Math. Program., Ser. A (2007) 110:475–499
DOI 10.1007/s10107-006-0010-7

FULL LENGTH PAPER

Underlying paths in interior point methods for the monotone semidefinite linear complementarity problem

Chee-Khian Sim · Gongyun Zhao

Received: 8 March 2005 / Accepted: 23 May 2006 /
Published online: 11 July 2006

Abstract An interior point method defines a search direction at each interior point of the feasible region. The search directions at all interior points together form a direction field, which gives rise to a system of ordinary differential equations (ODEs). Given an initial point in the interior of the feasible region, the unique solution of the ODE system is a curve passing through the point, with tangents parallel to the search directions along the curve. We call such curves off-central paths. We study off-central paths for the monotone semidefinite linear complementarity problem (SDLCP). We show that each off-central path is a well-defined analytic curve with parameter μ ranging over $(0, \infty)$ and any accumulation point of the off-central path is a solution to SDLCP. Through a simple example we show that the off-central paths are not analytic as a function of $\sqrt{\mu}$ and have first derivatives which are unbounded as a function of μ at $\mu = 0$ in general. On the other hand, for the same example, we can find a

This research was done during the author's PhD study at the Department of Mathematics, NUS and as a Research Engineer at the NUS Business School.

C-K. Sim
The Logistics Institute - Asia Pacific,
Block AS6, Level 5, 11 Law Link,
Singapore 119260, Singapore
e-mail: tlisck@nus.edu.sg

G. Zhao (✉)
Department of Mathematics,
National University of Singapore,
2 Science Drive 2, Singapore 117543,
Singapore
e-mail: matzgy@nus.edu.sg

subset of off-central paths which are analytic at $\mu = 0$. These "nice" paths are characterized by some algebraic equations.

1 Introduction

1.1 How are paths related to IPM and What roles do paths play in IPM?

The notion of a central path was introduced by Sonnevend [24] in 1985 with regard to interior point methods (IPMs). Since then, people have realized that an IPM is actually a homotopy method following underlying paths (central and off-central paths) and that many remarkable properties of IPMs are attributed to the nice geometry of the underlying paths. Readers who are interested in the basic geometry of underlying paths may refer to [3].

In [25,26,36,37,39] 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 of the complexity of IPMs. On the other hand, in [32] the authors showed 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. This topic is further studied in [13] and [19].

Another important role underlying paths play in the study of IPMs is to show 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 central path in an IPM may be singular at the optimal solution. Thus traditional approach of local convergence analysis does not work for IPMs. The fast local convergence of IPMs has instead been successfully proved by relating it to the boundedness of derivatives of the underlying paths in, e.g., [14,29,33,34].

The study of fast local convergence is particularly important for the semidefinite linear complementarity problem (SDLCP), with the semidefinite programming (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.

1.2 How are the underlying paths defined?

We assume that the reader is familiar with the definition of underlying paths for LP and LCP. Here we will concentrate on SDLCP (and view SDP as a special case).

The central path, i.e., the set of analytic centers, cf.[24], is defined by the following system:

$$\begin{aligned} F(X, Y) &= 0 \\ XY &= \mu I, \end{aligned} \tag{1}$$

where $\mu > 0$ is the parameter, $X, Y \in S^n_+$, and $F : S^n_+ \times S^n_+ \to R^{\tilde{n}}$ is linear, $\tilde{n} = n(n+1)/2$. Since staying on the central path is not practical for computation, investigation of neighboring paths is necessary. A natual way to define neighboring paths, i.e., off-central paths, is to replace I in (1) with a positive definite matrix M. For SDLCP, symmetrization of search direction is required, cf. [10,15]. A straightforward way to extend the central path to symmetrized off-central paths is to define them by

$$\begin{aligned} F(X, Y) &= 0 \\ H_P(XY) &= \mu M, \end{aligned} \tag{2}$$

where $H_P(U) := 1/2(PUP^{-1} + (PUP^{-1})^T)$ and $P \in \Re^{n \times n}$ is an invertible matrix. Such paths have been studied in [11,17,20,22]. The paper [5] also proposes a definition of paths which requires a Cholesky factorization in addition to the algebraic Eq. (2).

While such an off-central path is a natual extension of the central path and imitates certain geometry of the central path, it does not seem to be directly related to the main ingredient of IPMs: search directions. The complexity and convergence speed of IPMs are determined by the search directions, in particular, affine-scaling directions through which the duality gap is reduced.

As mentioned before, an IPM is a homotopy method whose search directions are determined by the underlying paths. This motivates us to define paths in connection with the field of affine-scaling directions. That is, a path is a curve in the interior of the feasible region such that the tangent of the curve at any point coincides with the direction in the field at this point. More precisely, such paths are defined by a system of ordinary differential equations (ODEs), which will be presented in Sect. 2.

1.3 Main results

In this paper, we propose a new definition for the underlying paths of an IPM for solving SDLCP. Our underlying paths possess an important property: the tangent of each path at a point coincides with the symmetrized Newton direction at that point. This property is not found in any existing definitions of paths for SDLCP with symmetrization other than AHO.

The study of off-central paths can be divided into two parts: (a) a study of the paths for $\mu > 0$ and (b) a study of the limiting behavior of the paths.

In the first part, we present a few basic properties, including the well-posedness of the ODE system for defining the paths and the analyticity of the paths for $\mu > 0$. The smoothness of the paths that a homotopy method, like IPM, follows is essential to the efficiency of the method. For instance, the relation between the complexity of an IPM for finding an ϵ-approximate solution and the curvature integral of the central path [36,39] is based on the analyticity of the underlying paths (both central and off-central) for $\mu \in [\epsilon, \mu_0]$.

In the second part, we determine (a) whether a limit point of a path is a solution of SDLCP and (b) whether a path is analytic or has bounded derivatives at the limit point of the path if the path converges. We will show a confirmative answer to the first question, i.e., any accumulation point of a path is a solution to the SDLCP. For the second question, there is a rather surprising observation. While all existing study (no matter for LCP [27] or for SDLCP [11,12,22]), assuming the existence of strictly complementary solutions, has shown that any underlying path (with generic definition given by (2) in the case of SDLCP) is analytic as a function of μ or $\sqrt{\mu}$ at its limit point, we have observed, through an example, that an off-central path, no matter how close it is to the central path, need not be analytic at its limit point, even with respect to $\sqrt{\mu}$. We also show that these off-central paths have unbounded first derivatives as $\mu \to 0$. Some conclusions can be drawn from this observation. First of all, the AHO direction where $P = I$ and the other directions where P depends on X and Y are essentially different. Analyticity of off-central paths at a limit point for AHO-IPM has been shown by [12,22] while non-analyticity of off-central paths at a limit point for HKM-IPM is observed through this example. Secondly, the central path and off-central paths are essentally different. The central path has been proven to be analytic at its limit point while off-central paths are observed to be non-analytic even as a function of $\sqrt{\mu}$ at their limit points, no matter how close the path is to the central path. To date, such an essential difference between central path and off-central paths has not been observed in other conic programming problems.

1.4 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 semidefinite (resp., definite), we write $X \succeq 0$ (resp., $X \succ 0$). The cone of positive semidefinite (resp., definite) symmetric matrices is denoted by S^n_+ (resp., S^n_{++}). Either the identity matrix or operator will be denoted by I. $\|\cdot\|$ for a vector in $\Re^n$ refers the Euclidean norm and for a matrix in $\Re^{p\times q}$, it refers to the Frobenius norm.

Given function $f : \Omega \longrightarrow E$ 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 $w \in \widetilde{\Omega}$ and a 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 subset $\widetilde{\Omega}$ should be clear from the context whenever it is used. Usually, $\widetilde{\Omega} = (0, \bar{w})$ for a small $\bar{w} > 0$.

2 Definition and basic properties of off-central paths

In this section, we elaborate further on what was described in Sect. 1.2 on how we obtain our definition of an off-central path by considering path-following interior point algorithm. We also prove the existence of such off-central path.

Let us consider the following SDLCP:

$$\begin{aligned} XY &= \ 0 \\ A(X) + B(Y) &= \ q \\ X, Y &\in \mathcal{S}^n_+, \end{aligned} \tag{3}$$

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 assumptions on SDLCP:

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 exists $X^1, Y^1 \succ 0$ such that $A(X^1) + B(Y^1) = q$.

In the predictor step of the predictor–corrector path-following algorithm, the algorithm searches a new point in the *affine scaling direction*, which is defined as the symmetrized Newton direction for the system $XY = 0$ and $A(X) + B(Y) = q$, more precisely, by the system (cf. [35])

$$H_P(X\Delta Y + \Delta XY) = -H_P(XY) \tag{4}$$

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

where $H_P(U) := 1/2(PUP^{-1} + (PUP^{-1})^T)$ and $P \in \Re^{n\times n}$ is an invertible matrix.

Because the affine scaling direction is aiming at an optimal solution at which $XY = 0$, it is more convenient to consider the direction field as comprising of scaled affine scaling directions $-1/\mu(\Delta X, \Delta Y)$, where μ is a parameter proportional to $\mathrm{Tr}(XY)$. By letting the derivative of a path to coincide with this direction, i.e., $(X', Y') = -1/\mu(\Delta X, \Delta Y)$, we obtain from (4) to (5) the following ODE system

$$H_P(XY' + X'Y) = \frac{1}{\mu} H_P(XY) \tag{6}$$

$$A(X') + B(Y') = 0 \tag{7}$$

$$(X, Y)(1) = (X^0, Y^0), \tag{8}$$

where $X^0, Y^0 \succ 0$ and $A(X^0) + B(Y^0) = q$. We only consider solution, $(X(\mu), Y(\mu))$, to (6)–(8) such that $X(\mu), Y(\mu) \in S^n_{++}$. We called such solution the *off-central path* of SDLCP (3) with respect to P and passing through (X^0, Y^0).

For the AHO direction, $P = I$. Hence (6) reduces to

$$(XY + YX)' = \frac{1}{\mu}(XY + YX).$$

This and (7) with the initial condition at $\mu = 1$ yield the algebraic equations

$$(XY + YX) = \mu(X^0Y^0 + Y^0X^0)$$

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

For other directions, such as the HKM and NT directions, P is a function of (X, Y), thus it is not possible to solve (6)–(8) to get an algebraic expression. This is an aspect which distinguishes the other directions from the AHO direction. Significant distinctions between off-central paths for AHO direction and for the other directions can be observed by comparing results in [12,22] and this paper. In the rest of this paper we will consider the case in which P is a nonconstant analytic function of (X, Y), such as the HKM and NT directions.

It is well-known that the central path is defined by

$$(XY)(\mu) = \mu I \tag{9}$$

$$A(X(\mu)) + B(Y(\mu)) = q \tag{10}$$

in the literature of interior point methods. It is easy to see (by differentiating these equations) that the central path satisfies the ODE system (6)–(8) for any P and for the initial point (X^0, Y^0) such that $X^0Y^0 = I$. Thus, the path defined by the system of ODEs (6)–(8) with $X^0Y^0 = I$ is the *central path*. The existing research [based on (9)–(10)] has shown that this central path, if SDLCP (3) satisfies strict complementarity condition, possesses many nice properties, in particular, it can be analytically (with respect to μ) extended to $\mu = 0$ [7]. In general, however, solutions of the ODE system (6)–(8) with different initial points do not satisfy an algebraic system and cannot be analytically extended to $\mu = 0$ (even with respect to $\sqrt{\mu}$, as will be discussed later). This implies that the central path and off-central path are essentially different.

If $(X(\mu), Y(\mu))$ is the solution of the system of ODEs with initial point (X^0, Y^0), then for any scalar $\alpha > 0$, $(\alpha X(\mu), \alpha Y(\mu))$ is also the solution of the system of ODEs but with initial point $(\alpha X^0, \alpha Y^0)$. Thus, we can always scale an initial (X^0, Y^0) such that $\mathrm{Tr}(X^0 Y^0) = n$. The reason for such scaled initial points can be seen in Lemma 2.1 below.

As in [30], we only consider P such that $PXYP^{-1}$ is symmetric. We also assume P is an analytic function of $X, Y \succ 0$. Such P include the well-known directions like the HKM and NT directions. Therefore, we have the following further assumptions:

Assumption 2.2

(a) The initial data (X^0, Y^0) in (8) when $\mu = 1$ satisfies $Tr(X^0 Y^0) = n$.
(b) The matrix P in (6) is such that $PXYP^{-1}$ is symmetric and P is also an analytic function of $X, Y \succ 0$.

Lemma 2.1 *Under Assumption* 2.2(a)*, for any invertible matrix P, the solution of the ODE system* (6)–(7)*, $(X(\mu), Y(\mu))$, satisfies*

$$Tr(XY)(\mu) = n\mu. \tag{11}$$

Proof Because

$$\mathrm{Tr}(H_P(U)) = \mathrm{Tr}(U) \quad \text{and} \quad \mathrm{Tr}(U') = (\mathrm{Tr}(U))'$$

for any P and U, taking "Tr" on both sides of (6), we obtain

$$(\mathrm{Tr}(XY))' = \frac{1}{\mu}\mathrm{Tr}(XY).$$

This leads to

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

□

Remark 2.1 Under Assumption 2.2(a), we see from Lemma 2.1 that the parameter μ in the ODE system (6)–(7) actually represents the duality gap, $X(\mu) \bullet Y(\mu)$, at the point $(X(\mu), Y(\mu))$ on the path.

We will show next that, given the initial point $(X^0, Y^0) \in S^n_{++} \times S^n_{++}$, the solution to (6)–(8), $(X(\mu), Y(\mu)) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, exists over $\mu \in (0, \infty)$ and is unique and analytic. [Recall that a function $f = (f_1, \ldots, f_m)$ from a subset $\mathcal{O}$ of $\Re^k$ to $\Re^m$ is analytic at a point $x = (x_1, \ldots, x_k)$ if f is defined in an open neighborhood of x and each f_i, $i = 1, \ldots, m$, can be written as a convergent power series expansion about $(x_1, \ldots, x_k)$ in this open neighborhood.] Thus, it defines a path for the SDLCP.

We are going to use a result from ODE theory, taken from [2, p. 100, 4, p. 196], and their theorem and corollary are combined as a theorem below for completeness:

Theorem 2.1 *Assume that a function f is continuously differentiable from $J \times D$ to E, where $J \subset \Re$ is an open interval, E is a finite dimensional Banach space over $\Re$, $D \subset E$ is open. Then for every $(t_0, x_0) \in J \times D$, there exists an unique nonextensible solution*

$$u(\cdot; t_0, x_0) : J(t_0, x_0) \to D$$

of the IVP

$$\dot{x} = f(t, x), \;\; x(t_0) = x_0.$$

The maximal interval of existence $J(t_0, x_0) := (t^-, t^+)$ is open. We either have

$$t^- = \inf J, \text{ resp. } t^+ = \sup J,$$

or

$$\lim_{t \to t^{\pm}} \min\{\text{dist}(u(t, t_0, x_0), \partial D), \|u(t; t_0, x_0)\|^{-1}\} = 0.$$

(We use the convention: $\text{dist}(x, \emptyset) = \infty$.*)*

When f is analytic over $J \times D$, where $D \subset E = \Re^n$, the solution u is analytic over $J(t_0, x_0)$.

In order to use Theorem 2.1, we need to express (6)–(7) in the form of IVP as in the theorem.

Now, (6) can be written as

$$((PX) \otimes_s P^{-T})\text{svec}(Y') + (P \otimes_s (P^{-T}Y))\text{svec}(X') = \frac{1}{\mu}\text{svec}(H_P(XY)).$$

Remark 2.2 Note that 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 [30].

Writing (7) in a similar way using *svec*, we can rewrite (6)–(7) as

$$\begin{pmatrix} \text{svec}(A_1)^T & \text{svec}(B_1)^T \\ \vdots & \vdots \\ \text{svec}(A_{\tilde{n}})^T & \text{svec}(B_{\tilde{n}})^T \\ P \otimes_s (P^{-T}Y) & (PX) \otimes_s P^{-T} \end{pmatrix} \begin{pmatrix} \text{svec}(X') \\ \text{svec}(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} 0 \\ \text{svec}(H_P(XY)) \end{pmatrix}, \quad (12)$$

which is another form of (6)–(7).

Similar to [30], it can be shown that the matrix in (12) is invertible for all $X, Y \succ 0$ under Assumption 2.2(b).

Let the matrix in (12) be denoted by $\mathcal{A}(X, Y)$. Then $\mathcal{A}(X, Y)$ is invertible for all $X, Y \succ 0$. Therefore, we can write (12) in the IVP form as

$$\begin{pmatrix} \mathrm{svec}(X') \\ \mathrm{svec}(Y') \end{pmatrix} = \mathcal{F}(\mu, X, Y),$$

where

$$\mathcal{F}(\mu, X, Y) = \frac{1}{\mu}\mathcal{A}^{-1}(X, Y)\begin{pmatrix} 0 \\ \mathrm{svec}(H_P(XY)) \end{pmatrix}.$$

One can see that $\mathcal{F}$ is analytic on $\Re_{++} \times (\mathcal{S}^n_{++} \times \mathcal{S}^n_{++})$. Hence, by Theorem 2.1, given $(1, (X^0, Y^0)) \in \Re_{++} \times (\mathcal{S}^n_{++} \times \mathcal{S}^n_{++})$ (where $A(X^0) + B(Y^0) = q$), there exists a *maximal interval of existence*

$$J_0 = (\mu^-, \mu^+) \subseteq \Re_{++} \tag{13}$$

and an unique analytic solution $X, Y : J_0 \longmapsto \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ of the IVP

$$\begin{pmatrix} \mathrm{svec}(X') \\ \mathrm{svec}(Y') \end{pmatrix} = \mathcal{F}(\mu, X, Y), \quad X(1) = X^0, \quad Y(1) = Y^0. \tag{14}$$

We have either

$$\begin{gathered} \mu^- = 0, \quad \text{resp. } \mu^+ = +\infty \quad \text{or} \\ \lim_{\mu \to \mu^\pm} \min\{\mathrm{dist}((\mathrm{X}(\mu), \mathrm{Y}(\mu)), \partial(\mathcal{S}^{\mathrm{n}}_{++} \times \mathcal{S}^{\mathrm{n}}_{++})), \|(\mathrm{X}(\mu), \mathrm{Y}(\mu))\|^{-1}\} = 0. \end{gathered} \tag{15}$$

It is not apparent so far what values μ^- and μ^+ should take. We will determine them later. Now we first show that, although $X(\mu)$ and $Y(\mu)$ do not satisfy the algebraic equation $X(\mu)Y(\mu) = \mu X^0 Y^0$, the maximum and minimum eigenvalues of $X(\mu)Y(\mu)$ do. This result is useful as we will see later on.

Theorem 2.2 *For all $\mu \in J_0$, $\lambda_{\min}(XY)(\mu) = \lambda_{\min}(X^0 Y^0)\mu$ and $\lambda_{\max}(XY)(\mu) = \lambda_{\max}(X^0 Y^0)\mu$.*

Proof Recall that P in (6) is invertible and an analytic function of X, Y. Therefore, with $X(\mu), Y(\mu)$ analytic with respect to μ, we have $P = P(\mu)$ is analytic with respect to μ. Also, $P(\mu)$ satisfies $(PXYP^{-1})(\mu) = ((PXYP^{-1})(\mu))^T$. We are going to use the latter two facts in the proof here.

For $\mu \in J_0$. Let $v_0 \in \Re^n$, $\|v_0\| = 1$, be such that $H_{P(\mu)}((XY)(\mu))v_0 = \lambda_{\min}(H_{P(\mu)}((XY)(\mu)))v_0 = \lambda_{\min}(XY)(\mu)v_0$. (The last equality holds because $(PXYP^{-1})(\mu)$ is symmetric.)

Therefore, by (6) and this choice of v_0, we have

$$v_0^T H_{P(\mu)}((XY)'(\mu))v_0 = \frac{1}{\mu}\lambda_{\min}(XY)(\mu).$$

We now focus our attention on the left-hand expression of the above equality. We have

$$\begin{aligned}
&v_0^T H_{P(\mu)}((XY)'(\mu))v_0 \\
&\quad = \limsup_{h\to 0^+} v_0^T \left(\frac{H_{P(\mu)}((XY)(\mu+h)) - H_{P(\mu)}((XY)(\mu))}{h}\right) v_0 \\
&\quad = \limsup_{h\to 0^+} (v_0^T H_{P(\mu)}((XY)(\mu+h))v_0 - \lambda_{\min}(XY)(\mu))/h \\
&\quad \geq \liminf_{h\to 0^+} (v_0^T H_{P(\mu)}((XY)(\mu+h))v_0 - v_0^T H_{P(\mu+h)}((XY)(\mu+h))v_0)/h \\
&\qquad + \limsup_{h\to 0^+} (v_0^T H_{P(\mu+h)}((XY)(\mu+h))v_0 - \lambda_{\min}(XY)(\mu))/h \\
&\quad \geq \liminf_{h\to 0^+} (v_0^T H_{P(\mu)}((XY)(\mu+h))v_0 - v_0^T H_{P(\mu+h)}((XY)(\mu+h))v_0)/h \\
&\qquad + \limsup_{h\to 0^+} (\min_{\|v\|=1} v^T H_{P(\mu+h)}((XY)(\mu+h))v - \lambda_{\min}(XY)(\mu))/h \\
&\quad = \liminf_{h\to 0^+} (v_0^T H_{P(\mu)}((XY)(\mu+h))v_0 - v_0^T H_{P(\mu+h)}((XY)(\mu+h))v_0)/h \\
&\qquad + \limsup_{h\to 0^+} (\lambda_{\min}(XY)(\mu+h) - \lambda_{\min}(XY)(\mu))/h.
\end{aligned}$$

Let $f(\xi) = v_0^T P(\mu+\xi)(XY)(\mu+h)P^{-1}(\mu+\xi)v_0$.

Therefore,

$$\liminf_{h\to 0^+} (v_0^T H_{P(\mu)}((XY)(\mu+h))v_0 - v_0^T H_{P(\mu+h)}((XY)(\mu+h))v_0)/h$$

in above is equal to

$$-\liminf_{h\to 0^+} \frac{f(h)-f(0)}{h} = -\liminf_{h\to 0^+} f'(\xi_h),$$

where the last equality follows from the mean value theorem and $0 < \xi_h < h$.

Let us try to find the value of the last limit.

We have

$$\begin{aligned}
f'(\xi_h) &= v_0^T P'(\mu+\xi_h)(XY)(\mu+h)P^{-1}(\mu+\xi_h)v_0 \\
&\quad + v_0^T P(\mu+\xi_h)(XY)(\mu+h)(P^{-1})'(\mu+\xi_h)v_0 \\
&= v_0^T P'(\mu+\xi_h)(XY)(\mu+h)P^{-1}(\mu+\xi_h)v_0 \\
&\quad - v_0^T P(\mu+\xi_h)(XY)(\mu+h)P^{-1}(\mu+\xi_h)P'(\mu+\xi_h)P^{-1}(\mu+\xi_h)v_0.
\end{aligned}$$

Hence,

$$\begin{aligned}
\liminf_{h\to 0^+} f'(\xi_h) &= v_0^T P'(\mu)(XY)(\mu)P^{-1}(\mu)v_0 \\
&\quad -v_0^T P(\mu)(XY)(\mu)P^{-1}(\mu)P'(\mu)P^{-1}(\mu)v_0 \\
&= v_0^T P'(\mu)P^{-1}(\mu)(P(\mu)(XY)(\mu)P^{-1}(\mu)v_0) \\
&\quad -(P(\mu)(XY)(\mu)P^{-1}(\mu)v_0)^T P'(\mu)P^{-1}(\mu)v_0 \\
&= \lambda_{\min}(XY)(\mu)v_0^T P'(\mu)P^{-1}(\mu)v_0 \\
&\quad -\lambda_{\min}(XY)(\mu)v_0^T P'(\mu)P^{-1}(\mu)v_0 \\
&= 0,
\end{aligned}$$

where the second equality follows from $(PXYP^{-1})(\mu) = ((PXYP^{-1})(\mu))^T$ and the third equality follows from $(PXYP^{-1})(\mu)v_0 = H_{P(\mu)}(XY(\mu))v_0 = \lambda_{\min}(XY)(\mu)v_0$.

Therefore,

$$\frac{1}{\mu}\lambda_{\min}(XY)(\mu) \geq \limsup_{h\to 0^+} \frac{\lambda_{\min}(XY)(\mu+h) - \lambda_{\min}(XY)(\mu)}{h}.$$

On the other hand, consider (in what follows, in order to make reading easier, we suppress the dependence of P on μ)

$$\min_{\|v\|=1} v^T H_P((XY)'(\mu))v$$

which is equal to

$$\lim_{h\to 0^+}\left(\min_{\|v\|=1} v^T H_P\left(\frac{(XY)(\mu+h) - (XY)(\mu)}{h}\right)v\right).$$

Let $v_1 \in \Re^n$, $\|v_1\| = 1$ be such that

$$H_P((XY)(\mu+h))v_1 = \lambda_{\min}(H_P((XY)(\mu+h)))v_1.$$

Therefore, we have

$$\begin{aligned}
&\min_{\|v\|=1} v^T H_P\left(\frac{(XY)(\mu+h) - (XY)(\mu)}{h}\right)v \\
&\quad \leq v_1^T H_P\left(\frac{(XY)(\mu+h) - (XY)(\mu)}{h}\right)v_1 \\
&\quad = (\lambda_{\min}(H_P((XY)(\mu+h))) - v_1^T H_P((XY)(\mu))v_1)/h \\
&\quad \leq (\lambda_{\min}(XY)(\mu+h) - \lambda_{\min}(XY)(\mu))/h.
\end{aligned}$$

Taking limit infimum as h tends to 0^+ in above, we have

$$\min_{\|v\|=1} v^T H_P((XY)'(\mu))v = \lim_{h\to 0^+}\left(\min_{\|v\|=1} v^T H_P\left(\frac{(XY)(\mu+h)-(XY)(\mu)}{h}\right)v\right)$$
$$\le \liminf_{h\to 0^+} \frac{\lambda_{\min}(XY)(\mu+h)-\lambda_{\min}(XY)(\mu)}{h}.$$

But

$$\min_{\|v\|=1} v^T H_P((XY)'(\mu))v = \frac{1}{\mu}\lambda_{\min}(XY)(\mu).$$

This implies that $(1/\mu)\lambda_{\min}(XY)(\mu) \le \liminf_{h\to 0^+}(\lambda_{\min}(XY)(\mu+h) - \lambda_{\min}(XY)(\mu))/h$.

Hence $\lambda'_{\min}(XY)(\mu)$ exists for all $\mu \in J_0$ and

$$\lambda'_{\min}(XY)(\mu) = \frac{\lambda_{\min}(XY)(\mu)}{\mu}.$$

Therefore, integrating with respect to μ and using $(X(1), Y(1)) = (X^0, Y^0)$, we obtain $\lambda_{\min}(XY)(\mu) = \lambda_{\min}(X^0Y^0)\mu$.

Similarly, we can show that $\lambda_{\max}(XY)(\mu) = \lambda_{\max}(X^0Y^0)\mu$. □

Theorem 2.3 *Under Assumptions 2.1 and 2.2, the followings hold true:*

(a) *The system of ODEs (6)–(7) has an unique solution* $(X(\mu), Y(\mu)) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ *which is analytic on* $(0, +\infty)$.

(b) *For any* $0 < \bar{\mu} < +\infty$, *the solution* $(X(\mu), Y(\mu))$ *of the ODE system* (6)–(7) *is bounded on* $(0, \bar{\mu}]$.

Proof We first show that the solution is bounded on any bounded interval $(\mu^-, \bar{\mu}]$ of the maximal interval of existence $J_0 = (\mu^-, \mu^+)$. Secondly, we show that $\mu^- = 0$ and $\mu^+ = +\infty$. This shows (b), and together with Theorem 2.1 implies the property (a).

Let (X^1, Y^1) be a feasible point in $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ [which exists by Assumption 2.1(b)]. Then we have $\mathcal{A}(X(\mu)-X^1)+\mathcal{B}(Y(\mu)-Y^1) = 0$. By Assumption 2.1(a), SDLCP is monotone, thus $(X(\mu) - X^1) \bullet (Y(\mu) - Y^1) \ge 0$. From Lemma 2.1, it follows that

$$Y^1 \bullet X(\mu) + X^1 \bullet Y(\mu) \le X^1 \bullet Y^1 + n\mu.$$

Since $X^1, Y^1, X(\mu), Y(\mu) \in S^n_{++}$ and the products, $Y^1 \bullet X(\mu)$ and $X^1 \bullet Y(\mu)$, are bounded from above on $(\mu^-, \bar{\mu}]$, we have $X(\mu)$, $Y(\mu)$ are bounded on $(\mu^-, \bar{\mu}]$ for any $\bar{\mu} \in (\mu^-, \mu^+)$.

Now we show $\mu^- = 0$, $\mu^+ = +\infty$.

Suppose $\mu^- > 0$. Then since $X(\mu), Y(\mu)$ are bounded near $\mu = \mu^-$, we must have by (15), $\lim_{\mu\to\mu^-} \text{dist}((X(\mu), Y(\mu)), \partial(\mathcal{S}^n_{++} \times \mathcal{S}^n_{++})) = 0$. But by Theorem 2.2, we have $\lambda_{\min}(XY)(\mu) = \lambda_{\min}(X^0Y^0)\mu$. Therefore, with $X(\mu), Y(\mu) \in \mathcal{S}^n_{++}$ and bounded for all $\mu > \mu^-$ and close to μ^-, all accumulation points of $X(\mu), Y(\mu)$ as $\mu \to \mu^-$ are positive definite and bounded. Hence, we have $\lim_{\mu\to\mu^-} \text{dist}((X(\mu), Y(\mu)), \partial(\mathcal{S}^n_{++} \times \mathcal{S}^n_{++})) > 0$, which is a contradiction. Therefore, we have $\mu^- = 0$.

Similarly, $\mu^+ = +\infty$. □

We state in the theorem below, using Theorem 2.2, the relationship between any accumulation point of $(X(\mu), Y(\mu))$ as μ tends to zero and the original SDLCP.

Theorem 2.4 *Let (X^*, Y^*) be an accumulation point of the solution, $(X(\mu), Y(\mu))$, to the system of ODEs (6)–(7) as $\mu \to 0$. Then (X^*, Y^*) is a solution to the SDLCP (3).*

Proof Let (X^*, Y^*) be an accumulation point of $(X(\mu), Y(\mu))$ as μ tends to zero.

Then, by Theorem 2.2, $\lambda_{\min}(X^*Y^*) = \lambda_{\max}(X^*Y^*) = 0$ and $X^*, Y^* \in \mathcal{S}^n_+$ implies that $X^*Y^* = 0$. The latter together with $\mathcal{A}(X^*) + \mathcal{B}(Y^*) = q$, $X^*, Y^* \in \mathcal{S}^n_+$ implies that (X^*, Y^*) is a solution to the SDLCP (3). □

Corollary 2.1 *If the given SDLCP (3) has an unique solution, then every of its off-central paths will converge to the unique solution as μ approaches zero.*

Proof Since $X(\mu), Y(\mu)$ are bounded near $\mu = 0$ by Theorem 2.3(b), and all accumulation points of $(X(\mu), Y(\mu))$ is a solution to the SDLCP (3) by Theorem 2.4, we must have $(X(\mu), Y(\mu))$ converges as $\mu \to 0$ and its limit point is the unique solution to the SDLCP (3). □

Remark 2.3 When the SDLCP (3) has multiple solutions, then whether an off-central path converges is still an open question.

3 Limiting behavior of off-central paths

In this section, we show that an off-central path need not be analytic as a function of $\sqrt{\mu}$ at the limit point $\mu = 0$ and has unbounded first derivative as a function of μ, even if it is close to the central path. We observe this fact through an example. The example we choose has all the nice properties (e.g., primal and dual nondegeneracy) used in the literature and is representative of the common SDP (which is a special class of the monotone SDLCP) encountered in practice. This observation tells a bad news that interior point method with certain symmetrized directions, such as the HKM, for SDP and SDLCP cannot have fast local convergence in general. On a positive side, we will show, through the same example, that certain off-central paths, characterized by a condition, are analytic at the limit point.

We analyze the following primal-dual SDP pair:

$$(\mathcal{P})\ \min \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \bullet X$$
$$\text{subject to} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \bullet X = 2, \quad \begin{pmatrix} 0 & -1 \\ -1 & 2 \end{pmatrix} \bullet X = 0,\ X \in S^2_+$$

and

$$(\mathcal{D})\ \max \quad 2v_1$$
$$\text{subject to } v_1 \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 & -1 \\ -1 & 2 \end{pmatrix} + Y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad Y \in S^2_+.$$

This example is taken from [9]. Note that the example satisfies the standard assumptions for SDP that appear in the literature.

It has an unique solution, $\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$, which satisfies strict complementarity and nondegeneracy. (The concept of nondegeneracy is discussed for example in [9] and is widely used in the literature.) In this sense, the example is a nice, typical SDLCP example.

We choose this example from [9] mainly because it is simple and its nice properties. What we discussed below using this example, however, is not directly related to its discussion in [9].

Written as a SDLCP, the example can be expressed as

$$\begin{gathered} XY = 0 \\ \mathcal{A}\mathrm{svec}(X) + \mathcal{B}\mathrm{svec}(Y) = q \\ X, Y \in \mathcal{S}^2_+, \end{gathered}$$

where $\mathcal{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\sqrt{2} & 2 \end{pmatrix}$, $\mathcal{B} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sqrt{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$ and $q = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$. Note that $\mathcal{A}$ and $\mathcal{B}$ is the corresponding matrix representation of the linear operator A and B in (3).

We are going to analyze the asymptotic behavior of the off-central path $(X(\mu), Y(\mu))$ defined by the system of ODEs (12) (which is equivalent to (6)–(7)) for the example considered here. We specialized to the case when $P = Y^{1/2}$, that is, the dual HKM direction. In this case, (12) can be written as

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X \otimes_s Y^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(X') \\ \mathrm{svec}(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} 0 \\ \mathrm{svec}(X) \end{pmatrix} \tag{16}$$

with the initial conditions: $(X, Y)(1) = (X^0, Y^0)$, where (X^0, Y^0) satisfies

$$\mathcal{A}\mathrm{svec}(X^0) + \mathcal{B}\mathrm{svec}(Y^0) = q \tag{17}$$

$$\mathrm{Tr}(X^0 Y^0) = 2 \tag{18}$$

$$X^0, Y^0 \in \mathcal{S}^2_{++}. \tag{19}$$

Note that we obtain (18) from Assumption 2.2(a).
Equations (16) and (17) imply that $(X(\mu), Y(\mu))$ satisfies

$$\mathcal{A}\mathrm{svec}(X) + \mathcal{B}\mathrm{svec}(Y) = q.$$

From this equality, we see that

$$X(\mu) = \begin{pmatrix} 1 & x(\mu) \\ x(\mu) & x(\mu) \end{pmatrix} \quad \text{and} \quad Y(\mu) = \begin{pmatrix} y_1(\mu) & y_2(\mu) \\ y_2(\mu) & 1 - 2y_2(\mu) \end{pmatrix}$$

for some $x(\mu), y_1(\mu), y_2(\mu) \in \Re$.
By Assumption 2.1(a),

$$A\left(X(\mu) - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) + B\left(Y(\mu) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) = 0$$

implies that $x(\mu) + y_1(\mu) \leq \mathrm{Tr}(XY)(\mu)$. But $\mathrm{Tr}(XY)(\mu) = 2\mu$, by Lemma 2.1. Hence, with $x(\mu)$ and $y_1(\mu)$ positive for $\mu > 0$, we have $x(\mu) = O(\mu)$ and $y_1(\mu) = O(\mu)$. Also, $\det(Y(\mu)) > 0, 1 - 2y_2(\mu)$ bounded above by 1 near $\mu = 0$ and $y_1(\mu) = O(\mu)$ imply that $y_2(\mu) = O(\sqrt{\mu})$.

To analyze the asymptotic behavior of $(X(\mu), Y(\mu))$ as a function of $\sqrt{\mu}$ at $\mu = 0$, let us introduce a new variable $t = \sqrt{\mu}$ and write $\widetilde{X}(t) = X(t^2)$, $\widetilde{Y}(t) = Y(t^2)$, $\widetilde{x}(t) = (1/t^2)x(t^2)$, $\widetilde{y}_1(t) = (1/t^2)y_1(t^2)$ and $\widetilde{y}_2(t) = (1/t)y_2(t^2)$. (By asking whether $X(\mu), Y(\mu)$ are analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$, it is the same as asking whether $\widetilde{X}(t), \widetilde{Y}(t)$ are analytic at $t = 0$.) Then

$$\widetilde{X}(t) = \begin{pmatrix} 1 & t^2\widetilde{x}(t) \\ t^2\widetilde{x}(t) & t^2\widetilde{x}(t) \end{pmatrix} \quad \text{and} \quad \widetilde{Y}(t) = \begin{pmatrix} t^2\widetilde{y}_1(t) & t\widetilde{y}_2(t) \\ t\widetilde{y}_2(t) & 1 - 2t\widetilde{y}_2(t) \end{pmatrix}$$

and $\widetilde{x}(t)$, $\widetilde{y}_1(t)$, and $\widetilde{y}_2(t)$ are bounded near $t = 0$.

Expressing the ODE system (16) in terms of $\widetilde{X}(t)$ and $\widetilde{Y}(t)$, we have

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & \widetilde{X} \otimes_s \widetilde{Y}^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(\widetilde{X}') \\ \mathrm{svec}(\widetilde{Y}') \end{pmatrix} = \frac{2}{t} \begin{pmatrix} 0 \\ \mathrm{svec}(\widetilde{X}) \end{pmatrix} \tag{20}$$

with initial conditions: $(\widetilde{X}, \widetilde{Y})(1) = (X^0, Y^0)$ where (X^0, Y^0) satisfies (17)–(19).

First, we would like to simplify the above ODE system (20).

Proposition 3.1 *$(\widetilde{X}(t), \widetilde{Y}(t))$ satisfies the system of ODEs (20) and the initial conditions (17)–(19) if and only if*

$$(\widetilde{X}(t), \widetilde{Y}(t)) = \left(\begin{pmatrix} 1 & t^2(2-\widetilde{y}_1(t)) \\ t^2(2-\widetilde{y}_1(t)) & t^2(2-\widetilde{y}_1(t)) \end{pmatrix}, \begin{pmatrix} t^2\widetilde{y}_1(t) & t\widetilde{y}_2(t) \\ t\widetilde{y}_2(t) & 1-2t\widetilde{y}_2(t) \end{pmatrix} \right)$$

and $(\widetilde{y}_1(t), \widetilde{y}_2(t))$ satisfies the following equations:

$$\begin{pmatrix} 1-2t\widetilde{y}_2 & -\widetilde{y}_2+t(2-\widetilde{y}_1) \\ -\widetilde{y}_2+t(2-\widetilde{y}_1) & 2 \end{pmatrix} \begin{pmatrix} \widetilde{y}_1' \\ \widetilde{y}_2' \end{pmatrix} = \frac{1}{t} \begin{pmatrix} -\widetilde{y}_2(\widetilde{y}_2+t(2-\widetilde{y}_1)) \\ 2((\widetilde{y}_1-2)(\widetilde{y}_2+t\widetilde{y}_1)+\widetilde{y}_2) \end{pmatrix} \tag{21}$$

with the initial condition on $(\widetilde{y}_1(1), \widetilde{y}_2(1))$ such that

$$\begin{pmatrix} 1 & 2-\widetilde{y}_1(1) \\ 2-\widetilde{y}_1(1) & 2-\widetilde{y}_1(1) \end{pmatrix}, \begin{pmatrix} \widetilde{y}_1(1) & \widetilde{y}_2(1) \\ \widetilde{y}_2(1) & 1-2\widetilde{y}_2(1) \end{pmatrix} \in \mathcal{S}^2_{++}.$$

Proof Since $\mathrm{Tr}(XY)(\mu) = 2\mu$, that is, $\mathrm{Tr}(\widetilde{X}\widetilde{Y})(t) = 2t^2$, we have $x(\mu) = 2\mu - y_1(\mu)$ and $\widetilde{x}(t) = 2 - \widetilde{y}_1(t)$. Therefore,

$$X(\mu) = \begin{pmatrix} 1 & 2\mu-y_1(\mu) \\ 2\mu-y_1(\mu) & 2\mu-y_1(\mu) \end{pmatrix}, \quad Y(\mu) = \begin{pmatrix} y_1(\mu) & y_2(\mu) \\ y_2(\mu) & 1-2y_2(\mu) \end{pmatrix} \tag{22}$$

and

$$\widetilde{X}(t) = \begin{pmatrix} 1 & t^2(2-\widetilde{y}_1(t)) \\ t^2(2-\widetilde{y}_1(t)) & t^2(2-\widetilde{y}_1(t)) \end{pmatrix}, \quad \widetilde{Y}(t) = \begin{pmatrix} t^2\widetilde{y}_1(t) & t\widetilde{y}_2(t) \\ t\widetilde{y}_2(t) & 1-2t\widetilde{y}_2(t) \end{pmatrix}. \tag{23}$$

Since $(\widetilde{X}(t), \widetilde{Y}(t))$ has been expressed in terms of $\widetilde{y}_1(t)$ and $\widetilde{y}_2(t)$ in the form of (23), $(\widetilde{X}(t), \widetilde{Y}(t))$ satisfies (20) if and only if $\widetilde{y}_1(t)$ and $\widetilde{y}_2(t)$ satisfy

$$(1-2t\widetilde{y}_2)\widetilde{y}_1' + (-\widetilde{y}_2+t(2-\widetilde{y}_1))\widetilde{y}_2' = -\widetilde{y}_2(\widetilde{y}_2+t(2-\widetilde{y}_1))/t, \tag{24}$$

$$\begin{aligned} &(1-2t\widetilde{y}_2)(t(2-\widetilde{y}_1)-(2t\widetilde{y}_1+\widetilde{y}_2))\widetilde{y}_1' + 2(1-t(2-\widetilde{y}_1)(t\widetilde{y}_1+\widetilde{y}_2))\widetilde{y}_2' \\ &\quad = -(2-\widetilde{y}_1)(1-2t\widetilde{y}_2)(\widetilde{y}_2+2t\widetilde{y}_1)/t + \widetilde{y}_1\widetilde{y}_2(1+2t^2(2-\widetilde{y}_1))/t \end{aligned} \tag{25}$$

and

$$\begin{aligned} &(\widetilde{y}_1(1-3t\widetilde{y}_2)+\widetilde{y}_2(2t-\widetilde{y}_2))\widetilde{y}_1' + (2-\widetilde{y}_1)(t\widetilde{y}_1+\widetilde{y}_2)\widetilde{y}_2' \\ &\quad = -\widetilde{y}_2(2-\widetilde{y}_1)(\widetilde{y}_2+3t\widetilde{y}_1)/t. \end{aligned} \tag{26}$$

Adding Eq. (25) to $2t$ of Eq. (26) and simplifying, we obtain the following equation:

$$(2t-t\widetilde{y}_1-\widetilde{y}_2)\widetilde{y}_1' + 2\widetilde{y}_2' = 2((\widetilde{y}_1-2)(t\widetilde{y}_1+\widetilde{y}_2)+\widetilde{y}_2)/t. \tag{27}$$

From Eqs. (24) and (27), we obtain the desired system (21).

The initial condition on $(y_1(1), y_2(1))$ can be easily seen from (19) and (23). □

Remark 3.1 Since $\widetilde{X}(t), \widetilde{Y}(t) \in S^2_{++}$ for all $t > 0$, from their expressions in (23), we see that $0 < \widetilde{y}_1(t) < 2$ for all $t > 0$. Furthermore, if we define

$$\widetilde{X}_1(t) := \begin{pmatrix} 1 & 0 \\ 0 & 1/t \end{pmatrix} \widetilde{X}(t) \begin{pmatrix} 1 & 0 \\ 0 & 1/t \end{pmatrix} = \begin{pmatrix} 1 & t(2-\widetilde{y}_1(t)) \\ t(2-\widetilde{y}_1(t)) & 2-\widetilde{y}_1(t) \end{pmatrix}$$

and

$$\widetilde{Y}_1(t) := \begin{pmatrix} 1/t & 0 \\ 0 & 1 \end{pmatrix} \widetilde{Y}(t) \begin{pmatrix} 1/t & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \widetilde{y}_1(t) & \widetilde{y}_2(t) \\ \widetilde{y}_2(t) & 1-2t\widetilde{y}_2(t) \end{pmatrix},$$

from $\lambda_{\min}(\widetilde{X}\widetilde{Y})(t) = t^2\lambda_{\min}(X^0Y^0)$, we have

$$\lambda_{\min}(\widetilde{X}_1\widetilde{Y}_1)(t) = \lambda_{\min}(X^0Y^0) = \text{constant} > 0 \tag{28}$$

for all $t > 0$.

Since $\widetilde{X}_1(t), \widetilde{Y}_1(t) \in S^2_{++}$ are bounded for $t \in (0,1]$, it follows from (28) that $\lambda_{\min}(\widetilde{X}_1(t)), \lambda_{\min}(\widetilde{Y}_1(t)) \geq$ *constant* > 0 for all $t \in (0,1]$. Thus, $\widetilde{y}_1(t), 2-\widetilde{y}_1(t)$ are positive and bounded away from zero for $t > 0$ and as $t \to 0$.

We want to write the system of ODEs (21) in IVP form, for analysis. In order to do this, let us look at the determinant of the matrix in (21).

It can be seen easily that

$$\det\begin{pmatrix} 1-2t\widetilde{y}_2(t) & -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) \\ -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) & 2 \end{pmatrix}$$
$$= \det(\widetilde{X}_1(t)) + \det(\widetilde{Y}_1(t)) \neq 0 \ \text{ for all } t > 0,$$

where $\widetilde{X}_1$ and $\widetilde{Y}_1$ are defined in Remark 3.1. Therefore, we can invert the matrix in (21) to obtain the following:

$$\begin{pmatrix} \widetilde{y}_1' \\ \widetilde{y}_2' \end{pmatrix} = \frac{1}{t(\det(\widetilde{X}_1)+\det(\widetilde{Y}_1))} \times$$
$$\begin{pmatrix} 2(\widetilde{y}_1-2)(t\widetilde{y}_1(t\widetilde{y}_1-2t+2\widetilde{y}_2)+\widetilde{y}_2^2) \\ 2t\widetilde{y}_2(-\widetilde{y}_2+2t-t\widetilde{y}_1)+(t\widetilde{y}_1+\widetilde{y}_2)(-\widetilde{y}_2^2+(2-\widetilde{y}_1)(3t\widetilde{y}_2-2))+2\widetilde{y}_2 \end{pmatrix}. \tag{29}$$

Before analyzing the analyticity of off-central paths at the limit point, let us first state without proof the following lemma [8]:

Lemma 3.1 *Let f be a function defined on $[0,\infty)$. Suppose f is analytic on $[0,\infty)$ and $f(0)$ is not a nonnegative integer. Let z be a solution of $z'(\mu) = (z(\mu)/\mu)f(\mu)$ for $\mu > 0$ with $z(0) = 0$. If z is analytic at $\mu = 0$, then $z(\mu)$ is identically equal to zero for $\mu \geq 0$.*

We have the following main theorem for this section:

Theorem 3.1 *Let $\widetilde{X}(t)$ and $\widetilde{Y}(t)$, given by (23), be positive definite for $t > 0$. Then $(\widetilde{X}(t), \widetilde{Y}(t))$ is a solution to (20) for $t > 0$ and is analytic at $t = 0$ if and only if $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ for all $t > 0$, where $\widetilde{y}_1(t)$ satisfies $\widetilde{y}_1' = 2t\widetilde{y}_1(2-\widetilde{y}_1)/(1+2t^2(\widetilde{y}_1-1))$.*

Proof ($\Rightarrow$) Given that $(\widetilde{X}(t), \widetilde{Y}(t))$ is a solution to (20) for $t > 0$ and is analytic at $t = 0$.

We have $\widetilde{y}_1(t), \widetilde{y}_2(t)$ satisfy (29). From the first differential Eq. in (29), for $\lim_{t\to 0^+} \widetilde{y}_1'(t)$ to exists (which must be true since $(\widetilde{X}(t), \widetilde{Y}(t))$ is analytic at $t = 0$), we see that $\widetilde{y}_2(t)$ must approach zero as $t \to 0$. Therefore, since $\widetilde{y}_2(t)$ is analytic at $t = 0$, we have $\widetilde{y}_2(t) = tw(t)$, where $w(t)$ is analytic at $t = 0$. We want to show that $w(t) = -\widetilde{y}_1(t)$.

Now, from the first differential equation in (29), we have

$$\widetilde{y}_1' = \frac{2(\widetilde{y}_1 - 2)(t\widetilde{y}_1(t\widetilde{y}_1 - 2t + 2\widetilde{y}_2) + \widetilde{y}_2^2)}{t(2 - \widetilde{y}_1 - t^2(2 - \widetilde{y}_1)^2 + \widetilde{y}_1(1 - 2t\widetilde{y}_2) - \widetilde{y}_2^2)}.$$

Substituting $\widetilde{y}_2 = tw$ into the above equation and simplifying, we have

$$\widetilde{y}_1' = \frac{2t(\widetilde{y}_1 - 2)(\widetilde{y}_1(\widetilde{y}_1 - 2 + 2w) + w^2)}{2 - t^2((2 - \widetilde{y}_1)^2 + 2w\widetilde{y}_1 + w^2)}. \tag{30}$$

From the second differential equation in (29), we have

$$\widetilde{y}_2' = \frac{2t\widetilde{y}_2(-\widetilde{y}_2 + 2t - t\widetilde{y}_1) + (t\widetilde{y}_1 + \widetilde{y}_2)(-\widetilde{y}_2^2 + (2 - \widetilde{y}_1)(3t\widetilde{y}_2 - 2)) + 2\widetilde{y}_2}{t(2 - \widetilde{y}_1 - t^2(2 - \widetilde{y}_1)^2 + \widetilde{y}_1(1 - 2t\widetilde{y}_2) - \widetilde{y}_2^2)}.$$

Substituting tw for $\widetilde{y}_2$ and $tw' + w$ for $\widetilde{y}_2'$ into the above equation, we have, after bringing w to the right hand side of the resulting equation, dividing throughout by t and simplifying,

$$w' = \frac{2(2 - \widetilde{y}_1)((w + \widetilde{y}_1)(t^2w - 1) + 2t^2w)}{t(2 - t^2((2 - \widetilde{y}_1)^2 + 2w\widetilde{y}_1 + w^2))}. \tag{31}$$

Adding up Eqs. (30) and (31) and upon simplifications, we obtain

$$(\widetilde{y}_1 + w)'(t) = \frac{2(2 - \widetilde{y}_1(t))(t^2(2 - \widetilde{y}_1(t)) - 1)}{t(2 - t^2((2 - \widetilde{y}_1(t))^2 + 2w\widetilde{y}_1 + w^2))}(\widetilde{y}_1(t) + w(t)).$$

Let $z(t) = \widetilde{y}_1(t) + w(t)$. Then $z(t)$ is analytic at $t = 0$, since $\widetilde{y}_1(t)$ and $w(t)$ are analytic at $t = 0$. We have the following differential equation:

$$z'(t) = \frac{z(t)}{t}\left(\frac{2(2 - \widetilde{y}_1(t))(t^2(2 - \widetilde{y}_1(t)) - 1)}{2 - t^2((2 - \widetilde{y}_1(t))^2 + z^2 - \widetilde{y}_1^2)}\right). \tag{32}$$

Let $f(t) = 2(2-\widetilde{y}_1(t))(t^2(2-\widetilde{y}_1(t)) - 1)/(2 - t^2((2-\widetilde{y}_1(t))^2 + z^2 - \widetilde{y}_1^2))$. Then $f(t)$ is analytic for all $t \geq 0$. Also, $f(0) = -(2-\widetilde{y}_1(0))$, which is strictly less than zero by Remark 3.1.

From (32), we see that in order for $z'(t)$ to exist as t approaches zero, which should be the case since $z(t)$ is analytic at $t = 0$, we must have $z(0) = 0$, since $f(0)$ is nonzero. Now $z(t), f(t)$ here satisfy the conditions in Lemma 3.1. Therefore, by the lemma, $z(t)$ is identically equal to zero which implies that $w(t) = -\widetilde{y}_1(t)$.

Using $w(t) = -\widetilde{y}_1(t)$, expressing the differential equation (30) in terms of $\widetilde{y}_1$, we obtain the ODE of $\widetilde{y}_1$ in the theorem.

($\Leftarrow$) Suppose $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ for all $t > 0$, where $\widetilde{y}_1(t)$ satisfies

$$\widetilde{y}_1' = \frac{2t\widetilde{y}_1(2-\widetilde{y}_1)}{1+2t^2(\widetilde{y}_1 - 1)} \text{ for } t > 0. \tag{33}$$

Then, since the right-hand side of the above ODE of $\widetilde{y}_1$ is analytic at $t = 0$ and $\widetilde{y}_1 \in \Re$, we have, by Theorem 2.1, that $\widetilde{y}_1(t)$ can be analytically extended to $t = 0$. Hence $\widetilde{y}_2(t)$ can also be analytically extended to $t = 0$. These imply that $\widetilde{X}(t), \widetilde{Y}(t)$ are analytic at $t = 0$.

With $\widetilde{y}_2(t)$ related to $\widetilde{y}_1(t)$ by $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$, where $\widetilde{y}_1(t)$ satisfying the ODE (33), we can also check easily that $\widetilde{y}_1(t)$ and $\widetilde{y}_2(t)$ satisfy (21). Hence, by Proposition 3.1, $(\widetilde{X}(t), \widetilde{Y}(t))$ satisfies (20) for $t > 0$. □

Letting $(X(\mu), Y(\mu))$ be a solution of the ODE system (16) and using the relations $y_1(\mu) = \mu\widetilde{y}_1(\sqrt{\mu}), y_2(\mu) = \sqrt{\mu}\widetilde{y}_2(\sqrt{\mu})$, we have the following neat equivalences:

Theorem 3.2 *Let $(X(\mu), Y(\mu)) \in S^n_{++} \times S^n_{++}$, given by (22), be the solution to (16) for $\mu > 0$ with initial conditions given by (17)–(19), i.e., $(X(\mu), Y(\mu))$ is an off-central path. Then the followings are equivalent:*

(a) $y_2(\mu_0) = -y_1(\mu_0)$ *for some* $\mu_0 > 0$.
(b) $y_2(\mu) = -y_1(\mu)$ *for all* $\mu > 0$.
(c) $(X(\mu), Y(\mu))$ *can be extended analytically to* $\mu = 0$.
(d) $(\tilde{X}(t), \tilde{Y}(t))$ *can be extended analytically to* $t = \sqrt{\mu} = 0$.

Proof Note that conditions (a) and (b) can be written in terms of t, e.g., (b) is equivalent to

$$\tilde{y}_2(t) = -t\tilde{y}_1(t). \tag{34}$$

(a) $\Rightarrow$ (d). Suppose $\widetilde{y}_2(t_0) = -t_0\widetilde{y}_1(t_0)$ for some $t_0 > 0$.
Define

$$\widetilde{U}(t) := \begin{pmatrix} 1 & t^2(2-\widetilde{v}_1(t)) \\ t^2(2-\widetilde{v}_1(t)) & t^2(2-\widetilde{v}_1(t)) \end{pmatrix}, \quad \widetilde{V}(t) := \begin{pmatrix} t^2\widetilde{v}_1(t) & t\widetilde{v}_2(t) \\ t\widetilde{v}_2(t) & 1-2t\widetilde{v}_2(t) \end{pmatrix} \quad \text{for } t > 0,$$

such that $\widetilde{v}_1(t_0) = \widetilde{y}_1(t_0), \widetilde{v}_2(t_0) = \widetilde{y}_2(t_0), \widetilde{v}_2(t) = -t\widetilde{v}_1(t)$ for all $t > 0$ and $\widetilde{v}_1(t)$ satisfies $\widetilde{v}_1' = 2t\widetilde{v}_1(2-\widetilde{v}_1)/(1+2t^2(\widetilde{v}_1 - 1))$. We have $\widetilde{U}(t), \widetilde{V}(t) \in S^n_{++}$ in a neighborhood of $t = t_0$. Also $(\widetilde{U}(t), \widetilde{V}(t))$ satisfies (20) in this neighborhood. By uniqueness of off-central path, we then have $\widetilde{U}(t) = \widetilde{X}(t), \widetilde{V}(t) = \widetilde{Y}(t)$ in this neighborhood of $t = t_0$. Hence, by an analytic continuity argument, $(\widetilde{X}(t), \widetilde{Y}(t)) = (\widetilde{U}(t), \widetilde{V}(t))$ for all $t > 0$ and $(\widetilde{X}(t), \widetilde{Y}(t))$ satisfies the necessary conditions in Theorem 3.1 since $(\widetilde{U}(t), \widetilde{V}(t))$ is defined using these conditions. Therefore, by the theorem, $(\widetilde{X}(t), \widetilde{Y}(t))$ can be extended analytically to $t = 0$.

(d) $\Rightarrow$ (b). The result follows from (34) and Theorem 3.1.

(b) $\Rightarrow$ (a). Obvious.

(c) $\Rightarrow$ (d). This is clear.

(d) $\Rightarrow$ (c). Suppose $(X(\mu), Y(\mu))$ is analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$.

Then $(\widetilde{X}(t), \widetilde{Y}(t))$ is analytic at $t = 0$. Hence, by Thereom 3.1, we have $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ for all $t > 0$, where $\widetilde{y}_1(t)$ satisfies $\widetilde{y}_1' = 2t\widetilde{y}_1(2 - \widetilde{y}_1)/(1 + 2t^2(\widetilde{y}_1 - 1))$.

It is clear that $y_1(\mu) = \mu\widetilde{y}_1(\sqrt{\mu})$ and $y_2(\mu) = \sqrt{\mu}\widetilde{y}_2(\sqrt{\mu})$. Therefore $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ implies that $y_2(\mu) = -y_1(\mu)$. The special structure of the ODE satisfied by $\widetilde{y}_1$, i.e., $\widetilde{y}_1' = 2t\widetilde{y}_1(2 - \widetilde{y}_1)/(1 + 2t^2(\widetilde{y}_1 - 1))$, the relations $y_1 = \mu\widetilde{y}_1$ and $y_2 = -y_1$, turn out to be the keys in showing that $(X(\mu), Y(\mu))$ can be extended analytically to $\mu = 0$ as follows. Letting $\widetilde{\widetilde{y}}_1(\mu)$ to be $\widetilde{y}_1(\sqrt{\mu})$, we see that $y_1(\mu) = \mu\widetilde{\widetilde{y}}_1(\mu)$, where $\widetilde{\widetilde{y}}_1(\mu)$ satisfies $\widetilde{\widetilde{y}}_1' = \widetilde{\widetilde{y}}_1(2 - \widetilde{\widetilde{y}}_1)/(1 + 2\mu(\widetilde{\widetilde{y}}_1 - 1))$ since $\widetilde{\widetilde{y}}_1' = \widetilde{y}_1'/(2\sqrt{\mu})$ and $\widetilde{y}_1' = 2\sqrt{\mu}\widetilde{y}_1(2 - \widetilde{y}_1)/(1 + 2\mu(\widetilde{y}_1 - 1))$. Since the right-hand side of the ODE satisfied by $\widetilde{\widetilde{y}}_1(\mu)$ is analytic at $\mu = 0$ and $\widetilde{\widetilde{y}}_1 \in \Re$, we have, by Theorem 2.1, $\widetilde{\widetilde{y}}_1(\mu)$ can be extended analytically to $\mu = 0$. Therefore, $y_1(\mu)$ and $y_2(\mu)$ are analytically extensible to $\mu = 0$, which further implies that $(X(\mu), Y(\mu))$ is analytic w.r.t. μ at $\mu = 0$. □

We see from Theorem 3.2, that, unlike [11], where off-central paths are defined differently and are analytically extensible as a function of $\sqrt{\mu}$ to $\mu = 0$, no matter how close we consider a starting point (for the off-central path) to the central path of the SDP example, we can always start off with a point whose off-central path is not analytic as a function of μ or $\sqrt{\mu}$ at $\mu = 0$. On the other hand, if the initial point satisfies a certain condition, namely, $y_2 = -y_1$, its off-central path can be analytically extended to $\mu = 0$.

Remark 3.2 Similar theorems as Theorems 3.1 and 3.2 can be stated if we consider $P = X^{-1/2}$, which corresponds to the so-called HKM direction.

We know from Theorem 3.2 that given an off-central path, $(X(\mu), Y(\mu))$, of form (22), if $y_2(\mu_0) \neq -y_1(\mu_0)$ for some $\mu_0 > 0$, then $(X(\mu), Y(\mu))$ cannot be analytically extended to $\mu = 0$. One may ask what further asymptotic behavior does $(X(\mu), Y(\mu))$ have when $y_2(\mu_0) \neq -y_1(\mu_0)$ for some $\mu_0 > 0$, besides being not analytically extensible to $\mu = 0$?

In what follows, we show that the first derivative of such $(X(\mu), Y(\mu))$ is unbounded as $\mu \to 0$. From this, it clearly shows that although the example given satisfies all kinds of regularity conditions, its off-central paths defined using the dual HKM direction have bad behavior near the optimal solution of the example.

Theorem 3.3 *Let* $(X(\mu), Y(\mu))$*, given by* (22)*, be the solution to* (16) *for* $\mu > 0$ *with initial conditions given by* (17)–(19)*, i.e.,* $(X(\mu), Y(\mu))$ *is an off-central path. If* $y_2(\mu_0) \neq -y_1(\mu_0)$ *for some* $\mu_0 > 0$*, then the first derivative of* $(X(\mu), Y(\mu))$ *is unbounded as* $\mu \to 0$.

Proof Now, if $(X(\mu), Y(\mu))$ is an off-central path of the example, then

$$X(\mu) = \begin{pmatrix} 1 & 2\mu - y_1(\mu) \\ 2\mu - y_1(\mu) & 2\mu - y_1(\mu) \end{pmatrix}, \quad Y(\mu) = \begin{pmatrix} y_1(\mu) & y_2(\mu) \\ y_2(\mu) & 1 - 2y_2(\mu) \end{pmatrix} \quad \text{[see (22)]}.$$

We have $y_1(\mu) = \mu\widetilde{y}_1(\sqrt{\mu}), y_2(\mu) = \sqrt{\mu}\widetilde{y}_2(\sqrt{\mu})$, where $\widetilde{y}_1, \widetilde{y}_2$ satisfy the system of ODEs (29) with $t = \sqrt{\mu}$. Expressing (29) in terms of y_1, y_2 and their first derivatives, and μ, we obtain

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \frac{y_1(1-2y_2) - y_2^2}{\mu(2\mu(1-2y_2) - (2\mu - y_1 - y_2)^2)} \begin{pmatrix} 2\mu \\ -(2\mu - y_1 - y_2) \end{pmatrix}. \tag{35}$$

Let

$$y_1(\mu) := \mu\widetilde{\widetilde{y}}_1(\mu) \text{ for } \mu > 0. \tag{36}$$

We have, by Remark 3.1, that $\widetilde{\widetilde{y}}_1(\mu)$ and $2 - \widetilde{\widetilde{y}}_1(\mu)$ are positive and bounded away from zero as $\mu \to 0$.

From $y_2' = -(2\mu - y_1 - y_2)(y_1(1-2y_2) - y_2^2)/(\mu(2\mu(1-2y_2) - (2\mu - y_1 - y_2)^2))$ in (35), using $y_1(\mu) = \mu\widetilde{y}_1(\sqrt{\mu}), y_2(\mu) = \sqrt{\mu}\widetilde{y}_2(\sqrt{\mu})$ and the boundedness of $\widetilde{y}_1, \widetilde{y}_2$ near $\mu = 0$, we see that if $y_2(\mu) \neq O(\mu)$, then y_2' is unbounded as $\mu \to 0$. Hence $(X(\mu), Y(\mu))$ has unbounded first derivative as $\mu \to 0$.

Therefore assuming that $y_2(\mu) = O(\mu)$, we wish to show that if $y_2(\mu_0) \neq -y_1(\mu_0)$ for some $\mu_0 > 0$, then $(X(\mu), Y(\mu))$ has unbounded first derivative as $\mu \to 0$.

We do this by showing that y_2'' behaves like $1/\mu$ for μ close to zero. From this, we see that y_2' is unbounded as $\mu \to 0$.

Since $y_2(\mu)$ is assumed to be $O(\mu)$, let

$$y_2(\mu) := \mu\widetilde{w}(\mu) \text{ for } \mu > 0, \tag{37}$$

where $\widetilde{w}(\mu) = O(1)$.

Upon computing y_2'' using the second equation in (35), we obtain

$$y_2'' = \frac{gh}{\mu(2\mu(1-2y_2) - (2\mu - y_1 - y_2)^2)}, \tag{38}$$

where

$$g(\mu) := \frac{y_1(1-2y_2)-y_2^2}{\mu(2\mu(1-2y_2)-(2\mu-y_1-y_2)^2)},$$
$$h(\mu) := -4\mu(2\mu-y_1-y_2)^2-(2\mu-y_1-y_2)^2(-4\mu^2 g+(y_1+y_2)(1-2\mu g)) +2(2\mu-y_1-y_2)(1-2y_2)\mu-4\mu^2(1-2y_2) +(y_1+y_2)(y_1(1-2y_2)-y_2^2).$$

Let us now estimate g and h.

Substituting $y_1=\mu\widetilde{\widetilde{y}}_1$ and $y_2=\mu\widetilde{w}$ into the expressions for g and h above, we have

$$g = \frac{\mu\widetilde{\widetilde{y}}_1(1-2\mu\widetilde{w})-\mu^2\widetilde{w}^2}{\mu(2\mu(1-2\mu\widetilde{w})-(2\mu-\mu\widetilde{\widetilde{y}}_1-\mu\widetilde{w})^2)} = \Theta(1/\mu),$$

where we have used $\widetilde{\widetilde{y}}_1$ is bounded away from zero as μ approaches zero and $\widetilde{w}(\mu)=O(1)$ to obtain the estimate. And grouping terms containing μ^3 together in h and after simplifications, we obtain

$$h=\mu^2(\widetilde{\widetilde{y}}_1-2)(\widetilde{\widetilde{y}}_1+\widetilde{w})+O(\mu^3).$$

Therefore from (38)

$$y_2'' = \frac{gh}{\mu(2\mu(1-2y_2)-(2\mu-y_1-y_2)^2)} = \frac{\Theta(1/\mu)}{\mu^2(2(1-2\mu\widetilde{w})-\mu(2-\widetilde{\widetilde{y}}_1-\widetilde{w})^2)}(\mu^2(\widetilde{\widetilde{y}}_1-2)(\widetilde{\widetilde{y}}_1+\widetilde{w})+O(\mu^3)) = -\Theta(1/\mu)(\widetilde{\widetilde{y}}_1+\widetilde{w})+O(1), \tag{39}$$

where the last equality follows since $2-\widetilde{\widetilde{y}}_1$ is positive and bounded away from zero for μ close to zero and $\widetilde{w}(\mu)=O(1)$.

We want to investigate $\widetilde{\widetilde{y}}_1+\widetilde{w}$ further for μ close to zero. We use its derivative in this investigation.

Using (35) and the relations $y_1=\mu\widetilde{\widetilde{y}}_1, y_2=\mu\widetilde{w}$, we have

$$(\widetilde{\widetilde{y}}_1+\widetilde{w})' = \frac{(\widetilde{\widetilde{y}}_1+\widetilde{w})(2-\widetilde{\widetilde{y}}_1)((2-\widetilde{\widetilde{y}}_1)\mu-1)}{\mu(2(1-2\mu\widetilde{w})-\mu(2-\widetilde{\widetilde{y}}_1-\widetilde{w})^2)}. \tag{40}$$

Now, we are ready to show using (39) and (40) that if $y_2(\mu)=O(\mu)$, then $y_2(\mu_0)\neq -y_1(\mu_0)$ for some $\mu_0>0$ implies that $(X(\mu),Y(\mu))$ has unbounded first derivative as $\mu\to 0$.

Suppose $y_2(\mu_0)>-y_1(\mu_0)$ for some $\mu_0>0$. We recall that $\widetilde{\widetilde{y}}_1,\widetilde{w}$ are related to y_1 and y_2 by (36) and (37) respectively, hence there exists $\epsilon>0$ such that

$\widetilde{\widetilde{y}}_1(\mu_1) + \widetilde{w}(\mu_1) = \epsilon$ for some $0 < \mu_1 \le 2/(4+\epsilon^2)$. [If not, then we have $y_2(\mu_1) = -y_1(\mu_1)$ which implies that $y_2(\mu) = -y_1(\mu)$ for all $\mu > 0$ by Theorem 3.2, a contradiction to $y_2(\mu_0) > -y_1(\mu_0)$.] Then from (40), we see that $(\widetilde{\widetilde{y}}_1 + \widetilde{w})'(\mu_1) \le 0$. Therefore, $\widetilde{\widetilde{y}}_1(\mu) + \widetilde{w}(\mu) \ge \epsilon$ for $\mu \le \mu_1$ and close to μ_1. In fact, we see that by the continuity of $\widetilde{\widetilde{y}}_1(\mu) + \widetilde{w}(\mu)$ for $0 < \mu \le \mu_1$ and (40) that $\widetilde{\widetilde{y}}_1(\mu) + \widetilde{w}(\mu) \ge \epsilon$ for all μ between 0 and μ_1. Hence, from (39), $y_2'' = -\Theta(1/\mu)$ which implies that y_2' is unbounded as μ approaches zero, i.e., the first derivative of $(X(\mu), Y(\mu))$ is unbounded as $\mu \to 0$.

Similarly, if $y_2(\mu_0) < -y_1(\mu_0)$ for some $\mu_0 > 0$, we also have the first derivative of $(X(\mu), Y(\mu))$ unbounded near $\mu = 0$. □

4 Conclusions

An IPM is a homotopy method in which the search direction at a point is parallel to the tangent to the path at that point. This inspires the idea of defining underlying paths of an IPM (where search directions are known) as solutions of the ODE associated with the field of search directions of the IPM. Since the behavior of an IPM, such as convergence, complexity, etc, depends mainly on its search directions, the characteristics of the underlying paths, which we have proposed in this paper, are helpful in analyzing and anticipating the behavior of IPMs.

We have shown that the paths, i.e., the solutions of the ODE system, uniquely exist and are analytic on $(0, +\infty)$. We have also shown that any accumulation point of a path is a solution of SDLCP. A surprising observation by analyzing an example is that off-central paths under our definition are usually not analytic with respect to $\sqrt{\mu}$ at $\mu = 0$. This observation sharply contrasts with what has been reported by other researchers, namely that, all paths (central paths, off-central paths, associated with AHO, HKM, etc, symmetrization operators) under various definitions [11,12,22] are analytic with respect to $\sqrt{\mu}$ (and in the case of AHO direction, even analytic with respect to μ) at $\mu = 0$, under strict complementarity condition.

It is interesting to see that different definitions of paths lead to such sharp contrasting behaviors of paths and each definition has its own implications. Since our paths are defined by fields of search directions which decide the convergence behavior of an IPM, the bad behavior (not being analytic and having unbounded derivative if $y_2 \ne -y_1$ in the example) of the underlying paths strongly indicates slow convergence of IPM near a solution to a SDLCP. Kojima et al. [9] observed certain evidence that IPMs do not converge superlinearly without "shrinking" the neighborhood using the HKM direction. Through the analysis of underlying paths here, we provide further evidence which anticipates the slow local convergence of IPMs for SDP and SDLCP using the HKM direction.

On the other hand, our ability to find a set of nice off-central paths casts a ray of hope; starting from certain points (e.g., satisfying $y_2 = -y_1$), the IPM in our example problem can follow paths which are analytic at $\mu = 0$, and hence can converge superlinearly without "neighborhood shrinking" [23]. Whether this is true for general SDPs and SDLCPs is still an open problem.

In this paper we have defined off-central paths and established their basic properties. However, what we have seen about these paths is very limited. The real behavior and rich structure of these paths and their implications to the study of IPMs have yet to be unveiled.

References

1. Adler, I., Monteiro, R.D.C.: Limiting behavior of the affine scaling continuous trajectories for linear programming problems. Math. Program. **50**(1), Series A, 29–51 (1991)
2. Amann, H.: Ordinary Differential Equations : An Introduction to Nonlinear Analysis, (translated from German by Gerhard Metzen) de Gruyter Studies in Mathematics vol 13 (1990)
3. Bayer, D.A., Lagarias, J.C.: The nonlinear geometry of linear programming, I, II, III. Trans. Am. Math. Soc. **314**, 499–526, 527–581 (1989) and **320**, 193–225 (1990)
4. Birkhoff, G., Rota, G.-C.: Ordinary Differential Equations, 4th edn (1989)
5. Chua, C.B.: A new notion of weighted centers for semidefinite programming. SIAM J. Optimi. **16**(4), 1092–1109 (2006)
6. Güler, O.: Limiting behavior of weighted central paths in linear programming. Math. Program. **65**(3),Series A, 347–363 (1994)
7. Halická, M.: Analyticity of the central path at the boundary point in semidefinite programming. European J. Opera. Res. **143**, 311–324 (2002)
8. Ince, E.L.: Ordinary Differential Equations. Dover Publications (1956)
9. Kojima, M., Shida, M., Shindoh, S.: Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Math. Program. **80**(72), Series A, 129–160 (1998)
10. Kojima, M., Shindoh, S., Hara, S.: Interior-point methods for the monotone semidefinite linear complementarity problems. SIAM J. Optimi. **7**, 86–125 (1997)
11. 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. Optimi. **15**(2), 348–374 (2004)
12. Lu, Z., Monteiro, R.D.C.: Limiting behavior of the Alizadeh-Haeberly-Overton weighted paths in semidefinite programming, Preprint, July 24, 2003
13. Megiddo, N., Mizuno, S., Tsuchiya, T.: A modified layered-step interior-point algorithm for linear programming. Math. Program. **82**, 339–355 (1998)
14. Mehrotra, S.: Quadratic convergence in a primal-dual method. Math. Opera. Res. **18**, 741–751 (1993)
15. Monteiro, R.D.C.: Primal-dual path following algorithms for semidefinite programming. SIAM J. Optimi. **7**, 663–678 (1997)
16. Monteiro, R.D.C., Pang, J.-S.: Properties of an interior-point mapping for mixed complementarity problems. Math. Opera. Res. **21**(3), 629–654 (1996)
17. Monteiro, R.D.C., Pang, J.-S.: On two interior-point mappings for nonlinear semidefinite complementarity problems. Math. Opera. Res. **23**(1), 39–60 (1998)
18. Monteiro, R.D.C., Tsuchiya, T.: Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem. Math. Oper. Rese. **21**(4), 793–814 (1996)
19. Monteiro, R.D.C., Tsuchiya, T.: A variant of the Vavasis-Ye layered-step interior-point algorithm for linear programming. SIAM J. Optimi. **23**(4), 1054–1079 (2003)
20. Monteiro, R.D.C., Zanjácomo, P.R.: General interior-point maps and existence of weighted paths for nonlinear semidefinite complementarity problems. Math. Oper. Res. **25**(3), 381–399 (2000)
21. Potra, F.A., Sheng, R.: Superlinear convergence of interior-point algorithms for semidefinite programming. J. Optim. Theory Appl. **99**(1), 103–119 (1998)
22. Preiß M., Stoer, J.: Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems. Math. Program. **99**(3), Series A, 499–520 (2004)
23. Sim, C.-K.: Underlying Paths and Local Convergence Behaviour of Path-following Interior Point Algorithm for SDLCP and SOCP. Ph.D. Thesis, National University of Singapore, (2004)

24. 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 Heidelberg New York (1985)
25. Sonnevend, G., Stoer, J., Zhao, G.: On the complexity of following the central path of linear programs by linear extrapolation. Methods Opera. Res. **62**, 19–31 (1989)
26. 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)
27. Stoer, J., Wechs, M.: On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems. Nume. Mathe. **81**(4), 631–645 (1999)
28. Stoer, J., Wechs, M., Mizuno, S.: High order infeasible-interior-point methods for solving sufficient linear complementarity problems. Mathe. Opera. Res. **23**(4), 832–862 (1998)
29. Sturm, J.F.: Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists. Math. Oper. Res. **24**(1), 72–94 (1999)
30. Todd, M.J., Toh, K.C., Tütüncü, R.H.: On the Nesterov-Todd direction in semidefinite programming. SIAM J. Optimi. **8**(3), 769–796 (1998)
31. Tütüncü, R.H.: Asymptotic behavior of continuous trajectories for primal-dual potential-reduction methods. SIAM J. Optimi. **14**(2), 402–414 (2003)
32. Vavasis, S.A., Ye, Y.: A primal-dual interior point method whose running time depends only on the contraint matrix. Math. Program. **74**(1), 79–120 (1996)
33. 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)
34. 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)
35. Zhang, Y.: On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optimi. **8**(2), 365–386 (1998)
36. Zhao, G.: On the relationship between the curvature integral and the complexity of path-following methods in linear programming. SIAM J. Optimi. **6**(1), 57–73 (1996)
37. 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**(1), 85–103 (1993)
38. Zhao, G., Sun, J.: On the rate of local convergence of high-order-infeasible-path-following algorithms for P_*-linear complementarity problems. Computa. Optim. Appl. **14**, 293–307 (1999)
39. Zhao, G., Zhu, J.: The curvature integral and the complexity of linear complementarity problems. Math. Program. **70**, 107–122 (1996)