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

A Predictor-Corrector Interior-Point Algorithm for the Semidefinite Linear Complementarity Problem Using the Alizadeh-Haeberly-Overton Search Direction

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

January 1996, B-311
Revised August 1996

Abstract. This paper proposes a globally convergent predictor-corrector infeasible-interior-point algorithm for the monotone semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton search direction, and shows its quadratic local convergence under the strict complementarity condition.

Key words Semidefinite Linear Complementarity Problem, Semidefinite Programming, Interior-Point Algorithm, Predictor-Corrector Algorithm, Local Convergence, Quadratic Convergence

† Department 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

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

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

1. Introduction.

Several distinct search directions have been employed in many interior-point algorithms ([1, 2, 4, 7, 8, 12, 13, 14, 18, 19, 20, 21, 23, 26, 27, 29, etc.]) developed so far for the SDP (semidefinite program). They are roughly classified into two groups. The search directions in one group ([1, 4, 19, 20, 21, 26, 27, etc.]) are founded on the self-concordant barrier or potential function [19] for the SDP, while each search direction in the other group ([2, 7, 8, 12, 13, 14, 18, 23, 29, etc.]) is derived from a certain linearization of the optimality condition, which consists of the primal feasibility, the dual feasibility and the complementarity equations, for the SDP. This paper is concerned with the latter group of search directions.

Among the search directions in the latter group, the one independently proposed by Helmberg-Rendl-Vanderbei-Wolkowicz [7] and Kojima-Shindoh-Hara [12], which we will call the HRVW/KSH/M search direction, has been studied extensively in recent papers [7, 8, 13, 14, 18, 23, 29, etc.]. In particular, Monteiro [18] devised a new formulation of the KSH search direction. Many polynomial-time primal-dual interior-point algorithms for the LP (linear program), such as central trajectory following algorithms [9, 10, 17, 28, etc.], potential reduction algorithms [11, 15, etc.], predictor-corrector algorithms [16, etc.], were extended to the SDP, and similar global (polynomial-time) computational complexities for the extended algorithms were established in those papers. Also there are a few articles [13, 23] that investigate the local convergence of interior-point algorithms for the SDP. Potra and Sheng [23] provided a sufficient condition for the superlinear convergence of an extension, with the use of Monteiro's new formulation of the KSH search direction, of the Mizuno-Todd-Ye type predictor-corrector algorithm for the LP to the SDP. In their recent paper [13], Kojima, Shida and Shindoh presented an example of SDPs which exhibits a substantial difficulty in the local convergence analysis of the Potra-Sheng extension of the Mizuno-Todd-Ye type predictor-corrector algorithm. They deduced from the example that the Potra-Sheng extension needs an additional condition to attain the superlinear convergence. Their condition requires that the generated sequence converges a solution of the SDP tangentially to the central surface. This example gave the authors a motivation to explore the local convergence of interior-point algorithms using different search directions. It will be shown in the current paper that one of the search directions proposed by Alizadeh, Haeberly and Overton [2], which we will call the AHO search direction (see (3) below), fits quite well the quadratic convergence under the strict complementarity condition.

Besides what we have called the AHO search direction above, Alizadeh, Haeberly and Overton [2] derived some other primal-dual search directions from linearization of the optimality condition for the SDP, and reported some numerical result that a primal-dual Mehrotra type predictor-corrector interior-point algorithm using the AHO search direction worked more efficiently than the algorithms using the other primal-dual search directions. But any theoretical convergence analysis has not been done on the algorithm using the AHO search direction.

Let $\mathcal{S}$ denote the set of all $n \times n$ symmetric real matrices. We regard $\mathcal{S}$ an $n(n+1)/2$-dimensional linear space with the inner product $\boldsymbol{X} \bullet \boldsymbol{Y} = \mathrm{Tr}\ \boldsymbol{X}^T\boldsymbol{Y}$ of $\boldsymbol{X}$ and $\boldsymbol{Y}$ in $\mathcal{S}$ and the Frobenius norm $\|\boldsymbol{X}\|_F = (\boldsymbol{X} \bullet \boldsymbol{X})^{1/2}$ of $\boldsymbol{X} \in \mathcal{S}$, where Tr $\boldsymbol{A}$ denotes the trace of an $n \times n$ matrix $\boldsymbol{A}$. We write $\boldsymbol{X} \succ \boldsymbol{O}$ if $\boldsymbol{X} \in \mathcal{S}$ is positive definite, and $\boldsymbol{X} \succeq \boldsymbol{O}$ if $\boldsymbol{X} \in \mathcal{S}$ is positive semidefinite. Here $\boldsymbol{O}$ denotes the $n \times n$ zero matrix. We also use the symbols $\mathcal{S}_{++}$ and $\mathcal{S}_{+}$ for the set of positive definite

symmetric matrices and the set of positive semidefinite symmetric matrices, respectively;

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

Let $\mathcal{F}$ be an $n(n+1)/2$-dimensional affine subspace of $\mathcal{S} \times \mathcal{S}$, and

$$\mathcal{F}_{+} = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} \ : \ \boldsymbol{X} \succeq \boldsymbol{O}, \ \boldsymbol{Y} \succeq \boldsymbol{O}\}.$$

We are concerned with the SDLCP (semidefinite linear complementarity problem):

$$\text{Find an } (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{+} \ \text{ such that } \boldsymbol{X} \bullet \boldsymbol{Y} = 0. \tag{1}$$

We call an $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}_{+}$ a feasible solution of the SDLCP (1). Throughout the paper we assume the monotonicity of the $n(n+1)/2$-dimensional affine subspace $\mathcal{F}$;

$$(\boldsymbol{U}' - \boldsymbol{U}) \bullet (\boldsymbol{V}' - \boldsymbol{V}) \geq 0 \ \text{ for every } (\boldsymbol{U}', \boldsymbol{V}'), \ (\boldsymbol{U}, \boldsymbol{V}) \in \mathcal{F}. \tag{2}$$

The monotone SDLCP was introduced in the paper [12] by Kojima, Shindoh and Hara as an extension of the monotone LCP (linear complementarity problem) and a mathematical framework on which they founded interior-point algorithms. Besides the interior-point algorithms given in their paper [12], many of the primal-dual interior-point algorithms developed so far for the SDP can be extended to the monotone SDLCP.

If we adapt the AHO search direction [2] to the monotone SDLCP, we can describe it as a solution of the system of equations

$$\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{Y}\boldsymbol{X} + d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}d\boldsymbol{X} = 2\beta\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Y} - \boldsymbol{Y}\boldsymbol{X} \ \text{ and } (\boldsymbol{X} + d\boldsymbol{X}, \boldsymbol{Y} + d\boldsymbol{Y}) \in \mathcal{F}. \tag{3}$$

Here $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ denotes an iterate, $\beta \in [0, 1]$ a search direction parameter, and $\mu = \boldsymbol{X} \bullet \boldsymbol{Y}/n$. It was shown in the recent paper [25] that the system (3) of equations above has the unique solution $(d\boldsymbol{X}, d\boldsymbol{Y})$ whenever $\boldsymbol{X} \succ \boldsymbol{O}$, $\boldsymbol{Y} \succ \boldsymbol{O}$ and $\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X} \succeq \boldsymbol{O}$. See also Corollary 3.2 in Section 3.

The current paper has two purposes. The one is to propose a globally convergent Mizuno-Todd-Ye type predictor-corrector infeasible-interior-point algorithm, with the use of the AHO search direction, for the monotone SDLCP. The other purpose is to demonstrate its quadratic convergence under the strict complementarity condition. After writing out our first version of the paper, Potra and Sheng [24] showed the superlinear convergence of the primal-dual infeasible-interior-point path-following algorithm proposed by Kojima, Shida and Shindoh [13] without the assumption of nondegeneracy. (see Condition 6.1 of Section 6.) Our proof of quadratic convergence without nondegeneracy condition is heavily relied on the paper [24]. Although we will describe the algorithm for the monotone SDLCP, we can easily apply it to the primal-dual pair of SDPs. See the paper [12] for detailed relations between the primal-dual pair of SDPs and the monotone SDLCP.

In Section 2, we present a globally convergent Mizuno-Todd-Ye type predictor-corrector infeasible-interior-point algorithm using the AHO search direction for the SDLCP (1). Section 3 is devoted to fundamental lemmas which we will use in Sections 4 and 5. Based on those lemmas, we prove the global convergence of the algorithm in Section 4, and derive its quadratic convergence under the strict complementarity condition in Section 5. In Section 6, we will show further local convergence properties under an additional assumption of nondegeneracy.

2. A Predictor-Corrector Interior-Point Algorithm.

Throughout the paper we use the following notation:

$$\begin{aligned}
\zeta &: \text{ a constant not less than } 1/n,\\
\mathcal{F}_0 &= \{(\boldsymbol{U}',\boldsymbol{V}') - (\boldsymbol{U},\boldsymbol{V}) \ : \ (\boldsymbol{U}',\boldsymbol{V}'),\ (\boldsymbol{U},\boldsymbol{V}) \in \mathcal{F}\},\\
\widetilde{\mathcal{N}}(\gamma,\tau) &= \left\{(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \begin{array}{l}(\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X})/2 \succeq (1-\gamma)\tau\boldsymbol{I},\\ \boldsymbol{X}\bullet\boldsymbol{Y}/n \le (1+\zeta\gamma)\tau\end{array}\right\}\\
&\quad \text{for each } \gamma \in [0,\ 1] \text{ and each } \tau \ge 0.
\end{aligned}$$

By construction, we see that

$$\begin{aligned}
&(1-\gamma)\tau \le \boldsymbol{X}\bullet\boldsymbol{Y}/n\\
&\quad \text{if } (\boldsymbol{X},\boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma,\tau),\ \gamma \in [0,1] \text{ and } \tau \ge 0,\\
&\widetilde{\mathcal{N}}(0,\tau) = \{(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \boldsymbol{X}\boldsymbol{Y} = \tau\boldsymbol{I}\} \subset \widetilde{\mathcal{N}}(\gamma,\tau) \subset \widetilde{\mathcal{N}}(\gamma',\tau)\\
&\quad \text{if } 0 < \gamma < \gamma' \le 1 \ \text{ and } \tau > 0.
\end{aligned}$$

Let $0 < \gamma < 1$. Then the set $\{(\boldsymbol{X},\boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma,\tau) \ : \ \tau > 0\}$ forms a neighborhood of "the central surface" $\{(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \boldsymbol{X}\boldsymbol{Y} = \tau\boldsymbol{I} \text{ for some } \tau > 0\}$. We call γ *a neighborhood parameter*. This set serves as the admissible region in which we confine iterates $(\boldsymbol{X}^k,\boldsymbol{Y}^k)$ $(k = 0,1,2,\dots)$ of Algorithm 2.2 described below. More precisely, starting from an $(\boldsymbol{X}^0,\boldsymbol{Y}^0,\theta^0,\gamma^0) = (\sqrt{\mu^0}\boldsymbol{I},\sqrt{\mu^0}\boldsymbol{I},1,0)$ for some $\mu^0 > 0$, Algorithm 2.2 generates a sequence $\{(\boldsymbol{X}^k,\boldsymbol{Y}^k,\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k,\theta^k,\gamma^k)\}$ such that for every $k = 1,2,\dots,$

$$1 \ge \theta^k \ge 0,\ \gamma > \gamma^k \ge 0, \tag{4}$$

$$1 = \theta^0 > \theta^k > \theta^{k+1}, \tag{5}$$

$$(\boldsymbol{X}^k,\boldsymbol{Y}^k) \in \widetilde{\mathcal{N}}(\gamma^k,\theta^k\mu^0),\ (\boldsymbol{X}^k,\boldsymbol{Y}^k) \in \mathcal{F} + \theta^k\left((\boldsymbol{X}^0,\boldsymbol{Y}^0) - (\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right), \tag{6}$$

$$(\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k) \in \widetilde{\mathcal{N}}(\gamma,\theta^{k+1}\mu^0),\ (\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k) \in \mathcal{F} + \theta^{k+1}\left((\boldsymbol{X}^0,\boldsymbol{Y}^0) - (\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right). \tag{7}$$

Here $(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})$ denotes an arbitrary pair of matrices in $\mathcal{F}$; in particular, we can take any solution of the SDLCP (1) for $(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})$ when the SDLCP (1) has a solution. Note that

$$\begin{aligned}
&\mathcal{F} + \theta\left((\boldsymbol{X}^0,\boldsymbol{Y}^0) - (\boldsymbol{X}',\boldsymbol{Y}')\right)\\
&= \mathcal{F} + \theta\left((\boldsymbol{X}^0,\boldsymbol{Y}^0) - (\boldsymbol{X},\boldsymbol{Y})\right)\\
&\quad \text{for any } (\boldsymbol{X}',\boldsymbol{Y}') \in \mathcal{F},\ (\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{F} \ \text{ and } \theta \in [0,1].
\end{aligned}$$

Among the iterates $\boldsymbol{X}^k,\ \boldsymbol{Y}^k,\ \boldsymbol{X}_c^k,\ \boldsymbol{Y}_c^k,\ \theta^k,\ \gamma^k$, the triplet $(\boldsymbol{X}^k,\boldsymbol{Y}^k,\theta^k)$ is updated to $(\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k,\theta^{k+1})$ by the Predictor Step (Step 2), while the triplet $(\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k,\gamma^k)$ to $(\boldsymbol{X}^{k+1},\boldsymbol{Y}^{k+1},\gamma^{k+1})$ by the Corrector Step (Step 4). θ^{k+1} serves as a measure of both feasibility and optimality. Given an $\epsilon \ge 0$, the algorithm stops (at Step 3) when θ^{k+1} gets equal to or smaller than ϵ. In this case, we have an approximate solution $(\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k)$ of the SDLCP (1) such that

$$\left.\begin{array}{l}\epsilon \ge \theta^{k+1} \ge 0,\\ \boldsymbol{X}_c^k \succeq \boldsymbol{O},\ \boldsymbol{Y}_c^k \succeq \boldsymbol{O},\ \boldsymbol{X}_c^k\bullet\boldsymbol{Y}_c^k/n \le (1+\zeta\gamma)\theta^{k+1}\mu^0,\\ (\boldsymbol{X}_c^k,\boldsymbol{Y}_c^k) \in \mathcal{F} + \theta^{k+1}\left((\boldsymbol{X}^0,\boldsymbol{Y}^0) - (\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right).\end{array}\right\} \tag{8}$$

We call ϵ *an accuracy parameter.*

Before we run Algorithm 2.2, we build up the hypothesis below. When the algorithm detects (at Step 1 or Step 3) that the hypothesis is false, it stops.

Hypothesis 2.1. [12] Let $\omega^* \geq 1$. There exists a solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP (1) such that

$$\omega^* \boldsymbol{X}^0 \succeq \boldsymbol{X}^* \quad \text{and } \omega^* \boldsymbol{Y}^0 \succeq \boldsymbol{Y}^*. \tag{9}$$

Algorithm 2.2.

Step 0: Choose an accuracy parameter $\epsilon \geq 0$, a neighborhood parameter $\gamma \in (0,\ 1)$ and an initial point $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$ with some $\mu^0 > 0$. Let $\theta^0 = 1$, $\sigma = 2\omega^*/(1-\gamma) + 1$, $\gamma^0 = 0$ and $k = 0$.

Step 1: If the inequality

$$\theta^k(\boldsymbol{X}^0 \bullet \boldsymbol{Y}^k + \boldsymbol{X}^k \bullet \boldsymbol{Y}^0) \leq \sigma \boldsymbol{X}^k \bullet \boldsymbol{Y}^k \tag{10}$$

does not hold then stop.

Step 2: (Predictor Step) Compute a solution $(d\boldsymbol{X}_p^k, d\boldsymbol{Y}_p^k)$ of the system of equations

$$\left.\begin{array}{l} \boldsymbol{X}^k d\boldsymbol{Y}_p^k + d\boldsymbol{Y}_p^k \boldsymbol{X}^k + d\boldsymbol{X}_p^k \boldsymbol{Y}^k + \boldsymbol{Y}^k d\boldsymbol{X}_p^k = -\boldsymbol{X}^k\boldsymbol{Y}^k - \boldsymbol{Y}^k\boldsymbol{X}^k, \\ (\boldsymbol{X}^k + d\boldsymbol{X}_p^k, \boldsymbol{Y}^k + d\boldsymbol{Y}_p^k) \in \mathcal{F}. \end{array}\right\} \tag{11}$$

Let

$$\left.\begin{array}{rcl} \delta_p^k & = & \dfrac{\|d\boldsymbol{X}_p^k\|_F \|d\boldsymbol{Y}_p^k\|_F}{\theta^k \mu^0}, \\ \hat{\alpha}_p^k & = & \dfrac{2}{\sqrt{1 + 4\delta_p^k/(\gamma - \gamma^k)} + 1}, \\ \check{\alpha}_p^k & = & \max\left\{\alpha' \in [0,1] \ : \ \begin{array}{l} (\boldsymbol{X}^k + \alpha d\boldsymbol{X}_p^k, \boldsymbol{Y}^k + \alpha d\boldsymbol{Y}_p^k) \in \widetilde{\mathcal{N}}(\gamma, (1-\alpha)\theta^k\mu^0) \\ \text{for every } \alpha \in [0, \alpha'] \end{array}\right\}. \end{array}\right\} \tag{12}$$

Choose a step length $\alpha_p^k \in [\hat{\alpha}_p^k, \check{\alpha}_p^k]$. Let

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

Step 3: If $\theta^{k+1} \leq \epsilon$ then stop. If the inequality

$$\theta^{k+1}(\boldsymbol{X}^0 \bullet \boldsymbol{Y}_c^k + \boldsymbol{X}_c^k \bullet \boldsymbol{Y}^0) \leq \sigma \boldsymbol{X}_c^k \bullet \boldsymbol{Y}_c^k \tag{13}$$

does not hold then stop.

Step 4: (Corrector Step) Compute a solution $(d\boldsymbol{X}_c^k, d\boldsymbol{Y}_c^k)$ of the system of equations

$$\left.\begin{array}{l} \boldsymbol{X}_c^k d\boldsymbol{Y}_c^k + d\boldsymbol{Y}_c^k \boldsymbol{X}_c^k + d\boldsymbol{X}_c^k \boldsymbol{Y}_c^k + \boldsymbol{Y}_c^k d\boldsymbol{X}_c^k = 2\theta^{k+1}\mu^0 \boldsymbol{I} - \boldsymbol{X}_c^k\boldsymbol{Y}_c^k - \boldsymbol{Y}_c^k\boldsymbol{X}_c^k, \\ (d\boldsymbol{X}_c^k, d\boldsymbol{Y}_c^k) \in \mathcal{F}_0. \end{array}\right\} \tag{14}$$

Let

$$\left.\begin{array}{rcl}
\delta_c^k & = & \dfrac{\|d\boldsymbol{X}_c^k\|_F \|d\boldsymbol{Y}_c^k\|_F}{\theta^{k+1}\mu^0}, \\
\hat{\alpha}_c^k & = & \left\{ \begin{array}{ll} \gamma/(2\delta_c^k) & \text{if } \gamma \le 2\delta_c^k, \\ 1 & \text{if } \gamma > 2\delta_c^k, \end{array} \right. \\
\check{\gamma}^{k+1} & = & \left\{ \begin{array}{ll} \gamma(1-\gamma/(4\delta_c^k)) & \text{if } \gamma \le 2\delta_c^k, \\ \delta_c^k & \text{if } \gamma > 2\delta_c^k, \end{array} \right. \\
\hat{\gamma}^{k+1} & = & \min \left\{ \gamma' \in [0,1] \ : \ \begin{array}{l} (\boldsymbol{X} + \alpha d\boldsymbol{X}^k, \boldsymbol{Y} + \alpha d\boldsymbol{Y}^k) \in \widetilde{\mathcal{N}}(\gamma', \theta^{k+1}\mu^0), \\ \alpha \in [0,1] \end{array} \right\}.
\end{array}\right\} \tag{15}$$

Choose a step length $\alpha_c^k \in [0,1]$ and γ^{k+1} such that

$$\left.\begin{array}{l}
\hat{\gamma}^{k+1} \le \gamma^{k+1} \le \check{\gamma}^{k+1}, \\
(\boldsymbol{X}_c^k + \alpha_c^k d\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k + \alpha_c^k d\boldsymbol{Y}_c^k) \in \widetilde{\mathcal{N}}(\gamma^{k+1}, \theta^{k+1}\mu^0).
\end{array}\right\} \tag{16}$$

(It will be shown in Lemma 3.8 that the pair of $\alpha_c^k = \hat{\alpha}_c^k$ and $\gamma^{k+1} = \check{\gamma}^{k+1}$ satisfies the relations above). Let $(\boldsymbol{X}^{k+1}, \boldsymbol{Y}^{k+1}) = (\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k) + \alpha_c^k (d\boldsymbol{X}_c^k, d\boldsymbol{Y}_c^k)$.

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

A distinctive feature of Algorithm 2.2 lies in the sophisticated step length control rule at Step 4 (the Corrector Step). When $\delta_c^k = \|d\boldsymbol{X}_c^k\|_F \|d\boldsymbol{Y}_c^k\|_F / (\theta^{k+1}\mu^0) < \gamma/2$, we can take the unit step length $\alpha_c^k = 1$ and a $\gamma^{k+1} \le \gamma/2$; hence $(\boldsymbol{X}^{k+1}, \boldsymbol{Y}^{k+1}) \in \widetilde{\mathcal{N}}(\gamma/2, \theta^{k+1}\mu^0)$. When $\delta_c^k \ge \gamma/2$, however, we may not be able to take the unit step length, and as δ_c^k gets larger, we are forced to take a smaller step length α_c^k. On the other hand, the step length control rule at Step 2 (the Predictor Step) is based on and similar to the one used in the paper [23]. The theorem below summarizes the consistency and the global convergence of Algorithm 2.2. A proof of the theorem is given in Section 4.

Theorem 2.3. *(Global Convergence Theorem)*

(i) *Algorithm 2.2 consistently generates a sequence* $\{(\boldsymbol{X}^k, \boldsymbol{Y}^k, \boldsymbol{X}_c^k, \boldsymbol{Y}_c^k, \theta^k, \gamma^k)\}$ *satisfying (4), (5), (6) and (7).*

(ii) *If Algorithm 2.2 stops at Step 1 violating the inequality (10), then there is no solution of the SDLCP (1) satisfying (9).*

(iii) *If Algorithm 2.2 stops at Step 3 with* $\theta^{k+1} \le \epsilon$, *then* $(\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k)$ *gives an approximate solution of the SDLCP (1) satisfying (8)*

(iv) *If Algorithm 2.2 stops at Step 3 violating the inequality (13), then there is no solution of the SDLCP (1) satisfying (9).*

(v) *If* $\epsilon > 0$, *Algorithm 2.2 stops in a finite number of iterations at either Step 1 or Step 3.*

3. Lemmas.

In this section, we present a series of lemmas which we will utilize in proving both Theorem 2.3 (the Global Convergence Theorem) in Section 4 and Theorem 5.1 (the Local Convergence Theorem) in Section 5.

Lemma 3.1. *Assume that* $(\boldsymbol{X}, \boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma, \tau)$ *for some* $\tau > 0$ *and* $\gamma \in (0, 1)$.

(i) $\|(\boldsymbol{XY} + \boldsymbol{YX})/2\|_F \leq n(1 + \zeta\gamma)\tau$.

(ii) $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$.

(iii) *Let* $(d\boldsymbol{X}, d\boldsymbol{Y})$ *be a solution of the system of equations*

$$\boldsymbol{X}d\boldsymbol{Y} + d\boldsymbol{Y}\boldsymbol{X} + d\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}d\boldsymbol{X} = \boldsymbol{C} \quad \textit{and } (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}_0, \tag{17}$$

where $\boldsymbol{C} \in \mathcal{S}$ *is a constant matrix. Then*

$$\|d\boldsymbol{X}\|_F \leq \frac{\|\boldsymbol{X}\|_F \|\boldsymbol{C}\|_F}{((1-\gamma)\tau)} \quad \textit{and } \|d\boldsymbol{Y}\|_F \leq \frac{\|\boldsymbol{Y}\|_F \|\boldsymbol{C}\|_F}{((1-\gamma)\tau)}.$$

Proof: (i) Let ν_j $(j = 1, 2, \ldots, n)$ denote the eigenvalues of $(\boldsymbol{XY} + \boldsymbol{YX})/2$. Since the matrix $(\boldsymbol{XY} + \boldsymbol{YX})/2$ is symmetric and positive definite, we see that

$$\|(\boldsymbol{XY} + \boldsymbol{YX})/2\|_F = \left(\sum_{j=1}^{n} (\nu_j)^2\right)^{1/2} \leq \sum_{j=1}^{n} \nu_j = \boldsymbol{X} \bullet \boldsymbol{Y} \leq n(1 + \zeta\gamma)\tau.$$

Here the last inequality follows from $(\boldsymbol{X}, \boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma, \tau)$.

(ii) By definition, we know that $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+$. If $\boldsymbol{X}$ and/or $\boldsymbol{Y}$ was singular, we would have

$$\boldsymbol{u}^T(\boldsymbol{XY} + \boldsymbol{YX})\boldsymbol{u} = 0 \text{ for some nonzero } \boldsymbol{u} \in R^n,$$

which would contradict the assumption that $(\boldsymbol{X}, \boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma, \tau)$ with $\tau > 0$ and $\gamma \in (0, 1)$.

(iii) We will use the notation $\boldsymbol{A} \otimes \boldsymbol{B}$ for the Kronecker product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, and the notation

$$\textbf{vec } \boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{\cdot 1} \\ \boldsymbol{A}_{\cdot 2} \\ \cdot \\ \boldsymbol{A}_{\cdot n} \end{pmatrix} \in R^{mn},$$

where $\boldsymbol{A}_{\cdot j}$ denotes the jth column of an $m \times n$ matrix $\boldsymbol{A}$. See *e.g.* the book [5] for basic properties on the Kronecker product. Let

$$\boldsymbol{E} = (\boldsymbol{X} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{X}) \quad \text{and} \quad \boldsymbol{F} = (\boldsymbol{Y} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{Y}).$$

Then both $\boldsymbol{E}$ and $\boldsymbol{F}$ are symmetric and positive definite. Hence we can rewrite the system (17) of equations as

$$\textbf{vec } d\boldsymbol{Y} + \boldsymbol{E}^{-1}\boldsymbol{F}\boldsymbol{E}(\boldsymbol{E}^{-1}\textbf{vec } d\boldsymbol{X}) = \boldsymbol{E}^{-1}\textbf{vec } \boldsymbol{C} \text{ and } (d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{F}_0.$$

It follows that

$$
\begin{aligned}
&\|\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}\|\|\boldsymbol{C}\|_F \\
&\quad\geq\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T\mathbf{vec}\ \boldsymbol{C} \\
&\quad\geq\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T\boldsymbol{F}\boldsymbol{E}(\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}) \\
&\qquad (\text{since } (\mathbf{vec}\ d\boldsymbol{X})^T(\mathbf{vec}\ d\boldsymbol{Y}) = d\boldsymbol{X}\bullet d\boldsymbol{Y} \geq 0 \text{ by the assumption (2)}) \\
&\quad=\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T(\boldsymbol{Y}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\boldsymbol{Y})(\boldsymbol{X}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\boldsymbol{X})(\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}) \\
&\quad=\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T(\boldsymbol{Y}\boldsymbol{X}\otimes\boldsymbol{I}+\boldsymbol{X}\otimes\boldsymbol{Y}+\boldsymbol{X}\otimes\boldsymbol{Y}+\boldsymbol{I}\otimes\boldsymbol{Y}\boldsymbol{X})(\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}) \\
&\quad\geq\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T(\boldsymbol{Y}\boldsymbol{X}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\boldsymbol{Y}\boldsymbol{X})(\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}) \\
&\qquad (\text{since both } \boldsymbol{Y}\otimes\boldsymbol{X} \text{ and } \boldsymbol{X}\otimes\boldsymbol{Y} \text{ are symmetric and positive definite}) \\
&\quad=\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T\left(((\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X})/2)\otimes\boldsymbol{I}+\boldsymbol{I}\otimes((\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X})/2)\right)(\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}) \\
&\quad\geq\ (\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X})^T\left(2(1-\gamma)\tau\boldsymbol{I}\otimes\boldsymbol{I}\right)(\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}) \\
&\qquad (\text{since } (\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X})/2 \succeq (1-\gamma)\tau\boldsymbol{I}) \\
&\quad=\ 2(1-\gamma)\tau\|\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}\|^2.
\end{aligned}
$$

Therefore we obtain that

$$\|\boldsymbol{C}\|_F \geq 2(1-\gamma)\tau\|\boldsymbol{E}^{-1}\mathbf{vec}\ d\boldsymbol{X}\|.$$

On the other hand, if $\boldsymbol{W}\in\mathcal{S}$ is a solution of the system of equations $\boldsymbol{X}\boldsymbol{W}+\boldsymbol{W}\boldsymbol{X}=d\boldsymbol{X}$, then $\mathbf{vec}\ \boldsymbol{W}=\boldsymbol{E}^{-1}(\mathbf{vec}\ d\boldsymbol{X})$; hence

$$
\begin{aligned}
\|d\boldsymbol{X}\|_F &= \|\boldsymbol{X}\boldsymbol{W}+\boldsymbol{W}\boldsymbol{X}\|_F \\
&\leq 2\|\boldsymbol{X}\|_F\|\boldsymbol{W}\|_F \\
&= 2\|\boldsymbol{X}\|_F\|\boldsymbol{E}^{-1}(\mathbf{vec}\ d\boldsymbol{X})\| \\
&\leq \|\boldsymbol{X}\|_F\|\boldsymbol{C}\|_F/((1-\gamma)\tau).
\end{aligned}
$$

We can prove similarly the inequality $\|d\boldsymbol{Y}\|_F \leq \|\boldsymbol{Y}\|_F\|\boldsymbol{C}\|_F/((1-\gamma)\tau)$. ∎

The corollary below ensures the existence and the uniqueness of the solution of the systems (11) and (14) of equations. This result was shown in the paper [25], but we give a proof of the corollary to make the current paper self-contained. The assumption in the corollary is slightly stronger than the one in Theorem 3.1 of the paper [25].

Corollary 3.2. *Assume that* $(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{S}_+\times\mathcal{S}_+$ *and* $\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}\boldsymbol{X}\succ\boldsymbol{O}$. *For any* $\boldsymbol{C}\in\mathcal{S}$ *and any* $(\boldsymbol{X}',\boldsymbol{Y}')\in\mathcal{S}\times\mathcal{S}$, *there exists a unique solution* $(d\boldsymbol{X},d\boldsymbol{Y})$ *of the system of equations*

$$\boldsymbol{X}d\boldsymbol{Y}+d\boldsymbol{Y}\boldsymbol{X}+d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}d\boldsymbol{X}=\boldsymbol{C} \quad \textit{and}\ (d\boldsymbol{X}+\boldsymbol{X}',d\boldsymbol{Y}+\boldsymbol{Y}')\in\mathcal{F}. \tag{18}$$

Proof: First we take $\tau>0$ and $\gamma\in(0,1)$ such that $(\boldsymbol{X},\boldsymbol{Y})\in\widetilde{\mathcal{N}}(\gamma,\tau)$. Let $m=n(n+1)/2$. For every $\boldsymbol{C}\in\mathcal{S}$ and $(\boldsymbol{X}',\boldsymbol{Y}')\in\mathcal{S}\times\mathcal{S}$, define

$$
\begin{aligned}
\mathcal{G}(\boldsymbol{C}) &= \{(d\boldsymbol{X},d\boldsymbol{Y})\in\mathcal{S}\times\mathcal{S}\ :\ \boldsymbol{X}d\boldsymbol{Y}+d\boldsymbol{Y}\boldsymbol{X}+d\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Y}d\boldsymbol{X}=\boldsymbol{C}\}, \\
\mathcal{F}(\boldsymbol{X}',\boldsymbol{Y}') &= \mathcal{F}-(\boldsymbol{X}',\boldsymbol{Y}').
\end{aligned}
$$

Since $\boldsymbol{X} \in \mathcal{S}$ is positive definite by (ii) of Lemma 3.1, for every $d\boldsymbol{X} \in \mathcal{S}$ there exists a unique $d\boldsymbol{Y} \in \mathcal{S}$ such that $(d\boldsymbol{X}, d\boldsymbol{Y}) \in \mathcal{G}(\boldsymbol{O})$. This implies that $\mathcal{G}(\boldsymbol{O})$ forms an m-dimensional linear subspace of the $2m$-dimensional linear space $\mathcal{S} \times \mathcal{S}$. Take an $(\boldsymbol{X}", \boldsymbol{Y}") \in \mathcal{F}$. Then we see that $\mathcal{F}(\boldsymbol{X}", \boldsymbol{Y}") = \mathcal{F}_0$; hence $\mathcal{F}(\boldsymbol{X}", \boldsymbol{Y}")$ forms an m-dimensional linear subspace of the $2m$-dimensional linear space $\mathcal{S} \times \mathcal{S}$. Applying (iii) of Lemma 3.1 with $\boldsymbol{C} = \boldsymbol{O}$, we know that the m-dimensional linear subspaces $\mathcal{G}(\boldsymbol{O})$ and $\mathcal{F}(\boldsymbol{X}", \boldsymbol{Y}") = \mathcal{F}_0$ of the $2m$-dimensional linear space $\mathcal{S} \times \mathcal{S}$ transversely intersect at the single point $(\boldsymbol{O}, \boldsymbol{O})$. Therefore, for any $\boldsymbol{C} \in \mathcal{S}$ and $(\boldsymbol{X}', \boldsymbol{Y}') \in \mathcal{S} \times \mathcal{S}$, their parallel translations $\mathcal{G}(\boldsymbol{C})$ and $\mathcal{F}(\boldsymbol{X}', \boldsymbol{Y}')$ transversely intersect at a single point, and the desired result follows. ∎

Lemma 3.3. *Suppose that Hypothesis 2.1 is true. Let* $\gamma \in (0,1)$, $\mu^0 > 0$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$ *and* $\sigma = 2\omega^*/(1-\gamma) + 1$. *If* $(\boldsymbol{X}, \boldsymbol{Y}, \theta)$ *satisfies*

$$\begin{aligned} &1 \geq \theta \geq 0, \ (\boldsymbol{X}, \boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma, \theta\mu^0), \\ &(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} + \theta\left((\boldsymbol{X}^0, \boldsymbol{Y}^0) - (\overline{\boldsymbol{X}}, \overline{\boldsymbol{Y}})\right) \quad \textit{for some } (\overline{\boldsymbol{X}}, \overline{\boldsymbol{Y}}) \in \mathcal{F} \end{aligned} \tag{19}$$

then the inequality

$$\theta\left(\boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}^0\right) \leq \sigma \boldsymbol{X} \bullet \boldsymbol{Y} \tag{20}$$

holds.

Proof: This lemma follows directly from Lemma 7.6 of [12]. ∎

Lemma 3.4. *Let* $\gamma \in (0,1)$, $\mu^0 > 0$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$ *and* $\sigma > 0$. *Suppose that (19) and (20) hold. Then*

$$\|\boldsymbol{X}\|_F \leq n\sigma(1+\zeta\gamma)\sqrt{\mu^0} \ \textit{ and } \|\boldsymbol{Y}\|_F \leq n\sigma(1+\zeta\gamma)\sqrt{\mu^0} \tag{21}$$

hold.

Proof: We see from (19) and (20) that

$$\begin{aligned} \theta\sqrt{\mu^0}\,(\|\boldsymbol{Y}\|_F + \|\boldsymbol{X}\|_F) &\leq \theta\sqrt{\mu^0}\,(\mathrm{Tr}\ \boldsymbol{Y} + \mathrm{Tr}\ \boldsymbol{X}) \\ &= \theta(\boldsymbol{X}^0 \bullet \boldsymbol{Y} + \boldsymbol{X} \bullet \boldsymbol{Y}^0) \\ &\leq \sigma \boldsymbol{X} \bullet \boldsymbol{Y} \\ &\leq n\sigma(1+\zeta\gamma)\theta\mu^0. \end{aligned}$$

Thus (21) follows. ∎

Lemma 3.5. *Let* $\gamma \in (0,1)$, $\mu^0 > 0$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$ *and* $\sigma > 0$. *Define*

$$\kappa_c = 2\sigma(1+\zeta\gamma)(2+\zeta\gamma)\sqrt{\mu^0}/(1-\gamma). \tag{22}$$

If $(\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k, \theta^{k+1})$ *satisfies (4), (7) and (13), then the solution* $(d\boldsymbol{X}_c^k, d\boldsymbol{Y}_c^k)$ *of the system (14) of equations satisfies that* $\|d\boldsymbol{X}_c^k\|_F \leq n^2\kappa_c$ *and* $\|d\boldsymbol{Y}_c^k\|_F \leq n^2\kappa_c$.

Proof: By applying (iii) of Lemma 3.1 with $\tau = \theta^{r+1}\mu^0$ and $\boldsymbol{C} = 2\theta^{k+1}\mu^0\boldsymbol{I} - \boldsymbol{X}_c^k\boldsymbol{Y}_c^k - \boldsymbol{Y}_c^k\boldsymbol{X}_c^k$, we have that

$$\begin{aligned}\|d\boldsymbol{X}_c^k\|_F &\leq \|\boldsymbol{X}_c^k\|_F\|2\theta^{k+1}\mu^0\boldsymbol{I} - \boldsymbol{X}_c^k\boldsymbol{Y}_c^k - \boldsymbol{Y}_c^k\boldsymbol{X}_c^k\|_F/((1-\gamma)\theta^{k+1}\mu^0)\\ &\leq \|\boldsymbol{X}_c^k\|_F\left(2\theta^{k+1}\mu^0\sqrt{n} + \|\boldsymbol{X}_c^k\boldsymbol{Y}_c^k + \boldsymbol{Y}_c^k\boldsymbol{X}_c^k\|_F\right)/((1-\gamma)\theta^{k+1}\mu^0)\end{aligned}$$

By (i) of Lemma 3.1 and Lemma 3.4, we also know that

$$\|\boldsymbol{X}_c^k\boldsymbol{Y}_c^k + \boldsymbol{Y}_c^k\boldsymbol{X}_c^k\|_F \leq 2n(1+\zeta\gamma)\theta^{k+1}\mu^0 \ \text{ and } \|\boldsymbol{X}_c^k\|_F \leq n\sigma(1+\zeta\gamma)\sqrt{\mu^0}.$$

It follows that

$$\begin{aligned}\|d\boldsymbol{X}_c^k\|_F &\leq \left(n\sigma(1+\zeta\gamma)\sqrt{\mu^0}\right)\left(2\theta^{k+1}\mu^0\sqrt{n} + 2n(1+\zeta\gamma)\theta^{k+1}\mu^0\right)/((1-\gamma)\theta^{k+1}\mu^0)\\ &\leq n^2\kappa_c.\end{aligned}$$

We can similarly prove that $\|d\boldsymbol{Y}_c^k\|_F \leq n^2\kappa_c$. ∎

Lemma 3.6. *Let $\gamma \in (0,1)$, $\mu^0 > 0$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$, $(\overline{\boldsymbol{X}}, \overline{\boldsymbol{Y}}) \in \mathcal{F}$ and $\sigma > 0$. Define*

$$\left.\begin{aligned}\varphi &= 2(1+\zeta\gamma)\max\left\{\mu^0,\ \sigma\sqrt{\mu^0}\left(\|\boldsymbol{X}^0 - \overline{\boldsymbol{X}}\|_F + \|\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}}\|_F\right)\right\},\\ \kappa_p &= \frac{2\sigma(1+\zeta\gamma)\varphi}{(1-\gamma)\sqrt{\mu^0}} + \max\left\{\|\boldsymbol{X}^0 - \overline{\boldsymbol{X}}\|_F,\ \|\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}}\|_F\right\}.\end{aligned}\right\} \tag{23}$$

Suppose that $(\boldsymbol{X}^k, \boldsymbol{Y}^k, \theta^k)$ satisfies (5), (6) and (10). Define

$$\left.\begin{aligned}\boldsymbol{K}_1^k &= -\boldsymbol{X}^k\boldsymbol{Y}^k - \boldsymbol{Y}^k\boldsymbol{X}^k,\\ \boldsymbol{K}_2^k &= -\theta^k\left(\boldsymbol{X}^k(\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}}) + (\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}})\boldsymbol{X}^k + (\boldsymbol{X}^0 - \overline{\boldsymbol{X}})\boldsymbol{Y}^k + \boldsymbol{Y}^k(\boldsymbol{X}^0 - \overline{\boldsymbol{X}})\right).\end{aligned}\right\} \tag{24}$$

Let $(d\boldsymbol{X}_p^k, d\boldsymbol{Y}_p^k)$ be the solution of the system (11) of equations. For each $j = 1, 2$, let $(\boldsymbol{U}_j^k, \boldsymbol{V}_j^k)$ be the solution of the system

$$\boldsymbol{X}^k\boldsymbol{V}_j^k + \boldsymbol{V}_j^k\boldsymbol{X}^k + \boldsymbol{U}_j^k\boldsymbol{Y}^k + \boldsymbol{Y}^k\boldsymbol{U}_j^k = \boldsymbol{K}_j^k \ \text{ and } (\boldsymbol{U}_j^k, \boldsymbol{V}_j^k) \in \mathcal{F}_0. \tag{25}$$

(i) $\|\boldsymbol{K}_j^k\|_F \leq n\varphi\theta^k$ $(j = 1, 2)$.

(ii) $(d\boldsymbol{X}_p^k, d\boldsymbol{Y}_p^k) = \sum_{j=1}^{2}(\boldsymbol{U}_j^k, \boldsymbol{V}_j^k) + \theta^k\left((\boldsymbol{X}^0, \boldsymbol{Y}^0) - (\overline{\boldsymbol{X}}, \overline{\boldsymbol{Y}})\right)$.

(iii) $\|d\boldsymbol{X}_p^k\|_F \leq n^2\kappa_p$ *and* $\|d\boldsymbol{Y}_p^k\|_F \leq n^2\kappa_p$.

Proof: (i) By (i) of Lemma 3.1 and Lemma 3.4, we see that

$$\begin{aligned}\|\boldsymbol{K}_1^k\|_F &= \|\boldsymbol{X}^k\boldsymbol{Y}^k + \boldsymbol{Y}^k\boldsymbol{X}^k\|_F \leq 2n(1+\zeta\gamma)\theta^k\mu^0 \leq n\varphi\theta^k,\\ \|\boldsymbol{K}_2^k\|_F &= \theta^k\|\boldsymbol{X}^k(\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}}) + (\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}})\boldsymbol{X}^k + (\boldsymbol{X}^0 - \overline{\boldsymbol{X}})\boldsymbol{Y}^k + \boldsymbol{Y}^k(\boldsymbol{X}^0 - \overline{\boldsymbol{X}})\|_F\\ &\leq 2\theta^k\|\boldsymbol{X}^k(\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}}) + (\boldsymbol{X}^0 - \overline{\boldsymbol{X}})\boldsymbol{Y}^k\|_F\\ &\leq 2\theta^k(n\sigma(1+\zeta\gamma)\sqrt{\mu^0})\left(\|\boldsymbol{Y}^0 - \overline{\boldsymbol{Y}}\|_F + \|\boldsymbol{X}^0 - \overline{\boldsymbol{X}}\|_F\right)\\ &\leq n\varphi\theta^k.\end{aligned}$$

(ii) Let $(\boldsymbol{U}, \boldsymbol{V})$ denote the right hand side of the identity to be proved. It is easily verified that $(\boldsymbol{U}, \boldsymbol{V})$ is a solution of the system (11) of equations. Since $(d\boldsymbol{X}^k_p, d\boldsymbol{Y}^k_p)$ is the unique solution of the system (11) of equations, we obtain the desired identity.

(iii) By (iii) of Lemma 3.1, Lemma 3.4 and the assertion (i) above, we see that

$$\begin{aligned}\|\boldsymbol{U}^k_j\| &\leq \|\boldsymbol{X}^k\|_F \|\boldsymbol{K}^k_j\|_F / ((1-\gamma)\theta^k\mu^0) \\ &\leq \left(n\sigma(1+\zeta\gamma)\sqrt{\mu^0}\right)\left(n\varphi\theta^k\right)/((1-\gamma)\theta^k\mu^0) \\ &= n^2\sigma(1+\zeta\gamma)\varphi/\left((1-\gamma)\sqrt{\mu^0}\right) \quad \text{for } j=1,2.\end{aligned}$$

Hence we see by the assertion (ii) above that

$$\|d\boldsymbol{X}^k_p\|_F \leq \|\boldsymbol{U}^k_1\|_F + \|\boldsymbol{U}^k_2\|_F + \theta^k\|\boldsymbol{X}^0 - \overline{\boldsymbol{X}}\|_F \leq n^2\kappa_p.$$

We can similarly prove that $\|d\boldsymbol{Y}^k_p\|_F \leq n^2\kappa_p$. ∎

Lemma 3.7. *Let $\gamma \in (0,1)$, $\mu^0 > 0$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$. Suppose that $(\boldsymbol{X}^k, \boldsymbol{Y}^k, \theta^k, \gamma^k)$ satisfies (4) and (6). Let $(d\boldsymbol{X}^k_p, d\boldsymbol{Y}^k_p)$ be the solution of the system (11) of equations. Define δ^k_p, $\hat{\alpha}^k_p$ and $\check{\alpha}^k_p$ by (12). Then $0 < \hat{\alpha}^k_p \leq \check{\alpha}^k_p \leq 1$.*

Proof: By definition, we know that $0 < \hat{\alpha}^k_p \leq 1$ and $0 \leq \check{\alpha}^k_p \leq 1$. Hence it suffices to show that

$$(\boldsymbol{X}^k + \alpha d\boldsymbol{X}^k_p, \boldsymbol{Y}^k + \alpha d\boldsymbol{Y}^k_p) \in \widetilde{\mathcal{N}}(\gamma, (1-\alpha)\theta^k\mu^0) \text{ for every } \alpha \in [0, \hat{\alpha}^k_p]. \tag{26}$$

Assume that $0 \leq \alpha \leq \hat{\alpha}^k_p$. Then

$$\begin{aligned}&\frac{(\boldsymbol{X}^k + \alpha d\boldsymbol{X}^k_p)(\boldsymbol{Y}^k + \alpha d\boldsymbol{Y}^k_p) + (\boldsymbol{Y}^k + \alpha d\boldsymbol{Y}^k_p)(\boldsymbol{X}^k + \alpha d\boldsymbol{X}^k_p)}{2} - (1-\gamma)(1-\alpha)\theta^k\mu^0\boldsymbol{I} \\ &= \frac{(1-\alpha)(\boldsymbol{X}^k\boldsymbol{Y}^k + \boldsymbol{Y}^k\boldsymbol{X}^k) + \alpha^2(d\boldsymbol{X}^k_p d\boldsymbol{Y}^k_p + d\boldsymbol{Y}^k_p d\boldsymbol{X}^k_p)}{2} - (1-\gamma)(1-\alpha)\theta^k\mu^0\boldsymbol{I} \\ &\succeq (1-\alpha)(1-\gamma^k)\theta^k\mu^0\boldsymbol{I} - \alpha^2\|d\boldsymbol{X}^k_p\|_F\|d\boldsymbol{Y}^k_p\|_F\boldsymbol{I} - (1-\gamma)(1-\alpha)\theta^k\mu^0\boldsymbol{I} \\ &= \left(-\alpha^2\delta^k_p + (1-\alpha)(\gamma-\gamma^k)\right)\theta^k\mu^0\boldsymbol{I}. \end{aligned} \tag{27}$$

On the other hand, it follows from the definition of $\hat{\alpha}^k_p$ in (12) that

$$\hat{\alpha}^k_p = \frac{-(\gamma-\gamma^k) + \sqrt{(\gamma-\gamma^k)^2 + 4\delta^k_p(\gamma-\gamma^k)}}{2\delta^k_p},$$

which coincides with a positive root of the quadratic polynomial $\left(-\alpha^2\delta^k_p + (1-\alpha)(\gamma-\gamma^k)\right)$ in (27). Therefore

$$\frac{(\boldsymbol{X}^k + \alpha d\boldsymbol{X}^k_p)(\boldsymbol{Y}^k + \alpha d\boldsymbol{Y}^k_p) + (\boldsymbol{Y}^k + \alpha d\boldsymbol{Y}^k_p)(\boldsymbol{X}^k + \alpha d\boldsymbol{X}^k_p)}{2} - (1-\gamma)(1-\alpha)\theta^k\mu^0\boldsymbol{I} \succeq \boldsymbol{O}$$

for every $\alpha \in [0, \hat{\alpha}^k_p]$,

which together with $(\boldsymbol{X}^k, \boldsymbol{Y}^k) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ implies that

$$(\boldsymbol{X}^k + \alpha d\boldsymbol{X}_p^k, \boldsymbol{Y}^k + \alpha d\boldsymbol{Y}_p^k) \in \mathcal{S}_+ \times \mathcal{S}_+ \ \text{ for every } \alpha \in [0, \hat{\alpha}_p^k].$$

We also see that for every $\alpha \in [0,1]$,

$$\begin{aligned}
&(1+\zeta\gamma)(1-\alpha)\theta^k\mu^p - \frac{(\boldsymbol{X}^k + \alpha d\boldsymbol{X}_p^k) \bullet (\boldsymbol{Y}^k + \alpha d\boldsymbol{Y}_p^k)}{n} \\
&= \ (1+\zeta\gamma)(1-\alpha)\theta^k\mu^0 - \frac{(1-\alpha)\boldsymbol{X}^k \bullet \boldsymbol{Y}^k}{n} + \frac{\alpha^2 d\boldsymbol{X}_p^k \bullet d\boldsymbol{Y}_p^k}{n} \\
&\geq \ \Big((1+\zeta\gamma)(1-\alpha) - (1-\alpha)(1+\zeta\gamma^k) - \alpha^2\delta_p^k/n\Big)\,\theta^k\mu^0 \\
&= \ \Big(-\alpha^2\delta_p^k/n + \zeta(1-\alpha)(\gamma-\gamma^k)\Big)\,\theta^k\mu^0 \\
&\geq \ \Big(-\alpha^2\delta_p^k + (1-\alpha)(\gamma-\gamma^k)\Big)\,\zeta\theta^k\mu^0 \ \text{(since } \zeta \geq 1/n).
\end{aligned}$$

Note that the coefficient of $\zeta\theta^k\mu^0$ is the same quadratic polynomial as the one appeared in (27) above. Hence

$$(1+\zeta\gamma)(1-\alpha)\theta^k\mu^p - \frac{(\boldsymbol{X}^k + \alpha d\boldsymbol{X}_p^k) \bullet (\boldsymbol{Y}^k + \alpha d\boldsymbol{Y}_p^k)}{n} \geq 0 \ \text{ for every } \alpha \in [0, \hat{\alpha}_p^k].$$

Thus we have shown the desired relation (26). ∎

Lemma 3.8. *Let $\gamma \in (0,1)$, $\mu^0 > 0$, $(\boldsymbol{X}^0, \boldsymbol{Y}^0) = (\sqrt{\mu^0}\boldsymbol{I}, \sqrt{\mu^0}\boldsymbol{I})$ and $\sigma > 0$. Suppose that $(\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k)$ satisfies (7) for some $\theta^{k+1} \in [0,1]$. Let $(d\boldsymbol{X}_c^k, d\boldsymbol{Y}_c^k)$ be the solution of the system (14) of equations. Define δ_c^k, $\hat{\alpha}_c^k$, $\check{\gamma}^{k+1}$ and $\hat{\gamma}^k$ by (15). Then $0 \leq \hat{\gamma}^{k+1} \leq \check{\gamma}^{k+1} < \gamma$, and the pair of $\alpha_c^k = \hat{\alpha}_c^k$ and $\gamma^{k+1} = \check{\gamma}^{k+1}$ satisfies (16).*

Proof: We first observe that for every $\alpha \in [0,1]$,

$$\begin{aligned}
&\frac{(\boldsymbol{X}_c^k + \alpha d\boldsymbol{X}_c^k)(\boldsymbol{Y}_c^k + \alpha d\boldsymbol{Y}_c^k) + (\boldsymbol{Y}_c^k + \alpha d\boldsymbol{Y}_c^k)(\boldsymbol{X}_c^k + \alpha d\boldsymbol{X}_c^k)}{2} \\
&= \ \frac{(1-\alpha)(\boldsymbol{X}_c^k\boldsymbol{Y}_c^k + \boldsymbol{Y}_c^k\boldsymbol{X}_c^k) + 2\alpha\theta^{k+1}\mu^0\boldsymbol{I} + \alpha^2(d\boldsymbol{X}_c^k d\boldsymbol{Y}_c^k + d\boldsymbol{Y}_c^k d\boldsymbol{X}_c^k)}{2} \\
&\succeq \ (1-\alpha)(1-\gamma)\theta^{k+1}\mu^0\boldsymbol{I} + \alpha\theta^{k+1}\mu^0\boldsymbol{I} - \alpha^2\|d\boldsymbol{X}_c^k\|_F\|d\boldsymbol{Y}_c^k\|_F\boldsymbol{I} \\
&= \ \Big(-\alpha^2\delta_c^k + \gamma\alpha + (1-\gamma)\Big)\,\theta^{k+1}\mu^0\boldsymbol{I}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\frac{(\boldsymbol{X}_c^k + \alpha d\boldsymbol{X}_c^k)(\boldsymbol{Y}_c^k + \alpha d\boldsymbol{Y}_c^k) + (\boldsymbol{Y}_c^k + \alpha d\boldsymbol{Y}_c^k)(\boldsymbol{X}_c^k + \alpha d\boldsymbol{X}_c^k)}{2} \succeq \boldsymbol{O} \ \text{ for every } \alpha \in [0, \hat{\alpha}_c^k], \\
&\frac{(\boldsymbol{X}_c^k + \hat{\alpha}_c^k d\boldsymbol{X}_c^k)(\boldsymbol{Y}_c^k + \hat{\alpha}_c^k d\boldsymbol{Y}_c^k) + (\boldsymbol{Y}_c^k + \hat{\alpha}_c^k d\boldsymbol{Y}_c^k)(\boldsymbol{X}_c^k + \hat{\alpha}_c^k d\boldsymbol{X}_c^k)}{2} \succeq (1-\check{\gamma}^{k+1})\theta^{k+1}\mu^0\boldsymbol{I}.
\end{aligned}$$

We also see that for every $\alpha \in [0,1]$,

$$\begin{aligned}
&\frac{(\boldsymbol{X}_c^k + \alpha d\boldsymbol{X}_c^k) \bullet (\boldsymbol{Y}_c^k + \alpha d\boldsymbol{Y}_c^k)}{n} \\
&= (1-\alpha)\frac{\boldsymbol{X}_c^k \bullet \boldsymbol{Y}_c^k}{n} + \alpha\theta^{k+1}\mu^0 + \alpha^2 \frac{d\boldsymbol{X}_c^k \bullet d\boldsymbol{Y}_c^k}{n} \\
&\leq \left((1-\alpha)(1+\zeta\gamma) + \alpha + \alpha^2\delta_c^k/n\right)\theta^{k+1}\mu^0 \\
&= \left(\alpha^2\delta_c^k/n - \zeta\gamma\alpha + (1+\zeta\gamma)\right)\theta^{k+1}\mu^0 \\
&\leq \left(\zeta\alpha^2\delta_c^k - \zeta\gamma\alpha + (1+\zeta\gamma)\right)\theta^{k+1}\mu^0 \ \text{(since } \zeta \geq 1/n).
\end{aligned}$$

It follows that

$$\frac{(\boldsymbol{X}_c^k + \hat{\alpha}_c^k d\boldsymbol{X}_c^k) \bullet (\boldsymbol{Y}_c^k + \hat{\alpha}_c^k d\boldsymbol{Y}_c^k)}{n} \leq (1+\zeta\check{\gamma}^{k+1})\theta^{k+1}\mu^0.$$

Thus we have shown that $(\boldsymbol{X}_c^k + \hat{\alpha}_c^k d\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k + \hat{\alpha}_c^k d\boldsymbol{Y}_c^k) \in \widetilde{\mathcal{N}}(\check{\gamma}^{k+1}, \theta^{k+1}\mu^0)$, and the desired result follows. ∎

4. Proof of Theorem 2.3.

4.1. Proof of (i), (ii), (iii) and (iv) of Theorem 2.3

By definition, $(\boldsymbol{X}_c^0, \boldsymbol{Y}_c^0, \theta^0, \gamma^0)$ satisfies (4), (6) and $\theta^0 = 1$. Let $q \geq 0$. Assume that we have computed $(\boldsymbol{X}_c^q, \boldsymbol{Y}_c^q, \theta^q, \gamma^q)$ satisfying (4) and (6) for $k = q$, we will investigate each step of the algorithm.

Step 1: Applying Lemma 3.3, we know that if the inequality (10) with $\theta = \theta^q$ and $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^q, \boldsymbol{Y}^q)$ at Step 1 does not hold then there is no solution of the SDLCP (1) satisfying (9). Hence we have shown (ii) of Theorem 2.3.

Step 2 (Predictor Step): Now suppose that the inequality (10) holds for $k = q$. The validity of Step 2 follows from Corollary 3.2 and Lemma 3.7. Corollary 3.2 ensures the existence and the uniqueness of the solution $(d\boldsymbol{X}_p^q, d\boldsymbol{Y}_p^q)$ of the system (11) of equations with $k = q$. In view of Lemma 3.7, we can consistently take a step length α_p^q satisfying

$$\begin{aligned}
&0 < \hat{\alpha}_p^q \leq \alpha_p^q \leq \check{\alpha}_p^q \leq 1, \\
&0 \leq \theta^{q+1} = (1-\alpha_p^q)\theta^q < \theta^q, \\
&(\boldsymbol{X}_c^q, \boldsymbol{Y}_c^q) = (\boldsymbol{X}^q, \boldsymbol{Y}^q) + \alpha_p^q(d\boldsymbol{X}_p^q, d\boldsymbol{Y}_p^q) \in \widetilde{\mathcal{N}}(\gamma, \theta^{k+1}\mu^0).
\end{aligned}$$

Hence (5) and the first relation of (7) with $k = q$ follow. To derive the second relation of (7) with $k = q$, we observe that

$$\begin{aligned}
(\boldsymbol{X}_c^q, \boldsymbol{Y}_c^q) &= (\boldsymbol{X}^q, \boldsymbol{Y}^q) + \alpha_p^q(d\boldsymbol{X}_p^q, d\boldsymbol{Y}_p^q) \\
&= (1-\alpha_p^q)(\boldsymbol{X}^q, \boldsymbol{Y}^q) + \alpha_p^q(\boldsymbol{X}^q + d\boldsymbol{X}_p^q, \boldsymbol{Y}^q + d\boldsymbol{Y}_p^q)
\end{aligned}$$

$$
\begin{aligned}
\in & \quad (1-\alpha_p^q)\left(\mathcal{F}+\theta^q\left((\boldsymbol{X}^0,\boldsymbol{Y}^0)-(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right)\right)+\alpha_p^q\mathcal{F} \\
& \quad \text{(since (6) and (11) hold for } k=q) \\
= & \quad \mathcal{F}+(1-\alpha_p^q)\theta^k\left((\boldsymbol{X}^0,\boldsymbol{Y}^0)-(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right) \\
= & \quad \mathcal{F}+\theta^{k+1}\left((\boldsymbol{X}^0,\boldsymbol{Y}^0)-(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right).
\end{aligned}
$$

Thus we have shown the latter relation of (7) with $k=q$.

Step 3 : If $\theta^{q+1}\le\epsilon$ then $(\boldsymbol{X}_p^q,\boldsymbol{Y}_p^q)$ satisfies (8) for $k=q$; hence (iii) of Theorem 2.3 holds. If the inequality (13) with $k=q$ does not hold, then we have the same conclusion as in Step 1 by applying Lemma 3.3 with $\theta=\theta^{q+1}$ and $(\boldsymbol{X},\boldsymbol{Y})=(\boldsymbol{X}_c^q,\boldsymbol{Y}_c^q)$. Hence we have shown (iv) of Theorem 2.3.

Step 4 (Corrector Step): The validity of this step follows from Corollary 3.2 and Lemma 3.8. Corollary 3.2 ensures the existence and the uniqueness of the solution $(d\boldsymbol{X}_c^q,d\boldsymbol{Y}_c^q)$ of the system (14) of equations with $k=q$. By Lemma 3.8, the first relation of (6) holds for $k=q+1$. and we can consistently take a step length α_c^q and a γ^{q+1} satisfying (16) with $k=q$. Hence the second relation of (4) holds for $q=k+1$. We also see by (7) and (14) with $k=q$ that

$$
\begin{aligned}
(\boldsymbol{X}^{q+1},\boldsymbol{Y}^{q+1}) &= (\boldsymbol{X}_c^q,\boldsymbol{Y}_c^q)+\alpha_c^q(d\boldsymbol{X}_c^q,d\boldsymbol{Y}_c^q) \\
&\in \mathcal{F}+\theta^{q+1}\left((\boldsymbol{X}^0,\boldsymbol{Y}^0)-(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right)+\mathcal{F}_0 \\
&= \mathcal{F}+\theta^{q+1}\left((\boldsymbol{X}^0,\boldsymbol{Y}^0)-(\overline{\boldsymbol{X}},\overline{\boldsymbol{Y}})\right).
\end{aligned}
$$

Therefore the latter relation of (6) holds for $k=q+1$.

We have shown (4) and (6) with $k=q+1$ to proceed to the $(q+1)$th iteration consistently, and proved (i) of Theorem 2.3.

4.2. Proof of (v) of Theorem 2.3.

Now we assume that $\epsilon>0$. Then we have $\theta^k\ge\epsilon$, the inequalities (10) and (13) while Algorithm 2.2 is running. We see by Lemmas 3.5 and 3.6 that

$$
0\le\delta_p^k\le n^4\kappa_p^2/(\epsilon\mu^0) \ \text{ and } 0\le\delta_c^k\le n^4\kappa_c^2/(\epsilon\mu^0) \ \text{ for every } k=0,1,\dots\ .
$$

Hence, defining

$$
\bar\gamma=\max\{\gamma(1-\gamma\epsilon\mu^0/(4n^4\kappa_c^2)),\gamma/2\} \ \text{ and } \bar\alpha_p=\frac{2}{\sqrt{1+4n^4\kappa_p^2/\left(\epsilon\mu^0(\gamma-\bar\gamma)\right)}+1}, \tag{28}
$$

we obtain that

$$
0\le\gamma^k\le\check\gamma^k\le\bar\gamma<\gamma \ \text{ and } 0<\bar\alpha_p\le\hat\alpha_p^k\le\alpha_p^k \ \text{ for every } k=0,1,2,\dots\ ;
$$

hence

$$
\epsilon \ \le \ \theta^k=\theta^0\prod_{j=1}^{k-1}(1-\alpha_p^j)\le\theta^0(1-\bar\alpha_p)^{(k-1)} \ \text{ for every } k=0,1,2,\dots\ . \tag{29}
$$

Therefore Algorithm 2.2 must stop in a finite number of iterations. This completes the proof of Theorem 2.3.

Remark 4.1. If we regard κ_p, κ_c, γ and μ^0 as constant in (28), then we can take a positive constant δ such that $\delta\epsilon/n^4 \leq \bar{\alpha}_p$ for every sufficiently large n. In this case we can derive from (29) that Algorithm 2.2 stops in $O((n^4/\epsilon)\log(1/\epsilon))$ iterations; hence Algorithm 2.2 works as a fully polynomial-time approximation scheme [22].

5. Local Convergence.

In the remainder of the paper, we assume Hypothesis 2.1 and discuss the local convergence of the sequence generated by Algorithm 2.2 with taking taking $\epsilon = 0$. Hypothesis 2.1 ensures that (ii) and (iv) of Theorem 2.3 can not occur. Since $\epsilon = 0$, if Algorithm 2.2 stops in a finite number of iterations (*i.e.,* the sequence is finite) then we obtain an exact solution $(\boldsymbol{X}_c^{k+1}, \boldsymbol{Y}_c^{k+1})$ of the SDLCP (1) at Step 3 of the last iteration.

Assuming that the sequence is infinite, we establish:

Theorem 5.1. *(Local Convergence Theorem) Assume that Hypothesis 2.1 and Condition 5.2 below hold. Let* $\{(\boldsymbol{X}^k, \boldsymbol{Y}^k, \boldsymbol{X}_c^k, \boldsymbol{Y}_c^k, \theta^k, \gamma^k)\}$ *be the sequence generated by Algorithm 2.2 with taking* $\epsilon = 0$ *at Step 0.*

(i) *The* $\hat{\alpha}_p^k$ *defined in (12) satisfies that* $\hat{\alpha}_p^k \to 1$ *as* $k \to \infty$.

(ii) *There is a positive constant* η *such that* $\theta^{k+1} \leq \eta(\theta^k)^2$ *for every* $k = 0, 1, 2, \ldots$.

Condition 5.2. *(Strict Complementarity) There is a solution* $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ *of the SDLCP (1), whose existence is ensured by Hypothesis 2.1, such that* $\boldsymbol{X}^* + \boldsymbol{Y}^* \succ \boldsymbol{O}$.

The assertion (ii) of the theorem implies:

(ii)' The optimality and feasibility measure θ^k converges zero quadratically.

The remainder of this section is devoted to a proof of the theorem. The subsequent argument is heavily relied on Potra and Sheng [24].

We first observe that $\boldsymbol{X}^*$ and $\boldsymbol{Y}^*$ commute. Hence there exists an orthogonal matrix $\boldsymbol{Q}$ such that

$$\boldsymbol{Q}^\top \boldsymbol{X}^* \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{\Lambda}_B^* & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \quad \text{and} \quad \boldsymbol{Q}^\top \boldsymbol{Y}^* \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Lambda}_N^* \end{pmatrix},$$

where $\boldsymbol{\Lambda}_B^*$ and $\boldsymbol{\Lambda}_N^*$ are diagonal matrices with dimensions k and $n-k$ for some $k \in \{0, 1, 2, \ldots, n\}$, respectively. For each $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S} \times \mathcal{S}$, define

$$\boldsymbol{Q}^\top \boldsymbol{X} \boldsymbol{Q} \equiv \hat{\boldsymbol{X}} = \begin{pmatrix} \hat{\boldsymbol{X}}_B & \hat{\boldsymbol{X}}_J \\ \hat{\boldsymbol{X}}_J^\top & \hat{\boldsymbol{X}}_N \end{pmatrix} \quad \text{and} \quad \boldsymbol{Q}^\top \boldsymbol{Y} \boldsymbol{Q} \equiv \hat{\boldsymbol{Y}} = \begin{pmatrix} \hat{\boldsymbol{Y}}_B & \hat{\boldsymbol{Y}}_J \\ \hat{\boldsymbol{Y}}_J^\top & \hat{\boldsymbol{Y}}_N \end{pmatrix}. \tag{30}$$

Here, $\hat{\boldsymbol{X}}_B$ and $\hat{\boldsymbol{Y}}_B$ (resp. $\hat{\boldsymbol{X}}_N$ and $\hat{\boldsymbol{Y}}_N$) are matrices of same size with $\boldsymbol{\Lambda}_B^*$ (resp. $\boldsymbol{\Lambda}_N^*$).

The following Lemma is easily derived from Lemma 4.4 in [23].

Lemma 5.3. *Suppose that* $(\boldsymbol{X}, \boldsymbol{Y}) \in \tilde{\mathcal{N}}(\gamma, \tau)$. *Then, under Condition 5.2, we have*

$$\hat{\boldsymbol{X}} = \boldsymbol{Q}^\top \boldsymbol{X} \boldsymbol{Q} = \begin{pmatrix} O(1) & O(\sqrt{\tau}) \\ O(\sqrt{\tau}) & O(\tau) \end{pmatrix} \quad \textit{and} \quad \hat{\boldsymbol{Y}} = \boldsymbol{Q}^\top \boldsymbol{Y} \boldsymbol{Q} = \begin{pmatrix} O(\tau) & O(\sqrt{\tau}) \\ O(\sqrt{\tau}) & O(1) \end{pmatrix}. \tag{31}$$

Proof: By Lemma 2.2 of [23], we have

$$\lambda_{min}(\boldsymbol{X}^{\frac{1}{2}} \boldsymbol{Y} \boldsymbol{X}^{\frac{1}{2}}) \geq \frac{1}{2} \lambda_{min}(\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X}),$$

where $\lambda_{min}(\boldsymbol{A})$ is the smallest eigenvalue of $\boldsymbol{A}$. Hence we have $\boldsymbol{X}^{\frac{1}{2}} \boldsymbol{Y} \boldsymbol{X}^{\frac{1}{2}} \succeq (1-\gamma)\tau \boldsymbol{I}$. Using the fact, we can easily follow the proof of Lemma 4.4 of [23]. ∎

Define an affine subspace containing the solution set of the SDLCP;

$$\widetilde{\mathcal{M}} \equiv \left\{ (\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F} \cap (\mathcal{S} \times \mathcal{S}) : \begin{array}{l} \hat{\boldsymbol{X}} = \begin{pmatrix} \boldsymbol{M}_B & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \text{ and } \hat{\boldsymbol{Y}} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{M}_N \end{pmatrix}, \\ \text{where } \boldsymbol{M}_B \text{ (resp. } \boldsymbol{M}_N) \text{ is an arbitrary matrix} \\ \text{of same size with } \boldsymbol{\Lambda}_B^* \text{ (resp. } \boldsymbol{\Lambda}_N^*) \end{array} \right\},$$

where $\boldsymbol{M}_B$ (resp. $\boldsymbol{M}_N$) is a matrix of same size with $\mathbf{Lambda}_B^*$ (resp. $\mathbf{Lambda}_N^*$). Then:

Lemma 5.4. *Under Condition 5.2, every accumulation point of* $(\boldsymbol{X}^k, \boldsymbol{Y}^k)$ *is a strictly complementary solution such that* $\hat{\boldsymbol{X}}_B = \Omega(1)$ *and* $\hat{\boldsymbol{Y}}_N = \Omega(1)$.

Proof: See Lemma 4.6 of [23] ∎

Let $\rho = \max\{\|\hat{\boldsymbol{X}}_J\|_F, \|\hat{\boldsymbol{Y}}_J\|_F, \tau\}$.

Lemma 5.5. *Suppose that* $(\boldsymbol{X}, \boldsymbol{Y}) \in \tilde{\mathcal{N}}(\gamma, \tau)$. *Under Condition 5.2, we have*

$$\|(\boldsymbol{X} - \check{\boldsymbol{X}}, \boldsymbol{Y} - \check{\boldsymbol{Y}})\|_F = O(\rho), \tag{32}$$

where $(\check{\boldsymbol{X}}, \check{\boldsymbol{Y}})$ *is the solution of the minimization problem:*

$$\min\{\|(\boldsymbol{X} - \boldsymbol{X}', \boldsymbol{Y} - \boldsymbol{Y}')\|_F : (\boldsymbol{X}', \boldsymbol{Y}') \in \widetilde{\mathcal{M}}\}. \tag{33}$$

Proof: Suppose that (32) does not hold, *i.e.*, there exists a convergent sequence $(\boldsymbol{X}_k, \boldsymbol{Y}_k) \in \tilde{\mathcal{N}}(\gamma, \tau)$ such that

$$\xi_k = \|(\boldsymbol{X}_k - \check{\boldsymbol{X}}_k, \boldsymbol{Y}_k - \check{\boldsymbol{Y}}_k)\|_F / \rho_k \to \infty, \rho_k \to 0 \text{ as } k \to \infty. \tag{34}$$

Let define

$$(\Delta \boldsymbol{X}_k, \Delta \boldsymbol{Y}_k) = \frac{1}{\rho_k \xi_k} (\boldsymbol{X}_k - \check{\boldsymbol{X}}_k, \boldsymbol{Y}_k - \check{\boldsymbol{Y}}_k). \tag{35}$$

Since $\|(\Delta \boldsymbol{X}_k, \Delta \boldsymbol{Y}_k)\|_F = 1$, taking a convergent subsequence, we may assume that

$$(\Delta \boldsymbol{X}_k, \Delta \boldsymbol{Y}_k) \to (\Delta \boldsymbol{X}', \Delta \boldsymbol{Y}').$$

Then we have that $(\Delta \boldsymbol{X}', \Delta \boldsymbol{Y}') \in \mathcal{F}_0$. From (31), we have for any $i \in B$,

$$|\boldsymbol{q}_i^\top (\Delta \boldsymbol{Y}_k) \boldsymbol{q}_i| = |\boldsymbol{q}_i^\top \boldsymbol{Y}_k \boldsymbol{q}_i| / (\xi_k \rho_k) \leq \|(\hat{\boldsymbol{Y}}_k)_B\|_F / (\xi_k \rho_k) \leq O(\tau_k)/(\xi_k \rho_k) = O(1/\xi_k) = o(1),$$

which implies $\boldsymbol{q}_i^\top \Delta \boldsymbol{Y}' \boldsymbol{q}_i = 0$ for each $i \in B$. Similarly, $\boldsymbol{q}_i^\top \Delta \boldsymbol{X}' \boldsymbol{q}_i = 0$ for each $i \in N$. For each pair i, j where $i \in B, j \in N$ or $i \in N, j \in B$, we have

$$|\boldsymbol{q}_i^\top (\Delta \boldsymbol{Y}_k) \boldsymbol{q}_j| = |\boldsymbol{q}_i^\top \boldsymbol{Y}_k \boldsymbol{q}_j| / (\xi_k \rho_k) \leq \|(\hat{\boldsymbol{Y}}_k)_J\|_F / (\rho_k \xi_k) \leq 1/\xi_k = o(1),$$

which implies $\boldsymbol{q}_i^\top \Delta \boldsymbol{Y}' q_j = 0$. Similarly, we have $\boldsymbol{q}_i^\top \Delta \boldsymbol{X}' \boldsymbol{q}_j = 0$ for any pair i, j described above. Hence,

$$(\check{\boldsymbol{X}}_k, \check{\boldsymbol{Y}}_k) + \nu(\Delta \boldsymbol{X}', \Delta \boldsymbol{Y}') \in \widetilde{\mathcal{M}}, \text{ for all } \nu \in \mathcal{R}.$$

Therefore, we have

$$\begin{array}{rl} & \dfrac{\|(\boldsymbol{X}_k - (\check{\boldsymbol{X}}_k + \rho_k \xi_k \Delta \boldsymbol{X}'), (\boldsymbol{Y}_k - (\check{\boldsymbol{Y}}_k + \rho_k \xi_k \Delta \boldsymbol{Y}'))\|_F}{\|(\boldsymbol{X}_k - \check{\boldsymbol{X}}_k, \boldsymbol{Y}_k - \check{\boldsymbol{Y}}_k)\|_F} \\ = & \dfrac{1}{\rho_k \xi_k} \|(\boldsymbol{X}_k - (\check{\boldsymbol{X}}_k + \rho_k \xi_k \Delta \boldsymbol{X}'), (\boldsymbol{Y}_k - (\check{\boldsymbol{Y}}_k + \rho_k \xi_k \Delta \boldsymbol{Y}'))\|_F \\ = & \|(\Delta \boldsymbol{X}_k - \Delta \boldsymbol{X}', \Delta \boldsymbol{Y}_k - \Delta \boldsymbol{Y}')\|_F \\ \to & 0 \quad \text{as } k \to \infty. \end{array}$$

This contradicts to the fact that $(\check{\boldsymbol{X}}_k, \check{\boldsymbol{Y}}_k)$ is the solution of the minimization problem (33) for large k. ∎

Lemma 5.6. *Suppose that $(\boldsymbol{X}, \boldsymbol{Y}) \in \tilde{\mathcal{N}}(\gamma, \tau)$. Under condition 5.2, we have*

$$\hat{\boldsymbol{X}}_J = O(\tau), \ \hat{\boldsymbol{Y}}_J = O(\tau). \tag{36}$$

Proof: Suppose that (36) does not hold. We may assume there exists a subsequence $(\boldsymbol{X}_k, \boldsymbol{Y}_k) \in \tilde{\mathcal{N}}(\gamma, \tau_k)$ such that

$$\chi_k = \frac{\|(\hat{\boldsymbol{X}}_k)_J\|_F}{\tau_k} \to \infty, \tau_k \to 0, \quad \text{as } k \to \infty. \tag{37}$$

Since $(\boldsymbol{X}_k, \boldsymbol{Y}_k) \in \tilde{\mathcal{N}}(\gamma, \tau_k)$, we have

$$\boldsymbol{X}_k \boldsymbol{Y}_k + \boldsymbol{Y}_k \boldsymbol{X}_k = O(\tau_k) = o(\|(\hat{\boldsymbol{X}}_k)_J\|_F). \tag{38}$$

This yields

$$(\hat{\boldsymbol{X}}_k)_B (\hat{\boldsymbol{Y}}_k)_J + (\hat{\boldsymbol{X}}_k)_J (\hat{\boldsymbol{Y}}_k)_N = o(\|(\hat{\boldsymbol{X}}_k)_J\|_F) \tag{39}$$

Hence

$$(\hat{\boldsymbol{Y}}_k)_J = -(\hat{\boldsymbol{X}}_k)_B^{-1} (\hat{\boldsymbol{X}}_k)_J (\hat{\boldsymbol{Y}}_k)_N + (\hat{\boldsymbol{X}}_k)_B^{-1} o(\|(\hat{\boldsymbol{X}}_k)_J\|_F). \tag{40}$$

Since $(\hat{\boldsymbol{X}}_k)_B = \Omega(1)$, $(\hat{\boldsymbol{Y}}_k)_N = \Omega(1)$ from Lemma 5.4, we have that

$$\|(\hat{\boldsymbol{Y}}_k)_J\|_F = \Omega(\|(\hat{\boldsymbol{X}}_k)_J\|_F). \tag{41}$$

Taking a convergent subsequence $(\boldsymbol{X}_k, \boldsymbol{Y}_k)$ such that

$$(\boldsymbol{X}_k, \boldsymbol{Y}_k) \to (\boldsymbol{X}^\infty, \boldsymbol{Y}^\infty) \quad \text{and} \quad (\frac{(\hat{\boldsymbol{X}}_k)_J}{\|(\hat{\boldsymbol{X}}_k)_J\|_F}, \frac{(\hat{\boldsymbol{Y}}_k)_J}{\|(\hat{\boldsymbol{X}}_k)_J\|_F}) \to (\boldsymbol{X}'_J, \boldsymbol{Y}'_J)$$

as $k \to \infty$. By (39) and letting $k \to \infty$, we get

$$\hat{\boldsymbol{X}}_B^\infty \boldsymbol{Y}'_J + \boldsymbol{X}'_J \hat{\boldsymbol{Y}}_N^\infty = \boldsymbol{O}.$$

By Lemma 5.4, $\hat{\boldsymbol{X}}_B^\infty, \hat{\boldsymbol{Y}}_N^\infty$ are positive definite matrices. Note that $\|\boldsymbol{X}'_J\|_F = 1$. Hence,

$$\begin{array}{rcl} \boldsymbol{X}'_J \bullet \boldsymbol{Y}'_J & = & -\boldsymbol{X}'_J \bullet [(\boldsymbol{X}_B^\infty)^{-1} \boldsymbol{X}'_J \boldsymbol{Y}_N^\infty] \\ & = & -\mathrm{Tr}\ ([(\boldsymbol{X}_B^\infty)^{-\frac{1}{2}} \boldsymbol{X}'_J (\boldsymbol{Y}_N^\infty)^{\frac{1}{2}}][(\boldsymbol{X}_B^\infty)^{-\frac{1}{2}} \boldsymbol{X}'_J (\boldsymbol{Y}_N^\infty)^{\frac{1}{2}}]^\top) \\ & < & 0. \end{array} \tag{42}$$

Let $(\check{\boldsymbol{X}}_k, \check{\boldsymbol{Y}}_k) \in \widetilde{\mathcal{M}}$ be defined as in Lemma 5.5 and let

$$\boldsymbol{X}" = \boldsymbol{X}_k - \check{\boldsymbol{X}}_k + \theta(\boldsymbol{X}^\infty - \boldsymbol{X}^0), \quad \boldsymbol{Y}" = \boldsymbol{Y}_k - \check{\boldsymbol{Y}}_k + \theta(\boldsymbol{Y}^\infty - \boldsymbol{X}^0).$$

It is easily seen that $(\boldsymbol{X}", \boldsymbol{Y}") \in \mathcal{F}^0$ and therefore $\boldsymbol{X}" \bullet \boldsymbol{Y}" \geq 0$. By Lemma 5.5, we have

$$(\boldsymbol{X}_k - \check{\boldsymbol{X}}_k) \bullet (\boldsymbol{Y}_k - \check{\boldsymbol{Y}}_k) \geq O(\rho_k \tau_k),$$

i.e.,

$$((\hat{\boldsymbol{X}}_k)_B - (\hat{\check{\boldsymbol{X}}}_k)_B) \bullet (\hat{\boldsymbol{Y}}_k)_B + 2(\hat{\boldsymbol{X}}_k)_J \bullet (\hat{\boldsymbol{Y}}_k)_J + (\hat{\boldsymbol{X}}_k)_N \bullet ((\hat{\boldsymbol{Y}}_k)_N - (\hat{\check{\boldsymbol{Y}}}_k)_N) \geq O(\rho_k \tau_k). \tag{43}$$

From Lemmas 5.3 and 5.5, we have

$$(\hat{\boldsymbol{Y}}_k)_B = O(\tau_k), (\hat{\boldsymbol{X}}_k)_N = O(\tau_k), \tag{44}$$

$$(\hat{\boldsymbol{X}}_k)_B - (\hat{\check{\boldsymbol{X}}}_k)_B = O(\rho_k), (\hat{\boldsymbol{Y}}_k)_N - (\hat{\check{\boldsymbol{Y}}}_k)_N = O(\rho_k). \tag{45}$$

Therefore we have

$$(\hat{\boldsymbol{X}}_k)_J \bullet (\hat{\boldsymbol{Y}}_k)_J \geq O(\rho_k \tau_k). \tag{46}$$

Let observe that

$$\begin{array}{rcl} \dfrac{\rho_k \tau_k}{\|(\hat{\boldsymbol{X}}_k)_J\|_F^2} & = & \dfrac{\tau_k}{\|(\hat{\boldsymbol{X}}_k)_J\|_F^2} \max\{\|(\hat{\boldsymbol{X}}_k)_J\|_F, \|(\hat{\boldsymbol{Y}}_k)_J\|_F, \tau_k\} \\ & = & \dfrac{\tau_k}{\|(\hat{\boldsymbol{X}}_k)_J\|_F} \max\left\{1, \dfrac{\|(\hat{\boldsymbol{Y}}_k)_J\|_F}{\|(\hat{\boldsymbol{X}}_k)_J\|_F}, \dfrac{\tau_k}{\|(\hat{\boldsymbol{X}}_k)_J\|_F}\right\} \\ & = & o(1) \end{array} \tag{47}$$

Dividing both sides of (46) by $\|(\hat{\boldsymbol{X}}_k)_J\|_F^2$, recalling (47), and letting $k \to \infty$, we obtain

$$\boldsymbol{X}'_J \bullet \boldsymbol{Y}'_J \geq 0$$

which contradicts (42). ∎

From Lemma 5.6, we get $\rho_k = \Omega(\tau_k)$. Therefore, if $(\boldsymbol{X}, \boldsymbol{Y}) \in \tilde{\mathcal{N}}(\gamma, \tau)$, we have

$$\|(\boldsymbol{X} - \check{\boldsymbol{X}}, \boldsymbol{Y} - \check{\boldsymbol{Y}})\|_F = O(\tau). \tag{48}$$

Lemma 5.7. *Suppose that* $(\boldsymbol{X}, \boldsymbol{Y}) \in \tilde{\mathcal{N}}(\gamma, \tau)$. *Then we have* $\delta_p = O(\tau)$ *and* $\delta_c = O(1)$.

Proof: By Lemma 3.1 (iii) and the fact that $2\tau \boldsymbol{I} - (\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X}) = O(\tau)$, we have $\delta_c = O(1)$. We have only to show that $\delta_p = O(\tau)$. Let $\Delta \boldsymbol{X} = d\boldsymbol{X}_p + (\boldsymbol{X} - \check{\boldsymbol{X}})$, $\Delta \boldsymbol{Y} = d\boldsymbol{Y}_p + (\boldsymbol{Y} - \check{\boldsymbol{Y}})$. Note that $(\Delta \boldsymbol{X}, \Delta \boldsymbol{Y})$ is a solution of (17) with $\boldsymbol{C} = (\boldsymbol{X} - \check{\boldsymbol{X}})(\boldsymbol{Y} - \check{\boldsymbol{Y}}) + (\boldsymbol{Y} - \check{\boldsymbol{Y}})(\boldsymbol{X} - \check{\boldsymbol{X}})$. By (48), we have $\|\boldsymbol{C}\|_F = O(\tau^2)$. Hence, by Lemma 3.1 (iii), we have $\|\Delta \boldsymbol{X}\|_F = O(\tau)$ and $\|\Delta \boldsymbol{Y}\|_F = O(\tau)$. Therefore, using (48) again, we have $\delta_p = O(\tau)$. ∎

Proof of Theorem 5.1: By Lemma 5.7, we know that there exists a constant C such that $\delta_c \leq C$ for all iterate, which implies the existence of a positive constant $\bar{\gamma}$ such that

$$0 \leq \gamma^k \leq \bar{\gamma} < \gamma \quad \text{for every } k = 0, 1, 2, \cdots. \tag{49}$$

We see by Lemma 5.7 that

$$0 \leq \delta_p^k = \frac{\|d\boldsymbol{X}_p^k\|_F \|d\boldsymbol{Y}_p^k\|_F}{\theta^k \mu^0} \leq O(\theta^k) \quad \text{for every } k = 0, 1, 2, \cdots. \tag{50}$$

Specifically, $\delta_p^k \to 0$ as $k \to \infty$. Hence $\hat{\alpha}_p^k \to 1$ as $k \to \infty$. This means (i) holds. Using the inequalities (49) and (50), we see that for every $k = 0, 1, 2, \cdots$,

$$\begin{aligned}
1 - \alpha_p^k &\leq 1 - \hat{\alpha}_p^k \\
&= 1 - \frac{2}{\sqrt{1 + 4\delta_p^k/(\gamma - \gamma^k)} + 1} \\
&= \frac{4\delta_p^k/(\gamma - \gamma^k)}{(\sqrt{1 + 4\delta_p^k/(\gamma - \gamma^k)} + 1)^2} \\
&\leq \delta_p^k/(\gamma - \gamma^k) \\
&\leq \delta_p^k/(\gamma - \bar{\gamma}) \\
&= O(\theta^k).
\end{aligned}$$

Finally we obtain by the construction of θ^{k+1} that

$$\theta^{k+1} = (1 - \alpha_p^k)\theta^k = O((\theta^k)^2).$$

Thus θ^k converges zero quadratically. This is the statement (ii).

6. Local Convergence under Nondegeneracy.

Throughout this section, we assume Hypothesis 2.1, Condition 5.2 and

Condition 6.1. *(Nondegeneracy)* $(\boldsymbol{U}, \boldsymbol{V}) = (\boldsymbol{O}, \boldsymbol{O})$ *if* $\boldsymbol{X}^*\boldsymbol{V} + \boldsymbol{U}\boldsymbol{Y}^* = \boldsymbol{O}$ *and* $(\boldsymbol{U}, \boldsymbol{V}) \in \mathcal{F}_0$, *where* $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ *is a solution of the SDLCP (1) whose existence is ensured by Hypothesis 2.1.*

Under these assumptions, the solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ of the SDLCP (1) ensured by Hypothesis 2.1 is the unique one. See Section 5 of [13]. Hence the sequence $\{(\boldsymbol{X}^k, \boldsymbol{Y}^k, \boldsymbol{X}_c^k, \boldsymbol{Y}_c^k, \theta^k, \gamma^k)\}$ generated by Algorithm 2.2 with taking $\epsilon = 0$ at Step 0 satisfies

$$(\boldsymbol{X}^k, \boldsymbol{Y}^k) \to (\boldsymbol{X}^*, \boldsymbol{Y}^*) \ \text{ and } (\boldsymbol{X}_c^k, \boldsymbol{Y}_c^k) \to (\boldsymbol{X}^*, \boldsymbol{Y}^*) \ \text{ as } k \to \infty \tag{51}$$

in addition to (4), (5), (6) and (7).

Remark. Under the strict complementarity , Condition 6.1 is equivalent to the combination of the primal and the dual nondegeneracy conditions given in the paper [3]. This fact was due to Haeberly [6].

Assuming that the sequence is infinite, we establish:

Theorem 6.2. *Assume that Hypothesis 2.1, Conditions 5.2 and 6.1 hold. Let* $\{(\boldsymbol{X}^k, \boldsymbol{Y}^k, \boldsymbol{X}_c^k, \boldsymbol{Y}_c^k, \theta^k, \gamma^k)\}$ *be the sequence generated by Algorithm 2.2 with taking* $\epsilon = 0$ *at Step 0.*

(i) *The* $\hat{\alpha}_c^k$ *defined in (15) satisfies that* $\hat{\alpha}_c^k = 1$ *for every sufficiently large* k.

(ii) *The* $\hat{\gamma}^{k+1}$ *defined in (15) satisfies that* $\hat{\gamma}^{k+1} \to 0$ *as* $k \to \infty$.

The assertions (i) and (ii) of the theorem imply:

(i)' For every sufficiently large k, we can take the unit step length $\alpha_c^k = 1$ at the Corrector Step.

(ii)' The sequence $\{(\boldsymbol{X}^k, \boldsymbol{Y}^k)\}$ converges the solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ tangentially to the central surface in the sense that

$$\|(\boldsymbol{X}^k\boldsymbol{Y}^k + \boldsymbol{Y}^k\boldsymbol{X}^k)/2 - \left(\boldsymbol{X}^k \bullet \boldsymbol{Y}^k/n\right)\boldsymbol{I}\|_F/(\boldsymbol{X}^k \bullet \boldsymbol{Y}^k/n) \to 0 \ \text{ as } k \to \infty.$$

Remark The property (ii)' is assumed in the other papers [13, 23, 24].

We need some lemmas to prove the theorem.

Lemma 6.3. *Assume that*

$$\boldsymbol{X}^*\boldsymbol{V} + \boldsymbol{V}\boldsymbol{X}^* + \boldsymbol{U}\boldsymbol{Y}^* + \boldsymbol{Y}^*\boldsymbol{U} = \boldsymbol{O} \ \text{ and } (\boldsymbol{U}, \boldsymbol{V}) \in \mathcal{F}_0.$$

Then $(\boldsymbol{U}, \boldsymbol{V}) = (\boldsymbol{O}, \boldsymbol{O})$.

Proof: Since $\boldsymbol{X}^* \in \mathcal{S}_+$ and $\boldsymbol{Y}^* \in \mathcal{S}_+$ are commutative, we can take an orthogonal matrix $\boldsymbol{P}$ that diagonalizes $\boldsymbol{X}^*$ and $\boldsymbol{Y}^*$ simultaneously;

$$\boldsymbol{P}^T\boldsymbol{X}^*\boldsymbol{P} = \boldsymbol{\Gamma} \ \text{ and } \boldsymbol{P}^T\boldsymbol{Y}^*\boldsymbol{P} = \boldsymbol{\Delta}$$

for some $n \times n$ diagonal matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Delta}$. Since $0 = \boldsymbol{X}^* \bullet \boldsymbol{Y}^* = \boldsymbol{\Gamma} \bullet \boldsymbol{\Delta}$ and $\boldsymbol{\Gamma} + \boldsymbol{\Delta} = \boldsymbol{P}^T(\boldsymbol{X}^* + \boldsymbol{Y}^*)\boldsymbol{P} \succ \boldsymbol{O}$, we may assume without loss of generality that the diagonal matrix $\boldsymbol{\Gamma}$ and $\boldsymbol{\Delta}$ have the forms

$$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \ \text{ and } \boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Delta}_{22} \end{pmatrix},$$

respectively. Here $\boldsymbol{\Gamma}_{11}$ is an $m \times m$ positive diagonal matrix, $\boldsymbol{\Delta}_{22}$ an $(n-m) \times (n-m)$ positive diagonal matrix, and $0 \leq m \leq n$. Let

$$\boldsymbol{U}' = \boldsymbol{P}^T \boldsymbol{U} \boldsymbol{P} = \begin{pmatrix} \boldsymbol{U}'_{11} & \boldsymbol{U}'_{12} \\ (\boldsymbol{U}'_{12})^T & \boldsymbol{U}'_{22} \end{pmatrix} \quad \text{and } \boldsymbol{V}' = \boldsymbol{P}^T \boldsymbol{V} \boldsymbol{P} = \begin{pmatrix} \boldsymbol{V}'_{11} & \boldsymbol{V}'_{12} \\ (\boldsymbol{V}'_{12})^T & \boldsymbol{V}'_{22} \end{pmatrix}.$$

Then we have from the assumption that

$$\boldsymbol{\Gamma}_{11} \boldsymbol{V}'_{11} + \boldsymbol{V}'_{11} \boldsymbol{\Gamma}_{11} = \boldsymbol{O}, \ \boldsymbol{\Delta}_{22} \boldsymbol{U}'_{22} + \boldsymbol{U}'_{22} \boldsymbol{\Delta}_{22} = \boldsymbol{O} \ \text{ and } \boldsymbol{\Gamma}_{11} \boldsymbol{V}'_{12} + \boldsymbol{U}'_{12} \boldsymbol{\Delta}_{22} = \boldsymbol{O}. \tag{52}$$

Using the Kronecker product of matrices, we can rewrite the first and the second equalities as

$$(\boldsymbol{I} \otimes \boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{11} \otimes \boldsymbol{I})(\mathbf{vec}\ \boldsymbol{V}'_{11}) = \boldsymbol{0} \ \text{ and } (\boldsymbol{I} \otimes \boldsymbol{\Delta}_{22} + \boldsymbol{\Delta}_{22} \otimes \boldsymbol{I})(\mathbf{vec}\ \boldsymbol{U}'_{22}) = \boldsymbol{0}.$$

Since $\boldsymbol{\Gamma}_{11}$ and $\boldsymbol{\Delta}_{22}$ are positive definite, so are the matrices $(\boldsymbol{I} \otimes \boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{11} \otimes \boldsymbol{I})$ and $(\boldsymbol{I} \otimes \boldsymbol{\Delta}_{22} + \boldsymbol{\Delta}_{22} \otimes \boldsymbol{I})$. Hence we have that $\boldsymbol{V}'_{11} = \boldsymbol{O}$ and $\boldsymbol{U}'_{22} = \boldsymbol{O}$. It follows from $\boldsymbol{V}'_{11} = \boldsymbol{O}$, $\boldsymbol{U}'_{22} = \boldsymbol{O}$ and the last equality of (52) that

$$\begin{aligned} \boldsymbol{X}^* \boldsymbol{V} + \boldsymbol{U} \boldsymbol{Y}^* &= \boldsymbol{P} \left(\boldsymbol{\Gamma} \boldsymbol{V}' + \boldsymbol{U}' \boldsymbol{\Delta} \right) \boldsymbol{P}^T \\ &= \boldsymbol{P} \begin{pmatrix} \boldsymbol{\Gamma}_{11} \boldsymbol{V}'_{11} & \boldsymbol{\Gamma}_{11} \boldsymbol{V}'_{12} + \boldsymbol{U}'_{12} \boldsymbol{\Delta}_{22} \\ \boldsymbol{O} & \boldsymbol{U}'_{22} \boldsymbol{\Delta}_{22} \end{pmatrix} \boldsymbol{P}^T \\ &= \boldsymbol{O}. \end{aligned}$$

Therefore we obtain by Condition 6.1 that $(\boldsymbol{U}, \boldsymbol{V}) = (\boldsymbol{O}, \boldsymbol{O})$. ∎

Lemma 6.4. *There exists a positive number η_c such that*

$$\| (d\boldsymbol{X}^k_c, d\boldsymbol{Y}^k_c) \|_F \leq \eta_c \theta^{k+1} \quad \textit{for every } k = 0, 1, 2, \ldots, \ .$$

Proof: Since $(\boldsymbol{X}^k_c, \boldsymbol{Y}^k_c) \in \widetilde{\mathcal{N}}(\gamma, \theta^{k+1}\mu^0)$, we see by (i) of Lemma 3.1 that the right hand side $2\theta^{k+1}\mu^0 \boldsymbol{I} - \boldsymbol{X}^k \boldsymbol{Y}^k - \boldsymbol{Y}^k \boldsymbol{X}^k$ of the system (14) of equations satisfies

$$\frac{\|2\theta^{k+1}\mu^0 \boldsymbol{I} - \boldsymbol{X}^k \boldsymbol{Y}^k - \boldsymbol{Y}^k \boldsymbol{X}^k\|_F}{\theta^{k+1}} \leq \frac{\theta^{k+1}\mu^0 \sqrt{n} + \|\boldsymbol{X}^k \boldsymbol{Y}^k + \boldsymbol{Y}^k \boldsymbol{X}^k\|_F}{\theta^{k+1}} \leq 2\sqrt{n}(\mu^0 + n + n\zeta\gamma).$$

On the other hand, we see from (14) that

$$\left.\begin{aligned} &\boldsymbol{X}^k_c \frac{d\boldsymbol{Y}^k_c}{\theta^{k+1}} + \frac{d\boldsymbol{Y}^k_c}{\theta^{k+1}} \boldsymbol{X}^k_c + \frac{d\boldsymbol{X}^k_c}{\theta^{k+1}} \boldsymbol{Y}^k_c + \boldsymbol{Y}^k_c \frac{d\boldsymbol{X}^k_c}{\theta^{k+1}} = \frac{2\theta^{k+1}\mu^0 \boldsymbol{I} - \boldsymbol{X}^k \boldsymbol{Y}^k - \boldsymbol{Y}^k \boldsymbol{X}^k}{\theta^{k+1}}, \\ &\frac{(d\boldsymbol{X}^k_c, d\boldsymbol{Y}^k_c)}{\theta^{k+1}} \in \mathcal{F}_0. \end{aligned}\right\} \tag{53}$$

Assume on the contrary that the sequence $\left\{ \dfrac{(d\boldsymbol{X}^k_c, d\boldsymbol{Y}^k_c)}{\theta^{k+1}} \right\}$ is unbounded. Along a subsequence, we then have that

$$\frac{\|(d\boldsymbol{X}^k_c, d\boldsymbol{Y}^k_c)\|_F}{\theta^{k+1}} \to \infty \ \text{ and } \frac{(d\boldsymbol{X}^k_c, d\boldsymbol{Y}^k_c)}{\|(d\boldsymbol{X}^k_c, d\boldsymbol{Y}^k_c)\|_F} \to (\boldsymbol{U}, \boldsymbol{V}) \neq (\boldsymbol{O}, \boldsymbol{O})$$

for some $(\boldsymbol{U}, \boldsymbol{V}) \in \mathcal{F}_0$. Now, dividing the identity (53) by $\|(d\boldsymbol{X}_c^k, d\boldsymbol{Y}_c^k)\|_F/\theta^{k+1}$ and taking its limit along the subsequence, we obtain that

$$\boldsymbol{X}^*\boldsymbol{V} + \boldsymbol{V}\boldsymbol{X}^* + \boldsymbol{U}\boldsymbol{Y}^* + \boldsymbol{Y}^*\boldsymbol{U} = \boldsymbol{O} \ \text{ and } (\boldsymbol{O}, \boldsymbol{O}) \neq (\boldsymbol{U}, \boldsymbol{V}) \in \mathcal{F}_0.$$

This contradicts Lemma 6.3. ∎

Proof of Theorem 6.2: In view of Lemma 6.4, we know that $\delta_c^k \to 0$ as $k \to \infty$, which implies (i) and (ii) of the theorem.

7. Concluding Remarks.

The admissible region $\{(\boldsymbol{X}, \boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma, \tau) \ : \ \tau > 0\}$ in which we confine iterates $(\boldsymbol{X}^k, \boldsymbol{Y}^k)$ $(k = 0, 1, 2, \dots)$ becomes larger as we take larger $\gamma < 1$. Taking the limit as $\gamma \to 1$, we have the largest admissible region

$$\bigcup_{0<\gamma<1} \{(\boldsymbol{X}, \boldsymbol{Y}) \in \widetilde{\mathcal{N}}(\gamma, \tau) \ : \ \tau > 0\} = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X} \succ \boldsymbol{O}\}.$$

It is easily seen that this set is contained in $\mathcal{S}_{++} \times \mathcal{S}_{++}$. But the converse relation

$$\mathcal{S}_{++} \times \mathcal{S}_{++} \subset \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X} \succ \boldsymbol{O}\}$$

is not true. For a counter example to this relation, see the paper [25].

We can employ a different admissible region. For every $\gamma \in [0, 1]$ and $\tau \geq 0$, define

$$\widehat{\mathcal{N}}(\gamma, \tau) = \left\{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \left\| \frac{\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X}}{2} - \tau\boldsymbol{I} \right\|_F \leq \gamma\tau \right\} \tag{54}$$

We can easily verify that

$$\{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \boldsymbol{X}\boldsymbol{Y} = \tau\boldsymbol{I}\} \subset \widehat{\mathcal{N}}(\gamma, \tau) \subset \widetilde{\mathcal{N}}(\gamma, \tau) \text{ for every } \gamma \in (0, 1) \text{ and } \tau > 0.$$

To employ $\widehat{\mathcal{N}}(\gamma, \tau)$ in stead of $\widetilde{\mathcal{N}}(\gamma, \tau)$, we need some modification in Algorithm 2.2. Among others, we need to replace δ_p^k in (12) and δ_c^k in (15) by

$$\delta_p^k = \frac{\|\boldsymbol{X}^k\boldsymbol{Y}^k\|_F}{\theta^k\mu^0} \text{ and } \delta_c^k = \frac{\|\boldsymbol{X}_c^k\boldsymbol{Y}_c^k\|_F}{\theta^{k+1}\mu^0},$$

respectively.

It is interesting to compare our (modified) algorithm with the predictor-corrector infeasible-interior-point algorithm given by Potra and Sheng [23, 24].

- Our modified algorithm uses the combination of the AHO search direction [2] and the neighborhood $\widehat{\mathcal{N}}(\gamma, \tau)$ (given in (54)) of $\{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ : \boldsymbol{X}\boldsymbol{Y} = \tau\boldsymbol{I}\}$ while Potra and Sheng's algorithm uses the combination of the KSH search direction [12] and the neighborhood

  $$\mathcal{N}(\gamma, \tau) = \{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ \ : \ \|\boldsymbol{X}^{1/2}\boldsymbol{Y}\boldsymbol{X}^{1/2} - \tau\boldsymbol{I}\|_F \leq \gamma\tau\}$$

  of $\{(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{S}_+ \times \mathcal{S}_+ : \boldsymbol{X}\boldsymbol{Y} = \tau\boldsymbol{I}\}$. Using Lemma 3.3 of Monteiro [18], we can prove that $\widehat{\mathcal{N}}(\gamma, \tau) \subset \mathcal{N}(\gamma, \tau)$ for every $\gamma \in (0, 1)$ and $\tau > 0$.

- Our global convergence theorem, Theorem 2.3 has not guaranteed the global convergence at any linear rate (see also Remark 4.1) while Potra and Sheng proved the global convergence at a linear rate for their algorithm (Theorems 3.6, 3.7 and 3.8 of [23]).
- We have proved the quadratic local convergence under the strict complementarity condition while their algorithm requires an additional restriction (the size of the central path neighborhood approaches zero.) to attain the superlinear convergence.

Acknowledgement

The authors would like to thank Prof. Paul Tseng for valuable comments to improve the paper.

References

[1] W. F. Alizadeh, "Interior point methods in semidefinite programming with application to combinatorial optimization," *SIAM Journal on Optimization* **5** (1995) 13–51.

[2] W. F. Alizadeh, J. -P. A. Haeberly and M. L. Overton, "Primal-dual interior-point methods for semidefinite programming," Working Paper, 1994.

[3] W. F. Alizadeh, J. -P. A. Haeberly and M. L. Overton, , "Complementarity and nondegeneracy in semidefinite programming," Working Paper, 1995.

[4] R. M. Freund, "Complexity of an algorithm for finding an approximate solution of a semidefinite program with no regularity assumption," Technical report OR 302-94, Operations Research Center, MIT, 1994.

[5] A. Graham, *Kronecker Products and Matrix Calculus with Applications,* (Ellis Horwood Ltd., West Sussex, England, 1981).

[6] J. -P. A. Haeberly, Private communication.

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

[8] M. Kojima, S. Kojima and S. Hara, "Linear algebra for semidefinite programming," Technical Report B-290, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, October 1994, Revised August 1995.

[9] M. Kojima, S. Mizuno and A. Yoshise, "A primal-dual interior point algorithm for linear programming," in *Progress in Mathematical Programming: Interior Point and Related Methods,* N. Megiddo, ed., (Springer Verlag, New York, 1989) 29–47.

[10] M. Kojima, S. Mizuno and A. Yoshise, "A polynomial-time algorithm for a class of linear complementary problems," *Mathematical Programming* **44** (1989) 1–26.

[11] M. Kojima, S. Mizuno and A. Yoshise, "An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems," *Mathematical Programming* **50** (1991) 331–342.

[12] M. Kojima, S. Shindoh and S. Hara, "Interior-point methods for the monotone semidefinite linear complementarity problems," Research Report #282, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, April 1994, Revised October 1995, to appear in *SIAM Journal on Optimization.*

[13] M. Kojima, M. Shida and S. Shindoh, "Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs," Research Report #306, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, December 1995.

[14] C-J. Lin and R. Saigal, "A predictor-corrector method for semi-definite linear programming," Working paper, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2117, October 1995.

[15] S. Mizuno, M. Kojima and M. J. Todd, "Infeasible-interior-point primal-dual potential-reduction algorithms for linear programming," *SIAM Journal on Optimization* **5** (1995) 52–67.

[16] S. Mizuno, M. J. Todd and Y. Ye, On Adaptive-step primal-dual interior-point algorithms for linear programming," *Mathematics of Operations Research* **18** (1993) 964–981.

[17] R.D.C. Monteiro and I. Adler, "Interior path-following primal-dual algorithm. Part I: Linear Programming", *Mathematical Programming* **44** (1989) 27-41.

[18] R.D.C. Monteiro, "Primal-dual path following algorithms for semidefinite programming," Working Paper, School Industrial and Systems Engineering, Georgia Tech., Atlanta, GA 30332, September 1995.

[19] Ju. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications* (SIAM, Philadelphia, 1993).

[20] Ju. E. Nesterov and M. J. Todd, "Self-scaled cones and interior-point methods in nonlinear programming," Working Paper, CORE, Catholic University of Louvain, Louvain-la-Neuve, Belgium, April 1994.

[21] Ju. E. Nesterov and M. J. Todd, "Primal-dual interior-point methods for self-scaled cones," Technical Report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853-3801, USA, 1995.

[22] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization*, (Prentice-Hall, Englewood, Cliffs, New Jersey, 1982).

[23] F.A. Potra and R. Sheng, "A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming," Department of Mathematics, University of Iowa, Iowa City, IA 52242, October 1995.

[24] F.A. Potra and R. Sheng, "Superlinear convergence of infeasible-interior-point algorithms for semidefinite programming," Department of Mathematics, University of Iowa, Iowa City, IA 52242, April 1996.

[25] M. Shida, S. Shindoh and M. Kojima, "Existence of search directions in interior-point-algorithms for the SDP and the monotone SDLCP," Research Report #310, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, January 1996.

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

[27] L. Vandenberghe and S. Boyd, " Semidefinite programming," Informations Systems Laboratory, Stanford University, 1994.

[28] Y. Zhang, "On the convergence of a class of infeasible interior-point algorithms for the horizontal linear complementarity problem," *SIAM Journal on Optimization* **4** (1994) 208–227.

[29] Y. Zhang, "On extending primal-dual interior-point algorithms from linear programming to semidefinite programming," Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland 21228-5398, October 1995.