Superlinear Convergence of Infeasible Predictor-Corrector Path-Following Interior Point Algorithm for SDLCP using the HKM Direction

Chee-Khian Sim
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong
Email: macksim@inet.poly.edu.hk

December 6, 2009

Abstract

Interior point method (IPM) defines a search direction at each interior point of a region. These search directions form a direction field which in turn gives rise to a system of ordinary differential equations (ODEs). The solutions of the system of ODEs can be viewed as underlying paths in the interior of the region. In [31], these off-central paths are shown to be well-defined analytic curves and any of their accumulation points is a solution to a given monotone semi-definite linear complementarity problem (SDLCP). The study of these paths provides a way to understand how iterates generated by an interior point algorithm behave. In this paper, we give a weak sufficient condition using these off-central paths that guarantees superlinear convergence of a predictor-corrector path-following interior point algorithm for SDLCP using the HKM direction. This sufficient condition is implied by a currently known sufficient condition for superlinear convergence. Using this sufficient condition, we show that for any linear semi-definite feasibility problem, superlinear convergence using the interior point algorithm, with the HKM direction, can be achieved, for a suitable starting point. We work under the assumption of strict complementarity.

Keywords: Semi-definite linear complementarity problem; Linear semi-definite feasibility problem; Interior point method; Superlinear convergence; HKM direction.

AMS Subject Classification: 90C51; 90C22; 34C05

1 Introduction.

The notion of a central path is introduced by Sonnevend [34] in 1985 to interior point method (IPM). Since then, researchers realize that IPM is actually a homotopy method following underlying paths and that many remarkable properties of IPM are attributed to the nice geometry of these paths. Readers who are interested to know more about basic geometry of these paths may refer to [3].

An important role the underlying paths play in the study of IPM is to show its fast local convergence. The classical proof of local convergence of an iterative method, such as the Newton's method, for finding the solution of a system of equations relies on the nonsingularity of the Jacobian matrix. However, the Jacobian matrix of the equation system defining the search direction in IPM may be singular at the optimal solution. Thus, traditional approach of local convergence analysis does not work for IPM. Studying underlying paths which mimic the behavior of iterates generated by an interior point algorithm, especially when the iterates are close together at a solution of the given problem, provides an alternative approach. Fast local convergence of IPM has been successfully proved by relating it to the boundedness of derivatives of the underlying paths in [24, 37, 43, 44]. See also [35, 36].

Study of fast local convergence is particularly important for the class of monotone semi-definite linear complementarity problems (SDLCPs)[1]. This is because, in contrast to a monotone linear complementarity problem (LCP), the exact solution of a SDLCP cannot be obtained from an approximate solution by determining a complementary basis.

The convergence analysis using a path-following interior point algorithm on semi-definite programs (SDPs), and hence on SDLCPs, is considered to be more difficult than on linear programs (LPs). This arises mainly due to the difficulty in maintaining symmetry in the linearized complementarity [45]. Researchers working in the IPM area have proposed ways to overcome this problem, which result in different symmetrized search directions [2, 12, 16, 25, 27, 29, 30, 42], along which iterates generated by interior point algorithms move. Among these search directions, the HKM and NT directions have been implemented in existing SDP solvers.

There are various ways in which underlying paths, using these different search directions, for SDLCPs are defined in the literature [5, 17, 19, 23, 31, 41]. Paths arising from different search directions are likely to behave differently from each other. In [31], a definition of underlying paths for SDLCPs, using ordinary differential equations (ODEs), is proposed. The motivation for defining paths in this way is to relate these paths to the vector field of search directions of IPM.

The study of these paths provides a viable way to understand how iterates generated by an interior point algorithm behave. It has been shown that off-central paths corresponding to the AHO direction are analytic at solutions of a SDLCP [19, 23]. Superlinear convergence of iterates generated by existing interior point algorithms using the AHO direction (which does not, say, perform "narrowing" of the neighborhood) is hence possible [2, 15, 18, 22]. On the other hand, for the HKM direction, conditions [17, 32, 33, 41] are needed to ensure that off-central paths are analytic at solutions of a SDLCP, for different definitions of these paths. Superlinear convergence of iterates generated by an interior point algorithm using the HKM

[1]The class of semi-definite programs (SDPs) is a special, but important subclass of the class of SDLCPs.

direction can only be proven under further modifications of the algorithm, for example, by performing "narrowing" of the neighborhood [14]. Other sufficient conditions for superlinear convergence of algorithm have been suggested in [17, 20, 21].

Under strict complementarity assumption, in this paper, we give a sufficient condition that ensure superlinear convergence of an infeasible predictor-corrector primal-dual path-following interior point algorithm for HKM direction, using off-central paths. We do this in Section 4. In Section 2, we give precise definition of a SDLCP, show how a SDP can be written as a SDLCP, and also define the off-central path which we are studying in this paper. In Section 3, we describe a transformation on the ODE system that defines an off-central path. We end up with a transformed ODE system that we use in Section 4. In Section 5, we show that the class of linear semi-definite feasibility problems[2] always enjoy superlinear convergence with the infeasible predictor-corrector primal-dual path-following interior point algorithm without further modifications to the algorithm, for suitable starting points. We do this using the sufficient condition derived in Section 4.

1.1 Notations and Common Definitions.

The space of symmetric $n \times n$ matrices is denoted by S^n. Given matrices X and Y in $\Re^{p\times q}$, the standard inner product is defined by $X \bullet Y \equiv \mathrm{Tr}(X^T Y)$, where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix. If $X \in S^n$ is positive semi-definite (resp., positive definite), we write $X \succeq 0$ (resp., $X \succ 0$). The cone of positive semi-definite (resp., positive definite) symmetric matrices is denoted by S^n_+ (resp., S^n_{++}). The identity matrix is denoted by I_n, where n stands for the size of the matrix. In case the subscript is absent from I, the size of the identity matrix should be clear from the context.

$\| \cdot \|$ for a vector in $\Re^n$ refers to its Euclidean norm, and for a matrix in $\Re^{p\times q}$, it refers to its Frobenius norm.

For a matrix $X \in \Re^{p\times q}$, we denote its component at the i^{th} row and j^{th} column by X_{ij}. Also, $X_{i\cdot}$ denotes the i^{th} row of X, and $X_{\cdot j}$ the j^{th} column of X. In case X is partitioned into blocks of submatrices, then X_{ij} refers to the submatrix in the corresponding (i, j) position.

Given square matrices $A_i \in \Re^{n_i \times n_i}$, $i = 1, \ldots, m$, $\mathrm{Diag}(A_1, \ldots, A_m)$ is a square matrix with A_i as its diagonal blocks arranged in accordance to the way they are lined up in $\mathrm{Diag}(A_1, \ldots, A_m)$. All the other entries in $\mathrm{Diag}(A_1, \ldots, A_m)$ are taken to be zero.

Given functions $f : \Omega \longrightarrow E$ and $g : \Omega \longrightarrow \Re_{++}$, where Ω is an arbitrary set and E is a normed vector space, and a subset $\widetilde{\Omega} \subseteq \Omega$. We write $f(w) = O(g(w))$ for all $w \in \widetilde{\Omega}$ to mean that $\|f(w)\| \leq Mg(w)$ for all $w \in \widetilde{\Omega}$, where $M > 0$ is a positive constant. Suppose

[2]Note that the class of linear semi-definite feasibility problems is an important class of problems that has wide applicability in many areas, for example, in control theory [4], among others. A projective method is used in [8] to solve a linear semi-definite feasibility problem. This method is extended to solve the convex feasibility problem in [1] and the conic feasibility problem in [10]. Cutting plane method is also used to solve the convex feasibility problem, in particular, the semi-definite feasibility problem in [6, 9, 38, 40]. In all these works, the assumption that the interior of the feasible region is nonempty is always made. In this paper, we do not need such an assumption to show superlinear convergence using the existing infeasible predictor-corrector primal-dual path-following interior point algorithm on a linear semi-definite feasibility problem. Only strict complementarity assumption and a suitable initial point are needed.

we have $E = S^n_{++}$. Then we write $f(w) = \Theta(g(w))$ for all $w \in \widetilde{\Omega}$ if $f(w) = \mathcal{O}(g(w))$ and $f(w)^{-1} = \mathcal{O}(g(w))$ for all $w \in \widetilde{\Omega}$. What the subset $\widetilde{\Omega}$ is should be clear from the context. Usually, $\widetilde{\Omega} = (0, \bar{w})$ for a small $\bar{w} > 0$. Given $\widetilde{\Omega} = (0, \bar{w})$, we write $f(w) = o(g(w))$ to mean that $\|f(w)\|/g(w) \rightarrow 0$, as $w \rightarrow 0$.

2 Definitions of a SDLCP and an Off-central Path.

Let us consider the following system defined by:

$$XY = 0, \tag{1}$$

$$\mathcal{A}(X) + \mathcal{B}(Y) = q, \tag{2}$$

$$X, Y \in \mathcal{S}^n_+, \tag{3}$$

where $\mathcal{A}, \mathcal{B} : \mathcal{S}^n \longrightarrow \Re^{\tilde{n}}$ are linear operators mapping $\mathcal{S}^n$ to the space $\Re^{\tilde{n}}$, where $\tilde{n} := n(n+1)/2$. Hence $\mathcal{A}$ and $\mathcal{B}$ have the form $\mathcal{A}(X) = (A_1 \bullet X, \ldots, A_{\tilde{n}} \bullet X)^T$, resp. $\mathcal{B}(Y) = (B_1 \bullet Y, \ldots, B_{\tilde{n}} \bullet Y)^T$, where $A_i, B_i \in \mathcal{S}^n$ for all $i = 1, \ldots, \tilde{n}$. Also, $q \in \Re^{\tilde{n}}$.

The semi-definite linear complementarity problem (SDLCP) is to find a solution to system (1)-(3). We also called the system (1)-(3) SDLCP.

We have the following assumptions on system (1)-(3) throughout the paper:

Assumption 2.1 *(a) SDLCP (1)-(3) is monotone, i.e.* $\mathcal{A}(X) + \mathcal{B}(Y) = 0$ *for* $X, Y \in \mathcal{S}^n \Rightarrow X \bullet Y \geq 0$.

(b) There exists at least one solution to SDLCP (1)-(3).

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

The first assumption is satisfied for the class of semi-definite programs (SDPs), with equality for $X \bullet Y$, instead of inequality. The last assumption is a technical assumption that can be satisfied for any given SDP.

A SDP in its primal and dual form is given by:

$$(\mathcal{P}) \quad \begin{array}{ll} \min & C \bullet X \\ \text{subject to} & A_i \bullet X = b_i, \quad i = 1, \ldots, m, \\ & X \in S^n_+ \end{array}$$

.

$$(\mathcal{D}) \quad \begin{array}{ll} \max & \sum_{i=1}^m b_i y_i \\ \text{subject to} & \sum_{i=1}^m y_i A_i + Y = C, \\ & Y \in S^n_+ \end{array}$$

It is without loss of generality that we assume $A_i, i = 1, \ldots, m$, are linearly independent.

Written in the form (2), in this case, $A_i = 0, i = m+1, \ldots, \tilde{n}$, while $B_i = 0, i = 1, \ldots, m$, and $B_i, i = m+1, \ldots, \tilde{n}$, linearly independent, are obtained from the subspace in S^n orthogonal to the space spanned by $A_i, i = 1, \ldots, m$.

Let us now define the infeasible off-central path for SDLCP passing through a point (X_0, Y_0), $X_0, Y_0 \succ 0$.

Definition 2.1 *The solution* $(X(\mu), Y(\mu))$, $X(\mu)$, $Y(\mu) \succ 0$, *where* $\mu > 0$, *to*

$$H_P(XY' + X'Y) = \frac{1}{\mu} H_P(XY), \quad (4)$$

$$\mathcal{A}(X') + \mathcal{B}(Y') = \frac{1}{\mu}(\mathcal{A}(X) + \mathcal{B}(Y) - q), \quad (5)$$

with the initial condition $(X(\bar{\mu}_0), Y(\bar{\mu}_0)) = (X_0, Y_0)$, X_0, $Y_0 \succ 0$, $\bar{\mu}_0 = Tr(X_0 Y_0)/n$, *is the infeasible* **off-central path** *for SDLCP, corresponding to* P, *passing through* (X_0, Y_0).
Here, $H_P(U) := \frac{1}{2}(PUP^{-1} + (PUP^{-1})^T)$, *and* $P \in \Re^{n \times n}$ *is an invertible matrix.*

We use $\bar{\mu}_0$ to denote the initial starting point for μ in the ODE system (4), (5) to distinguish it from μ_0 which is the duality gap divided by n for the initial iterate using Algorithm 4.1.

Assuming P is an analytic function of X, Y and $PXYP^{-1}$ is always symmetric (such P include the well-known directions like the HKM (and its dual) and NT directions), it is proved in [31] that when $\mathcal{A}(X_0)+\mathcal{B}(Y_0) = q$, the off-central path, $(X(\mu), Y(\mu))$, is well-defined, unique, analytic over $\mu \in (0, \infty)$, and any of its accumulation points is a solution to (1)-(3). It is easy to see that these also hold when we consider infeasible[3] off-central path, when $\mathcal{A}(X_0)+\mathcal{B}(Y_0) \neq q$.

Remark 2.1 *Due to linearity, (5) can also be written as*

$$\mathcal{A}(X') + \mathcal{B}(Y') = r_0/\bar{\mu}_0, \quad (6)$$

where r_0 *is given by*

$$\mathcal{A}(X_0) + \mathcal{B}(Y_0) - q.$$

Hence, we have $(X(\mu), Y(\mu))$ *satisfies*

$$\mathcal{A}(X) + \mathcal{B}(Y) = q + \mu r_0/\bar{\mu}_0. \quad (7)$$

Remark 2.2 *It is easy to see, using (4), that the parameter* μ *in the ODE system (4)-(5) (or (4), (6)) is actually the duality gap,* $X(\mu) \bullet Y(\mu)$, *divided by* n, *at the point* $(X(\mu), Y(\mu))$ *on the path.*

3 An Investigation on ODE System (4), (6).

In this paper, we consider only the (dual) HKM search direction, where $P = Y^{1/2}$. Hence, (4) and (6) can be written as:

$$H_{Y^{1/2}}(XY' + X'Y) = \frac{1}{\mu} H_{Y^{1/2}}(XY), \quad (8)$$

$$\mathcal{A}(X') + \mathcal{B}(Y') = r_0/\bar{\mu}_0, \quad (9)$$

where $r_0 = \mathcal{A}(X_0) + \mathcal{B}(Y_0) - q.$ and $\mu > 0$.

Written in matrix-vector form, the above ODE system (8), (9) can be rewritten as:

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X \otimes_s Y^{-1} \end{pmatrix} \begin{pmatrix} \text{svec}(X') \\ \text{svec}(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} \mu r_0/\bar{\mu}_0 \\ \text{svec}(X) \end{pmatrix}. \quad (10)$$

[3] From now onwards, we omit the word "infeasible" when we mention off-central path.

The operation $\otimes_s$ and the map "svec" are used extensively in this paper. For their definitions and properties, the reader can refer to pp. $775 - 776$ and the appendix of [39].

The matrices $\mathcal{A}$ and $\mathcal{B}$ are derived from the operators A and B, respectively, and are given by:

$$\mathcal{A} = \begin{pmatrix} \text{svec}(A_1)^T \\ \vdots \\ \text{svec}(A_{\tilde{n}})^T \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} \text{svec}(B_1)^T \\ \vdots \\ \text{svec}(B_{\tilde{n}})^T \end{pmatrix}. \tag{11}$$

Note that in case of a SDP, $A_i = 0, i = m+1, \ldots, \tilde{n}$, and $B_i = 0, i = 1, \ldots, m$. Hence, for a SDP, we write $\mathcal{A}$ and $\mathcal{B}$ as

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_1 \\ 0 \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} 0 \\ \mathcal{B}_1 \end{pmatrix}, \tag{12}$$

where $\mathcal{A}_1$ consists of m rows and has full row rank, while $\mathcal{B}_1$ consists of $\tilde{n} - m$ rows and also has full row rank.

We perform block gaussian elimination operation on $(\mathcal{A}\ \mathcal{B})$, where $\mathcal{A}$ and $\mathcal{B}$ are given by (11), to obtain

$$\begin{pmatrix}
\text{svec}\begin{pmatrix} (A_1)_{11} & (A_1)_{12} \\ (A_1)_{12}^T & (A_1)_{22} \end{pmatrix}^T & \text{svec}\begin{pmatrix} (B_1)_{11} & (B_1)_{12} \\ (B_1)_{12}^T & (B_1)_{22} \end{pmatrix}^T \\
\vdots & \vdots \\
\text{svec}\begin{pmatrix} (A_{i_1})_{11} & (A_{i_1})_{12} \\ (A_{i_1})_{12}^T & (A_{i_1})_{22} \end{pmatrix}^T & \text{svec}\begin{pmatrix} (B_{i_1})_{11} & (B_{i_1})_{12} \\ (B_{i_1})_{12}^T & (B_{i_1})_{22} \end{pmatrix}^T \\
\text{svec}\begin{pmatrix} 0 & (A_{i_1+1})_{12} \\ (A_{i_1+1})_{12}^T & (A_{i_1+1})_{22} \end{pmatrix}^T & \text{svec}\begin{pmatrix} (B_{i_1+1})_{11} & (B_{i_1+1})_{12} \\ (B_{i_1+1})_{12}^T & 0 \end{pmatrix}^T \\
\vdots & \vdots \\
\text{svec}\begin{pmatrix} 0 & (A_{i_1+i_2})_{12} \\ (A_{i_1+i_2})_{12}^T & (A_{i_1+i_2})_{22} \end{pmatrix}^T & \text{svec}\begin{pmatrix} (B_{i_1+i_2})_{11} & (B_{i_1+i_2})_{12} \\ (B_{i_1+i_2})_{12}^T & 0 \end{pmatrix}^T \\
\text{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{i_1+i_2+1})_{22} \end{pmatrix}^T & \text{svec}\begin{pmatrix} (B_{i_1+i_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \\
\vdots & \vdots \\
\text{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{\tilde{n}})_{22} \end{pmatrix}^T & \text{svec}\begin{pmatrix} (B_{\tilde{n}})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T
\end{pmatrix}. \tag{13}$$

In case of a SDP, to make use of the special structure of $\mathcal{A}$ and $\mathcal{B}$ as in (12), instead of performing block gaussian elimination on $(\mathcal{A}\ \mathcal{B})$, we perform block gaussian elimination on $\mathcal{A}_1$

and $\mathcal{B}_1$ individually to obtain

$$\left(\begin{array}{l} \text{svec}\begin{pmatrix} (A_1)_{11} & (A_1)_{12} \\ (A_1)_{12}^T & (A_1)_{22} \end{pmatrix}^T \\ \vdots \\ \text{svec}\begin{pmatrix} (A_{j_1})_{11} & (A_{j_1})_{12} \\ (A_{j_1})_{12}^T & (A_{j_1})_{22} \end{pmatrix}^T \\ \text{svec}\begin{pmatrix} 0 & (A_{j_1+1})_{12} \\ (A_{j_1+1})_{12}^T & (A_{j_1+1})_{22} \end{pmatrix}^T \\ \vdots \\ \text{svec}\begin{pmatrix} 0 & (A_{j_1+j_2})_{12} \\ (A_{j_1+j_2})_{12}^T & (A_{j_1+j_2})_{22} \end{pmatrix}^T \\ \text{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{j_1+j_2+1})_{22} \end{pmatrix}^T \\ \vdots \\ \text{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_m)_{22} \end{pmatrix}^T \end{array}\right), \quad \left(\begin{array}{l} \text{svec}\begin{pmatrix} (B_1)_{11} & (B_1)_{12} \\ (B_1)_{12}^T & (B_1)_{22} \end{pmatrix}^T \\ \vdots \\ \text{svec}\begin{pmatrix} (B_{k_1})_{11} & (B_{k_1})_{12} \\ (B_{k_1})_{12}^T & (B_{k_1})_{22} \end{pmatrix}^T \\ \text{svec}\begin{pmatrix} (B_{k_1+1})_{11} & (B_{k_1+1})_{12} \\ (B_{k_1+1})_{12}^T & 0 \end{pmatrix}^T \\ \vdots \\ \text{svec}\begin{pmatrix} (B_{k_1+k_2})_{11} & (B_{k_1+k_2})_{12} \\ (B_{k_1+k_2})_{12}^T & 0 \end{pmatrix}^T \\ \text{svec}\begin{pmatrix} (B_{k_1+k_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \\ \vdots \\ \text{svec}\begin{pmatrix} (B_{\tilde{n}-m})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \end{array}\right), \tag{14}$$

respectively.

Henceforth, we assume the above form (13) for $(\mathcal{A}\ \mathcal{B})$. In case of a SDP, $(\mathcal{A}\ \mathcal{B})$ is of the form (12) with $\mathcal{A}_1$ and $\mathcal{B}_1$ taking the form in (14).

To analyze superlinear convergence, besides Assumptions 2.1(a)-(c), we need an additional assumption as stated below:

Assumption 3.1 *There exists a strictly complementary solution (X^*, Y^*) to the SDLCP (1)-(3).*

Assumption 3.1 is currently needed in the literature to show superlinear convergence using an interior point algorithm on a SDLCP or SDP.

Henceforth, we assume that Assumptions 2.1(a)-(c) and 3.1 always hold in this paper.

Since X^* and Y^* commute, they are jointly diagonalizable by some orthogonal matrix Q. So, using this orthogonal similarity transformation of the matrices in the SDLCP (1)-(3), we may assume without loss of generality, that

$$X^* = \begin{pmatrix} \Lambda_{11}^* & 0 \\ 0 & 0 \end{pmatrix}, \qquad Y^* = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_{22}^* \end{pmatrix}, \tag{15}$$

where $\Lambda_{11}^* = \text{Diag}(\lambda_1^*, \ldots, \lambda_{k_0}^*) \succ 0$ and $\Lambda_{22}^* = \text{Diag}(\lambda_{k_0+1}^*, \ldots, \lambda_n^*) \succ 0$. Here, $\lambda_1^*, \ldots, \lambda_n^*$ are real numbers greater than zero.

Note that similar conclusions in this paper also holds for SDLCP (1)-(3) with (X^*, Y^*) not of the form (15), although all our following discussions are on SDLCP (1)-(3) with (X^*, Y^*) of the form (15).

Hereafter, whenever we partition a matrix $S \in S^n$, we do it in a similar way; i.e., S is always partitioned as $\begin{pmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{pmatrix}$, where $S_{11} \in S^{k_0}$, $S_{22} \in S^{n-k_0}$ and $S_{12} \in \Re^{k_0 \times (n-k_0)}$.

First, let us make a few observations on the off-central paths defined in Definition 2.1, in the following propositions.

Proposition 3.1 *[[23], Lemma 3.7] The set $\mathcal{U}$ defined by*

$$\begin{aligned}\{(X(\mu), Y(\mu)) \in S^n_{++} \times S^n_{++} \quad ; \quad & 0 < \mu \leq \bar{\mu}_0, \lambda_{\min}(XY)(\bar{\mu}_0) \geq D, \\ & \|X(\bar{\mu}_0)\| \leq C, \|Y(\bar{\mu}_0)\| \leq C\},\end{aligned}$$

is bounded, for fixed $\bar{\mu}_0, C, D > 0$.

Proof: Consider $(X(\mu), Y(\mu)) \in \mathcal{U}$, with $(X(\bar{\mu}_0), Y(\bar{\mu}_0)) = (X_0, Y_0)$.

We have by (4), $X(\mu) \bullet Y(\mu) = (\mu/\bar{\mu}_0) X(\bar{\mu}_0) \bullet Y(\bar{\mu}_0) = n\mu$.

Note that (X^*, Y^*) satisfies $A(X^*) + B(Y^*) = q$, and $(X(\mu), Y(\mu))$ satisfies $A(X(\mu)) + B(Y(\mu)) = q + (\mu/\bar{\mu}_0) r_0$.

Let

$$\hat{X}(\mu) = (1 - \mu/\bar{\mu}_0) X^* + (\mu/\bar{\mu}_0) X_0, \quad \hat{Y}(\mu) = (1 - \mu/\bar{\mu}_0) Y^* + (\mu/\bar{\mu}_0) Y_0.$$

Then

$$\begin{aligned} A(\hat{X}(\mu) - X(\mu)) + B(\hat{Y}(\mu) - Y(\mu)) &= (1 - \mu/\bar{\mu}_0) q + (\mu/\bar{\mu}_0)(q + r_0) - q - (\mu/\bar{\mu}_0) r_0 \\ &= 0. \end{aligned}$$

Hence, by Assumption 2.1(a),

$$(\hat{X}(\mu) - X(\mu)) \bullet (\hat{Y}(\mu) - Y(\mu)) \geq 0.$$

That is,

$$\begin{aligned} & \hat{X}(\mu) \bullet \hat{Y}(\mu) - \hat{X}(\mu) \bullet Y(\mu) - X(\mu) \bullet \hat{Y}(\mu) + X(\mu) \bullet Y(\mu) \\ = \; & (1 - \mu/\bar{\mu}_0)(\mu/\bar{\mu}_0) X^* \bullet Y_0 + (1 - \mu/\bar{\mu}_0)(\mu/\bar{\mu}_0) X_0 \bullet Y^* + n\mu^2/\bar{\mu}_0 X_0 \bullet Y_0 \\ & -(1 - \mu/\bar{\mu}_0) X^* \bullet Y(\mu) - (\mu/\bar{\mu}_0) X_0 \bullet Y(\mu) - (1 - \mu/\bar{\mu}_0) X(\mu) \bullet Y^* - (\mu/\bar{\mu}_0) X(\mu) \bullet Y_0 \\ & + X(\mu) \bullet Y(\mu) \\ \geq \; & 0. \end{aligned}$$

Therefore,

$$\begin{aligned} & (1 - \mu/\bar{\mu}_0)(X^* \bullet Y(\mu) + X(\mu) \bullet Y^*) + (\mu/\bar{\mu}_0)(X_0 \bullet Y(\mu) + X(\mu) \bullet Y_0) \\ \leq \; & (1 - \mu/\bar{\mu}_0)(\mu/\bar{\mu}_0)(X^* \bullet Y_0 + X_0 \bullet Y^*) + n\mu^2/\bar{\mu}_0 + X(\mu) \bullet Y(\mu) \\ \leq \; & (1 - \mu/\bar{\mu}_0)(\mu/\bar{\mu}_0)(X^* \bullet Y_0 + X_0 \bullet Y^*) + n\mu^2/\bar{\mu}_0 + n\mu. \end{aligned}$$

Hence,

$$\begin{aligned} & (1/\mu - 1/\bar{\mu}_0)(X^* \bullet Y(\mu) + X(\mu) \bullet Y^*) + (1/\bar{\mu}_0)(X_0 \bullet Y(\mu) + X(\mu) \bullet Y_0) \\ \leq \; & (1 - \mu/\bar{\mu}_0)(1/\bar{\mu}_0)(X^* \bullet Y_0 + X_0 \bullet Y^*) + n\mu/\bar{\mu}_0 + n \\ \leq \; & M_1, \end{aligned}$$

for $0 < \mu \le \bar{\mu}_0$, where $M_1 > 0$ depends only on $\bar{\mu}_0, C$.

Therefore,

$$X_0 \bullet Y(\mu) \le \bar{\mu}_0 M_1, \quad X(\mu) \bullet Y_0 \le \bar{\mu}_0 M_1.$$

We have, for $0 < \mu \le \bar{\mu}_0$,

$$\|X(\mu)\| \le \bar{\mu}_0 M_1/\lambda_{\min}(Y_0), \quad \|Y(\mu)\| \le \bar{\mu}_0 M_1/\lambda_{\min}(X_0).$$

Since $\|X_0\| \le C$ and $\|Y_0\| \le C$, together with $\lambda_{\min}(X_0 Y_0) \ge D$, we must have $\lambda_{\min}(X_0)$ and $\lambda_{\min}(Y_0)$ are uniformly bounded from below by a positive constant independent of $(X(\mu), Y(\mu))$ chosen from $\mathcal{U}$. Hence, we are done. **QED**

Proposition 3.2 *[[32], Lemma 2.2] $Y_{11}(\mu)$ and $X_{22}(\mu)$ are equal to $\mathcal{O}(\mu)$, and $\|X_{12}(\mu)\|$ and $\|Y_{12}(\mu)\|$ are equal to $\mathcal{O}(\sqrt{\mu})$, where the bounds are not dependent on any off-central path $(X(\mu), Y(\mu))$, as long as, $\lambda_{\min}(XY)(\bar{\mu}_0) \ge D$, $\|X(\bar{\mu}_0)\| \le C$, $\|Y(\bar{\mu}_0)\| \le C$, for fixed $\bar{\mu}_0, C, D > 0$.*

Proposition 3.3 *[[32], Lemma 2.3] $X_{11}(\mu)$ and $Y_{22}(\mu)$ are equal to $\Theta(1)$, and $X_{22}(\mu)$ and $Y_{11}(\mu)$ are equal to $\Theta(\mu)$, where the bounds are not dependent on any off-central path $(X(\mu), Y(\mu))$, as long as, $\lambda_{\min}(XY)(\bar{\mu}_0) \ge D$, $\|X(\bar{\mu}_0)\| \le C$, $\|Y(\bar{\mu}_0)\| \le C$, for fixed $\bar{\mu}_0, C, D > 0$.*

The proof of Proposition 3.2 and that of Proposition 3.3 are similar to that of Lemma 2.2 and that of Lemma 2.3 in [32], respectively, and hence will not be shown here again. See also Lemmas 3.10 and 3.11 of [23].

From these propositions, we have that

$$X(\mu) = \begin{pmatrix} \Theta(1) & O(\sqrt{\mu}) \\ O(\sqrt{\mu}) & \Theta(\mu) \end{pmatrix}, \quad Y(\mu) = \begin{pmatrix} \Theta(\mu) & O(\sqrt{\mu}) \\ O(\sqrt{\mu}) & \Theta(1) \end{pmatrix}.$$

Hence, we can write

$$X(\mu) = \begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu} I \end{pmatrix} \widetilde{X}(\mu) \begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu} I \end{pmatrix},$$
$$Y(\mu) = \begin{pmatrix} \sqrt{\mu} I & 0 \\ 0 & I \end{pmatrix} \widetilde{Y}(\mu) \begin{pmatrix} \sqrt{\mu} I & 0 \\ 0 & I \end{pmatrix},$$

where

$$\widetilde{X}(\mu) = \begin{pmatrix} \Theta(1) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}, \quad \widetilde{Y}(\mu) = \begin{pmatrix} \Theta(1) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}.$$

Before we go on, let us define new matrices as follows:

$$\bar{\mathcal{A}}(t) := \begin{pmatrix} \mathrm{svec}\begin{pmatrix} (A_1)_{11} & t(A_1)_{12} \\ t(A_1)_{12}^T & t^2(A_1)_{22} \end{pmatrix}^T \\ \vdots \\ \mathrm{svec}\begin{pmatrix} (A_{i_1})_{11} & t(A_{i_1})_{12} \\ t(A_{i_1})_{12}^T & t^2(A_{i_1})_{22} \end{pmatrix}^T \\ \mathrm{svec}\begin{pmatrix} 0 & (A_{i_1+1})_{12} \\ (A_{i_1+1})_{12}^T & t(A_{i_1+1})_{22} \end{pmatrix}^T \\ \vdots \\ \mathrm{svec}\begin{pmatrix} 0 & (A_{i_1+i_2})_{12} \\ (A_{i_1+i_2})_{12}^T & t(A_{i_1+i_2})_{22} \end{pmatrix}^T \\ \mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{i_1+i_2+1})_{22} \end{pmatrix}^T \\ \vdots \\ \mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{\tilde{n}})_{22} \end{pmatrix}^T \end{pmatrix},$$

$$\bar{\mathcal{B}}(t) := \begin{pmatrix} \mathrm{svec}\begin{pmatrix} t^2(B_1)_{11} & t(B_1)_{12} \\ t(B_1)_{12}^T & (B_1)_{22} \end{pmatrix}^T \\ \vdots \\ \mathrm{svec}\begin{pmatrix} t^2(B_{i_1})_{11} & t(B_{i_1})_{12} \\ t(B_{i_1})_{12}^T & (B_{i_1})_{22} \end{pmatrix}^T \\ \mathrm{svec}\begin{pmatrix} t(B_{i_1+1})_{11} & (B_{i_1+1})_{12} \\ (B_{i_1+1})_{12}^T & 0 \end{pmatrix}^T \\ \vdots \\ \mathrm{svec}\begin{pmatrix} t(B_{i_1+i_2})_{11} & (B_{i_1+i_2})_{12} \\ (B_{i_1+i_2})_{12}^T & 0 \end{pmatrix}^T \\ \mathrm{svec}\begin{pmatrix} (B_{i_1+i_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \\ \vdots \\ \mathrm{svec}\begin{pmatrix} (B_{\tilde{n}})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \end{pmatrix},$$

$$
\bar{\mathcal{A}}_1(t) := \begin{pmatrix}
\mathrm{svec}\begin{pmatrix} (A_1)_{11} & t(A_1)_{12} \\ t(A_1)_{12}^T & t^2(A_1)_{22} \end{pmatrix}^T \\
\vdots \\
\mathrm{svec}\begin{pmatrix} (A_{j_1})_{11} & t(A_{j_1})_{12} \\ t(A_{j_1})_{12}^T & t^2(A_{j_1})_{22} \end{pmatrix}^T \\
\mathrm{svec}\begin{pmatrix} 0 & (A_{j_1+1})_{12} \\ (A_{j_1+1})_{12}^T & t(A_{j_1+1})_{22} \end{pmatrix}^T \\
\vdots \\
\mathrm{svec}\begin{pmatrix} 0 & (A_{j_1+j_2})_{12} \\ (A_{j_1+j_2})_{12}^T & t(A_{j_1+j_2})_{22} \end{pmatrix}^T \\
\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_{j_1+j_2+1})_{22} \end{pmatrix}^T \\
\vdots \\
\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (A_m)_{22} \end{pmatrix}^T
\end{pmatrix},
$$

$$
\bar{\mathcal{B}}_1(t) := \begin{pmatrix}
\mathrm{svec}\begin{pmatrix} t^2(B_1)_{11} & t(B_1)_{12} \\ t(B_1)_{12}^T & (B_1)_{22} \end{pmatrix}^T \\
\vdots \\
\mathrm{svec}\begin{pmatrix} t^2(B_{k_1})_{11} & t(B_{k_1})_{12} \\ t(B_{k_1})_{12}^T & (B_{k_1})_{22} \end{pmatrix}^T \\
\mathrm{svec}\begin{pmatrix} t(B_{k_1+1})_{11} & (B_{k_1+1})_{12} \\ (B_{k_1+1})_{12}^T & 0 \end{pmatrix}^T \\
\vdots \\
\mathrm{svec}\begin{pmatrix} t(B_{k_1+k_2})_{11} & (B_{k_1+k_2})_{12} \\ (B_{k_1+k_2})_{12}^T & 0 \end{pmatrix}^T \\
\mathrm{svec}\begin{pmatrix} (B_{k_1+k_2+1})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \\
\vdots \\
\mathrm{svec}\begin{pmatrix} (B_{\tilde{n}-m})_{11} & 0 \\ 0 & 0 \end{pmatrix}^T
\end{pmatrix},
$$

where $t^2 = \mu$.

The following proposition relates the above defined new matrices with $\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}_1$ and $\mathcal{B}_1$:

Proposition 3.4

$$
\begin{aligned}
\mathcal{A}\left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) &= \mathrm{Diag}(I_{i_1}, tI_{i_2}, t^2 I_{\tilde{n}-i_1-i_2})\bar{\mathcal{A}}(t), \\
\mathcal{B}\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) &= \mathrm{Diag}(I_{i_1}, tI_{i_2}, t^2 I_{\tilde{n}-i_1-i_2})\bar{\mathcal{B}}(t),
\end{aligned}
$$

$$\mathcal{A}_1\left(\left(\begin{array}{cc} I & 0 \\ 0 & tI \end{array}\right)\otimes_s\left(\begin{array}{cc} I & 0 \\ 0 & tI \end{array}\right)\right) = \mathrm{Diag}(I_{j_1}, tI_{j_2}, t^2 I_{m-j_1-j_2})\bar{\mathcal{A}}_1(t),$$

$$\mathcal{B}_1\left(\left(\begin{array}{cc} tI & 0 \\ 0 & I \end{array}\right)\otimes_s\left(\begin{array}{cc} tI & 0 \\ 0 & I \end{array}\right)\right) = \mathrm{Diag}(I_{k_1}, tI_{k_2}, t^2 I_{\tilde{n}-m-k_1-k_2})\bar{\mathcal{B}}_1(t).$$

Proof: The proposition is clear from the definitions of the new matrices, and (13), (14). **QED**

Remark 3.1 *In case we are considering a SDP, we let*

$$\bar{\mathcal{A}}(t) := \left(\begin{array}{c} \bar{\mathcal{A}}_1(t) \\ 0 \end{array}\right), \quad \bar{\mathcal{B}}(t) := \left(\begin{array}{c} 0 \\ \bar{\mathcal{B}}_1(t) \end{array}\right),$$

then

$$\mathcal{A}\left(\left(\begin{array}{cc} I & 0 \\ 0 & tI \end{array}\right)\otimes_s\left(\begin{array}{cc} I & 0 \\ 0 & tI \end{array}\right)\right) = \mathrm{Diag}(I_{j_1}, tI_{j_2}, t^2 I_{m-j_1-j_2}, I_{k_1}, tI_{k_2}, t^2 I_{\tilde{n}-m-k_1-k_2})\bar{\mathcal{A}}(t),$$

$$\mathcal{B}\left(\left(\begin{array}{cc} tI & 0 \\ 0 & I \end{array}\right)\otimes_s\left(\begin{array}{cc} tI & 0 \\ 0 & I \end{array}\right)\right) = \mathrm{Diag}(I_{j_1}, tI_{j_2}, t^2 I_{m-j_1-j_2}, I_{k_1}, tI_{k_2}, t^2 I_{\tilde{n}-m-k_1-k_2})\bar{\mathcal{B}}(t).$$

Proposition 3.5 *We have*

$$q = \left(\begin{array}{c} \overline{q} \\ 0 \end{array}\right),$$

where $\overline{q} \in \Re^{i_1}$.
In case we are considering a SDP,

$$q = \left(\begin{array}{c} \hat{q} \\ 0 \\ \hat{\hat{q}} \\ 0 \end{array}\right),$$

where $\hat{q} \in \Re^{j_1}$, $\hat{\hat{q}} \in \Re^{k_1}$ *(whose first entry starts at the* $m+1$ *entry of* q*).*

Proof: Observe that $(X(\mu), Y(\mu))$ satisfies

$$\mathcal{A}\mathrm{svec}(X) + \mathcal{B}\mathrm{svec}(Y) = q + \mu r_0/\bar{\mu}_0.$$

Hence,

$$\mathrm{Diag}(I_{i_1}, tI_{i_2}, t^2 I_{\tilde{n}-i_1-i_2})(\bar{\mathcal{A}}(t)\mathrm{svec}(\widetilde{X}) + \bar{\mathcal{B}}(t)\mathrm{svec}(\widetilde{Y})) = q + t^2 r_0/\bar{\mu}_0, \tag{16}$$

where $\mu = t^2$.

We see from (16) by taking $t \rightarrow 0$ that besides the first i_1 entries, the rest of entries in q is zero.

In a similar fashion, for the case of a SDP, results in proposition follows from

$$\mathrm{Diag}(I_{j_1}, tI_{j_2}, t^2 I_{m-j_1-j_2}, I_{k_1}, tI_{k_2}, t^2 I_{\tilde{n}-m-k_1-k_2})(\bar{\mathcal{A}}(t)\mathrm{svec}(\widetilde{X}) + \bar{\mathcal{B}}(t)\mathrm{svec}(\widetilde{Y})) = q + t^2 r_0/\bar{\mu}_0.$$

QED

Now, inverse of the matrix on the extreme left in the ODE system (10) is given by

$$\begin{pmatrix} -(X \otimes_s Y^{-1})\mathcal{G}^{-1} & I + (X \otimes_s Y^{-1})\mathcal{G}^{-1}\mathcal{A} \\ \mathcal{G}^{-1} & -\mathcal{G}^{-1}\mathcal{A} \end{pmatrix},$$

where $\mathcal{G}(\mu) := \mathcal{B} - \mathcal{A}(X \otimes_s Y^{-1})$.

Therefore, (10) can be written as

$$\begin{pmatrix} \mathrm{svec}(X') \\ \mathrm{svec}(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} -\mu(X \otimes_s Y^{-1})\mathcal{G}^{-1}(r_0/\bar{\mu}_0) + \mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}\mathcal{A}\mathrm{svec}(X) \\ \mu\mathcal{G}^{-1}(r_0/\bar{\mu}_0) - \mathcal{G}^{-1}\mathcal{A}\mathrm{svec}(X) \end{pmatrix}, \quad (17)$$

where $\mu > 0$.

We have the following proposition:

Proposition 3.6

$$\begin{aligned} -\mu(X \otimes_s Y^{-1})\mathcal{G}^{-1}(r_0/\bar{\mu}_0) &+ \mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}\mathcal{A}\mathrm{svec}(X) \\ &= \frac{1}{2}((X \otimes_s Y^{-1})\mathcal{G}^{-1}(-\mu r_0/\bar{\mu}_0 + q) + \mathrm{svec}(X)), \\ \mu\mathcal{G}^{-1}(r_0/\bar{\mu}_0) - \mathcal{G}^{-1}\mathcal{A}\mathrm{svec}(X) &= \frac{1}{2}(\mathcal{G}^{-1}(\mu r_0/\bar{\mu}_0 - q) + \mathrm{svec}(Y)), \end{aligned}$$

where (X, Y) *satisfies* $\mathcal{A}\mathrm{svec}(X) + \mathcal{B}\mathrm{svec}(Y) = q + \mu r_0/\bar{\mu}_0$.

Proof: We need only show the first equality. The second equality can be shown in a similar manner.

The equality is shown by showing that

$$\mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}\mathcal{A}\mathrm{svec}(X) = \frac{1}{2}(\mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}(q + \mu r_0/\bar{\mu}_0)).$$

This is true since

$$\begin{aligned} & \mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}\mathcal{A}\mathrm{svec}(X) \\ = \; & \mathrm{svec}(X) + (\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})^{-1}\mathcal{A}\mathrm{svec}(X) \\ = \; & \frac{1}{2}(\mathrm{svec}(X) + (\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})^{-1}q - (\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})^{-1}\mathcal{B}\mathrm{svec}(Y) \\ & + \mathrm{svec}(X) + (\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})^{-1}\mathcal{A}\mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}(\mu r_0/\bar{\mu}_0)) \\ = \; & \frac{1}{2}(\mathrm{svec}(X) + (\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})^{-1}q - (\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})^{-1}(\mathcal{B}(X \otimes_s Y^{-1})^{-1} - \mathcal{A})\mathrm{svec}(X) \\ & + \mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}(\mu r_0/\bar{\mu}_0)) \\ = \; & \frac{1}{2}(\mathrm{svec}(X) + (X \otimes_s Y^{-1})\mathcal{G}^{-1}(q + \mu r_0/\bar{\mu}_0)), \end{aligned}$$

where the second equality follows from $\mathcal{A}\mathrm{svec}(X) + \mathcal{B}\mathrm{svec}(Y) = q + \mu r_0/\bar{\mu}_0$. **QED**

Using (17) and Proposition 3.6, (10) can be written as

$$\begin{pmatrix} \text{svec}(X') \\ \text{svec}(Y') \end{pmatrix} = \frac{1}{2\mu} \begin{pmatrix} (X \otimes_s Y^{-1})\mathcal{G}^{-1}(-\mu r_0/\bar{\mu}_0 + q) + \text{ svec}(X) \\ \mathcal{G}^{-1}(\mu r_0/\bar{\mu}_0 - q) + \text{ svec}(Y) \end{pmatrix}. \tag{18}$$

Observe using Propositions 3.4 and 3.5 (in case we are considering a SDP, we use Remark 3.1, instead of Proposition 3.4) that

$$\begin{aligned} (X \otimes_s Y^{-1})\mathcal{G}^{-1}(-\mu r_0/\bar{\mu}_0 + q) &= \left(\begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu} I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu} I \end{pmatrix} \right) (\widetilde{X} \otimes_s \widetilde{Y}^{-1})\widetilde{\mathcal{G}}^{-1}(-\sqrt{\mu}\widetilde{r}_0/\bar{\mu}_0 + q), \\ \mathcal{G}^{-1}(\mu r_0/\bar{\mu}_0 - q) &= \left(\begin{pmatrix} \sqrt{\mu} I & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} \sqrt{\mu} I & 0 \\ 0 & I \end{pmatrix} \right) \widetilde{\mathcal{G}}^{-1}(\sqrt{\mu}\widetilde{r}_0/\bar{\mu}_0 - q). \end{aligned}$$

Here, $\widetilde{\mathcal{G}}(\mu) := \bar{\mathcal{B}}(\sqrt{\mu}) - \bar{\mathcal{A}}(\sqrt{\mu})(\widetilde{X} \otimes_s \widetilde{Y}^{-1})$, and

$$\widetilde{r}_0 = \text{Diag}\left(\sqrt{\mu} I_{i_1}, I_{i_2}, \frac{1}{\sqrt{\mu}} I_{\tilde{n}-i_1-i_2} \right) r_0.$$

In case we are considering a SDP,

$$\widetilde{r}_0 = \text{Diag}\left(\sqrt{\mu} I_{j_1}, I_{j_2}, \frac{1}{\sqrt{\mu}} I_{m-j_1-j_2}, \sqrt{\mu} I_{k_1}, I_{k_2}, \frac{1}{\sqrt{\mu}} I_{\tilde{n}-m-k_1-k_2} \right) r_0. \tag{19}$$

Hence, (18) can be written as

$$\begin{aligned} &\begin{pmatrix} \text{svec}(X') \\ \text{svec}(Y') \end{pmatrix} \\ = \ & \frac{1}{2\mu} \begin{pmatrix} \left(\begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu} I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu} I \end{pmatrix} \right) ((\widetilde{X} \otimes_s \widetilde{Y}^{-1})\widetilde{\mathcal{G}}^{-1}(-\sqrt{\mu}\widetilde{r}_0/\bar{\mu}_0 + q) + \text{ svec}(\widetilde{X})) \\ \left(\begin{pmatrix} \sqrt{\mu} I & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} \sqrt{\mu} I & 0 \\ 0 & I \end{pmatrix} \right) (\widetilde{\mathcal{G}}^{-1}(\sqrt{\mu}\widetilde{r}_0/\bar{\mu}_0 - q) + \text{ svec}(\widetilde{Y})) \end{pmatrix}. \end{aligned} \tag{20}$$

We are going to use this ODE system (20) in our investigation on sufficient conditions for superlinear convergence using an interior point algorithm on a SDLCP.

4 Sufficient Conditions for Superlinear Convergence.

We consider an infeasible predictor-corrector primal-dual path-following interior point algorithm to solve a SDLCP. We only consider the (dual) HKM search direction, where $P = Y^{1/2}$, in this paper.

Notice that the infeasible central path, $(X^c(\mu), Y^c(\mu))$, $\mu > 0$, with $(X^c(\bar{\mu}_0), Y^c(\bar{\mu}_0)) = (X_0^c, Y_0^c)$, which is a special off-central path, satisfies:

$$XY = \mu I, \quad A(X) + B(Y) = q + (\mu/\bar{\mu}_0) r_0, \quad X, Y \in S_{++}^n, \tag{21}$$

where

$$X_0^c Y_0^c = \bar{\mu}_0 I, \quad X_0^c, Y_0^c \in S_{++}^n. \tag{22}$$

The existence of the infeasible central path is guaranteed, and has been discussed in [23, 26, 28].

We consider the following neighborhood of the infeasible central path:

$$\mathcal{N}(\beta, \tau) := \{(X, Y) \in S_{++}^n \times S_{++}^n;\ \|Y^{1/2} X Y^{1/2} - \tau I\| \leq \beta\tau\}.$$

Let us now describe the infeasible predictor-corrector primal-dual path-following interior point algorithm for the sake of completeness. The algorithm is from [21].

In the algorithm, solving the following linear system for $(U, V) \in S^n \times S^n$ plays an important role:

$$\begin{aligned} Y^{1/2}(XV + UY)Y^{-1/2} + Y^{-1/2}(VX + YU)Y^{1/2} &= 2(\sigma\tau I - Y^{1/2} X Y^{1/2}) && (23) \\ A(U) + B(V) &= -\bar{r}. && (24) \end{aligned}$$

The algorithm is described as follows:

Algorithm 4.1 *Choose $\beta_1 < \beta_2$, with $\beta_2^2/(2(1-\beta_2)^2) \leq \beta_1 < \beta_2 < \beta_2/(1-\beta_2) < 1$. Choose $(X_0, Y_0) \in \mathcal{N}(\beta_1, \tau_0)$ with $\tau_0 = \mu_0 = X_0 \bullet Y_0/n$, and set $\psi_0 = 1$. For $k = 0, 1, \ldots,$ do* (a1) *through* (a5)*:*

(*a*1) *Set $X = X_k$, $Y = Y_k$, $\psi = \psi_k$, $\tau = \tau_k$, and define*

$$r = A(X) + B(Y) - q.$$

(*a*2) *If $\max\{X \bullet Y, \|r\|\} \leq \epsilon$, then report (X, Y) as an approximate solution to (1)-(3), and terminate.*

(*a*3) *Find the solution U, V of the linear system (23), (24), with $\sigma = 0$, $\bar{r} = r$. Define*

$$\overline{X} = X + \overline{\alpha} U, \quad \overline{Y} = Y + \overline{\alpha} V,$$

where the steplength $\overline{\alpha}$ satisfies

$$\alpha_1 \leq \overline{\alpha} \leq \alpha_2. \tag{25}$$

Here,

$$\begin{aligned} \alpha_1 &= \frac{2}{\sqrt{1 + 4\delta/(\beta_2 - \beta_1)} + 1}, && (26) \\ \delta &= \frac{1}{\tau}\|Y^{1/2} U V Y^{-1/2}\|, && (27) \end{aligned}$$

and

$$\alpha_2 = \max\{\tilde{\alpha} \in [0, 1]\ ;\ (X + \alpha U, Y + \alpha V) \in \mathcal{N}(\beta_2, (1-\alpha)\tau)\ \forall\ \alpha \in [0, \tilde{\alpha}]\}.$$

Set $\psi_+ = (1 - \overline{\alpha})\psi$. If $\overline{\alpha} = 1$, then $(\overline{X}, \overline{Y})$ solves (1)-(3) and terminate.

($a4$) *Find the solution U_1, V_1 of the linear system (23), (24), with $\sigma = (1-\overline{\alpha})$, $\overline{r} = 0$. Set*

$$X_+ = \overline{X} + U_1, \quad Y_+ = \overline{Y} + V_1,$$
$$\tau_+ = (1-\overline{\alpha})\tau.$$

($a5$) *Set*

$$X_{k+1} = X_+, \quad Y_{k+1} = Y_+,$$
$$\tau_{k+1} = \tau_+, \quad \psi_{k+1} = \psi_+.$$

We have the following theorem on the above algorithm:

Theorem 4.1 *[[21], Theorem 2.6] For any integer $0 \le k < K_0$, where $K_0 = \infty$ if $\overline{\alpha} \ne 1$ for all iterations, Algorithm 4.1 defines a pair*

$$(X_k, Y_k) \in \mathcal{N}(\beta_1, \tau_k),$$

and

$$r_k := A(X_k) + B(Y_k) - q = \psi_k r_0,$$
$$\tau_k = \psi_k \tau_0,$$
$$(1-\beta_1)\tau_k \le \mu_k = (X_k \bullet Y_k)/n \le (1+\beta_1)\tau_k,$$

where

$$\psi_0 = 1, \quad \psi_k = \prod_{j=0}^{k-1}(1-\overline{\alpha}_j),$$

and $\overline{\alpha}_j$ is defined by (25).

Observe from the theorem, that we have $\tau_{k+1} = (1-\overline{\alpha}_k)\tau_k$. Convergence of τ_k to zero implies that any accumulation point of (X_k, Y_k) is a solution to (1)-(3).

The following theorem states the complexity results of Algorithm 4.1:

Theorem 4.2 *[[21], Corollary 3.9] Assume that in Algorithm 4.1, we choose a starting point of the form $X_0 = Y_0 = \rho I$, where $\rho > 0$ is a constant. Let*

$$\epsilon_0 = \max\{X_0 \bullet Y_0, \|r_0\|\},$$

and let $\epsilon > 0$ be arbitrary. Then the following statements hold:

(i) *The algorithm terminates with an ϵ-approximate solution $(X_k, Y_k) \in S^n_+ \times S^n_+$ with*

$$0 \le X_k \bullet Y_k \le \epsilon, \quad \|r_k\| \le \epsilon,$$

in a finite number of steps $k = K_\epsilon < \infty$.

(ii) *If $\max\{\|X^*\|, \|Y^*\|\} \le \rho$, where (X^*, Y^*) solves (1)-(3), then*

$$K_\epsilon = O(n \ln(\epsilon_0/\epsilon)).$$

(iii) *For any choice of $\rho > 0$, there is an index $k = \hat{K}_\epsilon = O(n \ln(\epsilon_0/\epsilon))$ such that either*

(iiia) $(X_k, Y_k) \in S^n_+ \times S^n_+$ *satisfies* $0 \leq X_k \bullet Y_k \leq \epsilon$, $\|r_k\| \leq \epsilon$, *or,*

(iiib) $\overline{\alpha} \leq 1/[n(1 + 29/\sqrt{\beta_2 - \beta_1})]$, *and there is no solution (X^*, Y^*) to (1)-(3) with* $\max\{\|X^*\|, \|Y^*\|\} \leq \rho$.

Superlinear convergence of iterates (X_k, Y_k) generated by Algorithm 4.1 means:

$$\frac{\tau_{k+1}}{\tau_k} = (1 - \overline{\alpha}_k) \rightarrow 0. \tag{28}$$

From Step (a3) of the algorithm, we observe that if $(\alpha_1)_k \rightarrow 1$, then superlinear convergence of the algorithm occurs. Now $(\alpha_1)_k \rightarrow 1$ if and only if

$$\delta_k = \frac{1}{\tau_k}\|Y_k^{1/2} U_k V_k Y_k^{-1/2}\| \rightarrow 0. \tag{29}$$

(29) is equivalent to

$$\delta'_k = \frac{1}{\mu_k}\|Y_k^{1/2} U_k V_k Y_k^{-1/2}\| \rightarrow 0. \tag{30}$$

Here, the subscript "k" in the symbols above stands for the k^{th} iteration of the algorithm.

A currently weak sufficient condition for superlinear convergence (28) using Algorithm 4.1, as given in [20] (see also [17]), is

$$\lim_{k \rightarrow \infty} X_k Y_k / \sqrt{\tau_k} = \lim_{k \rightarrow \infty} X_k Y_k / \sqrt{\mu_k} = 0, \tag{31}$$

when applied to SDPs. This follows by showing that (29) or (30) holds true. In Proposition 4.5 below, we give another proof to show that (31) implies superlinear convergence using Algorithm 4.1 for the more general class of SDLCPs.

Without any sufficient conditions, evidences, for example, [31], have suggested that superlinear convergence using Algorithm 4.1 is not likely to be possible. Superlinear convergence using the algorithm with further modifications has been achieved in [14].

In the rest of this section, we discuss sufficient conditions that ensure superlinear convergence using Algorithm 4.1.

Given iterates (X_k, Y_k) generated by Algorithm 4.1. We have $(X_k, Y_k) \in \mathcal{N}(\beta_1, \tau_k)$. Let the off-central path passing through (X_k, Y_k) when $\mu = \mu_k$ be denoted by $(X^k(\mu), Y^k(\mu))$. Also, let $(X_0^k, Y_0^k) = (X^k(\bar{\mu}_0), Y^k(\bar{\mu}_0))$, where $\bar{\mu}_0 > 0$ is fixed to be less than $(1 - \beta_1)\tau_0$.

We have the following proposition on (X_0^k, Y_0^k):

Proposition 4.1 *We have (X_0^k, Y_0^k) satisfies $\lambda_{\min}(X_0^k Y_0^k) \geq D$, $\|X_0^k\| \leq C$, $\|Y_0^k\| \leq C$, for all $k \geq 0$. Here, C and D are some fixed numbers greater than zero.*

Proof: We have

$$(X_k, Y_k) \in \mathcal{N}(\beta_1, \tau_k),$$

by Theorem 4.1. Hence, it is clear that

$$|\lambda_{\min}(X_kY_k) - \tau_k| \leq \beta_1\tau_k.$$

Therefore,

$$\begin{aligned} \lambda_{\min}(X_kY_k) &\geq (1-\beta_1)\tau_k \\ &\geq \frac{1-\beta_1}{1+\beta_1}\mu_k, \end{aligned} \tag{32}$$

where the second inequality follows from Theorem 4.1.

On the other hand,

$$\begin{aligned} \lambda_{\min}(X_kY_k) &= \lambda_{\min}(X^k(\mu_k)Y^k(\mu_k)) \\ &= \frac{\mu_k}{\bar{\mu}_0}\lambda_{\min}(X_0^kY_0^k). \end{aligned}$$

Together with (32), we have

$$\lambda_{\min}(X_k^0Y_k^0) \geq \frac{(1-\beta_1)\bar{\mu}_0}{1+\beta_1}.$$

Now, by Theorem 4.1,

$$A(X_k) + B(Y_k) = q + \frac{\tau_k}{\tau_0}(A(X_0) + B(Y_0) - q). \tag{33}$$

Also, we have

$$\begin{aligned} A(X_k) + B(Y_k) &= A(X^k(\mu_k)) + B(Y^k(\mu_k)) \\ &= q + \frac{\mu_k}{\bar{\mu}_0}(A(X_0^k) + B(Y_0^k) - q). \end{aligned} \tag{34}$$

Putting (33) and (34) together, and upon manipulations, we get

$$\begin{aligned} A\left(X_0^k - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}X_0\right) + B\left(Y_0^k - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}Y_0\right) &= \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)q \\ &= \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)(A(X^*) + B(Y^*)). \end{aligned}$$

Therefore,

$$A\left(X_0^k - \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}X_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)X^*\right)\right) + B\left(Y_0^k - \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}Y_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)Y^*\right)\right) = 0.$$

By Assumption 2.1(a), we have

$$\left(X_0^k - \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}X_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)X^*\right)\right) \bullet \left(Y_0^k - \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}Y_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)Y^*\right)\right) \geq 0.$$

Expanding the expression above, and upon manipulations, we obtain

$$\begin{aligned} & X_0^k \bullet \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}Y_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)Y^*\right) + \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}X_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)X^*\right) \bullet Y_0^k \\ \leq\ & X_0^k \bullet Y_0^k + \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}X_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)X^*\right) \bullet \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}Y_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)Y^*\right) \\ =\ & n\bar{\mu}_0 + \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}X_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)X^*\right) \bullet \left(\frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}Y_0 + \left(1 - \frac{\tau_k\bar{\mu}_0}{\mu_k\tau_0}\right)Y^*\right). \end{aligned} \tag{35}$$

From our choice of $\bar{\mu}_0$ and that $1/(1+\beta_1) \le \tau_k/\mu_k \le 1/(1-\beta_1)$, we have

$$
\begin{aligned}
& X_0^k \bullet \left(\frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0} Y_0 + \left(1 - \frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0}\right) Y^* \right) \\
\le \quad & n\bar{\mu}_0 + \left(\frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0} X_0 + \left(1 - \frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0}\right) X^* \right) \bullet \left(\frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0} Y_0 + \left(1 - \frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0}\right) Y^* \right),
\end{aligned}
$$

which further implies that $\|X_0^k\| \le C$ for some positive constant C, $\forall\ k \ge 0$. The same holds for $\|Y_0^k\|$. **QED**

Hence, using Proposition 4.1, by Propositions 3.1 - 3.3, the off-central paths $(X^k(\mu), Y^k(\mu))$ generated from iterates (X_k, Y_k) satisfies

$$
X^k(\mu) = \begin{pmatrix} \Theta(1) & O(\sqrt{\mu}) \\ O(\sqrt{\mu}) & \Theta(\mu) \end{pmatrix}, \quad Y^k(\mu) = \begin{pmatrix} \Theta(\mu) & O(\sqrt{\mu}) \\ O(\sqrt{\mu}) & \Theta(1) \end{pmatrix} \tag{36}
$$

independent of k.

As a special case, we have

$$
X_k = \begin{pmatrix} \Theta(1) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(\mu_k) \end{pmatrix}, \quad Y_k = \begin{pmatrix} \Theta(\mu_k) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(1) \end{pmatrix}. \tag{37}
$$

Note that (37) has been established for example in [14, 17, 20, 23]. In (36), we establish a stronger result than (37) in that a similar property as (37) also holds for off-central paths derived from these iterates.

We have the following proposition:

Proposition 4.2 *[[14], Lemma 6.2]*

$$
Y_k^{1/2} = \begin{pmatrix} \Theta(\sqrt{\mu_k}) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(1) \end{pmatrix}, \quad Y_k^{-1/2} = \begin{pmatrix} \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}.
$$

Proof: Let

$$
Y_k^{1/2} = \begin{pmatrix} \sqrt{Y}_{11}(\mu_k) & \sqrt{Y}_{12}(\mu_k) \\ \sqrt{Y}_{12}^T(\mu_k) & \sqrt{Y}_{22}(\mu_k) \end{pmatrix}.
$$

Since $Y_k^{1/2} Y_k^{1/2} = Y_k$, we have from

$$
\sqrt{Y}_{11}(\mu_k)^2 + \sqrt{Y}_{12}(\mu_k)\sqrt{Y}_{12}^T(\mu_k) = (Y_k)_{11} = \Theta(\mu_k),
$$

that

$$
\sqrt{Y}_{11}(\mu_k) = O(\sqrt{\mu_k}), \quad \sqrt{Y}_{12}(\mu_k) = O(\sqrt{\mu_k}).
$$

Also, from

$$
\sqrt{Y}_{12}^T(\mu_k)\sqrt{Y}_{12}(\mu_k) + \sqrt{Y}_{22}(\mu_k)^2 = \Theta(1),
$$

and $\sqrt{Y}_{12}(\mu_k) = O(\sqrt{\mu_k})$,

$$\sqrt{Y}_{22}(\mu_k) = \Theta(1).$$

Therefore,

$$Y_k^{1/2} = \begin{pmatrix} O(\sqrt{\mu_k}) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(1) \end{pmatrix}.$$

Now,

$$Y_k = \begin{pmatrix} \sqrt{\mu_k} I & 0 \\ 0 & I \end{pmatrix} \widetilde{Y}_k \begin{pmatrix} \sqrt{\mu_k} I & 0 \\ 0 & I \end{pmatrix}, \tag{38}$$

where $\widetilde{Y}_k$ remains symmetric, positive definite as k tends to infinity, by (37). Hence,

$$\widetilde{Y}_k^{-1} = \begin{pmatrix} \Theta(1) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}.$$

Using (38), we obtain

$$Y_k^{-1} = \begin{pmatrix} \Theta\left(\frac{1}{\mu_k}\right) & O\left(\frac{1}{\sqrt{\mu_k}}\right) \\ O\left(\frac{1}{\sqrt{\mu_k}}\right) & \Theta(1) \end{pmatrix}.$$

By $Y_k^{-1/2} Y_k^{-1/2} = Y_k^{-1}$, similar approach as above yields the following:

$$Y_k^{-1/2} = \begin{pmatrix} \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) & O(1) \\ O(1) & O(1) \end{pmatrix}.$$

Now, $Y_k^{1/2} Y_k^{-1/2} = I$ implies that

$$Y_k^{1/2} = \begin{pmatrix} \Theta(\sqrt{\mu_k}) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(1) \end{pmatrix}, \quad Y_k^{-1/2} = \begin{pmatrix} \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}.$$

QED

The proof for the above proposition is slightly different from that in [14], and hence we show it here.

Let us now look more closely at (30), which is the same as

$$Y_k^{1/2} U_k V_k Y_k^{-1/2} = o(\mu_k). \tag{39}$$

We give below an equivalent formulation of (39), using Proposition 4.2:

Proposition 4.3 *We have*

$$Y_k^{1/2} U_k V_k Y_k^{-1/2} = o(\mu_k)$$

if and only if

$$U_k V_k = \mu_k^2 \begin{pmatrix} o\left(\frac{1}{\mu_k}\right) & o\left(\frac{1}{\mu_k^{3/2}}\right) \\ o\left(\frac{1}{\sqrt{\mu_k}}\right) & o\left(\frac{1}{\mu_k}\right) \end{pmatrix}.$$

Proof: Let

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \frac{1}{\mu_k^2} U_k V_k.$$

Then, by Proposition 4.2,

$$\begin{aligned} & Y_k^{1/2} U_k V_k Y_k^{-1/2} \\ = \; & \mu_k^2 \begin{pmatrix} \Theta(1) & O(1) \\ O(1) & \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} \Theta(1) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(\sqrt{\mu_k}) \end{pmatrix} \\ = \; & \mu_k^2 \begin{pmatrix} \Theta(1)H_{11} + O(1)H_{21} & \Theta(1)H_{12} + O(1)H_{22} \\ O(1)H_{11} + \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) H_{21} & O(1)H_{12} + \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) H_{22} \end{pmatrix} \begin{pmatrix} \Theta(1) & O(\sqrt{\mu_k}) \\ O(\sqrt{\mu_k}) & \Theta(\sqrt{\mu_k}) \end{pmatrix} \\ = \; & \mu_k^2 \left(\begin{matrix} (\Theta(1)H_{11} + O(1)H_{21})\Theta(1) + (\Theta(1)H_{12} + O(1)H_{22})O(\sqrt{\mu_k}) \\ \left(O(1)H_{11} + \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) H_{21}\right) \Theta(1) + \left(O(1)H_{12} + \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) H_{22}\right) O(\sqrt{\mu_k}) \end{matrix} \right. \\ & \left. \begin{matrix} (\Theta(1)H_{11} + O(1)H_{21})O(\sqrt{\mu_k}) + (\Theta(1)H_{12} + O(1)H_{22})\Theta(\sqrt{\mu_k}) \\ \left(O(1)H_{11} + \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) H_{21}\right) O(\sqrt{\mu_k}) + \left(O(1)H_{12} + \Theta\left(\frac{1}{\sqrt{\mu_k}}\right) H_{22}\right) \Theta(\sqrt{\mu_k}) \end{matrix} \right). \end{aligned}$$

It can be verified easily using the above expression for $Y_k^{1/2} U_k V_k Y_k^{-1/2}$ that if

$$H_{11} = o\left(\frac{1}{\mu_k}\right), \tag{40}$$

$$H_{12} = o\left(\frac{1}{\mu_k^{3/2}}\right), \tag{41}$$

$$H_{21} = o\left(\frac{1}{\sqrt{\mu_k}}\right), \tag{42}$$

$$H_{22} = o\left(\frac{1}{\mu_k}\right), \tag{43}$$

then

$$Y_k^{1/2} U_k V_k Y_k^{-1/2} = o(\mu_k).$$

On the other hand, suppose

$$Y_k^{1/2} U_k V_k Y_k^{-1/2} = o(\mu_k).$$

Then,

$$U_k V_k = Y_k^{-1/2} \begin{pmatrix} o(\mu_k) & o(\mu_k) \\ o(\mu_k) & o(\mu_k) \end{pmatrix} Y_k^{1/2}.$$

Using Proposition 4.2, we show that (40)-(43) hold. **QED**

Hence, we have the following theorem which gives a sufficient condition for superlinear convergence using Algorithm 4.1:

Theorem 4.3 *Let (X_k, Y_k) be generated by Algorithm 4.1, and $(X^k(\mu), Y^k(\mu))$ be its corresponding off-central path, $k = 0, 1, \ldots$. If*

$$(X^k)'(\mu_k)(Y^k)'(\mu_k) = \begin{pmatrix} o\left(\frac{1}{\mu_k}\right) & o\left(\frac{1}{\mu_k^{3/2}}\right) \\ o\left(\frac{1}{\sqrt{\mu_k}}\right) & o\left(\frac{1}{\mu_k}\right) \end{pmatrix}, \tag{44}$$

then (X_k, Y_k) converges superlinearly.

Proof: Observe by comparing (4), (5) with (23), (24) (when $\sigma = 0$ and $\bar{r} = A(X_k) + B(Y_k) - q$) that

$$(X^k)'(\mu_k) = -\frac{1}{\mu_k}U_k, \quad (Y^k)'(\mu_k) = -\frac{1}{\mu_k}V_k.$$

The result then follows from (25)-(28), (30), and Proposition 4.3. **QED**

The currently known sufficient condition (31) for superlinear convergence of iterates (X_k, Y_k) generated by Algorithm 4.1 is equivalent to

$$(X_k Y_k)_{12} = o(\sqrt{\mu_k}). \tag{45}$$

In the below proposition, we show that (45) can be written in terms of convergence rate of block entries of X_k and Y_k individually.

Proposition 4.4

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

if and only if

$$(X_k)_{12} = o(\sqrt{\mu_k}) \quad \text{and} \quad (Y_k)_{12} = o(\sqrt{\mu_k}). \tag{46}$$

Proof: Suppose (X_k, Y_k) satisfies

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

Let $(\widetilde{X}_k, \widetilde{Y}_k)$, where

$$\begin{aligned} X_k &= \begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu_k} I \end{pmatrix} \widetilde{X}_k \begin{pmatrix} I & 0 \\ 0 & \sqrt{\mu_k} I \end{pmatrix}, \\ Y_k &= \begin{pmatrix} \sqrt{\mu_k} I & 0 \\ 0 & I \end{pmatrix} \widetilde{Y}_k \begin{pmatrix} \sqrt{\mu_k} I & 0 \\ 0 & I \end{pmatrix}, \end{aligned}$$

be the sequence of iterates corresponding to (X_k, Y_k), with any accumulation point $(\widetilde{X}^*, \widetilde{Y}^*)$. We have

$$(\widetilde{X}_k \widetilde{Y}_k)_{12} = o(1).$$

Hence,

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

We have

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

Hence,

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

Therefore, $(\widetilde{Y}^*)_{12} \bullet (\widetilde{X}^*)_{12} = -\mathrm{Tr}((\widetilde{Y}^*)_{12}^T(\widetilde{X}^*)_{11}(\widetilde{Y}^*)_{12}(\widetilde{Y}^*)_{22}^{-1}) \leq 0$.

On the other hand,

$$\mathcal{A}\mathrm{svec}(X_k) + \mathcal{B}\mathrm{svec}(Y_k) = q + \frac{\mu_k}{\bar{\mu}_0}(\mathcal{A}\mathrm{svec}(X_0^k) + \mathcal{B}\mathrm{svec}(Y_0^k) - q)$$

implies that

$$\bar{\mathcal{A}}(\sqrt{\mu_k})\mathrm{svec}(\widetilde{X}_k) + \bar{\mathcal{B}}(\sqrt{\mu_k})\mathrm{svec}(\widetilde{Y}_k) = q + \ \mathrm{Diag}\left(I_{i_1}, \frac{1}{\sqrt{\mu_k}}I_{i_2}, \frac{1}{\mu_k}I_{\widetilde{n-i_1-i_2}}\right)\frac{\mu_k}{\bar{\mu}_0}(\mathcal{A}\mathrm{svec}(X_0^k) + \mathcal{B}\mathrm{svec}(Y_0^k)).$$

As $\mu_k \to 0$, we have

$$\bar{\mathcal{A}}(0)\mathrm{svec}(\widetilde{X}^*) + \bar{\mathcal{B}}(0)\mathrm{svec}(\widetilde{Y}^*) = q + \frac{1}{\bar{\mu}_0}\left(\bar{\mathcal{A}}(0)\mathrm{svec}\begin{pmatrix} 0 & 0 \\ 0 & (X_0^*)_{22}\end{pmatrix} + \bar{\mathcal{B}}(0)\mathrm{svec}\begin{pmatrix} (Y_0^*)_{11} & 0 \\ 0 & 0\end{pmatrix}\right),$$

where (X_0^k, Y_0^k) converges to (X_0^*, Y_0^*).

That is,

$$\bar{\mathcal{A}}(0)\mathrm{svec}\left(\widetilde{X}^* - \frac{1}{\bar{\mu}_0}\begin{pmatrix} 0 & 0 \\ 0 & (X_0^*)_{22}\end{pmatrix}\right) + \bar{\mathcal{B}}(0)\mathrm{svec}\left(\widetilde{Y}^* - \frac{1}{\bar{\mu}_0}\begin{pmatrix} (Y_0^*)_{11} & 0 \\ 0 & 0\end{pmatrix}\right) = q.$$

Hence, we have

$$\bar{\mathcal{A}}(0)\mathrm{svec}\begin{pmatrix} 0 & (\widetilde{X}^*)_{12} \\ (\widetilde{X}^*)_{12}^T & (\widetilde{X}^*)_{22} - (1/\bar{\mu}_0)(X_0^*)_{22}\end{pmatrix} + \bar{\mathcal{B}}(0)\mathrm{svec}\begin{pmatrix} (\widetilde{Y}^*)_{11} - (1/\bar{\mu}_0)(Y_0^*)_{11} & (\widetilde{Y}^*)_{12} \\ (\widetilde{Y}^*)_{12}^T & 0\end{pmatrix} = 0,$$

using the structure of q in Proposition 3.5, and by monotonicity[4],

$$\begin{pmatrix} 0 & (\widetilde{X}^*)_{12} \\ (\widetilde{X}^*)_{12}^T & (\widetilde{X}^*)_{22} - (1/\bar{\mu}_0)(X_0^*)_{22}\end{pmatrix} \bullet \begin{pmatrix} (\widetilde{Y}^*)_{11} - (1/\bar{\mu}_0)(Y_0^*)_{11} & (\widetilde{Y}^*)_{12} \\ (\widetilde{Y}^*)_{12}^T & 0\end{pmatrix} \geq 0.$$

Hence, $(\widetilde{Y}^*)_{12} \bullet (\widetilde{X}^*)_{12} \geq 0$.

With $(\widetilde{Y}^*)_{12} \bullet (\widetilde{X}^*)_{12} = -\mathrm{Tr}((\widetilde{Y}^*)_{12}^T(\widetilde{X}^*)_{11}(\widetilde{Y}^*)_{12}(\widetilde{Y}^*)_{22}^{-1}) \leq 0$, we have $(\widetilde{Y}^*)_{12} = 0$, and we are done.

The if direction follows trivially. **QED**

We now give an alternative proof to show that (31) is a sufficient condition for superlinear convergence, by showing that (46) implies (44).

[4] See proof of Proposition 2.4 of [32].

Proposition 4.5 *Suppose the iterates (X_k, Y_k) generated by Algorithm 4.1 satisfies*

$$X_k = \begin{pmatrix} \Theta(1) & o(\sqrt{\mu_k}) \\ o(\sqrt{\mu_k}) & \Theta(\mu_k) \end{pmatrix}, \quad Y_k = \begin{pmatrix} \Theta(\mu_k) & o(\sqrt{\mu_k}) \\ o(\sqrt{\mu_k}) & \Theta(1) \end{pmatrix}, \tag{47}$$

that is, (31) is satisfied. Then

$$(X^k)'(\mu_k) = \begin{pmatrix} o\left(\frac{1}{\mu_k}\right) & o\left(\frac{1}{\sqrt{\mu_k}}\right) \\ o\left(\frac{1}{\sqrt{\mu_k}}\right) & O(1) \end{pmatrix}, \quad (Y^k)'(\mu_k) = \begin{pmatrix} O(1) & o\left(\frac{1}{\sqrt{\mu_k}}\right) \\ o\left(\frac{1}{\sqrt{\mu_k}}\right) & o\left(\frac{1}{\mu_k}\right) \end{pmatrix}. \tag{48}$$

Hence, superlinear convergence of (X_k, Y_k) follows.

Proof: Suppose (47) holds.
Let $(\widetilde{X}^*, \widetilde{Y}^*)$ be any accumulation point of $(\widetilde{X}_k, \widetilde{Y}_k)$.
We have $(\widetilde{X}^*)_{12} = (\widetilde{Y}^*)_{12} = 0$.
Let $(\widetilde{X}_k, \widetilde{Y}_k) \to (\widetilde{X}^*, \widetilde{Y}^*)$, as $k \to \infty$.
We have

$$(\widetilde{X}_k \otimes_s \widetilde{Y}_k^{-1})\widetilde{\mathcal{G}}_k^{-1} q \to \begin{pmatrix} -(\widetilde{X}^*)_{11} & 0 \\ 0 & (\widetilde{X}^*)_{22} \end{pmatrix},$$

and

$$\widetilde{\mathcal{G}}_k^{-1} q \to \begin{pmatrix} -(\widetilde{Y}^*)_{11} & 0 \\ 0 & (\widetilde{Y}^*)_{22} \end{pmatrix}.$$

Hence,

$$(\widetilde{X}_k \otimes_s \widetilde{Y}_k^{-1})\widetilde{\mathcal{G}}_k^{-1} q + \operatorname{svec}(\widetilde{X}_k) = \operatorname{svec}\begin{pmatrix} o(1) & o(1) \\ o(1) & O(1) \end{pmatrix}, \tag{49}$$

$$-\widetilde{\mathcal{G}}_k^{-1} q + \operatorname{svec}(\widetilde{Y}_k) = \operatorname{svec}\begin{pmatrix} O(1) & o(1) \\ o(1) & o(1) \end{pmatrix}. \tag{50}$$

(48) then follows using (49), (50) on (20).
Superlinear convergence of (X_k, Y_k) can be easily seen as (48) implies (44). **QED**

5 Superlinear Convergence for a Class of SDPs.

Consider the class of linear semi-definite feasibility problems, when we either have $C = 0$ in $(\mathcal{P})$, or $b_i = 0$, $i = 1, \ldots, m$, in $(\mathcal{D})$.

As mentioned in Section 2, Assumption 2.1(a) is satisfied for any SDP and hence for a linear semi-definite feasibility problem. Assumption 2.1(b) can also be satisfied for any given linear semi-definite feasibility problem as long as the feasible regions of $(\mathcal{P})$ and $(\mathcal{D})$ are nonempty. This can be seen easily by choosing an optimal solution to be any feasible $X \in S^n_+$, and $y_i = 0, i = 1, \ldots, m$, $Y = 0$ in case $C = 0$, and $X = 0$, and any feasible $(y_1, \ldots, y_m, Y), Y \in S^n_+$ in case $b_i = 0$, $i = 1, \ldots, m$.

A question arises whether Assumption 3.1 is automatically satisfied for any given linear semi-definite feasibility problem. In the below example, we show that it is not necessarily true:

Example 5.1 *Let $n = 4$ and $m = 4$.*
Define $C = 0$, $b_1 = 1$, $b_i = 0, i = 2, 3, 4$,

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$
$$A_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

It can be shown easily that the optimal solution to ($\mathcal{P}$) are of the form:

$$X^* = \begin{pmatrix} x_{11} & x_{12} & 0 & 0 \\ x_{12} & 1 - x_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \succeq 0,$$

and the optimal solution to ($\mathcal{D}$) are of the form:

$$y_1^* = 0, y_2^* \leq 0, y_3^* = 0, y_4^* = 0,$$
$$Y^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -y_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \succeq 0.$$

Hence, strict complementarity condition does not hold for this example.

Therefore, for a linear semi-definite feasibility problem, Assumptions 2.1(a)-(c) are not really necessary, since they can be satisfied, while Assumption 3.1 is needed.

For a linear semi-definite feasibility problem, we have q in (2) has its last $\tilde{n} - m$ entries equal to zero in case $C = 0$, and its first m entries equal to zero in case $b_i = 0$, $i = 1, \ldots, m$. In the former, $\hat{\hat{q}} = 0$, and in the latter, $\hat{q} = 0$, in Proposition 3.5.

We have the following theorem which follows by applying Theorem 4.3 from Section 4:

Theorem 5.1 *Assume Assumption 3.1. Let (X_k, Y_k) be iterates generated using Algorithm 4.1 on a linear semi-definite feasibility problem with the initial iterate (X_0, Y_0) satisfying*

$$\begin{pmatrix} svec\begin{pmatrix} 0 & 0 \\ 0 & (A_{j_1+j_2+1})_{22} \end{pmatrix}^T \\ \vdots \\ svec\begin{pmatrix} 0 & 0 \\ 0 & (A_m)_{22} \end{pmatrix}^T \end{pmatrix} svec(X_0) = 0, \tag{51}$$

in case $C = 0$*, and*

$$\left(\begin{array}{l} svec\left(\begin{array}{cc} (B_{k_1+k_2+1})_{11} & 0 \\ 0 & 0 \end{array} \right)^T \\ \vdots \\ svec\left(\begin{array}{cc} (B_{\tilde{n}-m})_{11} & 0 \\ 0 & 0 \end{array} \right)^T \end{array} \right) svec(Y_0) = 0, \tag{52}$$

in case $b_i = 0, i = 1, \ldots, m$.
Then the iterates converge superlinearly.

Proof: We need only prove this for a linear semi-definite feasibility problem when $C = 0$. The proof for the case when $b_i = 0$, $i = 1, \ldots, m$, is similar.

Consider the corresponding $(\widetilde{X}_k, \widetilde{Y}_k)$ to (X_k, Y_k). Let $(\widetilde{X}^*, \widetilde{Y}^*)$ be any accumulation point of $(\widetilde{X}_k, \widetilde{Y}_k)$. Without loss of generality, let $(\widetilde{X}_k, \widetilde{Y}_k) \rightarrow (\widetilde{X}^*, \widetilde{Y}^*)$. Also, let $(X_0^k, Y_0^k) \rightarrow (X_0^*, Y_0^*)$.

Note that by (7) and Theorem 4.1,

$$\begin{aligned} \mathcal{A}\text{svec}(X_k) + \mathcal{B}\text{svec}(Y_k) &= q + \frac{\mu_k}{\bar{\mu}_0} r_0^k \\ &= q + \frac{\tau_k}{\tau_0} r_0, \end{aligned}$$

where

$$\begin{aligned} r_0^k &= \mathcal{A}\text{svec}(X_0^k) + \mathcal{B}\text{svec}(Y_0^k) - q, \\ r_0 &= \mathcal{A}\text{svec}(X_0) + \mathcal{B}\text{svec}(Y_0) - q. \end{aligned}$$

We therefore have

$$r_0^k = \frac{\tau_k \bar{\mu}_0}{\mu_k \tau_0} r_0. \tag{53}$$

From (51), we have using (53) that

$$\left(\begin{array}{l} \text{svec}\left(\begin{array}{cc} 0 & 0 \\ 0 & (A_{j_1+j_2+1})_{22} \end{array} \right)^T \\ \vdots \\ \text{svec}\left(\begin{array}{cc} 0 & 0 \\ 0 & (A_m)_{22} \end{array} \right)^T \end{array} \right) \text{svec}(X_0^k) = 0. \tag{54}$$

Now, observe that

$$\mathcal{A}\text{svec}(X_k) + \mathcal{B}\text{svec}(Y_k) = q + \frac{\mu_k}{\bar{\mu}_0} r_0^k.$$

Hence,

$$\text{Diag}(I_{j_1}, \sqrt{\mu_k} I_{j_2}, \mu_k I_{m-j_1-j_2}, I_{k_1}, \sqrt{\mu_k} I_{k_2}, \mu_k I_{\tilde{n}-m-k_1-k_2})[\bar{\mathcal{A}}(\sqrt{\mu_k})(\widetilde{X}_k \otimes_s \widetilde{Y}_k^{-1}) + \bar{\mathcal{B}}(\sqrt{\mu_k})]\text{svec}(\widetilde{Y}_k) = q + \frac{\mu_k}{\bar{\mu}_0} r_0^k.$$

Therefore,

$$(\bar{\mathcal{A}}(\sqrt{\mu_k})(\widetilde{X}_k \otimes_s \widetilde{Y}_k^{-1}) + \bar{\mathcal{B}}(\sqrt{\mu_k}))\mathrm{svec}(\widetilde{Y}_k) = q + \frac{\sqrt{\mu_k}}{\bar{\mu}_0}\widetilde{r}_0^k,$$

where we have used the fact that besides the first j_1 entries of q, the rest of entries of q are zero. $\widetilde{r}_0^k$ is defined by (19).

Let $k \to \infty$. Then, using (54), we obtain

$$(\bar{\mathcal{A}}(0)(\widetilde{X}^* \otimes_s (\widetilde{Y}^*)^{-1}) + \bar{\mathcal{B}}(0))\mathrm{svec}(\widetilde{Y}^*) = q + \frac{1}{\bar{\mu}_0}\mathrm{Diag}(0,0,0,0,0,I_{\tilde{n}-m-k_1-k_2})r_0^*,$$

where $r_0^* = \mathcal{A}\mathrm{svec}(X_0^*) + \mathcal{B}\mathrm{svec}(Y_0^*) - q$.

Therefore,

$$(\bar{\mathcal{B}}(0) - \bar{\mathcal{A}}(0)(\widetilde{X}^* \otimes_s (\widetilde{Y}^*)^{-1}))\mathrm{svec}(\widetilde{Y}^*) = -q + \frac{1}{\bar{\mu}_0}\mathrm{Diag}(0,0,0,0,0,I_{\tilde{n}-m-k_1-k_2})r_0^*.$$

Hence,

$$\begin{aligned} (\widetilde{X}_k \otimes_s \widetilde{Y}_k^{-1})\widetilde{\mathcal{G}}_k^{-1}(-\sqrt{\mu_k}\widetilde{r}_0^k/\bar{\mu}_0 + q) &\to -\mathrm{svec}(\widetilde{X}^*), \\ \widetilde{\mathcal{G}}_k^{-1}(\sqrt{\mu_k}\widetilde{r}_0^k/\bar{\mu}_0 - q) &\to \mathrm{svec}(\widetilde{Y}^*), \end{aligned}$$

as $k \to \infty$, where we recall that $\widetilde{\mathcal{G}}_k = \bar{\mathcal{B}}(\sqrt{\mu_k}) - \bar{\mathcal{A}}(\sqrt{\mu_k})(\widetilde{X}_k \otimes_s \widetilde{Y}_k^{-1})$.

Hence, from (20), we see that

$$(X^k)'(\mu_k) = \begin{pmatrix} o\left(\frac{1}{\mu_k}\right) & o\left(\frac{1}{\sqrt{\mu_k}}\right) \\ o\left(\frac{1}{\sqrt{\mu_k}}\right) & o(1) \end{pmatrix}, \quad (Y^k)'(\mu_k) = \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}\left(\frac{1}{\sqrt{\mu_k}}\right) \\ \mathcal{O}\left(\frac{1}{\sqrt{\mu_k}}\right) & \mathcal{O}\left(\frac{1}{\mu_k}\right) \end{pmatrix}.$$

Therefore, (44) holds, and superlinear convergence of (X_k, Y_k) then follows from Theorem 4.3.
QED

In particular, the following corollary of Theorem 5.1 follows immediately:

Corollary 5.1 *If the interior of the primal feasible region is nonempty in case $C = 0$, or the interior of the dual feasible region is nonempty in case $b_i = 0, i = 1, \ldots, m$, then the iterates generated by Algorithm 4.1 converge superlinearly.*

We have the following final remark:

Remark 5.1 *We tested out examples of linear semi-definite feasibility problem (where $C = 0$) with $n = 4$ and $m = 4$ using the SDP solver, SDPT3. We tested different scenarios: with and without strict complementarity condition, different initial iterates (including those that satisfy the condition in Theorem 5.1). In all these cases (which also include Example 5.1), only linear convergence is observed, where μ_{k+1}/μ_k lies between 0.01 and 0.4 eventually. We set the tolerance for termination of algorithm to be 1×10^{-15} or smaller.*

References

[1] M. Ait Ramit, U. Helmke and J.B. Moore, *A finite steps algorithm for solving convex feasibility problems,* Journal of Global Optimization, Vol. 38(2007), pp. 143-160.

[2] F. Alizadeh, J.A. Haeberly and M. Overton, *Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results,* SIAM Journal on Optimization, Vol. 8(1998), pp. 746-768.

[3] D.A. Bayer and J.C. Lagarias, *The nonlinear geometry of linear programming, I, II, III,* Transactions of the American Mathematical Society, Vol. 314(1989), pp. 499-526, pp. 527-581, and Vol. 320(1990), pp. 193-225.

[4] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory,* SIAM Studies in Applied Mathematics, Philadelphia, 1994.

[5] C.B. Chua, *Analyticity of weighted central paths and error bounds for semidefinite programming,* Mathematical Programming, Series A, Vol. 115(2008), pp. 239-271.

[6] S.K. Chua, K.-C. Toh & G. Zhao, *An analytic center cutting plane method with deep cuts for semidefinite feasibility problems,* Journal of Optimization Theory and Applications, Vol. 123(2004), pp. 291-318.

[7] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations,* International Series in Pure and Applied Mathematics, 1955.

[8] P. Gahinet & A. Nemirovski, *The projective method for solving linear matrix inequalities,* Mathematical Programming, Series A, Vol. 77(1997), pp. 163-190.

[9] J.-L. Goffin, Z. Luo & Y. Ye, *Complexity analysis of an interior cutting plane method for convex feasibility problems,* SIAM Journal on Optimization, Vol. 6(1996), pp. 638-652.

[10] D. Henrion & J. Malick, *Projection methods for conic feasibility problems, applications to polynomial sum-of-squares decompositions,* preprint, 2009.

[11] J.X.da Cruz Neto, O.P. Ferreira and R.D.C. Monteiro, *Asymptotic behavior of the central path for a special class of degenerate SDP problems,* Mathematical Programming, Series A, Vol. 103(2005), pp. 487-514.

[12] C. Helmberg, F. Rendl, R.J. Vanderbei and H. Wolkowicz, *An interior-point method for semidefinite programming,* SIAM Journal on Optimization, Vol. 6(1996), pp. 342-361.

[13] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, New York, 1991.

[14] M. Kojima, M. Shida and S. Shindoh, *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs,* Mathematical Programming, Series A, Vol. 80(1998), pp. 129-160.

[15] M. Kojima, M. Shida and S. Shindoh, *A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the alizadeh-haeberly-overton search direction,* SIAM Journal on Optimization, Vol. 9(1999), pp. 444-465.

[16] M. Kojima, S. Shindoh and S. Hara, *Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices,* SIAM Journal on Optimization, Vol. 7(1997), pp. 86-125.

[17] Z. Lu and R.D.C. Monteiro, *Error bounds and limiting behavior of weighted paths associated with the SDP map* $X^{1/2}SX^{1/2}$, SIAM Journal on Optimization, Vol. 15(2004), pp. 348-374.

[18] Z. Lu and R.D.C. Monteiro, *A note on the local convergence of a predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem based on the alizadeh-haeberly-overton search direction,* SIAM Journal on Optimization, Vol. 15(2005), pp. 1147-1154.

[19] Z. Lu and R.D.C. Monteiro, *Limiting behavior of the alizadeh-haeberly-overton weighted paths in semidefinite programming,* Optimization Methods and Software, Vol. 22(2007), pp. 849-870.

[20] F.A. Potra and R. Sheng, *Superlinear convergence of interior-point algorithms for semidefinite programming,* Journal of Optimization Theory and Applications, Vol. 99(1998), pp. 103-119.

[21] F.A. Potra and R. Sheng, *A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming,* SIAM Journal on Optimization, Vol. 8(1998), pp. 1007-1028.

[22] M. Preiß and J. Stoer, *Analysis of infeasible-interior-point paths and high-order methods for solving SDLCPs,* Parametric Optimization and Related Topics, VII. Aportaciones Matemáticas: Investigación, 18(2004), pp. 235-251.

[23] M. Preiß and J. Stoer, *Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems,* Mathematical Programming, Series A, Vol. 99(2004), pp. 499-520.

[24] S. Mehrotra, *Quadratic convergence in a primal-dual method,* Mathematics of Operations Research, Vol. 18(1993), pp. 741-751.

[25] R.D.C. Monteiro, *Primal-dual path following algorithms for semidefinite programming,* SIAM Journal on Optimization, Vol. 7(1997), pp. 663-678.

[26] R.D.C. Monteiro and J.-S. Pang, *On two interior-point mappings for nonlinear semidefinite complementarity problem,* Mathematics of Operations Research, Vol. 23(1998), pp. 39-60.

[27] R.D.C. Monteiro and T. Tsuchiya, *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming,* SIAM Journal on Optimization, Vol. 9(1999), pp. 551-577.

[28] R.D.C. Monteiro and P.R. Zanjácomo, *General interior-point maps and existence of weighted paths for nonlinear semidefinite complementarity problem,* Mathematics of Operations Research, Vol. 25(2000), pp. 381-399.

[29] Y. Nesterov and M. Todd, *Self-scaled barriers and interior-point methods for convex programming,* Mathematics of Operations Research, Vol. 22(1997), pp. 1-42.

[30] Y. Nesterov and M. Todd, *Primal-dual interior-point methods for self-scaled cones,* SIAM Journal on Optimization, Vol. 8(1998), pp. 324-364.

[31] C.-K. Sim and G. Zhao, *Underlying paths in interior point methods for the monotone semidefinite linear complementarity problem,* Mathematical Programming, Series A, Vol. 110(2007), pp. 475-499.

[32] C.-K. Sim and G. Zhao, *Asymptotic behavior of helmberg-kojima-monteiro (HKM) paths in interior point methods for monotone semidefinite linear complementarity problems: general theory,* Journal of Optimization Theory and Applications, Vol. 137(2008), pp. 11-25.

[33] C.-K. Sim, *On the analyticity of underlying HKM path at a solution of monotone semidefinite linear complementarity problem,* Journal of Optimization Theory and Applications, Vol. 141(2009), pp. 193-215.

[34] G. Sonnevend, *An analytic center for polyhedrons and new classes for linear programming.* In: System Modelling and Optimization (A. Prekopa, Ed.), Lecture Notes in Control and Information Sciences, Vol. 84, Springer-Verlag, Berlin, 1985, pp. 866-876.

[35] J. Stoer and M. Wechs, *On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems,* Numerische Mathematik, Vol. 81(1999), pp. 631-645.

[36] J. Stoer, M. Wechs and S. Mizuno, *High order infeasible-interior-point methods for solving sufficient linear complementarity problems,* Mathematics of Operations Research, Vol. 23(1998), pp. 832-862.

[37] J.F. Sturm, *Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists,* Mathematics of Operations Research, Vol. 24(1999), pp. 72-94.

[38] J. Sun, K.-C. Toh & G. Zhao, *An analytic center cutting plane method for semidefinite feasibility problems,* Mathematics of Operations Research, Vol. 27(2002), pp. 332-346.

[39] M.J. Todd, K.C. Toh and R.H. Tütüncü, *On the nesterov-todd direction in semidefinite programming,* SIAM Journal on Optimization, Vol. 8(1998), pp. 769-796.

[40] K.-C. Toh, G. Zhao & J. Sun, *A multiple-cut analytic center cutting plane method for semidefinite feasibility problems,* SIAM Journal on Optimization, Vol. 12(2002), pp. 1126-1146.

[41] M. Trnovská and M. Halická, *Limiting behavior and analyticity of weighted central paths in semidefinite programming,* preprint.

[42] P. Tseng, *Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP,* Optimization Methods and Software, Vol. 9(1998), pp. 245-268.

[43] Y. Ye and K. Anstreicher, *On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP,* Mathematical Programming, Series A, Vol. 62(1993), pp. 537-551.

[44] Y. Ye, O. Güler, R.A. Tapia and Y. Zhang, *A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming,* Mathematical Programming, Series A, Vol. 59(1993), pp. 151-162.

[45] Y. Zhang, *On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming,* SIAM Journal on Optimization, Vol. 8(1998), pp. 365-386.