of Lemma 7.1 with } \boldsymbol{Q} = \beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}\text{, 3 of Lemma 7.6,} \\
&\qquad \text{and } \beta = n/(n+\nu))
\end{aligned}
$$

$$
\begin{aligned}
&\le -\frac{\alpha}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 \cdot \left(1 - \frac{\alpha}{4}\right) + \frac{\alpha^2(\zeta n\mu)^2)}{(1-1/20)\lambda_{min}^2} \\
&\le -\frac{\lambda_{min}}{5\zeta^2\beta(n+\nu)^2\mu} \cdot \frac{1}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F^2 \cdot \left(1 - \frac{1}{320}\right) \\
&\quad + \left(\frac{\lambda_{min}}{5\zeta^2 n(n+\nu)\mu}\right)^2 \cdot \frac{20(\zeta n\mu)^2}{19\lambda_{min}^2} \quad \left(\text{since } \alpha = \frac{\lambda_{min}}{5\zeta^2 n(n+\nu)\mu} \le \frac{1}{80}\right) \\
&= -\frac{1}{5\zeta^2(n+\nu)^2} \cdot \left(\frac{\sqrt{\lambda_{min}}}{\beta\mu}\|\boldsymbol{H}(\beta)\|_F\right)^2 \cdot \frac{319}{320} \\
&\quad + \frac{1}{5^2\zeta^2(n+\nu)^2} \cdot \frac{20}{19} \\
&\le -\frac{1}{5\zeta^2(n+\nu)^2} \cdot \frac{3}{4} \cdot \frac{319}{320} + \frac{1}{25\zeta^2(n+\nu)^2} \cdot \frac{20}{19} \text{ (by 2 of Lemma 7.4)} \\
&\le -\frac{1}{10\zeta^2(n+\nu)^2}.
\end{aligned}
$$

Thus we have shown the inequality (66).

Assertion 2 follows directly from 1 of Lemma 7.6. This completes the proof of Theorem 8.3.

Assume that the hypothesis in section 7 is true. Then Theorem 8.3 ensures that the IIP potential-reduction method generates an infinite sequence $\{(\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{S}^2_{++}\}$ satisfying

$$
\left\{
\begin{array}{l}
f(\boldsymbol{X}^r, \boldsymbol{Y}^r) \le f(\boldsymbol{X}^0, \boldsymbol{Y}^0) - r\delta, \\
\theta^r \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 \le \xi \boldsymbol{X}^r \bullet \boldsymbol{Y}^r, \\
(\boldsymbol{X}^r, \boldsymbol{Y}^r) \in \mathcal{F} + \theta^r\left((\boldsymbol{X}^0, \boldsymbol{Y}^0) - (\boldsymbol{X}', \boldsymbol{Y}')\right)
\end{array}
\right. \tag{71}
$$

for every $r = 0, 1, 2, \ldots$. Thus we may regard $(\boldsymbol{X}^r, \boldsymbol{Y}^r)$ with any sufficiently large r as an approximate solution of the SDLCP (1). More precisely, for any given $\epsilon > 0$ we have $\boldsymbol{X}^r \bullet \boldsymbol{Y}^r \le \epsilon$ and $\theta^r \le \xi\epsilon/(\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0)$ if $r \ge (f(\boldsymbol{X}^0, \boldsymbol{Y}^0) - \nu \log \epsilon)/\delta$. If in addition $\nu = \sqrt{n}$ and $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = \rho(\boldsymbol{I}, \boldsymbol{I})$ for some $\rho > 0$, then the right-hand side of the inequality above is of $O(n^{2.5}\log(n\rho/\epsilon))$, the same order as the constrained potential-reduction algorithm (Algorithm I) proposed by Mizuno, Kojima, and Todd [32]. Besides the constrained potential-reduction algorithm, Mizuno, Kojima, and Todd [32] also presented a pure potential-reduction algorithm (Algorithm II) and its $O(nL)$-iteration variant (Algorithm III). We could also modify the IIP potential-reduction method to develop such variants, but the details are omitted here.

9. Concluding remarks. There remain many theoretical and practical issues to be studied further on the monotone SDLCP (1) in symmetric matrices and interior-point methods for solving it. In particular, the authors are interested in feasible and infeasible interior-point methods using a wide neighborhood of the central trajectory. The central trajectory following method in section 8.1 is mainly of theoretical importance. We need to prepare an initial feasible interior point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ in the narrow neighborhood $\mathcal{N}(\gamma)$ with $\gamma = 0.1$ and confine the generated sequence in the neighborhood. Even when we know such an initial feasible interior point, we should take a smaller search direction parameter β and a larger step size parameter α to increase the computational efficiency. Instead of $\mathcal{N}(\gamma)$, we could consider a wider neighborhood

$$
\mathcal{N}_\infty(\pi) = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}^2_{++} : \begin{array}{l} \lambda_{min} \ge \pi \mathrm{Tr}\ \boldsymbol{X}\boldsymbol{Y}/n, \text{ where } \lambda_{min} \text{ denotes} \\ \text{the minimum eigenvalue of } \boldsymbol{X}\boldsymbol{Y} \end{array} \right\}
$$

(for infeasible interior-point methods)

or

$\mathcal{N}_\infty(\pi) \cap \mathcal{F}_{++}$ (for feasible interior-point methods)

of the central trajectory. Here $\pi > 0$. This type of neighborhood has been successfully utilized in many feasible and infeasible interior-point methods ([18, 20, 26, 33, 48], etc.) for linear programs in the Euclidean space. The authors tried to extend the Kojima–Megiddo–Mizuno infeasible interior-point method [18] to the SDLCP (1) in symmetric matrices but encountered some difficulties in analyzing the step length parameter α, which keeps the next iterate remaining in the neighborhood $\mathcal{N}_\infty(\pi)$.

Nesterov and Nemirovskii [36] discussed variational inequalities with monotone operators and presented a path-following method for solving them.

In their recent paper [37], Nesterov and Todd presented a quite general theoretical foundation for interior-point algorithms for a wide class of nonlinear programs in conic form including a primal-dual pair (2) of semidefinite programs as a special case. Among others, they proposed a joint scaling primal-dual interior-point method for linear programs in conic form, which is an extension of the $O(\sqrt{n}L)$ iteration potential-reduction algorithm given by Kojima–Mizuno–Yoshise [22]. Our current paper has been written independently of their paper [37].

Acknowledgment. The authors wish to thank Dr. M. Shida and the anonymous referees for very thorough reading of the manuscript and many valuable comments and suggestions to improve the paper.

REFERENCES

[1] F. Alizadeh, *Optimization over the positive-definite cone: Interior point methods and combinatorial applications*, in Advances in Optimization and Parallel Computing, P. Pardalos, ed., North-Holland, Amsterdam, 1992, pp. 1–25.

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

[3] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, *Primal-dual interior-point methods for semidefinite programming*, Presented at the Mathematical Programming Symposium, Ann Arbor, MI, 1994, manuscript.

[4] J. F. Bonnans and C. C. Gonzaga, *Convergence of interior point algorithms for the monotone linear complementarity problem*, Technical Report 2074, INRIA, Rocquencourt, France, 1993.

[5] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, PA, 1994.

[6] R. W. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, New York, 1992.

[7] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley & Sons, New York, 1968.

[8] G. H. Golub and C. H. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, London, 1989.

[9] C. C. Gonzaga and M. J. Todd, *An $O(\sqrt{n}L)$-iteration large-step primal-dual affine algorithm for linear programming problems*, SIAM J. Optim., 2 (1992), pp. 349–359.

[10] M. S. Gowda and T. I. Seidman, *Generalized linear complementarity problems*, Math. Programming, 46 (1990), pp. 329–340.

[11] O. Güller, *Generalized linear complementarity problems and interior point algorithms for their solutions*, Internal Report, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, CA, 1993.

[12] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, Program in Statistics and Operations Research, Princeton University, Princeton, NJ, 1994.

[13] R. D. Hill and S. R. Waters, *On the cone of positive semidefinite matrices*, Linear Algebra Appl., 90 (1987), pp. 81–88.
[14] M. W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, New York, 1974.
[15] F. Jarre, *An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices*, SIAM J. Control Optim., 31 (1993), pp. 1360–1377.
[16] N. Karmarkar, *A New Polynomial-Time Algorithm for Linear Programming*, Combinatorica, 4 (1984), pp. 373–395.
[17] M. Kojima, *Basic lemmas in polynomial-time infeasible-interior-point methods for linear programs*, Ann. Oper. Res., 62 (1996), pp. 1–28.
[18] M. Kojima, N. Megiddo, and S. Mizuno, *A primal-dual infeasible-interior-point algorithm for linear programming*, Math. Programming, 61 (1993), pp. 263–280.
[19] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise, *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Lecture Notes in Computer Science 538, Springer-Verlag, New York, 1991.
[20] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point algorithm for linear programming*, in Progress in Mathematical Programming, Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989, pp. 29–47.
[21] M. Kojima, S. Mizuno, and A. Yoshise, *A polynomial-time algorithm for a class of linear complementary problems*, Math. Programming, 44 (1989), pp. 1–26.
[22] M. Kojima, S. Mizuno, and A. Yoshise, *An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems*, Math. Programming, 50 (1991), pp. 331–342.
[23] M. Kojima, T. Noma, and A. Yoshise, *Global convergence in infeasible-interior-point algorithms*, Math. Programming, 65 (1994), pp. 43–72.
[24] M. Kojima, M. Shida, and S. Shindoh, *Reduction of monotone linear complementarity problems over cones to linear programs over cones*, Research Report #296, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, 1995.
[25] I. J. Lustig, *Feasibility issues in a primal-dual interior-point method for linear programming*, Math. Programming, 49 (1990/91), pp. 145–162.
[26] I. J. Lustig, R. E. Marsten, and D. F. Shanno, *Computational experience with a primal-dual interior point method for linear programming*, Linear Algebra Appl., 152 (1991), pp. 191–222.
[27] R. Marsten, R. Subramanian, M. Saltzman, I. J. Lustig, and D. Shanno, *Interior point methods for linear programming: Just call Newton, Lagrange, and Fiacco and McCormick!*, Interfaces, 20 (1990), pp. 105–116.
[28] L. McLinden, *The complementarity problem for maximal monotone multifunctions*, in Variational Inequalities and Complementarity Problems, R. W. Cottle, F. Giannessi, and J.-L. Lions, eds., John Wiley & Sons, New York, 1980, pp. 251–270.
[29] N. Megiddo, *A monotone complementarity problem with feasible solutions but not complementarity solutions*, Math. Programming, 12 (1977), pp. 131–132.
[30] N. Megiddo, *Pathways to the optimal set in linear programming*, in Progress in Mathematical Programming, Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989, pp. 131–158.
[31] S. Mizuno, F. Jarre, and J. Stoer, *A unified approach to infeasible-interior-point algorithms via geometrical linear complementarity problems*, Inst. Statist. Math., 1994.
[32] S. Mizuno, M. Kojima, and M. J. Todd, *Infeasible-interior-point primal-dual potential-reduction algorithms for linear programming*, SIAM J. Optim., 5 (1995), pp. 52–67.
[33] S. Mizuno, M. J. Todd, and Y. Ye, *On adaptive-step primal-dual interior-point algorithms for linear programming*, Math. Oper. Res., 18 (1993), pp. 964–981.
[34] R. D. C. Monteiro and I. Adler, *Interior path following primal-dual algorithms. Part* I: *Linear programming*, Math. Programming, 44 (1989), pp. 27–41.
[35] Ju. E. Nesterov and A. S. Nemirovskii, *Self-concordant functions and polynomial-time methods in convex programming*, Report, Central Economical and Mathematical Institute, USSR Acad. Sci., Moscow, USSR, 1989.
[36] Ju. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, SIAM, Philadelphia, PA, 1994.
[37] Ju. E. Nesterov and M. J. Todd, *Self-Scaled Cones and Interior-Point Methods in Nonlinear Programming*, Working Paper, CORE, Catholic University of Louvain, Louvain-la-Neuve, Belgium, 1994.
[38] A. M. Ostrowski and O. Taussky, *On the variation of the determinant of a positive definite matrix*, Nederl. Akad. Wetensch, Proc. Ser. A, 54 (1951), pp. 383–385.

[39] R. Sznajder and M. S. Gowda, *Generalizations of P_0- and P- properties; extended vertical and horizontal LCP's*, Linear Algebra Appl., to appear.

[40] K. Tanabe, *Centered Newton method for mathematical programming*, in System Modeling and Optimization, M. Iri and K. Yajima, eds., Springer-Verlag, New York, 1988, pp. 197–206.

[41] M. J. Todd, *Recent developments and new directions in linear programming*, in Recent Developments and Applications, M. Iri and K. Tanabe, eds., Kluwer Academic Publishers, London, 1989, pp. 109–157.

[42] M. J. Todd and Y. Ye, *A centered projective algorithm for linear programming*, Math. Oper. Res., 15 (1990), pp. 508–529.

[43] R. H. Tütüncü and M. J. Todd, *Reducing horizontal linear complementarity problems*, Technical Report, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1993.

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

[45] L. Vandenberghe and S. Boyd, *Semidefinite programming*, Working Paper, Information Systems Laboratory, Stanford University, Stanford, CA, 1995.

[46] Y. Ye, *An $O(n^3L)$ potential reduction algorithm for linear programming*, Math. Programming, 50 (1991), pp. 239–258.

[47] Y. Ye, *A fully polynomial-time approximation algorithm for computing a stationary point of the general linear complementarity problem*, Math. Oper. Res., 18 (1993), pp. 334–345.

[48] Y. Zhang, *On the convergence of a class of infeasible interior-point algorithms for the horizontal linear complementarity problem*, SIAM J. Optim., 4 (1994), pp. 208–227.

[49] Y. Zhang, R. A. Tapia, and F. Potra, *On the superlinear convergence of interior point algorithms for a general class of problems*, SIAM J. Optim., 3 (1993), pp. 413–422.