# Limiting behavior and analyticity of weighted central paths in semidefinite programming

M. Trnovská[*†] M. Halická[*‡]

December 31, 2007

**Abstract**

In this paper we analyze the limiting behavior of infeasible weighted central paths in semidefinite programming under the assumption that a strictly complementary solution exists. We show that the paths associated with the "square root"symmetrization of the weighted centrality condition are analytic functions of the barrier parameter $\mu$ even at $\mu = 0$ if and only if the weight matrix is block diagonal in terms of optimal block partition of variables. This result strengthens some recent result by Lu and Monteiro establishing the analyticity of the paths as functions of $\sqrt{\mu}$ at $\mu = 0$. Moreover, in this paper we study the analytical properties of the paths associated with the "Cholesky factor"symmetrization. We show that the paths exhibit the same analytical behavior at $\mu = 0$ as the paths corresponding to the square root symmetrization.

**Key words:** Semidefinite programming, interior-point methods, weighted central path, analyticity.

**AMS subject classification:** 90C51, 90C22

## 1 Introduction

By $S^n$ we denote the space of all real symmetric $n \times n$ matrices, and by $S^n_+$ and $S^n_{++}$ the subsets of positive semidefinite and positive definite matrices, respectively. If $\mathbf{X} \in S^n_+$, or $\mathbf{X} \in S^n_{++}$ we write $\mathbf{X} \succeq 0$ or $\mathbf{X} \succ 0$, respectively. By $L^n$ we denote the space of all real lower triangular $n \times n$ matrices, and by $L^n_+$ and $L^n_{++}$ the subsets of lower triangular matrices with nonnegative and positive diagonal entries, respectively. For matrices $\mathbf{X}$ and $\mathbf{Y}$ in $R^{p\times q}$ the standard inner product is defined by $\mathbf{X} \bullet \mathbf{Y} :=\mathrm{tr}(\mathbf{X}^T\mathbf{Y})$, where tr(.) denotes the trace of a matrix.

[*]Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynska dolina, 842 48 Bratislava, Slovakia.

[†]E–mail: trnovska@fmph.uniba.sk

[‡]Corresponding author. E–mail: halicka@fmph.uniba.sk

We consider the following primal-dual pair SDP problems in the standard form

$$\begin{array}{lllll} \text{minimize} & \mathbf{X} \bullet \mathbf{C} & & & \\ \text{subject to} & \mathbf{A}^i \bullet \mathbf{X} & = & b_i, & \text{for all } i = 1, \ldots, m, \\ & \mathbf{X} & \succeq & 0, & \end{array} \qquad (P)$$

and

$$\begin{array}{llll} \text{maximize} & b^T y & & \\ \text{subject to} & \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} & = & \mathbf{C}, \\ & \mathbf{S} & \succeq & 0, \end{array} \qquad (D)$$

where the data consists of $\mathbf{C} \in S^n$, $b \in R^m$ and $\mathbf{A}_i \in S^n$ $(i = 1, ..., m)$. The primal variable is $\mathbf{X} \in S^n$ and the dual variable consists of $(y, \mathbf{S}) \in R^m \times S^n$.

A primal-dual optimal solution $(\mathbf{X}, y, \mathbf{S})$ is called complementary, if $\mathbf{XS} = 0$. A strictly complementary optimal solution is defined as a complementary optimal solution $(\mathbf{X}, y, \mathbf{S})$ satisfying $\mathbf{X} + \mathbf{S} \succ 0$. Contrary to the case of linear programming, the existence of a strictly complementary solution is not generally ensured in SDP, even if the complementary optimal solution exists.

Given fixed $\mathbf{W} \in S^n_{++}$, $\Delta b \in R^m$ and $\Delta \mathbf{C} \in S^n$ the weighted central path is implicitly defined by the following $\mu > 0$ parameterized system of nonlinear conditions

$$\mathbf{A}^i \bullet \mathbf{X} = b_i + \mu \triangle b_i, \quad i = 1, \ldots, m, \quad \mathbf{X} \succ 0, \tag{1}$$

$$\sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C} + \mu \triangle \mathbf{C}, \qquad \mathbf{S} \succ 0, \tag{2}$$

$$\Phi_j(\mathbf{X}, \mathbf{S}) = \mu \mathbf{W}. \tag{3}$$

Here $\Phi_j(\mathbf{X}, \mathbf{S})$ , $j \in \{1, 2.3\}$, is a symmetrization map $\Phi_j : S^n_{++} \times S^n_{++} \to S^n$ defined by

$$\Phi_1(\mathbf{X}, \mathbf{S}) = (\mathbf{XS} + \mathbf{SX})/2, \tag{4}$$

$$\Phi_2(\mathbf{X}, \mathbf{S}) = \mathbf{X}^{\frac{1}{2}} \mathbf{S} \mathbf{X}^{\frac{1}{2}}, \tag{5}$$

$$\Phi_3(\mathbf{X}, \mathbf{S}) = \mathbf{L_X}^T \mathbf{S} \mathbf{L_X}, \tag{6}$$

where $\mathbf{X}^{\frac{1}{2}}$ and $\mathbf{L_X}$ denote the square root and the lower Cholesky factor of the positive definite matrix $\mathbf{X}$ respectively. That is, $\mathbf{X}^{\frac{1}{2}}$ and $\mathbf{L_X}$ are the unique matrices in $S^n_{++}$ and $L^n_{++}$ such that $\mathbf{X}^{\frac{1}{2}} \mathbf{X}^{\frac{1}{2}} = X$ and $\mathbf{L_X} \mathbf{L_X}^T = X$, respectively. The symmetrization $\Phi_1$, defined by (4), is the most treated one in SDP and is called the AHO symmetrization. The symmetrizations $\Phi_2$ and $\Phi_3$ given by (5) and (6) are also used in SDP and we will refer to them as to the square root and Cholesky factor symmetrization, respectively.

For each $j \in \{1, 2, 3\}$ suitable conditions on $(\mathbf{W}, \Delta \mathbf{C}, \Delta b)$ have been established in Monteiro and Zanjcomo [15] such that the system (1)-(3) has a unique solution $p_j(\mu) = (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ for any $\mu > 0$. Hence the mapping $\mu > 0 \to p_j(\mu)$ is well defined and we refer to it as to the weighted central path associated with the symmetrization $\Phi_j$. For brevity, we also call it the AHO (square root, or Cholesky) path.

When $(\mathbf{W}, \Delta \mathbf{C}, \Delta b) = (\mathbf{I}, 0, 0)$ the path $p_j(\mu)$, for any $j \in \{1, 2, 3\}$, is the usual central path. The central path is a key concept in the methodology of interior point methods. The geometric view of the central path is that of an analytic curve which converges to an optimal solution pair. Its properties and limiting behavior have been intensively studied and used in the design of algorithms. For the references on early works dealing with the well-definedness, differentiability and limiting behavior of the central path in the context of linear programming and the monotone complementarity see for example [9].

Limiting behavior of the central path in SDP depends on the existence of a strictly complementary optimal solution. Already the usual proof of the convergence of the central path by a characterization of the limit point does not work when no strictly complementary solution exists (see [7]). Nevertheless, it was shown by using some ideas from algebraic geometry that the central paths in SDP always converges to an optimal solution [7] and some kinds of partial characterization of the limit point were given in [8] and [18]. When a strictly complementary solution exists, the central path has the nice properties known from linear programming: it converges to the so-called analytic center of the optimal solution set (see [4, 11]) and can be analytically extended to a neighborhood of $\mu = 0$ [6].

A generalization of the notion of the weighted central path from linear programming to SDP is a delicate issue. In linear programming, a weighted central path consists of optimal solutions of certain weighted logarithmic barrier problems, or equivalently, of solutions of properly perturbed optimality conditions. Since the barrier problems possess unique optimal solutions, this equivalence yields the well-definedness of the weighted central path in linear programming. Unfortunately, it seems that in SDP there are no such barrier problems associated with the perturbed optimality conditions (1) – (3). Hence the question arises how to prove the existence of solutions to (1) – (3). This question was resolved by Monteiro and Zanjácomo [15] (see also [12]) by means of abstract theory of local homeomorphic maps in nonlinear analysis. As a result, for any symmetrization $\Phi_i$, $i = 1, 2, 3$, some conditions on $(\mathbf{W}, \Delta\mathbf{C}, \Delta b)$ have been introduced under which the system (1) – (3) has a unique solution. Similar results, however by a simpler technique of analytic continuation, have been proved by Preiss and Stoer (only for the AHO symmetrization) in [16] and by Trnovská in [20].

Having a well-defined weighted central path we can study the properties of these paths. Some of them have already been described under the assumptions of strict complementarity.

The most appealing properties are exhibited by the paths that are associated with the AHO symmetrization $\Phi_1$. It was shown independently by Preiss and Stoer in [17] and Lu and Monteiro in [10] that each AHO path can be extended as an analytic function of the barrier $\mu$ to $\mu = 0$. From this fact not only the convergence of the AHO path to an optimal solution follows, but we obtain also the convergence of its derivatives of all orders.

The paths associated with the square root symmetrization $\Phi_2$ were analyzed by Lu and Monteiro in [9]. It was shown that each square root path can be extended as a function of $\sqrt{\mu}$ to $\mu = 0$. From this fact the convergence of the paths follows as well, however, the derivatives only with respect to $\sqrt{\mu}$ are bounded as $\mu \to 0$. Moreover, it was shown in [9] that if the weight matrix $\mathbf{W}$ is not block diagonal (in terms of an optimal block partition) then the first order derivative of the path with respect to $\mu$ is not bounded as $\mu \to 0$. From this it follows that if $\mathbf{W}$ is not block diagonal, then the square root path considered as a function of $\mu$ is not analytic at $\mu = 0$. In this paper we prove that if $\mathbf{W}$ is block diagonal, then the associated square root path is analytic function of $\mu$ at $\mu = 0$.

Moreover, and this is the main goal of the paper, we show that the weighted central paths associated with the Cholesky factor symmetrization $\Phi_3$ exhibit the same limiting behavior as the paths associated with $\Phi_2$. That is, we show that each Cholesky path considered as a function of $\sqrt{\mu}$ can be analytically extended to $\mu = 0$ and, the path as a function of $\mu$ is analytic at $\mu = 0$ if and only if the corresponding weight matrix $\mathbf{W}$ is block diagonal.

To prove the analyticity of the paths we will use the implicit function theorem technique that was developed by Stoer and Wechs [19] in the context of a linear complementarity problem. This technique was also used by many other researchers (see [2, 3, 5, 6, 9, 10, 17]) in the analysis of the limiting behavior of the central and weighted central paths in linear and semidefinite programming.

The technique uses the fact that each weighted central path is implicitly defined by the system of conditions (1-3) involving solely analytic functions. Hence, if the Frechet derivative of the defining function with respect to $(\mathbf{X}, y, \mathbf{S})$ were nonsingular along the path, then the application of the analytic version of the implicit function theorem would yield the analyticity of the path. A drawback is however that this is possible only for $\mu > 0$ since as $\mu \to 0$ the Jacobian may vanish. Nevertheless, in many cases the implicit function can be used. Sometimes a detailed limiting analysis of particular blocks of the path allows to introduce normalized paths and a normalized system of equations has nonsingular Jacobian at any limit point.

In this paper we will use this technique twice. First, we will use a normalization that will yield the analyticity of the paths as functions of $\sqrt{\mu}$ at $\mu = 0$. This Phase I is described in Section 3. Then, under the assumption that the weight matrix is block diagonal, we introduce a new normalization that will allow to prove the analyticity of the paths as functions of $\mu$ at $\mu = 0$. This Phase II and described in Section 4.

## 2 Preliminaries

We make the following two assumptions throughout the paper.

**Assumption 2.1** $\mathbf{A}^i$ $(i = 1, \ldots, m)$ *are linearly independent.*

**Assumption 2.2** *There exists a strictly complementary optimal solution for (P) and (D).*

Assumption 2.1 is only a technical one, enforcing a one-to-one correspondence between $y$ and $\mathbf{S}$ in the dual feasible pairs $(\mathbf{S}, y)$. On the other hand, Assumption 2.2 is restrictive, but is commonly used in the analysis of superlinear convergence of interior-point algorithms. This assumption also plays an crucial role in our analysis. Moreover, the results of the paper [3] regarding the usual central path indicate that without this assumption the analytical properties of weighted central paths could be very complicated and difficult to describe.

Define the map $\mathcal{A} : S^n \to R^m$ as $\mathcal{A}(\mathbf{X}) = [\mathbf{A}^1 \bullet \mathbf{X}, \ldots, \mathbf{A}^m \bullet \mathbf{X}]$. Its adjoint map is $\mathcal{A}^* : R^m \to S^n$, $\mathcal{A}^*(y) = \sum_{i=1}^m \mathbf{A}^i y_i$. It can be easily seen that for $(\mathbf{X}, y, \mathbf{S}) \in S^n \times R^m \times S^n$ the following orthogonality property holds: if $\mathcal{A}(\mathbf{X}) = 0$ and $\mathcal{A}^*(y) + \mathbf{S} = 0$, then $\mathbf{X} \bullet \mathbf{S} = 0$.

Thanks to Assumption 2.2 we can use the standard procedure (described for example in [6, 9]) yielding an optimal partition of any $\mathbf{M} \in S^n$ into

$$\mathbf{M} = \begin{pmatrix} M_B & M_V \\ M_V^T & M_N \end{pmatrix}. \tag{7}$$

where $M_B$ and $M_N$ are square blocks of dimensions $|B| \times |B|$ and $|N| \times |N|$, respectively. In this partition any optimal solution pair is of the form

$$\hat{\mathbf{X}} = \begin{pmatrix} \hat{X}_B & 0 \\ 0 & 0 \end{pmatrix}, \qquad \hat{\mathbf{S}} = \begin{pmatrix} 0 & 0 \\ 0 & \hat{S}_N \end{pmatrix},$$

where $X_B \succeq 0$ and $S_N \succeq 0$ and for a strictly complementary optimal solution it holds $X_B \succ 0$ and $S_N \succ 0$. The usage of this optimal partition is standard in SDP when analyzing the limiting behavior of paths under strict complementarity (see [6, 9–11, 17]).

We now describe the conditions for welldefinedness of weighted central paths. For $\varepsilon \in (0, 1)$ denote

$$\mathcal{M}_\varepsilon = \{\mathbf{W} \in S^n_{++};\ \exists \nu : \|\mathbf{W} - \nu I\| < \varepsilon \nu\},$$

where $\|.\|$ means the spectral norm, i.e. $\|\mathbf{A}\| = \max\{\sqrt{\lambda}; \lambda \textit{ is an eigenvalue of } \mathbf{A}^T\mathbf{A}\}$. It can be easily seen that $\mathcal{M}_\varepsilon$ is a convex cone and $\mathbf{W} \in \mathcal{M}_\varepsilon$ if and only if

$$\lambda_{max}(\mathbf{W})/\lambda_{min}(\mathbf{W}) < (1 + \varepsilon)/(1 - \varepsilon).$$

Let $\triangle b, \triangle \mathbf{C}$ be such, that there exists $W_0 \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ and $\mu_0 > 0$ such that the system $(1) - (3)$ is solvable for $\mathbf{W} = \mathbf{W}_0$ and $\mu = \mu_0$. The existence result concerning the weighted central paths associated with the symmetrizations $\Phi_2$ and $\Phi_3$ defined by (5) and (6), respectively, is stated in the following proposition.

**Proposition 2.1** *Let* $\mu \in (0, \mu_0\rangle$ *and*
- $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$, *in the case of* $\Phi_2$;
- $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ *or* $\mathbf{W} \in D^n_{++}$ *in the case of* $\Phi_3$.

*Then there exists a unique solution of the system* $(1) - (3)$ *denoted by* $(X(\mu), y(\mu), S(\mu))$. *Moreover, the path* $\mu \to (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ *is an analytic function for* $\mu \in (0, \mu_0\rangle$.

*Proof.* The statement for $\mathcal{M}_{\frac{1}{\sqrt{2}}}$ (and the symmetrizations $\Phi_2, \Phi_3$) is a consequence of Corollary 1 of [15]. The statement for $D^n_{++}$ (and the symmetrization $\Phi_3$) is proved in Theorem 2 of [20] and for $\Delta b = 0$ and $\Delta \mathbf{C} = 0$ follows from Theorem 1 of [1]. ■

## 3 Phase I: paths as functions of $\sqrt{\mu}$

### 3.1 Introduction of normalized matrices in phase I

In this section we introduce normalized matrices. To this aim we summarize the results concerning the asymptotic behavior of the weighted path matrices and its square root and lower Cholesky factor. We use the customary $\mathcal{O}-$ and $o-$ notation for matrix valued function $\mathbf{A} : \mathbb{R}_{++} \to \mathbb{R}^{m \times n}$. Moreover, for $\mathbf{A}(\mu) \in S^n(L^n)$ we will write $\mathbf{A}(\mu) = \Theta(f(\mu))$ (where $f : \mathbb{R}_{++} \to \mathbb{R}_{++}$) if and only if there exists a constant $\alpha > 0$ such that $\mathbf{A}(\mu)/f(\mu) - 1/\alpha \mathbf{I} \in S^n_{++}$ and $\alpha \mathbf{I} - \mathbf{A}(\mu)/f(\mu) \in S^n_{++}$ ($\mathbf{A}(\mu)/f(\mu) - 1/\alpha \mathbf{I} \in L^n_{++}$ and $\alpha \mathbf{I} - \mathbf{A}(\mu)/f(\mu) \in L^n_{++}$).

The statement in Proposition 3.1 was proved for $j = 2$ in [9] (see Lemma 2.2 and Lemma 2.3). The statement for $j = 3$ can be shown analogously, therefore the proof of the proposition is omitted.

**Proposition 3.1** *Let $j \in \{2, 3\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small the path matrices posses the following asymptotic behavior*

$$\mathbf{X}(\mu) = \begin{pmatrix} \Theta(1) & \mathcal{O}(\sqrt{\mu}) \\ \mathcal{O}(\sqrt{\mu}) & \Theta(\mu) \end{pmatrix}, \quad \mathbf{S}(\mu) = \begin{pmatrix} \Theta(\mu) & \mathcal{O}(\sqrt{\mu}) \\ \mathcal{O}(\sqrt{\mu}) & \Theta(1) \end{pmatrix}. \tag{8}$$

*Moreover, the square root (in the case $j = 2$) and the lower Cholesky factor (in the case $j = 3$) posses the following asymptotic behavior*

$$\mathbf{Y}(\mu) = [\mathbf{X}(\mu)]^{\frac{1}{2}} = \begin{pmatrix} \Theta(1) & \mathcal{O}(\sqrt{\mu}) \\ \mathcal{O}(\sqrt{\mu}) & \Theta(\sqrt{\mu}) \end{pmatrix}, \quad \mathbf{L}(\mu) = \mathbf{L}_{\mathbf{X}(\mu)} = \begin{pmatrix} \Theta(1) & 0 \\ \mathcal{O}(\sqrt{\mu}) & \Theta(\sqrt{\mu}) \end{pmatrix}. \tag{9}$$

Put $\rho := \sqrt{\mu}$ and define the normalized matrices $\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{S}}(\rho)$ in the following way:

$$\begin{array}{ll} \mathbf{X}_B(\mu) = \tilde{\mathbf{X}}_B(\rho), & \mathbf{S}_B(\mu) = \rho^2 \tilde{\mathbf{S}}_B(\rho), \\ \mathbf{X}_V(\mu) = \rho \tilde{\mathbf{X}}_V(\rho), & \mathbf{S}_V(\mu) = \rho \tilde{\mathbf{S}}_V(\rho), \\ \mathbf{X}_N(\mu) = \rho^2 \tilde{\mathbf{X}}_N(\rho), & \mathbf{S}_N(\mu) = \tilde{\mathbf{S}}_N(\rho) \end{array} \tag{10}$$

and $y(\mu) = \tilde{y}(\rho)$. Similarly we can define the matrices $\tilde{\mathbf{Y}}(\rho)$ and $\tilde{\mathbf{L}}(\rho)$ with the equalities

$$\begin{array}{ll} \mathbf{Y}_B(\mu) = \tilde{\mathbf{Y}}_B(\rho), & \mathbf{L}_B(\mu) = \tilde{\mathbf{L}}_B(\rho), \\ \mathbf{Y}_V(\mu) = \rho \tilde{\mathbf{Y}}_V(\rho), & \mathbf{L}_V(\mu) = \rho \tilde{\mathbf{L}}_V(\rho), \\ \mathbf{Y}_N(\mu) = \rho \tilde{\mathbf{Y}}_N(\rho), & \mathbf{L}_N(\mu) = \rho \tilde{\mathbf{L}}_N(\rho). \end{array} \tag{11}$$

From Proposition 3.1 it follows that the functions $\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{S}}(\rho), \tilde{\mathbf{Y}}(\rho), \tilde{\mathbf{L}}(\rho)$ are bounded and thanks to Assumption 2.1 also $\tilde{y}(\rho)$ is bounded. Therefore there exists a sequence

$$\{\rho_k\}_{k=1}^{\infty} \to 0, \ \mu_k = \rho_k^2, \tag{12}$$

such that $\tilde{\mathbf{X}}(\rho_k), \tilde{\mathbf{S}}(\rho_k), \tilde{\mathbf{Y}}(\rho_k), \tilde{\mathbf{L}}(\rho_k)$ converge–and hence there exist limits

$$\begin{array}{ll} \lim_{k\to\infty} \tilde{\mathbf{X}}(\rho_k) =: \tilde{\mathbf{X}}^*, & \lim_{k\to\infty} \tilde{\mathbf{S}}(\rho_k) =: \tilde{\mathbf{S}}^*, \\ \lim_{k\to\infty} \tilde{\mathbf{Y}}(\rho_k) =: \tilde{\mathbf{Y}}^*, & \lim_{k\to\infty} \tilde{\mathbf{L}}(\rho_k) =: \tilde{\mathbf{L}}^* \end{array}$$

and $\lim_{k\to\infty} \tilde{y}(\rho_k) =: \tilde{y}^*$. Moreover, it holds that $\tilde{\mathbf{X}}_B^*, \tilde{\mathbf{X}}_N^*, \tilde{\mathbf{S}}_B^*, \tilde{\mathbf{S}}_N^*, \tilde{\mathbf{Y}}_B^*, \tilde{\mathbf{Y}}_N^* \in S_{++}^n$ and $\tilde{\mathbf{L}}_B^*, \tilde{\mathbf{L}}_N^* \in L_{++}^n$.
The statement in the next lemma was shown in [17] (see proof of the Lemma 3.12) for the weighted paths associated with AHO-symmetrization. The proof for the case $j = 2, 3$ can be done similar way and therefore it is omitted.

**Lemma 3.1** *Let $j \in \{2, 3\}$. Then*

$$\tilde{\mathbf{X}}_V^* \bullet \tilde{\mathbf{S}}_V^* = 0.$$

The following proposition deals with the asymptotic behavior of the path matrices (and its square root and lower Cholesky factor, respectively) under the assumption that the weight matrix is block diagonal.

**Proposition 3.2** *Let $j \in \{2, 3\}$. If $\mathbf{W}_V = 0$, then for $\mu \in (0, \mu_0\rangle$ sufficiently small the path matrices posses the following asymptotic behavior*

$$\mathbf{X}(\mu) = \begin{pmatrix} \Theta(1) & o(\sqrt{\mu}) \\ o(\sqrt{\mu}) & \Theta(\mu) \end{pmatrix}, \quad \mathbf{S}(\mu) = \begin{pmatrix} \Theta(\mu) & o(\sqrt{\mu}) \\ o(\sqrt{\mu}) & \Theta(1) \end{pmatrix}, \tag{13}$$

*moreover, the square root (in the case $j = 2$) and the lower Cholesky factor (in the case $j = 3$) posses the following asymptotic behavior*

$$\mathbf{Y}(\mu) = \mathbf{X}^{\frac{1}{2}}(\mu) = \begin{pmatrix} \Theta(1) & o(\sqrt{\mu}) \\ o(\sqrt{\mu}) & \Theta(\mu) \end{pmatrix}, \quad \mathbf{L}(\mu) = \mathbf{L}_{\mathbf{X}(\mu)} = \begin{pmatrix} \Theta(1) & 0 \\ o(\sqrt{\mu}) & \Theta(\sqrt{\mu}) \end{pmatrix}. \tag{14}$$

*Conversely, if $\mathbf{X}_V(\mu) = o(\sqrt{\mu})$ and $\mathbf{S}_V(\mu) = o(\sqrt{\mu})$, then $\mathbf{W}_V = 0$.*

*Proof.* Assume $j = 2$ The equality

$$\mathbf{Y}(\mu)\mathbf{S}(\mu)\mathbf{Y}(\mu) = \mu\mathbf{W},$$

implies

$$\begin{aligned} &\mathbf{Y}_B(\mu)\mathbf{S}_B(\mu)\mathbf{Y}_V(\mu) + \mathbf{Y}_V(\mu)\mathbf{S}_V^T(\mu)\mathbf{Y}_V(\mu) + \\ &+\mathbf{Y}_B(\mu)\mathbf{S}_V(\mu)\mathbf{Y}_N(\mu) + \mathbf{Y}_V(\mu)\mathbf{S}_N(\mu)\mathbf{Y}_N(\mu) = \mu\mathbf{W}_V. \end{aligned} \tag{15}$$

If we divide (15) by $\mu$ and put $\mu = \rho_k^2$, we obtain

$$\rho_k\tilde{\mathbf{Y}}_B(\rho_k)\tilde{\mathbf{S}}_V(\rho_k)\tilde{\mathbf{Y}}_V(\rho_k) + \rho_k\tilde{\mathbf{Y}}_V(\rho_k)\tilde{\mathbf{S}}_V^T(\rho_k)\tilde{\mathbf{Y}}_V(\rho_k) +$$

$$+\tilde{\mathbf{Y}}_B(\rho_k)\tilde{\mathbf{S}}_V(\rho_k)\tilde{\mathbf{Y}}_N(\rho_k) + \tilde{\mathbf{Y}}_V(\rho_k)\tilde{\mathbf{S}}_N(\rho_k)\tilde{\mathbf{Y}}_N(\rho_k) = \mathbf{W}_V.$$

By taking the limit $k \to \infty$ we have

$$\tilde{\mathbf{Y}}_B^*\tilde{\mathbf{S}}_V^*\tilde{\mathbf{Y}}_N^* + \tilde{\mathbf{Y}}_V^*\tilde{\mathbf{S}}_N^*\tilde{\mathbf{Y}}_N^* = \mathbf{W}_V. \tag{16}$$

If $\mathbf{X}_V(\mu) = o(\sqrt{\mu})$ then $\tilde{\mathbf{Y}}_V^* = 0$, similarly if $\mathbf{S}_V(\mu) = o(\sqrt{\mu})$ then $\tilde{\mathbf{S}}_V^*$ and therefore $\mathbf{W}_V = 0$. Now assume $\mathbf{W}_V = 0$. Because $\tilde{\mathbf{Y}}_N^* \succ 0$, from (16) it follows, that

$$\tilde{\mathbf{Y}}_B^*\tilde{\mathbf{S}}_V^* + \tilde{\mathbf{Y}}_V^*\tilde{\mathbf{S}}_N^* = 0. \tag{17}$$

From Lemma 3.1 and (17) we obtain that

$$0 = \tilde{\mathbf{X}}_V^* \bullet \tilde{\mathbf{S}}_V^* = tr[(\tilde{\mathbf{X}}_V^*)^T \tilde{\mathbf{S}}_V^*] = tr[(\tilde{\mathbf{Y}}_V^*)^T \tilde{\mathbf{Y}}_B^* \tilde{\mathbf{S}}_V^*] = -tr[(\tilde{\mathbf{Y}}_V^*)^T \tilde{\mathbf{Y}}_V^* \tilde{\mathbf{S}}_N^*].$$

Since $\tilde{\mathbf{S}}_N^* \succ 0$ we have that $\tilde{\mathbf{Y}}_V^* = 0$ and therefore also

$$\tilde{\mathbf{X}}_V^* = \tilde{\mathbf{Y}}_B^* \tilde{\mathbf{Y}}_V^* = 0, \qquad \tilde{\mathbf{S}}_V^* = -(\tilde{\mathbf{Y}}_B^*)^{-\frac{1}{2}} \tilde{\mathbf{Y}}_V^* \tilde{\mathbf{S}}_N^* = 0.$$

The statement for $j = 3$ can be shown analogously. ■

## 3.2 Transformation of feasibility conditions in phase I

In order to separate the blocks of the path matrices that posses different types of asymptotic behavior, we need to rewrite the equations in (1) and (2).

We first apply the optimal partition (defined in (7)) to any symmetric $n \times n$ matrix in equations (1) and (2). We obtain:

$$\begin{array}{r} \mathbf{A}_B^i \bullet \mathbf{X}_B + 2\mathbf{A}_V^i \bullet \mathbf{X}_V + \mathbf{A}_N^i \bullet \mathbf{X}_N = b_i + \mu\triangle b_i, \ i = 1, \ldots, m, \\ \sum_{i=1}^m \mathbf{A}_B^i y_i + \mathbf{S}_B = \mathbf{C}_B + \mu\triangle\mathbf{C}_B, \\ \sum_{i=1}^m \mathbf{A}_V^i y_i + \mathbf{S}_V = \mathbf{C}_V + \mu\triangle\mathbf{C}_V, \\ \sum_{i=1}^m \mathbf{A}_N^i y_i + \mathbf{S}_N = \mathbf{C}_N + \mu\triangle\mathbf{C}_N. \end{array} \tag{18}$$

Now, we define the matrices

$$\mathbb{A}_B = \begin{bmatrix} svec(\mathbf{A}_B^1) \\ \vdots \\ svec(\mathbf{A}_B^m) \end{bmatrix}, \quad \mathbb{A}_V = \begin{bmatrix} vec(\mathbf{A}_V^1) \\ \vdots \\ vec(\mathbf{A}_V^m) \end{bmatrix}, \quad \mathbb{A}_N = \begin{bmatrix} svec(\mathbf{A}_N^1) \\ \vdots \\ svec(\mathbf{A}_N^m) \end{bmatrix},$$

where $\mathbb{A}_B \in R^{m \times \bar{B}}$, $\mathbb{A}_V \in R^{m \times \bar{V}}$, $\mathbb{A}_N \in R^{m \times \bar{N}}$ and

$$\bar{B} := |B|(|B|+1)/2, \quad \bar{V} := |B||N|, \quad \bar{N} := |N|(|N|+1)/2.$$

Obviously, $\bar{B} + \bar{N} = \bar{n} - |B||N|$. The system (18) has the matrix-vector form

$$\begin{bmatrix} \mathbb{A}_B & 2\mathbb{A}_V & \mathbb{A}_N & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\mathbb{A}_B)^T & \mathbf{I}_{\bar{B}} & 0 & 0 \\ 0 & 0 & 0 & (\mathbb{A}_V)^T & 0 & \mathbf{I}_{\bar{V}} & 0 \\ 0 & 0 & 0 & (\mathbb{A}_N)^T & 0 & 0 & \mathbf{I}_{\bar{N}} \end{bmatrix} \begin{bmatrix} svec(\mathbf{X}_B) \\ vec(\mathbf{X}_V) \\ svec(\mathbf{X}_N) \\ y \\ svec(\mathbf{S}_B) \\ vec(\mathbf{S}_V) \\ svec(\mathbf{S}_N) \end{bmatrix} = \begin{bmatrix} b + \mu\triangle b \\ svec(\mathbf{C}_B + \mu\triangle\mathbf{C}_B) \\ vec(\mathbf{C}_V + \mu\triangle\mathbf{C}_V) \\ svec(\mathbf{C}_N + \mu\triangle\mathbf{C}_N) \end{bmatrix}.$$

Rewrite the system above once more as

$$\mathbb{P}v + \mathbb{Q}w + \mathbb{R}z = d + \mu\triangle d, \tag{19}$$

where

$$\mathbb{P} = \begin{bmatrix} \mathbb{A}_B & 0 & 0 \\ 0 & \mathbb{A}_B^T & 0 \\ 0 & \mathbb{A}_V^T & 0 \\ 0 & \mathbb{A}_N^T & \mathbf{I}_{\bar{N}} \end{bmatrix}, \quad \mathbb{Q} = \begin{bmatrix} 2\mathbb{A}_V & 0 \\ 0 & 0 \\ 0 & \mathbf{I}_{\bar{V}} \\ 0 & 0 \end{bmatrix}, \quad \mathbb{R} = \begin{bmatrix} \mathbb{A}_N & 0 \\ 0 & \mathbf{I}_{\bar{B}} \\ 0 & 0 \\ 0 & 0 \end{bmatrix};$$

$$v = \begin{bmatrix} svec(\mathbf{X}_B) \\ y \\ svec(\mathbf{S}_N) \end{bmatrix}, \quad w = \begin{bmatrix} vec(\mathbf{X}_V) \\ vec(\mathbf{S}_V) \end{bmatrix}, \quad z = \begin{bmatrix} svec(\mathbf{X}_N) \\ svec(\mathbf{S}_B) \end{bmatrix}$$

and

$$d = \begin{bmatrix} b \\ svec(\mathbf{C}_B) \\ vec(\mathbf{C}_V) \\ svec(\mathbf{C}_N) \end{bmatrix}, \qquad \triangle d = \begin{bmatrix} \triangle b \\ svec(\triangle \mathbf{C}_B) \\ vec(\triangle \mathbf{C}_V) \\ svec(\triangle \mathbf{C}_N) \end{bmatrix}.$$

Here $\mathbf{I}_{\bar{B}}$, $\mathbf{I}_{\bar{V}}$ and $\mathbf{I}_{\bar{N}}$ are identity matrices of dimensions $\bar{B} \times \bar{B}$, $\bar{V} \times \bar{V}$ and $\bar{N} \times \bar{N}$, respectively.
Denote $\bar{n} = \dim(S^n) = n(n+1)/2$. Then $\mathbb{P}$, $\mathbb{Q}$ and $\mathbb{R}$ are real matrices of dimensions $(m+\bar{n}) \times k_1$, $(m+\bar{n}) \times k_2$ and $(m+\bar{n}) \times k_3$, where

$$k_1 = m + \bar{n} - |B||N|, \; k_2 = 2|B||N|, \; k_3 = \bar{n} - |B||N|.$$

The following lemma is a simple consequence of known linear algebra results.

**Lemma 3.2** *Let* $\mathbf{A}$ *be an* $(l \times m)$ *matrix,* $rank(\mathbf{A}) = s$. *Then there exists a nonsingular* $(l \times l)$ *matrix* $\mathbf{M}$ *such that*

$$\mathbf{MA} = \begin{bmatrix} \mathbf{M}_1\mathbf{A} \\ \mathbf{M}_2\mathbf{A} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1\mathbf{A} \\ \mathbf{0} \end{bmatrix},$$

*where* $\mathbf{M}_1$ *is* $s \times l$. *Moreover* $rank(\mathbf{M}_1\mathbf{A}) = s$.

Let

$$s := rank(\mathbb{P}) \leq \min\{k_1, m + \bar{n}\} = k_1. \tag{20}$$

Now from Lemma 3.2 it follows that there exists a nonsingular $(m+\bar{n}) \times (m+\bar{n})$ matrix $\mathbf{M}$ such that

$$\mathbf{M}\mathbb{P} = \begin{bmatrix} \mathbf{M}_1\mathbb{P} \\ \mathbf{M}_2\mathbb{P} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1\mathbb{P} \\ 0 \end{bmatrix},$$

where $\mathbf{M}_1$ is of dimension $s \times (m+\bar{n})$ and $\mathbf{M}_2$ is of dimension $(m+\bar{n}-s) \times (m+\bar{n})$. Moreover $rank(\mathbf{M}_1\mathbb{P}) = s$. By multiplying (19) by $\mathbf{M}$ from the left we obtain an equivalent system

$$\begin{aligned} \mathbf{M}_1\mathbb{P}v + \mathbf{M}_1\mathbb{Q}w + \mathbf{M}_1\mathbb{R}z &= \mathbf{M}_1(d + \mu\triangle d), \\ \mathbf{M}_2\mathbb{Q}w + \mathbf{M}_2\mathbb{R}z &= \mathbf{M}_2(d + \mu\triangle d). \end{aligned} \tag{21}$$

Now let

$$t-s := rank(\mathbf{M}_2\mathbb{Q}) \leq \min\{m+\bar{n}-s, k_2\}. \tag{22}$$

Then again, Lemma 3.2 implies that there exists a nonsingular $(m+\bar{n}-s)\times(m+\bar{n}-s)$ matrix $\mathbf{N}$ such that

$$\mathbf{N}\mathbf{M}_2\mathbb{Q} = \begin{bmatrix} \mathbf{N}_1\mathbf{M}_2\mathbb{Q} \\ \mathbf{N}_2\mathbf{M}_2\mathbb{Q} \end{bmatrix} = \begin{bmatrix} \mathbf{N}_1\mathbf{M}_2\mathbb{Q} \\ 0 \end{bmatrix},$$

where $\mathbf{N}_1$, $\mathbf{N}_2$ have dimensions $(t-s)\times(m+\bar{n}-s)$ and $(m+\bar{n}-t)\times(m+\bar{n}-s)$ and $rank(\mathbf{N}_1\mathbf{M}_2\mathbb{Q}) = t-s$. Therefore, the system (21) is equivalent to

$$\begin{array}{rcl} \mathbf{M}_1\mathbb{P}v + \mathbf{M}_1\mathbb{Q}w + \mathbf{M}_1\mathbb{R}z & = & \mathbf{M}_1(d+\mu\triangle d), \\ \mathbf{N}_1\mathbf{M}_2\mathbb{Q}w + \mathbf{N}_1\mathbf{M}_2\mathbb{R}z & = & \mathbf{N}_1\mathbf{M}_2(d+\mu\triangle d), \\ \mathbf{N}_2\mathbf{M}_2\mathbb{R}z & = & \mathbf{N}_2\mathbf{M}_2(d+\mu\triangle d). \end{array}$$

If we denote $\mathbf{M}_1\mathbb{P} = \tilde{\mathbb{P}}_1$, $\mathbf{M}_1\mathbb{Q} = \tilde{\mathbb{Q}}_1$, $\mathbf{M}_1\mathbb{R} = \tilde{\mathbb{R}}_1$, $\mathbf{N}_1\mathbf{M}_2\mathbb{Q} = \tilde{\mathbb{Q}}_2$, $\mathbf{N}_1\mathbf{M}_2\mathbb{R} = \tilde{\mathbb{R}}_2$ and $\mathbf{N}_2\mathbf{M}_2\mathbb{R} = \tilde{\mathbb{R}}_3$, then the last system can be rewritten in the form

$$\begin{array}{rcl} \tilde{\mathbb{P}}_1 v + \tilde{\mathbb{Q}}_1 w + \tilde{\mathbb{R}}_1 z & = & \tilde{d}_1 + \mu\triangle\tilde{d}_1, \\ \tilde{\mathbb{Q}}_2 w + \tilde{\mathbb{R}}_2 z & = & \tilde{d}_2 + \mu\triangle\tilde{d}_2, \\ \tilde{\mathbb{R}}_3 z & = & \tilde{d}_3 + \mu\triangle\tilde{d}_3. \end{array} \tag{23}$$

It is well-known that if $\mathbf{A}$ is $n\times m$ matrix and $n \leq m$, then the map $x \mapsto \mathbf{A}x$ is surjective if and only if $rank(\mathbf{A}) = n$. This fact, together with relations (20), (22) and Assumption 2.1 can be used to prove the following lemma.

**Lemma 3.3** *The linear maps*

$$v \mapsto \tilde{\mathbb{P}}_1 v, \quad w \mapsto \tilde{\mathbb{Q}}_2 w, \quad z \mapsto \tilde{\mathbb{R}}_3 z$$

*are surjective.*

**Proposition 3.3** *Let*

$$\triangle v = \begin{bmatrix} svec(\triangle\mathbf{X}_B) \\ \triangle y \\ svec(\triangle\mathbf{S}_N) \end{bmatrix}, \quad \triangle w = \begin{bmatrix} vec(\triangle\mathbf{X}_V) \\ vec(\triangle\mathbf{S}_V) \end{bmatrix}, \triangle z = \begin{bmatrix} svec(\triangle\mathbf{X}_N) \\ svec(\triangle\mathbf{S}_B) \end{bmatrix}.$$

*(a) If*

$$\tilde{\mathbb{P}}_1\triangle v = 0, \quad \tilde{\mathbb{Q}}_2\triangle w = 0, \quad \tilde{\mathbb{R}}_3\triangle z = 0,$$

*then*

$$\triangle\mathbf{X}_B \bullet \triangle\mathbf{S}_B = 0, \quad \triangle\mathbf{X}_V \bullet \triangle\mathbf{S}_V = 0, \quad \triangle\mathbf{X}_N \bullet \triangle\mathbf{S}_N = 0.$$

*(b) If*

$$\tilde{\mathbb{P}}_1\triangle v = 0, \quad \tilde{\mathbb{Q}}_2\triangle w + \tilde{\mathbb{R}}_2\triangle z = 0, \quad \tilde{\mathbb{R}}_3\triangle z = 0,$$

*then*

$$\triangle\mathbf{X}_B \bullet \triangle\mathbf{S}_B = 0, \quad \triangle\mathbf{X}_N \bullet \triangle\mathbf{S}_N = 0.$$

*Moreover, if $\triangle\mathbf{X}_N = 0$ and $\triangle\mathbf{S}_B = 0$, then $\triangle\mathbf{X}_V \bullet \triangle\mathbf{S}_V = 0$.*

*Proof.* (a) Because of the surjectivity of $\tilde{\mathbb{Q}}_2$ (stated in Lemma 3.3) we have that there exist matrices $\mathbf{V}_1, \mathbf{V}_2$ of dimension $|B| \times |N|$ such that

$$\tilde{\mathbb{Q}}_2 \begin{bmatrix} vec(\mathbf{V}_1) \\ vec(\mathbf{V}_2) \end{bmatrix} + \tilde{\mathbb{R}}_2 \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0. \tag{24}$$

Similarly, because of the surjectivity of $\tilde{\mathbb{P}}_1$ we have, that there exist symmetric matrices $\mathbf{U}_1, \mathbf{U}_2$ of dimensions $|B| \times |B|$ and $|N| \times |N|$ and a vector $u \in R^m$ such that

$$\tilde{\mathbb{P}}_1 \begin{bmatrix} svec(\mathbf{U}_1) \\ u \\ svec(\mathbf{U}_2) \end{bmatrix} + \tilde{\mathbb{Q}}_1 \begin{bmatrix} vec(\mathbf{V}_1) \\ vec(\mathbf{V}_2) \end{bmatrix} + \tilde{\mathbb{R}}_1 \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0. \tag{25}$$

The equations $(24), (25)$ together with the equation $\tilde{\mathbb{R}}_3 \triangle z = 0$ are equivalent to

$$\mathbb{P} \begin{bmatrix} svec(\mathbf{U}_1) \\ u \\ svec(\mathbf{U}_2) \end{bmatrix} + \mathbb{Q} \begin{bmatrix} vec(\mathbf{V}_1) \\ vec(\mathbf{V}_2) \end{bmatrix} + \mathbb{R} \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0.$$

which is the same as

$$\mathcal{A}\left( \begin{bmatrix} \mathbf{U}_1 & \mathbf{V}_1 \\ \mathbf{V}_1^T & \triangle \mathbf{X}_N \end{bmatrix} \right) = 0, \quad \mathcal{A}^*(u) + \begin{bmatrix} \triangle \mathbf{S}_B & \mathbf{V}_2 \\ \mathbf{V}_2^T & \mathbf{U}_2 \end{bmatrix} = 0. \tag{26}$$

The assumption

$$\tilde{\mathbb{P}}_1 \triangle v = \tilde{\mathbb{P}}_1 \begin{bmatrix} svec(\triangle \mathbf{X}_B) \\ \triangle y \\ svec(\triangle \mathbf{S}_N) \end{bmatrix} = 0$$

is equivalent to

$$\mathcal{A}\left( \begin{bmatrix} \triangle \mathbf{X}_B & 0 \\ 0 & 0 \end{bmatrix} \right) = 0, \quad \mathcal{A}^*(\triangle y) + \begin{bmatrix} 0 & 0 \\ 0 & \triangle \mathbf{S}_N \end{bmatrix} = 0. \tag{27}$$

From (26) and (27) it follows that

$$\begin{bmatrix} \mathbf{U}_1 & \mathbf{V}_1 \\ \mathbf{V}_1^T & \triangle \mathbf{X}_N \end{bmatrix} \bullet \begin{bmatrix} 0 & 0 \\ 0 & \triangle \mathbf{S}_N \end{bmatrix} = \triangle \mathbf{X}_N \bullet \triangle \mathbf{S}_N = 0$$

and

$$\begin{bmatrix} \triangle \mathbf{S}_B & \mathbf{V}_2 \\ \mathbf{V}_2^T & \mathbf{U}_2 \end{bmatrix} \bullet \begin{bmatrix} \triangle \mathbf{X}_B & 0 \\ 0 & 0 \end{bmatrix} = \triangle \mathbf{X}_B \bullet \triangle \mathbf{S}_B = 0$$

due to the orthogonality property. Finally we have to show that $\triangle \mathbf{S}_V \bullet \triangle \mathbf{X}_V = 0$. From the surjectivity of $\tilde{\mathbb{P}}_1$ (see Lemma 3.3) we have that there exist symmetric matrices $\mathbf{V}_3, \mathbf{V}_4$ of dimensions $|B| \times |B|$ and $|N| \times |N|$ and a vector $v \in R^m$ such that

$$\tilde{\mathbb{P}}_1 \begin{bmatrix} svec\mathbf{V}_3 \\ v \\ svec\mathbf{V}_4 \end{bmatrix} + \tilde{\mathbb{Q}}_1 \begin{bmatrix} vec\triangle \mathbf{X}_V \\ vec\triangle \mathbf{S}_V \end{bmatrix} + \tilde{\mathbb{R}}_1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0.$$

This equation, together with $\tilde{\mathbb{Q}}_2 \triangle w + \tilde{\mathbb{R}}_3 0 = 0, \tilde{\mathbb{R}}_3 0 = 0$ implies

$$\mathbb{P} \begin{bmatrix} svec(\mathbf{V}_3) \\ v \\ svec(\mathbf{V}_4) \end{bmatrix} + \mathbb{Q} \begin{bmatrix} vec(\triangle \mathbf{X}_V) \\ vec(\triangle \mathbf{S}_V) \end{bmatrix} + \mathbb{R} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0.$$

The equation above can be rewritten as

$$\mathcal{A}\left( \begin{bmatrix} \mathbf{V}_3 & \triangle \mathbf{X}_V \\ \triangle \mathbf{X}_V^T & 0 \end{bmatrix} \right) = 0, \qquad \mathcal{A}^*(v) + \begin{bmatrix} 0 & \triangle \mathbf{S}_V \\ \triangle \mathbf{S}_V^T & \mathbf{V}_4 \end{bmatrix} = 0.$$

The orthogonality property implies

$$\begin{bmatrix} \mathbf{V}_3 & \triangle \mathbf{X}_V \\ \triangle \mathbf{X}_V^T & 0 \end{bmatrix} \bullet \begin{bmatrix} 0 & \triangle \mathbf{S}_V \\ \triangle \mathbf{S}_V^T & \mathbf{V}_4 \end{bmatrix} = 0$$

and hence $\triangle \mathbf{X}_V \bullet \triangle \mathbf{S}_V = 0$.

(b) Because of the surjectivity of $\tilde{\mathbb{P}}_1$ we have, that there exist symmetric matrices $\mathbf{W}_1, \mathbf{W}_2$ of dimensions $|B| \times |B|$ and $|N| \times |N|$ and a vector $w \in R^m$ such that

$$\tilde{\mathbb{P}}_1 \begin{bmatrix} svec(\mathbf{W}_1) \\ w \\ svec(\mathbf{W}_2) \end{bmatrix} + \tilde{\mathbb{Q}}_1 \begin{bmatrix} vec(\triangle \mathbf{X}_V) \\ vec(\triangle \mathbf{S}_V) \end{bmatrix} + \tilde{\mathbb{R}}_1 \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0. \tag{28}$$

The equation (28) together with $\tilde{\mathbb{Q}}_2 \triangle w + \tilde{\mathbb{R}}_2 \triangle z = 0, \tilde{\mathbb{R}}_3 \triangle z = 0$ is equivalent to

$$\mathbb{P} \begin{bmatrix} svec(\mathbf{W}_1) \\ u \\ svec(\mathbf{W}_2) \end{bmatrix} + \mathbb{Q} \begin{bmatrix} vec(\triangle \mathbf{X}_V) \\ vec(\triangle \mathbf{S}_V) \end{bmatrix} + \mathbb{R} \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0.$$

which is the same as

$$\mathcal{A}\left( \begin{bmatrix} \mathbf{W}_1 & \triangle \mathbf{X}_V \\ \triangle \mathbf{X}_V^T & \triangle \mathbf{X}_N \end{bmatrix} \right) = 0, \quad \mathcal{A}^*(u) + \begin{bmatrix} \triangle \mathbf{S}_B & \triangle \mathbf{S}_V \\ \triangle \mathbf{S}_V^T & \mathbf{W}_2 \end{bmatrix} = 0. \tag{29}$$

The equalities $\triangle \mathbf{X}_B \bullet \triangle \mathbf{S}_B = 0$, $\triangle \mathbf{X}_N \bullet \triangle \mathbf{S}_N = 0$ can be proved similarly as in the case (a). Assume $\triangle \mathbf{S}_B = 0$ and $\triangle \mathbf{X}_N = 0$. The orthogonality property and (29) yield

$$0 = \begin{bmatrix} \mathbf{W}_1 & \triangle \mathbf{X}_V \\ \triangle \mathbf{X}_V^T & \triangle \mathbf{X}_N \end{bmatrix} \bullet \begin{bmatrix} \triangle \mathbf{S}_B & \triangle \mathbf{S}_V \\ \triangle \mathbf{S}_V^T & \mathbf{W}_2 \end{bmatrix} =$$

$$= \mathbf{W}_1 \bullet \triangle \mathbf{S}_B + 2 \triangle \mathbf{X}_V \bullet \triangle \mathbf{S}_V + \triangle \mathbf{X}_N \bullet \mathbf{W}_2 = 2 \triangle \mathbf{X}_V \bullet \triangle \mathbf{S}_V.$$

■

### 3.3 Normalization of feasibility conditions in phase I

Define

$$\tilde{v}(\rho) = \begin{bmatrix} svec(\tilde{\mathbf{X}}_B(\rho)) \\ \tilde{y}(\rho) \\ svec(\tilde{\mathbf{S}}_N(\rho)) \end{bmatrix}, \quad \tilde{w}(\rho) = \begin{bmatrix} vec(\tilde{\mathbf{X}}_V(\rho)) \\ vec(\tilde{\mathbf{S}}_V(\rho)) \end{bmatrix}, \quad \tilde{z}(\rho) = \begin{bmatrix} svec(\tilde{\mathbf{X}}_N(\rho)) \\ svec(\tilde{\mathbf{S}}_B(\rho)) \end{bmatrix},$$

(where $\tilde{y}(\rho) = y(\mu)$). By inserting the normalized matrices into the system (23) we obtain

$$\begin{aligned} \tilde{\mathbb{P}}_1\tilde{v}(\rho) + \rho\tilde{\mathbb{Q}}_1\tilde{w}(\rho) + \rho^2\tilde{\mathbb{R}}_1\tilde{z}(\rho) &= \tilde{d}_1 + \rho^2\triangle\tilde{d}_1, \\ \rho\tilde{\mathbb{Q}}_2\tilde{w}(\rho) + \rho^2\tilde{\mathbb{R}}_2\tilde{z}(\rho) &= \tilde{d}_2 + \rho^2\triangle\tilde{d}_2, \\ \rho^2\tilde{\mathbb{R}}_3\tilde{z}(\rho) &= \tilde{d}_3 + \rho^2\triangle\tilde{d}_3. \end{aligned} \tag{30}$$

Inserting $\rho = \rho_k$, $\tilde{\mathbf{X}}(\rho) = \tilde{\mathbf{X}}(\rho_k)$, $\tilde{y}(\rho) = \tilde{y}(\rho_k)$ and $\tilde{\mathbf{S}}(\rho) = \tilde{\mathbf{S}}(\rho_k)$ into the system (30) and letting $\rho_k \to 0$ we find that $\tilde{d}_2 = 0, \tilde{d}_3 = 0$.
Define the map $\Psi$ in the following way:

$$\Psi(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = \begin{bmatrix} \tilde{\mathbb{P}}_1\tilde{v} + \rho\tilde{\mathbb{Q}}_1\tilde{w} + \rho^2\tilde{\mathbb{R}}_1\tilde{z} - \tilde{d}_1 + \rho^2\triangle\tilde{d}_1 \\ \tilde{\mathbb{Q}}_2\tilde{w} + \rho\tilde{\mathbb{R}}_2\tilde{z} - \rho\triangle\tilde{d}_2 \\ \tilde{\mathbb{R}}_3\tilde{z} - \triangle\tilde{d}_3 \end{bmatrix}, \tag{31}$$

where

$$\tilde{v} = \begin{bmatrix} svec(\tilde{\mathbf{X}}_B) \\ \tilde{y} \\ svec(\tilde{\mathbf{S}}_N) \end{bmatrix}, \quad \tilde{w} = \begin{bmatrix} vec(\tilde{\mathbf{X}}_V) \\ vec(\tilde{\mathbf{S}}_V) \end{bmatrix}, \quad \tilde{z} = \begin{bmatrix} svec(\tilde{\mathbf{X}}_N) \\ svec(\tilde{\mathbf{S}}_B) \end{bmatrix}.$$

It can be easily seen, that the Fréchet derivative of $\Psi$ is

$$D\Psi(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] = \begin{bmatrix} \tilde{\mathbb{P}}_1\triangle\tilde{v} \\ \tilde{\mathbb{Q}}_2\triangle\tilde{w} \\ \tilde{\mathbb{R}}_3\triangle\tilde{z} \end{bmatrix},$$

where

$$\triangle\tilde{v} = \begin{bmatrix} svec(\triangle\tilde{\mathbf{X}}_B) \\ \triangle\tilde{y} \\ svec(\triangle\tilde{\mathbf{S}}_N) \end{bmatrix}, \quad \triangle\tilde{w} = \begin{bmatrix} vec(\triangle\tilde{\mathbf{X}}_V) \\ vec(\triangle\tilde{\mathbf{S}}_V) \end{bmatrix}, \quad \triangle\tilde{z} = \begin{bmatrix} svec(\triangle\tilde{\mathbf{X}}_N) \\ svec(\triangle\tilde{\mathbf{S}}_B) \end{bmatrix}.$$

### 3.4 Analyticity of the paths as functions of $\sqrt{\mu}$ at $\mu = 0$

**Proposition 3.4** *Let $j \in \{2, 3\}$. Then the weighted path $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ is an analytic function of $\rho = \sqrt{\mu}$ for all $\mu \geq 0$ (sufficiently small).*

*Proof.* The statement for $j = 2$ was proved in [9] (see Theorem 3.3). We now prove the statement for $j = 3$, i.e. for the Cholesky paths. We fix a weight $\mathbf{W}$ from $\mathcal{M}_{\frac{1}{\sqrt{2}}}$ or $D^n_{++}$. Then the associated weighted central path satisfies (3) and (6)

$$\mathbf{L}_{\mathbf{X}}(\mu)^T\mathbf{S}\mathbf{L}_{\mathbf{X}}(\mu) = \mu\mathbf{W}$$

which can be rewritten equivalently as the pair

$$\begin{array}{c}\mathbf{L}(\mu)^T\mathbf{S}(\mu)\mathbf{L}(\mu) = \mu\mathbf{W}, \\ \mathbf{L}(\mu)^T\mathbf{L}(\mu) = \mathbf{X}(\mu).\end{array} \tag{32}$$

Setting the normalized matrices $\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{S}}(\rho), \tilde{\mathbf{L}}(\rho)$ defined in (10) and (11) into (32) and dividing some block equation by $\rho^2$ we obtain

$$\left.\begin{array}{rcl}\tilde{\mathbf{L}}(\rho)^T\tilde{\mathbf{S}}(\rho)\tilde{\mathbf{L}}(\rho) & = & \mathbf{W} \\ \tilde{\mathbf{L}}(\rho)\tilde{\mathbf{L}}(\rho)^T & = & \tilde{\mathbf{X}}(\rho)\end{array}\right\} \tag{33}$$

which holds for any $\rho = \sqrt{\mu} > 0$ sufficiently small. Considering (33) along the sequence $\{\rho_k\}$ defined in (12) and taking the limits $\rho_k \to 0$ we obtain that

$$\begin{array}{c}(\tilde{\mathbf{L}}^*)^T\tilde{\mathbf{S}}^*\tilde{\mathbf{L}}^* = \mathbf{W}, \\ \tilde{\mathbf{L}}^*(\tilde{\mathbf{L}}^*)^T = \tilde{\mathbf{X}}^*.\end{array} \tag{34}$$

Recall that $\tilde{\mathbf{L}}^* \in L^n_{++}$ and $\tilde{\mathbf{S}}^*, \tilde{\mathbf{X}}^* \succ 0$.

Define the map $\tilde{F}^3 : S^n \times L^n \times R^m \times S^n \times R \to R^m \times S^n \times S^n \times S^n$ in the following way:

$$\tilde{F}^3(\tilde{\mathbf{X}}, \tilde{\mathbf{L}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = \begin{bmatrix}\Psi(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) \\ \tilde{\mathbf{L}}^T\tilde{\mathbf{S}}\tilde{\mathbf{L}} - \mathbf{W} \\ \tilde{\mathbf{L}}\tilde{\mathbf{L}}^T - \tilde{\mathbf{X}}\end{bmatrix},$$

where $\Psi$ is defined by (31). Obviously $\tilde{F}^3$ is an analytic function of $(\tilde{\mathbf{X}}, \tilde{\mathbf{L}}, \tilde{y}, \tilde{\mathbf{S}}, \rho)$. Moreover, for $\rho > 0$ (sufficiently small) it holds

$$\tilde{F}^3(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{L}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho), \rho) = 0$$

and

$$\tilde{F}^3(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0) = 0.$$

Here it is important that for any $\rho > 0$ sufficiently small, $(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{L}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ is the unique solution to $\tilde{F}^3(\tilde{\mathbf{X}}, \tilde{\mathbf{L}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = 0$ on an open set $S^n_{++} \times L^n_{++} \times R^m \times S^n_{++}$. Moreover, $(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*) \in S^n_{++} \times L^n_{++} \times R^m \times S^n_{++}$. Note that the uniqueness is a consequence of the uniqueness declared in Proposition 2.1.

The Fréchet derivative of $\tilde{F}^3$ with respect to $(\tilde{\mathbf{X}}, \tilde{\mathbf{L}}, \tilde{y}, \tilde{\mathbf{S}})$ at the point $(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ is the linear map given as

$$D\tilde{F}^3(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{\mathbf{L}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] =$$

$$= \begin{bmatrix} D\Psi(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] \\ \triangle\tilde{\mathbf{L}}\tilde{\mathbf{S}}^*(\tilde{\mathbf{L}}^*)^T + \tilde{\mathbf{L}}^*\triangle\tilde{\mathbf{S}}(\tilde{\mathbf{L}}^*)^T + \tilde{\mathbf{L}}^*\tilde{\mathbf{S}}^*(\triangle\tilde{\mathbf{L}})^T \\ (\triangle\tilde{\mathbf{L}})^T\tilde{\mathbf{L}}^* + (\tilde{\mathbf{L}}^*)^T\triangle\tilde{\mathbf{L}} - \triangle\tilde{\mathbf{X}} \end{bmatrix}.$$

We now show that $D\tilde{F}^3(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ is a nonsingular linear map. To this aim we assume

$$D\tilde{F}^3(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{\mathbf{L}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] = 0.$$

From Proposition 3.3 we have that $\triangle\tilde{\mathbf{X}} \bullet \triangle\tilde{\mathbf{S}} = 0$. Due to (34), the nonsingularity of $D\tilde{F}^3$ for the case $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ follows from Lemma 2.4 of [14] and for $\mathbf{W} \in D^n_{++}$ was shown in [20] (see the proof of Theorem 1).

Having the nonsingularity of $D\tilde{F}^3$ we are ready to apply the implicit function theorem to $\tilde{F}^3(\tilde{\mathbf{X}}, \tilde{\mathbf{L}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = 0$ at the point $(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$. We obtain that there exist neighborhoods $\mathcal{I}$ and $\mathcal{U}$ of $\rho = 0$ and $(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*)$ respectively and an analytic function

$$(\hat{\mathbf{X}}, \hat{\mathbf{L}}, \hat{y}, \hat{\mathbf{S}}) : \mathcal{I} \to \mathcal{U}$$

such that

$$(\hat{\mathbf{X}}(0), \hat{\mathbf{L}}(0), \hat{y}(0), \hat{\mathbf{S}}(0)) = (\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*)$$

and

$$\tilde{F}^3(\hat{\mathbf{X}}(\rho), \hat{\mathbf{L}}(\rho), \hat{y}(\rho), \hat{\mathbf{S}}(\rho), \rho) = 0 \quad \forall \rho \in \mathcal{I}. \tag{35}$$

Since $(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*)$ belongs to the open set $S^n_{++} \times L^n_{++} \times R^m \times S^n_{++}$ we can restrict the neighborhoods $\mathcal{U}$ and $\mathcal{I}$ if necessary in such a way that $\mathcal{U} \subset S^n_{++} \times L^n_{++} \times R^m \times S^n_{++}$. Now both $(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{L}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ and $(\hat{\mathbf{X}}(\rho), \hat{L}(\rho), \hat{y}(\rho), \hat{\mathbf{S}}(\rho))$ are solutions to (35) for $\rho > 0$ and whence $(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{L}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho)) = (\hat{\mathbf{X}}(\rho), \hat{L}(\rho), \hat{y}(\rho), \hat{\mathbf{S}}(\rho))$ at any $\rho \in \mathcal{I} \cap (0, \infty)$ by the uniqueness of solutions on $S^n_{++} \times L^n_{++} \times R^m \times S^n_{++}$. Thus the function $(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{L}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ is analytically extendable to $\rho = 0$ by prescription

$$(\tilde{\mathbf{X}}(0), \tilde{\mathbf{L}}(0), \tilde{y}(0), \tilde{\mathbf{S}}(0)) = (\hat{\mathbf{X}}(0), \hat{L}(0), \hat{y}(0), \hat{\mathbf{S}}(0)) = (\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*).$$

Due to (10) also the function $(\mathbf{X}(\rho), \mathbf{L}(\rho), y(\rho), \mathbf{S}(\rho))$ is analytically extendable to $\rho = 0$. ■

# 4 Phase II: paths as functions of $\mu$

## 4.1 Introduction of new normalized matrices in phase II

**Lemma 4.1** *Let $j \in \{2, 3\}$ be arbitrary and assume $\mathbf{W}_V = 0$. Then*

$$\mathbf{X}_V(\mu) = \mathcal{O}(\mu), \ \mathbf{S}_V(\mu) = \mathcal{O}(\mu).$$

*Moreover, if $j = 2$, then $\mathbf{Y}_V(\mu) = \mathcal{O}(\mu)$ and if $j = 3$, then $\mathbf{L}_V(\mu) = \mathcal{O}(\mu)$.*

*Proof.* Since $\lim_{\rho \to 0}(\tilde{\mathbf{X}}_V(\rho), \tilde{\mathbf{S}}_V(\rho)) = (0, 0)$, the Taylor series expansions of $\tilde{\mathbf{X}}_V(\rho)$, $\tilde{\mathbf{S}}_V(\rho)$, which are analytic functions of $\rho$ for $\rho \geq 0$ sufficiently small, have the form

$$\tilde{\mathbf{X}}_V(\rho) = \rho \sum_{i=0}^{\infty} \mathbf{P}_i \rho^i, \qquad \tilde{\mathbf{S}}_V(\rho) = \rho \sum_{i=0}^{\infty} \mathbf{Q}_i \rho^i.$$

This implies

$$\mathbf{X}_V(\mu) = \rho \tilde{\mathbf{X}}_V(\rho) = \rho \mathcal{O}(\rho) = \mathcal{O}(\rho^2) = \mathcal{O}(\mu). \tag{36}$$

Similarly, it can be shown that $\mathbf{S}_V(\mu) = \mathcal{O}(\mu)$.

Moreover, assume $j = 2$. It holds

$$\mathbf{X}_V(\mu) = \mathbf{Y}_B(\mu)\mathbf{Y}_V(\mu) + \mathbf{Y}_V(\mu)\mathbf{Y}_N(\mu).$$

From the asymptotic behavior given in Proposition 3.1 ($\mathbf{Y}_B(\mu) = \Theta(1)$) it follows that $\mathbf{Y}_B(\mu)^{-1} = \mathcal{O}(1)$. Therefore we have

$$\|\mathbf{Y}_V(\mu)\|_F = \|\mathbf{Y}_B(\mu)^{-1}\mathbf{X}_V(\mu) - \mathbf{Y}_B(\mu)^{-1}\mathbf{Y}_V(\mu)\mathbf{Y}_N(\mu)\|_F \leq$$

$$\leq \|\mathbf{Y}_B(\mu)^{-1}\mathbf{X}_V(\mu)\|_F + \|\mathbf{Y}_B(\mu)^{-1}\mathbf{Y}_V(\mu)\mathbf{Y}_N(\mu)\|_F = \mathcal{O}(\mu)$$

where the last equality follows from (36) and Proposition 3.1.

Finally, if $j = 3$, by $\mathbf{X}_V(\mu) = \mathbf{L}_B(\mu)\mathbf{L}_V(\mu)$ and the asymptotic behavior $\mathbf{X}_V(\mu) = \mathcal{O}(\mu)$ and $\mathbf{L}_B(\mu) = \Theta(1)$ we obtain $\mathbf{L}_V(\mu) = \mathcal{O}(\mu)$. ∎

From now we will assume that $\mathbf{W}_V = 0$.

From Lemma 4.1 it follows that the path matrices posses the following asymptotic behavior:

$$\mathbf{X}(\mu) = \begin{pmatrix} \Theta(1) & \mathcal{O}(\mu) \\ \mathcal{O}(\mu) & \Theta(\mu) \end{pmatrix}, \quad \mathbf{S}(\mu) = \begin{pmatrix} \Theta(\mu) & \mathcal{O}(\mu) \\ \mathcal{O}(\mu) & \Theta(1) \end{pmatrix}. \tag{37}$$

Moreover, for $\mathbf{L}(\mu) = \mathbf{L}_{\mathbf{X}(\mu)}$ and $\mathbf{Y}(\mu) = [\mathbf{X}(\mu)]^{\frac{1}{2}}$ we obtain

$$\mathbf{L}(\mu) = \begin{pmatrix} \Theta(1) & 0 \\ \mathcal{O}(\mu) & \Theta(\sqrt{\mu}) \end{pmatrix}, \quad \mathbf{Y}(\mu) = \begin{pmatrix} \Theta(1) & \mathcal{O}(\mu) \\ \mathcal{O}(\mu) & \Theta(\sqrt{\mu}) \end{pmatrix}. \tag{38}$$

This asymptotic behavior naturally implies the following definition new normalized matrices:

$$\bar{\mathbf{X}}(\mu) := \begin{pmatrix} \mathbf{X}_B(\mu) & \mathbf{X}_V(\mu)/\mu \\ \mathbf{X}_V(\mu)^T/\mu & \mathbf{X}_N(\mu)/\mu \end{pmatrix}, \quad \bar{\mathbf{S}}(\mu) = \begin{pmatrix} \mathbf{S}_B(\mu)/\mu & \mathbf{S}_V(\mu)/\mu \\ \mathbf{S}_V(\mu)^T/\mu & \mathbf{S}_N(\mu) \end{pmatrix}$$

and

$$\bar{\mathbf{L}}(\mu) := \begin{pmatrix} \mathbf{L}_B(\mu) & 0 \\ \mathbf{L}_V(\mu)^T/\mu & \mathbf{L}_N(\mu)/\sqrt{\mu} \end{pmatrix}, \quad \bar{\mathbf{Y}}(\mu) := \begin{pmatrix} \mathbf{Y}_B(\mu) & \mathbf{Y}_V(\mu)/\mu \\ \mathbf{Y}_V(\mu)^T/\mu & \mathbf{Y}_N(\mu)/\sqrt{\mu} \end{pmatrix}.$$

## 4.2 Normalization of feasibility conditions in phase II

Define

$$\bar{v} = \begin{bmatrix} svec(\bar{\mathbf{X}}_B) \\ \bar{y} \\ svec(\bar{\mathbf{S}}_N) \end{bmatrix}, \quad \bar{w} = \begin{bmatrix} vec(\bar{\mathbf{X}}_V) \\ vec(\bar{\mathbf{S}}_V) \end{bmatrix}, \quad \bar{z} = \begin{bmatrix} svec(\bar{\mathbf{X}}_N) \\ svec(\bar{\mathbf{S}}_B) \end{bmatrix}$$

and rewrite the system (23) using the new normalized matrices:

$$\begin{aligned} \tilde{\mathbb{P}}_1\bar{v} + \mu\tilde{\mathbb{Q}}_1\bar{w} + \mu\tilde{\mathbb{R}}_1\bar{z} &= \tilde{d}_1 + \mu\triangle\tilde{d}_1, \\ \mu\tilde{\mathbb{Q}}_2\bar{w} + \mu\tilde{\mathbb{R}}_2\bar{z} &= \tilde{d}_2 + \mu\triangle\tilde{d}_2, \\ \mu\tilde{\mathbb{R}}_3\bar{z} &= \tilde{d}_3 + \mu\triangle\tilde{d}_3. \end{aligned} \tag{39}$$

From the asymptotic behavior given in (37) it follows that

$$\bar{\mathbf{X}}(\mu) = \mathcal{O}(1), \qquad \bar{\mathbf{S}}(\mu) = \mathcal{O}(1)$$

and therefore for any sequence $\{\mu_k\} \to 0$ the matrices $\bar{\mathbf{X}}(\mu_k), \bar{\mathbf{S}}(\mu_k)$ and also the associated vector $y(\mu_k) = \bar{y}(\mu_k)$ are bounded, so we may assume that the limit

$$\lim_{k\to\infty} (\bar{\mathbf{X}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k)) = (\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*)$$

exists. Inserting $\mu = \mu_k$, $\bar{\mathbf{X}} = \bar{\mathbf{X}}(\mu_k)$, $\bar{y} = \bar{y}(\mu_k)$, $\bar{\mathbf{S}} = \bar{\mathbf{S}}(\mu_k)$ into the system (39) and letting $k \to \infty$ we obtain that $\tilde{d}_2 = 0$ and $\tilde{d}_3 = 0$. Hence the system (39) has the form

$$\begin{aligned} \tilde{\mathbb{P}}_1\bar{v} + \mu\tilde{\mathbb{Q}}_1\bar{w} + \mu\tilde{\mathbb{R}}_1\bar{z} &= \tilde{d}_1 + \mu\triangle\tilde{d}_1, \\ \tilde{\mathbb{Q}}_2\bar{w} + \tilde{\mathbb{R}}_2\bar{z} &= \triangle\tilde{d}_2, \\ \tilde{\mathbb{R}}_3\bar{z} &= \triangle\tilde{d}_3. \end{aligned}$$

Define the map $\bar{\Psi}$ in the following way

$$\bar{\Psi}(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}}, \mu) = \begin{bmatrix} \tilde{\mathbb{P}}_1\bar{v} + \mu\tilde{\mathbb{Q}}_1\bar{w} + \mu\tilde{\mathbb{R}}_1\bar{z} - \tilde{d}_1 + \mu\triangle\tilde{d}_1 \\ \tilde{\mathbb{Q}}_2\bar{w} + \tilde{\mathbb{R}}_2\bar{z} - \triangle\tilde{d}_2 \\ \tilde{\mathbb{R}}_3\bar{z} - \triangle\tilde{d}_3 \end{bmatrix}. \tag{40}$$

The Fréchet derivative of $\bar{\Psi}$ with respect to variables $(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}})$ at the point $(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)$ is

$$D\bar{\Psi}(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = \begin{bmatrix} \tilde{\mathbb{P}}_1\triangle\bar{v} \\ \tilde{\mathbb{Q}}_2\triangle\bar{w} + \tilde{\mathbb{R}}_2\triangle\bar{z} \\ \tilde{\mathbb{R}}_3\triangle\bar{z} \end{bmatrix}$$

where

$$\triangle\bar{v} = \begin{bmatrix} svec(\triangle\bar{\mathbf{X}}_B) \\ \triangle\bar{y} \\ svec(\triangle\bar{\mathbf{S}}_N) \end{bmatrix}, \quad \triangle\bar{w} = \begin{bmatrix} vec(\triangle\bar{\mathbf{X}}_V) \\ vec(\triangle\bar{\mathbf{S}}_V) \end{bmatrix}, \quad \triangle\bar{z} = \begin{bmatrix} svec(\triangle\bar{\mathbf{X}}_N) \\ svec(\triangle\bar{\mathbf{S}}_B) \end{bmatrix}.$$

### 4.3 Nonsingularity of the Fréchet derivatives in phase II

In this section we define new normalized maps associated with the both symmetrizations $\Phi_2$ and $\Phi_3$. Under the assumption that the weight matrix is block diagonal we show that these maps are nonsingular at the limit point which will be useful to prove the analyticity of the paths as a function of $\mu$ at this point.

Firstly, we will consider the condition

$$\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}} = \mu\mathbf{W}$$

given in (3), which can be equivalently rewritten as the pair

$$\begin{aligned} \mathbf{YSY} &= \mu\mathbf{W}, \\ \mathbf{Y}^2 &= \mathbf{X}. \end{aligned}$$

The block form of this equalities is:

$$\begin{gathered}
\mathbf{Y}_B\mathbf{S}_B\mathbf{Y}_B + \mathbf{Y}_V\mathbf{S}_V^T\mathbf{Y}_B + \mathbf{Y}_B\mathbf{S}_V\mathbf{Y}_V^T + \mathbf{Y}_V\mathbf{S}_N\mathbf{Y}_V^T = \mu\mathbf{W}_B,\\
\mathbf{Y}_B\mathbf{S}_B\mathbf{Y}_V + \mathbf{Y}_V\mathbf{S}_V^T\mathbf{Y}_V + \mathbf{Y}_B\mathbf{S}_V\mathbf{Y}_N + \mathbf{Y}_V\mathbf{S}_N\mathbf{Y}_N = 0,\\
\mathbf{Y}_V^T\mathbf{S}_B\mathbf{Y}_V + \mathbf{Y}_N\mathbf{S}_V^T\mathbf{Y}_V + \mathbf{Y}_V^T\mathbf{S}_V\mathbf{Y}_N + \mathbf{Y}_N\mathbf{S}_N\mathbf{Y}_N = \mu\mathbf{W}_N,\\
\mathbf{Y}_B^2 + \mathbf{Y}_V\mathbf{Y}_V^T = \mathbf{X}_B,\\
\mathbf{Y}_B\mathbf{Y}_V + \mathbf{Y}_V\mathbf{Y}_N = \mathbf{X}_V,\\
\mathbf{Y}^2 + \mathbf{Y}_V^T\mathbf{Y}_V = \mathbf{X}_N.
\end{gathered}$$

Because $(\mathbf{X}(\mu), \mathbf{Y}(\mu), y(\mu), \mathbf{S}(\mu))$ satisfies the system above for $\mu > 0$ (sufficiently small), we have that the triple $(\bar{\mathbf{X}}(\mu), \bar{\mathbf{Y}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ satisfies

$$\begin{gathered}
\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{Y}}_B(\mu) + \mu\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{Y}}_B(\mu)\\
+\mu\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{Y}}_V(\mu)^T + \mu\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{Y}}_V(\mu)^T = \mathbf{W}_B,\\
\sqrt{\mu}\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{Y}}_V(\mu) + \mu\sqrt{\mu}\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{Y}}_V(\mu)+\\
\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{Y}}_N(\mu) + \bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{Y}}_N(\mu) = 0,\\
\mu^2\bar{\mathbf{Y}}_V(\mu)^T\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{Y}}_V(\mu) + \mu\bar{\mathbf{Y}}_N(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{Y}}_V(\mu)+\\
\mu\bar{\mathbf{Y}}_V(\mu)^T\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{Y}}_N(\mu) + \bar{\mathbf{Y}}_N(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{Y}}_N(\mu) = \mathbf{W}_N,\\
\bar{\mathbf{Y}}_B(\mu)^2 + \mu^2\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{Y}}_V(\mu)^T = \bar{\mathbf{X}}_B(\mu),\\
\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{Y}}_V(\mu) + \sqrt{\mu}\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{Y}}_N(\mu) = \bar{\mathbf{X}}_V(\mu),\\
\bar{\mathbf{Y}}_N(\mu)^2 + \mu\bar{\mathbf{Y}}_V(\mu)^T\bar{\mathbf{Y}}_V(\mu) = \bar{\mathbf{X}}_N(\mu),
\end{gathered}\tag{41}$$

From (37) and (38) it follows that for any sequence $\{\mu_k\} \to 0$ the sequence

$$(\bar{\mathbf{X}}(\mu_k), \bar{\mathbf{Y}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k))$$

is bounded, so we may assume that the limit

$$\lim_{k\to\infty}(\bar{\mathbf{X}}(\mu_k), \bar{\mathbf{Y}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k)) = (\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*)$$

exists. Define the map $\bar{F}^2$ in the following way

$$\bar{F}^2(\bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{y}, \bar{\mathbf{S}}, \mu) = \begin{bmatrix}
\bar{\Psi}(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}}, \mu)\\
\bar{\mathbf{Y}}_B\bar{\mathbf{S}}_B\bar{\mathbf{Y}}_B + \mu\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_V^T\bar{\mathbf{Y}}_B + \mu\bar{\mathbf{Y}}_B\bar{\mathbf{S}}_V\bar{\mathbf{Y}}_V^T + \mu\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N\bar{\mathbf{Y}}_V^T - \mathbf{W}_B\\
\sqrt{\mu}\bar{\mathbf{Y}}_B\bar{\mathbf{S}}_B\bar{\mathbf{Y}}_V + \mu\sqrt{\mu}\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_V^T\bar{\mathbf{Y}}_V + \bar{\mathbf{Y}}_B\bar{\mathbf{S}}_V\bar{\mathbf{Y}}_N + \bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N\bar{\mathbf{Y}}_N\\
\mu^2\bar{\mathbf{Y}}_V^T\bar{\mathbf{S}}_B\bar{\mathbf{Y}}_V + \mu\bar{\mathbf{Y}}_N\bar{\mathbf{S}}_V^T\bar{\mathbf{Y}}_V + \mu\bar{\mathbf{Y}}_V^T\bar{\mathbf{S}}_V\bar{\mathbf{Y}}_N + \bar{\mathbf{Y}}_N\bar{\mathbf{S}}_N\bar{\mathbf{Y}}_N - \mathbf{W}_N\\
\bar{\mathbf{Y}}_B^2 + \mu^2\bar{\mathbf{Y}}_V\bar{\mathbf{Y}}_V^T - \bar{\mathbf{X}}_B\\
\bar{\mathbf{Y}}_B\bar{\mathbf{Y}}_V + \sqrt{\mu}\bar{\mathbf{Y}}_V\bar{\mathbf{Y}}_N - \bar{\mathbf{X}}_V\\
\bar{\mathbf{Y}}_N^2 + \mu\bar{\mathbf{Y}}_V^T\bar{\mathbf{Y}}_V - \bar{\mathbf{X}}_N
\end{bmatrix},$$

where $\bar{\Psi}$ is defined in (40). For $\mu > 0$ sufficiently small it holds that

$$\bar{F}^2(\bar{\mathbf{X}}(\mu), \bar{\mathbf{Y}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu), \mu) = 0 \tag{42}$$

and

$$\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0) = 0. \tag{43}$$

The Fréchet derivative of $\bar{F}^2$ with respect to $(\bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{y}, \bar{\mathbf{S}})$ at the point $(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)$ is the linear map given by

$$D\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{Y}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] =$$

$$\begin{bmatrix} D\bar{\Phi}((\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] \\ \triangle\bar{\mathbf{Y}}_B\bar{\mathbf{S}}_B^*\bar{\mathbf{Y}}_B^* + \bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{S}}_B\bar{\mathbf{Y}}_B^* + \bar{\mathbf{Y}}_B^*\bar{\mathbf{S}}_B^*\triangle\bar{\mathbf{Y}}_B \\ \triangle\bar{\mathbf{Y}}_B\bar{\mathbf{S}}_V^*\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{S}}_V\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_B^*\bar{\mathbf{S}}_V^*\triangle\bar{\mathbf{Y}}_N + \triangle\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N^*\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_V^*\triangle\bar{\mathbf{S}}_N\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_V^*\bar{\mathbf{S}}_N^*\triangle\bar{\mathbf{Y}}_N \\ \triangle\bar{\mathbf{Y}}_N\bar{\mathbf{S}}_N^*\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_N^*\triangle\bar{\mathbf{S}}_N\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_N^*\bar{\mathbf{S}}_N^*\triangle\bar{\mathbf{Y}}_N \\ \triangle\bar{\mathbf{Y}}_B\bar{\mathbf{Y}}_B^* + \bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{Y}}_B - \triangle\bar{\mathbf{X}}_B \\ \triangle\bar{\mathbf{Y}}_B\bar{\mathbf{Y}}_V^* + \bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{Y}}_V - \triangle\bar{\mathbf{X}}_V \\ \triangle\bar{\mathbf{Y}}_N\bar{\mathbf{Y}}_N^* + \bar{\mathbf{Y}}_N^*\triangle\bar{\mathbf{Y}}_N - \triangle\bar{\mathbf{X}}_N \end{bmatrix}$$

**Lemma 4.2** *$D\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{Y}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}]$ is a nonsingular linear map.*

*Proof.* Assume

$$D\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{Y}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = 0. \tag{44}$$

We will show that $[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{Y}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = [0, 0, 0, 0]$. If we put $\mu = \mu_k$ in the system (41), then by taking the limit $k \to \infty$ we obtain that

$$\bar{\mathbf{Y}}_B^*\bar{\mathbf{S}}_B^*\bar{\mathbf{Y}}_B^* = \mathbf{W}_B, \qquad \bar{\mathbf{Y}}_N^*\bar{\mathbf{S}}_N^*\bar{\mathbf{Y}}_N^* = \mathbf{W}_N. \tag{45}$$

Because $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$, we have also that $\mathbf{W}_B \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ and $\mathbf{W}_N \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$. From the first equation of (44) and Proposition 3.3 (b) we have that

$$\triangle\bar{\mathbf{X}}_B \bullet \triangle\bar{\mathbf{S}}_B = 0, \quad \triangle\bar{\mathbf{X}}_N \bullet \triangle\bar{\mathbf{S}}_N = 0. \tag{46}$$

It follows from the proof of Lemma 2.3 of [13] that $\triangle\bar{\mathbf{X}}_B = 0$, $\triangle\bar{\mathbf{X}}_N = 0$, $\triangle\bar{\mathbf{Y}}_B = 0$, $\triangle\bar{\mathbf{Y}}_N = 0$, $\triangle\bar{\mathbf{S}}_B = 0$, $\triangle\bar{\mathbf{S}}_N = 0$. These equalities together with (44) imply

$$\bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{S}}_V\bar{\mathbf{Y}}_N^* + \triangle\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N^*\bar{\mathbf{Y}}_N^* = 0, \qquad \bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{Y}}_V - \triangle\bar{\mathbf{X}}_V = 0$$

or, since $\bar{\mathbf{Y}}_N^* \succ 0$,

$$\bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{S}}_V + \triangle\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N^* = 0, \qquad \bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{Y}}_V = \triangle\bar{\mathbf{X}}_V.$$

From Proposition 3.3 (b) we now have that $\triangle\bar{\mathbf{X}}_V \bullet \triangle\bar{\mathbf{S}}_V = 0$, which, together with the equalities above, yields

$$0 = -\triangle\bar{\mathbf{X}}_V \bullet \triangle\bar{\mathbf{S}}_V = -tr(\triangle\bar{\mathbf{X}}_V^T\triangle\bar{\mathbf{S}}_V) = -tr(\triangle\bar{\mathbf{Y}}_V^T\bar{\mathbf{Y}}_B^*\triangle\bar{\mathbf{S}}_V) =$$

$$= tr(\triangle\bar{\mathbf{Y}}_V^T\bar{\mathbf{Y}}_B^*(\bar{\mathbf{Y}}_B^*)^{-1}\triangle\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N^*) = tr(\triangle\bar{\mathbf{Y}}_V\bar{\mathbf{S}}_N^*\triangle\bar{\mathbf{Y}}_V^T).$$

The matrix in the last brace is positive semidefinite and hence from positive definitness of $\bar{\mathbf{S}}_N^*$ it follows that $\triangle\bar{\mathbf{Y}}_V = 0$. Therefore also $\triangle\bar{\mathbf{X}}_V = \triangle\bar{\mathbf{S}}_V = 0$. Assumption (A1) gives $\triangle\bar{y} = 0$. ■

Now we will use the same procedure to show the nonsingularity of the normalized map associated with $\Phi_3$. Consider the last condition from the system (1) – (3):

$$\mathbf{L_X}^T \mathbf{S} \mathbf{L_X} = \mu \mathbf{W}.$$

It can be equivalently replaced by the pair

$$\begin{gathered} \mathbf{L}^T \mathbf{S} \mathbf{L} = \mu \mathbf{W}, \\ \mathbf{L}\mathbf{L}^T = \mathbf{X}, \end{gathered}$$

which can be rewritten in the following block form:

$$\begin{gathered} \mathbf{L}_B^T \mathbf{S}_B \mathbf{L}_B + \mathbf{L}_V \mathbf{S}_V^T \mathbf{L}_B + \mathbf{L}_B^T \mathbf{S}_V^T \mathbf{L}_V^T + \mathbf{L}_V \mathbf{S}_N \mathbf{L}_V^T = \mu \mathbf{W}_B, \\ \mathbf{L}_B^T \mathbf{S}_V \mathbf{L}_N + \mathbf{L}_V \mathbf{S}_N \mathbf{L}_N = 0, \\ \mathbf{L}_N^T \mathbf{S}_N \mathbf{L}_N = \mu \mathbf{W}_N, \\ \mathbf{L}_B \mathbf{L}_B^T = \mathbf{X}_B, \\ \mathbf{L}_B \mathbf{L}_V = \mathbf{X}_V, \\ \mathbf{L}_N \mathbf{L}_N^T + \mathbf{L}_V^T \mathbf{L}_V = \mathbf{X}_N. \end{gathered}$$

Because $(\mathbf{X}(\mu), \mathbf{L}(\mu), y(\mu), \mathbf{S}(\mu))$ satisfies the system above for $\mu > 0$ (sufficiently small), we have that the triple $(\bar{\mathbf{X}}(\mu), \bar{\mathbf{L}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ satisfies

$$\begin{gathered} \bar{\mathbf{L}}_B(\mu)^T \bar{\mathbf{S}}_B(\mu) \bar{\mathbf{L}}_B(\mu) + \mu \bar{\mathbf{L}}_V(\mu) \bar{\mathbf{S}}_V(\mu)^T \bar{\mathbf{L}}_B(\mu) + \\ \mu \bar{\mathbf{L}}_B(\mu)^T \bar{\mathbf{S}}_V(\mu)^T \bar{\mathbf{L}}_V(\mu)^T + \mu \bar{\mathbf{L}}_V(\mu) \bar{\mathbf{S}}_N(\mu) \bar{\mathbf{L}}_V(\mu)^T = \mathbf{W}_B, \\ \bar{\mathbf{L}}_B(\mu)^T \bar{\mathbf{S}}_V(\mu) \bar{\mathbf{L}}_N(\mu) + \bar{\mathbf{L}}_V(\mu) \bar{\mathbf{S}}_N(\mu) \bar{\mathbf{L}}_N(\mu) = 0, \\ \bar{\mathbf{L}}_N(\mu)^T \bar{\mathbf{S}}_N(\mu) \bar{\mathbf{L}}_N(\mu) = \mu \mathbf{W}_N, \\ \bar{\mathbf{L}}_B(\mu) \bar{\mathbf{L}}_B(\mu)^T = \mathbf{X}_B(\mu), \\ \bar{\mathbf{L}}_B(\mu) \bar{\mathbf{L}}_V(\mu) = \mathbf{X}_V(\mu), \\ \bar{\mathbf{L}}_N(\mu) \bar{\mathbf{L}}_N(\mu)^T + \mu \bar{\mathbf{L}}_V(\mu)^T \bar{\mathbf{L}}_V(\mu) = \mathbf{X}_N(\mu). \end{gathered} \tag{47}$$

From (37) and (37) it follows that for any sequence $\{\mu_k\} \to 0$ the sequence

$$(\bar{\mathbf{X}}(\mu_k), \bar{\mathbf{L}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k))$$

is bounded, so we may assume that the limit

$$\lim_{k \to \infty} (\bar{\mathbf{X}}(\mu_k), \bar{\mathbf{L}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k)) = (\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*)$$

exists. Define the map $\bar{F}^3$ in the following way

$$\bar{F}^3(\bar{\mathbf{X}}, \bar{\mathbf{L}}, \bar{y}, \bar{\mathbf{S}}, \mu) = \begin{bmatrix} \bar{\Psi}(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}}, \mu) \\ \bar{\mathbf{L}}_B^T \bar{\mathbf{S}}_B \bar{\mathbf{L}}_B + \mu \bar{\mathbf{L}}_V \bar{\mathbf{S}}_V^T \bar{\mathbf{L}}_B + \mu \bar{\mathbf{L}}_B^T \bar{\mathbf{S}}_V^T \bar{\mathbf{L}}_V^T + \mu \bar{\mathbf{L}}_V \bar{\mathbf{S}}_N \bar{\mathbf{L}}_V^T - \mathbf{W}_B \\ \bar{\mathbf{L}}_B^T \bar{\mathbf{S}}_V \bar{\mathbf{L}}_N + \bar{\mathbf{L}}_V \bar{\mathbf{S}}_N \bar{\mathbf{L}}_N \\ \bar{\mathbf{L}}_N^T \bar{\mathbf{S}}_N \bar{\mathbf{L}}_N - \mu \mathbf{W}_N \\ \bar{\mathbf{L}}_B \bar{\mathbf{L}}_B^T - \mathbf{X}_B \\ \bar{\mathbf{L}}_B \bar{\mathbf{L}}_V - \mathbf{X}_V \\ \bar{\mathbf{L}}_N \bar{\mathbf{L}}_N^T + \mu \bar{\mathbf{L}}_V^T \bar{\mathbf{L}}_V - \mathbf{X}_N \end{bmatrix}$$

where $\bar{\Psi}$ is defined by (40). For $\mu > 0$ sufficiently small it holds, that

$$\bar{F}^3(\bar{\mathbf{X}}(\mu), \bar{\mathbf{L}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu), \mu) = 0 \tag{48}$$

and

$$\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0) = 0. \tag{49}$$

The Fréchet derivative of $\bar{F}^3$ with respect to $(\bar{\mathbf{X}}, \bar{\mathbf{L}}, \bar{y}, \bar{\mathbf{S}})$ at the point $\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)$ is the linear map given by

$$D\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{L}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] =$$

$$\begin{bmatrix} D\bar{\Phi}((\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] \\ \triangle\bar{\mathbf{L}}_B^T\bar{\mathbf{S}}_B^*\bar{\mathbf{L}}_B^* + (\bar{\mathbf{L}}_B^*)^T\triangle\bar{\mathbf{S}}_B\bar{\mathbf{L}}_B^* + (\bar{\mathbf{L}}_B^*)^T\bar{\mathbf{S}}_B^*\triangle\bar{\mathbf{L}}_B \\ \triangle\bar{\mathbf{L}}_B\bar{\mathbf{S}}_V^*\bar{\mathbf{L}}_N^* + (\bar{\mathbf{L}}_B^*)^T\triangle\bar{\mathbf{S}}_V\bar{\mathbf{L}}_N^* + (\bar{\mathbf{L}}_B^*)^T\bar{\mathbf{S}}_V^*\triangle\bar{\mathbf{L}}_N + \triangle\bar{\mathbf{L}}_V\bar{\mathbf{S}}_N^*\bar{\mathbf{L}}_N^* + (\bar{\mathbf{L}}_V^*)^T\triangle\bar{\mathbf{S}}_N\bar{\mathbf{L}}_N^* + (\bar{\mathbf{L}}_V^*)^T\bar{\mathbf{S}}_N^*\triangle\bar{\mathbf{L}}_N \\ \triangle\bar{\mathbf{L}}_N\bar{\mathbf{S}}_N^*\bar{\mathbf{L}}_N^* + (\bar{\mathbf{L}}_N^*)^T\triangle\bar{\mathbf{S}}_N\bar{\mathbf{L}}_N^* + (\bar{\mathbf{L}}_N^*)^T\bar{\mathbf{S}}_N^*\triangle\bar{\mathbf{L}}_N \\ \triangle\bar{\mathbf{L}}_B(\bar{\mathbf{L}}_B^*)^T + \bar{\mathbf{L}}_B^*\triangle\bar{\mathbf{L}}_B^T - \triangle\bar{\mathbf{X}}_B \\ \triangle\bar{\mathbf{L}}_B\bar{\mathbf{L}}_V^* + \bar{\mathbf{L}}_B^*\triangle\bar{\mathbf{L}}_V - \triangle\bar{\mathbf{X}}_V \\ \triangle\bar{\mathbf{L}}_N(\bar{\mathbf{L}}_N^*)^T + \bar{\mathbf{L}}_N^*\triangle\bar{\mathbf{L}}_N - \triangle\bar{\mathbf{X}}_N \end{bmatrix}$$

**Lemma 4.3** *$D\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{L}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}]$ is a nonsingular linear map.*

*Proof.* Assume

$$D\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{L}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = 0. \tag{50}$$

We will show that $[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{L}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = [0, 0, 0, 0]$. If we put $\mu = \mu_k$ in the system (47), then by taking the limit $k \to \infty$ we obtain that

$$(\bar{\mathbf{L}}_B^*)^T\bar{\mathbf{S}}_B^*\bar{\mathbf{L}}_B = \mathbf{W}_B, \qquad (\bar{\mathbf{L}}_N^*)^T\bar{\mathbf{S}}_N^*\bar{\mathbf{L}}_N = \mathbf{W}_N. \tag{51}$$

Since if $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$, we have also that $\mathbf{W}_B \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ and $\mathbf{W}_N \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$. Similarly, if $\mathbf{W} \in D_{++}^n$, then the blocks $\mathbf{W}_B, \mathbf{W}_N$ are also positive diagonal matrices. From the first equation of (50) and Proposition 3.3 (b) we have that

$$\triangle\bar{\mathbf{X}}_B \bullet \triangle\bar{\mathbf{S}}_B = 0, \quad \triangle\bar{\mathbf{X}}_N \bullet \triangle\bar{\mathbf{S}}_N = 0. \tag{52}$$

By applying Lemma 2.4 of [14] in the case of $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ or Theorem 1 of [20] in the case $\mathbf{W} \in D_{++}^n$ we obtain that $\triangle\bar{\mathbf{X}}_B = \triangle\bar{\mathbf{X}}_N = 0$, $\triangle\bar{\mathbf{L}}_B = \triangle\bar{\mathbf{L}}_N = 0$, $\triangle\bar{\mathbf{S}}_B = \triangle\bar{\mathbf{S}}_N = 0$. These equalities, together with (50) imply

$$(\bar{\mathbf{L}}_B^*)^T\triangle\bar{\mathbf{S}}_V\bar{\mathbf{L}}_N^* + \triangle\bar{\mathbf{L}}_V\bar{\mathbf{S}}_N^*\bar{\mathbf{L}}_N^* = 0, \qquad \bar{\mathbf{L}}_B^*\triangle\bar{\mathbf{L}}_V - \triangle\bar{\mathbf{X}}_V = 0$$

or, since $\bar{\mathbf{L}}_N^* \in L_{++}^n$,

$$(\bar{\mathbf{L}}_B^*)^T\triangle\bar{\mathbf{S}}_V + \triangle\bar{\mathbf{L}}_V\bar{\mathbf{S}}_N^* = 0, \qquad \bar{\mathbf{L}}_B^*\triangle\bar{\mathbf{L}}_V = \triangle\bar{\mathbf{X}}_V.$$

From Proposition 3.3 (b) we now have that $\triangle\bar{\mathbf{X}}_V \bullet \triangle\bar{\mathbf{S}}_V = 0$, which together with the equalities above yield

$$0 = -\triangle\bar{\mathbf{X}}_V \bullet \triangle\bar{\mathbf{S}}_V = -tr(\triangle\bar{\mathbf{X}}_V^T\triangle\bar{\mathbf{S}}_V) = -tr(\triangle\bar{\mathbf{L}}_V^T(\bar{\mathbf{L}}_B^*)^T\triangle\bar{\mathbf{S}}_V) =$$

$$= tr(\triangle\bar{\mathbf{L}}_V^T(\bar{\mathbf{L}}_B^*)^T(\bar{\mathbf{L}}_B^*)^{-T}\triangle\bar{\mathbf{L}}_V\bar{\mathbf{S}}_N^*) = tr(\triangle\bar{\mathbf{L}}_V\bar{\mathbf{S}}_N^*\triangle\bar{\mathbf{L}}_V^T).$$

Because the matrix in the last brace is positive semidefinite and $\bar{\mathbf{S}}_N^* \succ 0$ we have that $\triangle\bar{\mathbf{L}}_V = 0$. Therefore also $\triangle\bar{\mathbf{X}}_V = \triangle\bar{\mathbf{S}}_V = 0$. Assumption 2.1 gives $\triangle\bar{y} = 0$. ■

### 4.4 Analyticity of the paths as functions of $\mu$ at $\mu = 0$

**Lemma 4.4** *Let $j \in \{2,3\}$ and assume $\mathbf{W}_V \neq 0$. Then $\frac{d\mathbf{X}_V(\mu)}{d\mu}$ and $\frac{d\mathbf{S}_V(\mu)}{d\mu}$ are not bounded as $\mu \to 0$.*

*Proof.* Compute

$$\frac{d\mathbf{X}_V(\mu)}{d\mu} = \frac{d[\rho\tilde{\mathbf{X}}_V(\rho)]}{d\mu} = \frac{d\rho}{d\mu}\tilde{\mathbf{X}}_V(\rho) + \rho\frac{d\tilde{\mathbf{X}}_V(\rho)}{d\mu} = \frac{1}{2}\left[\frac{\tilde{\mathbf{X}}_V(\rho)}{\rho} + \frac{d\tilde{\mathbf{X}}_V(\rho)}{d\rho}\right]. \tag{53}$$

Since $\tilde{\mathbf{X}}_V(\rho)$ is an analytic function of $\rho$ at $\rho = 0$ (see the proof of Proposition 3.4), $\frac{d\tilde{\mathbf{X}}_V(\rho)}{d\rho}$ is bounded as $\rho \to 0$. Hence if $\frac{d\mathbf{X}_V(\mu)}{d\mu}$ were bounded (as $\mu \to 0$), then from (53) we would have that $\lim_{\rho\to 0}\tilde{\mathbf{X}}_V(\rho) = 0$. But this would imply $\mathbf{W}_V = 0$ (see Proposition 3.2), which would contradict to the assumption. The statement for $\frac{d\mathbf{S}_V(\mu)}{d\mu}$ can be proved similarly. ■

**Proposition 4.1** *Let $j \in \{2,3\}$. Then the associated weighted path $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ is an analytic function of $\mu$ for all $\mu \geq 0$ if and only if $\mathbf{W}_V = 0$.*

*Proof.* Assume $\mathbf{W}_V = 0$. From (42), (43) and Lemma 4.2 it follows that the implicit function theorem can be applied to get the analyticity of the square root path. By similar arguments as in Proposition 3.4 we obtain that the function $(\bar{\mathbf{X}}(\mu), \bar{\mathbf{Y}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ is analytically extendable to $\mu = 0$ by prescription

$$(\bar{\mathbf{X}}(\mu), \bar{\mathbf{Y}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu)) = (\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*).$$

Therefore also the path function $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ associated with the symmetrization $\Phi_2$ is analytically extendable to $\mu = 0$.
Analogously, from (48), (49) and Lemma 4.3 we obtain for the Cholesky path that the function $(\bar{\mathbf{X}}(\mu), \bar{\mathbf{L}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ can be analytically extended to $\mu = 0$ by prescription

$$(\bar{\mathbf{X}}(0), \bar{\mathbf{L}}(0), \bar{y}(0), \bar{\mathbf{S}}(0)) = (\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*).$$

Therefore also the path function $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ associated with the symmetrization $\Phi_3$ is analytically extendable to $\mu = 0$.

The reverse implication follows from Lemma 4.4 and properties of analytic functions.
■

## 5 Conclusion

Let us mention that some kind of weighted central paths associated with a Cholesky factor symmetrization was introduced and analyzed by Chua in [1, 2] where in [2] the analyticity of the paths as functions of $\mu$ at $\mu = 0$ was established. The concept of the weighted central path, as introduced in [1], coincides with the concept treating in this paper when the weight matrix is diagonal. Whence, in the case of diagonal weight matrix, the results of this paper are in agreement with those from [2].

## References


[1] C. B. Chua, *A new notion of weighted centers for semidefinite programming*, SIAM Journal on Optimization **16** (2006), 1092–1109.

[2] ———, *Analyticity of weighted central path and error bound for semidefinite programming*, to appear in Mathematical Programming.

[3] J. X. da Cruz Neto, O. P. Ferreira, and R. D. C. Monteiro, *Asymptotic behavior of the central path for a special class of degenerate sdp problems*, Mathematical Programming **103** (2005), no. 3, 487–514.

[4] E. de Klerk, C. Roos, and T. Terlaky, *Infeasible–start semidefinite programming algorithms via self dual embeddings*, Fields Institute Communications **18** (1998), 215–236.

[5] M. Halická, *Analytical properties of the central path at boundary point in linear programming*, Mathematical Programming **84** (1999), no. 2, 335–355.

[6] ———, *Analyticity of the central path at the boundary point in semidefinite programming*, EJOR **143** (2002), 311–324.

[7] M. Halická, E. deKlerk, and C. Roos, *On the convergence of the central path in semidefinite optimization*, SIAM Journal on Optimization **12** (2002), 1090–1099.

[8] ———, *Limiting behavior of the central path in semidefinite optimization*, Optimization Methods and Software **20** (2005), no. 1, 99–113.

[9] Z. Lu and R. D. C. Monteiro, *Error bounds and limiting behavior of weighted paths associated with the sdp map* $x^{\frac{1}{2}}sx^{\frac{1}{2}}$, SIAM J. on Optimization **15** (2004), 348–374.

[10] ———, *Limiting behavior of the alizadeh-haeberly-overton weighted paths in semidefinite programming*, Optimization Methods and Software **22** (2007), 849–870.

[11] Z.-Q. Luo, J. F.Sturm, and S. Zhang, *Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming*, SIAM Journal on Optimization **8** (1998), 59–81.

[12] R. D. C. Monteiro and J. S. Pang, *On two interior-point mappings for nonlinear semidefinite complementarity problems*, Math. Oper. Res. **23** (1998), 39–60.

[13] R. D. C. Monteiro and T. Tsuchiya, *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming*, SIAM Journal on Optimization **9(3)** (1999), 551–577.

[14] R. D. C. Monteiro and P. R. Zanjacomo, *Implementation of primal-dual methods for semidefinite programming based on monteiro and tsuchiya newton directions and their variants*, Optimization Methods and Software **11** (1999), 91–140.

[15] ———, *General interior-point maps and existence of weighted paths for nonlinear semidefinite complementarity problems*, Mathematics of Operations Research **25** (2000), no. 3, 381–399.

[16] M. Preiss and J. Stoer, *Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems*, Institut für angewandte Mathematik und Statistik, Universität Würzburg, working paper, 2003.

[17] ———, *Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems*, Mathematical Programming **99** (2004), no. 3.

[18] G. Sporre and A. Forsgren, *Characterization of the limit point of the central path in semidefinite programming*, 2002, Report TRITA-MAT-2002-OS12, Department of Mathematics, Royal Institute of Technology.

[19] J. Stoer and M. Wechs, *Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity*, Mathematical Programming **83** (1998), 403–423.

[20] M. Trnovská, *Existence of weighted interior-point paths in semidefinite programming*, 2007, Proceedings of 15th International Scientific Conference of Mathematical Methods in Economics and Industry, Herlany, Slovakia.