Research Reports on Mathematical and Computing Sciences
Series B : Operations Research

Department of Mathematical and Computing Sciences
Tokyo Institute of Technology
2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan

Search Directions in the SDP and the Monotone SDLCP: Generalization and Inexact Computation

Masakazu Kojima[†]
Masayuki Shida[‡]
Susumu Shindoh[♯]

March 1997, B-327

Abstract. We propose a family of directions that generalizes many directions proposed so far in interior-point methods for the SDP (semidefinite programming) and for the monotone SDLCP (semidefinite linear complementarity problem). We derive the family from the Helmberg-Rendl-Vanderbei-Wolkowicz/Kojima-Shindoh-Hara/Monteiro direction by relaxing its "centrality equation" into a "centrality inequality." Using this family, we present a predictor-corrector infeasible-interior-point method to provide a theoretical basis for inexact computation of directions in primal-dual interior-point methods for the SDP.

Key words Interior-Point Methods, Semidefinite Programming, Semidefinite Linear Complementarity Problem, Inexact Computation, Search Direction

† Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan. e-mail: kojima@is.titech.ac.jp

‡ Dept. of Mathematics, Faculty of Engineering, Kanagawa University, Rokkakubashi, Kanagawa-ku, Yokohama 221, Japan. e-mail: shida@is.titech.ac.jp

♯ Dept. of Mathematics and Physics, The National Defense Academy, Hashirimizu 1-10-20, Yokosuka 239, Japan. e-mail: shindoh@cc.nda.ac.jp

1. Introduction.

Recently various search directions have been proposed in primal-dual interior-point methods for solving the SDP (semidefinite programming). We are concerned with

- the AHO direction given by Alizadeh-Haeberly-Overton [1],
- the HRVW/KSH/M direction proposed independently by Helmberg-Rendl-Vanderbei-Wolkowicz [7] and Kojima-Shindoh-Hara [9], and later rediscovered by Monteiro [13] via a scaled symmetrization operator,
- the NT direction given by Nesterov-Todd [18], and reformulated via the V-space transformation by Sturm-Zhang [21],
- the KSH family of directions given by Kojima-Shindoh-Hara [9],
- the Tseng family of directions given by [23],
- the MZ family of directions given by Monteiro-Zhang [17] (see also Monteiro [13], Zhang [24] and Monteiro [14]),
- the MT family of directions given by Monteiro-Tsuchiya [15].

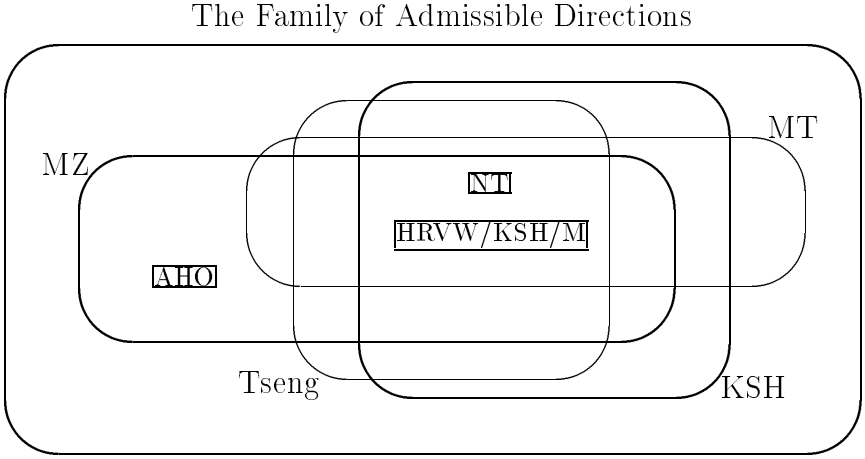
Using these directions, numerous primal-dual interior-point methods such as potential-reduction methods, path-following methods and predictor-corrector methods have been developed and their global and local convergence have been studied ([1, 2, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, etc.]). This raises a few natural questions like why there are so many directions and why all the directions listed above attain the same best known computational complexity.

Usually we describe each direction in terms of three kinds of linear equations, namely, primal-feasibility, dual-feasibility and centrality equations. All the directions share the primal-feasibility and the dual-feasibility equations. In general, the other centrality equation, which is derived through linearlization of the central trajectory equation $\boldsymbol{XY} = \mu \boldsymbol{I}$, differs from one direction to another. When the current point lies on the central trajectory, however, all the centrality equations are equivalent to each other, that is, they define a common solution set. Furthermore the coefficient matrices of each centrality equation change continuously as the current point changes. This implies that the solution set of one centrality equation and the solution set of another centrality equation are "close" to each other whenever the current point lies in a narrow neighborhood of the central trajectory in which short-step path-following methods and predictor-corrector methods generate all iterates. Therefore there must be a nice way of describing a larger family of directions that attains the same best known computational complexity as all the directions listed above did. This gives one motivation of the current paper. Using "a centrality inequality" instead of a centrality equation, we propose such a larger family of directions that we will call *admissible directions*. Figure 1. illustrates relations among the directions listed above and our admissible directions when the current point is feasible and lies in a neighborhood of the central trajectory.

The other motivation of this paper is to introduce an inexact computation of a direction, which was proposed by Mizuno and Jarre [12] for primal-dual interior-point methods for LPs, into SDPs. From our computational experience on the SDPA [3], which is a C++ program of a primal-dual interior-point algorithm for SDPs, the authors learned that an exact computation of the HRVW/KSH/M, the NT and the AHO directions in an SDP is much more expensive than an exact computation of a standard primal-dual direction in an LP having the same number of constraints. When we employ the HRVW/KSH/M direction, the NT direction or the AHO direction, we can reduce its computation to a system of linear equations $\boldsymbol{B}d\boldsymbol{y} = \boldsymbol{d}$ in $d\boldsymbol{y} \in R^m$ with an $m \times m$ matrix $\boldsymbol{B}$ and $\boldsymbol{d} \in R^m$ as in the LP case, where m denotes the number of equality constraints of the primal SDP to be solved. In the LP case, $\boldsymbol{B}$ corresponds to $\boldsymbol{AD}^2\boldsymbol{A}^T$, where $\boldsymbol{A}$ denotes the $m \times n$ matrix associated with the equality constraint $\boldsymbol{Ax} = \boldsymbol{b}$ of the primal problem and $\boldsymbol{D}$ the primal-dual diagonal scaling matrix, and $\boldsymbol{B}$ is generally sparse when $\boldsymbol{A}$ is sparse. But, in the SDP case,

(i) the computation of the coefficient matrix $\boldsymbol{B}$ is much more expensive than in the LP case, and

Figure 1: Relation between directions for the SDP and the monotone SDLCP when the current point lies in a neighborhood of the central trajectory.



(ii) $\boldsymbol{B}$ is generally full dense even if data matrices of the SDP to be solved are sparse.

See also the papers [2, 22]. These are critical disadvantages that we need to take account of when we develop a practically efficient primal-dual interior-point algorithm for solving large scale SDPs. To overcome the first disadvantage, Fujisawa, Kojima and Nakata proposed an efficient method for computing the coefficient matrix $\boldsymbol{B}$ in their recent paper [4]. The second disadvantage is more serious when we apply a direct method such as the Cholesky factorization (in the NT direction and the HRVW/KSH/M direction cases) and the LU factorization (in the AHO direction case) to solving the full dense system of equations $\boldsymbol{B}d\boldsymbol{y} = \boldsymbol{d}$; if its size m is large, say more than a few thousands, it is quite difficult even to store $\boldsymbol{B}$ in the memory of a standard workstation. In such a case, we are forced to utilize an iteration method such as the conjugate gradient method to compute an approximate or inexact direction. In fact, the conjugate gradient method has been applied to large scale SDPs arising from relaxation of the quadratic assignment problem [11, 25]. The inequality used to describe our family of admissible directions provides a sufficient condition for an inexact direction to enjoy the same best known theoretical computational complexity as in the case of exact computation of direction.

Our major concern is the SDP, but we deal with the monotone SDLCP (semidefinite linear complementarity problem), which was introduced by Kojima-Shindoh-Hara [9]. There are two reasons. First, the monotone SDLCP is a general mathematical model for a wide class of problems; in particular, we can directly apply the results given in this paper to not only the SDP but also the monotone linear complementarity problem and the convex quadratic program. Second, interior-point methods for the monotone SDLCP and their complexity analysis require shorter and simpler description than primal-dual interior-point methods for the SDP and their analysis.

An important issue on interior-point methods for the SDP and the monotone SDLCP is how we determine a step length. Many practical implementations of primal-dual interior-point methods for the LP take a step length very near to the boundary (for example, 99.995 percent of the step length to the boundary) of the nonnegative orthant from the current point which lies in its interior. From some computational results reported so far [2, 22], however, we know that such an aggressive step length often causes a serious numerical instability in the SDP case. Such a numerical instability is expected to be more serious if we allow an inaccurate direction such as our admissible directions for large scale SDPs. Also we have to take account of the cost and the numerical instability of computing a step length. In order to measure the step length along a direction from the current point which lies in the interior of the set of positive semidefinite

matrices to its boundary, we need to compute the eigenvalues (or the eigenvalue decomposition) of some dense symmetric matrices generated from the current point and a direction. We learned from computational experience on the SDPA [3] that the numerical computation of those eigenvalues becomes more expensive as an SDP to be solved is larger, and that it becomes unstable as the current iterate approaches too close to the boundary of the set of positive semidefinite matrices before it approaches to a solution of the SDP. Thus we want to develop an interior-point algorithm using inaccurate directions that stably generates all iterates in a neighborhood of the central trajectory. One algorithm which is expected to satisfy this requirement is the short-step path-following interior-point algorithm. But it is not attractive in practice. Instead, we propose a variant of the Mizuno-Todd-Ye type predictor-corrector infeasible-interior-point algorithm using the standard Frobenius norm neighborhood of the infeasible central trajectory. A significant feature of this algorithm is a step length computation based on the Cholesky factorization of a symmetric positive definite current primal iterate; the algorithm requires neither eigenvalues nor the eigenvalue decomposition of any symmetric matrix to compute a step length.

The paper is organized as follows: After introducing notation and symbols, Section 2 presents some basic materials such as the monotone SDLCP, the infeasible central trajectory and its Frobenius norm neighborhood that are necessary in our succeeding sections. We propose a family of directions, which we have named admissible directions above, in Section 3, and make detailed discussion on how to choose a step length along an admissible direction in Section 4. In Section 5, we state a predictor-corrector infeasible-interior-point algorithm for the monotone SDLCP using admissible directions, and a main theorem which establishes the global convergence of the algorithm. Section 6 is devoted to a proof of a key lemma used in the proof of the main theorem.

2. Preliminaries.

2.1 Symbols and Notation.

Let R be the set of real numbers, R^m the m dimensional Euclidean space, $R^{n\times n}$ the set of $n\times n$ real matrices, $\mathcal{S}\subset R^{n\times n}$ the set of $n\times n$ real symmetric matrices, $\mathcal{S}_+\subset\mathcal{S}$ the set of positive semidefinite matrices, and $\mathcal{S}_{++}\subset\mathcal{S}_+$ the set of positive definite matrices. If $\boldsymbol{A}-\boldsymbol{B}\in\mathcal{S}_+$ holds for matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in $\mathcal{S}$, we write $\boldsymbol{A}\succeq\boldsymbol{B}$ or $\boldsymbol{B}\preceq\boldsymbol{A}$. We denote the standard inner product of $\boldsymbol{A}$ and $\boldsymbol{B}$ in $R^{n\times n}$ by $\boldsymbol{A}\bullet\boldsymbol{B}=\sum_{i=1}^n\sum_{j=1}^n A_{ij}B_{ij}$, where A_{ij} (B_{ij}, respectively) denotes the (i,j)th element of $\boldsymbol{A}$ ($\boldsymbol{B}$, respectively), and the Frobenius norm of $\boldsymbol{A}\in R^{n\times n}$ by $\|\boldsymbol{A}\|_F=(\boldsymbol{A}\bullet\boldsymbol{A})^{1/2}=\left(\sum_{i=1}^n\sum_{j=1}^n A_{ij}^2\right)^{1/2}$. We use the notation $\boldsymbol{A}^{1/2}$ for the symmetric square root of a matrix $\boldsymbol{A}\in\mathcal{S}_+$, $\boldsymbol{I}$ for the $n\times n$ identity matrix, and $\boldsymbol{O}$ for the $n\times n$ zero matrix.

2.2 The Monotone SDLCP and a Primal-Dual Pair of SDPs.

The SDLCP (semidefinite linear complementarity problem) is the problem of finding a point $(\boldsymbol{X},\boldsymbol{Y})$ such that

$$(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{F},\ \boldsymbol{X}\in\mathcal{S}_+,\ \boldsymbol{Y}\in\mathcal{S}_+\ \text{ and }\boldsymbol{X}\bullet\boldsymbol{Y}=0. \tag{1}$$

Here $\mathcal{F}$ denotes an $n(n+1)/2$ dimensional affine subspace of $\mathcal{S}\times\mathcal{S}$. Throughout the paper we assume that $\mathcal{F}$ is monotone;

$$(\boldsymbol{X}-\boldsymbol{X}')\bullet(\boldsymbol{Y}-\boldsymbol{Y}')\geq 0\ \text{ if }(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{F}\ \text{ and }(\boldsymbol{X}',\boldsymbol{Y}')\in\mathcal{F}. \tag{2}$$

Now we relate the primal-dual pair of SDPs $\mathcal{P}$ and $\mathcal{D}$ to the monotone SDLCP (1). Let $a_i\in R$ $(i=1,2,\ldots,m)$ and $\boldsymbol{A}_i\in\mathcal{S}$ $(i=0,1,2,\ldots,m)$. Consider the SDP and its dual:

$$\left.\begin{array}{lll}
\mathcal{P}: & \text{minimize} & \displaystyle\sum_{i=1}^m a_i x_i \\
 & \text{subject to} & \displaystyle\sum_{i=1}^m \boldsymbol{A}_i x_i-\boldsymbol{X}=\boldsymbol{A}_0,\ \boldsymbol{X}\in\mathcal{S}_+. \\
\mathcal{D}: & \text{maximize} & \boldsymbol{A}_0\bullet\boldsymbol{Y} \\
 & \text{subject to} & \boldsymbol{A}_i\bullet\boldsymbol{Y}=a_i\ (i=1,2,\ldots,m),\ \boldsymbol{Y}\in\mathcal{S}_+.
\end{array}\right\}$$

Let

$$\mathcal{F} = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S} \times \mathcal{S} \ : \ \begin{array}{l} \boldsymbol{X} = \sum_{i=1}^{m} \boldsymbol{A}_i x_i - \boldsymbol{A}_0 \ \text{ for some } \boldsymbol{x} \in R^m, \\ \boldsymbol{A}_i \bullet \boldsymbol{Y} = a_i \ (i = 1, 2, \ldots, m) \end{array} \right\}. \tag{3}$$

Then $\mathcal{F}$ forms an $n(n+1)/2$ dimensional affine subspace of $\mathcal{S} \times \mathcal{S}$ satisfying the self-orthogonality relation

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

which is a special case of monotonicity (2). We may regard $\mathcal{F} \cap (\mathcal{S}_+ \times \mathcal{S}_+)$ as the combined feasible region of the primal-dual pair of SDPs $\mathcal{P}$ and $\mathcal{D}$; $(\boldsymbol{x}, \boldsymbol{X})$ is a feasible solution of $\mathcal{P}$ and $\boldsymbol{Y}$ a feasible solution of $\mathcal{D}$ if and only if $(\boldsymbol{X}, \boldsymbol{Y})$ lies in $\mathcal{F} \cap (\mathcal{S}_+ \times \mathcal{S}_+)$. We see that $\sum_{i=1}^{m} a_i x_i - \boldsymbol{A}_0 \bullet \boldsymbol{Y} = \boldsymbol{X} \bullet \boldsymbol{Y} \geq 0$ holds for every feasible solution $(\boldsymbol{x}, \boldsymbol{X})$ of $\mathcal{P}$ and every feasible solution $\boldsymbol{Y}$ of $\mathcal{D}$. This result is known as the weak duality. The left hand side of the identity above corresponds to the duality gap. Thus $(\boldsymbol{x}, \boldsymbol{X})$ is a minimum solution of $\mathcal{P}$ and $\boldsymbol{Y}$ a maximum solution of $\mathcal{D}$ if $(\boldsymbol{X}, \boldsymbol{Y})$ is a solution of the monotone SDLCP (1) . The converse (the strong duality) is known to be true whenever there is an interior feasible solution $(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}})$ of the monotone SDLCP (1);

$$(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}}) \in \mathcal{F}, \ \hat{\boldsymbol{X}} \in \mathcal{S}_{++} \ \text{ and } \hat{\boldsymbol{Y}} \in \mathcal{S}_{++} \tag{4}$$

See the papers [9, etc.] for more detailed relation between the primal-dual pair of SDPs $\mathcal{P}$ and $\mathcal{D}$ and the monotone SDLCP (1).

2.3 The Infeasible Central Trajectory.

Let $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ and $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) \in \mathcal{F}$. The point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ serves as an initial point. Define

$$\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \equiv \mathcal{F} + \theta(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}). \tag{5}$$

Obviously we have $\mathcal{F}(1, \boldsymbol{X}^0, \boldsymbol{Y}^0) \ni (\boldsymbol{X}^0, \boldsymbol{Y}^0)$ and $\mathcal{F}(0, \boldsymbol{X}^0, \boldsymbol{Y}^0) = \mathcal{F}$. It should be noted that the right hand side of (5) does not depend on any specific choice of $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) \in \mathcal{F}$;

$$\mathcal{F} + \theta(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}) = \mathcal{F} + \theta(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}', \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}') \text{ for any } (\bar{\boldsymbol{X}}', \bar{\boldsymbol{Y}}') \in \mathcal{F}.$$

For each $\theta \in R$, the set $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ forms an $n(n+1)/2$ dimensional monotone affine subspace of $\mathcal{S} \times \mathcal{S}$.

When $\mathcal{F}$ is of the form (3) induced from the primal-dual pair of SDPs $\mathcal{P}$ and $\mathcal{D}$, we need to choose a triplet $(\boldsymbol{x}^0, \boldsymbol{X}^0, \boldsymbol{Y}^0) \in R^m \times \mathcal{S}_{++} \times \mathcal{S}_{++}$ as an initial point. In this case, we see that

$$\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) = \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S} \times \mathcal{S} \ : \ \begin{array}{l} \sum_{i=1}^{m} \boldsymbol{A}_i x_i - \boldsymbol{X} = (1-\theta)\boldsymbol{A}_0 + \theta \left(\sum_{i=1}^{m} \boldsymbol{A}_i x_i^0 - \boldsymbol{X}^0 \right) \\ \text{for some } \boldsymbol{x} \in R^m, \\ \boldsymbol{A}_i \bullet \boldsymbol{Y} = (1-\theta) a_i + \theta \boldsymbol{A}_i \bullet \boldsymbol{Y}^0 \ (i = 1, 2, \ldots, m) \end{array} \right\}.$$

Hence we need no $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}) \in \mathcal{F}$ in the representation of $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$.

Let

$$\begin{aligned} \overline{\theta} &\equiv \sup \left\{ \theta \geq 0 : \mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset \right\}, \\ \underline{\theta} &\equiv \inf \left\{ \theta \geq 0 : \mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset \right\}. \end{aligned}$$

Lemma 2.1.

(i) $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) = \mathcal{F}$ *for every* $\theta \geq 0$ *if and only if the initial point* $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ *is feasible, i.e.,* $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F}$.

(ii) $0 \leq \underline{\theta} < 1 < \overline{\theta}$.

(iii) $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset$ *for every* $\theta \in (\underline{\theta}, \overline{\theta})$.

Proof: Assertions (i) and (ii) are straightforward from the definition of $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$. To prove assertion (iii), assume that $\theta \in (\underline{\theta}, \overline{\theta})$. Then there exists θ' and θ'' such that

$$\theta' < \theta < \theta'', \ \mathcal{F}(\theta', \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset \text{ and } \mathcal{F}(\theta'', \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset.$$

Let $(\boldsymbol{X}', \boldsymbol{Y}') \in \mathcal{F}(\theta', \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++})$, $(\boldsymbol{X}'', \boldsymbol{Y}'') \in \mathcal{F}(\theta'', \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++})$ and $\zeta \equiv (\theta'' - \theta)/(\theta'' - \theta')$. Then

$$\zeta(\boldsymbol{X}', \boldsymbol{Y}') + (1 - \zeta)(\boldsymbol{X}'', \boldsymbol{Y}'') \quad \in \quad \mathcal{S}_{++} \times \mathcal{S}_{++}.$$

We also see that

$$\begin{aligned} \zeta(\boldsymbol{X}', \boldsymbol{Y}') + (1 - \zeta)(\boldsymbol{X}'', \boldsymbol{Y}'') \quad &\in \quad \zeta\mathcal{F}(\theta', \boldsymbol{X}^0, \boldsymbol{Y}^0) + (1 - \zeta)\mathcal{F}(\theta'', \boldsymbol{X}^0, \boldsymbol{Y}^0) \\ &= \quad \mathcal{F} + \zeta\theta'(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}) + (1 - \zeta)\theta''(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}) \\ &= \quad \mathcal{F} + \theta(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}) \\ &= \quad \mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0). \end{aligned}$$

Thus we have shown that $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset$. ∎

For every $\theta \in (\underline{\theta}, \overline{\theta})$, define

$$\mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \quad \equiv \quad \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) : \boldsymbol{X}\boldsymbol{Y} = \frac{\theta \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0}{n} \boldsymbol{I} \right\}.$$

When $\mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) = \mathcal{F}$ for every $\theta \in R$ (or $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ is feasible), $\underline{\theta} = 0$ and $\overline{\theta} = \infty$. In this case the set $\{(\theta, \mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)) : 0 < \theta < \infty\}$ is known as the *central trajectory* (or the *central path*), and $\mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ converges to a solution of the monotone SDLCP (1) as $\theta \to 0$ [9]. These results were extended by Shida-Shindoh [20] and Goldfarb-Scheinberg [5] to the case where the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ is not necessarily feasible.

Lemma 2.2.

(i) The set $\{(\theta, \mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)) : \underline{\theta} < \theta < \overline{\theta}\}$ forms a smooth trajectory in $R \times \mathcal{S}_{++} \times \mathcal{S}_{++}$.

(ii) If the set $\{(\theta, \mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)) : 0 < \theta \leq 1\}$ is bounded then $\underline{\theta} = 0$ and the point $\mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ converges to a solution of the monotone SDLCP (1) as $\theta \to 0$.

(iii) If $\mathcal{F} \cap (\mathcal{S}_{++} \times \mathcal{S}_{++}) \neq \emptyset$ then $\underline{\theta} = 0$ and the set $\{(\theta, \mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)) : 0 < \theta \leq 1\}$ is bounded.

Proof: See [20, Theorems 2.2 and 2.3], and [5, Theorem 5.2]. ∎

We will call the set $\{(\theta, \mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)) : \underline{\theta} < \theta < \overline{\theta}\}$ the *infeasible central trajectory.*

2.4 Neighborhood of the Infeasible Central Trajectory.

As many short step path-following methods and predictor-corrector methods developed so far for the SDP and the monotone SDLCP (1), the predictor-corrector algorithm given in Section 5 numerically traces the infeasible central trajectory $\{(\theta, \mathcal{C}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)) : \underline{\theta} < \theta < \overline{\theta}\}$. For this purpose, we first impose an additional assumption on the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$:

$$(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++}, \ \mu^0 \equiv \frac{\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0}{n} \ \text{ and } \ \left\| (\boldsymbol{X}^0)^{1/2} \boldsymbol{Y}^0 (\boldsymbol{X}^0)^{1/2} - \mu^0 \boldsymbol{I} \right\|_F \leq \gamma \mu^0, \tag{6}$$

where $\gamma \in [0, 1)$. For example, we take an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\rho \boldsymbol{I}, \rho \boldsymbol{I})$ for some $\rho > 0$. Then, define

$$\mathcal{N}(\gamma, \theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \equiv \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \cap (\mathcal{S}_+ \times \mathcal{S}) : \| \boldsymbol{X}^{1/2} \boldsymbol{Y} \boldsymbol{X}^{1/2} - \theta \mu^0 \boldsymbol{I} \|_F \leq \gamma \theta \mu^0 \right\} \tag{7}$$

for every $\theta > 0$. It should be noted that if $\gamma \in [0,1)$ and $\theta > 0$ then

$$\left.\begin{array}{l}(1-\gamma)\theta\mu^0 \boldsymbol{I} \preceq \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} \preceq (1+\gamma)\theta\mu^0\boldsymbol{I},\\ n(1-\gamma)\theta\mu^0 \leq \boldsymbol{X}\bullet\boldsymbol{Y} \leq n(1+\gamma)\theta\mu^0, \text{ and}\\ \boldsymbol{X}\in\mathcal{S}_{++},\ \boldsymbol{Y}\in\mathcal{S}_{++}\end{array}\right\} \text{ for any } (\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0); \tag{8}$$

hence $\mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)$ can be rewritten as

$$\mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0) = \left\{(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F}(\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0) \cap (\mathcal{S}_{++}\times\mathcal{S}_{++}) : \|\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} - \theta\mu^0\boldsymbol{I}\|_F \leq \gamma\theta\mu^0\right\}. \tag{7}'$$

The set $\bigcup_{\underline{\theta}<\theta<\overline{\theta}} \{(\theta,\boldsymbol{X},\boldsymbol{Y}) : (\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)\}$ forms a neighborhood of the infeasible central trajectory. This type of neighborhood has been used in many papers [8, 9, 10, 14, 15, 16, 19, 21, 23, 24, etc.]. The parameter $\gamma \in [0,1)$ determines the width of the neighborhood of the infeasible central trajectory. If we took $\gamma = 0$, the neighborhood would coincide with the infeasible central trajectory.

We generate a sequence $\{(\theta^k,\boldsymbol{X}^k,\boldsymbol{Y}^k)\}$ in the neighborhood such that

$$1 = \theta^0 \geq \theta^k \geq \theta^{k+1} \geq 0 \ \text{ and } (\boldsymbol{X}^k,\boldsymbol{Y}^k) \in \mathcal{N}(\gamma,\theta^k,\boldsymbol{X}^0,\boldsymbol{Y}^0)\ (k = 1,2,\ldots,\ ). \tag{9}$$

If in addition $\theta^k \to 0$ as $k \to \infty$ then any accumulation point of the sequence $\{(\boldsymbol{X}^k,\boldsymbol{Y}^k)\}$ is a solution of the monotone SDLCP (1).

2.5 Getting Information on Solvability.

Practically, if θ^k gets smaller than a given accuracy parameter ϵ, we have an approximate solution $(\boldsymbol{X}^k,\boldsymbol{Y}^k)$ of the monotone SDLCP (1) satisfying

$$\left.\begin{array}{l}0 \leq \theta^k < \epsilon,\ (\boldsymbol{X}^k,\boldsymbol{Y}^k) \in \mathcal{F} + \theta^k(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}),\\ (\boldsymbol{X}^k,\boldsymbol{Y}^k) \in \mathcal{S}_{++}\times\mathcal{S}_{++},\ \boldsymbol{X}^k\bullet\boldsymbol{Y}^k \leq n\mu^0\theta^k \leq n\mu^0\epsilon.\end{array}\right\} \tag{10}$$

But this does not always cover the case. The monotone SDLCP (1) may not have a solution or the magnitude of a solution may be too large to approximate. We can get such information by using the lemma below.

Lemma 2.3. *[9, Lemma 7.6] Let $\gamma \in [0,1)$, $\omega \geq 1$ and $\varphi_0(\gamma,\omega) = (2\omega + 1 - \gamma)/(1-\gamma)$. Suppose that the initial point $(\boldsymbol{X}^0,\boldsymbol{Y}^0)$ satisfies (6), and that there exists a solution $(\boldsymbol{X}^*,\boldsymbol{Y}^*)$ of the monotone SDLCP (1) satisfying*

$$\boldsymbol{X}^* \preceq \omega\boldsymbol{X}^0 \quad \text{and} \quad \boldsymbol{Y}^* \preceq \omega\boldsymbol{Y}^0, \tag{11}$$

If $(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)$ for some $\theta \in (0,1]$, then

$$\theta(\boldsymbol{X}^0\bullet\boldsymbol{Y} + \boldsymbol{X}\bullet\boldsymbol{Y}^0) \leq \varphi_0(\gamma,\omega)\boldsymbol{X}\bullet\boldsymbol{Y}. \tag{12}$$

We will utilize inequality (12) to get information on solvability of the monotone SDLCP (1) in Section 5: If a current point $(\boldsymbol{X},\boldsymbol{Y}) = (\boldsymbol{X}^k,\boldsymbol{Y}^k) \in \mathcal{N}(\gamma,\theta^k,\boldsymbol{X}^0,\boldsymbol{Y}^0)$ does not satisfy inequality (12), we can conclude that the monotone SDLCP (1) has no solution $(\boldsymbol{X}^*,\boldsymbol{Y}^*)$ satisfying (11). We remark that (12) always holds at the initial point $(\boldsymbol{X},\boldsymbol{Y}) = (\boldsymbol{X}^0,\boldsymbol{Y}^0)$ with $\theta = 1$ regardless of the existence of such a solution.

2.6 Newton Direction towards the Infeasible Central Trajectory.

Let $\theta^0 = 1$. It follows from (6) that $(\theta^0,\boldsymbol{X}^0,\boldsymbol{Y}^0)$ satisfies (9). Now, assuming that a current point $(\theta,\boldsymbol{X},\boldsymbol{Y}) = (\theta^k,\boldsymbol{X}^k,\boldsymbol{Y}^k)$ satisfies (9), we outline the idea on how we generate a new point. We first choose a direction parameter $\beta \in [0,1]$ which determine a target point $(\theta',\boldsymbol{X}',\boldsymbol{Y}')$;

$$\theta' = \beta\theta^k,$$
$$(\boldsymbol{X}',\boldsymbol{Y}') \in \mathcal{F}(\theta',\boldsymbol{X}^0,\boldsymbol{Y}^0),\ \boldsymbol{X}'\boldsymbol{Y}' = \theta'\mu^0\boldsymbol{I}, \tag{13}$$
$$(\boldsymbol{X}',\boldsymbol{Y}') \in \mathcal{S}_+\times\mathcal{S}_+. \tag{14}$$

If $\beta = 0$ then $(\boldsymbol{X}', \boldsymbol{Y}')$ satisfying (13) and (14) is a solution of the monotone SDLCP (1), otherwise the target point $(\theta', \boldsymbol{X}', \boldsymbol{Y}')$ lies on the infeasible central trajectory.

How we choose the direction parameter $\beta \in [0, 1]$ differs from an algorithm to an algorithm. One iteration of the Mizuno-Todd-Ye type predictor-corrector algorithm consists of two distinct steps, a predictor step in which we take $\beta = 0$ to reduce the value of θ, and a corrector step in which we take $\beta = 1$ to locate the next point nearer to the infeasible central trajectory. Ideally we want to choose a direction $(d\boldsymbol{X}, d\boldsymbol{Y})$ such that $(\boldsymbol{X}', \boldsymbol{Y}') = (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y})$ satisfies the relations (13) and (14) in both steps. Substituting $(\boldsymbol{X}', \boldsymbol{Y}') = (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y})$ into (13) and eliminating the quadratic term $d\boldsymbol{X}d\boldsymbol{Y}$, we obtain a system of linear equations

$$\left.\begin{array}{l} (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) \in \mathcal{F} + \beta\theta(\boldsymbol{X}_0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}_0 - \bar{\boldsymbol{Y}}), \\ \boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y} - \beta\theta\mu^0\boldsymbol{I} + \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{O}. \end{array}\right\} \tag{15}$$

We may interpret $(d\boldsymbol{X}, d\boldsymbol{Y})$ as "a Newton direction" for approximating a solution $(\boldsymbol{X}', \boldsymbol{Y}')$ of the system (13) of equations at $(\boldsymbol{X}, \boldsymbol{Y})$. It is known, however, that (15) does not have a solution $(d\boldsymbol{X}, d\boldsymbol{Y})$ with both $d\boldsymbol{X}$ and $d\boldsymbol{Y}$ symmetric in general. To avoid this difficulty, we need appropriate modification in the system (15) of equations. All the directions listed in the Introduction share the first equation of (15) when they are applied to the monotone SDLCP (1), but they employ different modifications in the second equation of (15). Our admissible direction which we describe in Section 3.1 may be regarded as an inexact computation of the HRVW/KSH/M direction.

3. Admissible Directions.

3.1 Derivation and Definition.

For each $n \times n$ nonsingular matrix $\boldsymbol{P} \in R^{n\times n}$, define $\boldsymbol{H}_P : R^{n\times n} \to \mathcal{S}$ by

$$\boldsymbol{H}_P(\boldsymbol{A}) \equiv \frac{\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} + (\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^T}{2} \quad \text{for every } \boldsymbol{A} \in R^{n\times n}.$$

The matrix $\boldsymbol{P}$ serves as a scaling matrix, and the image $\boldsymbol{H}_P(\boldsymbol{A})$ of $\boldsymbol{A} \in R^{n\times n}$ is the symmetrization of the scaled matrix $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. Note that $\boldsymbol{H}_P(\boldsymbol{A})$ is linear in $\boldsymbol{A}$; $\boldsymbol{H}_P(\lambda_1\boldsymbol{A}_1 + \lambda_2\boldsymbol{A}_2) = \lambda_1\boldsymbol{H}_P(\boldsymbol{A}_1) + \lambda_2\boldsymbol{H}_P(\boldsymbol{A}_2)$ for every $\lambda_1 \in R$, $\lambda_2 \in R$, $\boldsymbol{A}_1 \in R^{n\times n}$ and $\boldsymbol{A}_2 \in R^{n\times n}$. Using the mapping $\boldsymbol{H}_P$, Monteiro-Zhang [17] (see also Monteiro [13], Zhang [24] and Monteiro [14]) proposed a family of directions, which we have called the MZ family in the Introduction. Let $\theta \in (0, 1]$, $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}(\theta, \boldsymbol{X}^0, \boldsymbol{Y}^0) \bigcap (\mathcal{S}_{++} \times \mathcal{S}_{++})$ and $\beta \in [0, 1]$. For the monotone SDLCP (1), the MZ family is described as the family of directions $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying the system of linear equations (16) and (17) below for some $n \times n$ nonsingular matrix $\boldsymbol{P}$.

$$(\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) \in \mathcal{F} + \beta\theta(\boldsymbol{X}^0 - \bar{\boldsymbol{X}}, \boldsymbol{Y}^0 - \bar{\boldsymbol{Y}}), \tag{16}$$

$$\boldsymbol{H}_P\left(\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y}\right) = \boldsymbol{H}_P\left(\beta\theta\mu^0\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}\right). \tag{17}$$

As special cases, we have the HRVW/KSH/M direction [7, 9, 13], the NT direction [18] and the AHO direction [1, 2] if we take $\boldsymbol{P} = \boldsymbol{X}^{1/2}$, $\boldsymbol{P} = \left(\boldsymbol{X}^{1/2}(\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2})^{-1/2}\boldsymbol{X}^{1/2}\right)^{-1/2}$ and $\boldsymbol{P} = \boldsymbol{I}$, respectively (see, for example, [22]).

Now we introduce an inexact computation of a direction in the MZ family by relaxing (17) into an inequality:

$$\|\text{the left hand side of (17)} - \text{the right hand side of (17)}\|_F \le \kappa\|\text{the right hand side of (17)}\|_F,$$

i.e.,

$$\left\|\boldsymbol{H}_P\left(\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y}\right) - \boldsymbol{H}_P\left(\beta\theta\mu^0\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}\right)\right\|_F \le \kappa\left\|\boldsymbol{H}_P\left(\beta\theta\mu^0\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y}\right)\right\|_F. \tag{18}$$

Here $\kappa \in [0, 1)$ denotes a tolerance parameter in computing an inexact direction.

In this paper we investigate the case where we take $\boldsymbol{P} = \boldsymbol{X}^{1/2}$; the system of equations (16) and (17) defines the HRVW/KSH/M direction. We also see that

$$\mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)=\{(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{F}(\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)\cap(\mathcal{S}_+\times\mathcal{S}_+):\left\|\boldsymbol{H}_{X^{1/2}}(\boldsymbol{XY}-\theta\mu^0\boldsymbol{I})\right\|_F\le\gamma\theta\mu^0\}. \qquad (7)'$$

The lemma below provides us with a bound $\kappa(\gamma+(1-\beta)\sqrt{n})\theta\mu^0$ for the right hand side of (18).

Lemma 3.1. *Suppose that* $(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)$ *for some* $\theta\in(0,1]$. *Let* $\beta\in[0,1]$. *Then*

$$\left\|\boldsymbol{H}_{X^{1/2}}\left(\beta\theta\mu^0\boldsymbol{I}-\boldsymbol{XY}\right)\right\|_F \quad\le\quad (\gamma+(1-\beta)\sqrt{n})\theta\mu^0.$$

Proof: By a simple calculation, we see that

$$\begin{aligned}\left\|\boldsymbol{H}_{X^{1/2}}\left(\beta\theta\mu^0\boldsymbol{I}-\boldsymbol{XY}\right)\right\|_F &= \left\|\boldsymbol{H}_{X^{1/2}}\left(\theta\mu^0\boldsymbol{I}-\boldsymbol{XY}\right)-\boldsymbol{H}_{X^{1/2}}\left((1-\beta)\theta\mu^0\boldsymbol{I}\right)\right\|_F\\ &\le \left(\left\|\boldsymbol{H}_{X^{1/2}}\left(\boldsymbol{XY}-\theta\mu^0\boldsymbol{I}\right)\right\|_F+(1-\beta)\theta\mu^0\left\|\boldsymbol{H}_{X^{1/2}}(\boldsymbol{I})\right\|_F\right)\\ &\le (\gamma+(1-\beta)\sqrt{n})\theta\mu^0.\end{aligned}$$

Here the last inequality follows from $(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)$, (7)' and $\|\boldsymbol{I}\|_F=\sqrt{n}$. ∎

Generalizing the bound $\kappa(\gamma+(1-\beta)\sqrt{n})\theta\mu^0$ for the right hand side of (18), we obtain the inequality

$$\left\|\boldsymbol{H}_{X^{1/2}}\left(\boldsymbol{X}d\boldsymbol{Y}+d\boldsymbol{X}\boldsymbol{Y}\right)-\boldsymbol{H}_{X^{1/2}}\left(\beta\theta\mu^0\boldsymbol{I}-\boldsymbol{XY}\right)\right\|_F\le\left(\kappa\gamma+(1-\beta)\chi(n)\right)\theta\mu^0, \qquad (19)$$

where χ denotes a nonnegative valued function of n. Note that the right hand side of (19) coincides with the bound $\kappa(\gamma+(1-\beta)\sqrt{n})\theta\mu^0$ for the right hand side of (18) if we take $\chi(n)=\kappa\sqrt{n}$. We call $(d\boldsymbol{X},d\boldsymbol{Y})$ satisfying (16) and (19) for some $\kappa\in[0,1)$ and some nonnegative valued function $\chi(n)$ of n an *admissible direction*.

Usually the symmetric square root $\boldsymbol{X}^{1/2}$ of $\boldsymbol{X}\in\mathcal{S}_{++}$ is represented as $\boldsymbol{X}^{1/2}=\boldsymbol{Q}\boldsymbol{\Lambda}^{1/2}\boldsymbol{Q}^T$, where $\boldsymbol{Q}$ denotes an $n\times n$ orthogonal matrix and $\boldsymbol{\Lambda}$ an $n\times n$ diagonal matrix of the positive eigenvalues of $\boldsymbol{X}\in\mathcal{S}_{++}$, *i.e.,* $\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T$ is the eigenvalue decomposition of $\boldsymbol{X}\in\mathcal{S}_{++}$. But if $\boldsymbol{L}$ is an $n\times n$ matrix satisfying $\boldsymbol{X}=\boldsymbol{L}\boldsymbol{L}^T$ then $\boldsymbol{X}^{1/2}=\boldsymbol{L}\boldsymbol{P}$ for some $n\times n$ orthogonal matrix $\boldsymbol{P}$, so that

$$\|\boldsymbol{H}_{X^{1/2}}(\boldsymbol{A})\|_F=\|\boldsymbol{H}_L(\boldsymbol{A})\|_F \quad\text{for every } \boldsymbol{A}\in R^{n\times n}. \qquad (20)$$

Therefore we can apply the Cholesky factorization of $\boldsymbol{X}$ instead of the eigenvalue decomposition of $\boldsymbol{X}$ to a test of whether a given $(d\boldsymbol{X},d\boldsymbol{Y})$ satisfies the inequality (19).

Remark 3.2. To derive admissible directions from the MZ family of directions characterized by (16) and (17), we have relaxed equation (17) into inequality (18), but we have left (16) as it is. Another approach to inexact computation of directions is to relax (16) instead of equation (17). Such a relaxation was done in inexact computation of directions proposed by Mizuno-Jarre [12] for linear programs. For example, we could relax (16) into

$$(\boldsymbol{X}+d\boldsymbol{X},\boldsymbol{Y}+d\boldsymbol{Y})\in\mathcal{F}+\beta\theta\left(\mathcal{U}-(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}})\right)$$

in the monotone SDLCP (1) case, where $\mathcal{U}$ denotes a compact subset of $\mathcal{S}_{++}\times\mathcal{S}_{++}$.

3.2 Special Cases.

Throughout this subsection we assume that the initial point $(\boldsymbol{X}^0,\boldsymbol{Y}^0)\in\mathcal{S}_{++}\times\mathcal{S}_{++}$ is feasible, *i.e.,* $(\boldsymbol{X}^0,\boldsymbol{Y}^0)\in\mathcal{F}\cap(\mathcal{S}_{++}\times\mathcal{S}_{++})$. Then the first condition (16) imposed on an admissible direction turns out to be

$$(\boldsymbol{X}+d\boldsymbol{X},\boldsymbol{Y}+d\boldsymbol{Y})\in\mathcal{F}. \qquad (16)'$$

Evaluating the left hand side of the other condition (19) imposed on an admissible direction, we see that

$$\begin{aligned}&\left\|\boldsymbol{H}_{X^{1/2}}\left(\boldsymbol{X}d\boldsymbol{Y}+d\boldsymbol{X}\boldsymbol{Y}\right)-\boldsymbol{H}_{X^{1/2}}\left(\beta\theta\mu^0\boldsymbol{I}-\boldsymbol{XY}\right)\right\|_F\\ &\quad\le\quad\left\|\boldsymbol{X}^{-1/2}\left(\boldsymbol{X}d\boldsymbol{Y}+d\boldsymbol{X}\boldsymbol{Y}\right)\boldsymbol{X}^{1/2}-\left(\beta\theta\mu^0\boldsymbol{I}-\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\right)\right\|_F.\end{aligned}$$

Table 1: Parameters with which the KSH, the MZ, the Tseng and the MT families are contained in the family of admissible directions .

Family	γ	c_1	c_2	References
KSH	$0 < \gamma < 1/(2(1+\sqrt{2}))$	1	$1+2\sqrt{2}$	[16, Section 3]
MZ	$0 < \gamma < 1/(1+2\sqrt{2})$	1	$2\sqrt{2}$	[14, Section 3]
Tseng	$0 < \gamma < 1/3$	1	2	[23, Section 4]
MT	$0 < \gamma < 1/(2\sqrt{2})$	$\sqrt{2}$	$\sqrt{2}$	[15, Section 2]

Hence if $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfies (16)' and

$$\left\| \boldsymbol{X}^{-1/2} \left(\boldsymbol{X} d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y} \right) \boldsymbol{X}^{1/2} - \left(\beta\theta\mu^0 \boldsymbol{I} - \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} \right) \right\|_F \leq \kappa(\gamma + (1-\beta)\sqrt{n})\theta\mu^0, \tag{19)'}$$

then it is an admissible direction with $\chi(n) = \kappa\sqrt{n}$.

We can prove that every direction in all the families listed in the Introduction, namely, the KSH, the MZ, the Tseng and the MT families satisfy (19)' if we choose a neighborhood parameter γ and a tolerance parameter $\kappa = c_2\gamma/(1 - c_1\gamma)$ as given in Table 1. So they may be regarded as admissible directions. Proof of this fact is given in Appendix. We note that if γ and $\kappa = c_2\gamma/(1 - c_1\gamma)$ are taken as in Table 1 then $0 < \gamma < 1$ and $0 < \kappa < 1$.

4. Step Length.

In this section, we present how we choose a step length along an admissible direction to generate a new point in the predictor and the corrector steps of the infeasible-interior-point algorithm given in Section 5. Let a tolerance parameter $\kappa \in [0, 1)$ and a nonnegative valued function $\chi(n)$ of n be fixed. Suppose that $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{N}(\gamma, \theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ for some $\theta \in (0, 1]$. Choose a direction parameter $\beta \in [0, 1]$; we take $\beta = 0$ for the predictor step and $\beta = 1$ for the corrector step. Let $(d\boldsymbol{X}, d\boldsymbol{Y})$ be an admissible direction. Define

$$\theta(\alpha, \beta) \equiv (1 - (1-\beta)\alpha)\theta \quad \text{and } (\boldsymbol{X}(\alpha), \boldsymbol{Y}(\alpha)) \equiv (\boldsymbol{X}, \boldsymbol{Y}) + \alpha(d\boldsymbol{X}, d\boldsymbol{Y}), \tag{21}$$

where $\alpha \in [0, 1]$ denotes a step length. We need to choose a step length $\alpha \in [0, 1]$ such that the new point $(\boldsymbol{X}(\alpha), \boldsymbol{Y}(\alpha))$ lies in $\mathcal{N}(\gamma, \theta(\alpha, \beta), \boldsymbol{X}^0, \boldsymbol{Y}^0)$, or equivalently it satisfies the relations

$$(\boldsymbol{X}(\alpha), \boldsymbol{Y}(\alpha)) \in \mathcal{F}(\theta(\alpha, \beta), \boldsymbol{X}^0, \boldsymbol{Y}^0), \tag{22}$$

$$\boldsymbol{X}(\alpha) \in \mathcal{S}_+ \quad \text{and } \sigma_0(\alpha, \beta) \leq 0, \tag{23}$$

where

$$\sigma_0(\alpha, \beta) \quad \equiv \quad \left\| \boldsymbol{H}_{X(\alpha)^{1/2}}(\boldsymbol{X}(\alpha)\boldsymbol{Y}(\alpha) - \theta(\alpha, \beta)\mu^0 \boldsymbol{I}) \right\|_F - \gamma\theta(\alpha, \beta)\mu^0.$$

Note that (23) implies $\boldsymbol{Y}(\alpha) \in \mathcal{S}_+$. See also (8).

The relation (22) holds for any $\alpha \in [0, 1]$. In fact, we see that

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

For each $\alpha \in [0, 1]$, we can compute $\sigma_0(\alpha, \beta)$ in $O(n^3)$ arithmetic operations. We first apply the Cholesky factorization to the $n \times n$ symmetric matrix $\boldsymbol{X}(\alpha)$. This requires $O(n^3)$ arithmetic operations. If we fail, we

know that $\boldsymbol{X}(\alpha)$ is not positive semidefinite. Now suppose that we have computed the Cholesky factorization $\boldsymbol{X}(\alpha) = \boldsymbol{L}(\alpha)\boldsymbol{L}(\alpha)^T$ where $\boldsymbol{L}(\alpha)$ is an $n \times n$ lower triangular matrix. Then

$$\sigma_0(\alpha,\beta) \quad = \quad \left\| \boldsymbol{H}_{L(\alpha)}(\boldsymbol{X}(\alpha)\boldsymbol{Y}(\alpha) - \theta(\alpha,\beta)\mu^0\boldsymbol{I}) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0.$$

Using this formula for $\sigma_0(\alpha,\beta)$, we can compute the value of $\sigma_0(\alpha,\beta)$ in $O(n^3)$ arithmetic operations. To choose a legitimate and desirable step length satisfying (23) in the predictor and corrector steps, we could repeatedly apply this procedure. But that would be expensive. To avoid this difficulty, we use the lemma below.

Let $\boldsymbol{L}\boldsymbol{L}^T$ be the Cholesky factorization of $\boldsymbol{X}$. For every $\alpha \in [0,1]$ and $\beta \in [0,1]$, define

$$\begin{aligned}
\sigma_1(\alpha,\beta) \quad &\equiv \quad \left\| \boldsymbol{H}_L \left(\boldsymbol{X}\boldsymbol{Y} + \alpha(\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y}) - \theta(\alpha,\beta)\mu^0\boldsymbol{I} \right) \right\|_F \\
&\quad +\alpha^2 \left\| \boldsymbol{H}_L \left(d\boldsymbol{X}d\boldsymbol{Y} \right) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0, \\
\sigma_2(\alpha,\beta) \quad &\equiv \quad (1-\alpha) \left\| \boldsymbol{H}_L \left(\boldsymbol{X}\boldsymbol{Y} - \theta\mu^0\boldsymbol{I} \right) \right\|_F + \alpha \left\| \boldsymbol{H}_L \left(\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y} - \beta\theta\mu^0\boldsymbol{I} + \boldsymbol{X}\boldsymbol{Y} \right) \right\|_F \\
&\quad +\alpha^2 \left\| \boldsymbol{H}_L \left(d\boldsymbol{X}d\boldsymbol{Y} \right) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0, \\
\sigma_3(\alpha,\beta) \quad &\equiv \quad (1-\alpha) \left\| \boldsymbol{H}_L \left(\boldsymbol{X}\boldsymbol{Y} - \theta\mu^0\boldsymbol{I} \right) \right\|_F + \alpha \left(\kappa\gamma + (1-\beta)\chi(n) \right) \theta\mu^0 \\
&\quad +\alpha^2 \left\| \boldsymbol{H}_L \left(d\boldsymbol{X}d\boldsymbol{Y} \right) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0. \qquad (24)
\end{aligned}$$

Lemma 4.1. *Let $\gamma \in [0,1)$, $\theta \in (0,1]$, and $(d\boldsymbol{X}, d\boldsymbol{Y})$ be an admissible direction at $(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)$.*

(i) For each $\beta \in [0,1]$, the function $\sigma_1(\cdot,\beta) : [0,1] \to R$ is convex.

(ii) For each $\beta \in [0,1]$, the functions $\sigma_2(\cdot,\beta) : [0,1] \to R$ and $\sigma_3(\cdot,\beta) : [0,1] \to R$ are convex and quadratic.

(iii) $\sigma_0(\alpha,\beta) \le \sigma_1(\alpha,\beta) \le \sigma_2(\alpha,\beta) \le \sigma_3(\alpha,\beta)$ for every $\alpha \in [0,1]$ and $\beta \in [0,1]$.

Proof: To prove (i), we only have to show that the first term in the definition of $\sigma_1(\alpha,\beta)$ is convex in α. It can be rewritten in the form $\|\alpha\boldsymbol{B} + \boldsymbol{C}\|_F$ for some $\boldsymbol{B} \in \mathcal{S}$ and $\boldsymbol{C} \in \mathcal{S}$. If $\bar{\alpha}\boldsymbol{B} + \boldsymbol{C} = \boldsymbol{O}$ for some $\bar{\alpha} \in R$, then $\|\alpha\boldsymbol{B} + \boldsymbol{C}\|_F = |\alpha - \bar{\alpha}|\|\boldsymbol{B}\|_F$ for every $\alpha \in R$; hence it is convex in $\alpha \in R$. Otherwise we can easily verify that its second derivative in $\alpha \in R$ is positive for every $\alpha \in R$; hence it is strictly convex in $\alpha \in R$. Assertion (ii) is straightforward. To derive the first inequality of (iii), we observe that

$$\begin{aligned}
\sigma_0(\alpha,\beta) \quad &= \quad \left\| \boldsymbol{H}_{X(\alpha)^{1/2}}(\boldsymbol{X}(\alpha)\boldsymbol{Y}(\alpha) - \theta(\alpha,\beta)\mu^0\boldsymbol{I}) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0 \\
&\le \quad \left\| \boldsymbol{H}_{X^{1/2}}(\boldsymbol{X}(\alpha)\boldsymbol{Y}(\alpha) - \theta(\alpha,\beta)\mu^0\boldsymbol{I}) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0 \\
&\quad \text{(by [24, Lemma 4.2])} \\
&= \quad \left\| \boldsymbol{H}_L(\boldsymbol{X}(\alpha)\boldsymbol{Y}(\alpha) - \theta(\alpha,\beta)\mu^0\boldsymbol{I}) \right\|_F - \gamma\theta(\alpha,\beta)\mu^0 \\
&\le \quad \sigma_1(\alpha,\beta).
\end{aligned}$$

The second inequality of (iii) is straightforward. The third inequality of (iii) follows from (19) and (20). ∎

4.1 Step Length in the Predictor Step.

We will show in Theorem 5.2 that the current iterate $(\boldsymbol{X},\boldsymbol{Y})$ lies in $\mathcal{N}(\tilde{\gamma},\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0)$ for some $\tilde{\gamma} \in (0,\gamma)$ when we apply the predictor step. Let $\beta = 0$. Ideally we want to find the largest step length $\alpha = {}^0\alpha_p \in [0,1]$ that satisfies (23). Using Lemma 4.1, we present some lower bounds for ${}^0\alpha_p$ which we can compute more cheaply than ${}^0\alpha_p$ itself. For each $\ell = 1,2,3$, define

$${}^\ell\alpha_p \equiv \max\left\{ \alpha \in [0,1] \ : \ \sigma_\ell(\alpha,0) \le 0 \right\}.$$

By $\sigma_\ell(0,0) \le 0$ and the convexity of $\sigma_\ell(\cdot,0) : [0,1] \to R$ ((i) and (ii) of Lemma 4.1), we see that $\sigma_\ell(\alpha,0) \le 0$ for every $\alpha \in [0, {}^\ell\alpha_p]$; hence by (iii) of Lemma 4.1, if $0 \le q < \ell \le 3$ then

$$\sigma_q(\alpha,1) \le \sigma_\ell(\alpha,1) \le 0 \ \text{ for every } \alpha \in [0, {}^\ell\alpha_p].$$

Since $(\boldsymbol{X}(0), \boldsymbol{Y}(0)) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$, we see by the continuity of $(\boldsymbol{X}(\alpha), \boldsymbol{Y}(\alpha))$ in $\alpha \in [0,1]$ and the inequality $\sigma_0(\alpha, 0) \leq 0$ for every $\alpha \in [0, {}^{\ell}\alpha_p]$ that $(\boldsymbol{X}(\alpha), \boldsymbol{Y}(\alpha)) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ for every $\alpha \in [0, {}^{\ell}\alpha_p]$. See also (8). Thus we have obtained that $0 \leq {}^{3}\alpha_p \leq {}^{2}\alpha_p \leq {}^{1}\alpha_p \leq {}^{0}\alpha_p$. In Section 5, we require to take a step length $\alpha_p \in [0,1]$ satisfying

$$\boldsymbol{X}(\alpha_p) \in \mathcal{S}_+, \ {}^{3}\alpha_p \leq \alpha_p \ \text{ and } \sigma_0(\alpha_p, 0) \leq 0 \tag{25}$$

to ensure the global convergence of the infeasible-interior-point algorithm presented there. As we have observed above, all the step lengths ${}^{0}\alpha_p$, ${}^{1}\alpha_p$, ${}^{2}\alpha_p$ and ${}^{3}\alpha_p$ fulfill the requirement (25).

To compute ${}^{1}\alpha_p$, we first rewrite the convex function $\sigma_1(\cdot, 0)$ in the form:

$$\sigma_1(\alpha, 0) = \left(a + b\alpha + c\alpha^2\right)^{1/2} + \left\|\boldsymbol{H}_L(d\boldsymbol{X}\, d\boldsymbol{Y})\right\|_F \alpha^2 - (1-\alpha)\gamma\theta\mu^0. \tag{26}$$

Here the a, b and c are determined by $\boldsymbol{X}$, $\boldsymbol{Y}$, $d\boldsymbol{X}$, $d\boldsymbol{Y}$, θ and μ^0 and their computation requires $O(n^3)$ arithmetic operations. Then we apply the binary search to the inequality $\sigma_1(\alpha, 0) \leq 0$ in the interval $[0,1]$ to approximate ${}^{1}\alpha_p$.

On the other hand, we can compute ${}^{2}\alpha_p$ and ${}^{3}\alpha_p$ more easily. In fact, if $\sigma_\ell(1,0) \leq 0$ ($\ell = 2$ or 3), let ${}^{\ell}\alpha_p = 1$; otherwise, let ${}^{\ell}\alpha_p$ ($\ell = 2$ or 3) be the positive root of the quadratic equation $\sigma_\ell(\alpha, 0) = 0$.

4.2 Step Length in the Corrector Step.

In the corrector step, we take $\beta = 1$. Ideally we want to find an $\alpha = {}^{0}\alpha_c \in [0,1]$ that minimize the left hand side $\sigma_0(\alpha, 1)$ of inequality in (23), but such a step length seems difficult to compute in general. For each $\ell = 1, 2, 3$, we define ${}^{\ell}\alpha_c$ to be a minimizer of $\sigma_\ell(\alpha, 1)$ subject to $\alpha \in [0,1]$. By Lemma 4.1 and $\sigma_\ell(0,1) \leq 0$, we then know that $\sigma_0({}^{0}\alpha_c, 1) \leq \sigma_q({}^{\ell}\alpha_c, 1) \leq \sigma_\ell({}^{\ell}\alpha_c, 1) \leq 0$ whenever $0 \leq q < \ell \leq 3$. In Section 5, we require to take a step length $\alpha_c \in [0,1]$ satisfying

$$\boldsymbol{X}(\alpha_c) \in \mathcal{S}_+ \ \text{ and } \sigma_0(\alpha_c, 1) \leq \sigma_3({}^{3}\alpha_c, 1)$$

or equivalently

$$\boldsymbol{X}(\alpha_c) \in \mathcal{S}_+ \ \text{ and } \left\|\boldsymbol{H}_{L(\alpha_c)}\left(\boldsymbol{X}(\alpha_c)\boldsymbol{Y}(\alpha_c) - \theta\mu^0 I\right)\right\|_F \leq \sigma_3({}^{3}\alpha_c, 1) + \gamma\theta\mu^0. \tag{27}$$

All the step lengths ${}^{0}\alpha_c$, ${}^{1}\alpha_c$, ${}^{2}\alpha_c$ and ${}^{3}\alpha_c$ fulfill the requirement (27).

To compute the step length ${}^{1}\alpha_c$, we rewrite the convex function $\sigma_1(\cdot, 1)$ in the form (26), and we then solve a simple one dimensional convex minimization problem:

$$\begin{array}{ll} \text{minimize} & \sigma_0(\alpha, 1) = (a + b\alpha + c\alpha^2)^{1/2} + \left\|\boldsymbol{H}_L(d\boldsymbol{X}\, d\boldsymbol{Y})\right\|_F \alpha^2 - \gamma(1-\alpha)\theta\mu^0 \\ \text{subject to} & \alpha \in [0,1]. \end{array}$$

Since both $\sigma_2(\cdot, 1)$ and $\sigma_3(\cdot, 1)$ are one dimensional convex quadratic functions, we can easily compute their minimizer ${}^{2}\alpha_c$ and ${}^{3}\alpha_c$ over the interval $[0,1]$.

5. Predictor-Corrector Infeasible-Interior-Point Algorithm.

Before we run the predictor-corrector infeasible-interior-point algorithm below, we choose some parameters; an accuracy parameter $\epsilon > 0$, a neighborhood parameter $\gamma \in (0,1)$, a tolerance parameter $\kappa \in [0,1)$ and a nonnegative valued function $\chi(n)$ of n. Define

$$\begin{aligned} \tau_c(\alpha) &\equiv 1 - \alpha(1-\kappa) + \alpha^2 \frac{(1+\gamma)^{1/2}(1+\kappa)^2\gamma}{2(1-\gamma)^{3/2}} \ \text{ for every } \alpha \in [0,1], \\ \hat{\tau}_c &\equiv \min\left\{\tau_c(\alpha) : \alpha \in [0,1]\right\}, \\ \hat{\gamma} &\equiv \hat{\tau}_c\gamma. \end{aligned}$$

Since $(1-\kappa) > 0$, $\tilde{\gamma} < \gamma$. Note that $\tilde{\gamma}$ depends on γ and κ but neither on ω nor on n. We choose an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ such that

$$(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++},\ \mu^0 \equiv \frac{\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0}{n} \ \text{ and } \ \left\| (\boldsymbol{X}^0)^{1/2} \boldsymbol{Y}^0 (\boldsymbol{X}^0)^{1/2} - \mu^0 \boldsymbol{I} \right\|_F \leq \tilde{\gamma} \mu^0. \tag{28}$$

We then formulate a hypothesis that the monotone SDLCP (1) has a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ satisfying (11) for some $\omega \geq 1$. For example, take $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\rho \boldsymbol{I}, \rho \boldsymbol{I})$ for a sufficiently large $\rho > 0$ and $\omega = 2$; then the inequality in (28) holds, and the hypothesis turns out to be: the monotone SDLCP (1) has a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ satisfying $\boldsymbol{X}^* \preceq 2\rho \boldsymbol{I}$ and $\boldsymbol{Y}^* \preceq 2\rho \boldsymbol{I}$.

Algorithm 5.1.

Step 0: Let $\theta^0 = 1$ and $k = 0$.

Step 1: Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^k, \boldsymbol{Y}^k)$ and $\theta = \theta^k$. If $\theta \leq \epsilon$ then stop. If inequality (12) is violated then stop.

Step 2 (Predictor Step): Let $\beta = 0$. Compute an admissible direction $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying (16) and (19). Choose a step length $\alpha_p^k = \alpha_p \in [0,1]$ satisfying (25) as we have mentioned in Section 4.1. Let

$$(\hat{\boldsymbol{X}}^k, \hat{\boldsymbol{Y}}^k) \equiv (\boldsymbol{X}, \boldsymbol{Y}) + \alpha_p^k (d\boldsymbol{X}, d\boldsymbol{Y}) \ \text{ and } \theta^{k+1} = (1 - \alpha_p^k)\theta.$$

Step 3: Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\hat{\boldsymbol{X}}^k, \hat{\boldsymbol{Y}}^k)$ and $\theta = \theta^{k+1}$.

Step 4 (Corrector Step): Let $\beta = 1$. Compute an admissible direction $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying (16) and (19). Choose a step length $\alpha_c^k = \alpha_c \in [0,1]$ satisfying (27) as we have mentioned in Section 4.2, and let

$$(\boldsymbol{X}^{k+1}, \boldsymbol{Y}^{k+1}) \equiv (\boldsymbol{X}, \boldsymbol{Y}) + \alpha_c^k (d\boldsymbol{X}, d\boldsymbol{Y}).$$

Step 5: Replace k by $k+1$. Go to Step 1. ∎

The algorithm generates a sequence $\{(\alpha_p^k, \boldsymbol{X}^k, \boldsymbol{Y}^k, \alpha_c^k, \hat{\boldsymbol{X}}^k, \hat{\boldsymbol{Y}}^k, \theta^k)\}$ such that for every $k = 0, 1, 2, \ldots,$

$$\left.\begin{array}{l} 0 \leq \alpha_p^k \leq 1,\ (\boldsymbol{X}^k, \boldsymbol{Y}^k) \in \mathcal{N}(\gamma, \theta^k, \boldsymbol{X}^0, \boldsymbol{Y}^0), \\ 0 \leq \alpha_c^k \leq 1,\ (\hat{\boldsymbol{X}}^k, \hat{\boldsymbol{Y}}^k) \in \mathcal{N}(\gamma, \theta^{k+1}, \boldsymbol{X}^0, \boldsymbol{Y}^0), \\ \theta^{k+1} = (1 - \alpha_p^k)\theta^k. \end{array}\right\} \tag{29}$$

If it stops at Step 1 with $\theta^k \leq \epsilon$ then $(\boldsymbol{X}^k, \boldsymbol{Y}^k)$ gives an approximate solution of the monotone SDLCP (1) satisfying (10). If it stops at Step 1 violating inequality (12), we know by Lemma 2.3 that the monotone SDLCP (1) does not have a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ satisfying (11).

Theorem 5.2. *Suppose that the monotone SDLCP (1) has a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ satisfying (11).*

(i) $(\boldsymbol{X}^k, \boldsymbol{Y}^k) \in \mathcal{N}(\tilde{\gamma}, \theta^k, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ $(k = 0, 1, \ldots,)$.

(ii) We can find a $\tilde{\alpha}_p \in (0,1]$ (depending on $\gamma,\ \kappa,\ \omega,\ \psi$ and $\chi(n)$) such that $\tilde{\alpha}_p \leq \alpha_p^k$ $(k = 0, 1, \ldots,)$.

(iii) If we take $\chi(n) = O(n)$ (i.e., $0 \leq \chi(n) \leq cn$ for some $c \geq 0$), then we can find a positive number δ (depending on $\gamma,\ \kappa,\ \psi$ and ω but not on n) such that $\delta/n \leq \alpha_p^k$ $(k = 0, 1, \ldots,)$.

(iv) If the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ is feasible and if we take $\chi(n) = O(\sqrt{n})$, then we can find a positive number δ (depending on γ and κ but neither on ω nor on n) such that $\delta/\sqrt{n} \leq \alpha_p^k$ $(k = 0, 1, \ldots,)$.

By the last equality of (29) and (ii) of the theorem, θ^k (hence $\boldsymbol{X}^k \bullet \boldsymbol{Y}^k$) converges at least linearly to zero with a ratio not greater that $(1 - \tilde{\alpha}_p)$. If we take $\chi(n) = O(n)$ then we can ensure that the linear convergence rate is not greater than $(1 - \delta/n)$ for some δ depending on $\gamma,\ \kappa,\ \psi$ and ω but not on n. If we take a feasible initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ and $\chi(n) = O(\sqrt{n})$ then the linear convergence rate is not greater than $(1 - \delta/\sqrt{n})$ for some δ depending on γ and κ but neither on ω nor on n.

We need the following lemma to prove the theorem. Proof of the lemma will be given in the next section.

Lemma 5.3. *Let $\gamma \in [0,1)$, $\kappa \in [0,1)$, and $\chi(n)$ be a nonnegative valued function of n. Suppose that the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ satisfies (28), and that there exists a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the monotone SDLCP (1) satisfying (11) for some $\omega \geq 1$. Let $\beta \in [0,1]$ and $(d\boldsymbol{X}, d\boldsymbol{Y})$ be an admissible direction at $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{N}(\gamma, \theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ for some $\theta \in (0,1]$. Let*

$$\left.\begin{array}{rcl} \varphi_0(\gamma,\omega) & = & \dfrac{(2\omega+1-\gamma)}{1-\gamma}, \\ \psi & = & \begin{cases} 0 & \textit{if } (\boldsymbol{X}^0, \boldsymbol{Y}^0) \textit{ is feasible,} \\ 1 & \textit{otherwise,} \end{cases} \\ \varphi_1(\gamma,\omega,\psi,n) & = & \chi(n) + \sqrt{n} + (\sqrt{2n} + 2\omega\varphi_0(\gamma,\omega)(1+\gamma)n)\psi, \\ \varphi_2(\beta,\gamma,\kappa,\omega,\psi,n) & = & \dfrac{(1+\gamma)^{1/2}}{2(1-\gamma)^{3/2}} \left((1+\kappa)\gamma + (1-\beta)\varphi_1(\gamma,\omega,\psi,n)\right)^2. \end{array}\right\} \tag{30}$$

Then

$$\|\boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}\, d\boldsymbol{Y})\|_F \leq \varphi_2(\beta,\gamma,\kappa,\omega,\psi,n)\theta\mu^0. \tag{31}$$

Proof of Theorem 5.2: (i) We have chosen the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ satisfying (28). Hence $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{N}(\tilde{\gamma}, \theta^0, \boldsymbol{X}^0, \boldsymbol{Y}^0)$. Let $k \geq 0$. We will show that $(\boldsymbol{X}^{k+1}, \boldsymbol{Y}^{k+1}) \in \mathcal{N}(\tilde{\gamma}, \theta^{k+1}, \boldsymbol{X}^0, \boldsymbol{Y}^0)$. Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\hat{\boldsymbol{X}}^k, \hat{\boldsymbol{Y}}^k)$, $\theta = \theta^{k+1}$, $\beta = 1$, and $(d\boldsymbol{X}, d\boldsymbol{Y})$ be an admissible direction. It follows from (29) that $(\boldsymbol{X}, \boldsymbol{Y}) = (\hat{\boldsymbol{X}}^k, \hat{\boldsymbol{Y}}^k) \in \mathcal{N}(\gamma, \theta^{k+1}, \boldsymbol{X}^0, \boldsymbol{Y}^0)$. Define $(\boldsymbol{X}(\alpha), \boldsymbol{Y}(\alpha))$ by (21), and $\sigma_3(\alpha,\beta)$ by (24). In view of the discussion of Section 4.2, it suffices to show that $\sigma_3({}^3\alpha_c, 1) + \gamma\theta\mu^0 \leq \tilde{\gamma}\theta\mu^0$. See (27). For every $\alpha \in [0,1]$, we have that

$$\begin{array}{rcl} \sigma_3({}^3\alpha_c,1) + \gamma\theta\mu^0 & \leq & \sigma_3(\alpha,1) + \gamma\theta\mu^0 \text{ (by the definition of } {}^3\alpha_c) \\ & \leq & (1-\alpha)\gamma\theta\mu^0 + \alpha\kappa\gamma\theta\mu^0 + \alpha^2\varphi_2(1,\gamma,\kappa,\omega,\psi,n)\theta\mu^0 \\ & & \text{(by } (\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{N}(\gamma,\theta,\boldsymbol{X}^0,\boldsymbol{Y}^0) \text{ and Lemma5.3)} \\ & = & \left((1-\alpha(1-\kappa)) + \alpha^2\dfrac{(1+\gamma)^{1/2}(1+\kappa)^2\gamma}{2(1-\gamma)^{3/2}}\right)\gamma\theta\mu^0 \\ & = & \tau_c(\alpha)\gamma\theta\mu^0. \end{array}$$

Hence we obtain that $\sigma_3({}^3\alpha_c, 1) + \gamma\theta\mu^0 \leq \min\left\{\tau_c(\alpha)\gamma\theta\mu^0 : \alpha \in [0,1]\right\} = \tilde{\gamma}\theta\mu^0$.

(ii) Let $\beta = 0$. Define

$$\begin{array}{rcl} \tau_p(\alpha) & \equiv & (1-\alpha)\tilde{\gamma} + \alpha(\kappa\gamma + \chi(n)) + \alpha^2\varphi_2(0,\gamma,\kappa,\omega,\psi,n) - (1-\alpha)\gamma \\ & = & -\gamma + \tilde{\gamma} + \alpha\left(\gamma - \tilde{\gamma} + \kappa\gamma + \chi(n)\right) + \alpha^2\varphi_2(0,\gamma,\kappa,\omega,\psi,n), \\ \tilde{\alpha}_p & \equiv & \max\left\{\alpha \in [0,1] : \tau_p(\alpha) \leq 0\right\}. \end{array}$$

Then $\tilde{\alpha}_p > 0$ because $\tau_p(0) = -\gamma + \tilde{\gamma} < 0$. Recall that $\tilde{\gamma}$ depends on γ and κ; hence $\tilde{\alpha}_p$ depends on γ, κ, ω, ψ and $\chi(n)$. Let $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^k, \boldsymbol{Y}^k)$, $\theta = \theta^k$, $\beta = 0$, and $(d\boldsymbol{X}, d\boldsymbol{Y})$ be an admissible direction. In view of the discussion of Section 4.2, it suffices to show that $\sigma_3(\tilde{\alpha}_p, 0) \leq 0$; then $\tilde{\alpha}_p \leq {}^3\alpha_3 \leq \alpha_p$ follows. By (i) of Theorem 5.2 and Lemma 5.3, we see that

$$\begin{array}{rcl} \sigma_3(\tilde{\alpha}_p,0) & \leq & (1-\tilde{\alpha}_p)\tilde{\gamma}\theta\mu^0 + \tilde{\alpha}_p(\kappa\gamma + \chi(n))\theta\mu^0 + \tilde{\alpha}_p^2\varphi_2(0,\gamma,\kappa,\omega,\psi,n)\theta\mu^0 - \gamma(1-\tilde{\alpha}_p)\theta\mu^0 \\ & = & \tau_p(\tilde{\alpha}_p)\theta\mu^0 \;\leq\; 0. \end{array}$$

(iii) Suppose that $\chi(n) = O(n)$. Then $\tau_p(\alpha)$ defined above is of the form

$$\tau_p(\alpha) = -\gamma + \tilde{\gamma} + \alpha O(n) + \alpha^2 O(n^2).$$

Here $O(n)$ and $O(n^2)$ depend on γ, κ, ψ and ω. Hence we can find a positive number δ (depending on γ, κ, ψ and ω) such that $\tilde{\alpha}_p \equiv \max\{\alpha \in [0,1] : \tau_p(\alpha) \leq 0\} \geq \delta/n$.

(iv) Suppose that $(\boldsymbol{X}^0, \boldsymbol{Y}^0)$ is feasible and that $\chi(n) = O(\sqrt{n})$. Then $\psi = 0$ and $\tau_p(\alpha)$ defined in (ii) is of the form

$$\tau_p(\alpha) = -\gamma + \tilde{\gamma} + \alpha O(\sqrt{n}) + \alpha^2 O(n).$$

Here $O(\sqrt{n})$ and $O(n)$ depend on γ and κ. Hence we can find a positive number δ (depending on γ and κ) such that $\tilde{\alpha}_p \equiv \max\{\alpha \in [0,1] : \tau_p(\alpha) \leq 0\} \geq \delta/\sqrt{n}$. ∎

Figure 2: Relation between Lemmas

$$\left.\begin{array}{r} \\ \\ \text{Lemma 3.1} \\ \text{[9, Lemma 7.5] and} \end{array}\right. \begin{array}{c} \\ \\ \Longrightarrow \\ \Longrightarrow \end{array} \left.\begin{array}{l} \text{Lemma 6.2} \\ \text{Lemma 6.3} \\ \text{Lemma 6.1} \\ \text{Lemma 2.3} \end{array}\right. \begin{array}{c} \\ \\ \Longrightarrow \\ \Longrightarrow \end{array} \left.\begin{array}{l} \\ \\ \text{Lemma 6.4} \\ \text{Lemma 6.5} \end{array}\right\} \Longrightarrow \text{Lemma 5.3}$$

Remark 5.4. Suppose that the neighborhood parameter $\gamma \in (0,1)$ and the tolerance parameter $\kappa \in [0,1)$ satisfies

$$\tau_c(1) \equiv \kappa + \frac{(1+\gamma)^{1/2}(1+\kappa)^2\gamma}{2(1-\gamma)^{3/2}} < 1.$$

If we replace $\tilde{\gamma} = \tilde{\tau}_c\gamma$ by $\tilde{\gamma} = \tau_c(1)\gamma$, then $\tilde{\gamma} < \gamma$. In this case, we can take the step length $\alpha_c^k = 1$ in Step 4 (Corrector Step). All the results in Theorem 5.2 remain valid.

6. Proof of Lemma 5.3.

We will state a series of lemmas, which leads to a proof of Lemma 5.3, under the assumption of Lemma 5.3. They are based on or closely related to lemmas given in the paper [24]; Lemma 6.1 is corresponding to [24, Lemma 3.2], Lemma 6.2 to [24, Lemma 3.3], Lemma 6.3 to [24, Lemma 3.5], Lemma 6.4 to [24, Lemma 3.6] and Lemma 6.5 to [24, Lemma 3.7], respectively. But our notation is slightly different from the one used in the paper [24]. Figure 2 shows logical relation between the lemmas, where "A $\Longrightarrow$ B" means that the proof of B relies on A.

In addition to the notation and symbol listed in (30) of Lemma 5.3, we use:

$$\begin{aligned}
&\boldsymbol{A} \otimes \boldsymbol{B} \ : \text{ the Kronecker product of } \boldsymbol{A}, \boldsymbol{B} \in R^{n\times n},\\
&\boldsymbol{A}_{\cdot i} \ : \text{ the } i\text{-th column of } \boldsymbol{A} \in R^{n\times n},\\
&\mathbf{vec}\boldsymbol{A} \equiv \begin{pmatrix} \boldsymbol{A}_{\cdot 1} \\ \boldsymbol{A}_{\cdot 2} \\ \vdots \\ \boldsymbol{A}_{\cdot n} \end{pmatrix} \in \mathcal{R}^{n^2} \text{ for every } \boldsymbol{A} \in R^{n\times n},\\
&\boldsymbol{R}_c \equiv \boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{X}d\boldsymbol{Y}), \quad \boldsymbol{E} \equiv \boldsymbol{X}^{1/2} \otimes \boldsymbol{X}^{1/2},\\
&\boldsymbol{S} \equiv \frac{1}{2}\left(\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\right), \quad \boldsymbol{F} \equiv \boldsymbol{S}\boldsymbol{E}^{-1}, \quad \boldsymbol{D} \equiv \boldsymbol{S}^{-1/2}\boldsymbol{F}.
\end{aligned}$$

It follows from the assumption $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{N}(\gamma, \theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ that all the eigenvalues of $\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}$ are bounded by the positive number $(1-\gamma)\theta\mu^0$ from below and by the number $(1+\gamma)\theta\mu^0$ from above (see (8)). This implies that all the eigenvalues of the symmetric matrix $\boldsymbol{S} = \frac{1}{2}\left(\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\right)$ are bounded by the same numbers from below and from above, respectively;

$$(1-\gamma)\theta\mu^0 \leq \text{all the eigenvalues of } \boldsymbol{S} \leq (1+\gamma)\theta\mu^0. \tag{32}$$

See the book [6] for basic properties on the Kronecker product.

Lemma 6.1. *(i)* $\|\mathbf{vec}\boldsymbol{R}_c\| \leq \left(\gamma(1+\kappa) + (1-\beta)(\sqrt{n} + \chi(n))\right)\theta\mu^0$.

(ii) $\left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2 \leq \dfrac{\theta\mu^0}{(1-\gamma)}\left(\gamma(1+\kappa) + (1-\beta)(\sqrt{n} + \chi(n))\right)^2$.

Proof: To prove (i), we see that

$$\begin{aligned}
\|\mathbf{vec}\boldsymbol{R}_c\| &= \|\boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{X}d\boldsymbol{Y})\|_F \\
&\le \left\|\boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{X}d\boldsymbol{Y})-\beta\theta\mu^0\boldsymbol{I}+\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\right\|_F+\left\|\beta\theta\mu^0-\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\right\|_F \\
&\le \left(\gamma(1+\kappa)+(1-\beta)(\sqrt{n}+\chi(n))\right)\theta\mu^0 \quad \text{(by (19) and Lemma 3.1).}
\end{aligned}$$

We then obtain that

$$\begin{aligned}
\left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2 &= (\mathbf{vec}\boldsymbol{R}_c)^T\boldsymbol{S}^{-1}(\mathbf{vec}\boldsymbol{R}_c) \\
&\le \frac{1}{(1-\gamma)\theta\mu^0}\|\mathbf{vec}\boldsymbol{R}_c\|^2 \quad \text{(by (32))} \\
&\le \frac{\theta\mu^0}{(1-\gamma)}\left(\gamma(1+\kappa)+(1-\beta)(\sqrt{n}+\chi(n))\right)^2.
\end{aligned}$$

∎

Lemma 6.2. $\|\boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}d\boldsymbol{Y})\|_F \le \frac{1}{2}\sqrt{\frac{1+\gamma}{1-\gamma}}\left(\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\|^2+\left\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right\|^2\right).$

Proof: By the definitions of the matrices $\boldsymbol{E}$, $\boldsymbol{F}$, $\boldsymbol{S}$ and $\boldsymbol{D}$, we see that

$$\begin{aligned}
\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\|^2 &= (\mathbf{vec}d\boldsymbol{X})^T\boldsymbol{F}^T\boldsymbol{S}^{-1}\boldsymbol{F}(\mathbf{vec}d\boldsymbol{X}) \\
&= (\mathbf{vec}d\boldsymbol{X})^T\boldsymbol{E}^{-T}\boldsymbol{S}\boldsymbol{E}^{-1}(\mathbf{vec}d\boldsymbol{X}) \\
&\ge (1-\gamma)\theta\mu^0\left\|\boldsymbol{E}^{-1}(\mathbf{vec}d\boldsymbol{X})\right\|^2 \quad \text{(by (32))} \\
&= (1-\gamma)\theta\mu^0\left\|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\right\|_F^2, \\
\left\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right\|^2 &= (\mathbf{vec}d\boldsymbol{Y})^T\boldsymbol{E}\boldsymbol{S}^{-1}\boldsymbol{E}(\mathbf{vec}d\boldsymbol{Y}) \\
&\ge \frac{1}{(1+\gamma)\theta\mu^0}\|\boldsymbol{E}(\mathbf{vec}d\boldsymbol{Y})\|^2 \quad \text{(by (32))} \\
&= \frac{1}{(1+\gamma)\theta\mu^0}\left\|\boldsymbol{X}^{1/2}d\boldsymbol{Y}\boldsymbol{X}^{1/2}\right\|_F^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}d\boldsymbol{Y})\|_F &= \frac{1}{2}\left\|\boldsymbol{X}^{-1/2}(d\boldsymbol{X}d\boldsymbol{Y})\boldsymbol{X}^{1/2}+\boldsymbol{X}^{1/2}(d\boldsymbol{Y}d\boldsymbol{X})\boldsymbol{X}^{-1/2}\right\|_F \\
&\le \left\|\boldsymbol{X}^{-1/2}(d\boldsymbol{X}d\boldsymbol{Y})\boldsymbol{X}^{1/2}\right\|_F \\
&\le \left\|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\right\|_F\left\|\boldsymbol{X}^{1/2}d\boldsymbol{Y}\boldsymbol{X}^{1/2}\right\|_F \\
&\le \sqrt{\frac{1+\gamma}{1-\gamma}}\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\|\left\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right\| \\
&\le \frac{1}{2}\sqrt{\frac{1+\gamma}{1-\gamma}}\left(\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\|^2+\left\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right\|^2\right).
\end{aligned}$$

∎

Lemma 6.3. *Define*

$$\begin{aligned}
\xi &\equiv (1-\beta)\psi\theta\left(\left\|\boldsymbol{D}\mathbf{vec}(\boldsymbol{X}^0-\boldsymbol{X}^*)\right\|+\left\|\boldsymbol{D}^{-T}\mathbf{vec}(\boldsymbol{Y}^0-\boldsymbol{Y}^*)\right\|\right), \\
\eta &\equiv \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2+2\left((1-\beta)\psi\theta\right)^2(\boldsymbol{X}^0-\boldsymbol{X}^*)\bullet(\boldsymbol{Y}^0-\boldsymbol{Y}^*).
\end{aligned}$$

Then $\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\|^2+\left\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right\|^2 \le \left(\xi+\sqrt{\xi^2+\eta}\right)^2.$

Proof: For simplicity of notation, we write $\zeta \equiv \left(\left\| \boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{X} \right\|^2 + \|\boldsymbol{D}\mathbf{vec}d\boldsymbol{Y}\|^2 \right)^{1/2}$. By assumption, we know that

$$\begin{aligned}(\boldsymbol{X}', \boldsymbol{Y}') &\equiv (\boldsymbol{X}, \boldsymbol{Y}) - \theta\psi((\boldsymbol{X}^0 - \boldsymbol{X}^*, \boldsymbol{Y}^0 - \boldsymbol{Y}^*) \in \mathcal{F}, \\ (\boldsymbol{X}'', \boldsymbol{Y}'') &\equiv (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) - \beta\theta\psi(\boldsymbol{X}^0 - \boldsymbol{X}^*, \boldsymbol{Y}^0 - \boldsymbol{Y}^*) \in \mathcal{F}.\end{aligned}$$

By the monotonicity condition (2) imposed on $\mathcal{F}$, we have that

$$\begin{aligned}0 &\leq (\boldsymbol{X}'' - \boldsymbol{X}') \bullet (\boldsymbol{Y}'' - \boldsymbol{Y}') \\ &= \left(d\boldsymbol{X} + (1-\beta)\psi\theta(\boldsymbol{X}^0 - \boldsymbol{X}^*)\right) \bullet \left(d\boldsymbol{Y} + (1-\beta)\psi\theta(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\right).\end{aligned}$$

Hence

$$\begin{aligned}&(\boldsymbol{D}\mathbf{vec}d\boldsymbol{X})^T(\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}) \\ &= d\boldsymbol{X} \bullet d\boldsymbol{Y} \\ &\geq -(1-\beta)\psi\theta\left(d\boldsymbol{X} \bullet (\boldsymbol{Y}^0 - \boldsymbol{Y}^*) + (\boldsymbol{X}^0 - \boldsymbol{X}^*) \bullet d\boldsymbol{Y}\right) \\ &\quad -\left((1-\beta)\psi\theta\right)^2 (\boldsymbol{X}^0 - \boldsymbol{X}^*) \bullet (\boldsymbol{Y}^0 - \boldsymbol{Y}^*) \quad \text{(by the inequality above)} \\ &= -(1-\beta)\psi\theta\left((\boldsymbol{D}\mathbf{vec}d\boldsymbol{X})^T \boldsymbol{D}^{-T}\mathbf{vec}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*) + \left(\boldsymbol{D}\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\right)^T \boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right) \\ &\quad -\frac{1}{2}\left(\eta - \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2\right) \quad \text{(by the definition of } \eta) \\ &\geq -(1-\beta)\psi\theta\left(\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\| \left\|\boldsymbol{D}^{-T}\mathbf{vec}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\right\| + \left\|\boldsymbol{D}\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\right\| \left\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\right\|\right) \\ &\quad -\frac{1}{2}\left(\eta - \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2\right) \\ &\geq -(1-\beta)\psi\theta\left(\left\|\boldsymbol{D}\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\right\| + \left\|\boldsymbol{D}^{-T}\mathbf{vec}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\right\|\right)\zeta \\ &\quad -\frac{1}{2}\left(\eta - \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2\right) \quad \text{(by the definition of } \zeta) \\ &= -\xi\zeta - \frac{1}{2}\left(\eta - \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2\right).\end{aligned}$$

On the other hand, we see from the definition of $\boldsymbol{R}_c$ that $\mathbf{vec}\boldsymbol{R}_c = \boldsymbol{E}\mathbf{vec}d\boldsymbol{Y} + \boldsymbol{F}\mathbf{vec}d\boldsymbol{X}$. Hence $\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c = \boldsymbol{D}\mathbf{vec}d\boldsymbol{X} + \boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}$. Therefore

$$\begin{aligned}\zeta^2 &= \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2 - 2(\boldsymbol{D}\mathbf{vec}d\boldsymbol{X})^T(\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}) \\ &\leq \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2 + 2\left(\xi\zeta + \frac{1}{2}\left(\eta - \left\|\boldsymbol{S}^{-1/2}\mathbf{vec}\boldsymbol{R}_c\right\|^2\right)\right) \\ &= 2\xi\zeta + \eta.\end{aligned}$$

This inequality implies that $\zeta \leq \xi + \sqrt{\xi^2 + \eta}$. ∎

In the following two lemmas, we evaluate the quantities η and ξ given in Lemma 6.3.

Lemma 6.4. *Define* $\bar{\eta} = \dfrac{\theta\mu^0}{1-\gamma}\left((1+\kappa)\gamma + (1-\beta)(\chi(n) + \sqrt{n} + \sqrt{2n}\psi)\right)^2$. *Then* $\eta \leq \bar{\eta}$.

Proof: First we observe that

$$\begin{aligned}\theta^2(\boldsymbol{X}^0 - \boldsymbol{X}^*) \bullet (\boldsymbol{Y}^0 - \boldsymbol{Y}^*) &= \theta^2(\boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 - \boldsymbol{X}^0 \bullet \boldsymbol{Y}^* - \boldsymbol{X}^* \bullet \boldsymbol{Y}^0) \\ &\leq \theta^2 \boldsymbol{X}^0 \bullet \boldsymbol{Y}^0 = n\theta^2\mu^0 \leq n\theta\mu^0.\end{aligned}$$

By the inequality above and Lemma 6.1, we have

$$\begin{aligned}
\eta &= \left\| \boldsymbol{S}^{-1/2} \mathbf{vec} \boldsymbol{R}_c \right\|^2 + 2\left((1-\beta)\psi\theta\right)^2 (\boldsymbol{X}^0 - \boldsymbol{X}^*) \bullet (\boldsymbol{Y}^0 - \boldsymbol{Y}^*) \\
&\leq \frac{1}{(1-\gamma)} \left((1+\kappa)\gamma + (1-\beta)(\chi(n) + \sqrt{n})\right)^2 \theta\mu^0 + 2\left((1-\beta)\psi\right)^2 n\theta\mu^0 \\
&\leq \frac{\theta\mu^0}{1-\gamma} \left((1+\kappa)\gamma + (1-\beta)(\chi(n) + \sqrt{n} + \sqrt{2n}\psi)\right)^2 \\
&= \bar{\eta}.
\end{aligned}$$

∎

Lemma 6.5. *Define* $\bar{\xi} = \sqrt{\dfrac{\theta\mu^0}{1-\gamma}}\omega\varphi_0(\gamma,\omega)(1+\gamma)(1-\beta)\psi n$. *Then* $\xi \leq \bar{\xi}$.

Proof: By the definition of $\boldsymbol{E}$, $\boldsymbol{S}$, $\boldsymbol{F}$, we see that

$$\begin{aligned}
&\left\| \boldsymbol{D}\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*) \right\|^2 \\
&= (\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*))^T \boldsymbol{F}^T \boldsymbol{S}^{-1} \boldsymbol{F}^T \mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*) \\
&= (\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*))^T \boldsymbol{E}^{-1} \boldsymbol{S} \boldsymbol{E}^{-1} \mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*) \\
&= (\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*))^T (\boldsymbol{Y}^{1/2} \otimes \boldsymbol{Y}^{1/2})(\boldsymbol{Y}^{-1/2}\boldsymbol{X}^{-1/2} \otimes \boldsymbol{Y}^{-1/2}\boldsymbol{X}^{-1/2})\boldsymbol{S} \\
&\quad (\boldsymbol{X}^{-1/2}\boldsymbol{Y}^{-1/2} \otimes \boldsymbol{X}^{-1/2}\boldsymbol{Y}^{-1/2})(\boldsymbol{Y}^{1/2} \otimes \boldsymbol{Y}^{1/2})(\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*)) \\
&= \frac{1}{2}(\mathbf{vec}\boldsymbol{Y}^{1/2}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\boldsymbol{Y}^{1/2})^T (\boldsymbol{I} \otimes \boldsymbol{Y}^{-1/2}\boldsymbol{X}^{-1}\boldsymbol{Y}^{-1/2} + \boldsymbol{Y}^{-1/2}\boldsymbol{X}^{-1}\boldsymbol{Y}^{-1/2} \otimes \boldsymbol{I}) \\
&\quad (\mathbf{vec}\boldsymbol{Y}^{1/2}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\boldsymbol{Y}^{1/2}) \\
&\leq \frac{1}{(1-\gamma)\theta\mu^0} \left\| \boldsymbol{Y}^{1/2}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\boldsymbol{Y}^{1/2} \right\|_F^2 .
\end{aligned}$$

Here the last inequality follows from the fact that all the eigenvalues of the matrix $\boldsymbol{Y}^{1/2}\boldsymbol{X}\boldsymbol{Y}^{1/2}$ are bounded by the positive number $(1-\gamma)\theta\mu^0$ from below. Similarly we see that

$$\begin{aligned}
\left\| \boldsymbol{D}^{-T} \mathbf{vec}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*) \right\|^2 &= (\mathbf{vec}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*))^T \boldsymbol{E}\boldsymbol{S}^{-1}\boldsymbol{E}\mathbf{vec}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*) \\
&\leq \frac{1}{(1-\gamma)\theta\mu^0} \left\| \boldsymbol{X}^{1/2}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\boldsymbol{X}^{1/2} \right\|_F^2 .
\end{aligned}$$

Thus we obtain that

$$\begin{aligned}
\xi &= (1-\beta)\psi\theta \left(\left\| \boldsymbol{D}\mathbf{vec}(\boldsymbol{X}^0 - \boldsymbol{X}^*) \right\| + \left\| \boldsymbol{D}^{-T}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*) \right\| \right) \\
&\leq (1-\beta)\psi\theta \frac{1}{\sqrt{(1-\gamma)\theta\mu^0}} \left(\left\| \boldsymbol{Y}^{1/2}(\boldsymbol{X}^0 - \boldsymbol{X}^*)\boldsymbol{Y}^{1/2} \right\|_F + \left\| \boldsymbol{X}^{1/2}(\boldsymbol{Y}^0 - \boldsymbol{Y}^*)\boldsymbol{X}^{1/2} \right\|_F \right) \\
&\leq \frac{(1-\beta)\psi\theta}{\sqrt{(1-\gamma)\theta\mu^0}} \left(\omega(\boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}^0) \right) \text{ (by [9, Lemma 7.5])} \\
&\leq \sqrt{\frac{\theta\mu^0}{1-\gamma}}(1-\beta)\psi\omega\varphi_0(\gamma,\omega)n(1+\gamma) \text{ (by Lemma2.3)} \\
&= \bar{\xi}.
\end{aligned}$$

∎

Now we are ready to prove Lemma 5.3:

$$\left\| \boldsymbol{H}_{X^{1/2}}(d\boldsymbol{X}\, d\boldsymbol{Y}) \right\|_F$$

$$
\begin{aligned}
&\leq \frac{1}{2}\sqrt{\frac{1+\gamma}{1-\gamma}}\left(\|\boldsymbol{D}\mathbf{vec}d\boldsymbol{X}\|^2+\|\boldsymbol{D}^{-T}\mathbf{vec}d\boldsymbol{Y}\|^2\right)\text{ (by Lemma 6.2)}\\
&\leq \frac{1}{2}\sqrt{\frac{1+\gamma}{1-\gamma}}\left(\xi+\sqrt{\xi^2+\eta}\right)^2\text{ (by Lemma 6.3)}\\
&\leq \frac{1}{2}\sqrt{\frac{1+\gamma}{1-\gamma}}\left(\bar{\xi}+\sqrt{\bar{\xi}^2+\bar{\eta}}\right)^2\text{ (by Lemmas 6.4 and 6.5)}\\
&\leq \frac{1}{2}\sqrt{\frac{1+\gamma}{1-\gamma}}\left(\bar{\xi}+\bar{\xi}+\sqrt{\bar{\eta}}\right)^2\text{ (since } 0<\bar{\xi} \text{ and } 0<\bar{\eta})\\
&= \frac{(1+\gamma)^{1/2}}{2(1-\gamma)^{3/2}}\left((1+\kappa)\gamma+(1-\beta)\left((\chi(n)+\sqrt{n}+(\sqrt{2n}+2\omega\varphi_0(\gamma,\omega)(1+\gamma)n)\psi\right)\right)^2\theta\mu^0\\
&= \varphi_2(\beta,\gamma,\kappa,\psi,n)\theta\mu^0.
\end{aligned}
$$

7. Concluding Remarks.

By definition, an admissible direction needs to satisfy (16). We present a method for forcing an inexact direction to satisfy (16) in the SDP case. We assume that the set of data matrices $\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots,\boldsymbol{A}_m$ is linearly independent. In the SDP case, (16) turns out to be

$$
\left.\begin{array}{l}
\displaystyle\sum_{i=1}^m \boldsymbol{A}_i(x_i+dx_i)-(\boldsymbol{X}+d\boldsymbol{X})=(1-\theta)\boldsymbol{A}_0+\theta\left(\sum_{i=1}^m\boldsymbol{A}_ix_i^0-\boldsymbol{X}^0\right),\\
\boldsymbol{A}_i\bullet(\boldsymbol{Y}+d\boldsymbol{Y})=(1-\theta)a_i+\theta\boldsymbol{A}_i\bullet\boldsymbol{Y}^0\ (i=1,2,\ldots,m).
\end{array}\right\}
\tag{33}
$$

Hence an admissible direction $(d\boldsymbol{x},d\boldsymbol{X},d\boldsymbol{Y})$ is required to exactly satisfy this system (33) of linear equations. Let $(d\boldsymbol{x}',d\boldsymbol{X}',d\boldsymbol{Y}')$ be an inexact direction which does not necessarily satisfy (33). Define

$$
d\boldsymbol{x}\equiv d\boldsymbol{x}'\ \text{ and }\ d\boldsymbol{X}\equiv\boldsymbol{X}-\sum_{i=1}^m\boldsymbol{A}_i(x_i+dx_i)+(1-\theta)\boldsymbol{A}_0+\theta\left(\sum_{i=1}^m\boldsymbol{A}_ix_i^0-\boldsymbol{X}^0\right).
$$

Then $(d\boldsymbol{x},d\boldsymbol{X})$ satisfies the first linear equation of (33). To compute $d\boldsymbol{Y}$ which not only satisfies the rest of the linear equations in (33) but also approximates $d\boldsymbol{Y}'$, we propose to solve the least square problem:

$$
\begin{array}{ll}
\text{minimize} & \left\|d\boldsymbol{Y}-d\boldsymbol{Y}'\right\|_F^2\\
\text{subject to} & \boldsymbol{A}_i\bullet(\boldsymbol{Y}+d\boldsymbol{Y})=(1-\theta)a_i+\theta\boldsymbol{A}_i\bullet\boldsymbol{Y}^0\ (i=1,2,\ldots,m).
\end{array}
$$

The computation of the minimizer $d\boldsymbol{Y}$ of this problem is reduced to the system of linear equations

$$
\boldsymbol{B}'\boldsymbol{\lambda}=\boldsymbol{d}'\ \text{ and }\ d\boldsymbol{Y}=d\boldsymbol{Y}'+\sum_{i=1}^m\lambda_i\boldsymbol{A}_i,
\tag{34}
$$

where $\boldsymbol{B}'$ denotes an $m\times m$ symmetric matrix whose (i,j)th element B'_{ij} is given by $\boldsymbol{A}_i\bullet\boldsymbol{A}_j$, and $\boldsymbol{d}'$ an m-dimensional vector determined by the $\boldsymbol{A}_i$, a_i, $\boldsymbol{Y}$, $\boldsymbol{Y}^0$, $d\boldsymbol{Y}'$ and θ. By the assumption that the set of data matrices $\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots,\boldsymbol{A}_m$ is linearly independent, the symmetric matrix $\boldsymbol{B}'$ is positive definite, so that we can apply the Cholesky factorization to the first equation of (34).

As we have mentioned in the Introduction, the computation of the HRVW/KSH/M and the NT directions is reduced to a linear equation $\boldsymbol{B}d\boldsymbol{y}=\boldsymbol{d}$ similar to the first equation of (34). But, even when the data matrices $\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots,\boldsymbol{A}_m$ are sparse, its $m\times m$ positive definite coefficient matrix $\boldsymbol{B}$ is completely dense in general; this makes the exact computation of the directions difficult when m is large. On the other hand, when the data matrices $\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots,\boldsymbol{A}_m$ are sparse, the $m\times m$ symmetric positive definite coefficient matrix $\boldsymbol{B}'$ in (34) is expected to be sparse; so solving (34) is expected to be much easier than the exact computation of the HRVW/KSH/M and the NT directions.

Acknowledgment. The authors would like to thank Renato Monteiro for many interesting discussions, which have been very useful in preparing this article.

References

[1] F. Alizadeh, J.-P.A. Haeberly, and M.L. Overton. Primal-dual interior-point methods for semidefinite programming. Technical Report 659, Computer Science Department, Courant Institute of Mathematical Sciences, New York University, 1994.

[2] F. Alizadeh, J.-P.A. Haeberly, and M.L. Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. Technical Report 721, Computer Science Dept., New York University, New York, 1996.

[3] K. Fujisawa, M. Kojima, and K. Nakata. SDPA (semidefinite programming algorithm) – user's manual –. Technical Report B-308, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1995, Revised August 1996.

[4] K. Fujisawa, M. Kojima, and K. Nakata. Exploiting sparsity in primal-dual interior-point-methods for semidefinite programming. Technical Report B-324, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, January 1997.

[5] D. Goldfarb and K. Scheinberg, Interior point trajectories in semidefinite programming. Working Paper, Dept. of IEOR, Columbia University, New York, NY, March 1996, Revised February 1997.

[6] A. Graham. *Kronecker Products and Matrix Calculus with Applications*. Ellis Horwood Limited, John Willey & Sons, New York, 1981.

[7] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM J. on Optimization, Vol.6*, pages 342–361, 1996.

[8] M. Kojima, M. Shida, and S. Shindoh. Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Technical Report B-306, Information Sciences, Tokyo Institute of Technology, December 1995, Revised October 1996.

[9] M. Kojima, S. Shindoh, and S. Hara. Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM J. on Optimization, Vol.7*, pages 86–125, 1997.

[10] C.-J. Lin and R. Saigal. A predictor-corrector method for semi-definite linear programming. Working paper, Dept. of Industrial Engineering and Operations Engineering, The University of Michigan, Ann Arbor, Michigan, October 1995.

[11] C.-J. Lin and R. Saigal. Predictor-corrector method for semidefinite programming. Informs Atlanta fall meeting, Novembers 3-6, 1996.

[12] S. Mizuno and F. Jarre. Global and polynomial-time convergence of an infeasible-interior-point algorithm using inexact computation. Technical Report 605, The Institute of Statistical Mathematics, Tokyo, 1996.

[13] R.D.C. Monteiro. Primal-dual path-following algorithms for semidefinite programming. *SIAM J. on Optimization*. To appear.

[14] R.D.C. Monteiro. Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions. Technical report, School of Industrial and Systems Engineering, Georgia Tech. USA, 1996.

[15] R.D.C. Monteiro and T. Tsuchiya. Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming. Technical Report 627, The Institute of Statistical Mathematics, Tokyo, Japan, 1996.

[16] R.D.C. Monteiro and T. Tsuchiya. Polynomiality of primal-dual algorithms for semidefinite linear complementarity problems based on the Kojima-Shindoh-Hara family of directions. Technical report, School of Industrial and Systems Engineering, Georgia Tech. USA, 1996.

[17] R.D.C. Monteiro and Y. Zhang. A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. Technical report, School of Industrial and Systems Engineering, Georgia Tech. USA, 1996.

[18] Yu.E. Nesterov and M.J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. on Optimization*. To appear.

[19] F.A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. Technical report, Computational Mathematics, Dept. of Mathematics, The University of Iowa, 1995.

[20] M. Shida and S. Shindoh, Monotone semidefinite complementarity problems. Technical Report B-312, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, 1996.

[21] J.F. Sturm and S. Zhang. Symmetric primal-dual path following algorithms for semidefinite programming. Technical Report 9554/A, Econometric Institute, Erasmus University Rotterdam, Netherlands, 1995.

[22] M.J. Todd, K.C. Toh, and R.H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. Technical report, School of Operations Research and Industrial Engineering, Cornell University, USA, 1996.

[23] P Tseng. Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP. Technical report, Department of Mathematics, University of Washington, Seattle, USA, 1996.

[24] Y. Zhang. On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM J. on Optimization*. To appear.

[25] Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite programming relaxations for the quadratic assignment problem. Technical Report 95-27, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada, 1996.

Appendix.

Based on the papers [16, 14, 23, 15], we show that if we take a feasible initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F} \cap (\mathcal{S}_{++} \times \mathcal{S}_{++})$ and if we take γ and $\kappa = c_2\gamma/(1 - c_1\gamma)$ as in Table 1, then every direction in the KSH, the MZ, the Tseng, and the MT families is an admissible direction. Although the MZ and the MT families were described for the SDP, they can be naturally extended to the monotone SDLCP (1). When the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in (\mathcal{S}_{++} \times \mathcal{S}_{++})$ is feasible, every direction in the families share (16)' with the family of admissible directions. So in order to prove that a direction in the families is an admissible direction, we need only to show that it satisfies (19)'. In the remainder of this appendix, we assume that $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{F} \cap (\mathcal{S}_{++} \times \mathcal{S}_{++})$ and that $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} \cap \mathcal{N}(\gamma, \theta, \boldsymbol{X}^0, \boldsymbol{Y}^0)$ for some $\theta \in (0, 1]$.

A1. The KSH Family of Directions.

The KSH family of directions [9] is defined as the set of $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying (16)' and

$$\boldsymbol{X}(d\boldsymbol{Y} + \widetilde{d\boldsymbol{Y}}) + (d\boldsymbol{X} + \widetilde{d\boldsymbol{X}})\boldsymbol{Y} = \beta\theta\mu^0\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y} \tag{35}$$

for some pair $(\widetilde{d\boldsymbol{X}}, \widetilde{d\boldsymbol{Y}})$ of $n \times n$ skew-symmetric matrices such that $\widetilde{d\boldsymbol{X}} \bullet \widetilde{d\boldsymbol{Y}} \geq 0$. Let $0 < \gamma < 1/\left(2(1+\sqrt{2})\right)$ and $\kappa = (1 + 2\sqrt{2})\gamma/(1 - \gamma)$. Then every direction in the KSH family satisfies the inequality (19)'. To prove this fact, define

$$\delta_x \equiv \|\boldsymbol{X}^{-1/2} d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F \quad \text{and} \quad \widetilde{\delta}_x \equiv \|\boldsymbol{X}^{-1/2} \widetilde{d\boldsymbol{X}}\boldsymbol{X}^{-1/2}\|_F$$

as in the paper [16]. Then we know by Lemma 3.6 of the paper [16] that

$$\max\{\delta_x, \tilde{\delta}_x\} \le \frac{\gamma + (1-\beta)\sqrt{n}}{1-\gamma};$$

hence

$$\begin{aligned}
&\|\boldsymbol{X}^{-1/2}\left(d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{X}d\boldsymbol{Y}\right)\boldsymbol{X}^{1/2} - \beta\theta\mu^0\boldsymbol{I} + \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\|_F \\
&\le \quad (1+2\sqrt{2})\max\{\delta_x, \tilde{\delta}_x\}\|\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} - \theta\mu^0\boldsymbol{I}\|_F \\
&\qquad \text{(see [16, the proof of Lemma 3.4])} \\
&\le \quad \frac{(1+2\sqrt{2})\gamma}{1-\gamma}(\gamma + (1-\beta)\sqrt{n})\theta\mu^0 \\
&\le \quad \kappa(\gamma + (1-\beta)\sqrt{n})\theta\mu^0.
\end{aligned}$$

Thus the inequality (19)' holds.

A2. The MZ Family of Directions.

When the initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ is feasible, the MZ family of directions [13, 14, 17, 24] is defined as the set of $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying (16)' and (17) for some $n \times n$ nonsingular matrix $\boldsymbol{P}$. Let $0 < \gamma < 1/(1+2\sqrt{2})$ and $\kappa = 2\sqrt{2}\gamma/(1-\gamma)$. Then every direction $(d\boldsymbol{X}, d\boldsymbol{Y})$ in the MZ family satisfies the inequality (19)'. Using the argument in [14, Section 3], we prove this fact. Define

$$\begin{aligned}
&\hat{\boldsymbol{X}} \equiv \boldsymbol{P}\boldsymbol{X}\boldsymbol{P}^T,\ \widehat{d\boldsymbol{X}} \equiv \boldsymbol{P}d\boldsymbol{X}\boldsymbol{P}^T,\ \hat{\boldsymbol{Y}} \equiv \boldsymbol{P}^{-T}\boldsymbol{Y}\boldsymbol{P}^{-1},\ \widehat{d\boldsymbol{Y}} \equiv \boldsymbol{P}^{-T}d\boldsymbol{Y}\boldsymbol{P}^{-1}, \\
&\widetilde{\boldsymbol{W}}_x \equiv \hat{\boldsymbol{X}}^{-1/2}\left(\widehat{d\boldsymbol{X}}\hat{\boldsymbol{Y}} + \hat{\boldsymbol{X}}\widehat{d\boldsymbol{Y}} + \hat{\boldsymbol{X}}\hat{\boldsymbol{Y}} - \beta\theta\mu^0\boldsymbol{I}\right)\hat{\boldsymbol{X}}^{1/2}, \\
&\hat{\delta}_x \equiv \|\hat{\boldsymbol{X}}^{-1/2}\widehat{d\boldsymbol{X}}\hat{\boldsymbol{Y}}^{1/2}\|_F,\ \Phi_\beta(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}}) \equiv \left\|\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}^{1/2} - \beta\theta\mu^0\hat{\boldsymbol{X}}^{-1/2}\hat{\boldsymbol{Y}}^{-1/2}\right\|_F.
\end{aligned}$$

Since $0 < \gamma \le 1/(1+2\sqrt{2})$, we know that

$$\begin{aligned}
\|\widetilde{\boldsymbol{W}}_x\|_F &\le \sqrt{2}\hat{\delta}_x\|\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}^{1/2} - \theta\mu^0\hat{\boldsymbol{X}}^{-1/2}\hat{\boldsymbol{Y}}^{-1/2}\|_F \quad \text{(by [14, Lemma 3.6])}, \\
\Phi_\nu(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}})^2 &\le \frac{\gamma^2 + (1-\nu)^2 n}{1-\gamma}\theta\mu^0 \quad \text{for every } \nu \in [0,1] \quad \text{(by [14, Lemma 3.7])}, \\
\hat{\delta}_x &\le 2\Phi_\beta(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}}) \quad \text{(by [14, Lemma 3.8])}.
\end{aligned}$$

Hence

$$\begin{aligned}
&\left\|\hat{\boldsymbol{X}}^{-1/2}(\hat{\boldsymbol{X}}\widehat{d\boldsymbol{Y}} + \widehat{d\boldsymbol{X}}\hat{\boldsymbol{Y}})\hat{\boldsymbol{X}}^{1/2} - \beta\theta\mu^0\boldsymbol{I} + \hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2}\right\|_F \\
&= \quad \|\widetilde{\boldsymbol{W}}_x\|_F \\
&\le \quad \sqrt{2}\hat{\delta}_x\|\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}^{1/2} - \theta\mu^0\hat{\boldsymbol{X}}^{-1/2}\hat{\boldsymbol{Y}}^{-1/2}\|_F \\
&\le \quad 2\sqrt{2}\Phi_\beta(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}})\Phi_1(\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}}) \\
&\le \quad 2\sqrt{2}\left(\frac{\gamma^2 + (1-\beta)^2 n}{1-\gamma}\theta\mu^0\right)^{1/2}\left(\frac{\gamma^2}{1-\gamma}\theta\mu^0\right)^{1/2} \\
&\le \quad \frac{2\sqrt{2}\gamma}{1-\gamma}\left(\gamma + (1-\beta)\sqrt{n}\right)\theta\mu^0 \\
&= \quad \kappa\left(\gamma + (1-\beta)\sqrt{n}\right)\theta\mu^0.
\end{aligned}$$

Thus the inequality (19)' holds.

A3. The Tseng Family of Directions.

For simplicity of notation, we use $\boldsymbol{Q} = \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}$. The Tseng family of directions [23] is defined as the set of $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying (19)' and

$$\frac{1}{2}\left(\boldsymbol{Q}^{\tau}(\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2})\boldsymbol{Q}^{1-\tau} + \boldsymbol{Q}^{1-\tau}(\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2})\boldsymbol{Q}^{\tau}\right) + \boldsymbol{X}^{1/2}d\boldsymbol{Y}\boldsymbol{X}^{1/2} = \beta\theta\mu^0\boldsymbol{I} - \boldsymbol{Q}$$

for some $\tau \in [0,1]$. Suppose that $0 < \gamma < 1/3$ and $\kappa = 2\gamma/(1-\gamma)$. Then every direction in the Tseng family satisfies (19)'. We prove this fact. Applying [23, Lemma 4.1] with $\boldsymbol{d} = \boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}$, we have that

$$\left\| \left(\theta\mu^0\boldsymbol{I}\otimes\boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{Q}^{\tau}\otimes\boldsymbol{Q}^{1-\tau} + \boldsymbol{Q}^{1-\tau}\otimes\boldsymbol{Q}^{\tau}\right)\right)\mathbf{vec}(\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2})\right\| \le \gamma\theta\mu^0\|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F. \quad (36)$$

We also see by [23, the proof of Lemma 4.2 (b)] and Lemma 3.1 that

$$\begin{aligned}(1-\gamma)\theta\mu^0\|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F^2 &\le \|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F\|\beta\theta\mu^0\boldsymbol{I} - \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\|_F \\ &\le \|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F(\gamma + (1-\beta)\sqrt{n})\theta\mu^0,\end{aligned}$$

which implies that

$$\|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F \le \frac{1}{1-\gamma}(\gamma + (1-\beta)\sqrt{n}). \quad (37)$$

Hence

$$\begin{aligned}&\|\boldsymbol{X}^{-1/2}(\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{X}\boldsymbol{Y})\boldsymbol{X}^{1/2} - \beta\theta\mu^0\boldsymbol{I} + \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\|_F \\ &= \left\| \left(\boldsymbol{Q}\otimes\boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{Q}^{\tau}\otimes\boldsymbol{Q}^{1-\tau} + \boldsymbol{Q}^{1-\tau}\otimes\boldsymbol{Q}^{\tau}\right)\right)\mathbf{vec}(\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2})\right\| \\ &\quad \left(\text{by the definition of the Tseng direction and } \boldsymbol{Q} = \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2}\right) \\ &\le \|(\theta\mu^0\boldsymbol{I} - \boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2})\otimes\boldsymbol{I}\mathbf{vec}(\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2})\| \\ &\quad + \left\|\left\{\theta\mu^0\boldsymbol{I}\otimes\boldsymbol{I} - 1/2\left(\boldsymbol{Q}^{\tau}\otimes\boldsymbol{Q}^{1-\tau} + \boldsymbol{Q}^{1-\tau}\otimes\boldsymbol{Q}^{\tau}\right)\right\}\mathbf{vec}(\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2})\right\| \\ &\le 2\gamma\theta\mu^0\|\boldsymbol{X}^{-1/2}d\boldsymbol{X}\boldsymbol{X}^{-1/2}\|_F \quad (\text{by (36)}) \\ &\le \frac{2\gamma}{1-\gamma}(\gamma + (1-\beta)\sqrt{n})\theta\mu^0 \quad (\text{by (37)}) \\ &= \kappa(\gamma + (1-\beta)\sqrt{n})\theta\mu^0.\end{aligned}$$

Thus the inequality (19)' holds.

A4. The MT Family of Directions.

The MT family of directions [15] is defined as the set of $(d\boldsymbol{X}, d\boldsymbol{Y})$ satisfying (16)' and

$$\left.\begin{array}{l}\hat{\boldsymbol{X}} = \boldsymbol{P}\boldsymbol{X}\boldsymbol{P}^T, \widehat{d\boldsymbol{X}} = \boldsymbol{P}d\boldsymbol{X}\boldsymbol{P}^T,\ \hat{\boldsymbol{Y}} = \boldsymbol{P}^{-T}\boldsymbol{Y}\boldsymbol{P}^{-1}, \widehat{d\boldsymbol{Y}} = \boldsymbol{P}^{-T}d\boldsymbol{Y}\boldsymbol{P}^{-1}, \\ \hat{\boldsymbol{X}}^{1/2}\boldsymbol{U} + \boldsymbol{U}\hat{\boldsymbol{X}}^{1/2} = \widehat{d\boldsymbol{X}}, \\ \boldsymbol{U}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2} + \hat{\boldsymbol{X}}^{1/2}\widehat{d\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2} + \hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\boldsymbol{U} = \beta\theta\mu^0\boldsymbol{I} - \hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2}\end{array}\right\} \quad (38)$$

for some $n \times n$ nonsingular matrix $\boldsymbol{P}$. Suppose that $0 < \gamma < 1/(2\sqrt{2})$ and that $\kappa = \sqrt{2}\gamma/(1-\sqrt{2}\gamma)$. Then every direction in the MT family satisfies the inequality (19)'. We prove this fact.

Let $\widehat{\boldsymbol{C}} \equiv \hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2} - \theta\mu^0\boldsymbol{I}$. Then

$$\left\|\widehat{\boldsymbol{C}}\right\|_F = \left\|\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2} - \theta\mu^0\boldsymbol{I}\right\|_F = \left\|\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} - \theta\mu^0\boldsymbol{I}\right\|_F \le \gamma\theta\mu^0. \quad (39)$$

By [15, Lemma 2.3 and the inequality (18)], we also see that

$$\sqrt{2}\left\|\boldsymbol{U}\hat{\boldsymbol{X}}^{-1/2}\right\|_F \le \left\|\hat{\boldsymbol{X}}^{-1/2}\widehat{d\boldsymbol{X}}\hat{\boldsymbol{X}}^{-1/2}\right\|_F \le \frac{\gamma+(1-\beta)\sqrt{n}}{1-\sqrt{2}\gamma}. \tag{40}$$

Hence

$$\begin{aligned}
&\left\|\boldsymbol{X}^{-1/2}\left(\boldsymbol{X}d\boldsymbol{Y}+d\boldsymbol{X}\boldsymbol{Y}\right)\boldsymbol{X}^{1/2}-(\beta\theta\mu^0\boldsymbol{I}-\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2})\right\|_F \\
&= \left\|\hat{\boldsymbol{X}}^{-1/2}(\hat{\boldsymbol{X}}\widehat{d\boldsymbol{Y}}+\widehat{d\boldsymbol{X}}\hat{\boldsymbol{Y}})\hat{\boldsymbol{X}}^{1/2}-(\beta\theta\mu^0\boldsymbol{I}-\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2})\right\|_F \\
&= \left\|\hat{\boldsymbol{X}}^{-1/2}\widehat{d\boldsymbol{X}}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2}-(\boldsymbol{U}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2}+\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\boldsymbol{U})\right\|_F \quad (\text{ by (38)}) \\
&= \left\|(\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\otimes\hat{\boldsymbol{X}}^{-1/2})\mathbf{vec}\widehat{d\boldsymbol{X}}-(\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}})\mathbf{vec}\boldsymbol{U}\right\| \\
&= \left\|\left((\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\otimes\hat{\boldsymbol{X}}^{-1/2})(\hat{\boldsymbol{X}}^{1/2}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\hat{\boldsymbol{X}}^{1/2})-(\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}})\right)\mathbf{vec}\boldsymbol{U}\right\| \\
&\quad (\text{by (38)}) \\
&= \left\|\left(\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2}\otimes\hat{\boldsymbol{X}}^{-1/2}-\boldsymbol{I}\otimes\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\right)\mathbf{vec}\boldsymbol{U}\right\| \\
&= \left\|\left((\widehat{\boldsymbol{C}}+\theta\mu^0\boldsymbol{I})\otimes\hat{\boldsymbol{X}}^{-1/2}-\boldsymbol{I}\otimes(\widehat{\boldsymbol{C}}+\theta\mu^0\boldsymbol{I})\hat{\boldsymbol{X}}^{-1/2}\right)\mathbf{vec}\boldsymbol{U}\right\| \\
&\quad \left(\text{since } \widehat{\boldsymbol{C}}=\hat{\boldsymbol{X}}^{1/2}\hat{\boldsymbol{Y}}\hat{\boldsymbol{X}}^{1/2}-\theta\mu^0\boldsymbol{I}\right) \\
&= \left\|\left(\widehat{\boldsymbol{C}}\otimes\hat{\boldsymbol{X}}^{-1/2}-\boldsymbol{I}\otimes\widehat{\boldsymbol{C}}\hat{\boldsymbol{X}}^{-1/2}\right)\mathbf{vec}\boldsymbol{U}\right\| \\
&= \left\|\hat{\boldsymbol{X}}^{-1/2}\boldsymbol{U}\widehat{\boldsymbol{C}}-\widehat{\boldsymbol{C}}\hat{\boldsymbol{X}}^{-1/2}\boldsymbol{U}\right\|_F \\
&\le 2\left\|\widehat{\boldsymbol{C}}\right\|_F\left\|\boldsymbol{U}\hat{\boldsymbol{X}}^{-1/2}\right\|_F \\
&\le \frac{\sqrt{2}\gamma}{1-\sqrt{2}\gamma}(\gamma+(1-\beta)\sqrt{n})\theta\mu^0 \quad (\text{by (39) and (40)}) \\
&= \kappa(\gamma+(1-\beta)\sqrt{n})\theta\mu^0.
\end{aligned}$$

Thus the inequality (19)' holds.