J Optim Theory Appl (2008) 137: 11–25
DOI 10.1007/s10957-007-9280-3

Asymptotic Behavior of Helmberg-Kojima-Monteiro (HKM) Paths in Interior-Point Methods for Monotone Semidefinite Linear Complementarity Problems: General Theory

C.K. Sim · G. Zhao

Published online: 8 December 2007

Abstract An interior-point method (IPM) defines a search direction at each interior point of the feasible region. These search directions form a direction field, which in turn gives rise to a system of ordinary differential equations (ODEs). Thus, it is natural to define the underlying paths of the IPM as the solutions of the system of ODEs. In Sim and Zhao (Math. Program. Ser. A, 2007, to appear), these off-central paths are shown to be well-defined analytic curves and any of their accumulation points is a solution to the given monotone semidefinite linear complementarity problem (SDLCP). Off-central paths for a simple example are also studied in Sim and Zhao (Math. Program. Ser. A, 2007, to appear) and their asymptotic behavior near the solution of the example is analyzed. In this paper, which is an extension of Sim and Zhao (Math. Program. Ser. A, 2007, to appear), we study the asymptotic behavior of the off-central paths for general SDLCPs using the dual HKM direction. We give a necessary and sufficient condition for when an off-central path is analytic as a function of $\sqrt{\mu}$ at a solution of the SDLCP. Then, we show that, if the given SDLCP has a unique solution, the first derivative of its off-central path, as a function of $\sqrt{\mu}$, is bounded. We work under the assumption that the given SDLCP satisfies the strict complementarity condition.

Communicated by P. Tseng.

This research was done during the first author's Ph.D. study at the Department of Mathematics, National University of Singapore, and as a Research Engineer at the NUS Business School. The second author's work was supported in part by NUS Academic Research Grant R-146-000-084-112.

For technical details of the paper, please refer to The Logistics Institute—Asia Pacific Internal Report. See also http://www.math.nus.edu.sg/~matzgy/publist.html.

C.K. Sim (✉)
Department of Decision Sciences, National University of Singapore Business School, Singapore, Singapore
e-mail: bizsck@nus.edu.sg

G. Zhao
Department of Mathematics, National University of Singapore, Singapore, Singapore

Keywords Semidefinite linear complementarity problems · Ordinary differential equations · Off-central paths · HKM direction · Analyticity

1 Introduction

The notion of central path was introduced by Sonnevend [1] in 1985 for the interior point method (IPM). Since then, people realized that the IPM is actually a homotopy method following the underlying paths and that many remarkable properties of the IPM can be attributed to the nice geometry of these paths. Readers interested in knowing more about the basic geometry of these paths may refer to [2–4].

In [5–9], it was found that, for solving a linear program (LP) or a linear complementarity problem (LCP), the number of iterations needed by a predictor-corrector path-following algorithm to reduce the duality gap μ from μ_0 to $\epsilon > 0$ is equivalent to the integral of the curvature of the central path from μ_0 to ϵ. This equivalence relates a discrete analysis (complexity analysis) to a continuous analysis (curvature of path) and thus opens a new way to estimate upper and lower bounds for the complexity of IPMs. In [10], it is shown that the complexity of their layered least squares path-following LP algorithm depends only on the constraint matrix, by observing those regions where the central path is straight or crossing over.

Another important role the underlying paths play in the study of IPMs is to show its fast local convergence. Classical proof of the 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 central path in an IPM may be singular at the optimal solution. Thus traditional approach of local convergence analysis does not work for IPMs. Fast local convergence of IPMs has instead been successfully proved by relating it to the boundedness of derivatives of the underlying paths in [11–14].

The study of fast local convergence is particularly important for the semidefinite linear complementarity problem (SDLCP), with the semidefinite program (SDP) as a special case, because, in contrast to LCP, the exact solution of a SDLCP cannot be obtained from an approximate solution by determining a complementary basis.

There are various ways in which the underlying paths, using different search directions, for SDLCPs are defined in the literature [15–18]. In [18], a new definition of the underlying paths of IPMs for SDLCPs, using ordinary differential equations (ODEs), is proposed. The motivation for defining paths in this way is to relate the paths to the vector field of search directions of the IPM (see more details in [18]). In this paper, we use this definition of paths for SDLCPs to study the asymptotic behavior of the paths for general SDLCPs. As mentioned in earlier paragraphs, studying the asymptotic behavior of paths is important in the investigation of local convergence of IPMs for SDLCPs.

Throughout what follows, we restrict ourselves to the dual HKM direction and assume that the SDLCP satisfies strict complementarity. The HKM direction and its dual are among the most used directions in designing interior point algorithms, besides, the AHO and NT directions. The asymptotic analyticity behavior of off-central paths for the SDLCP using the AHO direction has been studied in [16, 17]. In [15], the asymptotic analyticity behavior of off-central paths for the SDP using the HKM direction is studied in general. The authors in [15] show that the off-central paths are analytic as a function of $\sqrt{\mu}$ in the limit. The off-central paths that the authors in [15–17] used are defined by algebraic equations and are not directly related to search directions of the IPM, while in [18] and this paper, they are defined using ODEs obtained from search directions. In [18], the asymptotic analyticity behavior of off-central paths, using the dual HKM direction, is investigated for a simple example. In this paper, we attempt to investigate the asymptotic behavior of off-central paths, using the dual HKM direction, for general SDLCPs.

In [18], it is shown, through an example, that there are two sets of off-central paths: paths in one set are analytic at $\mu = 0$ and those in the other set are not. For that example, the authors find a condition which characterizes analytic and nonanalytic paths. For general problems, similar conditions have not been found. In this paper, we show that an off-central path $(X(\mu), Y(\mu))$ (whose definition is given in Definition 2.1) is analytic with respect to $\sqrt{\mu}$ if and only if an off-diagonal submatrix of Y (or X) is analytic with respect to $\sqrt{\mu}$ and the submatrix is equal to $O(\mu)$ as $\mu \to 0$. This result is interesting on its own.

Another phenomenon observed in [18], again by an example, is that the first derivative of an off-central path with respect to μ is unbounded as $\mu \to 0$. A natural question is whether the first order derivative of an off-central path with respect to $\sqrt{\mu}$ is bounded as $\mu \to 0$. One may guess that the first order derivatives of those nonanalytic paths are likely to be unbounded near $\mu = 0$ even as a function of $\sqrt{\mu}$. Our study in this paper shows a fact contrary to this intuition.

In Sect. 2, we first define SDLCPs and off-central paths for SDLCPs. We also describe in detail a reformulation of the ODE system that described an off-central path for the SDLCP. The main result in Sect. 3 is a necessary and sufficient condition for when an off-central path, as a function of $\sqrt{\mu}$ (where μ is the parameter of the path, and proportional to the duality gap between the primal and dual variables), is analytic at a solution of the SDLCP. This condition is not intuitively obvious and may provide some insight into the study of asymptotic analyticity behavior of off-central paths. We also derive in this section a weak sufficient condition for convergence of an off-central path. In Sect. 4, we show that if the given SDLCP has a unique solution, then the first derivative of any off-central path, as a function of $\sqrt{\mu}$, is bounded. Finally, we give some concluding remarks and future directions in Sect. 5.

1.1 Notations and Common Definitions

The space of symmetric $n \times n$ matrices is denoted by S^n. Given the 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 semidefinite (resp., positive definite), we write $X \succeq 0$ (resp., $X \succ 0$). The cone of positive semidefinite (resp., positive definite) symmetric matrices is denoted by S^n_+ (resp., S^n_{++}). Either the identity matrix or the identity operator are denoted by I.

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

For a matrix $X \in \Re^{p\times q}$, we denote its component at the intersection of the ith row and jth column by X_{ij}. Also, $X_{i\cdot}$ denotes the ith row of X and $X_{\cdot j}$ the jth 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 the 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 the functions $f : \Omega \longrightarrow E$ and $g : \Omega \longrightarrow \Re_{++}$, with Ω an arbitrary set and E a normed vector space, and given 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 $w \in \widetilde{\Omega}$ and constant $M > 0$. Moreover, for a function $U : \Omega \longrightarrow S^n_{++}$, we write $U(w) = \Theta(g(w))$ for all $w \in \widetilde{\Omega}$ if $U(w) = O(g(w))$ and $U(w)^{-1} = O(g(w)^{-1})$ for all $w \in \widetilde{\Omega}$. The latter condition is equivalent to the existence of a constant $M > 0$ such that

$$\frac{1}{M} I \preceq \frac{1}{g(w)} U(w) \preceq MI, \quad \forall\, w \in \widetilde{\Omega}.$$

The meaning of the subset $\widetilde{\Omega}$ should be clear from the context whenever it is used. Usually, $\widetilde{\Omega} = (0, \bar{w})$ for a small $\bar{w} > 0$.

A function $f = (f_1, \ldots, f_m)$ from an open subset $\mathcal{O}$ of $\Re^k$ to $\Re^m$ is analytic at a point $x = (x_1, \ldots, x_k) \in \mathcal{O}$ if each f_i, $i = 1, \ldots, m$, can be written as a convergent power series expansion about $(x_1, \ldots, x_k)$ in an open neighborhood of x. Furthermore, if $x^0 \in \Re^k$ is on the boundary of $\mathcal{O}$, we say that f is analytic at x^0 (or can be extended analytically to x^0); we let $f(x^0) = \lim_{x \to x^0} f(x)$, if there exists an analytic function g which is analytic at x^0 and coincides with f wherever both are defined.

Note that the above applies also if an argument of f is a symmetric matrix, in which case, we consider the variable to lie in an Euclidean space of appropriate dimension. If the range of f is in the space of matrices, we consider also it to be in an appropriate Euclidean space when considering analyticity, so that the above applies.

2 Formulation and Reformulation of ODEs for the HKM Off-Central Path

Let us consider the following SDLCP:

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

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

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

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

We have the following assumption on the SDLCP throughout the paper:

Assumption 2.1

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

(b) *There exist* $X^1, Y^1 \succ 0$ *such that* $A(X^1) + B(Y^1) = q$.
(c) $\{A(X) + B(Y) : X, Y \in S^n\} = \Re^{\tilde{n}}$.

The above are basic assumptions used in the literature when SDLCP is studied in the context of IPMs. Besides Assumption 2.1, we need also another assumption in this paper, given below.

Assumption 2.2 *There exists a strictly complementary solution* (X^*, Y^*) *to the SDLCP* (1–3).

The analysis of the asymptotic behavior of an off-central path for a general SDLCP is considered to be difficult without this assumption (Assumption 2.2). However, we note that there has been some work done in this area for special classes of SDLCP without the assumption; see for example [19].

Let us now define the off-central path for SDLCP passing through a point (X^0, Y^0), $X^0, Y^0 \succ 0$, satisfying $A(X) + B(Y) = q$.

Definition 2.1 The solution $(X(\mu), Y(\mu))$, $\mu > 0$, to the equations

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

$$A(X') + B(Y') = 0, \tag{5}$$

with the initial condition $(X(1), Y(1)) = (X^0, Y^0)$, $X^0, Y^0 \succ 0$, is the 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.

Assuming that P is an analytic function of X, Y and that the matrix $PXYP^{-1}$ is always symmetric (such P include well-known directions like the HKM (and its dual) and NT directions), it is proved in [18] that the above definition is well-defined, and that $(X(\mu), Y(\mu))$, $X(\mu), Y(\mu) \succ 0$, is unique, analytic over $\mu \in (0, \infty)$. The motivation for defining an off-central path as in Definition 2.1 is also given in [18].

Remark 2.1 The central path $(X_c(\mu), Y_c(\mu))$ for SDLCP, which satisfies $X_c(\mu)Y_c(\mu) = \mu I$, is a special example of off-central path for SDLCP. When $\mu = 1$, it satisfies

$$\mathrm{Tr}(X_c(1)Y_c(1)) = n.$$

Therefore, we require also the initial data (X^0, Y^0) when $\mu = 1$ in (4), (5) to satisfy

$$\mathrm{Tr}(X^0Y^0) = n.$$

In this case, it is easy to see, using (4), that the parameter μ in the ODE system (4), (5) actually represents the duality gap, $X(\mu) \bullet Y(\mu)$, at the point $(X(\mu), Y(\mu))$ on the path.

Using the operation $\otimes_s$ and the map svec (with inverse smat), whose properties are given on pp. 775–776 and the Appendix of [20] (we have listed the properties of the operation $\otimes_s$ and the map svec in the Appendix to this paper for the readers' convenience), we can rewrite (4), (5) as

$$\begin{pmatrix} \mathrm{svec}(A_1)^T & \mathrm{svec}(B_1)^T \\ \vdots & \vdots \\ \mathrm{svec}(A_{\tilde{n}})^T & \mathrm{svec}(B_{\tilde{n}})^T \\ P \otimes_s (P^{-T} Y) & (PX) \otimes_s P^{-T} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(X') \\ \mathrm{svec}(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} 0 \\ \mathrm{svec}(H_P(XY)) \end{pmatrix}, \quad (6)$$

where $\tilde{n} = n(n+1)/2$.

As mentioned in Introduction, we consider only the dual HKM direction in this paper. This corresponds to $P = Y^{1/2}$ [21]. Therefore, (6) becomes

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X \otimes_s Y^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(X') \\ \mathrm{svec}(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} 0 \\ \mathrm{svec}(X) \end{pmatrix}, \quad (7)$$

with

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

As $\mu \to 0$, $(X(\mu), Y(\mu))$ will tend to the boundary of the feasible region. Thus, they are expected to be singular at the limit. Therefore, the left-hand matrix in (7) is not invertible, and may not be defined, in the limit as $\mu \to 0$ on an off-central path for SDLCP. Hence using (7) is not likely to yield results on the asymptotic behavior of off-central paths for SDLCP. To overcome this, we will make a transformation to (7). We wish that in the transformed system the coefficient matrix on the left-hand side will be invertible at $\mu = 0$, and the original and new systems have the same solution for $\mu > 0$. If such a new system can be formulated and its solution can be shown to be analytic (with respect to μ or $\sqrt{\mu}$) at the $\mu = 0$, then the solution of the original system can be analytically extended to $\mu = 0$. Therefore, the system of ODEs obtained after the transformation will provide us an appropriate platform to answer the question when an off-central path $(X(\mu), Y(\mu))$ converges and is analytic at its limit point.

We attempt only to study the analyticity of an off-central path at its limit point with respect to $\sqrt{\mu}$ instead of μ in this paper because $\sqrt{\mu}$ appears naturally in the off-diagonal entries of $X(\mu), Y(\mu)$, as shown in (8) and (9) below. This leads us to naturally investigate asymptotic behavior of $X(\mu), Y(\mu)$ with respect to $\sqrt{\mu}$ first.

In what follows, we suppress occasionally the dependence of a vector or matrix on its parameters for the sake of clarity. Whether these matrices or vectors are dependent on a parameter and the parameter involved should be clear from the context.

Let (X^*, Y^*) be a strictly complementary solution to the SDLCP (1–3), which exists by Assumption 2.2.

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

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

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

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 $S_{11} \in S^m$, $S_{22} \in S^{n-m}$ and $S_{12} \in \Re^{m\times(n-m)}$.

In order to transform the ODE system (7) into a more "manageable" system of ODEs, we perform a transformation of variables. For this purpose, we prove first a few lemmas below. These lemmas are adapted from [17].

Lemma 2.1 On an off-central path, $X(\mu)$, $Y(\mu)$ are bounded for $\mu > 0$ near 0.

Proof See [22]. □

Lemma 2.2 (see [17], Lemma 3.10) *$Y_{11}(\mu)$ and $X_{22}(\mu)$ are equal to $\mathcal{O}(\mu)$ and $\|X_{12}(\mu)\|$ and $\|Y_{12}(\mu)\|$ are equal to $\mathcal{O}(\sqrt{\mu})$.*

Proof See [22]. □

Lemma 2.3 (see [17], Lemma 3.11) *$X_{11}(\mu)$ and $Y_{22}(\mu)$ are equal to $\Theta(1)$, and $X_{22}(\mu)$ and $Y_{11}(\mu)$ are equal to $\Theta(\mu)$.*

Proof See [22]. □

The above lemmas show that, for an off-central path $(X(\mu), Y(\mu))$, we have

$$X(\mu) = \begin{pmatrix} X_{11} & \sqrt{\mu}\widetilde{X}_{12} \\ \sqrt{\mu}\widetilde{X}_{12}^T & \mu\widetilde{X}_{22} \end{pmatrix} \tag{8}$$

and

$$Y(\mu) = \begin{pmatrix} \mu\widetilde{Y}_{11} & \sqrt{\mu}\widetilde{Y}_{12} \\ \sqrt{\mu}\widetilde{Y}_{12}^T & Y_{22} \end{pmatrix}, \tag{9}$$

where $X_{11}, Y_{22}, \widetilde{X}_{22}, \widetilde{Y}_{11}$ are equal to $\Theta(1)$ and $\|\widetilde{X}_{12}(\mu)\|, \|\widetilde{Y}_{12}(\mu)\|$ are equal to $\mathcal{O}(1)$.

Letting

$$\widetilde{X}(\mu) = \begin{pmatrix} X_{11} & \widetilde{X}_{12} \\ \widetilde{X}_{12}^T & \widetilde{X}_{22} \end{pmatrix} \quad \text{and} \quad \widetilde{Y}(\mu) = \begin{pmatrix} \widetilde{Y}_{11} & \widetilde{Y}_{12} \\ \widetilde{Y}_{12}^T & Y_{22} \end{pmatrix},$$

we can then write

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

Lemma 2.4 $\widetilde{X}(\mu)$ and $\widetilde{Y}(\mu)$ are positive definite for all $\mu > 0$ and any of their accumulation points are also positive definite.

Proof See [22]. □

Let

$$X_1(t)=X(t^2), \qquad Y_1(t)=Y(t^2).$$

Similarly, let

$$\widetilde{X}_1(t)=\widetilde{X}(t^2) \quad \text{and} \quad \widetilde{Y}_1(t)=\widetilde{Y}(t^2).$$

Then, X_1, $\widetilde{X}_1$ and Y_1, $\widetilde{Y}_1$ are related by

$$X_1(t)=\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\widetilde{X}_1(t)\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}, \tag{10}$$

$$Y_1(t)=\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\widetilde{Y}_1(t)\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}. \tag{11}$$

To study the analyticity of $(X(\mu), Y(\mu))$ with respect to $\sqrt{\mu}$ at $\mu = 0$ is the same as studying the analyticity of $(X_1(t), Y_1(t))$ when $t = 0$. The following proposition shows that it suffices to do this by studying the analyticity of $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ at $t = 0$.

Proposition 2.1 $X_1(t)$ is analytic at $t = 0$ if and only if $\widetilde{X}_1(t)$ is analytic at $t = 0$. Similarly, $Y_1(t)$ is analytic at $t = 0$ if and only if $\widetilde{Y}_1(t)$ is analytic at $t = 0$.

Proof See [22]. □

Therefore, by the above proposition, we need study only the analyticity of $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ at $t = 0$ to conclude the property for $X_1(t)$ and $Y_1(t)$. An advantage of using $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ rather than $X_1(t)$ and $Y_1(t)$ is because their accumulation points are positive definite, by Lemma 2.4, which is a desirable property.

Hence, we are going to express the system of ODEs (7) in terms of $\widetilde{X}_1$ and $\widetilde{Y}_1$.

In terms of X_1 and Y_1, (7) becomes

$$\frac{1}{2}\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X_1 \otimes_s Y_1^{-1} \end{pmatrix}\begin{pmatrix} \mathrm{svec}(X_1') \\ \mathrm{svec}(Y_1') \end{pmatrix} = \frac{1}{t}\begin{pmatrix} 0 \\ \mathrm{svec}(X_1) \end{pmatrix}. \tag{12}$$

Let us reiterate again that, if we consider X_1 and Y_1 on an off-central path, then the matrix on the extreme left in (12) is not invertible and may not even be defined as t tends to zero (since Y_1^{-1} does not exist in the limit); hence it is not possible to analyze the asymptotic behavior of $X_1(t)$ and $Y_1(t)$ if we just use (12). This provides the motivation for us to express (12) in terms of $\widetilde{X}_1$ and $\widetilde{Y}_1$, after which we will see that further analysis is possible.

We have the following proposition:

Proposition 2.2 *The off-central path for SDLCP, $(X(\mu), Y(\mu))$, $\mu > 0$, is the solution of the system of ODEs* (7) *with $(X(1), Y(1)) = (X^0, Y^0)$, if and only if $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$, $t > 0$, is the solution to the following system of ODEs*:

$$\begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(\widetilde{X}_1') \\ \mathrm{svec}(\widetilde{Y}_1') \end{pmatrix}$$
$$= \begin{pmatrix} -\mathcal{G}(t) & -\mathcal{H}(t) \\ \frac{1}{t}\left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right) \otimes_s I & -\frac{1}{t}(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\left(\left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right) \otimes_s I\right) \end{pmatrix} \begin{pmatrix} \mathrm{svec}(\widetilde{X}_1) \\ \mathrm{svec}(\widetilde{Y}_1) \end{pmatrix}, \quad (13)$$

with $(\widetilde{X}_1(1), \widetilde{Y}_1(1)) = (X^0, Y^0)$.

Here $X(\mu)(= X_1(t))$, $\widetilde{X}_1(t)$ and $Y(\mu)(= Y_1(t))$, $\widetilde{Y}_1(t)$ are related by (10) *and* (11) *respectively, where $\mu = t^2$, and*

$$\mathcal{A}(t)_{k\cdot} = \begin{cases} \left(\mathrm{svec}\left(\begin{smallmatrix} (A_k)_{11} & t(A_k)_{12} \\ t(A_k)_{12}^T & t^2(A_k)_{22} \end{smallmatrix}\right)\right)^T, & 1 \le k \le i_1, \\ \left(\mathrm{svec}\left(\begin{smallmatrix} 0 & (A_k)_{12} \\ (A_k)_{12}^T & t(A_k)_{22} \end{smallmatrix}\right)\right)^T, & i_1 + 1 \le k \le i_1 + i_2, \\ \left(\mathrm{svec}\left(\begin{smallmatrix} 0 & 0 \\ 0 & (A_k)_{22} \end{smallmatrix}\right)\right)^T, & i_1 + i_2 + 1 \le k \le \tilde{n}, \end{cases} \quad (14)$$

$$\mathcal{B}(t)_{k\cdot} = \begin{cases} \left(\mathrm{svec}\left(\begin{smallmatrix} t^2(B_k)_{11} & t(B_k)_{12} \\ t(B_k)_{12}^T & (B_k)_{22} \end{smallmatrix}\right)\right)^T, & 1 \le k \le i_1, \\ \left(\mathrm{svec}\left(\begin{smallmatrix} t(B_k)_{11} & (B_k)_{12} \\ (B_k)_{12}^T & 0 \end{smallmatrix}\right)\right)^T, & i_1 + 1 \le k \le i_1 + i_2, \\ \left(\mathrm{svec}\left(\begin{smallmatrix} (B_k)_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right)\right)^T, & i_1 + i_2 + 1 \le k \le \tilde{n}, \end{cases} \quad (15)$$

$$\mathcal{G}(t)_{k\cdot} := \begin{cases} \left(\mathrm{svec}\left(\begin{smallmatrix} 0 & (A_k)_{12} \\ (A_k)_{12}^T & 2t(A_k)_{22} \end{smallmatrix}\right)\right)^T, & 1 \le k \le i_1, \\ \left(\mathrm{svec}\left(\begin{smallmatrix} 0 & 0 \\ 0 & (A_k)_{22} \end{smallmatrix}\right)\right)^T, & i_1 + 1 \le k \le i_1 + i_2, \\ 0, & i_1 + i_2 + 1 \le k \le \tilde{n}, \end{cases} \quad (16)$$

$$\mathcal{H}(t)_{k\cdot} := \begin{cases} \left(\mathrm{svec}\left(\begin{smallmatrix} 2t(B_k)_{11} & (B_k)_{12} \\ (B_k)_{12}^T & 0 \end{smallmatrix}\right)\right)^T, & 1 \le k \le i_1, \\ \left(\mathrm{svec}\left(\begin{smallmatrix} (B_k)_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right)\right)^T, & i_1 + 1 \le k \le i_1 + i_2, \\ 0, & i_1 + i_2 + 1 \le k \le \tilde{n}. \end{cases} \quad (17)$$

Proof See [22]. □

The importance of this proposition is that the coefficient matrix on the left-hand side is nonsingular for all $t \geq 0$ (even at $t = 0$), as will be shown in Proposition 2.3 below. This enables us to investigate the asymptotic behavior of the off-central paths as $t \to 0$ (or $\mu \to 0$).

In the following proposition, we observe an important property of the matrix $\begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}$ on the left-hand side of the system of equations (13).

Proposition 2.3 $\begin{pmatrix} \beta\mathcal{A}(t) & \beta\mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}$, *where* $\beta \neq 0$, $\beta \in \Re$, *is invertible for all* $t \geq 0$ *and* $\widetilde{X}_1$, $\widetilde{Y}_1$ *positive definite.*

Proof See [22]. □

Note that the matrix $\begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}$ in (13) is invertible at any accumulation point of $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$. This follows from Proposition 2.3 since any accumulation point of $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ is positive definite, by Lemma 2.4. This fact implies that the matrix is still well-defined and invertible at the limit as t tends to zero and this enables us to study the asymptotic behavior of $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$.

Using (13), we can give a necessary and sufficient condition for the pair $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ of an off-central path to be analytic at $t = 0$. This will be studied in the next section.

3 Asymptotic Analyticity Behavior of a HKM Off-Central Path

First, we have the following technical proposition:

Proposition 3.1 *Let* $(\widetilde{X}_1^*, \widetilde{Y}_1^*)$ *be an accumulation point of* $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ *of an off-central path as* t *approaches zero. Then,*

$$(\widetilde{Y}_1^*)_{12} = 0 \quad \Longleftrightarrow$$
$$(\widetilde{Y}_1^*)^{-1}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\widetilde{Y}_1^*\widetilde{X}_1^* + \widetilde{X}_1^*\widetilde{Y}_1^*\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}(\widetilde{Y}_1^*)^{-1} = \begin{pmatrix} 2(\widetilde{X}_1^*)_{11} & 0 \\ 0 & -2(\widetilde{X}_1^*)_{22} \end{pmatrix}.$$

Proof See [22]. □

With this technical proposition, the following proposition follows immediately.

Proposition 3.2 *Let* $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ *be a solution to the system of ODEs* (13) *for* $t > 0$. *Suppose* $\widetilde{X}_1(t)$ *and* $\widetilde{Y}_1(t)$ *converge as* $t \to 0$. *Then,* $\lim_{t\to 0}(\widetilde{Y}_1)_{12}(t) = 0$.

Proof See [22]. □

We are now ready to state a necessary and sufficient condition for $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ to be analytic at $t = 0$. We have the following theorem.

Theorem 3.1 *Let $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ be a solution to the system of ODEs (13) for $t > 0$. Then $\widetilde{X}_1(t)$, $\widetilde{Y}_1(t)$ are analytic at $t = 0$ if and only if $(\widetilde{Y}_1)_{12}(t)$ converges to zero as $t \to 0$ and is analytic at $t = 0$.*

Proof See [22]. □

From the sufficiency proof of Theorem 3.1, we observe that a sufficient condition for $\widetilde{X}_1(t)$, $\widetilde{Y}_1(t)$, and hence for an off-central path $(X(\mu), Y(\mu))$ to converge as t (or μ) tends to zero is $(\widetilde{Y}_1)_{12}(t) = O(t^\alpha)$, that is, $Y_{12}(\mu) = O(\mu^{0.5(1+\alpha)})$, for any $\alpha > 0$. Therefore, we have the following corollary.

Corollary 3.1 *Let $(X(\mu), Y(\mu))$ be an off-central path for the SDLCP (1–3), $\mu > 0$, under Assumptions 2.1 and 2.2. Suppose that $Y_{12}(\mu) = O(\mu^{0.5(1+\alpha)})$ for some $\alpha > 0$. Then, $(X(\mu), Y(\mu))$ converges as $\mu \to 0$.*

Proof See [22]. □

Remark 3.1 For the special case of the central path, Corollary 3.1 gives a convergence proof of the path to a solution of the SDLCP using the ODE approach, under the assumption of strict complementarity. Using algebraic geometry results, in [23], a convergence result is obtained for the central paths of SDPs, without the strict complementarity assumption.

Using Theorem 3.1, we have the main theorem for the section.

Theorem 3.2 *Let $(X(\mu), Y(\mu))$ be an off-central path for the SDLCP (1–3), $\mu > 0$, under Assumptions 2.1 and 2.2. Then $X(\mu)$, $Y(\mu)$ are analytic as a function of $t = \sqrt{\mu}$ at $t = 0$ if and only if $\lim_{\mu \to 0} Y_{12}(\mu)/\mu$ exists and the analyticity of $Y_{12}(\mu)/\mu$ as a function of $t = \sqrt{\mu}$ can be extended to $t = 0$.*

Proof See [22]. □

From Theorem 3.2, we see that the asymptotic analyticity of an off-central path for SDLCP as a function of $\sqrt{\mu}$ depends on only the asymptotic analyticity of one of its off-diagonal entries. This is a rather surprising result. From [18], we know that not all off-central paths are analytic at the solution of SDLCP. The above theorem gives a criterion as to when an off-central path for a general SDLCP is analytic at the solution.

To end this section, we remark that a similar theorem to Theorem 3.2 can also be stated for the HKM direction.

4 Boundedness of First Derivative of a HKM Off-Central Path

In [18], the authors show through a simple example that most off-central paths for the SDLCP, $(X(\mu), Y(\mu))$, have unbounded first derivatives as μ tends to zero. This suggests an undesirable consequence on the local convergence behavior of the IPM, using the dual HKM direction, on the SDLCP given the close relation between the boundedness of the derivatives of the off-central paths and the local behavior of interior point path-following algorithm when the iterates are near the solution of SDLCP. It turns out that an off-central path for SDLCP, $(X(\mu), Y(\mu))$, does not behave too badly if we perform a slight transformation on the parameter μ. We show in this section that if we consider $(X_1(t), Y_1(t)) = (X(t^2), Y(t^2))$, where $t = \sqrt{\mu}$, then the first derivatives of $X_1(t)$ and $Y_1(t)$ are bounded as t approaches zero. Note that we consider only the case when the SDLCP (1–3) has a unique solution. That is, we have an additional assumption on the SDLCP (1–3).

Assumption 4.1 *The SDLCP* (1–3) *has a unique solution* (X^*, Y^*), *which is strictly complementary.*

In this section, we assume without loss of generality that the SDLCP (1–3) that we consider has already undergone the various equivalent transformations that we made in Sect. 2. Hence, the unique solution (X^*, Y^*) can be written as $\left(\begin{pmatrix} X_{11}^* & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & Y_{22}^* \end{pmatrix}\right)$, where $X_{11}^*, Y_{22}^* \succ 0$.

By uniqueness of the solution to the given SDLCP, we have the following lemma.

Lemma 4.1 *If* $(U_{11}, V_{22}) \in S^m \times S^{n-m}$ *is such that*

$$\begin{pmatrix} (A_1)_{11} \bullet U_{11} + (B_1)_{22} \bullet V_{22} \\ \vdots \\ (A_{i_1})_{11} \bullet U_{11} + (B_{i_1})_{22} \bullet V_{22} \end{pmatrix} = q_1,$$

then $U_{11} = X_{11}^*$ *and* $V_{22} = Y_{22}^*$.

Proof See [22]. □

The above lemma plays an important role in the proof of the boundedness of the first derivatives of $X_1(t)$ and $Y_1(t)$ for t close to zero.

We have in Sect. 2 an ODE system for $(X_1(t), Y_1(t))$ given by

$$\frac{1}{2}\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X_1 \otimes_s Y_1^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(X_1') \\ \mathrm{svec}(Y_1') \end{pmatrix} = \frac{1}{t}\begin{pmatrix} 0 \\ \mathrm{svec}(X_1) \end{pmatrix}.$$

To analyze the behavior of X_1' and Y_1' as $t \to 0$, let us first invert the matrix on the left-hand side of (12) or the above system.

Therefore, we have, after simplifications,

$$\begin{pmatrix} \mathrm{svec}(X_1') \\ \mathrm{svec}(Y_1') \end{pmatrix} = \frac{1}{t} \begin{pmatrix} ((\begin{smallmatrix} I & 0 \\ 0 & tI \end{smallmatrix}) \otimes_s (\begin{smallmatrix} I & 0 \\ 0 & tI \end{smallmatrix}))(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\widetilde{\mathcal{G}}_1^{-1} q + \mathrm{svec}(X_1) \\ -((\begin{smallmatrix} tI & 0 \\ 0 & I \end{smallmatrix}) \otimes_s (\begin{smallmatrix} tI & 0 \\ 0 & I \end{smallmatrix}))\widetilde{\mathcal{G}}_1^{-1} q + \mathrm{svec}(Y_1) \end{pmatrix}, \tag{18}$$

where

$$\widetilde{\mathcal{G}}_1 = \mathcal{B}(t) - \mathcal{A}(t)(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1}).$$

Note that it is advantageous to use (18) to analyze the behavior of X_1' and Y_1' near t equal to zero, since $\widetilde{\mathcal{G}}_1$ is invertible for all $t \geq 0$ and $\widetilde{X}_1, \widetilde{Y}_1$ positive definite, by Proposition 2.3. Hence, the vector on the right-hand side of (18) is defined in the limit as t tends to zero for $(X_1(t), Y_1(t))$ of an off-central path.

We are now ready to state and prove the main theorem in this section.

Theorem 4.1 *Under Assumptions 2.1 and 4.1, given an off-central path for the SDLCP (1–3), $(X(\mu), Y(\mu))$, let $X_1(t) = X(t^2)$ and $Y_1(t) = Y(t^2)$. We have that $X_1'(t), Y_1'(t)$ are bounded near $t = 0$.*

Proof See [22]. □

5 Conclusion and Future Directions

In this paper we study the asymptotic behavior of an off-central path for the SDLCP, using the dual HKM direction. The purpose of this paper is to provide a framework upon which the asymptotic behavior of an off-central path for SDLCP, using the dual HKM direction, can be analyzed, using (13) and (18), which can reveal more about the properties of off-central paths than (12) near $t = 0$. From a practical point of view, we are left with the following open questions:

(Q1) Given a problem in a specific class of SDLCP, how to determine if its paths are all analytic, all nonanalytic, or a mixture?

(Q2) If a problem has both analytic and nonanalytic paths, what are the practical conditions to distinguish them?

We do not attempt to answer these questions in this paper. In Sect. 3, we give a necessary and sufficient condition for when an off-central path is analytic as a function of $\sqrt{\mu}$ at a solution of SDLCP. This condition is closely related to the analysis of the asymptotic analytic behavior of the paths for the example in [18]. In [18], we obtain an algebraic condition for the asymptotic analyticity of the paths for the example considered there. Here, we are unable to obtain a similar algebraic condition and further analysis needs to be done in future to obtain a more practical necessary and sufficient condition for the asymptotic analyticity. The asymptotic analyticity of the off-central paths as a function of μ for general SDLCP will also be investigated as future work. In Sect. 4, we show that an off-central path for the SDLCP, when viewed as a function of $t = \sqrt{\mu}$, has bounded first derivative as t approaches zero.

We assume that the SDLCP has a unique solution which is strictly complementary in the section. Whether the same result holds without the uniqueness assumption is still an open question. In [18], it indicates, through an example, that the usual interior point path-following algorithm, based on paths as a function of μ (where μ represents the duality gap between the primal and dual variables), may not converge fast to a solution of the SDLCP in general, since the first derivatives of the paths for the example are unbounded as μ tends to zero. The results in this section suggests that it may be worthwhile to investigate and design interior point path-following algorithm, using underlying paths as a function of $\sqrt{\mu}$, instead of μ, whose iterates possibly converge rapidly to the unique solution of the SDLCP. A similar study on such new interior point path-following algorithm has been done for the LCPs in [14, 24, 25], where a parametrization different from the usual one is used for the underlying paths, as in this paper.

Appendix

If U is an $n \times n$ symmetric matrix, then $\mathrm{svec}(U)$ is defined by

$$\mathrm{svec}(U) := (u_{11}, \sqrt{2}u_{21}, \ldots, \sqrt{2}u_{n1}, u_{22}, \sqrt{2}u_{32}, \ldots, \sqrt{2}u_{n2}, \ldots, u_{nn})^T.$$

Properties of the symmetrized Kronecker product, $\otimes_s$, used here are:

- $(G \otimes_s K)\mathrm{svec}(H) = \frac{1}{2}\mathrm{svec}(KHG^T + GHK^T)$.
- $\mathrm{svec}(G)^T \mathrm{svec}(K) = G \bullet K$.
- $G \otimes_s K = K \otimes G$.
- $(G \otimes_s K)^T = G^T \otimes_s K^T$.
- $(G \otimes_s K)(H \otimes_s L) = \frac{1}{2}((GH) \otimes_s (KL) + (GL) \otimes_s (KH))$.
- If G and K are symmetric and positive definite, then so is $G \otimes_s K$.
- If G is invertible, then $G \otimes_s G$ is invertible and $(G \otimes_s G)^{-1} = G^{-1} \otimes_s G^{-1}$.

References

1. Sonnevend, G.: An analytic center for polyhedrons and new classes for linear programming. In: Prekopa, A. (ed.) System Modelling and Optimization. Lecture Notes in Control and Information Sciences, vol. 84, pp. 866–876. Springer, Berlin (1985)
2. Bayer, D.A., Lagarias, J.C.: The nonlinear geometry of linear programming, Part I. Trans. Am. Math. Soc. **314**, 499–526 (1989)
3. Bayer, D.A., Lagarias, J.C.: The nonlinear geometry of linear programming, Part II. Trans. Am. Math. Soc. **314**, 527–581 (1989)
4. Bayer, D.A., Lagarias, J.C.: The nonlinear geometry of linear programming, Part III. Trans. Am. Math. Soc. **320**, 193–225 (1990)
5. Sonnevend, G., Stoer, J., Zhao, G.: On the complexity of following the central path of linear programs by linear extrapolation. Method. Oper. Res. **62**, 19–31 (1989)
6. Sonnevend, G., Stoer, J., Zhao, G.: On the complexity of following the central path of linear programs by linear extrapolation II. Math. Program. **52**, 527–553 (1991)
7. Zhao, G., Stoer, J.: Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals. Appl. Math. Optim. **27**, 85–103 (1993)
8. Zhao, G.: On the relationship between the curvature integral and the complexity of path-following methods in linear programming. SIAM J. Optim. **6**, 57–73 (1996)

9. Zhao, G., Zhu, J.: The curvature integral and the complexity of linear complementarity problems. Math. Program. **70**, 107–122 (1996)
10. Vavasis, S.A., Ye, Y.: A primal-dual interior point method whose running time depends only on the contraint matrix. Math. Program. **74**, 79–120 (1996)
11. Mehrotra, S.: Quadratic convergence in a primal-dual method. Math. Oper. Res. **18**, 741–751 (1993)
12. Ye, Y., Anstreicher, K.: On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP. Math. Program. **62**, 537–551 (1993)
13. Ye, Y., Güler, O., Tapia, R.A., Zhang, Y.: A quadratically convergence $O(\sqrt{n}L)$-iteration algorithm for linear programming. Math. Program. **59**, 151–162 (1993)
14. Sturm, J.F.: Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists. Math. Oper. Res. **24**, 72–94 (1999)
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. Lu, Z., Monteiro, R.D.C.: Limiting behavior of the Alizadeh-Haeberly-Overton weighted paths in semidefinite programming. Optim. Method. Softw. **22**, 849–870 (2003)
17. Preiß, M., Stoer, J.: Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems Math. Program. A **99**, 499–520 (2004)
18. Sim, C.-K., Zhao, G.: Underlying paths in interior point methods for the monotone semidefinite linear complementarity problem. Math. Program. A **110**, 475–499 (2007)
19. 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. A **103**, 487–514 (2005)
20. Todd, M.J., Toh, K.C., Tütüncü, R.H.: On the Nesterov-Todd direction in semidefinite programming. SIAM J. Optim. **8**, 769–796 (1998)
21. Zhang, Y.: On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optim. **8**, 365–386 (1998)
22. Asymptotic behavior of HKM paths in interior point methods for monotone semidefinite linear complementarity problems: General theory. The Logistics Institute—Asia Pacific Internal Report. See also http://www.math.nus.edu.sg/~matzgy/publist.html
23. Halická, M., De Klerk, E., Roos, C.: On the convergence of the central path in semidefinite optimization. SIAM J. Optim. **12**, 1090–1099 (2002)
24. Stoer, J., Wechs, M., Mizuno, S.: High order infeasible-interior-point methods for solving sufficient linear complementarity problems. Math. Oper. Res. **23**, 832–862 (1998)
25. Zhao, G., Sun, J.: On the rate of local convergence of high-order-infeasible-path-following algorithms for P_*-linear complementarity problems. Comput. Optim. Appl. **14**, 293–307 (1999)