UNDERLYING PATHS AND LOCAL CONVERGENCE BEHAVIOUR OF PATH-FOLLOWING INTERIOR POINT ALGORITHM FOR SDLCP AND SOCP

SIM CHEE KHIAN

M.Sc(U.Wash.,Seattle),DipSA(NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

NATIONAL UNIVERSITY OF SINGAPORE

2004

Acknowledgements

I am deeply indebted to my thesis advisor, Associate Professor Zhao Gong Yun, for the time spent in discussion with me while doing this project. I would also like to express indebtedness to my parents and other family members for the care that they have shown me during my PhD programme. This research was conducted while I was supported by an NUS Research Scholarship in my first year of PhD study and a A*STAR graduate fellowship after the first year.

Contents

Summary

In this dissertation, we define a new way to view off-central path for semidefinite linear complementarity problem (SDLCP) and second order cone programming (SOCP). They are defined using a system of ordinary differential equations (ODEs). Asymptotic behaviour of these off-central paths is directly related to the local convergence behaviour of path-following interior point algorithm [26, 22].

In Chapter 2, we consider off-central path for SDLCP. We show the existence of off-central path (starting from any interior point) for general direction for all $\mu > 0$. Also, as is expected, any accumulation point of an off-central path is a solution to the SDLCP. We then restrict our attention to the dual HKM direction and show using a "nice" example that not all off-central paths are analytic w.r.t $\sqrt{\mu}$ at the limit when $\mu = 0$. We derive a simple necessary and sufficient condition to when an off-central path is analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$. It also turns out that for this example, an off-central path is analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$ if and only if it is analytic w.r.t μ at $\mu = 0$. Using the example on the predictor-corrector algorithm, we show that if an iterate lies on an off-central path which is analytic at $\mu = 0$, then after the predictor and corrector step, the next iterate will also lie on an off-central path which is analytic at $\mu = 0$. This implies that if we have a suitably chosen initial iterate, then using the feasible predictor-corrector algorithm, the iterates will converge superlinearly to the solution of SDLCP. Next, we work on the general SDLCP. Assuming strict complementarity and carefully transforming the system of ODEs defining the off-central path to an equivalent set of ODEs, we are able to extract more asymptotic properties of the off-central path. More importantly, we give a necessary and sufficient condition to when an off-central path in general is analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$.

In Chapter 3, we consider off-central path for multiple cone SOCP. We first define off-central path for SOCP for general direction and then restrict our attention to the AHO direction. We show using an example that off-central path defined using the AHO direction may not exist if we start from some interior point. Based on this example, we then give a region, which is possibly the largest, in which off-central path, starting from any point in this region, is well-defined for all $\mu > 0$. By further restricting the region to a smaller one and assuming strict complementarity, we are able to show that any off-central path in this smaller region converges to a strictly complementary optimal solution. We prove this by giving a characterization of the relative interior of the optimal solution set and then relate it to the set of strict complementary optimal solutions. The usefulness of strict complementarity on asymptotic analyticity of off-central path is shown for 1-cone SOCP.

Chapter 1

Introduction

In path-following interior point algorithms, the central path plays an important role. These algorithms (for example, the predictor-corrector algorithm) are such that the iterates try to "follow" the central path closely. Ideally, we would want the iterates to stay on the central path (which leads to the optimal solution of the optimization problem under consideration). However, this is usually not practical. Hence there arises a need to study "nearby" paths on which the iterates lie, besides the central path, that also lead to the optimal solution. In this respect, there are a number of papers in the literature, [17, 21, 9, 10, 24, 13, 5, 11, 12, 15] and the references therein, that discuss these so-called off-central paths.

In [15], the authors considered the existence and uniqueness of off-central paths for nonlinear semidefinite complementarity problems, which include the semidefinite linear complementarity problem and semidefinite programming as special cases. The nonlinear semidefinite complementarity problem that they considered is to find a triple $(X, Y, z) \in S^n \times S^n \times \Re^m$ such that

$$F(X, Y, z) = 0, \quad XY = 0, \quad X, Y \in S^n_+,$$

where $F : S^n_+ \times S^n_+ \times \Re^m \longrightarrow S^n \times \Re^m$ is a continuous map. Here S^n stands for the space of $n \times n$ symmetric matrices while S^n_+ stands for the space of $n \times n$

symmetric positive semidefinite matrices.

By representing the complementarity condition, $XY = 0, X, Y \in S^n_+$, in several equivalent forms, the authors defined interior-point maps using which off-central paths are defined. An example of an interior-point map considered in [15] is $H : S^n_+ \times S^n_+ \times \Re^m \longrightarrow S^n \times \Re^m \times S^n$ defined by

$$H(X, Y, z) = \begin{pmatrix} F(X, Y, z) \\ X^{1/2} Y X^{1/2} \end{pmatrix}.$$

Clearly, (X, Y, z) is a solution of the nonlinear semidefinite complementarity problem if and only if it satisfies $H(X, Y, z) = 0$. Under appropriate assumptions on F (which we will not elaborate here), it was shown that, given M in a certain set in S^n_{++}, $H(X, Y, z) = (0, \mu M)$ has a unique solution for every $\mu \in (0, 1]$. These solutions, as μ varies, define an off-central path, which is based on the given interior-point map H, for the nonlinear semidefinite complementarity problem.

In [17], the authors also considered the question of existence and uniqueness of off-central paths, but for a more specified algebraic system:

$$\begin{array}{rcl} A(X) + B(Y) & = & q + \mu \bar{q} \\ \frac{1}{2}(XY + YX) & = & \mu M \\ X, Y & \in & S^n_{++} \end{array} \tag{1.1}$$

where $M \in S^n_{++}$ is fixed. Here A, B are linear operators from S^n to $\Re^{\widetilde{n}}$, where $\widetilde{n} = \frac{n(n+1)}{2}$.

Their result about existence and uniqueness of the off-central path $(X, Y)(.)$ as a function of $\mu > 0$ is not new. It was proven in [12, 15] by means of deep results from nonlinear analysis. However, the proof in [17] is more elementary, essentially relying only on the Implicit Function Theorem.

The study of off-central paths is especially important in the limit as the paths approach the optimal solution. For example, the analyticity of these paths at the limit point, when $\mu = 0$, has an effect on the rate of convergence of path-following algorithms (See [26]). For linear programming and linear complementarity problem, the asymptotic behaviour of off-central paths is discussed in [21, 24, 13, 5]. As for second order cone programming (SOCP), as far as we know, there have not been any discussion on the local behaviour of off-central path at the limit point in the literature.

Here we will discuss, in more detail, the literature on the limiting behaviour of off-central paths for semidefinite programming (SDP) and semidefinite linear complementarity problem (SDLCP).

A semidefinite linear complementarity problem is to find a pair $(X, Y) \in S^n_+ \times S^n_+$ such that

$$XY = 0$$
$$A(X) + B(Y) = q,$$

where A, B are linear operators from S^n to $\Re^{\widetilde{n}}$, $\widetilde{n} = \frac{n(n+1)}{2}$.

As noted earlier, the complementarity condition, $XY = 0, X, Y \in S^n_+$, can be represented in several equivalent forms. The reason we need to work on these equivalent forms instead of the original complementarity condition, $XY = 0, X, Y \in S^n_+$, itself is because we have to ensure that the search directions in interior-point algorithms are symmetric (see, for example, [25]). The common equivalent forms used are $(XY + YX)/2 = 0$, $X^{1/2}YX^{1/2} = 0$, $Y^{1/2}XY^{1/2} = 0$ and $W^{1/2}XYW^{1/2} = 0$ where W is such that $WXW = Y$. The first equivalent form results in the AHO direction, while the second and third equivalent forms result in the HKM direction and its dual and the last equivalent form results in the NT direction.

In [17], the authors considered off-central paths for SDLCP corresponding to the AHO direction. To them, an off-central path is the solution to the following set of algebraic equations

$$
\begin{aligned}
A(X) + B(Y) &= q + \mu\bar{q} \\
\frac{1}{2}(XY + YX) &= \mu M \\
X, Y &\in S^n_{++}
\end{aligned}
$$

where $M \in S^n_{++}$ is fixed and $\mu > 0$.

Assuming strict complementarity solution of the SDLCP, the authors were able to show, in [17], that the off-central path is analytic at $\mu = 0$, with respect to μ, for any $M \in S^n_{++}$. In the same spirit, the authors in [10] shows the same result, but for the case of SDP and also assuming strict complementarity.

The authors of [10] also show in another paper, [9], the asymptotic behaviour of off-central paths for SDP corresponding to another direction (the HKM direction), different from the AHO direction. They considered an off-central path which is the solution to the following system of algebraic equations

$$
\begin{aligned}
A(X) &= b + \mu\Delta b \\
A^* y + Y &= C + \mu\Delta C \\
X^{1/2} Y X^{1/2} &= \mu M \\
X, Y &\in S^n_{++}
\end{aligned}
$$

where $M \in S^n_{++}$, $\Delta b \in \Re^m$ and $\Delta C \in S^n$ are fixed.

Assuming strict complementarity, the authors in [9] show that an off-central path, as a function of $\sqrt{\mu}$, can be extended analytically beyond 0 and as a corollary, they show that the path converges as μ tends to zero.

There are also some work done in the literature that study the analyticity at the limit point of off-central paths, without assuming strict complementarity, for certain class of SDP. See, for example, [16]. However, it is generally believed that it is difficult to analyse the analyticity of off-central paths at the limit point for general SDLCP or SDP without assuming strict complementarity.

In our current work, we have a different viewpoint to define off-central path for SDLCP/SDP and SOCP. We use the concept of direction field. We will only consider the 2-dimensional case to describe this concept, since higher dimensions are similar. Let us consider the 2-dimensional plane. At each point on the plane or an open subset of the plane, we can associated with it a 2-dimensional vector. The set of such 2-dimensional vectors then constitutes a direction field on the plane or open subset (One can similarly imagine a direction field defined in $\Re^n$ for general $n \geq 3$). To be meaningful, however, the direction field must be such that we can "draw" smooth curves on the plane or in the open subset with each element of a direction field along the tangent line to a curve. An area of mathematics where direction field arises naturally is in the area of differential equations. The solution curves to a system of ordinary differential equations made up the smooth curves that we are considering. The first derivatives of these curves are then elements of a direction field.

The concept of direction field can be applied to the predictor-corrector algorithm for SDLCP and SOCP. It induces a system of ordinary differential equations (ODEs) whose solution is the off-central path for SDLCP and SOCP (Notice the difference between our definition of off-central path as compared to that in the literature described earlier where off-central path is the solution to an algebraic system of equations. There are also works done in the literature concerning linear programming where off-central path is defined as a solution of ODE system, see for example, [24] and the references therein). We believe that our definition of off-central path is more natural since it is directly derived from algorithmic aspect of SDLCP and SOCP, that is, from the search directions in path-following interior point algorithm.

In our current work, we are going to study the off-central paths defined in the "ODE" way for SDLCP and SOCP. This study is directly related to the asymptotic behaviour of path-following interior point algorithm.

1.1 Notations

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 $X \bullet Y \equiv Tr(X^T Y)$, where $Tr(\cdot)$ 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) 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 maximum norm. $\|\cdot\|_F$ for a matrix in $\Re^{p\times q}$ refers to the Frobenius norm.

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

Given a square matrix X with real eigenvalues, $\lambda_i(X)$ refers to the i^{th} eigenvalue of X arranged in decreasing order, $\lambda_{max}(X)$ refers to the maximum eigenvalue of X while $\lambda_{min}(X)$ refers to the minimum eigenvalue of X.

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

For a function, $f(\cdot)$, of one variable analytic at a point μ_0, we denote its kth derivative at μ_0 by $f^{(k)}(\mu_0)$.

Also, $\begin{pmatrix} n \\ k \end{pmatrix}$ stands for $\frac{n!}{k!(n-k)!}$.

Given a differentiable function Φ from an open set $\mathcal{O}$ in $\Re^{n_1} \times \ldots \times \Re^{n_k} \times \Re^m$ to $\Re$. Suppose $(z_1, \ldots, z_k, w) \in \mathcal{O}$ where $z_1 \in \Re^{n_1}$, …, $z_k \in \Re^{n_k}$ and $w \in \Re^m$. Then $D_{(z_1,\ldots,z_k)}\Phi$ is the derivative row vector of Φ w.r.t the component $(z_1, \ldots, z_k)$ of $(z_1, \ldots, z_k, w)$. If the codomain of Φ is $\Re^n$ for $n \geq 2$, then $D_{(z_1,\ldots,z_k)}\Phi$ is defined in a similar manner.

Relative interior of a convex set C, denoted by riC, is defined as the interior which results when C is regarded as a subset of its affine hull.

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 $f(w) = \mathcal{O}(g(w))$ for all $w \in \widetilde{\Omega}$ to mean that $\|f(w)\| \leq Mg(w)$ for all $w \in \widetilde{\Omega}$ and $M > 0$ is a constant; Moreover, for a function $U : \Omega \longrightarrow S^n_{++}$, we write $U(w) = \Theta(g(w))$ for all $w \in \widetilde{\Omega}$ if $U(w) = \mathcal{O}(g(w))$ and $U(w)^{-1} = \mathcal{O}(g(w))$ for all $w \in \widetilde{\Omega}$. The latter condition is equivalent to the existence of a constant $M > 0$ such that

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

If $\{u(\nu) : \nu > 0\}$ and $\{v(\nu) : \nu > 0\}$ are real sequences with $v(\nu) > 0$, then $u(\nu) = o(v(\nu))$ means that $\lim_{\nu \to 0} \frac{u(\nu)}{v(\nu)} = 0$. If $u(\nu)$ is a matrix or vector, then $u(\nu) = o(v(\nu))$ means that $\lim_{\nu \to 0} \frac{\|u(\nu)\|}{v(\nu)} = 0$.

Chapter 2

Analysis of Off-Central Paths for SDLCP

Using our definition of off-central path, $(X(\mu), Y(\mu))$, we show that this path is well-behaved in the sense that it is well defined and analytic for all $\mu > 0$ and any of its acummulation point as $\mu \to 0$ is a solution to the SDLCP. This is done in Section 2.1. In Section 2.2, we show, using a simple example, that the off-central paths are not analytic at $\mu = 0$ in general. In fact, we show a stronger result that the off-central paths are not analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$ in general. This finding surprised us, because all off-central paths studied in the literature up to date —[21] for LCP, [17, 10] for SDLCP/SDP with AHO direction, and [9] in which off-central paths associated with HKM direction are defined by a system of algebraic equations— are analytic w.r.t μ or $\sqrt{\mu}$ at $\mu = 0$. This observation also unveils a substantial difference between AHO direction and other directions. On the other hand, for the same example, there exists a subset of off-central paths which are analytic at $\mu = 0$. These "nice" paths are characterized by some algebraic equations. Then, in Section 2.2.1, we show that by applying the predictor-corrector path-following algorithm to this example and starting from a point on any such a "nice" path, superlinear convergence can be achieved. Finally, in Section 2.3, we give a necessary and sufficient condition for an off-central path of a general SDLCP, satisfying the strict complementarity condition, to be analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$.

2.1 Off-Central Path for SDLCP

In this section, we define a direction field associated to the predictor-corrector algorithm for semidefinite linear complementarity problem (SDLCP). This gives rise to a system of ordinary differential equations (ODEs) whose solution is the off-central path for SDLCP.

Let us consider the following SDLCP:

$$\begin{aligned} XY &= 0 \\ A(X) + B(Y) &= q \\ X, Y &\in \mathcal{S}^n_+ \end{aligned} \tag{2.1}$$

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 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* $A(X^1) + B(Y^1) = q$.

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

In the predictor step of the predictor-corrector path-following algorithm, the algorithm searches a new point in the affine direction, which is defined as the Newton direction for the system $XY = 0$. Let (X, Y) be the current point and

$$(X^+, Y^+) = (X, Y) + (\Delta X, \Delta Y). \tag{2.2}$$

From the equation $X^+Y^+ = 0$, we obtain

$$XY + X\Delta Y + \Delta XY + \Delta X\Delta Y = 0.$$

The linear part is the Newton equation, i.e.,

$$X\Delta Y + \Delta XY = -XY.$$

For SDLCP, we make certain symmetrization [25]

$$H_P(X\Delta Y + \Delta XY) = -H_P(XY). \tag{2.3}$$

where $H_P(U) := \frac{1}{2}(PUP^{-1} + (PUP^{-1})^T)$ and $P \in \Re^{n\times n}$ is an invertible matrix. (2.3) defines the *affine direction* $(\Delta X, \Delta Y)$ at (X, Y).

It then follows that the *direction field* comprises, at each point $(X, Y) \succ 0$, the direction $(\Delta X, \Delta Y)$ defined by (2.3).

A path in $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ passing through (X^0, Y^0) and having its tangent vectors elements of this direction field is then determined by

$$\begin{aligned} H_P(XY' + X'Y) &= -H_P(XY), & (2.4)\\ A(X') + B(Y') &= 0 & (2.5) \end{aligned}$$

with the initial condition $(X, Y)(0) = (X^0, Y^0)$ where $X^0, Y^0 \succ 0$. Here equation (2.5) arises out of the feasibility equation in (2.1).

Without loss of generality, we can make a parameter transformation $\mu = \exp(-t)$, where t is the parameter in (2.4)-(2.5). Then we have (we still use the notation (X, Y) for the path with the new parameter μ)

$$\begin{aligned} H_P(XY' + X'Y) &= \frac{1}{\mu}H_P(XY), & (2.6)\\ A(X') + B(Y') &= 0 & (2.7) \end{aligned}$$

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

We will now show that, given the initial condition $(X, Y)(1) = (X^0, Y^0)$, the solution to (2.6)-(2.7), $(X(\mu), Y(\mu))$, $X(\mu), Y(\mu) \succ 0$, is unique, analytic and exists over $\mu \in (0, \infty)$. We called this solution the **off-central path** for SDLCP.

Remark 2.1 *The central path $(X_c(\mu), Y_c(\mu))$ for SDLCP, which satisfies $(X_cY_c)(\mu) = \mu I$, is a special example of off-central path for SDLCP. When $\mu = 1$, it satisfies $Tr((X_cY_c)(1)) = n$. Therefore, we also require the initial data (X^0, Y^0) when $\mu = 1$ in (2.6)-(2.7) to satisfy $Tr(X^0Y^0) = n$.*

As in [23], 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.

For the AHO direction, $P = I$. Hence (2.6) reduces to

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

This and (2.7) with the initial condition at $\mu = 1$ yield the algebraic equations (1.1). For other directions, such as HKM and NT directions, P is a function of (X, Y), thus it is not possible to solve (2.6)-(2.7) 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 [17] and this chapter.

We are going to use a result from ODE theory, taken from [2] pp. 100 and [3] pp.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 a unique nonextensible solution*

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

of the IVP

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

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

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

(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 (2.6)-(2.7) in the form of IVP as in the theorem.

Now, (2.6) can be written as

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

Remark 2.2 *Note that the operation* $\otimes_s$ *and the map svec are used extensively in this chapter. For their definitions and properties, the reader can refer to pp. 775-776 and the appendix of [23].*

Writing (2.7) in a similar way using *svec*, we can rewrite (2.6)-(2.7) as

$$\begin{pmatrix} svec(A_1)^T & svec(B_1)^T \\ \vdots & \vdots \\ svec(A_{\tilde{n}})^T & svec(B_{\tilde{n}})^T \\ P \otimes_s P^{-T} Y & PX \otimes_s P^{-T} \end{pmatrix} \begin{pmatrix} svec(X') \\ svec(Y') \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\mu} svec(H_P(XY)) \end{pmatrix}, \quad (2.8)$$

which is another form of (2.6)-(2.7).

In the following proposition, we show that the matrix in (2.8) is invertible for all $X, Y \succ 0$ and hence, we can express (2.6)-(2.7) in the IVP form of Theorem 2.1 and the theorem is then applicable for our case.

Proposition 2.1

$$\begin{pmatrix} svec(A_1)^T & svec(B_1)^T \\ \vdots & \vdots \\ svec(A_{\tilde{n}})^T & svec(B_{\tilde{n}})^T \\ P \otimes_s P^{-T}Y & PX \otimes_s P^{-T} \end{pmatrix}$$

is nonsingular for all $X, Y \succ 0$*.*

Proof. Since the given matrix is square, it suffices to show that it is one-to-one.

Therefore, given the below matrix-vector equation,

$$\begin{pmatrix} svec(A_1)^T & svec(B_1)^T \\ \vdots & \vdots \\ svec(A_{\tilde{n}})^T & svec(B_{\tilde{n}})^T \\ P \otimes_s P^{-T}Y & PX \otimes_s P^{-T} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

we need to show that $u = v = 0$.

We have $(P \otimes_s P^{-T}Y)u + (PX \otimes_s P^{-T})v = 0$ implies that $(PX \otimes_s P^{-T})v = -(P \otimes_s P^{-T}Y)u$. But $PX \otimes_s P^{-T} = (PXP^T \otimes_s I)(P \otimes_s P)^{-T}$ and $P \otimes_s P^{-T}Y = (I \otimes_s P^{-T}YP^{-1})(P \otimes_s P)$. Therefore

$$\begin{aligned} & (PX \otimes_s P^{-T})v = -(P \otimes_s P^{-T}Y)u \\ \Longrightarrow \quad & v = -(P \otimes_s P)^T(PXP^T \otimes_s I)^{-1}(I \otimes_s P^{-T}YP^{-1})(P \otimes_s P)u \end{aligned}$$

Note that $(PXP^T \otimes_s I)^{-1}$ and $I \otimes_s P^{-T}YP^{-1}$ are symmetric, positive definite and they commute (since $PXYP^{-1}$ is symmetric). Therefore, $(P \otimes_s P)^T(PXP^T \otimes_s I)^{-1}(I \otimes_s P^{-T}YP^{-1})(P \otimes_s P)$ is symmetric, positive definite.

Now, $svec(A_i)^Tu + svec(B_i)^Tv = 0$ for $i = 1, \ldots, \tilde{n}$ implies that $u^Tv \geq 0$, by Assumption 2.1(a). That is,

$$u^T(P \otimes_s P)^T(PXP^T \otimes_s I)^{-1}(I \otimes_s P^{-T}YP^{-1})(P \otimes_s P)u \leq 0.$$

But with $(P \otimes_s P)^T(PXP^T \otimes_s I)^{-1}(I \otimes_s P^{-T}YP^{-1})(P \otimes_s P)$ symmetric, positive definite, we must have $u = 0$. And hence, $v = 0$. **QED**

Let the matrix in Proposition 2.1 be denoted by $\mathcal{A}(X, Y)$. We have shown that $\mathcal{A}(X, Y)$ is invertible for all $X, Y \succ 0$. Therefore, we can write (2.8) in the IVP form as

$$\begin{pmatrix} svec(X') \\ svec(Y') \end{pmatrix} = \mathcal{F}(\mu, X, Y) \tag{2.9}$$

where

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

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

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

The maximal interval of existence J_0 is open:

$$J_0 = (\mu^-, \mu^+), \tag{2.10}$$

where either we have

$$\mu^- = 0, \quad \text{resp. } \mu^+ = +\infty \qquad \text{or}$$

$$\lim_{\mu \to \mu^+ (\mu \to \mu^-)} \min\{\mathrm{dist}((\mathrm{X}(\mu), \mathrm{Y}(\mu)), \partial(\mathcal{S}^{\mathrm{n}}_{++} \times \mathcal{S}^{\mathrm{n}}_{++})), \|(\mathrm{svec}(\mathrm{X}(\mu)), \mathrm{svec}(\mathrm{Y}(\mu)))\|^{-1}\} = 0.$$

Also, since $\mathcal{F}(\mu, X, Y)$ is analytic over $\Re_{++} \times (\mathcal{S}^n_{++} \times \mathcal{S}^n_{++})$, by the same theorem, we have $(X(\mu), Y(\mu))$ is analytic over J_0.

We want to determine the value of μ^- and μ^+ in (2.10). We do this by stating and proving the following theorem:

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

Proof. Note that since $\lambda_{min}(XY)(\mu)$ is locally lipschitz continuous on J_0, by Theorem 7.20 in [19], $\lambda'_{min}(XY)(\mu)$ exists almost everywhere. We first show that whenever it exists, $\lambda'_{min}(XY)(\mu) = \lambda_{min}(XY)(\mu)/\mu$ for $\mu \in J_0$. Hence, $\lambda_{min}(XY)(\mu)$ is monotonic on J_0.

Recall that P in (2.6) 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

$$\begin{aligned} H_{P(\mu)}((XY)(\mu))v_0 &= \lambda_{min}(H_{P(\mu)}((XY)(\mu)))v_0 \\ &= \lambda_{min}(XY)(\mu)v_0 \end{aligned}$$

The last equality holds because $(PXYP^{-1})(\mu)$ is symmetric.
Therefore, by (2.6) 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

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

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

$$\begin{aligned} & -\liminf_{h\to 0^+}\frac{f(h)-f(0)}{h} \\ = & -\liminf_{h\to 0^+} f'(\xi_h) \end{aligned}$$

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

Note that $P(\mu+\xi)P^{-1}(\mu+\xi) = I$ implies that $P'(\mu+\xi)P^{-1}(\mu+\xi) + P(\mu+\xi)(P^{-1})'(\mu+\xi) = 0$.

Hence $(P^{-1})'(\mu+\xi) = -P^{-1}(\mu+\xi)P'(\mu+\xi)P^{-1}(\mu+\xi)$.

Therefore, $f'(\xi_h) = v_0^T P'(\mu+\xi_h)(XY)(\mu+h)P^{-1}(\mu+\xi_h)v_0 - v_0^T P(\mu+\xi_h)(XY)(\mu+h)P^{-1}(\mu+\xi_h)P'(\mu+\xi_h)P^{-1}(\mu+\xi_h)v_0$.

Hence,

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

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

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

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

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

But

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

This implies that $\frac{1}{\mu}\lambda_{min}(XY)(\mu) \leq \liminf_{h\to 0^+} \frac{\lambda_{min}(XY)(\mu+h)-\lambda_{min}(XY)(\mu)}{h}$.
Hence $\lambda'_{min}(XY)(\mu)$ whenever it exists has

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

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 $\lambda_{max}(XY)(\mu) = \lambda_{max}(X^0Y^0)\mu$. **QED**

Remark 2.3 *We can also see easily that* $Tr(XY)(\mu) = Tr(X^0Y^0)\mu = n\mu$ *for all* $\mu \in J_0$*, using (2.6). Here the last equality follows from Remark 2.1.*

Also, we have the following remark which is used in the proofs of Corollaries 2.1 and 2.2.

Remark 2.4 *On an off-central path,* $X(\mu)$*,* $Y(\mu)$ *are bounded near* $\mu = 0$*. This can be easily seen using Thereom 2.2 and from* $(X(\mu) - X^1) \bullet (Y(\mu) - Y^1) \geq 0$*, which follows from Assumption 2.1(a) and (b).*

As an immediate consequence of the above theorem, we have

Corollary 2.1 $\mu^- = 0$*,* $\mu^+ = +\infty$ *in (2.10). Therefore, the solution* $(X(\mu), Y(\mu))$ *to (2.6)-(2.7) in* $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ *is unique and analytic for* $\mu \in (0, +\infty)$*.*

Proof. By Theorem 2.2, it is clear that for all $\mu > 0$, $X(\mu), Y(\mu) \in S^n_{++}$. Hence $\mu^- = 0$ and $\mu^+ = +\infty$. **QED**

We also 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 2.3 *Let* (X^*, Y^*) *be an accumulation point of the solution,* $(X(\mu), Y(\mu))$*, to the system of ODEs (2.6)-(2.7) as* $\mu \to 0$*. Then* (X^*, Y^*) *is a solution to the SDLCP (2.1).*

Proof. Let (X^*, Y^*) be an accumulation point of $(X(\mu), Y(\mu))$ as μ tends to zero. Then, by Theorem 2.2, $\lambda_{min}(X^*Y^*) = \lambda_{max}(X^*Y^*) = 0$, which implies that $X^*Y^* = 0$. Together with $A(X^*) + B(Y^*) = q,\ X^*, Y^* \in S^n_+$, we have (X^*, Y^*) is a solution to the SDLCP (2.1). **QED**

Corollary 2.2 *If the given SDLCP (2.1) has a unique solution, then every of its off-central paths will converge to the unique solution as* μ *approaches zero.*

Remark 2.5 *When the SDLCP (2.1) has multiple solutions, then whether an off-central path converges is still an open question.*

2.2 Investigation of Asymptotic Analyticity of Off-Central Path for SDLCP using a "Nice" Example

In this section, we show that an off-central path need not be analytic w.r.t $\sqrt{\mu}$ at the limit point, even if it is close to the central path. We observe this fact through an example. The example we choose has all nice properties (e.g. primal and dual nondegeneracy) and thus is representative of the common SDP (which is a special class of monotone SDLCP) encountered in practice. This observation tells a bad news which is that interior point method with certain symmetrized directions 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. Moreover, this condition can be sustained by the predictor-corrector interior point method, i.e., starting from a point satisfying this condition, after the predictor and corrector step, the new point will also satisfy this condition. This means that if we can choose a starting point satisfying this condition, then the predictor-corrector algorithm will converge superlinearly/quadratically.

Consider the following primal-dual SDP pair:

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

and

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

This example is taken from [8]. 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 non-degeneracy (The concept of non-degeneracy is discussed for example in [8] and is widely used in the literature). In this sense, the example is a nice, typical SDLCP example.

We choose this example from [8] 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 [8].

Written as a SDLCP, the example can be expressed as

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

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 (2.1).

We are going to analyse the asymptotic behaviour of the off-central path $(X(\mu), Y(\mu))$ defined by the system of ODEs (2.8). We specialized to the case when $P = Y^{1/2}$, that is, the dual HKM direction. In this case, (2.8) can be written as

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X \otimes_s Y^{-1} \end{pmatrix} \begin{pmatrix} svec(X') \\ svec(Y') \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\mu} svec(X) \end{pmatrix}. \tag{2.11}$$

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

$$\mathcal{A}svec(X^0) + \mathcal{B}svec(Y^0) = q \tag{2.12}$$

$$Tr(X^0Y^0) = 2 \tag{2.13}$$

$$X^0, Y^0 \in \mathcal{S}^2_{++}, \tag{2.14}$$

Note that we obtained (2.13) from Remark 2.1.

We are going to analyse the asymptotic behaviour of $(X(\mu), Y(\mu))$ w.r.t $\sqrt{\mu}$ at $\mu = 0$. To make presentation easier, let us introduce the matrices $\widetilde{X}(t)$ and $\widetilde{Y}(t)$ to be $X(t^2)$ and $Y(t^2)$ respectively.

Equations (2.11) and (2.12) imply that $(X(\mu), Y(\mu))$ satisfies

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

From this equality, we see that

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

for some $x(\mu), y_1(\mu), y_2(\mu) \in \Re$.
Now

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

implies that $x(\mu)+y_1(\mu) \leq Tr(XY)(\mu)$, by Assumption 2.1(a). But $Tr(XY)(\mu) = 2\mu$, by Remark 2.3 and (2.13). Hence, with $x(\mu)$ and $y_1(\mu)$ positive for $\mu > 0$, we have $x(\mu) = \mathcal{O}(\mu)$ and $y_1(\mu) = \mathcal{O}(\mu)$. Also, determinant of $Y(\mu)$ positive for all $\mu > 0$, $1 - 2y_2(\mu)$ bounded above by 1 and $y_1(\mu) = \mathcal{O}(\mu)$ implies that $y_2(\mu) = \mathcal{O}(\sqrt{\mu})$. Therefore, we can write $\widetilde{X}(t)$ and $\widetilde{Y}(t)$ as

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

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

Expressing the ODE system (2.11) in terms of $\widetilde{X}(t)$ and $\widetilde{Y}(t)$, we have

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & \widetilde{X} \otimes_s \widetilde{Y}^{-1} \end{pmatrix} \begin{pmatrix} svec(\widetilde{X}') \\ svec(\widetilde{Y}') \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{t} svec(\widetilde{X}) \end{pmatrix}. \tag{2.15}$$

with initial conditions: $(\widetilde{X}, \widetilde{Y})(1) = (X^0, Y^0)$ where (X^0, Y^0) satisfies (2.12)-(2.14).

Note that to investigate the asymptotic analyticity of $(X(\mu), Y(\mu))$ for the example w.r.t $\sqrt{\mu}$ at $\mu = 0$, we need only study the asymptotic property of $(\widetilde{X}(t), \widetilde{Y}(t))$.

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

Proposition 2.2 *$(\widetilde{X}(t), \widetilde{Y}(t))$ satisfies the system of ODEs (2.15) and the initial conditions (2.12)-(2.14) 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:

$$\begin{pmatrix} 1-2t\widetilde{y}_2 & -\widetilde{y}_2 + t(2-\widetilde{y}_1) \\ -\widetilde{y}_2 + t(2-\widetilde{y}_1) & 2 \end{pmatrix} \begin{pmatrix} \widetilde{y}_1' \\ \widetilde{y}_2' \end{pmatrix} = \frac{1}{t} \begin{pmatrix} -\widetilde{y}_2(\widetilde{y}_2 + t(2-\widetilde{y}_1)) \\ 2((\widetilde{y}_1 - 2)(\widetilde{y}_2 + t\widetilde{y}_1) + \widetilde{y}_2) \end{pmatrix} \tag{2.16}$$

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. For the second equation in the system (2.15), we write explicitly

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

$$(\widetilde{X} \otimes_s \widetilde{Y}^{-1})(t) = \frac{1}{\det(\widetilde{Y})} \times$$
$$\begin{pmatrix} 1-2t\widetilde{y}_2 & \frac{t}{\sqrt{2}}(t\widetilde{x}(1-2t\widetilde{y}_2) - \widetilde{y}_2) & -t^3\widetilde{x}\widetilde{y}_2 \\ \frac{t}{\sqrt{2}}(-\widetilde{y}_2 + t\widetilde{x}(1-2t\widetilde{y}_2)) & \frac{t^2}{2}(\widetilde{x}(1-4t\widetilde{y}_2) + \widetilde{y}_1) & \frac{t^3}{\sqrt{2}}\widetilde{x}(t\widetilde{y}_1 - \widetilde{y}_2) \\ -t^3\widetilde{x}\widetilde{y}_2 & \frac{t^3}{\sqrt{2}}\widetilde{x}(t\widetilde{y}_1 - \widetilde{y}_2) & t^4\widetilde{x}\widetilde{y}_1 \end{pmatrix},$$

Note that the dependence of $\widetilde{y}_1$ and $\widetilde{y}_2$ on t is omitted from the last expression for easy readability.

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

$$X(\mu) = \begin{pmatrix} 1 & 2\mu - y_1(\mu) \\ 2\mu - y_1(\mu) & 2\mu - y_1(\mu) \end{pmatrix}, \quad Y(\mu) = \begin{pmatrix} y_1(\mu) & y_2(\mu) \\ y_2(\mu) & 1 - 2y_2(\mu) \end{pmatrix} \tag{2.17}$$

and

$$\widetilde{X}(t) = \begin{pmatrix} 1 & t^2(2 - \widetilde{y}_1(t)) \\ t^2(2 - \widetilde{y}_1(t)) & t^2(2 - \widetilde{y}_1(t)) \end{pmatrix}, \quad \widetilde{Y}(t) = \begin{pmatrix} t^2\widetilde{y}_1(t) & t\widetilde{y}_2(t) \\ t\widetilde{y}_2(t) & 1 - 2t\widetilde{y}_2(t) \end{pmatrix} \tag{2.18}$$

With the above expression for $(\widetilde{X} \otimes_s \widetilde{Y}^{-1})(t)$ and $\widetilde{X}(t)$, $\widetilde{Y}(t)$ in (2.18), we have expressing (2.15) in terms of $\widetilde{y}_1(t)$ and $\widetilde{y}_2(t)$ that $(\widetilde{X}(t), \widetilde{Y}(t))$, of form (2.18), satisfies (2.15) if and only if $\widetilde{y}_1(t)$ and $\widetilde{y}_2(t)$ satisfy

$$(1 - 2t\widetilde{y}_2)\widetilde{y}_1' + (-\widetilde{y}_2 + t(2 - \widetilde{y}_1))\widetilde{y}_2' = -\widetilde{y}_2(\widetilde{y}_2 + t(2 - \widetilde{y}_1))/t, \tag{2.19}$$

$$\begin{gathered}(1 - 2t\widetilde{y}_2)(t(2 - \widetilde{y}_1) - (2t\widetilde{y}_1 + \widetilde{y}_2))\widetilde{y}_1' + 2(1 - t(2 - \widetilde{y}_1)(t\widetilde{y}_1 + \widetilde{y}_2))\widetilde{y}_2' = \\ -(2 - \widetilde{y}_1)(1 - 2t\widetilde{y}_2)(\widetilde{y}_2 + 2t\widetilde{y}_1)/t + \widetilde{y}_1\widetilde{y}_2(1 + 2t^2(2 - \widetilde{y}_1))/t\end{gathered} \tag{2.20}$$

and

$$\begin{gathered}(\widetilde{y}_1(1 - 3t\widetilde{y}_2) + \widetilde{y}_2(2t - \widetilde{y}_2))\widetilde{y}_1' + (2 - \widetilde{y}_1)(t\widetilde{y}_1 + \widetilde{y}_2)\widetilde{y}_2' = \\ -\widetilde{y}_2(2 - \widetilde{y}_1)(\widetilde{y}_2 + 3t\widetilde{y}_1)/t.\end{gathered} \tag{2.21}$$

Adding equation (2.20) to $2t$ of equation (2.21) and simplifying, we obtain the following equation:

$$(2t - t\widetilde{y}_1 - \widetilde{y}_2)\widetilde{y}_1' + 2\widetilde{y}_2' = 2((\widetilde{y}_1 - 2)(t\widetilde{y}_1 + \widetilde{y}_2) + \widetilde{y}_2)/t \tag{2.22}$$

From equations (2.19) and (2.22), we obtain the desired system (2.16).

The initial condition on $(y_1(1), y_2(1))$ can be easily seen from (2.14) and (2.18). **QED**

We want to write the system of ODEs (2.16) in IVP form, for analysis. In order to do this, let us look at the determinant of the matrix on the extreme left in (2.16).

We have the following technical proposition:

Proposition 2.3

$$det\begin{pmatrix} 1-2t\widetilde{y}_2(t) & -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) \\ -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) & 2 \end{pmatrix}$$

is nonzero for $t>0$. *Here* $\widetilde{y}_1(t),\ \widetilde{y}_2(t)$ *appear in Proposition 2.2 where* $(\widetilde{X}(t),\widetilde{Y}(t))$ *is the solution to (2.15) for* $t>0$.

Proof. Now, $\lambda_{min}(XY)(\mu)=\lambda_{min}(X^0Y^0)\mu$ by Theorem 2.2. Hence $\lambda_{min}(\widetilde{X}\widetilde{Y})(t)=\lambda_{min}(X^0Y^0)t^2$. Therefore,

$$\lambda_{min}(\widetilde{X}_1\widetilde{Y}_1)(t)=\lambda_{min}(X^0Y^0)$$

where $\widetilde{X}_1(t)=\begin{pmatrix} 1 & t(2-\widetilde{y}_1(t)) \\ t(2-\widetilde{y}_1(t)) & 2-\widetilde{y}_1(t) \end{pmatrix}$ and $\widetilde{Y}_1(t)=\begin{pmatrix} \widetilde{y}_1(t) & \widetilde{y}_2(t) \\ \widetilde{y}_2(t) & 1-2t\widetilde{y}_2(t) \end{pmatrix}$.

We have $det(\widetilde{X}_1(t))$ and $det(\widetilde{Y}_1(t))$ are positive for $t>0$. We are going to use this latter fact in the proof of the proposition.

We have

$$det\begin{pmatrix} 1-2t\widetilde{y}_2(t) & -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) \\ -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) & 2 \end{pmatrix}=$$

$$2(1-2t\widetilde{y}_2(t))-(-\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)))^2.$$

Expressing the last expression in terms of $det(\widetilde{X}_1(t))$ and $det(\widetilde{Y}_1(t))$, we have

$$det\begin{pmatrix} 1-2t\widetilde{y}_2(t) & -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) \\ -\widetilde{y}_2(t)+t(2-\widetilde{y}_1(t)) & 2 \end{pmatrix}=det(\widetilde{X}_1(t))+det(\widetilde{Y}_1(t)).$$

Now, we know that $det(\widetilde{X}_1(t))$ and $det(\widetilde{Y}_1(t))$ are positive for all $t > 0$ by above. Hence we are done. **QED**

Therefore, we can invert the matrix in (2.16) to obtain the following:

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

where $\widetilde{X}_1$ and $\widetilde{Y}_1$ are defined in the proof of Proposition 2.3.
Upon simplifying the right-hand side of the ODEs, we have

$$\begin{pmatrix} \widetilde{y}_1' \\ \widetilde{y}_2' \end{pmatrix} = \frac{1}{t(det(\widetilde{X}_1)+det(\widetilde{Y}_1))}\times \tag{2.23}$$
$$\begin{pmatrix} 2(\widetilde{y}_1-2)(t\widetilde{y}_1(t\widetilde{y}_1-2t+2\widetilde{y}_2)+\widetilde{y}_2^2) \\ 2t\widetilde{y}_2(-\widetilde{y}_2+2t-t\widetilde{y}_1)+(t\widetilde{y}_1+\widetilde{y}_2)(-\widetilde{y}_2^2+(2-\widetilde{y}_1)(3t\widetilde{y}_2-2))+2\widetilde{y}_2 \end{pmatrix}.$$

Before analyzing the analyticity of off-central paths at the limit point, let us first state and prove a lemma:

Lemma 2.1 *Let f be a function defined on $[0,\infty)$. Suppose f is analytic at 0 and $f(0)$ is not a positive integer. Let z be a solution of $z'(\mu) = \frac{z(\mu)}{\mu}f(\mu)$ for $\mu > 0$ with $z(0) = 0$. If z is analytic at $\mu = 0$, then $z(\mu)$ is identically equal to zero for $\mu \geq 0$.*

Proof. Consider $z'(\mu) = \frac{z(\mu)}{\mu}f(\mu)$. We will now show that $z^{(n)}(0) = 0$ and $\lim_{\mu\to 0}\left(\frac{z}{\mu}\right)^{(n-1)} = 0$ for all $n \geq 1$ by induction on n.
For $n = 1$. We have by L'Hopital's Rule that $\lim_{\mu\to 0}\frac{z(\mu)}{\mu} = z'(0)$. Therefore, from $z'(\mu) = \frac{z(\mu)}{\mu}f(\mu)$, we obtained $z'(0) = z'(0)f(0)$ by taking limit of μ to zero. But $f(0)$ is not a positive integer implies that $z'(0) = 0$. Hence induction hypothesis is true for $n = 1$.
Now, suppose that $z^{(k)}(0) = 0$ and $\lim_{k\to 0}\left(\frac{z}{\mu}\right)^{(k-1)} = 0$ for $k \leq n$.

Consider $z^{(n+1)}(\mu) = \left(\frac{z(\mu)}{\mu} f(\mu)\right)^{(n)}$.

We have

$$\begin{aligned} z^{(n+1)}(\mu) &= \left(\frac{z(\mu)}{\mu} f(\mu)\right)^{(n)} \\ &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} f^{(n-k)}(\mu) \left(\frac{z}{\mu}\right)^{(k)} \\ &= \left(\frac{z}{\mu}\right)^{(n)} f(\mu) + \sum_{k=0}^{n-1} \begin{pmatrix} n \\ k \end{pmatrix} f^{(n-k)}(\mu) \left(\frac{z}{\mu}\right)^{(k)}. \end{aligned}$$

Note that the second equality in above follows from product rule for derivatives.

Now, by induction hypothesis and because f is analytic at $\mu = 0$,

$$\lim_{\mu \to 0} \sum_{k=0}^{n-1} \begin{pmatrix} n \\ k \end{pmatrix} f^{(n-k)}(\mu) \left(\frac{z}{\mu}\right)^{(k)} = 0.$$

Therefore,

$$\lim_{\mu \to 0} z^{(n+1)}(\mu) = \lim_{\mu \to 0} \left(\frac{z}{\mu}\right)^{(n)} f(\mu).$$

By applying product rule for derivatives repeatedly on $\left(\frac{z}{\mu}\right)^{(n)}$, we have

$$\begin{aligned} \left(\frac{z}{\mu}\right)^{(n)} &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} z^{(n-k)}(\mu) \left(\frac{1}{\mu}\right)^{(k)} \\ &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} z^{(n-k)}(\mu) \left(\frac{(-1)^k k!}{\mu^{k+1}}\right). \end{aligned}$$

Applying L'Hopital's Rule on the last expression, we have

$$\begin{aligned} \lim_{\mu \to 0} \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} z^{(n-k)}(\mu) \left(\frac{(-1)^k k!}{\mu^{k+1}}\right) &= \lim_{\mu \to 0} \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \frac{(-1)^k k!}{(k+1)!} z^{(n+1)}(\mu) \\ &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \frac{(-1)^k}{k+1} z^{(n+1)}(0). \end{aligned}$$

Now since $\sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \frac{(-1)^k}{k+1} = \frac{1}{n+1}$ for all $n \geq 0$, we have that the last expression in above is equal to $\frac{z^{(n+1)}(0)}{n+1}$.

Substituting $\lim_{\mu\to 0}\left(\frac{z}{\mu}\right)^{(n)} = \frac{z^{(n+1)}(0)}{n+1}$ and $\lim_{\mu\to 0} f(\mu) = f(0)$ into $\lim_{\mu\to 0} z^{(n+1)}(\mu) = \lim_{\mu\to 0}\left(\frac{z}{\mu}\right)^{(n)} f(\mu)$. We have

$$z^{(n+1)}(0) = \frac{f(0)}{(n+1)} z^{(n+1)}(0).$$

which implies that $z^{(n+1)}(0) = 0$ and $\lim_{\mu\to 0}\left(\frac{z}{\mu}\right)^{(n)} = 0$ since $f(0)$ is not a positive integer.

Hence, by induction, $z^{(n)}(0) = 0$ and $\lim_{\mu\to 0}\left(\frac{z}{\mu}\right)^{(n-1)} = 0$ for all $n \geq 1$. Therefore, with $z(0)$ also equals to zero and $z(\mu)$ is analytic at $\mu = 0$, we have $z(\mu)$ is identically zero. **QED**

Remark 2.6 *Note that the result in Lemma 2.1 is a classical result and can be found for example in [7]. We include its proof here because it is elementary and does not require deep theoretical background to understand it.*

We have the following main theorem for this section:

Theorem 2.4 *Let $\widetilde{X}(t)$ and $\widetilde{Y}(t)$, given by (2.18), be positive definite for $t > 0$. Then $(\widetilde{X}(t), \widetilde{Y}(t))$ is a solution to (2.15) for $t > 0$ and is analytic at $t = 0$ if and only if $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ for all $t \geq 0$, where $\widetilde{y}_1(t)$ satisfies $\widetilde{y}_1' = \frac{2t\widetilde{y}_1(2-\widetilde{y}_1)}{1+2t^2(\widetilde{y}_1-1)}$.*

Proof. ($\Rightarrow$) Suppose $(\widetilde{X}(t), \widetilde{Y}(t))$ is a solution to (2.15) for $t > 0$ and is analytic at $t = 0$.

Then, from the first differential equation in (2.23), 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 (2.23), 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

$$\widetilde{y}_1' = \frac{2t(\widetilde{y}_1 - 2)(\widetilde{y}_1(\widetilde{y}_1 - 2 + 2w) + w^2)}{2 - t^2((2 - \widetilde{y}_1)^2 + 2w\widetilde{y}_1 + w^2)}. \tag{2.24}$$

From the second differential equation in (2.23), 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,

$$w' = \frac{2(2 - \widetilde{y}_1)((w + \widetilde{y}_1)(t^2 w - 1) + 2t^2 w)}{t(2 - t^2((2 - \widetilde{y}_1)^2 + 2w\widetilde{y}_1 + w^2))}. \tag{2.25}$$

Adding up equations (2.24) and (2.25) 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:

$$z'(t) = \frac{z(t)}{t}\left(\frac{2(2 - \widetilde{y}_1(t))(t^2(2 - \widetilde{y}_1(t)) - 1)}{2 - t^2((2 - \widetilde{y}_1(t))^2 + z^2 - \widetilde{y}_1^2)}\right). \tag{2.26}$$

Let $f(t) = \frac{2(2-\widetilde{y}_1(t))(t^2(2-\widetilde{y}_1(t))-1)}{2-t^2((2-\widetilde{y}_1(t))^2+z^2-\widetilde{y}_1^2)}$. Then $f(t)$ is analytic at $t = 0$. Also, $f(0) = -(2 - \widetilde{y}_1(0))$, which is strictly less than zero since $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$, in the proof of Proposition 2.3, are positive definite even in the limit as t approaches zero.

From (2.26), 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 2.1. 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 (2.24) in terms of $\widetilde{y}_1$, we obtained the ODE of $\widetilde{y}_1$ in the theorem.

($\Leftarrow$) Suppose $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ for all $t \geq 0$, where $\widetilde{y}_1(t)$ satisfies $\widetilde{y}_1' = \frac{2t\widetilde{y}_1(2-\widetilde{y}_1)}{1+2t^2(\widetilde{y}_1-1)}$. Then, since the right-hand side of the ODE of $\widetilde{y}_1$ is analytic at $t = 0$ and $\widetilde{y}_1 \in \Re$, we have, by Theorem 2.1, that $\widetilde{y}_1(t)$ is analytic at $t = 0$. Hence $\widetilde{y}_2(t)$ is also analytic at $t = 0$. 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 in the theorem, we can also check easily that $\widetilde{y}_1(t)$ and $\widetilde{y}_2(t)$ satisfy (2.16). Hence, by Proposition 2.2, $(\widetilde{X}(t), \widetilde{Y}(t))$ satisfies (2.15) for $t > 0$. **QED**

Using Theorem 2.4, we have the following interesting result:

Corollary 2.3 *Let $X(\mu)$, $Y(\mu)$, given by (2.17), be positive definite for $\mu > 0$. Suppose $(X(\mu), Y(\mu))$ is a solution to (2.11) for $\mu > 0$ with initial conditions given by (2.12)-(2.14). Then $(X(\mu), Y(\mu))$ is analytic w.r.t μ at $\mu = 0$ if and only if it is analytic w.r.t $\sqrt{\mu}$ at $\mu = 0$.*

Proof. ($\Rightarrow$) This is clear.

($\Leftarrow$) 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 2.4, we have $\widetilde{y}_2(t) = -t\widetilde{y}_1(t)$ for all $t \geq 0$, where $\widetilde{y}_1(t)$ satisfies $\widetilde{y}_1' = \frac{2t\widetilde{y}_1(2-\widetilde{y}_1)}{1+2t^2(\widetilde{y}_1-1)}$.
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 $y_2(\mu) = -y_1(\mu)$. 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 $\widetilde{\widetilde{y}}_1' = \frac{\widetilde{\widetilde{y}}_1(2-\widetilde{\widetilde{y}}_1)}{1+2\mu(\widetilde{\widetilde{y}}_1-1)}$ since $\widetilde{y}_1(t)$ satisfies $\widetilde{y}_1' = \frac{2t\widetilde{y}_1(2-\widetilde{y}_1)}{1+2t^2(\widetilde{y}_1-1)}$. Since the right-hand side of the ODE satisfied by $\widetilde{\widetilde{y}}_1(\mu)$ is analytic at $\mu = 0$, we have, by Theorem 2.1, $\widetilde{\widetilde{y}}_1(\mu)$ is also analytic at $\mu = 0$. Therefore, $y_1(\mu)$ and $y_2(\mu)$ are analytic at $\mu = 0$, which further implies that $(X(\mu), Y(\mu))$ is analytic at $\mu = 0$. Hence, we are done. **QED**

Remark 2.7 *From the proof of Corollary 2.3, we see that we have a result similar to Theorem 2.4 which is that $(X(\mu), Y(\mu))$, given by (2.17), is a solution to (2.11) for $\mu > 0$ and is analytic at $\mu = 0$ if and only if $y_2(\mu) = -y_1(\mu)$ for all $\mu \geq 0$, where $y_1(\mu) = \mu\widetilde{\widetilde{y}}_1(\mu)$ and $\widetilde{\widetilde{y}}_1(\mu)$ satisfies $\widetilde{\widetilde{y}}_1' = \frac{\widetilde{\widetilde{y}}_1(2-\widetilde{\widetilde{y}}_1)}{1+2\mu(\widetilde{\widetilde{y}}_1-1)}$.*

We also have:

Remark 2.8 *We see, from Theorem 2.4, that 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 w.r.t* μ *or* $\sqrt{\mu}$ *at* $\mu = 0$*. On the other hand, if the initial point satisfies a certain condition, its off-central path can be analytic at* $\mu = 0$*. In the next section, we will see how this latter fact can be used to ensure superlinear convergence of the first-order predictor-corrector algorithm.*

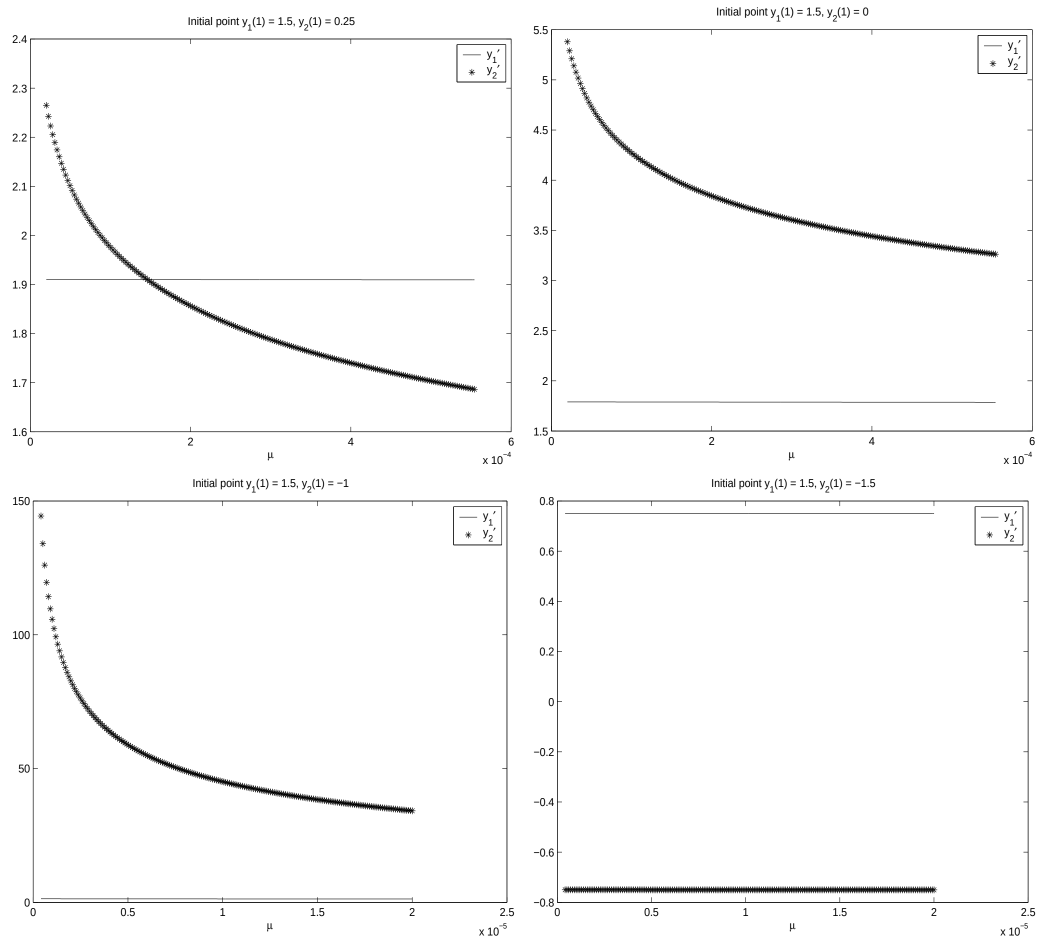
To end this section, we have the below final remark:

Remark 2.9 *If we consider* $P = X^{-1/2}$*, which corresponds to the so-called HKM direction, then by performing manipulations similar to the above (and hence will not be shown here), Theorem 2.4 also holds. In particular, we also have the interesting relation* $y_2 = -y_1$*, as in Remark 2.7. We do not know about the case of NT direction since manipulations for NT direction on this example proved to be too complicated. Finally, we remark that we choose the dual HKM direction over the HKM direction to show the results above because it is computationally advantageous to use this direction when we compute the iterates of path-following algorithm in general (see [25]). Hence it is more meaningful to show results using the dual HKM direction.*

2.2.1 Implications to Predictor-Corrector Algorithm

From the previous section, Remark 2.7, we note that not all off-central paths of the given example are analytic at the limit as μ approaches zero. In fact, we see that only if we start an off-central path, $(X(\mu), Y(\mu))$, at a point (X_0, Y_0) with $X_0 = \begin{pmatrix} 1 & 2\mu_0 - y_1^0 \\ 2\mu_0 - y_1^0 & 2\mu_0 - y_1^0 \end{pmatrix}$, $Y_0 = \begin{pmatrix} y_1^0 & y_2^0 \\ y_2^0 & 1 - 2y_2^0 \end{pmatrix}$ such that $y_2^0 = -y_1^0$, then it is analytic at the limit as $\mu \to 0$. This is a very restrictive condition for an off-central path to be analytic at $\mu = 0$.

One may ask whether for other starting points, the off-central path may have at least bounded first derivatives as μ approaches zero. We have written codes in matlab to see how the first derivatives of $y_1(\mu)$ and $y_2(\mu)$ behave, as μ approaches zero, for different starting points. The results are shown in the figures below:



We see from these figures that indeed, without $y_2 = -y_1$, the first derivatives of the off-central path do not seems to be bounded in the limit as μ approaches zero. From [26], it then suggests that we cannot conclude superlinear convergence of the predictor-corrector algorithm using this example, if we choose any point as the initial iterate. However, in what follows, we will show that by choosing suitable initial iterate, superlinear convergence of first-order predictor-corrector algorithm on this example is still possible.

Let us first define a set $\mathcal{S}$, for the given example, which is the collection of all off-central paths in $S^2_{++} \times S^2_{++}$ which are analytic at their limit point as $\mu \to 0$.

We have the following observation on the structure of $\mathcal{S}$:

Proposition 2.4

$$\mathcal{S} = \{(X,Y)\ :\ X,Y \in S^2_{++},\ \mathcal{A}svec(X) + \mathcal{B}svec(Y) = q,\ Y_{12} = -Y_{11}\}$$

Proof. ($\subseteq$) Let $(X,Y) \in \mathcal{S}$. Clearly $(X,Y) \in \{(X,Y)\ ;\ X,Y \in S^2_{++},\ \mathcal{A}svec(X) + \mathcal{B}svec(Y) = q,\ Y_{12} = -Y_{11}\}$.

($\supseteq$) Let $(X,Y) \in \{(X,Y)\ ;\ X,Y \in S^2_{++},\ \mathcal{A}svec(X) + \mathcal{B}svec(Y) = q,\ Y_{12} = -Y_{11}\}$. Then $\mathcal{A}svec(X) + \mathcal{B}svec(Y) = q,\ Y_{12} = -Y_{11} \Rightarrow X = \begin{pmatrix} 1 & x^0 \\ x^0 & x^0 \end{pmatrix}, Y = \begin{pmatrix} y_1^0 & y_2^0 \\ y_2^0 & 1-2y_2^0 \end{pmatrix}$ for some $x^0, y_1^0, y_2^0 \in \Re$ and $y_2^0 = -y_1^0$. Define $\mu_0 = Tr(XY)/2$. Then $x^0 = 2\mu_0 - y_1^0$. By Remark 2.7, the ODE of $\widetilde{\widetilde{y}}_1$ there has a solution with initial point $\widetilde{\widetilde{y}}_1(\mu_0) = y_1^0/\mu_0$ and its resulting off-central path $(X(\mu), Y(\mu))$, using the ODE solution $\widetilde{\widetilde{y}}_1(\mu)$, is analytic at $\mu = 0$. This off-central path has (X,Y) as the point at μ_0. Hence $(X,Y) \in \mathcal{S}$. **QED**

In the first-order predictor-corrector algorithm, the predictor and corrector steps are obtained by solving the following system of equations:

$$\begin{aligned} H_P(X\Delta Y + \Delta XY) &= \sigma\mu I - H_P(XY) \\ A(\Delta X) + B(\Delta Y) &= 0 \end{aligned}$$

where (X,Y) is the current iterate and for $\sigma = 0$, $(\Delta X, \Delta Y)$ corresponds to the predictor step, $(\Delta_p X, \Delta_p Y)$, and for $\sigma = 1$, $(\Delta X, \Delta Y)$ corresponds to the corrector step, $(\Delta_c X, \Delta_c Y)$. Also, $\mu = Tr(XY)/n$ where n is the matrix size of X (or Y).

The intermediate iterate, (X^p, Y^p), after the predictor step, is obtained by adding suitable scalar multiple of $(\Delta_p X, \Delta_p Y)$ to (X,Y). The next iterate (X^+, Y^+) of the algorithm is then obtained by adding $(\Delta_c X, \Delta_c Y)$ to (X^p, Y^p).

We want to show for the example that if $(X, Y) \in \mathcal{S}$, then the next iterate, (X^+, Y^+), also belongs to $\mathcal{S}$. It then follows that if our initial feasible iterate $(X_0, Y_0) \in \mathcal{S}$, then any iterate generated by the first-order predictor-corrector algorithm lies on an off-central path which is analytic at the optimal solution (since it also belongs to $\mathcal{S}$) and hence, by [26], the iterates converge quadratically to the optimal solution.

We have the following proposition:

Proposition 2.5 *If* $(X, Y) \in \mathcal{S}$, *then* $(X^p, Y^p) \in \mathcal{S}$.

Proof. We know that the derivative at the point where the off-central path passes through (X, Y) is along the same direction as $(\Delta_p X, \Delta_p Y)$. Therefore, $(\Delta_p X, \Delta_p Y)$ has the form

$$\Delta_p X = -\mu \begin{pmatrix} 0 & 2 - \Delta_p y_1 \\ 2 - \Delta_p y_1 & 2 - \Delta_p y_1 \end{pmatrix}, \quad \Delta_p Y = -\mu \begin{pmatrix} \Delta_p y_1 & \Delta_p y_2 \\ \Delta_p y_2 & -2\Delta_p y_2 \end{pmatrix}$$

where $\Delta_p y_2 = -\Delta_p y_1$.
Therefore, $(X^p, Y^p) = (X, Y) + \alpha(\Delta_p X, \Delta_p Y)$ for some $\alpha > 0$ implies that $(Y^p)_{11} = Y_{11} - \alpha\mu\Delta_p y_1$ and $(Y^p)_{12} = Y_{12} - \alpha\mu\Delta_p y_2$. Clearly, $(Y^p)_{12} = -(Y^p)_{11}$. Also, since $\mathcal{A}svec(X^p) + \mathcal{B}svec(Y^p) = q$, we are done, by Proposition 2.4. **QED**

Next, we show that if $(X^p, Y^p) \in \mathcal{S}$, then (X^+, Y^+) also belongs to $\mathcal{S}$ and we would have shown that all iterates generated by the first-order predictor-corrector algorithm, if suitably initialized, have the nice property as stated above.
We do this by studying the path corresponding to the corrector step, which is the solution of the following system of ODEs:

$$H_P((XY)') = \frac{Tr(XY)}{2} I - H_P(XY) \tag{2.27}$$

$$A(X') + B(Y') = 0 \tag{2.28}$$

where $P = Y^{1/2}$. We denote the parameter in (2.27)-(2.28) by the variable t. Note that taking trace on both sides of (2.27) and integrating w.r.t t, we see that on a solution of the system of ODEs (2.27)-(2.28), $Tr(XY)/2$ is equal to a constant μ^+ for all t. Therefore, we will write (2.27) as

$$H_P((XY)') = \mu^+ I - H_P(XY) \tag{2.29}$$

from now onwards, where μ^+ is a constant.

For the solution curve of (2.28)-(2.29), $(X(t), Y(t))$, passing through (X^p, Y^p) (and hence satisfying $\mathcal{A}svec(X) + \mathcal{B}svec(Y) = q$ and $Tr(XY) = \mu^+$), we see that it is of the form $\left(\begin{pmatrix} 1 & 2\mu^+ - w_1(t) \\ 2\mu^+ - w_1(t) & 2\mu^+ - w_1(t) \end{pmatrix}, \begin{pmatrix} w_1(t) & w_2(t) \\ w_2(t) & 1 - 2w_2(t) \end{pmatrix}\right)$, which satisfies (2.28) automatically.

We have the following proposition:

Proposition 2.6 *Let*

$$(X(t), Y(t)) = \left(\begin{pmatrix} 1 & 2\mu^+ - w_1(t) \\ 2\mu^+ - w_1(t) & 2\mu^+ - w_1(t) \end{pmatrix}, \begin{pmatrix} w_1(t) & w_2(t) \\ w_2(t) & 1 - 2w_2(t) \end{pmatrix}\right)$$

where $w_2(t) = -w_1(t)$ *with* $w_1(t)$ *satisfying* $(1 + 2w_1 - 2\mu^+)w_1' = \mu^+ + (2\mu^+ - 1)w_1 - w_1^2$, $w_1(0) = (Y^p)_{11}$. *Then* $(X(t), Y(t))$ *is the unique solution of (2.28)-(2.29) passing through* (X^p, Y^p).

Proof. Suppose $(X(t), Y(t))$ satisfies the conditions in the proposition.

Then we first observe that $(X(t), Y(t))$ of the given form satisfies (2.28) automatically. This is noted in the discussion before the proposition. Therefore, we only need to show that $(X(t), Y(t))$ satisfies (2.29) and then by Theorem 2.1, it is the unique solution of (2.28)-(2.29) passing through (X^p, Y^p).

Note that (2.29) can be written as $(Y^{1/2} \otimes_s Y^{1/2})svec(X') + ((Y^{1/2}X) \otimes_s Y^{-1/2})svec(Y') = \mu^+ svec(I) - (Y^{1/2} \otimes_s Y^{1/2})svec(X)$ using $svec$ and $\otimes_s$ notations. Taking the inverse of $Y^{1/2} \otimes_s Y^{1/2}$ on both sides of this equation and using the properties of $\otimes_s$, we get

$$svec(X') + (X \otimes_s Y^{-1})svec(Y') = \mu^+ svec(Y^{-1}) - svec(X). \tag{2.30}$$

Substituting

$$(X(t),Y(t)) = \left(\begin{pmatrix} 1 & 2\mu^+ - w_1(t) \\ 2\mu^+ - w_1(t) & 2\mu^+ - w_1(t) \end{pmatrix}, \begin{pmatrix} w_1(t) & w_2(t) \\ w_2(t) & 1-2w_2(t) \end{pmatrix}\right)$$

and expressions for X', Y', Y^{-1} and $X \otimes_s Y^{-1}$ (this expression in terms of $w_1(t)$ and $w_2(t)$ can be easily derived from a similar expression in Section 2.2) in terms of $w_1(t)$ and $w_2(t)$ into (2.30) and upon simplification, we get the following three equations:

$$(1-2w_2)w_1' + ((2\mu^+ - w_1) - w_2)w_2' = (1-2w_2)(\mu^+ - w_1) + w_2^2, \tag{2.31}$$

$$\begin{gathered}(1-2w_2)(2\mu^+ - 3w_1 - w_2)w_1' + (2\mu^+ - 2(2\mu^+ - w_1)(w_1+w_2))w_2' = \\ -2\mu^+ w_2 - 2(w_1(1-2w_2) - w_2^2)(2\mu^+ - w_1)\end{gathered} \tag{2.32}$$

and

$$\begin{gathered}(w_2(2\mu^+ - w_1 - w_2) + w_1(1-2w_2))w_1' + (2\mu^+ - w_1)(w_1+w_2)w_2' = \\ -\mu^+ w_1 + (w_1(1-2w_2) - w_2^2)(2\mu^+ - w_1).\end{gathered} \tag{2.33}$$

We can easily check that if $w_1(t)$ and $w_2(t)$ of $(X(t),Y(t))$ are given by the conditions in the proposition, then they satisfy (2.31)-(2.33). Hence $(X(t),Y(t))$ in the proposition satisfies (2.29). Therefore, we are done. **QED**

As in the proof of Proposition 2.5, we observe that the derivative of the solution $(X(t),Y(t))$ to (2.28)-(2.29) passing through (X^p,Y^p) is along the same direction as $(\Delta_c X, \Delta_c Y)$. Therefore, by Proposition 2.6, $(\Delta_c X, \Delta_c Y)$ has the form

$$\Delta_c X = \begin{pmatrix} 0 & -\Delta_c w_1 \\ -\Delta_c w_1 & -\Delta_c w_1 \end{pmatrix}, \quad \Delta_c Y = \begin{pmatrix} \Delta_c w_1 & \Delta_c w_2 \\ \Delta_c w_2 & -2\Delta_c w_2 \end{pmatrix}$$

where $\Delta_c w_2 = -\Delta_c w_1$. Adding this to (X^p,Y^p) (which satisfies $(Y^p)_{12} = -(Y^p)_{11}$ and $\mathcal{A}svec(X^p) + \mathcal{B}svec(Y^p) = q$), we see that (X^+,Y^+) also satisfies $(Y^+)_{12} = -(Y^+)_{11}$ and $\mathcal{A}svec(X^+) + \mathcal{B}svec(Y^+) = q$. Therefore, $(X^+,Y^+) \in \mathcal{S}$.

In conclusion, in this section, we show that for the example under consideration, if the initial iterate for the first-order predictor-corrector algorithm lies on an off-central path which is analytic at its limit point, then all iterates generated by the algorithm also lies on some off-central path analytic at its limit point. Hence, these iterates converge superlinearly to the optimal solution.

2.3 General Theory for Asymptotic Analyticity of Off-Central Path for SDLCP

In Section 2.1, we shown that any accumulation point of $(X(\mu), Y(\mu))$, the solution to (2.6)-(2.7) in $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$, is a solution to (2.1), as μ tends to zero. In this section, the asymptotic behaviour of $(X(\mu), Y(\mu))$ will be analysed. Instead of studying the limiting behaviour of $(X(\mu), Y(\mu))$ for general P, which is too daunting a task, we will do so only for the case when $P = Y^{1/2}$, the so-called dual HKM direction. Note that the case when $P = I$ has already been studied in [17] and hence will not be discussed here.

We first make a few transformations to (2.8) which is an equivalent form of (2.6)-(2.7). The system of ODEs obtained after these transformations allows us to give a necessary and sufficient condition to when an off-central path $(X(\mu), Y(\mu))$ is analytic at its limit point with respect to $t = \sqrt{\mu}$. We only attempt to study the analyticity of the off-central path at its limit point with respect to $\sqrt{\mu}$ instead of μ because $\sqrt{\mu}$ naturally appears in the off diagonal entries of $X(\mu), Y(\mu)$, as shown in (2.34) and (2.35) below. This leads us to naturally investigate asymptotic behaviour of $X(\mu), Y(\mu)$ w.r.t $\sqrt{\mu}$.

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

We need an additional assumption besides Assumption 2.1 before we proceed. The analysis of the asymptotic behaviour of an off-central path for a general SDLCP is difficult without this assumption although there are some recent work done in this area for special classes of SDLCP without the assumption. See for example [16].

Here, we will discuss the case of SDLCP (2.1) with the assumption (in addition to Assumption 2.1), which is stated below.

Assumption 2.2 *There exists a strictly complementary solution,* (X^*, Y^*), *to SDLCP (2.1).*

Since X^* and Y^* commutes, they are jointly diagonalizable by some orthogonal matrix. So, using a suitable orthogonal similarity transformation of the matrices in SDLCP (2.1), we may assume, without loss of generality, that

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

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

Hereafter, whenever we partitioned 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 (2.8) into a more "manageable" system of ODEs, we now claim that for $(X(\mu), Y(\mu))$ on an off-central path, we have

$$X(\mu) = \begin{pmatrix} X_{11} & \sqrt{\mu}\widetilde{X}_{12} \\ \sqrt{\mu}\widetilde{X}_{12}^T & \mu\widetilde{X}_{22} \end{pmatrix} \tag{2.34}$$

and

$$Y(\mu) = \begin{pmatrix} \mu\widetilde{Y}_{11} & \sqrt{\mu}\widetilde{Y}_{12} \\ \sqrt{\mu}\widetilde{Y}_{12}^T & Y_{22} \end{pmatrix} \tag{2.35}$$

where $X_{11}, Y_{22}, \widetilde{X}_{22}$ and $\widetilde{Y}_{11}$ are equal to $\Theta(1)$ and $\|\widetilde{X}_{12}(\mu)\|_F, \|\widetilde{Y}_{12}(\mu)\|_F$ are equal to $\mathcal{O}(1)$. We proved this in a few propositions below. These propositions are adapted from [17].

Proposition 2.7 *([17] Lemma 3.10) $Y_{11}(\mu)$ and $X_{22}(\mu)$ are equal to $\mathcal{O}(\mu)$ and $\|X_{12}(\mu)\|_F$ and $\|Y_{12}(\mu)\|_F$ are equal to $\mathcal{O}(\sqrt{\mu})$.*

Proof. Now, $A(X(\mu) - X^*) + B(Y(\mu) - Y^*) = 0$ implies, by Assumption 2.1(a), that $(X(\mu)-X^*)\bullet(Y(\mu)-Y^*) \geq 0$. Hence $X(\mu)\bullet Y^* + X^*\bullet Y(\mu) \leq X(\mu)\bullet Y(\mu) = Tr(XY)(\mu)$.

Note that by Remark 2.3, $Tr(XY)(\mu) = Tr(X^0Y^0)\mu = n\mu$. Hence, $X(\mu) \bullet Y^* + X^* \bullet Y(\mu) = \mathcal{O}(\mu)$. That is, $\sum_{i=m+1}^n \lambda_i^* x_{ii}(\mu) + \sum_{i=1}^m \lambda_i^* y_{ii}(\mu) = \mathcal{O}(\mu)$, where $x_{ii}(\mu), y_{ii}(\mu)$ are the diagonal elements of $X(\mu)$ and $Y(\mu)$ respectively. This implies that $X_{22}(\mu) = \mathcal{O}(\mu)$ and $Y_{11}(\mu) = \mathcal{O}(\mu)$.

Also, we have $\|X_{12}(\mu)\|_F^2 \leq Tr(X_{11}(\mu))Tr(X_{22}(\mu))$ (by Lemma 2.2 of [17]), together with the fact that $X(\mu)$ is bounded near μ equal to zero (by Remark 2.4) and $X_{22}(\mu) = \mathcal{O}(\mu)$, implies that $\|X_{12}(\mu)\|_F = \mathcal{O}(\sqrt{\mu})$.

Similarly, we can show that $\|Y_{12}(\mu)\|_F = \mathcal{O}(\sqrt{\mu})$. **QED**

Proposition 2.8 *([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. Now, $det\left(\frac{X(\mu)Y(\mu)}{\mu}\right) = \prod_{i=1}^n \frac{\lambda_i(XY)(\mu)}{\mu} \geq \lambda_{min}(X^0Y^0)^n$, where the inequality follows from Theorem 2.2.

On the other hand, $det\left(\frac{X(\mu)Y(\mu)}{\mu}\right) = \frac{1}{\mu^n} det(X(\mu))det(Y(\mu)) \leq det(X_{11}(\mu))\, det\left(\frac{X_{22}(\mu)}{\mu}\right) \times det(Y_{22}(\mu))\, det\left(\frac{Y_{11}(\mu)}{\mu}\right)$ (where the inequality follows from Theorem 2.4 in [17]). Therefore, we have $\lambda_{min}(X^0Y^0)^n \leq det(X_{11}(\mu))det\left(\frac{X_{22}(\mu)}{\mu}\right) det(Y_{22}(\mu))det\left(\frac{Y_{11}(\mu)}{\mu}\right)$. Taking log on both sides of the inequality, we have

$$\begin{aligned} n\log\lambda_{min}(X^0Y^0) &\leq \log det(X_{11}(\mu)) + \log det\left(\frac{X_{22}(\mu)}{\mu}\right) + \log det(Y_{22}(\mu)) \\ &\quad + \log det\left(\frac{Y_{11}(\mu)}{\mu}\right). \end{aligned}$$

Since $X_{22}(\mu)$ and $Y_{11}(\mu)$ are equal to $\mathcal{O}(\mu)$ (by the previous proposition) and $X(\mu), Y(\mu)$ are bounded (by Remark 2.4), we must have, from the above logarithmic inequality, that $X_{11}(\mu)$ and $Y_{22}(\mu)$ are equal to $\Theta(1)$ and $X_{22}(\mu)$ and $Y_{11}(\mu)$ are equal to $\Theta(\mu)$. **QED**

Therefore, our claim on $(X(\mu), Y(\mu))$ of an off-central path having form (2.34) and (2.35) is true.

Letting $\widetilde{X}(\mu) = \begin{pmatrix} X_{11} & \widetilde{X}_{12} \\ \widetilde{X}_{12}^T & \widetilde{X}_{22} \end{pmatrix}$ and $\widetilde{Y}(\mu) = \begin{pmatrix} \widetilde{Y}_{11} & \widetilde{Y}_{12} \\ \widetilde{Y}_{12}^T & Y_{22} \end{pmatrix}$, we can then write

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

and

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

Remark 2.10 *It can be seen easily that since $\lambda_{min}(XY)(\mu) = \lambda_{min}(X^0Y^0)\mu$ and $\lambda_{max}(XY)(\mu) = \lambda_{max}(X^0Y^0)\mu$ (Thereom 2.2), the above relationship between $\widetilde{X}$, X and $\widetilde{Y}$, Y implies that $\lambda_{min}(\widetilde{X}\widetilde{Y})(\mu) = \lambda_{min}(X^0Y^0)$ and $\lambda_{max}(\widetilde{X}\widetilde{Y})(\mu) = \lambda_{max}(X^0Y^0)$. Hence $\widetilde{X}(\mu)$ and $\widetilde{Y}(\mu)$ are positive definite for all $\mu > 0$ and any of their accumulation points are also positive definite.*

Let $X_1(t) = X(t^2)$, $Y_1(t) = Y(t^2)$. Similarly, let $\widetilde{X}_1(t) = \widetilde{X}(t^2)$ and $\widetilde{Y}_1(t) = \widetilde{Y}(t^2)$. Then $X_1, \widetilde{X}_1, Y_1, \widetilde{Y}_1$ are related by

$$X_1(t) = \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \widetilde{X}_1(t) \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \tag{2.36}$$

and

$$Y_1(t) = \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \widetilde{Y}_1(t) \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}. \tag{2.37}$$

To study the analyticity of $(X(\mu), Y(\mu))$ w.r.t $\sqrt{\mu}$ at $\mu = 0$, it 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 2.9 *$X_1(t)$ 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. This is clear since by (2.36) and (2.37), $(X_1)_{11}(t) = (\widetilde{X}_1)_{11}(t)$, $(X_1)_{12}(t) = t(\widetilde{X}_1)_{12}(t)$, $(X_1)_{22}(t) = t^2(\widetilde{X}_1)_{22}(t)$, $(Y_1)_{11}(t) = t^2(\widetilde{Y}_1)_{11}(t)$, $(Y_1)_{12}(t) = t(\widetilde{Y}_1)_{12}(t)$ and $(Y_1)_{22}(t) = (\widetilde{Y}_1)_{22}(t)$. We also need the fact that $\widetilde{X}_1$ and $\widetilde{Y}_1$ are bounded near $t = 0$, which follows from Propositions 2.7 and 2.8. **QED**

Therefore, by this proposition, we need only study 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 for studying the asymptotic behaviour of $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ than that of $X_1(t)$ and $Y_1(t)$ is because their accumulation points are positive definite, by Remark 2.10 (which is a desirable property), unlike that of $X_1(t)$ and $Y_1(t)$.

Hence, we are going to express the system of ODEs (2.8) in terms of $\widetilde{X}_1$ and $\widetilde{Y}_1$.

First, we observe that by letting $P = Y^{1/2}$ in (2.8), inverting $P \otimes_s P^{-T}Y$ and observing that $(P \otimes_s P^{-T}Y)^{-1}(PX \otimes_s P^{-T}) = X \otimes_s Y^{-1}$ and $(P \otimes_s P^{-T}Y)^{-1}$ $svec(H_P(XY)) = svec(X)$ when $P = Y^{1/2}$, (2.8) becomes

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X \otimes_s Y^{-1} \end{pmatrix} \begin{pmatrix} svec(X') \\ svec(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} 0 \\ svec(X) \end{pmatrix} \tag{2.38}$$

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

In terms of X_1 and Y_1, (2.38) becomes

$$\frac{1}{2}\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X_1 \otimes_s Y_1^{-1} \end{pmatrix}\begin{pmatrix} svec(X_1') \\ svec(Y_1') \end{pmatrix} = \frac{1}{t}\begin{pmatrix} 0 \\ svec(X_1) \end{pmatrix} \tag{2.39}$$

If we consider X_1 and Y_1 to be on an off-central path, then the matrix on the extreme left in (2.39) is not invertible and may not even be defined as t tends to zero (since Y_1^{-1} may not exist in the limit) and hence it is not possible to analyse the asymptotic behaviour of $X_1(t)$ and $Y_1(t)$ if we just use (2.39). By expressing (2.39) in terms of $\widetilde{X}_1$ and $\widetilde{Y}_1$, we will see that further analysis is possible.

In what follows, as in Section 2.1, the properties of the operation $\otimes_s$ and the map *svec* are used extensively. Also, following [23], the inverse map of *svec* is denoted by *smat*.

Note that since

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

we have

$$\begin{aligned} X_1'(t) \quad = \quad & \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \widetilde{X}_1(t) \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \widetilde{X}_1'(t) \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} + \\ & \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \widetilde{X}_1(t) \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}. \end{aligned}$$

Therefore,

$$\begin{aligned} svec(X_1'(t)) \quad = \quad & 2\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) svec(\widetilde{X}_1(t)) + \\ & \left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) svec(\widetilde{X}_1'(t)). \end{aligned}$$

Similarly,

$$svec(Y_1'(t)) = 2\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) svec(\widetilde{Y}_1(t)) + \left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) svec(\widetilde{Y}_1'(t)).$$

Substituting these into (2.39), we have

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X_1 \otimes_s Y_1^{-1} \end{pmatrix} \begin{pmatrix} \left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) svec(\widetilde{X}_1) \\ \left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) svec(\widetilde{Y}_1) \end{pmatrix} +$$

$$\frac{1}{2}\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & X_1 \otimes_s Y_1^{-1} \end{pmatrix} \begin{pmatrix} \left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) svec(\widetilde{X}_1') \\ \left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) svec(\widetilde{Y}_1') \end{pmatrix}$$

$$= \frac{1}{t}\begin{pmatrix} 0 \\ \left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) svec(\widetilde{X}_1) \end{pmatrix}$$

Therefore,

$$\frac{1}{2}\mathcal{M}_1 \begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \frac{1}{t}\begin{pmatrix} 0 \\ \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} svec(\widetilde{X}_1) \end{pmatrix} - \mathcal{M}_2 \begin{pmatrix} svec(\widetilde{X}_1) \\ svec(\widetilde{Y}_1) \end{pmatrix}, \tag{2.40}$$

where

$$\mathcal{M}_1 := \begin{pmatrix} \mathcal{A}\left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) & \mathcal{B}\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) \\ \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} & (X_1 \otimes_s Y_1^{-1})\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) \end{pmatrix}$$

and

$$\mathcal{M}_2 := \begin{pmatrix} \mathcal{A}\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) & \mathcal{B}\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) \\ \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} & (X_1 \otimes_s Y_1^{-1})\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) \end{pmatrix}.$$

Let us look more closely at $\mathcal{M}_1$. We will show that it can be written as a product of two matrices where one of the matrices is invertible for all $t \geq 0$ and $\widetilde{X}_1$, $\widetilde{Y}_1$ positive definite.

First, consider the matrices $\mathcal{A}, \mathcal{B}$ in $\mathcal{M}_1$ (and also in $\mathcal{M}_2$). We have the following lemma (note that the lemma is inspired by a similar result in [17], see also [9]):

Lemma 2.2 *There exists an invertible matrix T such that*

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

$$\begin{pmatrix} \left(svec\begin{pmatrix} (\widetilde{A_1})_{11} & (\widetilde{A_1})_{12} \\ (\widetilde{A_1})_{12}^T & (\widetilde{A_1})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_1})_{11} & (\widetilde{B_1})_{12} \\ (\widetilde{B_1})_{12}^T & (\widetilde{B_1})_{22} \end{pmatrix}\right)^T \\ \vdots & \vdots \\ \left(svec\begin{pmatrix} (\widetilde{A_{i_1}})_{11} & (\widetilde{A_{i_1}})_{12} \\ (\widetilde{A_{i_1}})_{12}^T & (\widetilde{A_{i_1}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{i_1}})_{11} & (\widetilde{B_{i_1}})_{12} \\ (\widetilde{B_{i_1}})_{12}^T & (\widetilde{B_{i_1}})_{22} \end{pmatrix}\right)^T \\ \left(svec\begin{pmatrix} 0 & (\widetilde{A_{i_1+1}})_{12} \\ (\widetilde{A_{i_1+1}})_{12}^T & (\widetilde{A_{i_1+1}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{i_1+1}})_{11} & (\widetilde{B_{i_1+1}})_{12} \\ (\widetilde{B_{i_1+1}})_{12}^T & 0 \end{pmatrix}\right)^T \\ \vdots & \vdots \\ \left(svec\begin{pmatrix} 0 & (\widetilde{A_{i_1+i_2}})_{12} \\ (\widetilde{A_{i_1+i_2}})_{12}^T & (\widetilde{A_{i_1+i_2}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{i_1+i_2}})_{11} & (\widetilde{B_{i_1+i_2}})_{12} \\ (\widetilde{B_{i_1+i_2}})_{12}^T & 0 \end{pmatrix}\right)^T \\ \left(svec\begin{pmatrix} 0 & 0 \\ 0 & (\widetilde{A_{i_1+i_2+1}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{i_1+i_2+1}})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T \\ \vdots & \vdots \\ \left(svec\begin{pmatrix} 0 & 0 \\ 0 & (\widetilde{A_{\tilde{n}}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{\tilde{n}}})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T \end{pmatrix} \quad (2.41)$$

where $0 \leq i_1, i_2 \leq \tilde{n}$ *- how* i_1 *and* i_2 *are defined should be clear from the proof of the lemma.*

Proof. In order to prove this, imagine that $svec(A_i)^T$ is written as $(\widehat{(A_i)_{11}}\ \ \widehat{(A_i)_{12}}\ \ \widehat{(A_i)_{22}})$ where $\widehat{(A_i)_{11}}$ is a row of vector corresponding to the upper left hand block $(A_i)_{11}$ of A_i, $\widehat{(A_i)_{12}}$ corresponds to the upper right hand block $(A_i)_{12}$ of A_i and $\widehat{(A_i)_{22}}$ corresponds to the lower right hand block $(A_i)_{22}$ of A_i. Similarly, $svec(B_i)^T$ is written as $(\widehat{(B_i)_{11}}\ \ \widehat{(B_i)_{12}}\ \ \widehat{(B_i)_{22}})$.

Let

$$i_1 = rank \begin{pmatrix} \widehat{(A_1)_{11}} & \widehat{(B_1)_{22}} \\ \vdots & \vdots \\ \widehat{(A_{\tilde{n}})_{11}} & \widehat{(B_{\tilde{n}})_{22}} \end{pmatrix},$$

$$i_2 = rank \begin{pmatrix} \widehat{(A_1)_{11}} & \widehat{(A_1)_{12}} & \widehat{(B_1)_{12}} & \widehat{(B_1)_{22}} \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{(A_{\tilde{n}})_{11}} & \widehat{(A_{\tilde{n}})_{12}} & \widehat{(B_{\tilde{n}})_{12}} & \widehat{(B_{\tilde{n}})_{22}} \end{pmatrix} - i_1,$$

$$i_3 = rank \begin{pmatrix} svec(A_1)^T & svec(B_1)^T \\ \vdots & \vdots \\ svec(A_{\tilde{n}})^T & svec(B_{\tilde{n}})^T \end{pmatrix} - (i_1 + i_2),$$

where $i_1 + i_2 + i_3 = \tilde{n}$ (by Assumption 2.1(c)). Then the lemma holds by applying block Gaussian elimination method to

$$\begin{pmatrix} svec(A_1)^T & svec(B_1)^T \\ \vdots & \vdots \\ svec(A_{\tilde{n}})^T & svec(B_{\tilde{n}})^T \end{pmatrix}.$$

QED

Remark 2.11 *It should be noted, by construction, that in the above*

$$i_1 = rank \begin{pmatrix} \left(svec\begin{pmatrix} \widetilde{(A_1)}_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} 0 & 0 \\ 0 & \widetilde{(B_1)}_{22} \end{pmatrix}\right)^T \\ \vdots & \vdots \\ \left(svec\begin{pmatrix} \widetilde{(A_{i_1})}_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} 0 & 0 \\ 0 & \widetilde{(B_{i_1})}_{22} \end{pmatrix}\right)^T \end{pmatrix},$$

$$i_2 = rank \begin{pmatrix} \left(svec\begin{pmatrix} 0 & \widetilde{(A_{i_1+1})}_{12} \\ \widetilde{(A_{i_1+1})}_{12}^T & 0 \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} 0 & \widetilde{(B_{i_1+1})}_{12} \\ \widetilde{(B_{i_1+1})}_{12}^T & 0 \end{pmatrix}\right)^T \\ \vdots & \vdots \\ \left(svec\begin{pmatrix} 0 & \widetilde{(A_{i_1+i_2})}_{12} \\ \widetilde{(A_{i_1+i_2})}_{12}^T & 0 \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} 0 & \widetilde{(B_{i_1+i_2})}_{12} \\ \widetilde{(B_{i_1+i_2})}_{12}^T & 0 \end{pmatrix}\right)^T \end{pmatrix},$$

$i_3 =$

$$rank \begin{pmatrix} \left(svec\begin{pmatrix} 0 & 0 \\ 0 & (\widetilde{A_{i_1+i_2+1}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{i_1+i_2+1}})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T \\ \vdots & \vdots \\ \left(svec\begin{pmatrix} 0 & 0 \\ 0 & (\widetilde{A_{\tilde{n}}})_{22} \end{pmatrix}\right)^T & \left(svec\begin{pmatrix} (\widetilde{B_{\tilde{n}}})_{11} & 0 \\ 0 & 0 \end{pmatrix}\right)^T \end{pmatrix},$$

where $i_1 + i_2 + i_3 = \tilde{n}$.

From now onwards, we can assume, without loss of generality, that $\mathcal{A} = \begin{pmatrix} svec(A_1)^T \\ \vdots \\ svec(A_{\tilde{n}})^T \end{pmatrix}$ and $\mathcal{B} = \begin{pmatrix} svec(B_1)^T \\ \vdots \\ svec(B_{\tilde{n}})^T \end{pmatrix}$ are given by (2.41). In these forms, again, $(\mathcal{A}\ \mathcal{B})$ have full row rank and

$$\mathcal{A}u + \mathcal{B}v = 0 \quad \Rightarrow \quad u^T v \geq 0. \tag{2.42}$$

Now, for each $i = 1, \ldots, \tilde{n}$,

$$svec(A_i)^T \left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \right) = \left(svec \begin{pmatrix} (A_i)_{11} & t(A_i)_{12} \\ t(A_i)_{12}^T & t^2(A_i)_{22} \end{pmatrix} \right)^T.$$

Together with form (2.41) for $\mathcal{A}$, we can see easily that

$$\mathcal{A} \left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \right) =$$

$$diag(I, tI, t^2 I) \begin{pmatrix} \left(svec \begin{pmatrix} (A_1)_{11} & t(A_1)_{12} \\ t(A_1)_{12}^T & t^2 (A_1)_{22} \end{pmatrix} \right)^T \\ \vdots \\ \left(svec \begin{pmatrix} (A_{i_1})_{11} & t(A_{i_1})_{12} \\ t(A_{i_1})_{12}^T & t^2 (A_{i_1})_{22} \end{pmatrix} \right)^T \\ \left(svec \begin{pmatrix} 0 & (A_{i_1+1})_{12} \\ (A_{i_1+1})_{12}^T & t(A_{i_1+1})_{22} \end{pmatrix} \right)^T \\ \vdots \\ \left(svec \begin{pmatrix} 0 & (A_{i_1+i_2})_{12} \\ (A_{i_1+i_2})_{12}^T & t(A_{i_1+i_2})_{22} \end{pmatrix} \right)^T \\ \left(svec \begin{pmatrix} 0 & 0 \\ 0 & (A_{i_1+i_2+1})_{22} \end{pmatrix} \right)^T \\ \vdots \\ \left(svec \begin{pmatrix} 0 & 0 \\ 0 & (A_{\tilde{n}})_{22} \end{pmatrix} \right)^T \end{pmatrix}.$$

Let the matrix on the extreme right in the above expression be $\mathcal{A}(t)$.

Remark 2.12 *Note that in this section, $diag(I, tI, t^2 I)$ or $diag(I, tI, t^2 I, C)$ where C is a matrix, whenever it appears, has its first diagonal block of dimension i_1, its second diagonal block of dimension i_2 and its third diagonal block of dimension $\widetilde{n} - i_1 - i_2 = i_3$.*

In a similar fashion, we have

$$\mathcal{B}\left(\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \right) \right) = diag(I, tI, t^2 I)\mathcal{B}(t),$$

where

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

Therefore, we have the following lemma,

Lemma 2.3

$$\mathcal{A}\left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) = diag(I, tI, t^2I)\mathcal{A}(t)$$

and

$$\mathcal{B}\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) = diag(I, tI, t^2I)\mathcal{B}(t),$$

where $\mathcal{A}(t)$ and $\mathcal{B}(t)$ are defined as above.

Proof. As above. **QED**

Remark 2.13 *Since $\mathcal{A}svec(X_1(t)) + \mathcal{B}svec(Y_1(t)) = q$, we have*

$$\mathcal{A}\left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) svec(\widetilde{X}_1(t)) +$$
$$\mathcal{B}\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) svec(\widetilde{Y}_1(t)) = q.$$

Hence, by Lemma 2.3,

$$diag(I, tI, t^2 I)\left(\mathcal{A}(t)svec(\widetilde{X}_1(t)) + \mathcal{B}(t)svec(\widetilde{Y}_1(t))\right) = q.$$

This implies that q is equal to $(q_1^T, 0, 0)^T$ where $q_1 \in \Re^{i_1}$, which can be seen by letting t tends to zero in above. Therefore,

$$\mathcal{A}(t)svec(\widetilde{X}_1(t)) + \mathcal{B}(t)svec(\widetilde{Y}_1(t)) = \begin{pmatrix} q_1 \\ 0 \\ 0 \end{pmatrix}. \tag{2.43}$$

By (2.36), (2.37) and using the properties of $\otimes_s$, we have

$$(X_1 \otimes_s Y_1^{-1})\left(\begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right) = \left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right)(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})$$

Therefore, we can write $\mathcal{M}_1$ as

$$\mathcal{M}_1 = diag\left(I, tI, t^2 I, \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right)\begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}. \tag{2.44}$$

In a similar manner, we can express $\mathcal{M}_2$ in terms of $\mathcal{A}(t)$, $\mathcal{B}(t)$, $\widetilde{X}_1$ and $\widetilde{Y}_1$ as follows:

Using Lemma 2.3, we have

$$\mathcal{A}\left(\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right)\right) =$$
$$diag(I, tI, t^2 I)\mathcal{A}(t)\left(\left(\begin{pmatrix} I & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & \frac{1}{t}I \end{pmatrix}\right)\right)\left(\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right)\right) =$$
$$diag(I, tI, t^2 I)\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I\right).$$

Also,

$$\mathcal{B}\left(\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right)\right) = diag(I, tI, t^2 I)\mathcal{B}(t)\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right)$$

and

$$(X_1 \otimes_s Y_1^{-1})\left(\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} tI & 0 \\ 0 & I \end{pmatrix}\right)\right) =$$
$$\left(\left(\begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right)\right)(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \quad := \mathcal{M}_3.$$

where the last expression follows from (2.36), (2.37) and using the properties of $\otimes_s$.

We have

$$\mathcal{M}_2 = \left(\begin{array}{c|c} diag(I, tI, t^2 I)\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I\right) & diag(I, tI, t^2 I)\mathcal{B}(t)\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} & \mathcal{M}_3 \end{array}\right). \tag{2.45}$$

Substituting (2.44) and (2.45) into the system of ODEs (2.40) and simplifying the right hand side of its equality sign, we have

$$\frac{1}{2}diag\left(I, tI, t^2 I, \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix}\right) \begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix} \begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \mathcal{G} \begin{pmatrix} svec(\widetilde{X}_1) \\ svec(\widetilde{Y}_1) \end{pmatrix} \tag{2.46}$$

where

$$\mathcal{G} = \left(\begin{array}{c|c} -diag(I, tI, t^2 I)\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I\right) & -diag(I, tI, t^2 I)\mathcal{B}(t)\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s \begin{pmatrix} I & 0 \\ 0 & tI \end{pmatrix} & -\mathcal{M}_3 \end{array}\right).$$

Taking the inverse of the matrix on the extreme left in (2.46) and simplifying, we finally obtain

$$\frac{1}{2}\begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix} \begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \tag{2.47}$$

$$\begin{pmatrix} -\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(t)\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I & -(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \end{pmatrix} \times \begin{pmatrix} svec(\widetilde{X}_1) \\ svec(\widetilde{Y}_1) \end{pmatrix}.$$

Remark 2.14 *Instead of inverting $P \otimes_s P^{-T}Y$ in (2.8) to obtain the system of ODEs (2.38), we can also invert $PX \otimes_s P^{-T}$ in (2.8), when $P = Y^{1/2}$, to obtain*

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ (X \otimes_s Y^{-1})^{-1} & I \end{pmatrix} \begin{pmatrix} svec(X') \\ svec(Y') \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} 0 \\ svec(Y) \end{pmatrix}. \tag{2.48}$$

Proceeding in a similar manner as what was described above to obtain (2.47) from (2.38), we obtained from (2.48), the following system of ODEs,

$$\frac{1}{2}\begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ (\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})^{-1} & I \end{pmatrix}\begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \tag{2.49}$$

$$\begin{pmatrix} -\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(t)\left(\begin{pmatrix} \frac{1}{t}I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ -(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})^{-1}\left(\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I\right) & \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t}I \end{pmatrix} \otimes_s I \end{pmatrix} \times$$

$$\begin{pmatrix} svec(\widetilde{X}_1) \\ svec(\widetilde{Y}_1) \end{pmatrix}.$$

We have the following result by combining the systems of ODEs (2.47) and (2.49):

Proposition 2.10 *Given $(X(\mu), Y(\mu))$, $\mu > 0$, an off-central path of SDLCP (2.1) with $(X(1), Y(1)) = (X^0, Y^0)$. Let $X_1(t) = X(t^2)$ and $Y_1(t) = Y(t^2)$. Then $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ is a solution to the following system of ODEs*

$$\begin{pmatrix} \frac{1}{2}\mathcal{A}(t) & \frac{1}{2}\mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}\begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \tag{2.50}$$

$$\frac{1}{t}\begin{pmatrix} -\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(t)\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I & -(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) \end{pmatrix} \times$$

$$\begin{pmatrix} svec(\widetilde{X}_1) \\ svec(\widetilde{Y}_1) \end{pmatrix}.$$

Here $X(\mu), \widetilde{X}_1(t)$ and $Y(\mu), \widetilde{Y}_1(t)$ are related by (2.36) and (2.37) respectively where $\mu = t^2$.

Proof. Suppose $(X(\mu), Y(\mu))$ is an off-central path of SDLCP (2.1), then it is clear that $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$, $t > 0$, is a solution to the systems of ODE (2.47) and

(2.49). We have, from (2.49),

$$\tfrac{1}{2}(svec(\widetilde{X}_1') + (\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})svec(\widetilde{Y}_1')) =$$

$$\tfrac{1}{t}\left(-\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) svec(\widetilde{X}_1) + (\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) svec(\widetilde{Y}_1)\right).$$

By adding this to a similar equation from (2.47) and keeping the other half of the system of equalities in (2.47), (2.49) unchanged, we obtained the system of ODEs (2.50). Clearly, from the way (2.50) is obtained from (2.47), (2.49), we have $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ is also its solution. **QED**

Note that we will use (2.50) in the analysis that follows since it is more "symmetric" than (2.47) or (2.49).

We observe, in the following proposition, 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 equation (2.50).

Proposition 2.11 $\begin{pmatrix} \beta\mathcal{A}(t) & \beta\mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}$, *where* $\beta \neq 0$, $\beta \in \Re$, *is invertible for all* $t \geq 0$ *and* $\widetilde{X}_1$, $\widetilde{Y}_1$ *positive definite.*

Proof. To prove the proposition, it suffices to show that

$$\begin{pmatrix} \beta\mathcal{A}(t) & \beta\mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \Rightarrow \quad u = v = 0,$$

for $t \geq 0$ and $\widetilde{X}_1$, $\widetilde{Y}_1$ positive definite.

A sufficient condition for this to hold is to show that

$$\mathcal{A}(t)u + \mathcal{B}(t)v = 0 \quad \Rightarrow \quad u^T v \geq 0 \tag{2.51}$$

Now, for $t > 0$, (2.51) is true by Lemma 2.3 and since (2.42) holds.

Therefore, we need only show (2.51) for the case $t = 0$.

Suppose $\mathcal{A}(0)u + \mathcal{B}(0)v = 0$. We want to show that $u^T v \geq 0$ (The idea to prove this follows the proof of Theorem 3.13 in [17]).

Let $u = svec \begin{pmatrix} U_{11} & U_{12} \\ U_{12}^T & U_{22} \end{pmatrix}$ and $v = svec \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}$.

We have $\mathcal{A}\ svec \begin{pmatrix} U_{11} & 0 \\ 0 & 0 \end{pmatrix} + \mathcal{B}\ svec \begin{pmatrix} 0 & 0 \\ 0 & V_{22} \end{pmatrix} = 0$ since $\mathcal{A}(0)u + \mathcal{B}(0)v = 0$.

Also, $\mathcal{A}\ svec \begin{pmatrix} W_1 & Z_1 \\ Z_1^T & U_{22} \end{pmatrix} + \mathcal{B}\ svec \begin{pmatrix} V_{11} & Z_2 \\ Z_2^T & W_2 \end{pmatrix} = 0$ for some $W_1 \in S^m$, $W_2 \in S^{n-m}$ and $Z_1, Z_2 \in \Re^{m \times (n-m)}$. This is possible because $\mathcal{A}(0)u + \mathcal{B}(0)v = 0$ and by Remark 2.11.

Letting $X(s) = \begin{pmatrix} W_1 & Z_1 \\ Z_1^T & U_{22} \end{pmatrix} + s \begin{pmatrix} U_{11} & 0 \\ 0 & 0 \end{pmatrix}$ and $Y(s) = \begin{pmatrix} V_{11} & Z_2 \\ Z_2^T & W_2 \end{pmatrix} + s \begin{pmatrix} 0 & 0 \\ 0 & V_{22} \end{pmatrix}$, we have $\mathcal{A}\ svec(X(s)) + \mathcal{B}\ svec(Y(s)) = 0$ for all $s \in \Re$. Therefore, by (2.42), $X(s) \bullet Y(s) \geq 0$ for all $s \in \Re$. Expanding $X(s) \bullet Y(s)$, we have $W_1 \bullet V_{11} + U_{22} \bullet W_2 + 2Z_1 \bullet Z_2 + s(U_{11} \bullet V_{11} + U_{22} \bullet V_{22}) \geq 0$ for all $s \in \Re$. This must imply that $U_{11} \bullet V_{11} + U_{22} \bullet V_{22} = 0$.

We are done if we can show that $U_{12} \bullet V_{12} \geq 0$. This is true since there exist $W_3 \in S^m$ and $W_4 \in S^{n-m}$ such that $\mathcal{A}\ svec \begin{pmatrix} W_3 & U_{12} \\ U_{12}^T & 0 \end{pmatrix} + \mathcal{B}\ svec \begin{pmatrix} 0 & V_{12} \\ V_{12}^T & W_4 \end{pmatrix} = 0$ (the reason for this is because $\mathcal{A}(0)u + \mathcal{B}(0)v = 0$ and by Remark 2.11.) and then by (2.42).

Therefore, we have $u^T v = \begin{pmatrix} U_{11} & U_{12} \\ U_{12}^T & U_{22} \end{pmatrix} \bullet \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \geq 0$. **QED**

Note that the matrix $\begin{pmatrix} \frac{1}{2}\mathcal{A}(t) & \frac{1}{2}\mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}$ in (2.50) is invertible at any accumulation point of $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ (This follows from Proposition 2.11 since any accumulation point of $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ is positive definite, by Remark 2.10). This fact implies that we can invert the matrix at the limit as t tends to zero and this enables us to study the asymptotic behaviour of $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$.

Using (2.50), we can give a necessary and sufficient condition for $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ of an off-central path to be analytic at $t = 0$.

First, we have the following technical proposition:

Proposition 2.12 *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*

$$(\widetilde{Y}_1^*)_{12} = 0 \Longleftrightarrow$$

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

Proof. ($\Rightarrow$) Clear.

($\Leftarrow$) Suppose

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

Then we have

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

Now,

$$\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \widetilde{Y}_1^* \widetilde{X}_1^* \widetilde{Y}_1^* + \widetilde{Y}_1^* \widetilde{X}_1^* \widetilde{Y}_1^* \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right)_{11} =$$
$$2\Big((\widetilde{Y}_1^*)_{11} (\widetilde{X}_1^*)_{11} (\widetilde{Y}_1^*)_{11} + (\widetilde{Y}_1^*)_{12} (\widetilde{X}_1^*)_{12}^T (\widetilde{Y}_1^*)_{11} +$$
$$(\widetilde{Y}_1^*)_{11} (\widetilde{X}_1^*)_{12} (\widetilde{Y}_1^*)_{12}^T + (\widetilde{Y}_1^*)_{12} (\widetilde{X}_1^*)_{22} (\widetilde{Y}_1^*)_{12}^T \Big)$$

and

$$2\left(\widetilde{Y}_1^* \begin{pmatrix} (\widetilde{X}_1^*)_{11} & 0 \\ 0 & -(\widetilde{X}_1^*)_{22} \end{pmatrix} \widetilde{Y}_1^* \right)_{11} = 2(\widetilde{Y}_1^*)_{11} (\widetilde{X}_1^*)_{11} (\widetilde{Y}_1^*)_{11} - 2(\widetilde{Y}_1^*)_{12} (\widetilde{X}_1^*)_{22} (\widetilde{Y}_1^*)_{12}^T.$$

Equating them together, we have

$$(\widetilde{Y}_1^*)_{12}(\widetilde{X}_1^*)_{12}^T(\widetilde{Y}_1^*)_{11} + (\widetilde{Y}_1^*)_{11}(\widetilde{X}_1^*)_{12}(\widetilde{Y}_1^*)_{12}^T + 2(\widetilde{Y}_1^*)_{12}(\widetilde{X}_1^*)_{22}(\widetilde{Y}_1^*)_{12}^T = 0.$$

Hence,

$$(\widetilde{Y}_1^*)_{12}(\widetilde{X}_1^*)_{12}^T + (\widetilde{Y}_1^*)_{11}(\widetilde{X}_1^*)_{12}(\widetilde{Y}_1^*)_{12}^T(\widetilde{Y}_1^*)_{11}^{-1} = -2(\widetilde{Y}_1^*)_{12}(\widetilde{X}_1^*)_{22}(\widetilde{Y}_1^*)_{12}^T(\widetilde{Y}_1^*)_{11}^{-1}.$$

Therefore, $(\widetilde{Y}_1^*)_{12} \bullet (\widetilde{X}_1^*)_{12} = -Tr((\widetilde{Y}_1^*)_{12}(\widetilde{X}_1^*)_{22}(\widetilde{Y}_1^*)_{12}^T(\widetilde{Y}_1^*)_{11}^{-1}) \leq 0$.

On the other hand, consider $X_1(t)$ and $Y_1(t)$ where $X_1(t), \widetilde{X}_1(t)$ and $Y_1(t), \widetilde{Y}_1(t)$ are related by (2.36) and (2.37) respectively. Let $\{t_k\}$ be a sequence tending to zero such that $(X_1(t_k), Y_1(t_k))$ approaches (X^*, Y^*) and $(\widetilde{X}_1(t_k), \widetilde{Y}_1(t_k))$ approaches $(\widetilde{X}_1^*, \widetilde{Y}_1^*)$. Note that (X^*, Y^*) is a solution to SDLCP (2.1) (Hence $X^* \bullet Y^* = 0$). Also, $(X^*)_{11} = (\widetilde{X}_1^*)_{11}$ and $(Y^*)_{22} = (\widetilde{Y}_1^*)_{22}$.

Note also that since $(X_1(t_k), Y_1(t_k))$ and (X^*, Y^*) satisfy $\mathcal{A}(X) + \mathcal{B}(Y) = q$, we have, by Assumption 2.1(a), $(X_1(t_k) - X^*) \bullet (Y_1(t_k) - Y^*) \geq 0$.

Therefore, $X_1(t_k) \bullet Y_1(t_k) \geq X_1(t_k) \bullet Y^* + X^* \bullet Y_1(t_k)$, where we have used $X^* \bullet Y^* = 0$.

Note that $X_1(t_k) \bullet Y_1(t_k) = t_k^2 \widetilde{X}_1(t_k) \bullet \widetilde{Y}_1(t_k)$, $X_1(t_k) \bullet Y^* = t_k^2 (\widetilde{X}_1(t_k))_{22} \bullet (\widetilde{Y}_1^*)_{22}$ and $X^* \bullet Y_1(t_k) = t_k^2 (\widetilde{X}_1^*)_{11} \bullet (\widetilde{Y}_1(t_k))_{11}$ by (2.36), (2.37) and $(Y^*)_{22} = (\widetilde{Y}_1^*)_{22}$, $(X^*)_{11} = (\widetilde{X}_1^*)_{22}$. Hence $\widetilde{X}_1(t_k) \bullet \widetilde{Y}_1(t_k) \geq (\widetilde{X}_1(t_k))_{22} \bullet (\widetilde{Y}_1^*)_{22} + (\widetilde{X}_1^*)_{11} \bullet (\widetilde{Y}_1(t_k))_{11}$. Letting t_k tends to zero, we have $\widetilde{X}_1^* \bullet \widetilde{Y}_1^* \geq (\widetilde{X}_1^*)_{22} \bullet (\widetilde{Y}_1^*)_{22} + (\widetilde{X}_1^*)_{11} \bullet (\widetilde{Y}_1^*)_{11}$. Since $\widetilde{X}_1^* \bullet \widetilde{Y}_1^* = (\widetilde{X}_1^*)_{11} \bullet (\widetilde{Y}_1^*)_{11} + 2(\widetilde{X}_1^*)_{12} \bullet (\widetilde{Y}_1^*)_{12} + (\widetilde{X}_1^*)_{22} \bullet (\widetilde{Y}_1^*)_{22}$, we have $(\widetilde{X}_1^*)_{11} \bullet (\widetilde{Y}_1^*)_{11} + 2(\widetilde{X}_1^*)_{12} \bullet (\widetilde{Y}_1^*)_{12} + (\widetilde{X}_1^*)_{22} \bullet (\widetilde{Y}_1^*)_{22} \geq (\widetilde{X}_1^*)_{22} \bullet (\widetilde{Y}_1^*)_{22} + (\widetilde{X}_1^*)_{11} \bullet (\widetilde{Y}_1^*)_{11}$. This implies that $(\widetilde{X}_1^*)_{12} \bullet (\widetilde{Y}_1^*)_{12} \geq 0$.

Combining with $(\widetilde{Y}_1^*)_{12} \bullet (\widetilde{X}_1^*)_{12} \leq 0$ obtained earlier, we have $Tr((\widetilde{Y}_1^*)_{12}(\widetilde{X}_1^*)_{22} (\widetilde{Y}_1^*)_{12}^T(\widetilde{Y}_1^*)_{11}^{-1}) = 0$ which means that $(\widetilde{Y}_1^*)_{12} = 0$, since $(\widetilde{X}_1^*)_{22}$, $(\widetilde{Y}_1^*)_{11}$ are symmetric, positive definite. Hence we are done. **QED**

With this technical proposition, the following proposition follows almost immediately.

Proposition 2.13 *Let $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ be a solution to the system of ODEs (2.50) for $t > 0$. Suppose $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ converges as $t \longrightarrow 0$. Then* $\lim_{t\to 0}(\widetilde{Y}_1)_{12}(t) = 0$.

Proof. Suppose $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ converge as $t \longrightarrow 0$, then it is easy to see that $\widetilde{X}_1'(t), \widetilde{Y}_1'(t) = o\left(\frac{1}{t}\right)$. Therefore, if $\widetilde{X}_1(t) \longrightarrow \widetilde{X}_1^*, \widetilde{Y}_1(t) \longrightarrow \widetilde{Y}_1^*$ as $t \longrightarrow 0$, we must have

$$\begin{pmatrix} -\mathcal{A}(0)\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(0)\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I & -(\widetilde{X}_1^* \otimes_s (\widetilde{Y}_1^*)^{-1})\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) \end{pmatrix} \begin{pmatrix} svec(\widetilde{X}_1^*) \\ svec(\widetilde{Y}_1^*) \end{pmatrix}$$

is equal to zero. Therefore,

$$\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) svec(\widetilde{X}_1^*) - (\widetilde{X}_1^* \otimes_s (\widetilde{Y}_1^*)^{-1})\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) svec(\widetilde{Y}_1^*) = 0.$$

Using the properties of $\otimes_s$, we have

$$svec\begin{pmatrix} (\widetilde{X}_1^*)_{11} & 0 \\ 0 & -(\widetilde{X}_1^*)_{22} \end{pmatrix} - \frac{1}{2}\left(\left(\widetilde{X}_1^*\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right)\right) \otimes_s (\widetilde{Y}_1^*)^{-1} + \left((\widetilde{Y}_1^*)^{-1}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right)\right) \otimes_s \widetilde{X}_1^*\Bigg) svec(\widetilde{Y}_1^*) = 0,$$

which implies that

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

Hence $(\widetilde{Y}_1^*)_{12} = 0$, by Proposition 2.12. Therefore, $\lim_{t\to 0}(\widetilde{Y}_1)_{12}(t) = 0$. **QED**

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 2.5 *Let $(\widetilde{X}_1(t), \widetilde{Y}_1(t))$ be a solution to the system of ODEs (2.50) for $t > 0$. Then $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ are analytic at $t = 0$ if and only if $(\widetilde{Y}_1)_{12}(t)$ is analytic at $t = 0$ and $(\widetilde{Y}_1)_{12}(0) = 0$.*

Proof. ($\Rightarrow$) Suppose $\widetilde{X}_1(t)$ and $\widetilde{Y}_1(t)$ are analytic at $t = 0$. Then they converge to unique limit points as $t \longrightarrow 0$. Therefore, by Proposition 2.13, $(\widetilde{Y}_1)_{12}(0) = 0$. This, together with the analyticity of $(\widetilde{Y}_1)_{12}(t)$ at $t = 0$, implies our required result.

($\Leftarrow$) Suppose $(\widetilde{Y}_1)_{12}(t) = tW_1(t)$ for t (> 0) near 0, where $W_1(t)$ is analytic at $t = 0$.

From (2.50), we have

$$\begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \frac{\mathcal{F}_1(t, \widetilde{X}_1, \widetilde{Y}_1)}{t},$$

where

$$\mathcal{F}_1(t, \widetilde{X}_1, \widetilde{Y}_1) = \begin{pmatrix} \frac{1}{2}\mathcal{A}(t) & \frac{1}{2}\mathcal{B}(t) \\ I & \widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1} \end{pmatrix}^{-1} \times$$

$$\begin{pmatrix} -\mathcal{A}(t)\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(t)\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I & -(\widetilde{X}_1 \otimes_s \widetilde{Y}_1^{-1})\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) \end{pmatrix} \begin{pmatrix} svec(\widetilde{X}_1) \\ svec(\widetilde{Y}_1) \end{pmatrix}.$$

Note that $\mathcal{F}_1(t, \widetilde{X}_1, \widetilde{Y}_1)$ is analytic at $(0, \widetilde{X}_1, \widetilde{Y}_1)$, where $\widetilde{X}_1, \widetilde{Y}_1 \succ 0$. Therefore, we can write $\mathcal{F}_1(t, \widetilde{X}_1, \widetilde{Y}_1)$ as $\sum_{n=0}^{\infty} a_n(\widetilde{X}_1, \widetilde{Y}_1)t^n$ where a_n is analytic at $(\widetilde{X}_1, \widetilde{Y}_1)$, $\widetilde{X}_1, \widetilde{Y}_1 \succ 0$, for all $n \geq 0$.

Now, $\mathcal{F}_1(0, \widetilde{X}_1, \widetilde{Y}_1) = a_0(\widetilde{X}_1, \widetilde{Y}_1)$.

We want to show that $\mathcal{F}_1(0, \widetilde{X}_1(t), \widetilde{Y}_1(t)) = a_0(\widetilde{X}_1(t), \widetilde{Y}_1(t)) = t\,\widetilde{a}_0(t, \widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$, where $\widetilde{a}_0$ as a function of $(t, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$ is analytic at $(0, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$ where $\widetilde{X}_1$, $(\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22} \succ 0$.

We have

$$a_0(\widetilde{X}_1(t), \widetilde{Y}_1(t)) = \mathcal{F}_1(0, \widetilde{X}_1(t), \widetilde{Y}_1(t)) = \begin{pmatrix} \frac{1}{2}\mathcal{A}(0) & \frac{1}{2}\mathcal{B}(0) \\ I & \widetilde{X}_1(t) \otimes_s \widetilde{Y}_1^{-1}(t) \end{pmatrix}^{-1} \times$$

$$\begin{pmatrix} -\mathcal{A}(0)\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(0)\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I & -(\widetilde{X}_1(t) \otimes_s \widetilde{Y}_1^{-1}(t))\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) \end{pmatrix} \times$$

$$\begin{pmatrix} svec(\widetilde{X}_1(t)) \\ svec(\widetilde{Y}_1(t)) \end{pmatrix}.$$

Now, $\begin{pmatrix} \frac{1}{2}\mathcal{A}(0) & \frac{1}{2}\mathcal{B}(0) \\ I & \widetilde{X}_1(t) \otimes_s \widetilde{Y}_1^{-1}(t) \end{pmatrix}^{-1}$ is equal to $\widetilde{B}_0(t, \widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$, where $\widetilde{B}_0$ as a function of $(t, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$ is analytic at $(0, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$, with $\widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22} \succ 0$, since $(\widetilde{Y}_1)_{12}(t)$ is analytic at $t = 0$ and $(\widetilde{Y}_1)_{12}(0) = 0$. Next, let us consider

$$\begin{pmatrix} -\mathcal{A}(0)\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) & -\mathcal{B}(0)\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) \\ \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I & -(\widetilde{X}_1(t) \otimes_s \widetilde{Y}_1^{-1}(t))\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\right) \end{pmatrix} \times$$

$$\begin{pmatrix} svec(\widetilde{X}_1(t)) \\ svec(\widetilde{Y}_1(t)) \end{pmatrix}.$$

Let

$$c(t) := -\mathcal{A}(0)\left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \otimes_s I\right) svec(\widetilde{X}_1(t)) - \mathcal{B}(0)\left(\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \otimes_s I\right) svec(\widetilde{Y}_1(t)).$$

Therefore,

$$c(t) = -\mathcal{A}(0) svec \begin{pmatrix} 0 & (\widetilde{X}_1)_{12}(t) \\ (\widetilde{X}_1)_{12}^T(t) & 2(\widetilde{X}_1)_{22}(t) \end{pmatrix} - \mathcal{B}(0) svec \begin{pmatrix} 2(\widetilde{Y}_1)_{11}(t) & (\widetilde{Y}_1)_{12}(t) \\ (\widetilde{Y}_1)_{12}^T(t) & 0 \end{pmatrix}.$$

By definition of $\mathcal{A}(0)$ and $\mathcal{B}(0)$, we have for $i = 1, \ldots, i_1$, $c(t)_i = 0$. Also, for $i = i_1 + i_2 + 1, \ldots, \tilde{n}$, because $\mathcal{A}(t)svec(\widetilde{X}_1(t)) + \mathcal{B}(t)svec(\widetilde{Y}_1(t)) = \begin{pmatrix} q_1 \\ 0 \\ 0 \end{pmatrix}$ (Remark 2.13) , we have $c(t)_i = 0$.

Now, for $i = i_1+1, \ldots, i_1+i_2$, using $\mathcal{A}(t)svec(\widetilde{X}_1(t))+\mathcal{B}(t)svec(\widetilde{Y}_1(t)) = \begin{pmatrix} q_1 \\ 0 \\ 0 \end{pmatrix}$ again, we have

$$svec\begin{pmatrix} 0 & (A_i)_{12} \\ (A_i)_{12}^T & t(A_i)_{22} \end{pmatrix}^T svec(\widetilde{X}_1(t))+svec\begin{pmatrix} t(B_i)_{11} & (B_i)_{12} \\ (B_i)_{12}^T & 0 \end{pmatrix}^T svec(\widetilde{Y}_1(t)) = 0.$$

Therefore,

$$\begin{aligned} svec\begin{pmatrix} 0 & (A_i)_{12} \\ (A_i)_{12}^T & 0 \end{pmatrix}^T svec\begin{pmatrix} 0 & (\widetilde{X}_1)_{12}(t) \\ (\widetilde{X}_1)_{12}^T(t) & 2(\widetilde{X}_1)_{22}(t) \end{pmatrix} + \\ t\ svec\begin{pmatrix} 0 & 0 \\ 0 & (A_i)_{22} \end{pmatrix}^T svec\begin{pmatrix} 0 & 0 \\ 0 & (\widetilde{X}_1)_{22}(t) \end{pmatrix} + \\ svec\begin{pmatrix} 0 & (B_i)_{12} \\ (B_i)_{12}^T & 0 \end{pmatrix}^T svec\begin{pmatrix} 2(\widetilde{Y}_1)_{22}(t) & (\widetilde{Y}_1)_{12}(t) \\ (\widetilde{Y}_1)_{12}^T(t) & 0 \end{pmatrix} + \\ t\ svec\begin{pmatrix} (B_i)_{11} & 0 \\ 0 & 0 \end{pmatrix}^T svec\begin{pmatrix} (\widetilde{Y}_1)_{11}(t) & 0 \\ 0 & 0 \end{pmatrix} = 0. \end{aligned}$$

Hence, $c(t)_i = t\, \widetilde{c}_0^i(t, \widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$ where $\widetilde{c}_0^i$ as a function of $(t, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$ is analytic at $(0, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$, with $\widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22} \succ 0$, for $i = i_1 + 1, \ldots, i_1 + i_2$.

Consider

$$\begin{aligned} smat\Bigg(\Bigg(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\Bigg) svec(\widetilde{X}_1(t)) - \\ (\widetilde{X}_1(t) \otimes_s \widetilde{Y}_1^{-1}(t)) \Bigg(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \otimes_s I\Bigg) svec(\widetilde{Y}_1(t))\Bigg) \end{aligned}$$

which is equal to

$$\begin{pmatrix} (\widetilde{X}_1)_{11}(t) & 0 \\ 0 & -(\widetilde{X}_1)_{22}(t) \end{pmatrix} - \frac{1}{4}\left(\left(\begin{pmatrix} 2(\widetilde{X}_1)_{11}(t) & 0 \\ 0 & -2(\widetilde{X}_1)_{22}(t) \end{pmatrix} + \right.\right.$$
$$\left.\widetilde{Y}_1^{-1}(t)\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\widetilde{Y}_1(t)\widetilde{X}_1(t) + \widetilde{X}_1(t)\widetilde{Y}_1(t)\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\widetilde{Y}_1^{-1}(t)\right).$$

Let

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

We have

$$\widetilde{Y}_1(t)D(t)\widetilde{Y}_1(t) = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\widetilde{Y}_1(t)\widetilde{X}_1(t)\widetilde{Y}_1(t) +$$
$$\widetilde{Y}_1(t)\widetilde{X}_1(t)\widetilde{Y}_1(t)\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} - 2\widetilde{Y}_1(t)\begin{pmatrix} (\widetilde{X}_1)_{11}(t) & 0 \\ 0 & -(\widetilde{X}_1)_{22}(t) \end{pmatrix}\widetilde{Y}_1(t).$$

Let $\widetilde{Y}_1(t) = \widehat{Y}_1(t) + \bar{Y}_1(t)$ where

$$\widehat{Y}_1(t) = \begin{pmatrix} (\widetilde{Y}_1)_{11}(t) & 0 \\ 0 & (\widetilde{Y}_1)_{22}(t) \end{pmatrix}, \ \bar{Y}_1(t) = \begin{pmatrix} 0 & (\widetilde{Y}_1)_{12}(t) \\ (\widetilde{Y}_1)_{12}^T(t) & 0 \end{pmatrix}.$$

Then, noting that

$$\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\widehat{Y}_1(t)\widetilde{X}_1(t)\widehat{Y}_1(t) + \widehat{Y}_1(t)\widetilde{X}_1(t)\widehat{Y}_1(t)\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} =$$
$$2\widehat{Y}_1(t)\begin{pmatrix} (\widetilde{X}_1)_{11}(t) & 0 \\ 0 & -(\widetilde{X}_1)_{22}(t) \end{pmatrix}\widehat{Y}_1(t),$$

we observe that every term in the above expression for $\widetilde{Y}_1(t)D(t)\widetilde{Y}_1(t)$ involves at least a $(\widetilde{Y}_1)_{12}(t)$. Therefore, with $(\widetilde{Y}_1)_{12}(t) = tW_1(t)$ for t (> 0) near 0 and $W_1(t)$ analytic at $t = 0$, we have $D(t) = t\ \widetilde{D}_0(t, \widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$, where

$\widetilde{D}_0$ as a function of $(t, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$ is analytic at $(0, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$, with $\widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22} \succ 0$.

Hence, $a_0(\widetilde{X}_1(t), \widetilde{Y}_1(t)) = t\ \widetilde{a}_0(t, \widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$, where $\widetilde{a}_0$ as a function of $(t, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$ is analytic at $(0, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$, with $\widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22} \succ 0$, is true.

Therefore, we have $(t, \widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$, for t (> 0) near 0, satisfies the following system of ODEs,

$$\begin{pmatrix} svec(\widetilde{X}_1') \\ svec(\widetilde{Y}_1') \end{pmatrix} = \widetilde{a}_0(t, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22}) + \sum_{n=1}^{\infty} a_n((\widetilde{Y}_1)_{12}(t), \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})\, t^{n-1},$$

where its right-hand side is analytic at $(0, \widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22})$, where $\widetilde{X}_1, (\widetilde{Y}_1)_{11}, (\widetilde{Y}_1)_{22} \succ 0$.

Therefore, from Theorem 4.1 of [4], pp. 15 and Theorem 2.1 above, we have $(\widetilde{X}_1(t), (\widetilde{Y}_1)_{11}(t), (\widetilde{Y}_1)_{22}(t))$ is analytic at $t = 0$, which together with the analyticity of $(\widetilde{Y}_1)_{12}(t)$ at $t = 0$, implies our required result. **QED**

Using Theorem 2.5, we end this section by giving a necessary and sufficient condition for $(X(\mu), Y(\mu))$ to be analytic with respect to $t = \sqrt{\mu}$ at the limit point when μ tends to zero.

Theorem 2.6 *Let $(X(\mu), Y(\mu))$ be an off central path of SDLCP (2.1) for $\mu > 0$. Then $X(\mu), Y(\mu)$ are analytic with respect to $t = \sqrt{\mu}$ at the limit point as $\mu \to 0$ if and only if $Y_{12}(\mu) = \mu W(\mu)$, where $W(\mu)$ is analytic with respect to $t = \sqrt{\mu}$ at the limit point as $\mu \to 0$.*

Proof. Using (2.36), (2.37) and Theorem 2.5. **QED**

Chapter 3

Analysis of Off-Central Paths for SOCP

In this chapter, we consider off-central paths for second order cone programming (SOCP). We consider the general case of multiple cone SOCP in most of our discussions here. However, in the last part of the chapter, we will show asymptotic analyticity of off-central path for the AHO direction only for single cone SOCP. Although this is not interesting, since a closed form formula for the primal and dual optimal solution for single cone SOCP is already known (see [1]), we still state and prove the result here for the sake of completeness and also because the result and its proof may shed some light in showing asymptotic analyticity of off-central path for multiple cone SOCP, which is still an open question.

We first define off-central path for SOCP for a general direction. Then we restrict our attention to the AHO direction. We show the existence of off-central paths for possibly the largest domain for this direction. This is done in Section 3.1. Next, in Section 3.2, we provide a condition when off-central path for SOCP defined by the AHO direction will converge to a strictly complementary optimal solution. Finally, we prove asymptotic analyticity of off-central path for the 1-cone SOCP defined by the AHO direction.

As far as we know, off-central path for SOCP defined by the AHO direction has not been discussed in the literature and our discussions here make contributions to this area.

3.1 Off-Central Path for SOCP

In this section, we define a direction field associated with the predictor-corrector algorithm for second order cone programming (SOCP). This gives rise to a system of ordinary differential equations (ODEs) whose solution is the off-central path for SOCP.

Consider the following primal program $(\mathcal{P})$

$$(\mathcal{P})\qquad \begin{array}{ll} \min & \sum_{i=1}^{N} c_i^T x_i \\ \text{subject to} & \sum_{i=1}^{N} A_i x_i = b \\ & \|\overline{x_i}\| \leq (x_i)_0 \qquad i = 1, \ldots, N \end{array}$$

Here $x_i = ((x_i)_0, \overline{x_i}^T)^T \in \Re^{k_i+1}$ and $A_i \in \Re^{m \times (k_i+1)}$.

Its dual program $(\mathcal{D})$ is

$$(\mathcal{D})\qquad \begin{array}{lll} \max & b^T y & \\ \text{subject to} & A_i^T y + s_i = c_i & i = 1, \ldots, N \\ & \|\overline{s_i}\| \leq (s_i)_0 & i = 1, \ldots, N \end{array}$$

where $y \in \Re^m$ and $s_i = ((s_i)_0, \overline{s_i}^T)^T \in \Re^{k_i+1}$.

$(\mathcal{P})$-$(\mathcal{D})$ together formed a SOCP.

We have the following standard assumptions on $(\mathcal{P})$-$(\mathcal{D})$:

Assumption 3.1

(a) There exists a strictly feasible solution to $(\mathcal{P})$ *and* $(\mathcal{D})$*.*

(b) $(A_1, \ldots, A_N)$ *has full row rank.*

Under Assumption 3.1, it is well known that there exists an optimal solution to $(\mathcal{P})$-$(\mathcal{D})$, the optimal solution set is bounded and $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*)$ is an optimal solution to $(\mathcal{P})$-$(\mathcal{D})$ if and only if

$$\begin{array}{c} Arw(x_i^*)s_i^* = 0 \quad \text{for } i = 1, \ldots, N \\ \sum_{i=1}^N A_i x_i^* = b \\ A_i^T y^* + s_i^* = c_i \quad \text{for } i = 1, \ldots, N \\ \|\overline{x_i^*}\| \leq (x_i^*)_0, \|\overline{s_i^*}\| \leq (s_i^*)_0 \quad \text{for } i = 1, \ldots, N. \end{array} \tag{3.1}$$

Here, $Arw(u) := \begin{pmatrix} u_0 & \overline{u}^T \\ \overline{u} & u_0 I \end{pmatrix}$ where $u = (u_0, \overline{u}^T)^T \in \Re^{k+1}$.

We will now define an off-central path for SOCP using the system of equations (3.1). As in the case of SDLCP - in Section 2.1 - we consider the predictor step of the predictor-corrector path-following algorithm for SOCP, which is based on the linearization of (3.1), to define the off-central path. We will use the MZ-type family of directions on the predictor step to define the path.

Define $\mathcal{G}_i := \{\lambda \widetilde{T}_i : \lambda > 0, \widetilde{T}_i \in \Re^{(k_i+1)\times(k_i+1)}, \widetilde{T}_i^T J_{k_i} \widetilde{T}_i = J_{k_i}, (\widetilde{T}_i)_{11} > 0\}$, where $J_{k_i} = \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} \in \Re^{(k_i+1)\times(k_i+1)}$.

It is well-known that $\mathcal{G}_i$ is exactly the automorphism group of the second order cone $\mathcal{K}_i := \{x = (x_0, \overline{x}^T)^T \in \Re^{k_i+1} \ : \ \|\overline{x}\| \leq x_0\}$, namely, the set of all nonsingular matrices G_i such that $\mathcal{K}_i = G_i(\mathcal{K}_i)$.

For different $G_i \in \mathcal{G}_i$ applied to the predictor step of the predictor-corrector path-following algorithm, we obtain the AHO direction, the HKM directions and the NT direction. For more details on the MZ-type family of directions for SOCP, please refer to [14].

As in the case of SDLCP in Section 2.1, starting from the system of equations defining the predictor step of the predictor-corrector algorithm, we obtain the following system of ODEs for SOCP:

$$\begin{aligned} Arw(G_i^{-1}s_i)G_i^T x_i' + Arw(G_i^T x_i)G_i^{-1}s_i' &= -Arw(G_i^T x_i)G_i^{-1}s_i \quad \text{for } i = 1, \ldots, N \\ \textstyle\sum_{i=1}^N A_i x_i' &= 0 \\ A_i^T y' + s_i' &= 0 \quad \text{for } i = 1, \ldots, N. \end{aligned} \tag{3.2}$$

where $G_i \in \mathcal{G}_i$ and the initial point $(x_1(0), \ldots, x_N(0),\ y(0), s_1(0), \ldots, s_N(0))$ of the ODE system (3.2) satisfies the primal and dual feasibility conditions.

The solution to the system of ODEs (3.2) is called the **off-central path** for SOCP.

In the discussions that follow, we concentrate on the case when $G_i = I$ for all $i = 1, \ldots, N$, which corresponds to the AHO direction. In this case, (3.2) becomes

$$\begin{aligned} Arw(x_i)s_i' + Arw(s_i)x_i' &= -Arw(x_i)s_i \quad \text{for } i = 1, \ldots, N \\ \textstyle\sum_{i=1}^N A_i x_i' &= 0 \\ A_i^T y' + s_i' &= 0 \quad \text{for } i = 1, \ldots, N \end{aligned} \tag{3.3}$$

with the initial point $(x_1(0), \ldots, x_N(0), y(0), s_1(0), \ldots, s_N(0))$ of the ODE system (3.3) satisfying the primal and dual feasibility conditions.

Now the first equation in (3.3) can be written as $(Arw(x_i)s_i)' = -Arw(x_i)s_i$. Letting $z_i(t) = Arw(x_i)s_i$. We have $z_i'(t) = -z_i(t)$, from $(Arw(x_i)s_i)' = -Arw(x_i)s_i$, which implies that $z_i(t) = e^{-t}m_i$ for some $m_i \in \Re^{k_i+1}$.

By applying an appropriate change of variable from t to μ (to be precise, letting $\mu = \exp(-t)$), we have that any solution to (3.3) must satisfy the following algebraic system of equations:

$$\begin{aligned} Arw(x_i(\mu))s_i(\mu) &= \mu m_i \quad \text{for } i = 1, \ldots, N \\ \textstyle\sum_{i=1}^N A_i x_i(\mu) &= b \\ A_i^T y(\mu) + s_i(\mu) &= c_i \quad \text{for } i = 1, \ldots, N \end{aligned} \tag{3.4}$$

Remark 3.1 *Note that if $m_i = (1, 0, \ldots, 0)^T$, then (3.4), together with $\|\overline{x_i}(\mu)\| < (x_i)_0(\mu)$ and $\|\overline{s_i}(\mu)\| < (s_i)_0(\mu)$ for $i = 1, \ldots, N$, give rise to the equations defining the central path for SOCP for $\mu > 0$. Therefore, to be consistent with the way the central path is related to (3.4), we require that $\sum_{i=1}^N (m_i)_0 = N$ in (3.4).*

To be meaningful, we also require that an off-central path stays in the interior of the second order cones for all $\mu > 0$. Therefore, besides satisfying (3.4), an off-central path, $(x_1(\mu), \ldots, x_N(\mu), y(\mu), s_1(\mu), \ldots, s_N(\mu))$, must also satisfy

$$\|\overline{x_i}(\mu)\| < (x_i)_0(\mu), \quad \|\overline{s_i}(\mu)\| < (s_i)_0(\mu) \;\text{ for } i = 1, \ldots, N \tag{3.5}$$

for $\mu > 0$.

A question therefore arises as to under what conditions does such off-central path exists for all $\mu > 0$.

In the following, we give an example in which an off-central path does not exist for all $\mu > 0$. In particular, for this example, it does not satisfy (3.5) for all $\mu > 0$.

Example 3.1 *Consider* $N = 1$. *Let* $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *and let* $x^0 = \left(1 \; \frac{\sqrt{3}}{2} \; \frac{\sqrt{3}}{6}\right)^T$, $s^0 = \left(1 \; \frac{1}{2} \; \frac{1}{2}\right)^T$, $y^0 = (1 \; 1)^T$ *be a point that satisfied (3.4) when* $\mu = 1$. *Note that* x^0 *and* s^0 *both lie in the interior of the second order cone. The solution to (3.4) with these initial conditions is*

$$s(\mu) = \begin{pmatrix} s_0(\mu) \\ s_1(\mu) \\ s_2(\mu) \end{pmatrix} = \begin{pmatrix} \left(\frac{1}{2} + \frac{3\sqrt{3}}{22}\right)\mu + \frac{3}{11}\sqrt{\left(\frac{11}{6} + \frac{\sqrt{3}}{2}\right)^2 \mu^2 - \frac{11}{3}(1+\sqrt{3})\mu + \frac{11}{3}} \\ \frac{1}{2} \\ \left(\frac{19}{44} + \frac{\sqrt{3}}{12}\right)\mu - \frac{\sqrt{3}}{22}\sqrt{\left(\frac{11}{6} + \frac{\sqrt{3}}{2}\right)^2 \mu^2 - \frac{11}{3}(1+\sqrt{3})\mu + \frac{11}{3}} \end{pmatrix},$$

$$x(\mu) = \begin{pmatrix} 1 \\ \frac{1}{s_0(\mu)}\left[\left(\frac{1}{2} + \frac{\sqrt{3}}{2}\right)\mu - \frac{1}{2}\right] \\ \frac{\sqrt{3}}{6} \end{pmatrix},$$

$$y(\mu) = \begin{pmatrix} 2 - s_0(\mu) \\ \frac{3}{2} - s_2(\mu) \end{pmatrix}.$$

When $\mu = \frac{1}{\sqrt{3}}$, $s(\mu) = \begin{pmatrix} \frac{\sqrt{3}}{3} & \frac{1}{2} & \frac{\sqrt{3}}{6} \end{pmatrix}^T$ *and we have* $s_0(\mu) = \|\bar{s}(\mu)\|$. *Therefore,* $s(\frac{1}{\sqrt{3}})$ *no longer lies in the interior of the second order cone. Hence, for this example, it does not satisfy (3.5) for all* $\mu > 0$.

We observe that a property of the above example is that $Arw(x^0)s^0$ does not lie in the interior of the second order cone (in fact, it lies on the boundary of the second order cone).

We may ask whether if $Arw(x^0)s^0$ lies in the interior of the second order cone, then the off-central path exists for all $\mu > 0$ with $x(1) = x^0$ and $s(1) = s^0$? The answer is affirmative.

We have the following theorem:

Theorem 3.1 *Let* $(x_1^0, \ldots, x_N^0, y^0, s_1^0, \ldots, s_N^0)$ *satisfies (3.4) and (3.5) with* $\mu = 1$. *Suppose that for all* $i = 1, \ldots, N$, $Arw(x_i^0)s_i^0$ *lies in the interior of each of its second order cone, then there exists an unique analytic solution* $(x_1(\mu), \ldots, x_N(\mu), y(\mu), s_1(\mu), \ldots, s_N(\mu))$ *to (3.4) and (3.5) for* $\mu > 0$ *such that* $(x_1(1), \ldots, x_N(1), y(1), s_1(1), \ldots, s_N(1)) = (x_1^0, \ldots, x_N^0, y^0, s_1^0, \ldots, s_N^0)$.

Note that in Thereom 3.1, we show for the first time the existence of off-central path defined using the AHO direction, for, arguably, the largest domain possible. This domain is analogous to the domain for the existence of AHO search direction for SDP, as shown in [20].

In order to prove Theorem 3.1, we need to use the following two lemmas:

Lemma 3.1 *Suppose* X, S *are symmetric matrices with* S *invertible. If* $XS+SX$ *is positive definite, then* $WS^{-1}XW^T$ *is invertible, where* W *has full row rank.*

Proof. Let $XS + SX = C$. Then $WS^{-1}XW^T + WXS^{-1}W^T = WS^{-1}CS^{-1}W^T$.
Let $v \in$ Ker$(WS^{-1}XW^T)$. Then $WXS^{-1}W^Tv = WS^{-1}CS^{-1}W^Tv$.
Therefore, $WS^{-1}CS^{-1}W^Tv \in$ Range$(WXS^{-1}W^T)$. But note that

$$\text{Range}(WXS^{-1}W^T) = \text{Ker}((WXS^{-1}W^T)^T)^\perp = \text{Ker}(WS^{-1}XW^T)^\perp$$

Therefore, we must have $v^T W S^{-1} C S^{-1} W^T v = 0$. Since C is a symmetric positive definite matrix, we have $S^{-1} W^T v = 0$. This implies that $v = 0$ because W has full row rank. Hence, $WS^{-1}XW^T$ is invertible and the lemma is proved. **QED**

Lemma 3.2 *If $x, s, Arw(x)s$ are all in the interior of a second order cone, that is, $x, s, Arw(x)s \in \text{Int}(K)$ where K is a second order cone, then $Arw(x)Arw(s) + Arw(s)Arw(x)$ is positive definite.*

Proof. Suppose $x, s, Arw(x)s \in \text{Int}(K)$.
Let $x = (x_0, \overline{x}^T)^T$, $s = (s_0, \overline{s}^T)^T$. Then

$$\tfrac{1}{2}(Arw(x)Arw(s) + Arw(s)Arw(x))$$

$$= \begin{pmatrix} x^T s & x_0\overline{s}^T + s_0\overline{x}^T \\ x_0\overline{s} + s_0\overline{x} & \tfrac{1}{2}(\overline{x}\overline{s}^T + \overline{s}\overline{x}^T) + x_0 s_0 I \end{pmatrix}$$

It is easy to see that $Arw(x)Arw(s) + Arw(s)Arw(x)$ is positive definite if and only if

$$\begin{aligned} &\tfrac{1}{2}(\overline{x}\overline{s}^T + \overline{s}\overline{x}^T) + x_0 s_0 I - \tfrac{1}{x^T s}(x_0\overline{s} + s_0\overline{x})(x_0\overline{s}^T + s_0\overline{x}^T) \\ &= (\tfrac{1}{2} - \tfrac{x_0 s_0}{x^T s})(\overline{x}\overline{s}^T + \overline{s}\overline{x}^T) + x_0 s_0 I - \tfrac{x_0^2}{x^T s}\overline{s}\overline{s}^T - \tfrac{s_0^2}{x^T s}\overline{x}\overline{x}^T \end{aligned} \tag{3.6}$$

is positive definite.
Therefore, to prove the lemma, it suffices to show that (3.6) is positive definite.
We need to show that $\min_{\|v\|=1} v^T[(\frac{1}{2} - \frac{x_0 s_0}{x^T s})(\overline{x}\overline{s}^T + \overline{s}\overline{x}^T) + x_0 s_0 I - \frac{x_0^2}{x^T s}\overline{s}\overline{s}^T - \frac{s_0^2}{x^T s}\overline{x}\overline{x}^T]v > 0$.
Now,

$$\begin{aligned} &v^T[(\tfrac{1}{2} - \tfrac{x_0 s_0}{x^T s})(\overline{x}\overline{s}^T + \overline{s}\overline{x}^T) + x_0 s_0 I - \tfrac{x_0^2}{x^T s}\overline{s}\overline{s}^T - \tfrac{s_0^2}{x^T s}\overline{x}\overline{x}^T]v \\ &\quad = (v^T\overline{x})(v^T\overline{s}) + x_0 s_0\|v\|^2 - \tfrac{1}{x^T s}(x_0 v^T\overline{s} + s_0 v^T\overline{x})^2. \end{aligned}$$

Therefore, to prove the lemma, it is enough to show that

$$\min_{\|v\|=1} x^T s(v^T\overline{x})(v^T\overline{s}) - (x_0 v^T\overline{s} + s_0 v^T\overline{x})^2 > -(x_0 s_0)x^T s,$$

given $x, s, Arw(x)s \in \text{Int}(K)$.

However,

$$
\begin{aligned}
&\min_{\|v\|=1} x^T s(v^T\overline{x})(v^T\overline{s}) - (x_0 v^T\overline{s} + s_0 v^T\overline{x})^2 \\
&\geq \min\{-x_0^2\beta^2 - s_0^2\alpha^2 + (x^T s - 2x_0 s_0)\alpha\beta \;\;;\;\; |\alpha| \leq \|\overline{x}\|, |\beta| \leq \|\overline{s}\|, \\
&\qquad |x_0\beta + s_0\alpha| \leq \|x_0\overline{s} + s_0\overline{x}\|\}
\end{aligned}
$$

Given that $x, s, Arw(x)s \in \text{Int}(K)$, the possible "optimal value" candidates to the latter minimization problem are

$$
\begin{aligned}
&-x^T s\|\overline{x}\|\|\overline{s}\| - (x_0\|\overline{s}\| - s_0\|\overline{x}\|)^2, \\
&-\tfrac{s_0}{x_0}x^T s\|\overline{x}\|^2 + \tfrac{x^T s}{x_0}\|\overline{x}\|\|x_0\overline{s} + s_0\overline{x}\| - \|x_0\overline{s} + s_0\overline{x}\|^2 \text{ and} \\
&-\tfrac{x_0}{s_0}x^T s\|\overline{s}\|^2 + \tfrac{x^T s}{s_0}\|\overline{s}\|\|x_0\overline{s} + s_0\overline{x}\| - \|x_0\overline{s} + s_0\overline{x}\|^2.
\end{aligned}
$$

In all three cases, it can be shown easily that they are all greater than $-(x_0 s_0)x^T s$. Hence the lemma is proved. **QED**

Proof of Theorem 3.1. To show this, we basically rely on the Implicit Function Theorem.

First define a set $\mathcal{O}$ by

$$
\begin{aligned}
\mathcal{O} \;=\; &\{(x_1, \ldots, x_N, y, s_1, \ldots, s_N) \;:\; y \in \Re^m, (x_i, s_i) \in \text{Int}(K_i) \times \text{Int}(K_i), \\
&Arw(x_i)s_i \in \text{Int}(K_i) \text{ for } i = 1, \ldots, N\}
\end{aligned}
$$

where $K_i := \{x_i = ((x_i)_0, \overline{x_i}^T)^T \in \Re^{k_i+1} \;:\; \|\overline{x_i}\| \leq (x_i)_0\}$, which is a second order cone.

We consider the map $\Phi : \mathcal{O} \times \Re_{++} \longmapsto \Re^m \times (\Re^{k_1+1} \times \ldots \times \Re^{k_N+1}) \times (\Re^{k_1+1} \times \ldots \times \Re^{k_N+1})$ defined by

$$\Phi(x_1, \ldots, x_N, y, s_1, \ldots, s_N, \mu) := \begin{pmatrix} \sum_{i=1}^N A_i x_i - b \\ A_1^T y + s_1 - c_1 \\ \vdots \\ A_N^T y + s_N - c_N \\ Arw(x_1) s_1 - \mu Arw(x_1^0) s_1^0 \\ \vdots \\ Arw(x_N) s_N - \mu Arw(x_N^0) s_N^0 \end{pmatrix}.$$

Φ is clearly an analytic map and if we can show that for every $(x_1, \ldots, x_N, y, s_1, \ldots, s_N, \mu) \in \mathcal{O} \times \Re_{++}$ such that $\Phi(x_1, \ldots, x_N, y, s_1, \ldots, s_N, \mu) = 0$, $D_z\Phi(z, \mu)$, where $z = (x_1, \ldots, x_N, y, s_1, \ldots, s_N)$, is nonsingular, then the theorem is proved by applying the Implicit Function Theorem and a continuity argument.

Now,

$$D_z\Phi(z, \mu) = \begin{pmatrix} W & 0 & 0 \\ 0 & W^T & I \\ S & 0 & X \end{pmatrix}$$

where $W = (A_1 \ \ldots \ A_N)$, $X = diag(Arw(x_1), \ldots, Arw(x_N))$ and $S = diag(Arw(s_1), \ldots, Arw(s_N))$.

Consider $\begin{pmatrix} W & 0 & 0 \\ 0 & W^T & I \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta X \\ \Delta y \\ \Delta S \end{pmatrix} = 0.$

If we can show that $\Delta X, \Delta y$ and ΔS are equal to zero, then we are done.

We have $WS^{-1}XW^T\Delta y = 0$, $\Delta S = -W^T\Delta y$ and $\Delta X = S^{-1}XW^T\Delta y$. Therefore, if $WS^{-1}XW^T$ is invertible, then $(\Delta X, \Delta y, \Delta S) = 0$ and hence $D_z\Phi(z)$ is nonsingular. However, by Lemma 3.2 and Lemma 3.1 with Assumption 3.1(b) , we know that for $(x_1, \ldots, x_N, y, s_1, \ldots, s_N) \in \mathcal{O}$, $WS^{-1}XW^T$ is invertible. Hence we are done. **QED**

We have shown in Theorem 3.1 that for suitable initial point $(x_1^0, \ldots, x_N^0, y^0, s_1^0, \ldots, s_N^0)$, an off-central path, passing through this point and satisfying (3.4)-(3.5), exists. The region of initial points is, arguably, the largest possible, given Example 3.1. It is clear that every accumulation point of such off-central path as μ approaches zero is an optimal solution to $(\mathcal{P})$-$(\mathcal{D})$. In fact, by an analogous argument as in [6], the off-central path converges to an unique limit point, which is an optimal solution to $(\mathcal{P})$-$(\mathcal{D})$.

3.2 Asymptotic Properties of Off-Central Path for SOCP

In this section, we again restrict our attention to AHO direction, that is, to off-central path defined by (3.4) and (3.5).
We show that for off-central paths restricted to a neighbourhood of the central path, they converge to strictly complementary optimal solutions of $(\mathcal{P})-(\mathcal{D})$. We then show that when we consider the 1-cone SOCP, for an off-central path in this neighbourhood, it is analytic at the limit point when $\mu = 0$. This has an impact on the rate of convergence of path-following interior-point algorithms, see [22, 26].

Let us first define what we meant by strictly complementary optimal solutions:

Definition 3.1 *Let* $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*)$ *be an optimal solution to* $(\mathcal{P}) - (\mathcal{D})$. $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*)$ *is strictly complementary if and only if*
(a) $(x_i^*)_0 > \|\overline{x_i^*}\| \Longleftrightarrow s_i^* = 0$,
(b) $(s_i^*)_0 > \|\overline{s_i^*}\| \Longleftrightarrow x_i^* = 0$,
(c) $0 \neq (x_i^*)_0 = \|\overline{x_i^*}\| \Longleftrightarrow 0 \neq (s_i^*)_0 = \|\overline{s_i^*}\|$.

We assume from now onwards that for the SOCP that we consider, there always exists a strictly complementary optimal solution.

We know from the previous section that if $(x_1(\mu), \ldots, x_N(\mu), y(\mu), s_1(\mu), \ldots, s_N(\mu))$, $\mu > 0$, is an off-central path, then it converges to an unique limit point. We are going to show that if this off-central path is restricted to a certain neighbourhood of the central path, then the unique limit point is actually a strictly complementary optimal solution to $(\mathcal{P}) - (\mathcal{D})$.

The neighbourhood is derived from the following condition on the off-central path:

Assumption 3.2 *Let* $m_i = Arw(x_i^0)s_i^0$ *for* $i = 1, \ldots, N$, *where* $(x_1^0, \ldots, x_N^0, y^0, s_1^0, \ldots, s_N^0) = (x_1(1), \ldots, x_N(1), y(1), s_1(1), \ldots, s_N(1))$. *Then, we assume that* $(m_i)_0 > 3\|\overline{m_i}\|$ *for* $i = 1, \ldots, N$.

The ”restricted” neighbourhood that we consider is defined by:

$$\mathcal{N} := \left\{ (x_1, \ldots, x_N, y, s_1, \ldots, s_N) \ : \ \|Arw(x_i)s_i - x_i^T s_i e_i\| < \frac{1}{3} x_i^T s_i, i = 1, \ldots, N \right\}.$$

Here $e_i = (1, 0, \ldots, 0)^T \in \Re^{k_i+1}$.

Therefore, an off-central path satisfies Assumption 3.2 if and only if it stays in $\mathcal{N}$ for all $\mu > 0$.

Note that the central path belongs to $\mathcal{N}$. Hence $\mathcal{N}$ is a neighbourhood of the central path. Also, note that the neighbourhood of the central path defined here differs from the one defined in [14, 1].

We have the following theorem:

Theorem 3.2 *Suppose Assumption 3.2 holds. Then* $(x_1(\mu), \ldots, x_N(\mu), y(\mu), s_1(\mu), \ldots, s_N(\mu))$ *converges to a strictly complementary optimal solution of* $(\mathcal{P})$-$(\mathcal{D})$ *where* $(x_1(\mu), \ldots, x_N(\mu), y(\mu), s_1(\mu), \ldots, s_N(\mu))$ *is the solution to (3.4) and (3.5) for* $\mu > 0$.

We defer the proof of Theorem 3.2 as we need to use other results in its proof as discussed below.

Let $\mathcal{O}_{\mathcal{P}}$ = primal optimal solution set to $(\mathcal{P})$ and $\mathcal{O}_{\mathcal{D}}$ = dual optimal solution set to $(\mathcal{D})$.

Consider $\mathcal{O}_{\mathcal{P}}$.

Let $M_{\mathcal{P}} := \{i \in \{1, \ldots, N\} \ : \ \forall \ (x_1^*, \ldots, x_N^*) \in \mathcal{O}_{\mathcal{P}}, \ (x_i^*)_0 = \|\overline{x_i^*}\|\}$,
$M_{\mathcal{P}}^1 := \{i \in M_{\mathcal{P}} \ : \ \exists \ (x_1^*, \ldots, x_N^*) \in \mathcal{O}_{\mathcal{P}}$ with $x_i^* = 0$ and $\exists \ (\widehat{x_1}, \ldots, \widehat{x_N}) \in \mathcal{O}_{\mathcal{P}}$ with $\widehat{x_i} \neq 0\}$,
$M_{\mathcal{P}}^2 := \{i \in M_{\mathcal{P}} \ : \ \forall \ (x_1^*, \ldots, x_N^*) \in \mathcal{O}_{\mathcal{P}}, \ x_i^* = 0\}$,
$M_{\mathcal{P}}^3 := \{i \in M_{\mathcal{P}} \ : \ \forall \ (x_1^*, \ldots, x_N^*) \in \mathcal{O}_{\mathcal{P}}, \ x_i^* \neq 0\}$.
Therefore, $M_{\mathcal{P}}^c = \{i \in \{1, \ldots, N\} \ : \ \exists \ (x_1^*, \ldots, x_N^*) \in \mathcal{O}_{\mathcal{P}}, \ (x_i^*)_0 > \|\overline{x_i^*}\|\}$.

Note that if $(x_1^*, \ldots, x_N^*), \ (\widehat{x_1}, \ldots, \widehat{x_N}) \in \mathcal{O}_{\mathcal{P}}$, then for each $i \in M_{\mathcal{P}}$, $\widehat{x_i} = \alpha_i x_i^*$ for some $\alpha_i \geq 0$, assuming $x_i^* \neq 0$.

We have below a lemma that characterizes the relative interior of $\mathcal{O}_{\mathcal{P}}$.

Lemma 3.3 $\mathrm{ri}\mathcal{O}_{\mathcal{P}} = \{(x_1^*, \ldots, x_N^*) \in \mathcal{O}_{\mathcal{P}} \ : \ (x_i^*)_0 > \|\overline{x_i^*}\| \ \forall \ i \in M_{\mathcal{P}}^c$ and $x_i^* \neq 0 \ \forall \ i \in M_{\mathcal{P}}^1\}$.

Proof. Let us denote the set on the right hand side of the equality sign in the lemma by X. Therefore, we need to show that $\mathrm{ri}\mathcal{O}_{\mathcal{P}} = X$.
($\subseteq$) Let $(x_1^*, \ldots, x_N^*) \in \mathrm{ri}\mathcal{O}_{\mathcal{P}}$.
$\exists \ (\widehat{x_1}, \ldots, \widehat{x_N}) \in \mathcal{O}_{\mathcal{P}}$ such that $\forall \ i \in M_{\mathcal{P}}^c$, $(\widehat{x_i})_0 > \|\overline{\widehat{x_i}}\|$ (by taking convex combinations).
$\exists \ (\widetilde{x_1}, \ldots, \widetilde{x_N}) \in \mathcal{O}_{\mathcal{P}}$ such that $\forall \ i \in M_{\mathcal{P}}^1$, $\widetilde{x_i} \neq 0$ (again, by taking convex combinations).
Let $(\check{x_1}, \ldots, \check{x_N}) = \lambda(\widehat{x_1}, \ldots, \widehat{x_N}) + (1-\lambda)(\widetilde{x_1}, \ldots, \widetilde{x_N})$ with $0 < \lambda < 1$.
Then $(\check{x_1}, \ldots, \check{x_N}) \in \mathcal{O}_{\mathcal{P}}$ and $\forall \ i \in M_{\mathcal{P}}^c$, $(\check{x_i})_0 > \|\overline{\check{x_i}}\|$ and $\forall \ i \in M_{\mathcal{P}}^1$, $\check{x_i} \neq 0$.
Then, by Theorem 6.4 of [18], pp. 47, $\exists \ \mu > 1$ such that $\mu(x_1^*, \ldots, x_N^*) + (1 - \mu)(\check{x_1}, \ldots, \check{x_N}) \in \mathcal{O}_{\mathcal{P}}$, since $(x_1^*, \ldots, x_N^*) \in \mathrm{ri}\mathcal{O}_{\mathcal{P}}$.
Therefore, $(x_1^*, \ldots, x_N^*) = \alpha(\check{x_1}, \ldots, \check{x_N}) + (1-\alpha)(v_1^*, \ldots, v_N^*)$ for some $0 < \alpha < 1$ and $(v_1^*, \ldots, v_N^*) \in \mathcal{O}_{\mathcal{P}}$.
Hence, $\forall \ i \in M_{\mathcal{P}}^c$, $(x_i^*)_0 > \|\overline{x_i^*}\|$ and $\forall \ i \in M_{\mathcal{P}}^1$, $x_i^* \neq 0$.
This implies that $(x_1^*, \ldots, x_N^*) \in X$.
($\supseteq$) Let $(x_1^*, \ldots, x_N^*) \in X$.

Given $(\widehat{x_1}, \ldots, \widehat{x_N}) \in \mathcal{O}_{\mathcal{P}}$.
$\forall\ i \in M^c_{\mathcal{P}}$, $\exists\ \mu_i > 1$ such that $\|\overline{\mu_i x_i^* + (1-\mu_i)\widehat{x_i}}\| \le (\mu_i x_i^* + (1-\mu_i)\widehat{x_i})_0$.
$\forall\ i \in M^1_{\mathcal{P}}$, we have $x_i^* \ne 0$ and $(x_i^*)_0 > 0$. Hence, $\exists\ \mu_i > 1$ such that

$$\|\overline{\mu_i x_i^* + (1-\mu_i)\widehat{x_i}}\| = (\mu_i x_i^* + (1-\mu_i)\widehat{x_i})_0$$

with $(\mu_i x_i^* + (1-\mu_i)\widehat{x_i})_0 \ge 0$.
Similarly for $i \in M^2_{\mathcal{P}}$ and $i \in M^3_{\mathcal{P}}$.
Let $\mu = \min\{\mu_i\}$. Then $\mu > 1$ and $\mu(x_1^*, \ldots, x_N^*) + (1-\mu)(\widehat{x_1}, \ldots, \widehat{x_N}) \in \mathcal{O}_{\mathcal{P}}$.
Therefore, again, by Theorem 6.4 of [18], pp. 47, $(x_1^*, \ldots, x_N^*) \in$ ri$\mathcal{O}_{\mathcal{P}}$. **QED**

Next, we consider $\mathcal{O}_{\mathcal{D}}$. Again, we partition $\{1, \ldots, N\}$ into disjoint sets as follows:
$M_{\mathcal{D}} := \{i \in \{1, \ldots, N\}\ :\ \forall\ (y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{D}},\ (s_i^*)_0 = \|\overline{s_i^*}\|\}$,
$M^1_{\mathcal{D}} := \{i \in M_{\mathcal{D}}\ :\ \exists\ (y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{D}}$ with $s_i^* = 0$ and $\exists\ (\widehat{y}, \widehat{s_1}, \ldots, \widehat{s_N}) \in \mathcal{O}_{\mathcal{D}}$ with $\widehat{s_i} \ne 0\}$,
$M^2_{\mathcal{D}} := \{i \in M_{\mathcal{D}}\ :\ \forall\ (y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{D}},\ s_i^* = 0\}$,
$M^3_{\mathcal{D}} := \{i \in M_{\mathcal{D}}\ :\ \forall\ (y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{D}},\ s_i^* \ne 0\}$.
Therefore, $M^c_{\mathcal{D}} = \{i \in \{1, \ldots, N\}\ :\ \exists\ (y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{D}},\ (s_i^*)_0 > \|\overline{s_i^*}\|\}$.
We also have

$$\text{ri}\mathcal{O}_{\mathcal{D}} = \{(y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{D}}\ :\ (s_i^*)_0 > \|\overline{s_i^*}\|\ \forall\ i \in M^c_{\mathcal{D}} \text{ and } s_i^* \ne 0\ \forall\ i \in M^1_{\mathcal{D}}\}.$$

Remark 3.2 *Without assuming strict complementarity, it is easy to see from the above characterization of* ri$\mathcal{O}_{\mathcal{P}}$ *and* ri$\mathcal{O}_{\mathcal{D}}$ *that for the 2-cone SOCP under Assumption 3.1, if the primal and dual optimal solutions are both not unique, then* $(\mathcal{P}) - (\mathcal{D})$ *always has a strictly complementary optimal solution. This is analogous to the well-known existence result of strictly complementary optimal solution for linear programming. It is still an open question whether for* $n-$*cone SOCP,* $n \ge 3$*, when the primal and dual optimal solutions are both not unique, there always exists strictly complementary optimal solutions under Assumption 3.1. In the case when the primal or dual optimal solution is unique, it is easy to* *find an example to show that there does not exist a strictly complementary optimal solution under Assumption 3.1 alone. For example, one may consider a SOCP converted from a strongly convex quadratic programming problem whose unique solution does not satisfy the strict complementarity condition.*

We observe that if $(\mathcal{P})-(\mathcal{D})$ has a strictly complementary optimal solution and by the first equation in (3.1) (which is called the complementary slackness condition), we have

$$\begin{gathered} M_{\mathcal{P}}^2 \subseteq M_{\mathcal{D}}^c,\ M_{\mathcal{D}}^2 \subseteq M_{\mathcal{P}}^c \\ M_{\mathcal{P}}^3 \subseteq M_{\mathcal{D}}^1 \cup M_{\mathcal{D}}^3,\ M_{\mathcal{D}}^3 \subseteq M_{\mathcal{P}}^1 \cup M_{\mathcal{P}}^3 \\ M_{\mathcal{P}}^1 \subseteq M_{\mathcal{D}}^1 \cup M_{\mathcal{D}}^3,\ M_{\mathcal{D}}^1 \subseteq M_{\mathcal{P}}^1 \cup M_{\mathcal{P}}^3 \\ M_{\mathcal{P}}^c \subseteq M_{\mathcal{D}}^2,\ M_{\mathcal{D}}^c \subseteq M_{\mathcal{P}}^2. \end{gathered}$$

Therefore, $M_{\mathcal{P}}^c = M_{\mathcal{D}}^2,\ M_{\mathcal{P}}^1 \cup M_{\mathcal{P}}^3 = M_{\mathcal{D}}^1 \cup M_{\mathcal{D}}^3$ and $M_{\mathcal{P}}^2 = M_{\mathcal{D}}^c$.
We can easily see from these and the characterization of $\mathrm{ri}\mathcal{O}_{\mathcal{P}}$ and $\mathrm{ri}\mathcal{O}_{\mathcal{D}}$ above that $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*) \in \mathcal{O}_{\mathcal{P}} \times \mathcal{O}_{\mathcal{D}}$ is strictly complementary if and only if $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*) \in \mathrm{ri}\mathcal{O}_{\mathcal{P}} \times \mathrm{ri}\mathcal{O}_{\mathcal{D}} = \mathrm{ri}(\mathcal{O}_{\mathcal{P}} \times \mathcal{O}_{\mathcal{D}})$.
Using this latter fact, Theorem 3.2 can now be proved by showing that the limit point of an off-central path in $\mathcal{N}$ lies in the relative interior of the optimal solution set.

Proof of Theorem 3.2. Let $(x_1(\mu), \ldots, x_N(\mu), y(\mu), s_1(\mu), \ldots, s_N(\mu)) \to (x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*)$.
Consider any $(\widehat{x_1}, \ldots, \widehat{x_N}, \widehat{y}, \widehat{s_1}, \ldots, \widehat{s_N}) \in \mathrm{ri}\mathcal{O}_{\mathcal{P}} \times \mathrm{ri}\mathcal{O}_{\mathcal{D}}$.
We have

$$\begin{aligned} \sum_{i=1}^N (\widehat{x_i} - x_i(\mu))^T (\widehat{s_i} - s_i(\mu)) &= \sum_{i=1}^N (\widehat{x_i} - x_i(\mu))^T (c_i - A_i^T \widehat{y} + A_i^T y(\mu) - c_i) \\ &= \sum_{i=1}^N [A_i(\widehat{x_i} - x_i(\mu))]^T (y(\mu) - \widehat{y}) = 0. \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &= \sum_{i=1}^{N}(\widehat{x_i} - x_i(\mu))^T(\widehat{s_i} - s_i(\mu)) \\ &= \sum_{i=1}^{N}\widehat{x_i}^T\widehat{s_i} - \sum_{i=1}^{N}\widehat{x_i}^T s_i(\mu) - \sum_{i=1}^{N}\widehat{s_i}^T x_i(\mu) + \sum_{i=1}^{N} x_i(\mu)^T s_i(\mu). \end{aligned}$$

Now, $\sum_{i=1}^{N} x_i(\mu)^T s_i(\mu) = \mu \sum_{i=1}^{N}(m_i)_0$ and $\sum_{i=1}^{N}\widehat{x_i}^T\widehat{s_i} = 0$. Therefore,

$$\sum_{i=1}^{N}\widehat{x_i}^T s_i(\mu) + \sum_{i=1}^{N}\widehat{s_i}^T x_i(\mu) = \mu\sum_{i=1}^{N}(m_i)_0.$$

Note that for any $i = 1, \ldots, N$, $\widehat{x_i}^T s_i(\mu),\ \widehat{s_i}^T x_i(\mu) \geq 0$.

Hence

$$\widehat{x_i}^T s_i(\mu),\ \widehat{s_i}^T x_i(\mu) \leq \mu\sum_{i=1}^{N}(m_i)_0 \text{ for all } i = 1, \ldots, N.$$

Consider $\widehat{x_i}^T s_i(\mu)$.

We have $Arw(x_i(\mu))s_i(\mu) = \mu m_i$. Therefore

$$\begin{gathered} x_i(\mu)^T s_i(\mu) = \mu(m_i)_0 \\ (x_i(\mu))_0\overline{s_i(\mu)} + (s_i(\mu))_0\overline{x_i(\mu)} = \mu\overline{m_i} \end{gathered} \tag{3.7}$$

The second equation of (3.7) implies that

$$(s_i(\mu))_j = \frac{\mu(m_i)_j - (s_i(\mu))_0(x_i(\mu))_j}{(x_i(\mu))_0} \quad \text{for } j = 1, \ldots, k_i.$$

Substituting this into the first equation of (3.7), we get after some manipulations,

$$(s_i(\mu))_0 = \mu\frac{(m_i)_0(x_i(\mu))_0 - \sum_{j=1}^{k_i}(x_i(\mu))_j(m_i)_j}{(x_i(\mu))_0^2 - \|\overline{x_i(\mu)}\|^2},$$

from which,

$$(s_i(\mu))_j = \frac{\mu(m_i)_j}{(x_i(\mu))_0} - \frac{\mu(x_i(\mu))_j}{(x_i(\mu))_0}\left[\frac{(m_i)_0(x_i(\mu))_0 - \sum_{j=1}^{k_i}(x_i(\mu))_j(m_i)_j}{(x_i(\mu))_0^2 - \|\overline{x_i(\mu)}\|^2}\right]$$

Now, $\widehat{x}_i^T s_i(\mu) \leq \mu \sum_{i=1}^N (m_i)_0$ implies that

$$(\widehat{x}_i)_0 \left[\frac{(m_i)_0(x_i(\mu))_0 - \sum_{j=1}^{k_i}(x_i(\mu))_j(m_i)_j}{(x_i(\mu))_0^2 - \|\overline{x_i(\mu)}\|^2}\right] + \sum_{k=1}^{k_i}(\widehat{x}_i)_k \left\{\frac{(m_i)_k}{(x_i(\mu))_0} - \frac{(x_i(\mu))_k}{(x_i(\mu))_0}\left[\frac{(m_i)_0(x_i(\mu))_0 - \sum_{j=1}^{k_i}(x_i(\mu))_j(m_i)_j}{(x_i(\mu))_0^2 - \|\overline{x_i(\mu)}\|^2}\right]\right\} \leq \sum_{i=1}^N (m_i)_0$$

Upon manipulations, we have

$$\left[1 - \frac{\sum_{j=1}^{k_i}(m_i)_j(x_i(\mu))_j}{(m_i)_0(x_i(\mu))_0}\right]\left[\frac{(\widehat{x}_i)_0(x_i(\mu))_0 - \sum_{j=1}^{k_i}(\widehat{x}_i)_j(x_i(\mu))_j}{(x_i(\mu))_0^2 - \|\overline{x_i(\mu)}\|^2}\right] \leq \frac{\sum_{k=1}^N (m_k)_0}{(m_i)_0} - \frac{\sum_{j=1}^{k_i}(m_i)_j(\widehat{x}_i)_j}{(m_i)_0(x_i(\mu))_0} \tag{3.8}$$

Now,

$$\frac{\sum_{j=1}^{k_i}(m_i)_j(x_i(\mu))_j}{(m_i)_0(x_i(\mu))_0} \leq \frac{\|\overline{m_i}\|}{(m_i)_0}\frac{\|\overline{x_i}(\mu)\|}{(x_i(\mu))_0} \leq \frac{\|\overline{m_i}\|}{(m_i)_0}$$

Also,

$$-\frac{\sum_{j=1}^{k_i}(m_i)_j(\widehat{x}_i)_j}{(m_i)_0(x_i(\mu))_0} \leq \frac{\|\overline{m_i}\|}{(m_i)_0}\frac{\|\overline{\widehat{x}_i}\|}{(x_i(\mu))_0} \leq \frac{\|\overline{m_i}\|}{(m_i)_0}\frac{(\widehat{x}_i)_0}{(x_i(\mu))_0}$$

and

$$\begin{aligned}
\frac{(\widehat{x}_i)_0(x_i(\mu))_0 - \sum_{j=1}^{k_i}(\widehat{x}_i)_j(x_i(\mu))_j}{(x_i(\mu))_0^2 - \|\overline{x_i(\mu)}\|^2} &= \frac{1 - \sum_{j=1}^{k_i}\frac{(\widehat{x}_i)_j(x_i(\mu))_j}{(\widehat{x}_i)_0(x_i(\mu))_0}}{1 - \sum_{j=1}^{k_i}\frac{(x_i(\mu))_j^2}{(x_i(\mu))_0^2}}\frac{(\widehat{x}_i)_0}{(x_i(\mu))_0} \\
&\geq \frac{(\widehat{x}_i)_0}{(x_i(\mu))_0}\frac{1 - \left(\sum_{j=1}^{k_i}\frac{(\widehat{x}_i)_j^2}{(\widehat{x}_i)_0^2}\right)^{1/2}\left(\sum_{j=1}^{k_i}\frac{(x_i(\mu))_j^2}{(x_i(\mu))_0^2}\right)^{1/2}}{1 - \sum_{j=1}^{k_i}\frac{(x_i(\mu))_j^2}{(x_i(\mu))_0^2}} \\
&\geq \frac{(\widehat{x}_i)_0}{(x_i(\mu))_0}\frac{1}{1 + \left(\sum_{j=1}^{k_i}\frac{(\widehat{x}_i)_j^2}{(\widehat{x}_i)_0^2}\right)^{1/2}\left(\sum_{j=1}^{k_i}\frac{(x_i(\mu))_j^2}{(x_i(\mu))_0^2}\right)^{1/2}} \\
&\geq \frac{1}{2}\frac{(\widehat{x}_i)_0}{(x_i(\mu))_0}
\end{aligned}$$

Therefore, we have from (3.8),

$$\frac{1}{2}\left(1 - \frac{\|\overline{m_i}\|}{(m_i)_0}\right)\frac{(\widehat{x}_i)_0}{(x_i(\mu))_0} \leq \frac{\sum_{k=1}^N (m_k)_0}{(m_i)_0} + \frac{\|\overline{m_i}\|}{(m_i)_0}\frac{(\widehat{x}_i)_0}{(x_i(\mu))_0}.$$

That is,

$$\frac{1}{2}\left(1 - 3\frac{\|\overline{m_i}\|}{(m_i)_0}\right)\frac{(\widehat{x}_i)_0}{(x_i(\mu))_0} \leq \frac{\sum_{k=1}^N (m_k)_0}{(m_i)_0}. \tag{3.9}$$

Now, if $i \in M_{\mathcal{P}}^c$, then $(\widehat{x_i})_0 > 0$. Hence by (3.9) and Assumption 3.2, $(x_i^*)_0 > 0$. Also, since $\|\overline{\widehat{x_i}}\| < (\widehat{x_i})_0$ and $\|\overline{m_i}\| < (m_i)_0$, we have, by (3.8), $\|\overline{x_i^*}\| < (x_i^*)_0$.
If $i \in M_{\mathcal{P}}^1$, then $(\widehat{x_i})_0 > 0$, (3.9) and Assumption 3.2 implies that $(x_i^*)_0 > 0$. Thus $x_i^* \neq 0$.
Hence $(x_1^*, \ldots, x_N^*) \in$ ri$\mathcal{O}_{\mathcal{P}}$ by the above characterization of ri$\mathcal{O}_{\mathcal{P}}$.
By similar argument, we also have $(y^*, s_1^*, \ldots, s_N^*) \in$ ri$\mathcal{O}_{\mathcal{D}}$.
Therefore, $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*) \in$ ri$\mathcal{O}_{\mathcal{P}} \times$ ri$\mathcal{O}_{\mathcal{D}}$. That is, $(x_1^*, \ldots, x_N^*, y^*, s_1^*, \ldots, s_N^*)$ is strictly complementary. **QED**

It is important to determine whether the limit point of an off-central path is strictly complementary since we can then use it to analyze the analyticity of the off-central path at the limit when $\mu = 0$. This has an impact on the rate of convergence of interior-point algorithms, see [22, 26]. As an illustration of the use of strict complementarity on asymptotic analyticity, we have the following proposition:

Proposition 3.1 *Consider $N = 1$, that is, a 1-cone SOCP. Assume that the primal feasible set is not equal to the primal optimal solution set and the dual feasible set is not equal to the dual optimal solution set. If Assumption 3.2 holds for an off-central path $(x(\mu), y(\mu), s(\mu))$, then it is analytic at the limit point when $\mu = 0$.*

Proof. Suppose Assumption 3.2 holds for an off-central path $(x(\mu), y(\mu), s(\mu))$, $\mu > 0$.
Let $(x(\mu), y(\mu), s(\mu)) \longrightarrow (x^*, y^*, s^*)$ as $\mu \longrightarrow 0$.
Since the primal feasible set is not equal to the primal optimal solution set, the dual feasible set is not equal to the dual optimal solution set and (x^*, y^*, s^*) is strictly complementary (by Theorem 3.2), we must have $0 \neq (x^*)_0 = \|\overline{x^*}\|$, $0 \neq (s^*)_0 = \|\overline{s^*}\|$.
We want to show that $(x(\mu), y(\mu), s(\mu))$ is analytic at $\mu = 0$.

Consider the map $\Psi : \Re^{k+1} \times \Re^m \times \Re^{k+1} \times \Re \longmapsto \Re^m \times \Re^{k+1} \times \Re^{k+1}$ defined by

$$\Psi(x, y, s, \mu) := \begin{pmatrix} Ax - b \\ A^T y + s - c \\ Arw(x)s - \mu Arw(x^0)s^0 \end{pmatrix}.$$

If we can show that $D_z\Psi(x^*, y^*, s^*, 0)$, where $z = (x, y, s)$, is nonsingular, then we are done by the Implicit Function Theorem, since Ψ is analytic for all $(x, y, s, \mu) \in \Re^{k+1} \times \Re^m \times \Re^{k+1} \times \Re$.

Now,

$$D_z\Psi(x^*, y^*, s^*, 0) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Arw(s^*) & 0 & Arw(x^*) \end{pmatrix}.$$

Note that we can write $Arw(s^*) = QD_1Q^T$ and $Arw(x^*) = QD_2Q^T$ where $Q = (q_1, q_2, \ldots, q_{k+1})$, $q_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \frac{\bar{x}^*}{\|\bar{x}^*\|} \end{pmatrix}$, $q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\frac{\bar{x}^*}{\|\bar{x}^*\|} \end{pmatrix}$ $(QQ^T = I)$, $D_1 = diag(0, 2s_0, s_0, \ldots, s_0)$ and $D_2 = diag(2x_0, 0, x_0, \ldots, x_0)$. Therefore,

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Arw(s^*) & 0 & Arw(x^*) \end{pmatrix} =$$

$$diag(I, I, Q) \begin{pmatrix} AQ & 0 & 0 \\ 0 & A^T & Q \\ D_1 & 0 & D_2 \end{pmatrix} diag(Q^T, I, Q^T)$$

Hence, to show that $\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Arw(s^*) & 0 & Arw(x^*) \end{pmatrix}$ is nonsingular, we need only show that $\begin{pmatrix} AQ & 0 & 0 \\ 0 & A^T & Q \\ D_1 & 0 & D_2 \end{pmatrix}$ is nonsingular.

Consider

$$\begin{pmatrix} AQ & 0 & 0 \\ 0 & A^T & Q \\ D_1 & 0 & D_2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = 0. \tag{3.10}$$

If we can show that $u, v, w = 0$, then $\begin{pmatrix} AQ & 0 & 0 \\ 0 & A^T & Q \\ D_1 & 0 & D_2 \end{pmatrix}$ is nonsingular.

From (3.10), we have

$$\begin{aligned} AQu &= 0 \\ A^T v + Qw &= 0 \\ D_1 u + D_2 w &= 0. \end{aligned} \tag{3.11}$$

Observe that, except for one entry, all the diagonal entries of D_1 and D_2 are nonzero. Using this fact, we have, from the last equation in (3.11), that $u = (u_1, 0, u_3, \ldots, u_{k+1})^T$ and $w = (0, w_2, w_3, \ldots, w_{k+1})^T$, where $u_i = -\frac{x_0}{s_0} w_i, i = 3, \ldots, k+1$. Using the first two equations in (3.11) and $QQ^T = I$, we have $\sum_{i=3}^{k+1} u_i w_i = 0$. Hence, with $u_i = -\frac{x_0}{s_0} w_i, i = 3, \ldots, k+1$, $u_i = w_i = 0$ for $i = 3, \ldots, k+1$.

From $AQu = 0$, with $u = (u_1, 0, \ldots, 0)^T$, we have $u_1 A q_1 = 0$. That is, $\frac{1}{\sqrt{2}} \frac{u_1}{(x^*)_0} A x^* = 0$. But $Ax^* = b$, therefore, $\frac{u_1}{(x^*)_0} b = 0$. Now, since the dual feasible set is not equal to the dual optimal solution set, we must have $b \neq 0$. Therefore, $u_1 = 0$.

Similarly, the primal feasible set not equal to the primal optimal solution set implies that $w_2 = 0$. Also, $v = 0$, since A has full row rank.

Therefore, we have $u, v, w = 0$ and we are done. **QED**

It should be noted that the above proposition is not interesting since a closed form formula for the primal and dual optimal solution for 1-cone SOCP is already known, see [1]. Since it is generally believed that without strict complementarity, it is difficult to analyze the asymptotic analyticity behaviour of off-central path, we state this proposition here to illustrate that with only strict complementarity, it is possible to derive asymptotic analyticity behaviour of off-central path. Also, its proof is given since it is quite ”neat”.

Chapter 4

Future Directions

The work done in this dissertation is not really complete. First of all, for the asymptotic behaviour of off-central path for SDLCP, we have yet to show that it has a unique limit point as μ approaches zero under weak assumptions, although we believe that this should be true. Also, we state in Chapter 2, Section 2.3, a necessary and sufficient condition for an off-central path for SDLCP to be analytic w.r.t $\sqrt{\mu}$ at the limit when $\mu = 0$. This necessary and sufficient condition unfortunately is not very practical and we would like to find a more practical condition for analyticity of off-central path that can be "implemented", like the example that we analyzed in Section 2.2 and for which, we have an algebraic condition $y_2 = -y_1$ for analyticity. This algebraic condition proves to be useful when we consider local convergence behaviour of first-order predictor-corrector algorithm.

As for off-central path for SOCP, we only consider the existence of off-central path for $\mu > 0$ for the AHO direction. It would be interesting to consider this question in general for other directions. We would also like to further investigate the convergence to strictly complementary optimal solution and asymptotic analyticity of off-central path for multiple cone SOCP, for the AHO or other directions, which are still open questions. We believe that there are a lot more work that need to be done in the area of SOCP, in particular, multiple cone SOCP in relation to interior point algorithm and its underlying paths, and the work presented in this dissertation is only preliminary.

References

[1] F. Alizadeh and D. Goldfarb, *Second-Order Cone Programming,* Mathematical Programming, 95(2003), No. 1, Series B, pp. 3-51.

[2] Herbert Amann; translated from German by Gerhard Metzen, *Ordinary Differential Equations : An Introduction to Nonlinear Analysis,* de Gruyter Studies in Mathematics 13, 1990.

[3] Garrett Birkhoff and Gian-Carlo Rota, *Ordinary Differential Equations,* John Wiley & Sons, Fourth Edition, 1989.

[4] Earl A. Coddington and Norman Levinson, *Theory of Ordinary Differential Equations,* International Series in Pure and Applied Mathematics, 1955.

[5] Osman Güler, *Limiting Behavior of Weighted Central Paths in Linear Programming,* Mathematical Programming, 65(1994), No. 3, Series A, pp. 347-363.

[6] M. Halická, E. de Klerk and C. Roos, *On the Convergence of the Central Path in Semidefinite Optimization,* SIAM Journal on Optimization, 2002, Vol. 12, No. 4, pp. 1090-1099.

[7] E.L. Ince, *Ordinary Differential Equations,* Dover Publications, 1956.

[8] Masakazu Kojima, Masayuki Shida and Susumu Shindoh, *Local Convergence of Predictor-Corrector Infeasible-Interior-Point Algorithms for SDPs*

and SDLCPs, Mathematical Programming, 80(1998), No. 72, Series A, pp. 129-160.

[9] Zhaosong Lu and Renato D.C. Monteiro, *Error Bounds and Limiting Behavior of Weighted Paths associated with the SDP map* $X^{1/2}SX^{1/2}$, SIAM Journal on Optimization, 2004, Vol. 15, No. 2, pp. 348-374.

[10] Zhaosong Lu and Renato D.C. Monteiro, *Limiting Behavior of the Alizadeh-Haeberly-Overton Weighted Paths in Semidefinite Programming,* Preprint, July 24, 2003.

[11] Renato D.C. Monteiro and Jong-Shi Pang, *Properties of an Interior-Point Mapping for Mixed Complementarity Problems,* Mathematics of Operations Research, August 1996, Vol. 21, No. 3, pp. 629-654.

[12] Renato D.C. Monteiro and Jong-Shi Pang, *On Two Interior-Point Mappings for Nonlinear Semidefinite Complementarity Problems,* Mathematics of Operations Research, February 1998, Vol. 23, No. 1, pp. 39-60.

[13] Renato D.C. Monteiro and T. Tsuchiya, *Limiting Behavior of the Derivatives of certain trajectories associated with a Monotone Horizontal Linear Complementarity Problem,* Mathematics of Operations Research, November 1996, Vol. 21, No. 4, pp. 793-814.

[14] Renato D.C. Monteiro and T. Tsuchiya, *Polynomial Convergence of Primal-Dual Algorithms for the Second-Order Cone Program based on the MZ-family of directions,* Mathematical Programming, 88(2000), Series A, pp. 61-83.

[15] Renato D.C. Monteiro and Paulo R. Zanjácomo, *General Interior-Point Maps and Existence of Weighted Paths for Nonlinear Semidefinite Complementarity Problems,* Mathematics of Operations Research, August 2000, Vol. 25, No. 3, pp. 381-399.

[16] João X. da Cruz Neto, Orizon P. Ferreira and Renato D.C. Monteiro, *Asymptotic Behavior of the Central Path for a special class of Degenerate SDP Problems,* Preprint, July 7, 2003.

[17] Martin Preiß and Josef Stoer, *Analysis of Infeasible-Interior-Point Paths arising with Semidefinite Linear Complementarity Problems,* Mathematical Programming, 99(2004), No. 3, Series A, pp. 499-520.

[18] R. Tyrrell Rockafellar, *Convex Analysis,* Princeton Landmarks in Mathematics and Physics Series, 1997.

[19] Walter Rudin, *Real and Complex Analysis,* McGraw-Hill International Editions - Mathematics Series, Third Edition, 1987.

[20] Masayuki Shida, Susumu Shindoh and Masakazu Kojima, *Existence and Uniqueness of Search Directions in Interior-Point Algorithms for the SDP and the Monotone SDLCP,* SIAM Journal on Optimization, May 1998, Vol. 8, No. 2, pp. 387-396.

[21] Josef Stoer and Martin Wechs, *On the Analyticity Properties of Infeasible-Interior-Point Paths for Monotone Linear Complementarity Problems,* Numerische Mathematik, 81(1999), No. 4, pp. 631-645.

[22] Josef Stoer, Martin Wechs and S. Mizuno, *High Order Infeasible-Interior-Point Methods for solving Sufficient Linear Complementarity Problems,* Mathematics of Operations Research, 1998, Vol. 23, No. 4, pp. 832-862.

[23] M.J. Todd, K.C. Toh and R.H. Tütüncü, *On the Nesterov-Todd Direction in Semidefinite Programming,* SIAM Journal on Optimization, August 1998, Vol. 8, No. 3, pp. 769-796.

[24] Reha H. Tütüncü, *Asymptotic Behavior of Continuous Trajectories for Primal-Dual Potential-Reduction Methods,* SIAM Journal on Optimization, 2003, Vol. 14, No. 2, pp. 402-414.

[25] Yin Zhang, *On extending some Primal-Dual Interior-Point Algorithms from Linear Programming to Semidefinite Programming,* SIAM Journal on Optimization, May 1998, Vol. 8, No. 2, pp. 365-386.

[26] Gongyun Zhao and Jie Sun, *On the Rate of Local Convergence of High-Order-Infeasible-Path-Following Algorithms for P_*-Linear Complementarity Problems,* Computational Optimization and Applications, 1999, 14, pp. 293-307.