Interior point method on semi-definite linear complementarity problems using the Nesterov–Todd (NT) search direction: polynomial complexity and local convergence

Chee-Khian Sim[1]

Abstract

We consider in this paper an infeasible predictor–corrector primal–dual path following interior point algorithm using the Nesterov–Todd search direction to solve semi-definite linear complementarity problems. Global convergence and polynomial iteration complexity of the algorithm are established. Two sufficient conditions are also given for superlinear convergence of iterates generated by the algorithm. Preliminary numerical results are finally provided when the algorithm is used to solve semi-definite linear complementarity problems.

Keywords Nesterov–Todd (NT) direction · Predictor–corrector primal–dual path following interior point algorithm · Semi-definite linear complementarity problem · Polynomial complexity · Local convergence

1 Introduction

The class of semi-definite linear complementarity problems (SDLCPs), which contains the class of semi-definite programs (SDPs) as an important subclass, has many real life applications, for example, in optimal control, estimation and signal processing, communications and networks, statistics, finance [3]. Semi-definite programming has wide applications in NP-hard combinatorial problems [8] and global optimization, where it is used to find bounds on optimal values and to find approximate solutions. Interior point methods have been proven to be successful in solving linear programs, with many works in the literature devoted to its study since the 80s. Semi-definite programs are extensions of linear programs to the space of symmetric matrices, and

✉ Chee-Khian Sim
   chee-khian.sim@port.ac.uk

1   School of Mathematics and Physics, University of Portsmouth, Lion Gate Building, Lion Terrace, Portsmouth PO1 3HF, UK

                    ⌂ Springer

interior point methods have been successfully extended from solving linear programs to solving semi-definite programs with the same polynomial complexity results—see for example [1,14,17,45].

Among different interior point methods (IPMs), primal–dual path following interior point algorithms are the most successful and most widely studied. Due to the difficulty in maintaining symmetry in the linearized complementarity when using primal–dual path following interior point method to solve an SDP, researchers working in the IPM domain have proposed ways to overcome this problem, resulting in different symmetrized search directions [1,11,14,17,18,20,21,42] being introduced. Among these search directions, the Alizadeh–Haeberly–Overton (AHO), Helmberg–Kojima–Monteiro (HKM) and Nesterov–Todd (NT) directions are better known, with the latter two being implemented in SDP solvers, such as SeDuMi, SDPT3. Various studies have been conducted to analyze primal–dual path following interior point algorithm, using the HKM search direction, such as [25], and the NT search direction, such as [16], to solve semi-definite programs. Works [19,27,30] have also been done to give an unified polynomial complexity analysis of interior point algorithms on semi-definite programs using a commutative class of search directions, which include the HKM and NT directions. More recently, a stream of research [10,29,43,44] has appeared that derives polynomial complexity for a full NT step interior point method in solving linear programs, semi-definite programs and symmetric cone programs. Recent works on interior point methods include [2,6,28] that design new interior point algorithms, different from that considered in this paper, to solve symmetric cone programs.

The focus of this paper is on analyzing an infeasible predictor–corrector primal–dual path following interior point algorithm, using the NT search direction, to solve semi-definite linear complementarity problems. We consider an infeasible interior point algorithm since it is more practical than a feasible interior point algorithm, as it is usually difficult to find an initial interior point iterate which is also feasible. The algorithm considered in this paper was studied in [25] to solve semi-definite programs using the HKM search direction. The analysis in [25] to show global convergence and polynomial complexity cannot be carried over in a straightforward manner to analyze the algorithm using the search direction considered in this paper. Our contributions in this paper include showing polynomial complexity $\mathcal{O}(n \ln(\max\{n\tau_0, \|r_0\|_2\}/\epsilon))$ for the infeasible interior point algorithm using the NT search direction to solve a semi-definite linear complementarity problem. Our result complements the result obtained in [25] which considers the algorithm using the HKM search direction to solve semi-definite programs. Our iteration complexity bound is the best known iteration bound for infeasible interior point algorithms using the "narrow" neighborhood known so far in the literature. To the best of our knowledge, it is also the first time this polynomial complexity is derived for an infeasible predictor–corrector primal–dual path following interior point algorithm, using the NT search direction, on a semi-definite linear complementarity problem. Furthermore, under strict complementarity assumption, we provide local convergence results in Sect. 4 that are analogous to that in [34]. Note that superlinear convergence results using interior point methods are hard to obtain, to quote the opening sentences of [22]: "Local superlinear convergence is a natural and very desirable property of many methods in nonlinear optimization. However, for interior-point methods the corresponding analysis is not trivial". It is worthwhile

mentioning that among these local convergence results, we show that for the important class of linear semi-definite feasibility problems, only a suitably chosen initial iterate is needed for superlinear convergence, unlike what is generally believed to be needed to achieve superlinear convergence, which is for iterates to get close to the central path by repeatedly solving the corrector-step linear system in an iteration (see for example [13,22]). We should also mention that although local convergence results using the NT search direction has been established in [16] by "narrowing" the central path neighborhood, the algorithm considered there generates feasible iterates, while here we consider an infeasible algorithm, which is more practical, and the analysis for the infeasible case is more complicated than that for the feasible case. Finally, our results in this paper indicate that using the NT search direction in an interior point algorithm to solve semi-definite linear complementarity problems is as good as using the HKM search direction both from the "polynomial complexity" and the "local convergence" point of view.

1.1 Facts, notations and terminology

The space of symmetric $n \times n$ matrices is denoted by S^n. The cone of positive semi-definite (resp., postive definite) symmetric matrices is denoted by S^n_+ (resp., S^n_{++}). The identity matrix is denoted by $I_{n \times n}$, where n stands for the size of the matrix. We omit the subscript when the size of the identity matrix is clear from the context. Given a symmetric matrix G, $\lambda_{\min}(G)$ and $\lambda_{\max}(G)$ are denoted to be the minimum and maximum eigenvalue of G respectively.

Given matrices G and K in $\Re^{n_1 \times n_2}$, the inner product, $G \bullet K$, between the two matrices is defined to be $G \bullet K := \mathrm{Tr}(G^T K) = \mathrm{Tr}(G K^T)$, where $\mathrm{Tr}(\cdot)$ is the trace of a square matrix. $\| \cdot \|_2$ for a vector in $\Re^n$ refers to its Euclidean norm, and for a matrix in $\Re^{n_1 \times n_2}$, it refers to its operator norm. On the other hand, $\|G\|_F := \sqrt{G \bullet G}$, for $G \in \Re^{n_1 \times n_2}$, refers to the Frobenius norm of G.

For a matrix $G \in \Re^{n_1 \times n_2}$, we denote its component in the ith row and the jth column by G_{ij}. $G_{i\cdot}$ denotes the ith row of G and $G_{\cdot j}$ the jth column of G. In the case when G is partitioned into blocks of submatrices, then G_{ij} refers to the submatrix in the corresponding (i, j) position.

Given square matrices $G_i \in \Re^{n_i \times n_i}, i = 1, \ldots, N$, $\mathrm{Diag}(G_1, \ldots, G_N)$ is a square matrix with $G_i, i = 1, \ldots, N$, as its main diagonal blocks arranged in accordance to the way they are lined up in $\mathrm{Diag}(G_1, \ldots, G_N)$. All the other entries in $\mathrm{Diag}(G_1, \ldots, G_N)$ are zeroes.

Given $X \in S^n$, $\mathrm{svec}(X)$ is defined to be

$$\mathrm{svec}(X) := (X_{11}, \sqrt{2}X_{21}, \ldots, \sqrt{2}X_{n1}, X_{22}, \sqrt{2}X_{32}, \ldots, \sqrt{2}X_{n2}, \ldots, X_{nn})^T \in \Re^{\tilde{n}},$$

where $\tilde{n} := n(n + 1)/2$. $\mathrm{svec}(\cdot)$ sets up a one-to-one correspondence between S^n and $\Re^{\tilde{n}}$.

Note that for all $X, Y \in S^n$, $X \bullet Y = \mathrm{svec}(X)^T \mathrm{svec}(Y)$. Hence, $\|X\|_F = \|\mathrm{svec}(X)\|_2$ for $X \in S^n$.

Given $G, K \in \Re^{n \times n}$, $G \otimes_s K$ is a square matrix of size $\tilde{n}$ defined by

$$(G \otimes_s K)\text{svec}(H) := \frac{1}{2}\text{svec}(KHG^T + GHK^T), \ \forall \ H \in S^n.$$

Fact 1 (Appendix of [39]) *Let $G, K, L \in \Re^{n \times n}$.*

(a) $G \otimes_s K = K \otimes_s G$ *and* $(G \otimes_s K)^T = G^T \otimes_s K^T$.
(b) $(G \otimes_s K)(L \otimes_s L) = (GL) \otimes_s (KL)$ *and* $(L \otimes_s L)(G \otimes_s K) = (LG) \otimes_s (LK)$.
(c) *If G and K are commuting symmetric matrices. Let $\{x_i\}$ be their common basis of eigenvectors with corresponding eigenvalues λ_i^G and λ_i^K. Then $G \otimes_s K$ is symmetric and has the set of eigenvalues given by $\left\{ \frac{1}{2}(\lambda_i^G \lambda_j^K + \lambda_j^G \lambda_i^K) \right\}$. Also, svec $\left(\frac{1}{2}(x_i x_j^T + x_j x_i^T) \right)$ is an eigenvector corresponding to the eigenvalue $\frac{1}{2}(\lambda_i^G \lambda_j^K + \lambda_j^G \lambda_i^K)$ of $G \otimes_s K$.*

Fact 2 ([12]) *For $G, K \in \Re^{n \times n}$, $\|GK\|_F \leq \min\{\|G\|_F \|K\|_2, \|G\|_2 \|K\|_F\}$.*

Fact 3 *For $x \in \Re, x \geq 0$, we have*

$$\frac{\sqrt{1+x} - 1}{\sqrt{1+x} + 1} \leq \frac{\sqrt{x}}{1 + \sqrt{x}}.$$

Given functions $f : \Omega \to E$ and $g : \Omega \to \Re_{++}$, where Ω is an arbitrary set and E is a normed vector space with norm $\| \cdot \|$. For a subset $\hat{\Omega} \subseteq \Omega$, we write $f(w) = \mathcal{O}(g(w))$ for all $w \in \hat{\Omega}$ to mean that $\|f(w)\| \leq Mg(w)$ for all $w \in \hat{\Omega}$, where $M > 0$ is a positive constant. Suppose $E = S^n$. Then we write $f(w) = \Theta(g(w))$ if for all $w \in \hat{\Omega}$, $f(w) \in S_{++}^n$, $f(w) = \mathcal{O}(g(w))$ and $f(w)^{-1} = \mathcal{O}(1/g(w))$. The subset $\hat{\Omega}$ should be clear from the context. For example, $\hat{\Omega} = (0, \hat{w})$ for some $\hat{w} > 0$ or $\hat{\Omega} = \{w_k \ ; \ k \geq 0\}$, where $w_k \to 0$ as $k \to \infty$. In the latter case, we write $f(w_k) = o(g(w_k))$ to mean that $\|f(w_k)\|/g(w_k) \to 0$, as $k \to \infty$.

2 A primal–dual path following interior point algorithm on an SDLCP

We consider a semi-definite linear complementarity problem (SDLCP) which is the problem to find a solution, (X, Y), to the following system:

$$XY = 0, \tag{1}$$
$$\mathcal{A}(X) + \mathcal{B}(Y) = q, \tag{2}$$
$$X, Y \in S_+^n, \tag{3}$$

where $q \in \Re^{\tilde{n}}$ and $\mathcal{A}, \mathcal{B} : S^n \to \Re^{\tilde{n}}$ are linear operators mapping S^n to the space $\Re^{\tilde{n}}$, $\tilde{n} := n(n+1)/2$. $\mathcal{A}$ and $\mathcal{B}$ take the form

$$\mathcal{A}(X) = (A_1 \bullet X, \ldots, A_{\tilde{n}} \bullet X)^T, \ \mathcal{B}(Y) = (B_1 \bullet Y, \ldots, B_{\tilde{n}} \bullet Y)^T, \tag{4}$$

where A_i, $B_i \in S^n$ for all $i = 1, \ldots, \tilde{n}$. We also called the system (1)–(3) an SDLCP.

The following assumptions are assumed to hold for the system (1)–(3) in this and the next section, while we replace Assumption 1(b) by Assumption 2 in Sect. 4, although we still assume Assumptions 1(a), (c) in that section.

Assumption 1 (a) System (1)–(3) is monotone. That is, $\mathcal{A}(X) + \mathcal{B}(Y) = 0$ for $X, Y \in S^n \Rightarrow X \bullet Y \geq 0$.

(b) There exists $(X^1, Y^1) \in S_{++}^n \times S_{++}^n$ such that $\mathcal{A}(X^1) + \mathcal{B}(Y^1) = q$.

(c) $\{\mathcal{A}(X) + \mathcal{B}(Y) \, ; \, X, Y \in S^n\} = \Re^{\tilde{n}}$.

The first assumption [Assumption 1(a)] is satisfied for the class of semi-definite programs (SDPs), with equality for $X \bullet Y$ instead of inequality. The second assumption ensures that (2) is satisfied for some positive definite matrix pair, while the last assumption is a technical assumption that can be satisfied for any SDP. Note that Assumption 1(b) is only used in this paper to ensure the existence of a solution to the SDLCP (1)–(3).

An SDP in its primal and dual form is given by

$$
\begin{aligned}
(\mathcal{P}) \quad \min \quad & C \bullet X \\
\text{subject to } & A_i \bullet X = b_i, \quad i = 1, \ldots, m, \\
& X \in S_+^n, \\
(\mathcal{D}) \quad \max \quad & \sum_{i=1}^m b_i y_i \\
\text{subject to } & \sum_{i=1}^m y_i A_i + Y = C, \\
& Y \in S_+^n.
\end{aligned}
$$

In the above formulation of an SDP, it is without loss of generality to assume that $A_i, i = 1, \ldots, m$, are linearly independent.

An SDP is a special case of an SDLCP by letting $A_i = 0$ for $i = m + 1, \ldots, \tilde{n}$, $B_i = 0$ for $i = 1, \ldots, m$, in (4). $B_i, i = m + 1, \ldots, \tilde{n}$, in (4) are chosen to be linearly independent and belong to the subspace in S^n orthogonal to the space spanned by $A_i, i = 1, \ldots, m$.

Primal–dual path following interior point algorithms can be used to solve an SDLCP. We consider an infeasible predictor–corrector primal–dual path following interior point algorithm, as found in [25,34], in this paper. In [25,34], the search direction used is the Helmberg–Kojima–Monteiro (HKM) search direction [11,14,17], while in this paper, we consider the algorithm using the Nesterov–Todd (NT) search direction [20, 21]. The difference between the two search directions is the way "symmetrization" is being done on (1). For $P \in \Re^{n \times n}$ an invertible matrix, the similarly transformed symmetrization operator $H_P(\cdot)$, introduced in [45], is given by

$$
H_P(U) := \frac{1}{2}(PUP^{-1} + (PUP^{-1})^T),
$$

where $U \in \Re^{n \times n}$. Hence, $H_P(\cdot)$ is a map from $\Re^{n \times n}$ to S^n. Different search directions correspond to different P, as will be explained later.

Infeasible primal–dual path following interior point algorithm works on the principle that iterates generated by the algorithm "follows" an (infeasible) central path

$(X(\mu), Y(\mu))$, $\mu > 0$, which is the unique solution to

$$XY = \mu I,$$
$$\mathcal{A}(X) + \mathcal{B}(Y) = q + \frac{\mu}{\mu_0} r_0,$$
$$X, Y \in S_{++}^n,$$

where

$$r_0 := \mathcal{A}(X_0) + \mathcal{B}(Y_0) - q,$$

for some $X_0, Y_0 \in S_{++}^n$ and $\mu_0 > 0$. Here, X_0, Y_0 are such that $X_0 Y_0 = \mu_0 I$. The existence and uniqueness of this central path follows from Theorem 2.3 in [35]. It also follows from Theorem 2.4 in [35] that there exists a solution (X^*, Y^*) to the SDLCP (1)–(3). Although these theorems apply to the feasible central path when $r_0 = 0$ in [35], they can be easily shown to hold for an infeasible central path when $r_0 \neq 0$ by assuming Assumptions 1(a)–(c). We leave their proofs as exercises for the reader.

From now onwards, (X^*, Y^*) denotes a solution to the SDLCP (1)–(3).

Using $H_P(\cdot)$, the above central path $(X(\mu), Y(\mu))$, $\mu > 0$, is also the unique solution to

$$H_P(XY) = \mu I,$$
$$\mathcal{A}(X) + \mathcal{B}(Y) = q + \frac{\mu}{\mu_0} r_0,$$
$$X, Y \in S_{++}^n,$$

since we have for $X, Y \in S_{++}^n$ and $\mu > 0$,

$$H_P(XY) = \mu I \Leftrightarrow XY = \mu I.$$

Below we describe the infeasible predictor–corrector primal–dual path following interior point algorithm considered in this paper. This algorithm is the same as that in [25], although there is a wider choice for β_1, β_2 here. The key to the algorithm is solving the following system of linear equations:

$$H_P(X \Delta Y + \Delta X Y) = \sigma \tau I - H_P(XY), \tag{5}$$
$$\mathcal{A}(\Delta X) + \mathcal{B}(\Delta Y) = -\bar{r}, \tag{6}$$

for $\Delta X, \Delta Y \in S^n$, where $\tau > 0, 0 \leq \sigma \leq 1$ and $\bar{r} \in \Re^{\tilde{n}}$. Different choice of τ, σ and $\bar{r}$ results in different step in Algorithm 1 described below.

The following (narrow) neighborhood of the central path is used in this paper:

$$\mathcal{N}_1(\beta, \tau) = \{(X, Y) \in S_{++}^n \times S_{++}^n ; \|H_P(XY) - \tau I\|_F \leq \beta \tau\}, \tag{7}$$

where $\tau > 0$ and $0 < \beta < 1$.

Note that P that appears in (5) and (7) is not chosen arbitrary, but is related to $X, Y \in S_{++}^n$ as we will see next, after describing the algorithm we are analyzing in this paper.

Algorithm 1 *Given $\epsilon > 0$. Choose $\beta_1 < \beta_2$, with $\beta_2^2/(2(1 - \beta_2)) \leq \beta_1 < \beta_2 < \beta_2/(1 - \beta_2) < 1$. Choose $(X_0, Y_0) \in \mathcal{N}_1(\beta_1, \tau_0)$ with $\mu_0 = \tau_0 = X_0 \bullet Y_0/n$. For $k = 0, 1, \ldots,$ do (a1) through (a3):*

(a1) *If $\max\{X_k \bullet Y_k, \|r_k\|_2\} \leq \epsilon$, where $r_k = \mathcal{A}(X_k) + \mathcal{B}(Y_k) - q$, then report (X_k, Y_k) as an approximate solution to the system (1)–(3), and terminate.*

(a2) *(Predictor Step) Find the solution $(\Delta X_k^p, \Delta Y_k^p)$ of the linear system (5), (6), with $X = X_k, Y = Y_k, P = P_k, \sigma = 0, \tau = \tau_k$ and $\bar{r} = r_k$. Define*

$$\hat{X}_k = X_k + \hat{\alpha}_k \Delta X_k^p, \quad \hat{Y}_k = Y_k + \hat{\alpha}_k \Delta Y_k^p,$$

where the steplength $\hat{\alpha}_k$ satisfies

$$\alpha_{k,1} \leq \hat{\alpha}_k \leq \alpha_{k,2}. \tag{8}$$

Here,

$$\alpha_{k,1} = \frac{2}{\sqrt{1 + 4\delta_k/(\beta_2 - \beta_1)} + 1}, \tag{9}$$

$$\delta_k = \frac{1}{\tau_k} \|H_{P_k}(\Delta X_k^p \Delta Y_k^p)\|_F, \tag{10}$$

and

$$\alpha_{k,2} = \max\{\tilde{\alpha} \in [0, 1] \; ; \; (X_k + \alpha \Delta X_k^p, Y_k + \alpha \Delta Y_k^p)$$
$$\in \mathcal{N}_1(\beta_2, (1 - \alpha)\tau_k) \; \forall \, \alpha \in [0, \tilde{\alpha}]\}. \tag{11}$$

If $\hat{\alpha}_k = 1$, then $(\hat{X}_k, \hat{Y}_k)$ solves the system (1)–(3) and terminate.

(a3) *(Corrector Step) Find the solution $(\Delta X_k^c, \Delta Y_k^c)$ of the linear system (5), (6), with $X = \hat{X}_k, Y = \hat{Y}_k, P = \hat{P}_k, \sigma = (1 - \hat{\alpha}_k), \tau = \tau_k$ and $\bar{r} = 0$. Set*

$$X_{k+1} = \hat{X}_k + \Delta X_k^c, \quad Y_{k+1} = \hat{Y}_k + \Delta Y_k^c,$$
$$\tau_{k+1} = (1 - \hat{\alpha}_k)\tau_k.$$

Set $k + 1 \to k$ and go to Step (a1).

The above algorithm is an infeasible predictor–corrector primal–dual path following interior point algorithm. P_k in the algorithm is chosen such that $P_k X_k Y_k P_k^{-1} \in S^n$. Examples of P_k that satisfy this are $P_k = Y_k^{1/2}$, and P_k such that $P_k^T P_k = W_k^{-1}$ with $W_k Y_k W_k = X_k$. The former corresponds to the dual HKM search direction, while the latter corresponds to the NT search direction. $\hat{P}_k$ in the algorithm is also chosen to satisfy $\hat{P}_k \hat{X}_k \hat{Y}_k \hat{P}_k^{-1} \in S^n$.

We remark that $(\Delta X_k^p, \Delta Y_k^p)$ and $(\Delta X_k^c, \Delta Y_k^c)$ in Algorithm 1 exist and are unique, by observing that the left hand side of (5) and (6) together when written in matrix-vector product has the matrix invertible, which holds because $(X_k, Y_k) \in \mathcal{N}_1(\beta_1, \tau_k)$ (by Proposition 4) and $(\hat{X}_k, \hat{Y}_k) \in \mathcal{N}_1(\beta_2, (1 - \hat{\alpha}_k)\tau_k)$, respectively. (We leave the details to show the existence and uniqueness of $(\Delta X_k^p, \Delta Y_k^p)$ and $(\Delta X_k^c, \Delta Y_k^c)$ to the reader.) Furthermore, in the above algorithm, we note that there is a wider range of choice for β_1 and β_2 compared with the algorithm in [25].

Let us make an observation on our choice of P_k, $\hat{P}_k$ in the following proposition:

Proposition 1 *Suppose $P \in \Re^{n \times n}$ is an invertible matrix with $PXYP^{-1} \in S^n$, where $X, Y \in S_{++}^n$. Then PXP^T and $P^{-T}YP^{-1}$ have a common set of eigenvectors with corresponding real positive eigenvalues λ_i^X and λ_i^Y, $i = 1, \ldots, n$, respectively. Also, $PXYP^{-1}$ has the same set of eigenvectors with corresponding eigenvalues $\lambda_i^X \lambda_i^Y$, $i = 1, \ldots, n$.*

Proof Since $PXYP^{-1} \in S^n$, this implies that PXP^T, $P^{-T}YP^{-1} \in S_{++}^n$ commute. Hence they share a common set of eigenvectors with corresponding real positive eigenvalues $\lambda_i^X, \lambda_i^Y, i = 1, \ldots, n$. Furthermore, it is easy to see that $PXYP^{-1} = (PXP^T)(P^{-T}YP^{-1})$ has the same set of eigenvectors with corresponding eigenvalues $\lambda_i^X \lambda_i^Y, i = 1, \ldots, n$. □

From now onwards, we consider Algorithm 1 using the NT search direction in Steps (a2) and (a3) of the algorithm, which means that P_k and $\hat{P}_k$ in these steps satisfy $P_k^T P_k = W_k^{-1}$ with $W_k Y_k W_k = X_k$, and $\hat{P}_k^T \hat{P}_k = \hat{W}_k^{-1}$ with $\hat{W}_k \hat{Y}_k \hat{W}_k = \hat{X}_k$ respectively. Proposition 1 then applies to P_k and $\hat{P}_k$ given in this way since they satisfy $P_k X_k Y_k P_k^{-1}$, $\hat{P}_k \hat{X}_k \hat{Y}_k \hat{P}_k^{-1} \in S^n$, respectively. Furthermore, there are different ways in which P_k and $\hat{P}_k$ can be chosen to form the NT search direction in these steps of Algorithm 1. In Sect. 4, using a particular choice of P_k, we establish two sufficient conditions for superlinear convergence using the algorithm on SDLCPs that satisfy the strict complementarity assumption.

We also require that P that appears in (7) to be related to X, Y in (7) by $P^T P = W^{-1}$ with $WYW = X$.

The following are properties satisfied by (X_k, Y_k) and $(\hat{X}_k, \hat{Y}_k)$ in Algorithm 1, which are useful in the analysis given in the paper on the convergence behavior of iterates generated by the algorithm.

Proposition 2 *Let $(X, Y) \in \mathcal{N}_1(\beta, \tau)$, where $\tau > 0, 0 < \beta < 1$, and P satisfies $P^T P = W^{-1}$ with $WYW = X$, then*

$$PXP^T = P^{-T}YP^{-1} \quad (12)$$

$$X \bullet Y \leq n(1 + \beta)\tau, \quad (1 - \beta)\tau \leq \lambda_{\min}(XY) \leq \lambda_{\max}(XY) \leq (1 + \beta)\tau, \quad (13)$$

$$\|PXP^T\|_F \leq \sqrt{(1 + \beta)n\tau}, \quad \|P^{-T}YP^{-1}\|_F \leq \sqrt{(1 + \beta)n\tau}, \quad (14)$$

$$\|PXP^T\|_2 \leq \sqrt{(1 + \beta)\tau}, \quad \|P^{-T}YP^{-1}\|_2 \leq \sqrt{(1 + \beta)\tau}, \quad (15)$$

$$\|(PXP^T)^{-1/2}\|_2 \leq \frac{1}{((1-\beta)\tau)^{1/4}}, \quad \|(P^{-T}YP^{-1})^{-1/2}\|_2 \leq \frac{1}{((1-\beta)\tau)^{1/4}}, \tag{16}$$

$$\|[I \otimes_s (PXP^T)]^{-1}\|_2 \leq \frac{1}{\sqrt{(1-\beta)\tau}}. \tag{17}$$

Proof The relation (12) follows immediately from $P^T P = W^{-1}$, $WYW = X$ and then taking inverses.

Since $(X, Y) \in \mathcal{N}_1(\beta, \tau)$, we have $X, Y \in S^n_{++}$ and

$$\|H_P(XY) - \tau I\|_F \leq \beta\tau.$$

By Proposition 1 and observing that $H_P(XY) = PXYP^{-1}$, we have using the notations in Proposition 1 that

$$\sqrt{\sum_{i=1}^n (\lambda_i^X \lambda_i^Y - \tau)^2} \leq \beta\tau.$$

Hence, for all $i = 1, \ldots, n$,

$$(1-\beta)\tau \leq \lambda_i^X \lambda_i^Y \leq (1+\beta)\tau, \tag{18}$$

therefore (13) follows.

By (12), $\lambda_i^X = \lambda_i^Y$ for all $i = 1, \ldots, n$. The second inequality in (18) then implies (14), (15). Also, from (18) and $\lambda_i^X = \lambda_i^Y$, we have

$$\frac{1}{\lambda_i^X} = \frac{1}{\lambda_i^Y} \leq \frac{1}{\sqrt{(1-\beta)\tau}}$$

for all $i = 1, \ldots, n$. Inequalities in (16) then follow.

Using Fact 1(c), $I \otimes_s (PXP^T)$ is symmetric and

$$\lambda_{\min}(I \otimes_s (PXP^T)) = \min_{i,j=1,\ldots,n} \frac{1}{2}(\lambda_i^X + \lambda_j^X) \geq \sqrt{(1-\beta)\tau}.$$

Therefore,

$$\|[I \otimes_s (PXP^T)]^{-1}\|_2 = \frac{1}{\lambda_{\min}(I \otimes_s (PXP^T))} \leq \frac{1}{\sqrt{(1-\beta)\tau}},$$

which shows (17). □

The following technical result leads to Proposition 4, which ensures that the set in (11) is nonempty.

Proposition 3 *Given $(X, Y) \in \mathcal{N}_1(\beta, \tau)$, where $\tau > 0$, $0 < \beta < 1$, and P satisfies $P^T P = W^{-1}$ with $WYW = X$. Suppose $U, V \in S^n$ are such that*

$$\mathcal{A}(U) + \mathcal{B}(V) = 0, \tag{19}$$

$$H_P(XV + UY) = R, \tag{20}$$

then

$$\|(P \otimes_s P)\mathrm{svec}(U)\|_2^2 + \|(P^{-T} \otimes_s P^{-T})\mathrm{svec}(V)\|_2^2 \leq \frac{1}{(1-\beta)\tau} \|\mathrm{svec}(R)\|_2^2.$$

Proof Relation (20) can be written as

$$(P \otimes_s (P^{-T}Y))\mathrm{svec}(U) + ((PX) \otimes_s P^{-T})\mathrm{svec}(V) = \mathrm{svec}(R).$$

The latter can in turn be expressed as

$$[I \otimes_s (P^{-T}YP^{-1})](P \otimes_s P)\mathrm{svec}(U) + [I \otimes_s (PXP^T)](P^{-T} \otimes_s P^{-T})\mathrm{svec}(V)$$
$$= \mathrm{svec}(R), \tag{21}$$

by Fact 1(b).
Since P satisfies $P^T P = W^{-1}$ with $WYW = X$, by (12), we have $PXP^T = P^{-T}YP^{-1}$. Hence, from (21), taking the inverse of $I \otimes_s (PXP^T)$, we get

$$(P \otimes_s P)\mathrm{svec}(U) + (P^{-T} \otimes_s P^{-T})\mathrm{svec}(V) = [I \otimes_s (PXP^T)]^{-1}\mathrm{svec}(R). \tag{22}$$

Hence,

$$\|(P \otimes_s P)\mathrm{svec}(U)\|_2^2 + \|(P^{-T} \otimes_s P^{-T})\mathrm{svec}(V)\|_2^2$$
$$\leq \|(P \otimes_s P)\mathrm{svec}(U)\|_2^2 + 2\mathrm{svec}(U)^T \mathrm{svec}(V) + \|(P^{-T} \otimes_s P^{-T})\mathrm{svec}(V)\|_2^2$$
$$= \|(P \otimes_s P)\mathrm{svec}(U) + (P^{-T} \otimes_s P^{-T})\mathrm{svec}(V)\|_2^2$$
$$= \|[I \otimes_s (PXP^T)]^{-1}\mathrm{svec}(R)\|_2^2$$
$$\leq \|[I \otimes_s (PXP^T)]^{-1}\|_2^2 \|\mathrm{svec}(R)\|_2^2$$
$$\leq \frac{1}{(1-\beta)\tau} \|\mathrm{svec}(R)\|_2^2,$$

where the first inequality follows from (19) and Assumption 1(a), the first equality follows from (22), and the last inequality follows from Proposition 2. □

The following result shows that iterates (X_k, Y_k) generated by Algorithm 1 always belong to the narrow neighborhood of the central path (7).

Proposition 4 *For all $k \geq 0$, $(X_k, Y_k) \in \mathcal{N}_1(\beta_1, \tau_k)$.*

Proof We prove the proposition by induction. It is easy to see that $(X_0, Y_0) \in \mathcal{N}_1(\beta_1, \tau_0)$. Hence, the proposition holds for $k = 0$. Suppose the proposition holds for $k = k_0$, where $k_0 \geq 0$, that is, $(X_{k_0}, Y_{k_0}) \in \mathcal{N}_1(\beta_1, \tau_{k_0})$. We wish to show that

$$\|H_{P_{k_0+1}}(X_{k_0+1}Y_{k_0+1}) - \tau_{k_0+1}I\|_F \leq \beta_1 \tau_{k_0+1},$$

which then prove the proposition by induction. First observe that

$$
\begin{aligned}
&\| H_{P_{k_0+1}}(X_{k_0+1} Y_{k_0+1}) - \tau_{k_0+1} I \|_F \\
&= \| P_{k_0+1} X_{k_0+1} Y_{k_0+1} P_{k_0+1}^{-1} - \tau_{k_0+1} I \|_F \\
&= \| P_{k_0+1} [X_{k_0+1} Y_{k_0+1} - \tau_{k_0+1} I] P_{k_0+1}^{-1} \|_F \\
&\leq \| H_{\hat{P}_{k_0} P_{k_0+1}^{-1}} [P_{k_0+1}[X_{k_0+1} Y_{k_0+1} - \tau_{k_0+1} I] P_{k_0+1}^{-1}] \|_F \\
&= \| H_{\hat{P}_{k_0}}(X_{k_0+1} Y_{k_0+1}) - \tau_{k_0+1} I \|_F,
\end{aligned}
\tag{23}
$$

where the first equality follows since $P_{k_0+1} X_{k_0+1} Y_{k_0+1} P_{k_0+1}^{-1} \in S^n$, and the inequality follows from Lemma 2.2 in [25], again using $P_{k_0+1} X_{k_0+1} Y_{k_0+1} P_{k_0+1}^{-1} \in S^n$. Next, observe that

$$
\begin{aligned}
&H_{\hat{P}_{k_0}}(X_{k_0+1} Y_{k_0+1}) - \tau_{k_0+1} I \\
&= H_{\hat{P}_{k_0}}((\hat{X}_{k_0} + \Delta X_{k_0}^c)(\hat{Y}_{k_0} + \Delta Y_{k_0}^c)) - \tau_{k_0+1} I \\
&= H_{\hat{P}_{k_0}}(\hat{X}_{k_0} \hat{Y}_{k_0}) + H_{\hat{P}_{k_0}}(\hat{X}_{k_0} \Delta Y_{k_0}^c + \Delta X_{k_0}^c \hat{Y}_{k_0}) + H_{\hat{P}_{k_0}}(\Delta X_{k_0}^c \Delta Y_{k_0}^c) - \tau_{k_0+1} I \\
&= H_{\hat{P}_{k_0}}(\Delta X_{k_0}^c \Delta Y_{k_0}^c),
\end{aligned}
$$

where the third equality holds since $(\Delta X_{k_0}^c, \Delta Y_{k_0}^c)$ is the solution to the linear system (5), (6), in which $X = \hat{X}_{k_0}, Y = \hat{Y}_{k_0}, P = \hat{P}_{k_0}, \sigma = 1, \tau = \tau_{k_0+1}$ and $\bar{r} = 0$. Hence,

$$
\begin{aligned}
&\| H_{\hat{P}_{k_0}}(X_{k_0+1} Y_{k_0+1}) - \tau_{k_0+1} I \|_F \\
&= \| H_{\hat{P}_{k_0}}(\Delta X_{k_0}^c \Delta Y_{k_0}^c) \|_F \\
&\leq \| \hat{P}_{k_0} \Delta X_{k_0}^c \Delta Y_{k_0}^c \hat{P}_{k_0}^{-1} \|_F \\
&\leq \| \hat{P}_{k_0} \Delta X_{k_0}^c \hat{P}_{k_0}^T \|_F \| \hat{P}_{k_0}^{-T} \Delta Y_{k_0}^c \hat{P}_{k_0}^{-1} \|_F \\
&= \| \mathrm{svec}(\hat{P}_{k_0} \Delta X_{k_0}^c \hat{P}_{k_0}^T) \|_2 \| \mathrm{svec}(\hat{P}_{k_0}^{-T} \Delta Y_{k_0}^c \hat{P}_{k_0}^{-1}) \|_2 \\
&= \| (\hat{P}_{k_0} \otimes_s \hat{P}_{k_0}) \mathrm{svec}(\Delta X_{k_0}^c) \|_2 \| (\hat{P}_{k_0}^{-T} \otimes_s \hat{P}_{k_0}^{-T}) \mathrm{svec}(\Delta Y_{k_0}^c) \|_2 \\
&\leq \frac{1}{2} [\| (\hat{P}_{k_0} \otimes_s \hat{P}_{k_0}) \mathrm{svec}(\Delta X_{k_0}^c) \|_2^2 + \| (\hat{P}_{k_0}^{-T} \otimes_s \hat{P}_{k_0}^{-T}) \mathrm{svec}(\Delta Y_{k_0}^c) \|_2^2],
\end{aligned}
\tag{24}
$$

where the second inequality follows from Fact 2 and $\|U\|_2 \leq \|U\|_F$ for $U \in S^n$. Since $(\Delta X_{k_0}^c, \Delta Y_{k_0}^c)$ is the solution to the linear system (5), (6), in which $X = \hat{X}_{k_0}, Y = \hat{Y}_{k_0}, P = \hat{P}_{k_0}, \sigma = 1, \tau = \tau_{k_0+1}$ and $\bar{r} = 0$, where $(\hat{X}_{k_0}, \hat{Y}_{k_0}) \in \mathcal{N}_1(\beta_2, \tau_{k_0+1})$, it follows from Proposition 3 that

$$
\begin{aligned}
&\| (\hat{P}_{k_0} \otimes_s \hat{P}_{k_0}) \mathrm{svec}(\Delta X_{k_0}^c) \|_2^2 + \| (\hat{P}_{k_0}^{-T} \otimes_s \hat{P}_{k_0}^{-T}) \mathrm{svec}(\Delta Y_{k_0}^c) \|_2^2 \\
&\leq \frac{1}{(1-\beta_2)\tau_{k_0+1}} \| \mathrm{svec}(H_{\hat{P}_{k_0}}(\hat{X}_{k_0} \hat{Y}_{k_0}) - \tau_{k_0+1} I) \|_2^2.
\end{aligned}
$$

Therefore, by the above inequality, (23) and (24) leads to

$$\|H_{P_{k_0+1}}(X_{k_0+1}Y_{k_0+1}) - \tau_{k_0+1}I\|_F$$

$$\leq \frac{1}{2(1-\beta_2)\tau_{k_0+1}}\|\text{svec}(H_{\hat{P}_{k_0}}(\hat{X}_{k_0}\hat{Y}_{k_0}) - \tau_{k_0+1}I)\|_2^2$$

$$\leq \frac{\beta_2^2}{2(1-\beta_2)}\tau_{k_0+1},$$

where the last inequality holds since $(\hat{X}_{k_0}, \hat{Y}_{k_0}) \in \mathcal{N}_1(\beta_2, \tau_{k_0+1})$.
Since $\beta_1, \beta_2 \in (0,1)$ are chosen such that $\beta_2^2/(2(1-\beta_2)) \leq \beta_1$, we have from above, $(X_{k_0+1}, Y_{k_0+1}) \in \mathcal{N}_1(\beta_1, \tau_{k_0+1})$. By induction, the proposition is proved. □

As mentioned earlier, the above proposition ensures that $\alpha_{k,2}$ given by (11) is meaningful. The following result allows us to say more about $\alpha_{k,2}$ and also shows that we can always find $\hat{\alpha}_k$ that satisfies (8) in Algorithm 1.

Proposition 5 *We have $\alpha_{k,1} \leq \alpha_{k,2}$.*

Proof The proposition is proved by showing that for all $0 \leq \alpha \leq \alpha_{k,1}$,

$$(X_k + \alpha\Delta X_k^p, Y_k + \alpha\Delta Y_k^p) \in \mathcal{N}_1(\beta_2, (1-\alpha)\tau_k).$$

We have for all $0 \leq \alpha \leq \alpha_{k,1}$

$$\|H_{P_k}((X_k + \alpha\Delta X_k^p)(Y_k + \alpha\Delta Y_k^p)) - (1-\alpha)\tau_k I\|_F$$

$$= \|H_{P_k}(X_kY_k) + \alpha H_{P_k}(X_k\Delta Y_k^p + \Delta X_k^p Y_k) + \alpha^2 H_{P_k}(\Delta X_k^p \Delta Y_k^p) - (1-\alpha)\tau_k I\|_F$$

$$= \|(1-\alpha)[H_{P_k}(X_kY_k) - \tau_k I] + \alpha^2 H_{P_k}(\Delta X_k^p \Delta Y_k^p)\|_F$$

$$\leq (1-\alpha)\|H_{P_k}(X_kY_k) - \tau_k I\|_F + \alpha^2\|H_{P_k}(\Delta X_k^p \Delta Y_k^p)\|_F$$

$$\leq (1-\alpha)\beta_1\tau_k + \alpha^2\|H_{P_k}(\Delta X_k^p \Delta Y_k^p)\|_F$$

$$\leq \beta_2(1-\alpha)\tau_k,$$

where the second equality holds as $(\Delta X_k^p, \Delta Y_k^p)$ satisfies the linear system (5), (6), with $X = X_k, Y = Y_k, P = P_k, \sigma = 0, \tau = \tau_k$ and $\bar{r} = r_k$, the second inequality holds as $(X_k, Y_k) \in \mathcal{N}_1(\beta_1, \tau_k)$, and the last inequality follows from (9), (10). Hence, since $X_k, Y_k \in S_{++}^n$ and $\beta_2 < 1$, we see that if $0 < \alpha \leq \alpha_{k,1} < 1$, $X_k + \alpha\Delta X_k^p, Y_k + \alpha\Delta Y_k^p \in S_{++}^n$.
Now, since $X_k + \alpha\Delta X_k^p, Y_k + \alpha\Delta Y_k^p \in S_{++}^n$, the following holds.

$$\|H_{P_k(\alpha)}((X_k + \alpha\Delta X_k^p)(Y_k + \Delta Y_k^p)) - (1-\alpha)\tau_k I\|_F$$

$$= \|P_k(\alpha)[(X_k + \alpha\Delta X_k^p)(Y_k + \alpha\Delta Y_k^p) - (1-\alpha)\tau_k I]P_k(\alpha)^{-1}\|_F$$

$$\leq \|H_{P_kP_k(\alpha)^{-1}}(P_k(\alpha)[(X_k + \alpha\Delta X_k^p)(Y_k + \alpha\Delta Y_k^p) - (1-\alpha)\tau_k I]P_k(\alpha)^{-1})\|_F$$

$$= \|H_{P_k}((X_k + \alpha\Delta X_k^p)(Y_k + \alpha\Delta Y_k^p)) - (1-\alpha)\tau_k I\|_F,$$

where $P_k(\alpha)$ is such that $P_k(\alpha)^T P_k(\alpha) = W_k(\alpha)^{-1}$ with $W_k(\alpha)(Y_k + \alpha \Delta Y_k^p)W_k = X_k + \alpha \Delta X_k^p$ and hence $P_k(\alpha)(X_k + \alpha \Delta X_k^p)(Y_k + \alpha \Delta Y_k^p)P_k(\alpha)^{-1} \in S^n$, so that the above inequality holds by Lemma 2.2 in [25].

Putting everything together, if $0 \leq \alpha \leq \alpha_{k,1}$, we have $(X_k + \alpha \Delta X_k^p, Y_k + \alpha \Delta Y_k^p) \in \mathcal{N}_1(\beta_2, (1 - \alpha)\tau_k)$, as required. □

We remark that the above proposition implies that $\alpha_{k,2}$ given by (11) is always positive.

3 Global convergence and polynomial complexity of interior point algorithm

In this section, we show global convergence and iteration complexity results for iterates $\{(X_k, Y_k)\}$ generated by Algorithm 1, by considering the "duality gap", $\mu_k := X_k \bullet Y_k/n$, and the "feasibility gap", $r_k = \mathcal{A}(X_k) + \mathcal{B}(Y_k) - q$. The following proposition relates $\hat{\alpha}_k$, τ_k and (X_k, Y_k) generated by Algorithm 1 with μ_k and r_k. First, we have the following definition.

Definition 1 Define for $k \geq 0$, $\psi_k = \Pi_{j=0}^k (1 - \hat{\alpha}_j)$. Also, define $\psi_{-1} = 1$.

Proposition 6 For all $k \geq 0$, $(1 - \beta_1)\tau_k \leq \mu_k \leq (1 + \beta_1)\tau_k$ and $\mathcal{A}(X_k) + \mathcal{B}(Y_k) - q = r_k = \psi_{k-1}r_0$.

Proof By Proposition 4, for all $k \geq 0$, $(X_k, Y_k) \in \mathcal{N}_1(\beta_1, \tau_k)$. Hence,

$$\|H_{P_k}(X_k Y_k) - \tau_k I\|_F \leq \beta_1 \tau_k.$$

It then follows from Proposition 1 that

$$\sqrt{\sum_{i=0}^n (\lambda_i^{X_k} \lambda_i^{Y_k} - \tau_k)^2} \leq \beta_1 \tau_k,$$

where $\lambda_i^{X_k} \lambda_i^{Y_k}$, $i = 1, \ldots, n$ are the eigenvalues of $P_k X_k Y_k P_k^{-1}$. Therefore, for all $i = 1, \ldots, n$,

$$(1 - \beta_1)\tau_k \leq \lambda_i^{X_k} \lambda_i^{Y_k} \leq (1 + \beta_1)\tau_k.$$

The first result in the proposition then follows by noting that $X_k \bullet Y_k = \sum_{i=1}^n \lambda_i^{X_k} \lambda_i^{Y_k}$. Next, we show that

$$\mathcal{A}(X_k) + \mathcal{B}(Y_k) = q + \psi_{k-1}r_0 \tag{25}$$

by induction on $k \geq 0$.

Equality in (25) holds for $k = 0$. Suppose (25) holds for $k = k_0$ for some $k_0 \geq 0$. Then

$$
\begin{aligned}
\mathcal{A}(X_{k_0+1}) &+ \mathcal{B}(Y_{k_0+1}) \\
&= \mathcal{A}(\hat{X}_{k_0}) + \mathcal{B}(\hat{Y}_{k_0}) \\
&= \mathcal{A}(X_{k_0} + \hat{\alpha}_{k_0} \Delta X_{k_0}^p) + \mathcal{B}(Y_{k_0} + \hat{\alpha}_{k_0} \Delta Y_{k_0}^p) \\
&= \mathcal{A}(X_{k_0}) + \mathcal{B}(Y_{k_0}) + \hat{\alpha}_{k_0}(q - \mathcal{A}(X_{k_0}) - \mathcal{B}(Y_{k_0})) \\
&= \hat{\alpha}_{k_0} q + (1 - \hat{\alpha}_{k_0})(\mathcal{A}(X_{k_0}) + \mathcal{B}(Y_{k_0})) \\
&= \hat{\alpha}_{k_0} q + (1 - \hat{\alpha}_{k_0})(q + \psi_{k_0-1} r_0) \\
&= q + \psi_{k_0} r_0,
\end{aligned}
$$

where the first equality follows from $X_{k_0+1} = \hat{X}_{k_0} + \Delta X_{k_0}^c$, $Y_{k_0+1} = \hat{Y}_{k_0} + \Delta Y_{k_0}^c$ and $(\Delta X_{k_0}^c, \Delta Y_{k_0}^c)$ satisfying (6) with $\bar{r} = 0$, the third equality follows from $(\Delta X_{k_0}^p, \Delta Y_{k_0}^p)$ satistying (6) with $\bar{r} = r_{k_0}$, and the fifth equality follows by induction hypothesis. Hence, (25) holds for $k = k_0 + 1$, and by induction, (25) holds for all $k \geq 0$. □

To show global convergence and iteration complexity results using μ_k and r_k, we only need to investigate the behavior of ψ_{k-1}, since $\tau_k = \psi_{k-1}\tau_0$ and $\mu_k \leq (1+\beta_1)\tau_k$, and $r_k = \psi_{k-1} r_0$. By definition of ψ_k in Definition 1, this is achieved by analyzing $\hat{\alpha}_j$. We consider $\alpha_{j,1}$ instead, which is given by (9) since it is a lower bound to $\hat{\alpha}_j$. Since $\alpha_{j,1}$ given by (9) is expressed in terms of δ_j, to analyze $\alpha_{j,1}$, we only need to analyze δ_j. We have the following upper bound on δ_j.

Lemma 1 *For all $k \geq 0$, we have*

$$
\delta_k \leq L_x L_y,
$$

where

$$
L_x = \frac{1}{\sqrt{1 - \beta_1}} \left[\beta_1 + \sqrt{n} + (2 + \beta_1 + \varsigma)n \left(\varsigma_x + \frac{\sqrt{1 + \beta_1}}{\sqrt{1 - \beta_1}} (\varsigma_x + \varsigma_y) \right) \right], \quad (26)
$$

$$
L_y = \frac{1}{\sqrt{1 - \beta_1}} \left[\beta_1 + \sqrt{n} + (2 + \beta_1 + \varsigma)n \left(\varsigma_y + \frac{\sqrt{1 + \beta_1}}{\sqrt{1 - \beta_1}} (\varsigma_x + \varsigma_y) \right) \right]. \quad (27)
$$

Here,

$$
\begin{aligned}
\varsigma &:= \frac{X_0 \bullet Y^* + X^* \bullet Y_0}{X_0 \bullet Y_0}, \\
\varsigma_x &:= 1 + \| P_0 X^* P_0^T \|_F \| (P_0 X_0 P_0^T)^{-1/2} \|_2^2, \\
\varsigma_y &:= 1 + \| P_0^{-T} Y^* P_0^{-1} \|_F \| (P_0^{-T} Y_0 P_0^{-1})^{-1/2} \|_2^2.
\end{aligned}
$$

We prove the above lemma in Sect. 3.1, as its proof is quite involved.

Based on Lemma 1, we have the following global convergence theorem using Algorithm 1.

Theorem 1 *Given $(X_0, S_0) \in \mathcal{N}_1(\beta_1, \tau_0)$, we have $\mu_k \to 0$ and $r_k \to 0$ as $k \to \infty$, and hence any accumulation point of the sequence $\{(X_k, Y_k)\}$ generated by Algorithm 1 is a solution to the SDLCP (1)–(3).*

Proof Since $(X_0, S_0) \in S^n_{++}$ and for any (X^*, Y^*), a solution to the SDLCP (1)–(3), it is easy to see that L_x and L_y given by (26), (27) respectively are positive constants, since ς, ς_x and ς_y are constants. Therefore, by the relation between δ_j and L_x, L_y in Lemma 1, δ_j is bounded above by a positive constant (that depends on X_0, Y_0, X^*, Y^*), say, L, independent of $j \geq 0$. From (9), we therefore have

$$\alpha_{j,1} \geq \frac{2}{\sqrt{1 + 4L/(\beta_2 - \beta_1)} + 1},$$

for all $j \geq 0$. Since

$$\begin{aligned}
\psi_k &= \Pi_{j=0}^k (1 - \hat{\alpha}_j) \\
&\leq \Pi_{j=0}^k (1 - \alpha_{j,1}) \\
&\leq \left(1 - \frac{2}{\sqrt{1 + 4L/(\beta_2 - \beta_1)} + 1}\right)^{k+1},
\end{aligned}$$

ψ_k tends to zero as $k \to \infty$.

Therefore, the theorem is proved by applying Proposition 6. $\qquad\square$

Let us now state an iteration complexity result using Algorithm 1.

Theorem 2 *Let $(X_0, Y_0) \in \mathcal{N}_1(\beta_1, \tau_0)$ in Algorithm 1 be chosen such that there exists (X^*, Y^*) a solution to the SDLCP (1)–(3) with $\max\{\|P_0 X^* P_0^T\|_F, \|P_0^{-T} Y^* P_0^{-1}\|_F\} \leq \sqrt{\tau_0}$. Then given $\epsilon > 0$, if Algorithm 1 does not stop at Step (a2) before the kth iteration with a solution to the SDLCP (1)–(3), it stops at the kth iteration in Step (a1), where $k = \mathcal{O}(n \ln(\max\{n\tau_0, \|r_0\|_2\}/\epsilon))$, with an ϵ-approximate solution to the SDLCP (1)–(3).*

Proof We are given (X_0, Y_0) in $\mathcal{N}_1(\beta_1, \tau_0) \subset S^n_{++} \times S^n_{++}$, and (X^*, Y^*) in $S^n_+ \times S^n_+$ such that $\max\{\|P_0 X^* P_0^T\|_F, \|P_0^{-T} Y^* P_0^{-1}\|_F\} \leq \sqrt{\tau_0}$.
Observe that

$$\begin{aligned}
\varsigma &= \frac{X_0 \bullet Y^* + X^* \bullet Y_0}{X_0 \bullet Y_0} \\
&\leq \frac{\|P_0 X_0 P_0^T\|_F \|P_0^{-T} Y^* P_0^{-1}\|_F + \|P_0 X^* P_0^T\|_F \|P_0^{-T} Y_0 P_0^{-1}\|_F}{X_0 \bullet Y_0} \\
&\leq \frac{2\tau_0 \sqrt{(1 + \beta_1)n}}{n\tau_0} \\
&= 2\sqrt{\frac{1 + \beta_1}{n}},
\end{aligned}$$

where the second inequality holds by (14) (since $(X_0, Y_0) \in \mathcal{N}_1(\beta_1, \tau_0)$), $\max\{\|P_0 X^* P_0^T\|_F, \|P_0^{-T} Y^* P_0^{-1}\|_F\} \leq \sqrt{\tau_0}$ and $X_0 \bullet Y_0 = n\tau_0$.

Also,

$$\varsigma_x = 1 + \|P_0 X^* P_0^T\|_F \|(P_0 X_0 P_0^T)^{-1/2}\|_2^2 \leq 1 + \frac{1}{\sqrt{1 - \beta_1}},$$

$$\varsigma_y = 1 + \|P_0^{-T} Y^* P_0^{-1}\|_F \|(P_0^{-T} Y_0 P_0^{-1})^{-1/2}\|_2^2 \leq 1 + \frac{1}{\sqrt{1 - \beta_1}},$$

using (16) and $\max\{\|P_0 X^* P_0^T\|_F, \|P_0^{-T} Y^* P_0^{-1}\|_F\} \leq \sqrt{\tau_0}$.

Therefore, L_x, L_y given by (26), (27) respectively are less than or equal to $L_0 n$, where L_0 is a large enough number that depends only on β_1. Hence, by Lemma 1, (8) and (9),

$$\hat{\alpha}_j \geq \frac{2}{\sqrt{1 + 4L_0^2 n^2/(\beta_2 - \beta_1)} + 1}$$

for all $j \geq 0$.

Therefore,

$$\psi_{k-1} = \Pi_{j=0}^{k-1}(1 - \hat{\alpha}_j)$$

$$\leq \left(1 - \frac{2}{\sqrt{1 + 4L_0^2 n^2/(\beta_2 - \beta_1)} + 1}\right)^k$$

$$= \left(\frac{\sqrt{1 + 4L_0^2 n^2/(\beta_2 - \beta_1)} - 1}{\sqrt{1 + 4L_0^2 n^2/(\beta_2 - \beta_1)} + 1}\right)^k$$

$$\leq \left(\frac{2L_0 n}{\sqrt{\beta_2 - \beta_1} + 2L_0 n}\right)^k,$$

where the last inequality holds by Fact 3.

Since $\max\{X_k \bullet Y_k, \|r_k\|_2\} \leq \epsilon$ is needed for the algorithm to terminate at Step (a1), where $X_k \bullet Y_k/n = \mu_k \leq (1 + \beta_1)\tau_k = (1 + \beta_1)\psi_{k-1}\tau_0$ and $r_k = \psi_{k-1}r_0$, a sufficient condition for termination at Step (a1) at the kth iteration is when k satisfies

$$\left(\frac{2L_0 n}{\sqrt{\beta_2 - \beta_1} + 2L_0 n}\right)^k \max\{(1 + \beta_1)n\tau_0, \|r_0\|_2\} \leq \epsilon.$$

That is,

$$k \geq \ln\left(\frac{\max\{(1 + \beta_1)n\tau_0, \|r_0\|_2\}}{\epsilon}\right) \Big/ \ln\left(\frac{\sqrt{\beta_2 - \beta_1} + 2L_0 n}{2L_0 n}\right)$$

and the result then follows. □

3.1 Proof of Lemma 1

Note that

$$
\begin{aligned}
\delta_k &= \frac{1}{\tau_k} \| H_{P_k}(\Delta X_k^p \Delta Y_k^p) \|_F \\
&\le \frac{1}{\tau_k} \min\{ \| P_k \Delta X_k^p P_k^T \|_2 \| P_k^{-T} \Delta Y_k^p P_k^{-1} \|_F, \| P_k \Delta X_k^p P_k^T \|_F \| P_k^{-T} \Delta Y_k^p P_k^{-1} \|_2 \} \\
&\le \frac{1}{\tau_k} \| P_k \Delta X_k^p P_k^T \|_F \| P_k^{-T} \Delta Y_k^p P_k^{-1} \|_F,
\end{aligned}
\tag{28}
$$

where the first inequality holds by Fact 2. To show Lemma 1, we analyze $\| P_k \Delta X_k^p P_k^T \|_F$ and $\| P_k^{-T} \Delta Y_k^p P_k^{-1} \|_F$ that appear in (28) further by bounding them from above as given in the following proposition.

Proposition 7 *We have*

$$
\| P_k \Delta X_k^p P_k^T \|_F \le t_x + \frac{t}{\sqrt{(1 - \beta_1)\tau_k}},
\tag{29}
$$

$$
\| P_k^{-T} \Delta Y_k^p P_k^{-1} \|_F \le t_y + \frac{t}{\sqrt{(1 - \beta_1)\tau_k}},
\tag{30}
$$

where

$$
\begin{aligned}
t_x &:= \psi_{k-1} \| (P_k \otimes_s P_k) \mathrm{svec}(X_0 - X^*) \|_2, \\
t_y &:= \psi_{k-1} \| (P_k^{-T} \otimes_s P_k^{-T}) \mathrm{svec}(Y_0 - Y^*) \|_2, \\
t &:= \| \mathrm{svec}(-H_{P_k}(X_k Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k)) \|_2.
\end{aligned}
$$

Proof We first observe that $(\Delta X_k^p, \Delta Y_k^p)$ satisfies

$$
\mathcal{A}(\Delta X_k^p) + \mathcal{B}(\Delta Y_k^p) = q - \mathcal{A}(X_k) - \mathcal{B}(Y_k).
\tag{31}
$$

Now, by Proposition 6,

$$
\mathcal{A}(X_k) + \mathcal{B}(Y_k) = (1 - \psi_{k-1})q + \psi_{k-1}(\mathcal{A}(X_0) + \mathcal{B}(Y_0)),
$$

and with $\mathcal{A}(X^*) + \mathcal{B}(Y^*) = q$, from (31), we obtain

$$
\mathcal{A}(\Delta X_k^p) + \mathcal{B}(\Delta Y_k^p) = \psi_{k-1}(\mathcal{A}(X^* - X_0) + \mathcal{B}(Y^* - Y_0)).
$$

That is,

$$
\mathcal{A}(\Delta X_k^p + \psi_{k-1}(X_0 - X^*)) + \mathcal{B}(\Delta Y_k^p + \psi_{k-1}(Y_0 - Y^*)) = 0.
\tag{32}
$$

On the other hand,

$$
\begin{aligned}
H_{P_k}(X_k(\Delta Y_k^p &+ \psi_{k-1}(Y_0 - Y^*))) + (\Delta X_k^p + \psi_{k-1}(X_0 - X^*))Y_k) \\
&= H_{P_k}(X_k \Delta Y_k^p + \Delta X_k^p Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k) \\
&= -H_{P_k}(X_k Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k),
\end{aligned}
\tag{33}
$$

where the last equality holds since (X_k, Y_k) satisfies (5) with $X = X_k$, $Y = Y_k$, $\Delta X = \Delta X_k^p$, $\Delta Y = \Delta Y_k^p$, $\sigma = 0$. Since $(X_k, Y_k) \in \mathcal{N}_1(\beta_1, \tau_k)$, Proposition 3 can be applied to (32), (33). We have from the proposition

$$
\begin{aligned}
\|(P_k &\otimes_s P_k)\mathrm{svec}(\Delta X_k^p + \psi_{k-1}(X_0 - X^*))\|_2 \\
&\le \frac{1}{\sqrt{(1 - \beta_1)\tau_k}} \|\mathrm{svec}(-H_{P_k}(X_k Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) \\
&\quad + (X_0 - X^*)Y_k))\|_2,
\end{aligned}
\tag{34}
$$

$$
\begin{aligned}
\|(P_k^{-T} &\otimes_s P_k^{-T})\mathrm{svec}(\Delta Y_k^p + \psi_{k-1}(Y_0 - Y^*))\|_2 \\
&\le \frac{1}{\sqrt{(1 - \beta_1)\tau_k}} \|\mathrm{svec}(-H_{P_k}(X_k Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) \\
&\quad + (X_0 - X^*)Y_k))\|_2.
\end{aligned}
\tag{35}
$$

The proposition then follows by applying triangle inequality to (34), (35) and upon algebraic manipulations. $\qquad\square$

Our objective to prove Lemma 1 is achieved by bounding t_x, t_y and t that appear in the upper bounds to $\|P_k \Delta X_k^p P_k^T\|_F$, $\|P_k^{-T} \Delta Y_k^p P_k^{-1}\|_F$ in the above proposition appropriately. We need the following results to achieve this.

Proposition 8 *For all $k \ge 0$, $Y_k \bullet X_0 + Y_0 \bullet X_k \le (2 + \beta_1 + \varsigma)n\tau_0$.*

Proof By Proposition 6,

$$
\mathcal{A}(X_k) + \mathcal{B}(Y_k) = q + \psi_{k-1} r_0.
$$

Hence,

$$
\mathcal{A}(X_k - (1 - \psi_{k-1})X^* - \psi_{k-1}X_0) + \mathcal{B}(Y_k - (1 - \psi_{k-1})Y^* - \psi_{k-1}Y_0) = 0.
$$

Assumption 1(a) implies that

$$
(X_k - (1 - \psi_{k-1})X^* - \psi_{k-1}X_0) \bullet (Y_k - (1 - \psi_{k-1})Y^* - \psi_{k-1}Y_0) \ge 0,
$$

from which we have

$$
Y_k \bullet X_0 + Y_0 \bullet X_k \le \frac{1}{\psi_{k-1}} X_k \bullet Y_k + \psi_{k-1}X_0 \bullet Y_0 + (Y_0 \bullet X^* + Y^* \bullet X_0).
$$

Result then follows from the definition of ς, (13) (since $(X_k, Y_k) \in \mathcal{N}_1(\beta_1, \tau_k)$), $\tau_k = \psi_{k-1}\tau_0$, where $\psi_{k-1} \le 1$, and $X_0 \bullet Y_0 = n\mu_0 = n\tau_0$. $\qquad\square$

Remark 1 Since $(X_0, Y_0) \in S_{++}^n \times S_{++}^n$, we see easily from the above proposition that $\{(X_k, Y_k) ; \ k \ge 0\}$ is bounded.

Using Proposition 8, the following holds.

Proposition 9 *For all $k \ge 0$,*

$$\|P_k X_0 P_k^T\|_F \le (2 + \beta_1 + \varsigma)\|(P_k^{-T} Y_k P_k^{-1})^{-1/2}\|_2^2 n\tau_0,$$
$$\|P_k^{-T} Y_0 P_k^{-1}\|_F \le (2 + \beta_1 + \varsigma)\|(P_k X_k P_k^T)^{-1/2}\|_2^2 n\tau_0,$$
$$\|P_k X^* P_k^T\|_F \le (2 + \beta_1 + \varsigma)\|(P_k^{-T} Y_k P_k^{-1})^{-1/2}\|_2^2 \times$$
$$\|P_0 X^* P_0^T\|_F \|(P_0 X_0 P_0^T)^{-1/2}\|_2^2 n\tau_0,$$
$$\|P_k^{-T} Y^* P_k^{-1}\|_F \le (2 + \beta_1 + \varsigma)\|(P_k X_k P_k^T)^{-1/2}\|_2^2 \times$$
$$\|P_0^{-T} Y^* P_0^{-1}\|_F \|(P_0^{-T} Y_0 P_0^{-1})^{-1/2}\|_2^2 n\tau_0.$$

Proof It suffices to prove the first inequality since the proofs for the last three inequalities are similar.
We have

$$\begin{aligned}
\|P_k X_0 P_k^T\|_F &\le \|(P_k X_0 P_k^T)^{1/2}\|_F^2 \\
&= \|(P_k X_0 P_k^T)^{1/2}(P_k^{-T} Y_k P_k^{-1})^{1/2}(P_k^{-T} Y_k P_k^{-1})^{-1/2}\|_F^2 \\
&\le \|(P_k X_0 P_k^T)^{1/2}(P_k^{-T} Y_k P_k^{-1})^{1/2}\|_F^2 \|(P_k^{-T} Y_k P_k^{-1})^{-1/2}\|_2^2 \\
&= (X_0 \bullet Y_k)\|(P_k^{-T} Y_k P_k^{-1})^{-1/2}\|_2^2 \\
&\le (2 + \beta_1 + \varsigma)\|(P_k^{-T} Y_k P_k^{-1})^{-1/2}\|_2^2 n\tau_0,
\end{aligned}$$

where the first inequality holds since $P_k X_0 P_k^T \in S_{++}^n$, the second inequality holds by Fact 2, and the last inequality holds by Proposition 8. $\qquad\square$

With the above, we are ready to prove Lemma 1 by providing suitable upper bounds for t_x, t_y and t as given below.

Proposition 10

$$t \le \tau_k \left[\beta_1 + \sqrt{n} + \frac{\sqrt{1 + \beta_1}(2 + \beta_1 + \varsigma)n}{\sqrt{1 - \beta_1}}(\varsigma_x + \varsigma_y) \right], \tag{36}$$

$$t_x \le \frac{(2 + \beta_1 + \varsigma)n\sqrt{\tau_k}}{\sqrt{1 - \beta_1}}\varsigma_x, \tag{37}$$

$$t_y \le \frac{(2 + \beta_1 + \varsigma)n\sqrt{\tau_k}}{\sqrt{1 - \beta_1}}\varsigma_y. \tag{38}$$

Proof We first prove the upper bound on t. We have

$$
\begin{aligned}
t &= \|\mathrm{svec}(-H_{P_k}(X_k Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k))\|_2 \\
&= \| - H_{P_k}(X_k Y_k) + \psi_{k-1} H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k)\|_F \\
&\le \|H_{P_k}(X_k Y_k)\|_F + \psi_{k-1}\|H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k))\|_F .
\end{aligned}
\tag{39}
$$

Since $(X_k, Y_k) \in \mathcal{N}_1(\beta, \tau_k)$,

$$
\|H_{P_k}(X_k Y_k) - \tau_k I\|_F \le \beta_1 \tau_k.
$$

Hence,

$$
\|H_{P_k}(X_k Y_k)\|_F \le (\beta_1 + \sqrt{n})\tau_k.
$$

On the other hand,

$$
\begin{aligned}
&\|H_{P_k}(X_k(Y_0 - Y^*) + (X_0 - X^*)Y_k))\|_F \\
&\le \|P_k(X_k(Y_0 - Y^*))P_k^{-1}\|_F + \|P_k((X_0 - X^*)Y_k)P_k^{-1}\|_F .
\end{aligned}
$$

Now,

$$
\begin{aligned}
\|P_k(X_k(Y_0 - Y^*))P_k^{-1}\|_F &\le \|P_k X_k P_k^T\|_2 \|P_k^{-T}(Y_0 - Y^*)P_k^{-1}\|_F \\
&\le \|P_k X_k P_k^T\|_2 [\|P_k^{-T} Y_0 P_k^{-1}\|_F + \|P_k^{-T} Y^* P_k^{-1}\|_F] \\
&\le \|P_k X_k P_k^T\|_2 \|(P_k X_k P_k^T)^{-1/2}\|_2^2 (2 + \beta_1 + \varsigma)n\tau_0 \varsigma_y \\
&\le \frac{\sqrt{1 + \beta_1}(2 + \beta_1 + \varsigma)n\tau_0}{\sqrt{1 - \beta_1}} \varsigma_y,
\end{aligned}
$$

where the third inequality follows by Proposition 9, and the last inequality follows from (15), (16). Similarly,

$$
\|P_k((X_0 - X^*)Y_k)P_k^{-1}\|_F \le \frac{\sqrt{1 + \beta_1}(2 + \beta_1 + \varsigma)n\tau_0}{\sqrt{1 - \beta_1}} \varsigma_x.
$$

Putting everything together, we have (36).
In a similar way, we can show (37) and (38). □

From (28), (29), (30), using Proposition 10, Lemma 1 is proved.

4 Local convergence study of interior point algorithm

In this section, we investigate the local convergence behavior of iterates $\{(X_k, Y_k)\}$ generated by Algorithm 1.

We first need an assumption as follows.

Assumption 2 There exists a strictly complementary solution (X^*, Y^*) to the SDLCP (1)–(3). That is, (X^*, Y^*) satisfies $X^* + Y^* \in S_{++}^n$.

Assumption 2 is usually applied when we study the local convergence behavior of interior point algorithms on semi-definite linear complementary problems [13,16,24, 25,34]. Paper [5] consider asymptotic behavior of the central path for an SDP when this assumption is relaxed.

In this section, we consider an SDLCP (1)–(3) that satisfies Assumptions 1(a),(c) and Assumption 2.

(X^*, Y^*) now denotes a strictly complementary solution to the SDLCP (1)–(3). Since X^*, Y^* commute, they are jointly diagonalizable by some orthogonal matrix Q. Applying this orthogonal similarity transformation on the matrices in the SDLCP (1)–(3), we may assume without loss of generality that

$$X^* = \begin{bmatrix} \Lambda^{X^*} & 0 \\ 0 & 0 \end{bmatrix}, \quad Y^* = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda^{Y^*} \end{bmatrix},$$

where $\Lambda^{X^*} = \mathrm{Diag}(\lambda_1^{X^*}, \dots, \lambda_{k_0}^{X^*}) \in S_{++}^{k_0}$ and $\Lambda^{Y^*} = \mathrm{Diag}(\lambda_1^{Y^*}, \dots, \lambda_{n-k_0}^{Y^*}) \in S_{++}^{n-k_0}$.

Henceforth, whenever we partition a matrix $U \in S^n$, it is always understood that it is partitioned as $\begin{bmatrix} U_{11} & U_{12} \\ U_{12}^T & U_{22} \end{bmatrix}$, where $U_{11} \in S^{k_0}$, $U_{22} \in S^{n-k_0}$ and $U_{12} \in \Re^{k_0 \times (n-k_0)}$.

We study local superlinear convergence using Algorithm 1 in the sense of

$$\frac{\mu_{k+1}}{\mu_k} \to 0, \text{ as } k \to \infty. \tag{40}$$

This is equivalent to

$$\frac{\tau_{k+1}}{\tau_k} \to 0, \text{ as } k \to \infty, \tag{41}$$

by Proposition 6. Note that we have $\tau_k \to 0$ as $k \to \infty$, by Theorem 1 and Proposition 6.

Superlinear convergence in the sense of (40) is intimately related to local convergence behavior of iterates, as investigated for example in [23]. The following can be verified easily.

Proposition 11 *A sufficient condition for* (41) *to hold is*

$$\delta_k = \frac{1}{\tau_k} \| H_{P_k}(\Delta X_k^p \Delta Y_k^p) \|_F \to 0, \text{ as } k \to \infty, \text{ or } \| H_{P_k}(\Delta X_k^p \Delta Y_k^p) \|_F = o(\tau_k). \tag{42}$$

Proof Note that (41) holds if $\alpha_{k,1} \to 1$ as $k \to \infty$. By (9), this sufficient condition is equivalent to $\delta_k \to 0$ as $k \to \infty$, where δ_k is given by (10). Hence, the proposition is proved. □

In the rest of this section, we are going to show that if certain condition holds for (X_k, Y_k) and for certain subclass of SDLCPs, superlinear convergence using Algorithm 1 can be achieved. This is achieved by showing that the above sufficient condition (42) holds. Towards this end, we transform the system of equations, (5), (6), that relates $(\Delta X_k^P, \Delta Y_k^P)$ to (X_k, Y_k) to an equivalent system of equations, and then analyze the resulting system.

From Step (a1) of Algorithm 1, where the system of Eqs. (5), (6) is used with $\Delta X = \Delta X_k^P, \Delta Y = \Delta Y_k^P, X = X_k, Y = Y_k, P = P_k, \sigma = 0, \tau = \tau_k$ and $\bar{r} = r_k = (\tau_k/\tau_0)r_0$, we note that $(\Delta X_k^P, \Delta Y_k^P)$ satisfies

$$\mathcal{A}svec(\Delta X_k^P) + \mathcal{B}svec(\Delta Y_k^P) = -\frac{\tau_k}{\tau_0}r_0, \tag{43}$$

$$H_{P_k}(X_k \Delta Y_k^P + \Delta X_k^P Y_k) = -H_{P_k}(X_k Y_k), \tag{44}$$

where $\mathcal{A} : \Re^{\tilde{n}} \to \Re^{\tilde{n}}$ and $\mathcal{B} : \Re^{\tilde{n}} \to \Re^{\tilde{n}}$ are defined by $\mathcal{A}svec(U) := \mathcal{A}(U)$ and $\mathcal{B}svec(U) := \mathcal{B}(U)$, respectively, for $U \in S^n$. Here, $(\mathcal{A}\ \mathcal{B})$ has full row rank equal to $\tilde{n}$ by Assumption 1(c).

We observe that (43), (44) can be written in the following way.

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ I & W_k \otimes_s W_k \end{bmatrix} \begin{bmatrix} svec(\Delta X_k^P) \\ svec(\Delta Y_k^P) \end{bmatrix} = - \begin{bmatrix} \frac{\tau_k}{\tau_0}r_0 \\ svec(X_k) \end{bmatrix}, \tag{45}$$

by (12).

The task now is to transform (45) to an equivalent system of equations that allows us to show that if certain condition on iterates (X_k, Y_k) as given in Theorem 3 is satisfied, and for certain subclass of SDLCPs as given in Theorem 4, superlinear convergence using Algorithm 1 can be ensured. First, we observe the following.

Proposition 12 $X_k = \begin{bmatrix} \Theta(1) & \mathcal{O}(\sqrt{\tau_k}) \\ \mathcal{O}(\sqrt{\tau_k}) & \Theta(\tau_k) \end{bmatrix}$, $Y_k = \begin{bmatrix} \Theta(\tau_k) & \mathcal{O}(\sqrt{\tau_k}) \\ \mathcal{O}(\sqrt{\tau_k}) & \Theta(1) \end{bmatrix}$.

Since the proof of the above proposition follows from proofs of similar results in [26, 36] and the relation between μ_k and τ_k as stated in Proposition 6, it will not be shown here.

The new system of equations that we are going to derive involves "iterate" $(\bar{X}_k, \bar{Y}_k)$ corresponding to (X_k, Y_k), $\bar{W}_k$ corresponding to W_k, and corresponding "predicted step" $(\Delta \bar{X}_k^P, \Delta \bar{Y}_k^P)$ to $(\Delta X_k^P, \Delta Y_k^P)$.

Define

$$\bar{X}_k := \begin{bmatrix} I & 0 \\ 0 & \frac{1}{\sqrt{\tau_k}}I \end{bmatrix} X_k \begin{bmatrix} I & 0 \\ 0 & \frac{1}{\sqrt{\tau_k}}I \end{bmatrix},$$

$$\bar{Y}_k := \begin{bmatrix} \frac{1}{\sqrt{\tau_k}}I & 0 \\ 0 & I \end{bmatrix} Y_k \begin{bmatrix} \frac{1}{\sqrt{\tau_k}}I & 0 \\ 0 & I \end{bmatrix},$$

and

$$\bar{W}_k := \frac{1}{\sqrt{\tau_k}} \begin{bmatrix} \sqrt{\tau_k} I & 0 \\ 0 & I \end{bmatrix} W_k \begin{bmatrix} \sqrt{\tau_k} I & 0 \\ 0 & I \end{bmatrix}. \tag{46}$$

Proposition 13

$$\bar{X}_k = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix}, \bar{X}_k^{-1} = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix} \in S_{++}^n,$$

$$\bar{Y}_k = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix}, \bar{Y}_k^{-1} = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix} \in S_{++}^n,$$

$$\bar{W}_k = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix}, \bar{W}_k^{-1} = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix} \in S_{++}^n. \tag{47}$$

Also, any accumulation point of $\{\bar{X}_k\}, \{\bar{Y}_k\}, \{\bar{W}_k\}$, as k tends to infinity, are symmetric, positive definite matrices.

Proof Observe by Proposition 12 that

$$\bar{X}_k = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix} \in S_{++}^n,$$

$$\bar{Y}_k = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix} \in S_{++}^n,$$

and by (13) (from which we get $1 - \beta_1 \leq \lambda_{\min}(\bar{X}_k \bar{Y}_k) \leq \lambda_{\max}(\bar{X}_k \bar{Y}_k) \leq 1 + \beta_1$), any accumulation point of $\{\bar{X}_k\}, \{\bar{X}_k^{-1}\}$ or $\{\bar{Y}_k\}, \{\bar{Y}_k^{-1}\}$ are symmetric, positive definite as k tends to infinity, with

$$\bar{X}_k^{-1} = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix},$$

$$\bar{Y}_k^{-1} = \begin{bmatrix} \Theta(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix}.$$

By definition of $\bar{W}_k$ in (46), we have $\bar{W}_k \bar{Y}_k \bar{W}_k = \bar{X}_k$, which implies that $\bar{W}_k = \bar{X}_k^{1/2} (\bar{X}_k^{1/2} \bar{Y}_k \bar{X}_k^{1/2})^{-1/2} \bar{X}_k^{1/2}$, from which we see that $\{\bar{W}_k\}$ is a sequence of symmetric, positive definite matrices and has accumulation points which are symmetric, positive definite as k tends to infinity, since these are so for $\{\bar{X}_k\}$ and $\{\bar{Y}_k\}$. Therefore, (47) holds. $\square$

By partitioning matrices A_i and B_i in $\mathcal{A}, \mathcal{B}$, as appeared in (4), respectively into the 4 blocks format as discussed near the beginning of this section, we perform block

Gaussian elimination on $(\mathcal{A} \ \mathcal{B})$ so that $\mathcal{A}$, $\mathcal{B}$ can be rewritten as

$$
\begin{bmatrix}
\left(\operatorname{svec}\begin{pmatrix} (A_1)_{11} & (A_1)_{12} \\ (A_1)_{12}^T & (A_1)_{22} \end{pmatrix} \right)^T \\
\vdots \\
\left(\operatorname{svec}\begin{pmatrix} (A_{i_1})_{11} & (A_{i_1})_{12} \\ (A_{i_1})_{12}^T & (A_{i_1})_{22} \end{pmatrix} \right)^T \\
\left(\operatorname{svec}\begin{pmatrix} 0 & (A_{i_1+1})_{12} \\ (A_{i_1+1})_{12}^T & (A_{i_1+1})_{22} \end{pmatrix} \right)^T \\
\vdots \\
\left(\operatorname{svec}\begin{pmatrix} 0 & (A_{i_1+i_2})_{12} \\ (A_{i_1+i_2})_{12}^T & (A_{i_1+i_2})_{22} \end{pmatrix} \right)^T \\
\left(\operatorname{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{i_1+i_2+1})_{22} \end{pmatrix} \right)^T \\
\vdots \\
\left(\operatorname{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{\tilde{n}})_{22} \end{pmatrix} \right)^T
\end{bmatrix},
\begin{bmatrix}
\left(\operatorname{svec}\begin{pmatrix} (B_1)_{11} & (B_1)_{12} \\ (B_1)_{12}^T & (B_1)_{22} \end{pmatrix} \right)^T \\
\vdots \\
\left(\operatorname{svec}\begin{pmatrix} (B_{i_1})_{11} & (B_{i_1})_{12} \\ (B_{i_1})_{12}^T & (B_{i_1})_{22} \end{pmatrix} \right)^T \\
\left(\operatorname{svec}\begin{pmatrix} (B_{i_1+1})_{11} & (B_{i_1+1})_{12} \\ (B_{i_1+1})_{12}^T & 0 \end{pmatrix} \right)^T \\
\vdots \\
\left(\operatorname{svec}\begin{pmatrix} (B_{i_1+i_2})_{11} & (B_{i_1+i_2})_{12} \\ (B_{i_1+i_2})_{12}^T & 0 \end{pmatrix} \right)^T \\
\left(\operatorname{svec}\begin{pmatrix} (B_{i_1+i_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix} \right)^T \\
\vdots \\
\left(\operatorname{svec}\begin{pmatrix} (B_{\tilde{n}})_{11} & 0 \\ 0 & 0 \end{pmatrix} \right)^T
\end{bmatrix},
\tag{48}
$$

respectively. This technique has been used for example in [33,36]. See also [15,26]. We will take $\mathcal{A}$, $\mathcal{B}$ to be expressed as in (48) from now onwards. Note that $(\mathcal{A} \ \mathcal{B}) \in \mathfrak{R}^{\tilde{n} \times 2\tilde{n}}$ has full row rank by Assumption 1(c), and $\mathcal{A}$, $\mathcal{B}$ also satisfy

$$
\mathcal{A}u + \mathcal{B}v = 0 \text{ for } u, v \in \mathfrak{R}^{\tilde{n}} \Rightarrow u^T v \geq 0.
\tag{49}
$$

The implication in (49) holds by Assumption 1(a).

Remark 2 In the case of an SDP, $\mathcal{A}$ and $\mathcal{B}$ are written as

$$
\mathcal{A} = \begin{bmatrix} \mathcal{A}_1 \\ 0 \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 \\ \mathcal{B}_1 \end{bmatrix},
\tag{50}
$$

where $\mathcal{A}_1$ consists of m rows and $\mathcal{B}_1$ consists of $\tilde{n} - m$ rows. As discussed in [34], by performing block Gaussian elimination, $\mathcal{A}_1$ and $\mathcal{B}_1$ are given by

$$
\begin{bmatrix}
\left(\mathrm{svec}\begin{pmatrix} (A_1)_{11} & (A_1)_{12} \\ (A_1)_{12}^T & (A_1)_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} (A_{j_1})_{11} & (A_{j_1})_{12} \\ (A_{j_1})_{12}^T & (A_{j_1})_{22} \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} 0 & (A_{j_1+1})_{12} \\ (A_{j_1+1})_{12}^T & (A_{j_1+1})_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} 0 & (A_{j_1+j_2})_{12} \\ (A_{j_1+j_2})_{12}^T & (A_{j_1+j_2})_{22} \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{j_1+j_2+1})_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_m)_{22} \end{pmatrix}\right)^T
\end{bmatrix},
\quad
\begin{bmatrix}
\left(\mathrm{svec}\begin{pmatrix} (B_1)_{11} & (B_1)_{12} \\ (B_1)_{12}^T & (B_1)_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} (B_{k_1})_{11} & (B_{k_1})_{12} \\ (B_{k_1})_{12}^T & (B_{k_1})_{22} \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} (B_{k_1+1})_{11} & (B_{k_1+1})_{12} \\ (B_{k_1+1})_{12}^T & 0 \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} (B_{k_1+k_2})_{11} & (B_{k_1+k_2})_{12} \\ (B_{k_1+k_2})_{12}^T & 0 \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} (B_{k_1+k_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} (B_{\tilde{n}-m})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T
\end{bmatrix},
$$

respectively. Note that the way $\mathcal{A}$, $\mathcal{B}$ for an SDP are written in (50) is different from that for an SDLCP, see (48). We can however write them in the form of (48) by appropriately interchanging rows in $(\mathcal{A}\ \mathcal{B})$ for the SDP.

Now, in order to transform the equation system (45) to an equivalent system, let us define $\bar{\mathcal{A}}(\tau) \in \Re^{\tilde{n} \times \tilde{n}}$ and $\bar{\mathcal{B}}(\tau) \in \Re^{\tilde{n} \times \tilde{n}}$ to be

$$
\begin{bmatrix}
\left(\mathrm{svec}\begin{pmatrix} (A_1)_{11} & \sqrt{\tau}(A_1)_{12} \\ \sqrt{\tau}(A_1)_{12}^T & \tau(A_1)_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} (A_{i_1})_{11} & \sqrt{\tau}(A_{i_1})_{12} \\ \sqrt{\tau}(A_{i_1})_{12}^T & \tau(A_{i_1})_{22} \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} 0 & (A_{i_1+1})_{12} \\ (A_{i_1+1})_{12}^T & \sqrt{\tau}(A_{i_1+1})_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} 0 & (A_{i_1+i_2})_{12} \\ (A_{i_1+i_2})_{12}^T & \sqrt{\tau}(A_{i_1+i_2})_{22} \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{i_1+i_2+1})_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{\tilde{n}})_{22} \end{pmatrix}\right)^T
\end{bmatrix},
\quad
\begin{bmatrix}
\left(\mathrm{svec}\begin{pmatrix} \tau(B_1)_{11} & \sqrt{\tau}(B_1)_{12} \\ \sqrt{\tau}(B_1)_{12}^T & (B_1)_{22} \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} \tau(B_{i_1})_{11} & \sqrt{\tau}(B_{i_1})_{12} \\ \sqrt{\tau}(B_{i_1})_{12}^T & (B_{i_1})_{22} \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} \sqrt{\tau}(B_{i_1+1})_{11} & (B_{i_1+1})_{12} \\ (B_{i_1+1})_{12}^T & 0 \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} \sqrt{\tau}(B_{i_1+i_2})_{11} & (B_{i_1+i_2})_{12} \\ (B_{i_1+i_2})_{12}^T & 0 \end{pmatrix}\right)^T \\
\left(\mathrm{svec}\begin{pmatrix} (B_{i_1+i_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T \\
\vdots \\
\left(\mathrm{svec}\begin{pmatrix} (B_{\tilde{n}})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T
\end{bmatrix}, \quad (51)
$$

respectively, for $\tau \geq 0$.

The following proposition, whose proof can be found for example in [32,33,36], relates $\mathcal{A}$ with $\bar{\mathcal{A}}(\tau)$ and $\mathcal{B}$ with $\bar{\mathcal{B}}(\tau)$:

Proposition 14

$$\mathcal{A}\left(\begin{bmatrix} I & 0 \\ 0 & \sqrt{\tau}I \end{bmatrix} \otimes_s \begin{bmatrix} I & 0 \\ 0 & \sqrt{\tau}I \end{bmatrix}\right) = \mathrm{Diag}(I_{i_1 \times i_1}, \sqrt{\tau}I_{i_2 \times i_2}, \tau I_{(\tilde{n}-i_1-i_2) \times (\tilde{n}-i_1-i_2)}) \bar{\mathcal{A}}(\tau)$$

(52)

and

$$\mathcal{B}\left(\begin{bmatrix} \sqrt{\tau}I & 0 \\ 0 & I \end{bmatrix} \otimes_s \begin{bmatrix} \sqrt{\tau}I & 0 \\ 0 & I \end{bmatrix}\right) = \mathrm{Diag}(I_{i_1 \times i_1}, \sqrt{\tau}I_{i_2 \times i_2}, \tau I_{(\tilde{n}-i_1-i_2) \times (\tilde{n}-i_1-i_2)}) \bar{\mathcal{B}}(\tau).$$

(53)

An important property of $\bar{\mathcal{A}}(\tau)$, $\bar{\mathcal{B}}(\tau)$ is given below.

Proposition 15 *For $\tau \geq 0$, $(\bar{\mathcal{A}}(\tau)\ \bar{\mathcal{B}}(\tau))$ has full row rank, and*

$$\bar{\mathcal{A}}(\tau)u + \bar{\mathcal{B}}(\tau)v = 0 \ \text{for} \ u, v \in \Re^{\tilde{n}} \Rightarrow u^T v \geq 0.$$

(54)

Proof We note that $(\mathcal{A}(\tau)\ \mathcal{B}(\tau))$ has full row rank for all $\tau \geq 0$ follows from Assumption 1(c) and the way block Gaussian elimination is performed on $(\mathcal{A}\ \mathcal{B})$ to obtain $\mathcal{A}$, $\mathcal{B}$ in the form (48)—see explanations for example in [36].
The implication in (54) holds for $\tau > 0$ follows from Assumption 1(a) and (52), (53).
For $\tau = 0$, (54) holds by following the proof of Proposition 2.4 in [36] (see also the proof of Theorem 3.13 in [26]). $\qquad\square$

Because of the way $\mathcal{A}$, $\mathcal{B}$ are now structured, we have

Proposition 16 *q has the following form*

$$q = \begin{bmatrix} q^1 \\ 0 \\ 0 \end{bmatrix} \in \Re^{\tilde{n}},$$

where $q^1 \in \Re^{i_1}$.

Proof By Proposition 6, the following equation holds

$$\mathcal{A}\mathrm{svec}(X_k) + \mathcal{B}\mathrm{svec}(Y_k) = q + \frac{\tau_k}{\tau_0}r_0.$$

This equation is equivalent to

$$\bar{\mathcal{A}}(\tau_k)\mathrm{svec}(\bar{X}_k) + \bar{\mathcal{B}}(\tau_k)\mathrm{svec}(\bar{Y}_k)$$
$$= \mathrm{Diag}\left(I_{i_1 \times i_1}, \frac{1}{\sqrt{\tau_k}}I_{i_2 \times i_2}, \frac{1}{\tau_k}I_{(\tilde{n}-i_1-i_2) \times (\tilde{n}-i_1-i_2)}\right)\left(q + \frac{\tau_k}{\tau_0}r_0\right).$$

Since the left hand side of the above equation and $\frac{\tau_k}{\tau_0}\mathrm{Diag}\left(I_{i_1\times i_1}, \frac{1}{\sqrt{\tau_k}}I_{i_2\times i_2}, \frac{1}{\tau_k}\right.$ $\left.I_{(\tilde{n}-i_1-i_2)\times(\tilde{n}-i_1-i_2)}\right)r_0$ are bounded as k tends to infinity, we conclude that q must take the form as given in the proposition. $\qquad\square$

Remark 3 From the proof of the above proposition, we see that $(\bar{X}_k, \bar{Y}_k)$ satisfies

$$\bar{\mathcal{A}}(\tau_k)\mathrm{svec}(\bar{X}_k) + \bar{\mathcal{B}}(\tau_k)\mathrm{svec}(\bar{Y}_k) = q + \frac{\tau_k}{\tau_0}\bar{r}_{0,k},$$

where

$$\bar{r}_{0,k} := \mathrm{Diag}\left(I_{i_1\times i_1}, \frac{1}{\sqrt{\tau_k}}I_{i_2\times i_2}, \frac{1}{\tau_k}I_{(\tilde{n}-i_1-i_2)\times(\tilde{n}-i_1-i_2)}\right)r_0. \tag{55}$$

Finally, by defining $\Delta\bar{X}_k^p$ and $\Delta\bar{Y}_k^p$ to be

$$\Delta\bar{X}_k^p := \begin{bmatrix} I & 0 \\ 0 & \frac{1}{\sqrt{\tau_k}}I \end{bmatrix}\Delta X_k^p\begin{bmatrix} I & 0 \\ 0 & \frac{1}{\sqrt{\tau_k}}I \end{bmatrix}, \tag{56}$$

$$\Delta\bar{Y}_k^p := \begin{bmatrix} \frac{1}{\sqrt{\tau_k}}I & 0 \\ 0 & I \end{bmatrix}\Delta Y_k^p\begin{bmatrix} \frac{1}{\sqrt{\tau_k}}I & 0 \\ 0 & I \end{bmatrix}, \tag{57}$$

we obtain from (45) the following new system of equations derived from the original system (45)

$$\begin{bmatrix} \bar{\mathcal{A}}(\tau_k) & \bar{\mathcal{B}}(\tau_k) \\ I & \bar{W}_k\otimes_s\bar{W}_k \end{bmatrix}\begin{bmatrix} \mathrm{svec}(\Delta\bar{X}_k^p) \\ \mathrm{svec}(\Delta\bar{Y}_k^p) \end{bmatrix} = -\begin{bmatrix} \frac{\tau_k}{\tau_0}\bar{r}_{0,k} \\ \mathrm{svec}(\bar{X}_k) \end{bmatrix}, \tag{58}$$

Let us take the inverse of the matrix on the left hand side of the above equation, which can be shown to exist for $\tau_k > 0$. Define

$$\mathcal{G}_k := \mathcal{B} - \mathcal{A}(W_k\otimes_s W_k)$$

and

$$\bar{\mathcal{G}}_k := \bar{\mathcal{B}}(\tau_k) - \bar{\mathcal{A}}(\tau_k)(\bar{W}_k\otimes_s\bar{W}_k).$$

Note that $\mathcal{G}_k$ and $\bar{\mathcal{G}}_k$ are related by

$$\mathcal{G}_k\left(\begin{bmatrix} \sqrt{\tau_k}I & 0 \\ 0 & I \end{bmatrix}\otimes_s\begin{bmatrix} \sqrt{\tau_k}I & 0 \\ 0 & I \end{bmatrix}\right)$$
$$= \mathrm{Diag}\left(I_{i_1\times i_1}, \frac{1}{\sqrt{\tau_k}}I_{i_2\times i_2}, \frac{1}{\tau_k}I_{(\tilde{n}-i_1-i_2)\times(\tilde{n}-i_1-i_2)}\right)\bar{\mathcal{G}}_k.$$

We have from (58), by taking the inverse of the matrix on the left hand side in (58), the following

$$
\begin{bmatrix} \operatorname{svec}(\Delta \bar{X}_k^p) \\ \operatorname{svec}(\Delta \bar{Y}_k^p) \end{bmatrix} = - \begin{bmatrix} -(\bar{W}_k \otimes_s \bar{W}_k)\bar{\mathcal{G}}_k^{-1} & I + (\bar{W}_k \otimes_s \bar{W}_k)\bar{\mathcal{G}}_k^{-1}\mathcal{A} \\ \bar{\mathcal{G}}_k^{-1} & -\bar{\mathcal{G}}_k^{-1}\mathcal{A} \end{bmatrix} \begin{bmatrix} \frac{\tau_k}{\tau_0}\bar{r}_{0,k} \\ \operatorname{svec}(\bar{X}_k) \end{bmatrix}
$$

$$
= \frac{1}{2} \begin{bmatrix} -(\bar{W}_k \otimes_s \bar{W}_k)\bar{\mathcal{G}}_k^{-1}\left(q - \frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right) - \operatorname{svec}(\bar{X}_k) \\ \bar{\mathcal{G}}_k^{-1}\left(q - \frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right) - \operatorname{svec}(\bar{Y}_k) \end{bmatrix}, \tag{59}
$$

where to obtain the second equality, we use the identity

$$
-\bar{\mathcal{G}}_k^{-1}\left(\frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right) + \bar{\mathcal{G}}_k^{-1}\mathcal{A}\operatorname{svec}(\bar{X}_k) = \frac{1}{2}\bar{\mathcal{G}}_k^{-1}\left(q - \frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right) - \frac{1}{2}\operatorname{svec}(\bar{Y}_k).
$$

Given that we are using (59), which involves $\Delta \bar{X}_k^p$, $\Delta \bar{Y}_k^p$, to derive meaningful conditions for superlinear convergence using Algorithm 1, let us now express (42) in terms of them. We know that $P_k^T P_k = W_k^{-1}$ and by (46), we can let

$$
P_k = \frac{1}{\tau_k^{1/4}} \bar{W}_k^{-1/2} \begin{bmatrix} \sqrt{\tau_k}I & 0 \\ 0 & I \end{bmatrix}. \tag{60}
$$

In the following lemma, we provide the sufficient condition for superlinear convergence using Algorithm 1 in terms of $\Delta \bar{X}_k^p$, $\Delta \bar{Y}_k^p$.

Lemma 2 *If*

$$
\Delta \bar{X}_k^p \Delta \bar{Y}_k^p = o(1), \tag{61}
$$

then superlinear convergence in the sense of (40) using Algorithm 1 follows.

Proof A sufficient condition for superlinear convergence in the sense of (40) is (42). Now observe that for (42) to hold, it is sufficient to have

$$
P_k \Delta X_k^p \Delta Y_k^p P_k^{-1} = o(\tau_k). \tag{62}
$$

By (47), P_k given by (60) satisfies

$$
P_k = \frac{1}{\tau_k^{1/4}} \begin{bmatrix} \Theta(\sqrt{\tau_k}) & \mathcal{O}(1) \\ \mathcal{O}(\sqrt{\tau_k}) & \Theta(1) \end{bmatrix}, \ P_k^{-1} = \tau_k^{1/4} \begin{bmatrix} \Theta\left(\frac{1}{\sqrt{\tau_k}}\right) & \mathcal{O}\left(\frac{1}{\sqrt{\tau_k}}\right) \\ \mathcal{O}(1) & \Theta(1) \end{bmatrix}. \tag{63}
$$

By (63), (62) holds if and only if

$$
\Delta X_k^p \Delta Y_k^p = \begin{bmatrix} o(\tau_k) & o(\sqrt{\tau_k}) \\ o(\tau_k^{3/2}) & o(\tau_k) \end{bmatrix}. \tag{64}
$$

Hence, since (62) is sufficient for (40) to hold, the lemma follows by applying (56) and (57) on (64). □

The system of equations in (59) relates (X_k, Y_k) through $(\bar{X}_k, \bar{Y}_k)$ to $(\Delta \bar{X}_k^p, \Delta \bar{Y}_k^p)$, and allows us to have a way to validate conditions on (X_k, Y_k) for superlinear convergence using Algorithm 1 by showing that (61) holds. Before we provide such a condition in Theorem 3 below, let us observe the following.

Proposition 17

$$\lim_{k \to \infty} \frac{\tau_k}{\tau_0} \bar{r}_{0,k} = \frac{1}{\tau_0} \bar{\mathcal{A}}(0) \text{svec} \begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} + \frac{1}{\tau_0} \bar{\mathcal{B}}(0) \text{svec} \begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix}.$$

Proof We have

$$\lim_{k \to \infty} \frac{\tau_k}{\tau_0} \bar{r}_{0,k}$$

$$= \lim_{k \to \infty} \frac{1}{\tau_0} \text{Diag}(\tau_k I_{i_1 \times i_1}, \sqrt{\tau_k} I_{i_2 \times i_2}, I_{(\tilde{n}-i_1-i_2) \times (\tilde{n}-i_1-i_2)}) r_0$$

$$= \lim_{k \to \infty} \frac{1}{\tau_0} \bar{\mathcal{A}}(\tau_k) \left(\begin{bmatrix} \sqrt{\tau_k} I & 0 \\ 0 & I \end{bmatrix} \otimes_s \begin{bmatrix} \sqrt{\tau_k} I & 0 \\ 0 & I \end{bmatrix} \right) \text{svec}(X_0)$$

$$+ \frac{1}{\tau_0} \bar{\mathcal{B}}(\tau_k) \left(\begin{bmatrix} I & 0 \\ 0 & \sqrt{\tau_k} I \end{bmatrix} \otimes_s \begin{bmatrix} I & 0 \\ 0 & \sqrt{\tau_k} I \end{bmatrix} \right) \text{svec}(Y_0)$$

$$= \frac{1}{\tau_0} \bar{\mathcal{A}}(0) \text{svec} \begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} + \frac{1}{\tau_0} \bar{\mathcal{B}}(0) \text{svec} \begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix},$$

where the first equality holds by definition of $\bar{r}_{0,k}$ in (55), and the second equality holds by (52), (53) and the structure of q in Proposition 16. □

Theorem 3 *Suppose*

$$\frac{X_k Y_k}{\sqrt{\tau_k}} \to 0, \text{ as } k \to \infty, \tag{65}$$

then Algorithm 1 is a superlinearly convergent algorithm in the sense of (40).

Proof By Proposition 12, (65) implies that

$$(X_k Y_k)_{12} = o(\sqrt{\tau_k}),$$

which further implies that

$$(X_k)_{12} = o(\sqrt{\tau_k}), \ (Y_k)_{12} = o(\sqrt{\tau_k}), \tag{66}$$

by Claim 1, given in "Appendix".
Let $(\bar{X}^*, \bar{Y}^*)$ be any accumulation point of $\{(\bar{X}_k, \bar{Y}_k)\}$. Then (66) implies that

$$(\bar{X}^*)_{12} = (\bar{Y}^*)_{12} = 0.$$

We also have $\bar{W}^*$, which is the corresponding accumulation point of $\{\bar{W}_k\}$, with $(\bar{W}^*)_{12} = 0$.

Hence,

$$(\bar{\mathcal{A}}(0)(\bar{W}^* \otimes_s \bar{W}^*) + \bar{\mathcal{B}}(0))\text{svec}\begin{pmatrix} (\bar{Y}^*)_{11} & 0 \\ 0 & (\bar{Y}^*)_{22} \end{pmatrix}$$
$$= \bar{\mathcal{A}}(0)\text{svec}\begin{pmatrix} (\bar{X}^*)_{11} & 0 \\ 0 & (\bar{X}^*)_{22} \end{pmatrix} + \bar{\mathcal{B}}(0)\text{svec}\begin{pmatrix} (\bar{Y}^*)_{11} & 0 \\ 0 & (\bar{Y}^*)_{22} \end{pmatrix}. \qquad (67)$$

Furthermore, we have

$$q + \lim_{k \to \infty}\left(\frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right)$$
$$= q + \frac{1}{\tau_0}\bar{\mathcal{A}}(0)\text{svec}\begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} + \frac{1}{\tau_0}\bar{\mathcal{B}}(0)\text{svec}\begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix}, \qquad (68)$$

which follows from Proposition 17.

Note that

$$\bar{\mathcal{A}}(0)\text{svec}(\bar{X}^*) + \bar{\mathcal{B}}(0)\text{svec}(\bar{Y}^*) = q + \lim_{k \to \infty}\left(\frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right),$$

by Remark 3. From this equality, it then follows from (67), (68), $(\bar{X}^*)_{12} = (\bar{Y}^*)_{12} = 0$, the structure of q and that of $\bar{\mathcal{A}}(0)$, $\bar{\mathcal{B}}(0)$ that

$$-\bar{\mathcal{A}}(0)\text{svec}\begin{pmatrix} -(\bar{X}^*)_{11} & 0 \\ 0 & (\bar{X}^*)_{22} \end{pmatrix} + \bar{\mathcal{B}}(0)\text{svec}\begin{pmatrix} -(\bar{Y}^*)_{11} & 0 \\ 0 & (\bar{Y}^*)_{22} \end{pmatrix}$$
$$= q - \frac{1}{\tau_0}\bar{\mathcal{A}}(0)\text{svec}\begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} - \frac{1}{\tau_0}\bar{\mathcal{B}}(0)\text{svec}\begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

Therefore,

$$(\bar{\mathcal{B}}(0) - \bar{\mathcal{A}}(0)(\bar{W}^* \otimes \bar{W}^*))\text{svec}\begin{pmatrix} -(\bar{Y}^*)_{11} & 0 \\ 0 & (\bar{Y}^*)_{22} \end{pmatrix} = q - \lim_{k \to \infty}\left(\frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right).$$

Hence, in the limit as k tends to infinity, $\bar{\mathcal{G}}_k^{-1}\left(q - \frac{\tau_k}{\tau_0}\bar{r}_{0,k}\right)$ tends to

$$\text{svec}\begin{pmatrix} -(\bar{Y}^*)_{11} & 0 \\ 0 & (\bar{Y}^*)_{22} \end{pmatrix}.$$

Therefore, by (59),

$$\Delta\bar{Y}_k^p = \begin{pmatrix} -\Theta(1) & o(1) \\ o(1) & o(1) \end{pmatrix}.$$

In a similar way, we can show that

$$\Delta \bar{X}_k^p = \begin{pmatrix} o(1) & o(1) \\ o(1) & -\Theta(1) \end{pmatrix}.$$

Since (61) holds in this case, the theorem is proved. □

In Theorem 3, we provide a sufficient condition for superlinear convergence using Algorithm 1 on any SDLCP that satisfies Assumptions 1(a), (c) and Assumption 2. This sufficient condition is similar to that found in [24,34] using Algorithm 1 with the HKM search direction. We have shown in the above theorem that the same condition is also sufficient for superlinear convergence using Algorithm 1 with the NT search direction. Superlinear convergence result has been establised in [16] using the NT search direction by "narrowing" the neighborhood of the central path, although in [16], a feasible algorithm is considered, while here, we consider an infeasible algorithm, with more involved analysis.

In the following theorem, we give another sufficient condition for superlinear convergence using Algorithm 1 on SDLCPs that have certain structure

Theorem 4 *Let $\mathcal{A}, \mathcal{B}$ be such that for all $1 \leq i \leq \tilde{n}$, if $\mathcal{A}_{i\cdot} \neq 0$, then $\mathcal{B}_{i\cdot} = 0$ (or equivalently, if $\mathcal{B}_{i\cdot} \neq 0$, then $\mathcal{A}_{i\cdot} = 0$). Furthermore, let q satisfies either one of the following two conditions:*

1. *For all $1 \leq i \leq i_1$, if $\mathcal{B}_{i\cdot} \neq 0$, then $q_i = 0$.*
2. *For all $1 \leq i \leq i_1$, if $\mathcal{A}_{i\cdot} \neq 0$, then $q_i = 0$.*

Suppose (X_0, Y_0) is chosen such that

$$for \ all \ i_1 + i_2 + 1 \leq i \leq \tilde{n}, \ \mathcal{A}_{i\cdot}.\text{svec}(X_0) = 0,$$

if q satisfies the first condition, and

$$for \ all \ i_1 + i_2 + 1 \leq i \leq \tilde{n}, \ \mathcal{B}_{i\cdot}.\text{svec}(Y_0) = 0,$$

if q satisfies the second condition, then iterates generated by Algorithm 1 converge superlinearly in the sense of (40).

Proof We only need to show the theorem when the first condition (condition 1) on q is satisfied. The proof of the theorem when the second condition on q holds is similar. By Remark 3 and $\bar{W}_k \bar{Y}_k \bar{W}_k = \bar{X}_k$, we have

$$(\bar{\mathcal{A}}(\tau_k)(\bar{W}_k \otimes \bar{W}_k) + \bar{\mathcal{B}}(\tau_k))\text{svec}(\bar{Y}_k) = q + \frac{\tau_k}{\tau_0}\bar{r}_{0,k}$$

Let $(\bar{X}^*, \bar{Y}^*)$ be any accumulation point of $\{(\bar{X}_k, \bar{Y}_k)\}$ as k tends to infinity, with $\bar{W}^*$ the corresponding accumulation point of $\{\bar{W}_k\}$. Then,

$$q + \lim_{k \to \infty} \frac{\tau_k}{\tau_0}\bar{r}_{0,k} = (\bar{\mathcal{A}}(0)(\bar{W}^* \otimes \bar{W}^*) + \bar{\mathcal{B}}(0))\text{svec}(\bar{Y}^*). \tag{69}$$

On the other hand,

$$q + \lim_{k \to \infty} \frac{\tau_k}{\tau_0} \bar{r}_{0,k}$$

$$= q + \frac{1}{\tau_0} \bar{\mathcal{A}}(0) \text{svec} \begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} + \frac{1}{\tau_0} \bar{\mathcal{B}}(0) \text{svec} \begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix}, \qquad (70)$$

which follows from Proposition 17.

By conditions imposed on $(\mathcal{A}, \mathcal{B}, q)$ and (X_0, Y_0) in the theorem, (51), and the structure of q in Proposition 16, it follows from (69) and (70) that

$$(\bar{\mathcal{B}}(0) - \bar{\mathcal{A}}(0)(\bar{W}^* \otimes \bar{W}^*)) \text{svec}(\bar{Y}^*)$$

$$= -q + \frac{1}{\tau_0} \bar{\mathcal{A}}(0) \text{svec} \begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} + \frac{1}{\tau_0} \bar{\mathcal{B}}(0) \text{svec} \begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence, in the limit as k tends to infinity, $\bar{\mathcal{G}}_k^{-1} \left(q - \frac{\tau_k}{\tau_0} \bar{r}_{0,k} \right)$ tends to $-\text{svec}(\bar{Y}^*)$ and $(\bar{W}_k \otimes \bar{W}_k) \bar{\mathcal{G}}_k^{-1} \left(q - \frac{\tau_k}{\tau_0} \bar{r}_{0,k} \right)$ tends to $-\text{svec}(\bar{X}^*)$. Therefore, (59) implies that

$$\Delta \bar{X}_k^p = o(1), \ \Delta \bar{Y}_k^p = \mathcal{O}(1),$$

from which we conclude that (61) holds which then implies that Algorithm 1 is a superlinearly convergent algorithm. □

Theorem 4 tells us that for certain subclass of SDLCPs, if a suitable starting point (X_0, Y_0) is chosen, then we have superlinear convergence using Algorithm 1 on the SDLCP. We can apply this theorem to an important subclass of semi-definite programs known as linear semi-definite feasibility problems. A linear semi-definite feasibility problem is a semi-definite program where $C = 0$ in $(\mathcal{P})$ or $b_j = 0$, $j = 1, \ldots, m$, in $(\mathcal{D})$.

Corollary 1 *For a linear semi-definite feasibility problem, when $C = 0$ in $(\mathcal{P})$, if X_0 is chosen such that for all $i_1 + i_2 + 1 \leq i \leq \tilde{n}$, $\mathcal{A}_i.\text{svec}(X_0) = 0$, and when $b_j = 0$, $j = 1, \ldots, m$, in $(\mathcal{D})$, if Y_0 is chosen such that for all $i_1 + i_2 + 1 \leq i \leq \tilde{n}$, $\mathcal{B}_i.\text{svec}(Y_0) = 0$, then Algorithm 1 is a superlinearly convergent algorithm on the linear semi-definite feasibility problem.*

Proof When $C = 0$ in $(\mathcal{P})$, it is easy to check that $(\mathcal{A}, \mathcal{B}, q)$ satisfies the conditions in Theorem 4, with q satisfying the first condition there. When $b_j = 0$, $j = 1, \ldots, m$, in $(\mathcal{D})$, it is also easy to check that $(\mathcal{A}, \mathcal{B}, q)$ satisfies the conditions in Theorem 4, with q satisfying the second condition there. Corollary then follows from Theorem 4. □

Corollary 1 states a similar result as Theorem 5.1 in [34]. The latter holds for the HKM search direction, while Corollary 1 applies to the NT search direction. It is worthwhile to note that to solve the linear semi-definite feasibility problem, in the literature [4,7,9,38,41] the assumption that the interior of dual feasible region is nonempty is usually made. In the above corollary, we do not need such an assumption to show

superlinear convergence using Algorithm 1. Only strict complementarity assumption and a suitable initial iterate are needed. In fact, we see from Corollary 1 that if the linear semi-definite feasibility problem has primal feasible region with nonempty relative interior in the case when $C = 0$ in $(\mathcal{P})$, and has dual feasible region with nonempty interior in the case when $b_j = 0$, $j = 1, \ldots, m$, in $(\mathcal{D})$, then Algorithm 1 is always a superlinearly convergent algorithm, irrespective of starting point (X_0, Y_0). This is so because the conditions in the corollary are satisfied trivially, as $k_0 = n$ in the former case and $k_0 = 0$ in the latter case. That is, in these cases, the matrices that we considered are no longer partitioned into 4 blocks, and we have $i_1 = \tilde{n}$ and $i_2 = 0$ in both cases.

5 Numerical study

In this section, we report on the numerical results we obtained upon applying Algorithm 1 to solve instances of SDLCP (1)–(3). The algorithm is implemented by writing Matlab R2018a scripts and is run on a personal computer with an Intel(R) Core(TM) i5-4460 CPU and 8 GB of memory. In existing SDP solvers, such as SeDuMi [37], SDPT3 [40], there is an option for the equation system to find the corrector step in the interior point algorithm to have a second order term. Having the second order term tends to enhance practical efficiency of the algorithm. However, we decide to perform our numerical experiments without introducing a second order term in the equation system to find the corrector step in Algorithm 1. This is mainly due to more observed numerical warnings when we run our Matlab programs as iterates get closer to an optimal solution for instances of SDLCP (1)–(3) we tested when the algorithm is used with the second order term introduced. It is also worthwhile to note that with or without the second order term in the equation system to find the corrector step, the number of iterations to solve an instance of SDLCP (1)–(3), namely, a linear semi-definite feasibility problem (LSDFP) with

$$
A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$
$$
A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \tag{71}
$$
$$
C = 0, b_1 = 1, b_i = 0, \ i = 2, 3, 4, 5,
$$

using our implemented algorithm differs only by 1 iteration, when tolerance ϵ is set to 10^{-10}.

In all our numerical experiments, we set $\beta_1 = 0.3$, $\beta_2 = 0.45$ and the tolerance ϵ to be 10^{-10} in Algorithm 1.

We generate random instances of SDLCP (1)–(3) by first generating diagonal matrices $D_{\hat{\mathcal{A}}}, D_{\hat{\mathcal{B}}} \in \Re^{\tilde{n} \times \tilde{n}}$, where $\tilde{n} = n(n + 1)/2$, such that the main diagonal entries

of $D_{\hat{\mathcal{B}}}$ are randomly taken from the uniform distribution between -5 and -1, and the main diagonal entries of $D_{\hat{\mathcal{A}}}$ are zero or nonzero with equal probability. If a main diagonal entry of $D_{\hat{\mathcal{A}}}$ is nonzero, then it is randomly assigned a value from the uniform distribution between 0 and 4. We obtain the matrices $\hat{\mathcal{A}}, \hat{\mathcal{B}} \in \Re^{\tilde{n} \times \tilde{n}}$ by

$$\hat{\mathcal{A}} = V D_{\hat{\mathcal{A}}} U, \quad \hat{\mathcal{B}} = V D_{\hat{\mathcal{B}}} U,$$

where $U, V \in \Re^{\tilde{n} \times \tilde{n}}$ are randomly generated orthogonal matrices, see [31]. We have $\mathcal{A}, \mathcal{B} \in \Re^{\tilde{n} \times \tilde{n}}$ to be obtained from $\hat{\mathcal{A}}, \hat{\mathcal{B}}$ by interchanging corresponding columns in the latter matrices when a random number generated from the uniform (0, 1) distribution is less than 0.5, and keeping these columns when the random number is greater than or equal to 0.5. $\mathcal{A}, \mathcal{B}$ thus obtained satisfy Assumption 1(a),(c). On the other hand, Assumption 1(b) is satisfied by setting q to be

$$\mathcal{A}(\mathrm{svec}(I_{n \times n})) + \mathcal{B}(\mathrm{svec}(I_{n \times n})).$$

Hence, X^1, Y^1 in Assumption 1(b) are both equal to the identity matrix. We set

$$X_0 = \eta I_{n \times n}, \quad Y_0 = \eta I_{n \times n},$$

where

$$\eta = \max \left\{ 10, \sqrt{n}, n \max_{1 \leq i \leq \tilde{n}} \left\{ \frac{1 + |q_i|}{1 + \|\mathcal{A}_i\|_2}, \frac{1 + |q_i|}{1 + \|\mathcal{B}_i\|_2} \right\} \right\}.$$

This choice of initial iterate (X_0, Y_0) is motivated by [40].

For each n from 5 to 15, we attempt to solve, using Algorithm 1, 100 instances of SDLCP (1)–(3) of size n, randomly generated with given initial iterate (X_0, Y_0) as described in the last paragraph. For each n, we compute the average number of iterations taken and average runtime for the algorithm to terminate for those instances that give real valued (X_k, Y_k) upon termination of the algorithm. We denote I_n to be the number of these instances. Our results are given in Table 1.

Comparing our implemented algorithm with existing solvers, using our implemented algorithm on the LSDFP with data given by (71), we need 12 iterations and a runtime of 0.02 secs before termination, while SDPT3 needs 7 iterations with a runtime of 0.13 secs as reported by the software (OPTIONS.gaptol in SDPT3 is set to $1e - 10$), and SeDuMi needs 5 iterations with a runtime of 1.28 secs as reported by the software (pars.eps in SeDuMi is set to $1e - 14$). The initial iterate (X_0, Y_0) used for our implemented algorithm on this LSDFP is

$$X_0 = 10 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -0.5 \\ 0 & 0 & -0.5 & 1 \end{pmatrix}, \quad Y_0 = 10I. \tag{72}$$

The same initial iterate is also used when solving this problem using SDPT3.

Table 1 Performance of implemented interior point algorithm on SDLCP (1)–(3)

n	I_n	Average number of iterations to termination	Average runtime (s)
5	94	21.07	0.02
6	91	22.80	0.03
7	90	23.22	0.05
8	77	24.64	0.07
9	80	25.79	0.10
10	68	26.84	0.13
11	64	27.91	0.18
12	65	27.97	0.23
13	54	29.61	0.31
14	52	30.44	0.40
15	54	31.63	0.51

We now report on our numerical investigation on the local convergence of Algorithm 1 when it is used to solve the LSDFP with data given by (71). It is easy to check that the LSDFP satisfies Assumption 2 with

$$X^* = \begin{pmatrix} x_1 & x_2 & 0 & 0 \\ x_2 & 1 - x_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad Y^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & y_1 + y_2 & y_3 \\ 0 & 0 & y_3 & y_3 - y_2 \end{pmatrix},$$

where x_1, x_2, y_1, y_2, y_3 are constrained for these matrices to be positive semi-definite. It is also easy to check that when the initial iterate (X_0, Y_0) in Algorithm 1 is chosen to be given by (72), the condition in Corollary 1 is satisfied. As predicted by our theoretical result in the corollary, our numerical results given in Table 2 shows that superlinear convergence of iterates generated by the implemented algorithm when solving this LSDFP takes place. It is interesting to note that with other choices of initial iterate (X_0, Y_0), we also observe superlinear convergence of iterates generated by the implemented algorithm.

6 Conclusion

In this paper, we consider an infeasible predictor–corrector primal–dual path following interior point algorithm, using the Nesterov–Todd (NT) search direction, to solve a semi-definite linear complementarity problem (SDLCP). Global convergence is shown using the algorithm to solve an SDLCP and an iteration complexity bound which is polynomial in n, the size of the matrices involved, is also provided. This complexity bound is the best known so far for infeasible interior point algorithms when the "narrow" neighborhood is used for solving SDPs. Furthermore, we study superlinear convergence using the algorithm under strict complementarity assumption. Two suf-

Table 2 Convergence behavior of iterates generated by implemented interior point algorithm on LSDFP with data given by (71)

k	$X_k \bullet Y_k / X_{k-1} \bullet Y_{k-1}$
1	0.251141
2	0.433461
3	0.505825
4	0.527020
5	0.512764
6	0.464522
7	0.376313
8	0.242669
9	0.091191
10	0.010082
11	0.000104
12	0.000013

ficient conditions are provided for this to occur. The first sufficient condition is on the behavior of iterates generated by the algorithm, while the second sufficient condition is on the structure of SDLCPs. We finally report on preliminary numerical results we obtained upon implementing the interior point algorithm and using it to solve SDLCPs that are not necessarily SDPs.

Appendix

Claim 1 *We have*

$$(X_k Y_k)_{12} = o(\sqrt{\tau_k})$$

if and only if

$$(X_k)_{12} = o(\sqrt{\tau_k}), \ (Y_k)_{12} = o(\sqrt{\tau_k}).$$

Proof [1] We only need to show the if direction, since the only if direction follows trivially using Proposition 12.

Suppose $(X_k Y_k)_{12} = o(\sqrt{\tau_k})$. Then $(\bar{X}_k \bar{Y}_k)_{12} = o(1)$. Let $(\bar{X}^*, \bar{Y}^*)$ be any accumulation point of $(\bar{X}_k, \bar{Y}_k)$ as k tends to infinity. We have $(\bar{X}^* \bar{Y}^*)_{12} = 0$. If we can show

---

[1] The proof follows that of Proposition 4.4 in [34].

that $(\bar{X}^*)_{12} = (\bar{Y}^*)_{12} = 0$, then we show that $(\bar{X}_k)_{12} = o(1)$, $(\bar{Y}_k)_{12} = o(1)$, and hence $(X_k)_{12} = o(\sqrt{\tau_k})$, $(Y_k)_{12} = o(\sqrt{\tau_k})$.

Now, if we premultiply and postmultiply $(\bar{X}^*\bar{Y}^*)_{12}$ by $(\bar{Y}^*)_{12}^T$ and $(\bar{Y}^*)_{22}^{-1}$ respectively, and express $(\bar{X}^*\bar{Y}^*)_{12}$ in terms of block components of $\bar{X}^*$ and $\bar{Y}^*$, we obtain

$$(\bar{Y}^*)_{12}^T(\bar{X}^*)_{11}(\bar{Y}^*)_{12}(\bar{Y}^*)_{22}^{-1} + (\bar{Y}^*)_{12}^T(\bar{X}^*)_{12} = 0.$$

Therefore, we have

$$\begin{aligned}
(\bar{X}^*)_{12} \bullet (\bar{Y}^*)_{12} &= -\mathrm{Tr}((\bar{Y}^*)_{12}^T(\bar{X}^*)_{11}(\bar{Y}^*)_{12}(\bar{Y}^*)_{22}^{-1}) \\
&= -\|(\bar{X}^*)_{11}^{1/2}(\bar{Y}^*)_{12}(\bar{Y}^*)_{22}^{-1/2}\|_F^2 \le 0.
\end{aligned} \tag{73}$$

On the other hand, by Remark 3, $(\bar{X}_k, \bar{Y}_k)$ satisfies

$$\bar{\mathcal{A}}(\tau_k)\mathrm{svec}(\bar{X}_k) + \bar{\mathcal{B}}(\tau_k)\mathrm{svec}(\bar{Y}_k) = q + \frac{\tau_k}{\tau_0}\bar{r}_{0,k}. \tag{74}$$

In the limit as k tends to infinity, from (74), we obtain

$$\begin{aligned}
&\bar{\mathcal{A}}(0)\mathrm{svec}(\bar{X}^*) + \bar{\mathcal{B}}(0)\mathrm{svec}(\bar{Y}^*) \\
&= q + \frac{1}{\tau_0}\bar{\mathcal{A}}(0)\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix} + \frac{1}{\tau_0}\bar{\mathcal{B}}(0)\mathrm{svec}\begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix},
\end{aligned}$$

where we apply the result in Proposition 17.
That is,

$$\bar{\mathcal{A}}(0)\mathrm{svec}\left(\bar{X}^* - \frac{1}{\tau_0}\begin{pmatrix} 0 & 0 \\ 0 & (X_0)_{22} \end{pmatrix}\right) + \bar{\mathcal{B}}(0)\mathrm{svec}\left(\bar{Y}^* - \frac{1}{\tau_0}\begin{pmatrix} (Y_0)_{11} & 0 \\ 0 & 0 \end{pmatrix}\right) = q.$$

Using the structure of q in Proposition 16 and that of $\bar{\mathcal{A}}(0)$, $\bar{\mathcal{B}}(0)$, the following holds:

$$\begin{aligned}
&\bar{\mathcal{A}}(0)\mathrm{svec}\begin{pmatrix} 0 & (\bar{X}^*)_{12} \\ (\bar{X}^*)_{12}^T & (\bar{X}^*)_{22} - \frac{1}{\tau_0}(X_0)_{22} \end{pmatrix} \\
&+ \bar{\mathcal{B}}(0)\mathrm{svec}\begin{pmatrix} (\bar{Y}^*)_{11} - \frac{1}{\tau_0}(Y_0)_{11} & (\bar{Y}^*)_{12} \\ (\bar{Y}^*)_{12}^T & 0 \end{pmatrix} = 0.
\end{aligned}$$

By the monotonicity result in Proposition 15, we have

$$\begin{pmatrix} 0 & (\bar{X}^*)_{12} \\ (\bar{X}^*)_{12}^T & (\bar{X}^*)_{22} - \frac{1}{\tau_0}(X_0)_{22} \end{pmatrix} \bullet \begin{pmatrix} (\bar{Y}^*)_{11} - \frac{1}{\tau_0}(Y_0)_{11} & (\bar{Y}^*)_{12} \\ (\bar{Y}^*)_{12}^T & 0 \end{pmatrix} \ge 0.$$

Hence, $(\bar{X}^*)_{12} \bullet (\bar{Y}^*)_{12} \ge 0$. Combined with (73), we conclude from $\|(\bar{X}^*)_{11}^{1/2}(\bar{Y}^*)_{12}$ $(\bar{Y}^*)_{22}^{-1/2}\|_F^2 = 0$ that $(\bar{Y}^*)_{12} = 0$.

In a similar fashion, we can show that $(\bar{X}^*)_{12} = 0$. $\square$

References

1. Alizadeh, F., Haeberly, J.A., Overton, M.: Primal–dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. SIAM J. Optim. **8**, 746–768 (1998)
2. Alzalg, B.: A primal–dual interior-point method based on various selections of displacement step for symmetric optimization. Comput. Optim. Appl. **72**, 363–390 (2019)
3. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
4. Chua, S.K., Toh, K.-C., Zhao, G.: An analytic center cutting plane method with deep cuts for semidefinite feasibility problems. J. Optim. Theory Appl. **123**, 291–318 (2004)
5. da Cruz Neto, J.X., Ferreira, O.P., Monteiro, R.D.C.: Asymptotic behavior of the central path for a special class of degenerate SDP problems. Math. Program. Ser. A **103**, 487–514 (2005)
6. Faybusovich, L., Zhou, C.: Long-step path-following algorithm for solving symmetric programming problems with nonlinear objective functions. Comput. Optim. Appl. **72**, 769–795 (2019)
7. Gahinet, P., Nemirovski, A.: The projective method for solving linear matrix inequalities. Math. Program. Ser. B **77**, 163–190 (1997)
8. Goemans, M., Williamson, D.: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J. ACM **42**, 1115–1145 (1995)
9. Goffin, J.-L., Luo, Z., Ye, Y.: Complexity analysis of an interior cutting plane method for convex feasibility problems. SIAM J. Optim. **6**, 638–652 (1996)
10. Gu, G., Zangiabadi, M., Roos, C.: Full Nesterov–Todd step infeasible interior-point method for symmetric optimization. Eur. J. Oper. Res. **214**, 473–484 (2011)
11. Helmberg, C., Rendl, F., Vanderbei, R.J., Wolkowicz, H.: An interior-point method for semidefinite programming. SIAM J. Optim. **6**, 342–361 (1996)
12. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
13. Kojima, M., Shida, M., Shindoh, S.: Local convergence of predictor–corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Math. Program. Ser. A **80**, 129–160 (1998)
14. Kojima, M., Shindoh, S., Hara, S.: Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. SIAM J. Optim. **7**, 86–125 (1997)
15. Lu, Z., Monteiro, R.D.C.: Error bounds and limiting behavior of weighted paths associated with the SDP map $X^{1/2}SX^{1/2}$. SIAM J. Optim. **15**, 348–374 (2004)
16. Luo, Z.-Q., Sturm, J.F., Zhang, S.: Superlinear convergence of a symmetric primal–dual path following algorithm for semidefinite programming. SIAM J. Optim. **8**, 59–81 (1998)
17. Monteiro, R.D.C.: Primal–dual path following algorithms for semidefinite programming. SIAM J. Optim. **7**, 663–678 (1997)
18. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of a new family of primal–dual algorithms for semidefinite programming. SIAM J. Optim. **9**, 551–577 (1999)
19. Monteiro, R.D.C., Zhang, Y.: A unified analysis for a class of long-step primal–dual path-following interior-point algorithms for semidefinite programming. Math. Program. **81**, 281–299 (1998)
20. Nesterov, Y., Todd, M.: Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. **22**, 1–42 (1997)
21. Nesterov, Y., Todd, M.: Primal–dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**, 324–364 (1998)
22. Nesterov, Y., Tunçel, L.: Local superlinear convergence of polynomial-time interior-point methods for hyperbolicity cone optimization problems. SIAM J. Optim. **26**, 139–170 (2016)
23. Potra, F.A.: Q-superlinear convergence of the iterates in primal–dual interior-point methods. Math. Program. Ser. A **91**, 99–115 (2001)
24. Potra, F.A., Sheng, R.: Superlinear convergence of interior-point algorithms for semidefinite programming. J. Optim. Theory Appl. **99**, 103–119 (1998)
25. Potra, F.A., Sheng, R.: A superlinearly convergent primal–dual infeasible-interior-point algorithm for semidefinite programming. SIAM J. Optim. **8**, 1007–1028 (1998)
26. Preiß, M., Stoer, J.: Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems. Math. Program. Ser. A **99**, 499–520 (2004)
27. Rangarajan, B.: Polynomial convergence of infeasible-interior-point methods over symmetric cones. SIAM J. Optim. **16**, 1211–1229 (2006)
28. Rigó, P.R., Darvay, Z.: Infeasible interior-point method for symmetric optimization using a positive-asymptotic barrier. Comput. Optim. Appl. **71**, 483–508 (2018)

29. Roos, C.: A full-Newton step $\mathcal{O}(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. **16**, 1110–1136 (2006)
30. Schmieta, S., Alizadeh, F.: Extension of primal–dual interior point algorithm to symmetric cones. Math. Program. **96**, 409–438 (2003)
31. Shilon, O.: RandOrthMat. https://uk.mathworks.com/matlabcentral/fileexchange/11783-randorthmat (2006)
32. Sim, C.-K.: On the analyticity of underlying HKM paths for monotone semidefinite linear complementarity problems. J. Optim. Theory Appl. **141**, 193–215 (2009)
33. Sim, C.-K.: Asymptotic behavior of underlying NT paths in interior point methods for monotone semidefinite linear complementarity problems. J. Optim. Theory Appl. **148**, 79–106 (2011)
34. Sim, C.-K.: Superlinear convergence of an infeasible predictor–corrector path-following interior point algorithm for a semidefinite linear complementarity problem using the Helmberg–Kojima–Monteiro direction. SIAM J. Optim. **21**, 102–126 (2011)
35. Sim, C.-K., Zhao, G.: Underlying paths in interior point methods for the monotone semidefinite linear complementarity problem. Math. Program. Ser. A **110**, 475–499 (2007)
36. Sim, C.-K., Zhao, G.: Asymptotic behavior of Helmberg–Kojima–Monteiro (HKM) paths in interior point methods for monotone semidefinite linear complementarity problem: general theory. J. Optim. Theory Appl. **137**, 11–25 (2008)
37. Sturm, J.F.: Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones (Updated for Version 1.05) (1998–2001)
38. Sun, J., Toh, K.-C., Zhao, G.: An analytic center cutting plane method for semidefinite feasibility problems. Math. Oper. Res. **27**, 332–346 (2002)
39. Todd, M.J., Toh, K.C., Tütüncü, R.H.: On the Nesterov–Todd direction in semidefinite programming. SIAM J. Optim. **8**, 769–796 (1998)
40. Toh, K.-C., Todd, M.J., Tütüncü, R.H.: On the implementation and usage of SDPT3—a Matlab software package for semidefinite-quadratic-linear programming, Version 4.0. In: Anjos, M.F., Lasserre, J.B. (eds.) Handbook on Semidefinite, Conic and Polynomial Optimization, Vol. 166 of International Series in Operations Research & Management Science, pp. 715–754. Springer, Berlin (2012)
41. Toh, K.-C., Zhao, G., Sun, J.: A multiple-cut analytic center cutting plane method for semidefinite feasibility problems. SIAM J. Optim. **12**, 1126–1146 (2002)
42. Tseng, P.: Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP. Optim. Methods Softw. **9**, 245–268 (1998)
43. Wang, G.Q., Bai, Y.Q., Gao, X.Y., Wang, D.Z.: Improved complexity analysis of full Nesterov–Todd step interior-point methods for semidefinite optimization. J. Optim. Theory Appl. **165**, 242–262 (2015)
44. Wang, G.Q., Kong, L.C., Tao, J.Y., Lesaja, G.: Improved complexity analysis of full Nesterov–Todd step feasible interior-point method for symmetric optimization. J. Optim. Theory Appl. **166**, 588–604 (2015)
45. Zhang, Y.: On extending some primal–dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optim. **8**, 365–386 (1998)