Faculty of Mathematics, Physics and Informatics

Comenius University, Bratislava

Weighted Central Paths in Semidefinite Programming

Dissertation Thesis in Applied Mathematics

2007

RNDr. Mária Trnovská

Preface

If you optimize everything, you will always be unhappy.
Donald Knuth

Semidefinite programming is a special class of convex mathematical programming, which has been recently intensively studied because of its applicability to various areas, such a s combinatorial optimization, system and control theory or mechanical and electrical engineering. Moreover, semidefinite programming problems can be efficiently solved by interior point methods. The most important concept in the theory of interior point methods is the central path. It is an analytic curve in the interior of the feasible set which tends to an optimal point at the boundary. The properties of the central path are important for designing and analyzing interior point algorithms. In this thesis the existence, the asymptotic behavior and the analyticity properties of different types of weighted central paths are studied.

The thesis is organized as follows. In Chapter 1 we introduce the semidefinite programming, the interior point methods, the notion of the central path and the weighted central path. Also a short historical overview is given and semidefinite programming applications and algorithms are discussed. Through this chapter we refer to many related works and papers from this area. In Chapter 2 the basic and well-known facts about semidefinite programming are presented, including the definitions of the problems and the duality theory in semidefinite programming. The existence proof and properties of the central path are also included. Chapter 3 deals with the existence of several types of weighted central paths. In this chapter we present the results [70], [71] of the author and we prove in all details that the weighted paths can be well defined for appropriately chosen weights. The results included in this chapter are not new, however, a new and relatively simple proof of the existence is given. Chapter 4 contains the results about the limiting behavior of the weighted central paths. The first part deals with the asymptotic behavior of the paths and in the second part these results are used for analyzing the analyticity of the paths at the boundary point. This chapter contains new results and some of them were presented in [30]. Chapter 5 is the conclusion where the main results of the thesis are summarized.

The thesis includes several appendices on necessary facts from several areas: matrix theory (Appendix A), derivatives of matrix functions (Appendix B) and asymptotic notation (Appendix C). Moreover, in Appendix D we review the assumptions needed in the thesis.

Acknowledgment

O give thanks unto the Lord, for He is good.
I Chronicles 16:34

First, a special thank you to my supervisor Margaréta Halická, who was not only an excellent teacher always asking the right questions to make me move further in my knowledge and teaching me to be precise, but also a caring and friendly person who helped me and supported me every time. Further I am indebted to doc. Milan Hamala for his invaluable comments to my work, his attendance at the seminar and willingness to share his knowledge. I am very grateful to members of the Department of Applied Mathematics of Faculty of Mathematics, Physics and Informatics, Comenius University, for the amazing working atmosphere, especially to prof. Pavel Brunovský, the main former of the friendly and inspiring environment, to Daniel Ševčovič for his unbreakable good will to help with any problem and to my (present and past) colleagues Soňa Kilianová, Marianka Remešíková, Beáta Stehlíková and Anka Zaušková, for their support, friendship and for the fun I had with them during my PhD studies. Finally I owe special thank you to my parents, my husband Dušan and my son Michal for their patience and support.

September 11, 2007 *Mária Trnovská*

List of Notation

SDP	semidefinite programming
LP	linear programming
IPM	intrior point methods
R^m	m-dimensional real vector space
$R^{p\times q}$	vector space of $p \times q$ real matrices
$\dim V$	dimension of vector space V
S^n	vector space of $n \times n$ real symmetric matrices, $\dim S^n = n(n+1)/2$
S^n_+	closed convex cone of all $n \times n$ real symmetric positive semidefinite matrices
S^n_{++}	open convex cone of all $n \times n$ real symmetric positive definite matrices
L^n	vector space of $n \times n$ real lower triangular matrices, $\dim L^n = n(n+1)/2$
L^n_+	closed convex cone of all $n \times n$ real lower triangular matrices with nonnegative diagonal entires
L^n_{++}	open convex cone of all $n \times n$ real lower triangular matrices with positive diagonal entires
U^n	vector space of $n \times n$ real upper triangular matrices, $\dim U^n = n(n+1)/2$
U^n_+	closed convex cone of all $n \times n$ real upper triangular matrices with nonnegative diagonal entires
U^n_{++}	open convex cone of all $n \times n$ real upper triangular matrices with positive diagonal entires
D^n	vector space of $n \times n$ real diagonal matrices,
D^n_+	closed convex cone of all $n \times n$ real diagonal matrices with nonnegative diagonal entires
D^n_{++}	open convex cone of all $n \times n$ real diagonal matrices with positive diagonal entires

$\mathbf{A} = (\mathbf{A}_{ij})$	matrix $\mathbf{A}$ with entries $\mathbf{A}_{ij}$
$rank(\mathbf{A})$	rank of matrix $\mathbf{A}$
$tr(\mathbf{A})$	trace of matrix $\mathbf{A}$, $tr(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}_{ii}$
$\det \mathbf{A}$	determinant of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of matrix $\mathbf{A}$
$\mathbf{A}^{T}$	transpose of matrix $\mathbf{A}$
$\mathbf{A}^{*}$	conjugate transpose of matrix $\mathbf{A}$
$Ker(\mathbf{A})$	kernel or null of matrix $\mathbf{A}$, $Ker(\mathbf{A}) = \{x \mid \mathbf{A}x = 0\}$
$Im(\mathbf{A})$	image space of $\mathbf{A}$, $Im(\mathbf{A}) = \{\mathbf{A}x\}$
$\mathbf{I}$	identity matrix (the dimension is clear from the context)
0	zero number, vector or matrix, respectively (the dimension is clear from the context)
$\mathbf{X} \succeq 0$	$\mathbf{X} \in S^n_+$
$\mathbf{X} \succ 0$	$\mathbf{X} \in S^n_{++}$
$\mathbf{L}_{\mathbf{X}}$	lower Cholesky factor of a positive semidefinite matrix $\mathbf{X}$, $\mathbf{L}_{\mathbf{X}} \in L^n_+$, $\mathbf{X} = \mathbf{L}_{\mathbf{X}}\mathbf{L}_{\mathbf{X}}^T$.
$\mathbf{U}_{\mathbf{X}}$	upper Cholesky factor of a positive semidefinite matrix $\mathbf{X}$, $\mathbf{U}_{\mathbf{X}} \in U^n_+$, $\mathbf{X} = \mathbf{U}_{\mathbf{X}}\mathbf{U}_{\mathbf{X}}^T$.
$\mathbf{X} \bullet \mathbf{Y}$	inner product defined on $R^{p\times q}$ as $tr(\mathbf{X}^T\mathbf{Y})$
$vec(\mathbf{A})$	for $\mathbf{A} \in R^{p\times q}$, $vec(\mathbf{A}) = (\mathbf{A}_{11}, \ldots, \mathbf{A}_{1q}, \mathbf{A}_{21}, \ldots, \mathbf{A}_{2q}, \ldots, \mathbf{A}_{pq}) \in R^{pq}$
$svec(\mathbf{B})$	for $\mathbf{B} \in S^n$, $svec(\mathbf{B}) = (\mathbf{B}_{11}, \sqrt{2}\mathbf{B}_{12}, \ldots, \sqrt{2}\mathbf{B}_{1n}, \mathbf{B}_{22}, \sqrt{2}\mathbf{B}_{23}, \ldots$ $\ldots, \sqrt{2}\mathbf{B}_{2n} \ldots, \mathbf{B}_{nn})$
$\Vert\mathbf{A}\Vert_F$	Frobenius matrix norm defined as $\Vert\mathbf{A}\Vert_F = \sqrt{\mathbf{A} \bullet \mathbf{A}}$
$\Vert\mathbf{A}\Vert_2$	spectral matrix norm defined as $\Vert\mathbf{A}\Vert_2 = \max\{\sqrt{\lambda}; \lambda \textit{ is an eigenvalue of } \mathbf{A}^T\mathbf{A}\}$
$\langle\langle\mathbf{B}\rangle\rangle_{\mathbf{A}}$	unique symmetric matrix $\mathbf{H}$ which satisfies the equation $\mathbf{AH} + \mathbf{HA} = \mathbf{B}$ for $\mathbf{A} \in S^n_{++}$, $\mathbf{B} \in S^n$.
$[[\mathbf{B}]]_{\mathbf{L}}$	unique lower triangular matrix $\mathbf{H}$ which satisfies the equation $\mathbf{LH}^T + \mathbf{HL}^T = \mathbf{B}$ for $\mathbf{L} \in L^n_{++}$, $\mathbf{B} \in S^n$.
$\star$	symmetric Kronecker product

Contents

Chapter 1

Introduction

Semidefinite programming

Semidefinite programming (SDP) is a special class of mathematical programming, which became popular during the 1990's. It has attracted researchers from various areas of theoretical and applied mathematics, including convex programming, linear algebra, numerical optimization, combinatorial optimization, control theory and statistics. This interest had several reasons: SDP contains important classes of convex optimization problems (like linear, quadratic and so–called second order cone programming) as special cases; semidefinite constraints arise directly in many applications not only in the areas mentioned above, but also in approximation theory, spectral analysis and structural design (a rich overview of the applications is given in [78]); and finally, SDP problems can be solved efficiently using interior point methods.

Semidefinite programming problems can be characterized as optimization problems with linear objective function and linear constraints, where the variable is a symmetric matrix, which is required to be positive semidefinite or, alternatively (in terms of vector variable), as problems where a linear function is optimized subject to the constraint that an affine combination of matrices is positive semidefinite. It seems that the first formulation of the SDP problem (as a generalization of linear programming problem) was given by Bellman and Fan [5], however the importance of the constraints requiring a certain affine combination of matrices to be positive semidefinite (now known as linear matrix inequalities or LMIs) had been recognized much earlier in control theory.[1]

The main differences between semidefinite programming and linear programming (LP) are that the feasible set in SDP problem is indeed convex but no more polyhedral

[1]These constraints are involved in Lyapunov's characterization of stability of the solution of a linear differential equation, for more about LMIs see e.g. [6].

and therefore there is no practical simplex method for SDP programs [2], and although SDP has a rich duality theory, it is not as strong as in linear programming. Still the theory of SDP is close to the theory of LP and many results from LP were generalized to SDP. The early papers dealing with theory of SDP (optimality conditions) were [10], [19], [61], [77].

Interior point methods

Positive development in SDP started after 1984, when Karmarkar published his projective algorithm for linear programming [35], possessing polynomial time worst case complexity and excellent behavior in practice. Surprisingly, it was later uncovered, that this algorithm closely relates to the (logarithmic) barrier method developed in 1960s in the context of nonlinear optimization (see e.g. Fiacco and McCormick [17]). Several variants of Karmarkar's interior point method have been developed—a major survey was made by Gonzaga [21]. [3]

Consequently, many mathematicians tried to extend the interior point method to general convex programming problems. However, while applying the new techniques, several problems regarding the Newton method analysis appeared. These problems were resolved by Nesterov and Nemirovsky in 1988 (see [52], [54], [55]) by defining the so called self-concordant property and showing that for every convex set a self-concordant barrier exists (however it is not always readily computable). This result led to generalization of the IPM for LP into interior point methods for SDP, which was made independently by Nesterov and Nemirovsky [53] and Alizadeh [2]. After that many works and papers concerning IPM algorithms appeared, see e.g. [4], [38], [45], [57], [43] [66], [12], [62], [37], [56].

Central path

Besides the algorithmical aspects of interior point methods also theoretical aspects of IPM were studied. These involve the central path—an analytical curve of solutions of parameterized barrier problems. It lies in the interior of the feasible set and tends to the optimal solution of the original problem. Most IPM algorithms follow the central path and its analytical properties are important for the convergence analysis of the algorithms.

Numerous facts concerning the convergence, the limiting behavior and the analyticity properties of this curve have been established. The central path was firstly studied in the context of linear programming and the works dealing with the limiting behavior of

[2]However, there were some approaches of extentions of the simplex method for SDP, this fact and related papers are discussed in [73].

[3]Another books on this subject are e.g. [60], [80]; for historical survey see e.g. [79], [27].

the path in LP are e.g. [1], [44], [76]. The convergence of the derivatives of the central path was established by Güler [22] and it was proved independently by Halická [23], [24] and Wechs [75] that the central path in LP can be analytically extended to the boundary point.

The analysis of the properties of the central path in linear complementarity problems (LCP) is more complicated. It was shown by Kojima et al. [36] that the central path converges to a solution. However, in LCP the asymptotic behavior of the derivatives of the path depends on the existence of the so–called strictly complementary optimal solution, related results were established by Monteiro et al. [46], [48]. The results about the analytical extension of the path can be found in the papers of Stoer and Wechs [65], [64].

The behavior of the central path in semidefinite programming depends on whether the SDP has a strictly complementary optimal solution. Under the assumption of the existence of the strictly complementary optimal solution a characterization of the limit point of the central path was given and it was shown that the central path converges to the so–called analytic center of the optimal solution set (see [43], [13], [20]). Moreover, it was shown by Halická [25] that the central path of an SDP possesing strictly complementary optimal solution can be analytically extended to the boundary point.

In the absence of strict complementarity it was shown by de Klerk et al. [14] and Goldfarb and Scheinberg [20] that any limit point of the central path must be a maximally complementary solution. This property was used to prove that the central path in SDP converges. This result is included in works of Halická et al. [29], where the proof was based on a deep result from algebraic geometry, and Kojima et al. [38], who used similar arguments as the ones in [36] and showed the convergence of the central path in monotone semidefinite complementarity problems, which is equivalent to SDP. Partial characterization of the limit point of the central path as an analytic center of some convex subset of the optimal solution set was given by Halická et al [28] and Sporre and Forsgren [63]. The analyticity of the of the central path at the boundary point in the absence of strict complementarity was given by da Cruz Neto et al. [11] for a certain subclass of SDP problems, however for general SDP this property has not been proved yet.

Weighted central path

In linear programming, the notion of the central path can be easily extended to the notion of the weighted central path—by defining weighted logarithmic barrier functions, see [22]. However this characterization does not seem to be good method to obtain suitable notion of the weighted central path in SDP. [4] Several approaches of how to define the weighted central path in SDP appeared. One of the first approaches of how to define weighted centers for SDP was given by Sturm and Zhang [67]. An another one was developed by Monteiro et al. [47], [50], where the result [46] is extended to nonlinear semidefinite complementarity problems.[5] In this approach, the weighted central path is defined as the set of optimal solutions of a weighted centering system and according to the symmetrization map used in the (perturbed) complementarity condition, several types of weighted paths can be distinguished. In what follows only this notion of weighted paths will be discussed.

The existence of the weighted central paths in SDP was studied by the following authors. A general approach was presented in the work of Monteiro and Zanjacomo [50], where the existence of the weighted paths in nonlinear semidefinite complementarity problems associated with various types of symmetrizations was proved and to this aim deep results from nonlinear analysis and theory of local homeomorphic maps were used. The approach of Preiss and Stoer [58] was more elementary (based on the Implicit function theorem), however only one type of weighted paths in the linear semidefinite complementarity problems was considered—the one associated with the so called AHO-symmetrization (see [4]). The weighted path in SDP associated with the Cholesky-type-symmetrization and positive diagonal weight was studied in the paper of Chua [7] and the existence was shown by defining weighted logarithmic barrier functions. A simplified and unified existence proof of all types of weighted paths in terms of semidefinite programming is given in [71] and in this thesis.

Also the limiting behavior and the analyticity of the weighted central paths in SDP at the boundary point was studied and all related results are given under the assumption of the existence of the strictly complementary optimal solution. These results are included in the works of Lu and Monteiro [41], [42], Preiss and Stoer [59], Chua [8], and Halická and Trnovská—which are contained in this thesis (see also [30]).

Semidefinite programming applications

As it was mentioned above, there exist many important applications of semidefinite programming—various problems can be directly formulated as an SDP and some can be relaxed using SDP to obtain better approximations of the optimal value.

In many cases semidefinite programming arises in the form of optimizing a linear combination of eigenvalues of a matrix subject to the linear constraints on the matrix (the

[4] This approach appeared to be possible for a special type of weighted path in SDP, associated with so-called Cholesky type symmetrization and positive diagonal weight, see [7].

[5] SDP can be considered as a special case of the nonlinear semidefinite complementarity problem.

equivalence to SDP is shown e.g. in [3]). A special quasiconvex programming problem can be formulated as a SDP (see [68]) and some (nonconvex) quadratic programs can be relaxed using SDP. One can find special interpretation of these problems, besides the standard SDP problem, in various areas, like e.g. statistics, graph theory or engineering and using IPM for SDP can appear to be a big advantage.

Applications in combinatorial optimization include e.g. the Lovász ϑ-function, the MAX-CUT problem or the maximum satisfiability problem. These applications are described by de Klerk in [15]. More about combinatorial optimization applications see e.g. [3], a survey was done by Goemans and Rendl in [78].

As to the SDP engineering applications, on the first place one should put the system and control theory. A monograph on this area is given by Boyd et al. [6], several examples can be found also in [73] and a survey is given by Balakrishnan and Wang in [78]. Another engineering applications are e.g. pattern recognition [73], [74] or structural design (for survey see Ben-Tal and Nemirovsky in [78]). More examples are given in [74].

Statistical applications of SDP appear mostly in the area of experiment design. Some examples are given in [74] and a survey was done by Fedorov and Lee in [78].

Semidefinite programming algorithms

The developement of efficient polynomial time algorithms for SDP started at the end of 1980s, when Nesterov and Nemirovski [52], [54] showed that the problem of minimalization of a linear function over a convex set can be solved in polynomial time as long as a selfconcordant barrier function for the convex program is known. They showed that linear program, quadratic constrainned quadratic program and semidefinite program have explicit and easily computable selfconcordant barier.

On the other hand in [4] a potential reduction algorithm for LP was extended to SDP by mechanical way. It was also shown that this approach was succesful for many polynomial time algorithms which were mostly primal or mostly dual oriented.

Since then many authors proposed interior point algorithms for SDP. This includes works [2], [31], [43], [38], [45], [49], [51], [13], [53], [57], [56]. An overview of the interior point algorithms for SDP can be found e.g. in [15], [78].

SDP literature and internet resources

At the end of this chapter we would like to refer to several resources, which can offer a compact view on semidefinite programming, including the theory, the algorithms and the applications. A survey of duality results, facts about the central path and interior point algorithms is given in the monograph by de Klerk [15]. Here the selected applications related to combinatorial optimization. A huge survey of the results concerning the theory and applications of SDP up to the year 2000 is contained in the handbook of Semidefinite Programming [78] edited by Wolkowicz et al. Shorter, however excellent surveys are given by Vandenberghe and Boyd[6] [73] and Todd[7] [68].

As to related internet resources, many informations can be found on

`http://www-user.tu-chemnitz.de/~helmberg/semidef.html`

`http://liinwww.ira.uka.de/bibliography/Math/psd.html`

[6]available at www.stanford.edu/ boyd/sdp.html

[7]available at http://www.orie.cornell.edu/ miketodd/soa5.ps

Chapter 2

Semidefinite programming

In this chapter we present some basic and well-known facts about semidefinite programming, which come mostly from the following resources: [3], [15], [26], [73]. The first part contains the definitions of the primal and dual SDP problems and the SDP duality theory. In the second part we focus on the central path in SDP—the existence proof and several properties of the central path are included.

2.1 Semidefinite programming problems

2.1.1 Basic definitions and properties

Let $\mathbf{A}^1, \ldots, \mathbf{A}^m, \mathbf{C} \in S^n$ and $b \in R^m$ be given. We will consider the following primal semidefinite programming problem

$$\begin{array}{rrcll} \text{minimize} & \mathbf{C} \bullet \mathbf{X} & & & \\ \text{subject to} & \mathbf{A}^i \bullet \mathbf{X} & = & b_i, & \text{for all } i = 1, \ldots, m, \\ & \mathbf{X} & \succeq & 0, & \end{array} \tag{2.1}$$

where $\mathbf{X} \in S^n$ is a variable. The dual semidefinite programming problem can be expressed in the form

$$\begin{array}{rrcl} \text{maximize} & b^T y & & \\ \text{subject to} & \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} & = & \mathbf{C}, \\ & \mathbf{S} & \succeq & 0, \end{array} \tag{2.2}$$

where $(\mathbf{S}, y) \in S^n \times R^m$ are variables. We will denote

$$\mathcal{P} := \left\{ \mathbf{X} \in S^n \;\middle|\; \mathbf{A}^i \bullet \mathbf{X} = b_i, \ i = 1, \ldots, m; \ \mathbf{X} \succeq 0 \right\},$$

$$\mathcal{P}^0 := \{\mathbf{X} \in S^n \mid \mathbf{A}^i \bullet \mathbf{X} = b_i,\ i = 1, \ldots, m;\ \mathbf{X} \succ 0\}$$

the primal feasible and the primal strictly feasible set, respectively, and

$$\mathcal{D} := \left\{ (y, \mathbf{S}) \in R^m \times S^n \;\middle|\; \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C},\ \mathbf{S} \succeq 0 \right\},$$

$$\mathcal{D}^0 := \left\{ (y, \mathbf{S}) \in R^m \times S^n \;\middle|\; \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C},\ \mathbf{S} \succ 0 \right\}$$

the dual feasible and the dual strictly feasible set, respectively.
Let p^*, d^* be the primal and dual optimal value, that is,

$$p^* \in \langle -\infty, \infty), \qquad p^* = \inf\{\mathbf{C} \bullet \mathbf{X} \mid \mathbf{X} \in \mathcal{P}\},$$

$$d^* \in (-\infty, \infty\rangle, \qquad d^* = \sup\{b^T y \mid y \in \mathcal{D}\},$$

and define $p^* = +\infty$ if $\mathcal{P} = \emptyset$ and $d^* = -\infty$ if $\mathcal{D} = \emptyset$.
We will denote

$$\mathcal{P}^* := \{\mathbf{X} \in \mathcal{P} \mid \mathbf{C} \bullet \mathbf{X} = p^*\},$$

$$\mathcal{D}^* := \{y \in \mathcal{D} \mid b^T y = d^*\}$$

the primal and dual optimal solution sets.

Denote

$$\mathcal{N} := \left\{ (\mathbf{X}, y, \mathbf{S}) \in S^n \times R^m \times S^n \;\middle|\; \mathbf{A}^i \bullet \mathbf{X} = 0,\ i = 1, 2, \ldots, m;\ \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = 0 \right\},$$

$$\mathcal{R} := \left\{ \Big(\mathbf{A}^1 \bullet \mathbf{X}, \ldots, \mathbf{A}^m \bullet \mathbf{X}, \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S}\Big) \in R^m \times S^n \;\middle|\; \mathbf{X} \in S^n, y \in R^m, \mathbf{S} \in S^n \right\}.$$

Lemma 2.1.1 *For any* $(\mathbf{X}, y, \mathbf{S}) \in \mathcal{N}$ *we have* $\mathbf{X} \bullet \mathbf{S} = 0$.

Proof. If $(\mathbf{X}, y, \mathbf{S}) \in \mathcal{N}$, then clearly

$$\mathbf{X} \bullet \mathbf{S} = \mathbf{X} \bullet \Big(-\sum_{i=1}^m \mathbf{A}^i y_i \Big) = -\sum_{i=1}^m (\mathbf{A}^i \bullet \mathbf{X}) y_i = 0.$$

$\square$

Corollary 2.1.1 *Let $\mathbf{X}_1, \mathbf{X}_2$ be primal feasible and $(y_1, \mathbf{S}_1), (y_2, \mathbf{S}_2)$ be dual feasible solutions. Then*

$$(\mathbf{X}_1 - \mathbf{X}_2) \bullet (\mathbf{S}_1 - \mathbf{S}_2) = 0.$$

Lemma 2.1.2 *If $\mathbf{A}^1, \ldots, \mathbf{A}^m$ are linearly independent, then $\mathcal{R} = R^m \times S^n$.*

Proof. From the linear independence of the matrices $\mathbf{A}^1, \ldots, \mathbf{A}^m$ it follows that if $z \in R^m$ is arbitrary, then there exists $\mathbf{X} \in S^n$ so that $\mathbf{A}^i \bullet \mathbf{X} = z_i \ \forall i = 1, 2, \ldots, m$. Let $\mathbf{Z}$ be any symmetric matrix. The matrices $\mathbf{A}^1, \ldots, \mathbf{A}^m$ can be completed to the basis $[\mathbf{A}^1, \ldots, \mathbf{A}^{\bar{n}}]$ of S^n Then

$$\mathbf{Z} = \sum_{i=1}^{m} \mathbf{A}^i y_i + \sum_{j=m+1}^{\bar{n}} \mathbf{A}^j y_j = \sum_{i=1}^{m} \mathbf{A}^i y_i + \mathbf{S}.$$

□

2.1.2 Duality and complementarity

Semidefinite programming has a rich duality theory, however not as strong as linear programming. It is well-known that in LP we either have that there exist optimal solutions of the primal and dual LP problems and $p^* = d^*$, or that one of the problems is infeasible and the other one is unbounded, or the both—the primal and the dual problem is infeasible. But in SDP it can happen that $p^* > d^*$ even if the both—primal and dual optimal solution exists, or that one of the problems is unsolvable (the optimal solution set is empty), however the optimal value is finite (for examples see e.g. [73]). To ensure that the pair of problems (2.1), (2.2) possesses the same primal-dual relations like in LP, the assumption of the existence of an interior point in the primal and dual strictly feasible set is needed (see Theorem 2.1.2 in the next). More about the duality theory in SDP can be found in [3], [15], [68], [73], [78].

Theorem 2.1.1 *(Weak duality theorem) If $\mathbf{X} \in \mathcal{P}$ and $(y, \mathbf{S}) \in \mathcal{D}$, then*

$$\mathbf{C} \bullet \mathbf{X} \geq b^T y.$$

Proof. We have that

$$\mathbf{C} \bullet \mathbf{X} - b^T y = \mathbf{C} \bullet \mathbf{X} - \sum_{i=1}^{m} (\mathbf{A}^i \bullet \mathbf{X}) y_i = \mathbf{X} \bullet (\mathbf{C} - \sum_{i=1}^{m} \mathbf{A}^i y_i) = \mathbf{X} \bullet \mathbf{S}.$$

However from the positive semidefinitness of the matrices $\mathbf{X}$ and $\mathbf{S}$ it follows that $\mathbf{X} \bullet \mathbf{S} \geq 0$ (see Proposition A.2.3).

□

As a simple consequence we obtain the following corollary.

Corollary 2.1.2

$$p^* \geq d^*.$$

For $\mathbf{X}, y, \mathbf{S}$ feasible, the value $\mathbf{C} \bullet \mathbf{X} - b^T y$ is called the duality gap and the value $p^* - d^*$ is called the optimal duality gap. Weak duality theorem implies that if $\mathbf{X} \in \mathcal{P}$ and $(y, \mathbf{S}) \in \mathcal{D}$ are be optimal, then the duality gap is zero, that is, $\mathbf{X} \bullet \mathbf{S} = 0$. However, this condition is equivalent to $\mathbf{XS} = 0$, since $\mathbf{X} \succeq 0, \mathbf{S} \succeq 0$ (see Proposition A.2.4).

The following well known result can be proved using the so-called generalized Farkas lemma (see Lemma 2.3 of [3]).

Theorem 2.1.2 *(Duality theorem)*
a) If $\mathcal{D}^0 \neq \emptyset$, then $d^* = p^*$. Moreover, if p^* is finite, then $\mathcal{P}^* \neq \emptyset$.
b) If $\mathcal{P}^0 \neq \emptyset$, then $d^* = p^*$. Moreover, if d^* is finite, then $\mathcal{D}^* \neq \emptyset$.
c) If $\mathcal{P}^0 \neq \emptyset, \mathcal{D}^0 \neq \emptyset$, then $d^* = p^*$ and $\mathcal{P}^* \neq \emptyset, \mathcal{D}^* \neq \emptyset$.

The necessary and sufficient conditions of optimality are stated in the following theorem.

Theorem 2.1.3 *If $\mathcal{P}^0 \neq \emptyset, \mathcal{D}^0 \neq \emptyset$, then $(\mathbf{X}, y, \mathbf{S})$ is an optimal solution of (2.1), (2.2) if and only if it satisfies the following system:*

$$\begin{array}{rll} \mathbf{A}^i \bullet \mathbf{X} = b_i & i = 1, \ldots, m, & \mathbf{X} \succeq 0 \\ \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C}, & & \mathbf{S} \succeq 0, \\ \mathbf{XS} = \mathbf{0}. & & \end{array} \tag{2.3}$$

The first condition is the primal feasibility, the second condition is the dual feasibility and the third condition is called the complementarity condition. The optimal solution that satisfies the complementarity condition is called complementary.

Definition 2.1.1 *(Strictly complementary solution)*
The complementary solution $(\mathbf{X}^*, y^*, \mathbf{S}^*)$ of the problems $(P), (D)$ is strictly complementary if

$$\mathbf{X}^* + \mathbf{S}^* \succ 0.$$

In linear programming, the existence of an optimal solution implies the existence of a strictly complementary solution. However in semidefinite programming, (as well as in quadratic convex programming or for linear complementarity problems) this property is not satisfied in general.

2.2 Central path in semidefinite programming

In this section we will introduce the central path in semidefinite programming as the optimal solution set of a class of barrier problems or equivalently as the solution set of the perturbed system (2.3) of necessary and sufficient conditions for (2.1) and (2.2). Through this section we will consider the following assumptions:

Assumption (A1): The matrices $\mathbf{A}^1, \ldots, \mathbf{A}^m$ are linearly independent.

Assumption (A2): $\mathcal{P}^0 \neq \emptyset, \mathcal{D}^0 \neq \emptyset$.

Assumption (A1) ensures the one-to-one correspondence between the dual variables y and $\mathbf{S}$. Assumption (A2) (also referred to as the interior point assumption) follows from Duality theorem. Actually, both assumptions together are equivalent to the fact that the primal and dual optimal solution sets are nonempty and bounded (See [69]).

2.2.1 Existence of the central path

Choose a fixed $\mu > 0$ and define the barrier functions:

$$f_\mu^P : S_{++}^n \to R, \qquad f_\mu^P(\mathbf{X}) = \mathbf{C} \bullet \mathbf{X} - \mu \ln \det(\mathbf{X})$$

and

$$f_\mu^D : R^m \times S_{++}^n \to R, \qquad f_\mu^D(y, \mathbf{S}) = b^T y + \mu \ln \det(\mathbf{S}).$$

The associated primal and dual barrier problems are

$$\left.\begin{array}{rl} \text{minimize} & f_\mu^P(\mathbf{X}) \\ \text{subject to} & \mathbf{A}^i \bullet \mathbf{X} = b_i, \quad i = 1, \ldots, m \\ & \mathbf{X} \succ 0 \end{array}\right\} \tag{2.4}$$

and

$$\left.\begin{array}{rl} \text{maximize} & f_\mu^D(y, \mathbf{S}) \\ \text{subject to} & \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C} \\ & \mathbf{S} \succ 0 \end{array}\right\} \tag{2.5}$$

where the variables are $\mathbf{X} \in S^n$ and $(y, \mathbf{S}) \in R^m \times S^n$, respectively.

Proposition 2.2.1 *For any $\mu > 0$ the both problems* (2.4) *and* (2.5) *have at most one solution.*

Proof. The statement follows from the fact that the primal barrier function f_μ^P is strictly convex and that the dual barrier function f_μ^D is strictly concave on $\mathcal{D}^0$.

$\square$

In what follows, we prove the existence of the solutions of the problems (2.4), (2.5). This result is well-known in the theory of interior point methods for SDP and various proofs can be found e.g. in [15], [68], [78]. Nevertheless, the whole proof is presented here, since this statement is fundamental for further explanation. Our proof will be based on the approach of [26] where, besides the well-known Weierstrass theorem, also the following characterization of an unbounded convex set is used.

Lemma 2.2.1 *Let $K \subseteq S^n_{++}$ be a nonempty, convex, closed and unbounded set. Then for any $\mathbf{V} \in K$ there exists $\mathbf{W} \in K$ such that*

$$\mathbf{W} - \mathbf{V} \succeq 0, \quad \mathbf{W} - \mathbf{V} \neq 0$$

and

$$\{\mathbf{V}_t \mid \mathbf{V}_t = \mathbf{V} + t(\mathbf{W} - \mathbf{V}),\ t \geq 0\} \subseteq K.$$

Proof. From the assumptions of the lemma we have that from any point in K one can lead a ray that lies in K. More exactly, for any $\mathbf{V} \in K$ there exists $\mathbf{A} \in S^n, \mathbf{A} \neq 0$ such that

$$\mathbf{V} + t\mathbf{A} \in K, \qquad \forall\, t \geq 0. \tag{2.6}$$

Put $\mathbf{W} = \mathbf{V} + \mathbf{A}$. It remains to show that $\mathbf{A} \succeq 0$. Let $\mathbf{Q}$ and $\mathbf{D}$ be the orthogonal and diagonal matrix, respectively, for which $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{D}$. Suppose there is a negative number on the diagonal of $\mathbf{D}$. Then for some positive t the matrix $\mathbf{Q}^T\mathbf{V}\mathbf{Q} + t\mathbf{D}$ has a negative number on diagonal. But this contradicts (2.6) and therefore $\mathbf{D} \succeq 0$ and obviously also $\mathbf{A} \succeq 0$.

$\square$

Theorem 2.2.1 *For any $\mu > 0$ the both problems* (2.4) *and* (2.5) *have a unique solution.*

Proof. We will prove the statement for the primal barrier problem. For the dual barrier problem it can be done similarly.

Because of Proposition 2.2.1 , it remains to show the existence of the optimal solution of the problem (2.4). From Assumption (A2) it follows that there exists $\mathbf{X}^0 \succ 0$ such that for all $i = 1, \ldots, m$

$$\mathbf{A}^i \bullet \mathbf{X}^0 = b_i.$$

Define the set

$$\mathcal{P}(\mathbf{X}^0) = \{\mathbf{X} \in \mathcal{P}^0 \mid f_\mu^P(\mathbf{X}) \leq f_\mu^P(\mathbf{X}^0)\}.$$

Now, it is easy to see that it sufficies to show that there exists an optimal solution of the following problem

$$\begin{array}{rl} \text{minimize} & f_\mu^P(\mathbf{X}) \\ \text{subject to} & \mathbf{X} \in \mathcal{P}(\mathbf{X}^0). \end{array}$$

The function $f_\mu^P(\mathbf{X})$ is continuous and the set $\mathcal{P}(\mathbf{X}^0)$ is nonempty. Therefore, in order to apply the well-known Weierstrass result, we only have to show that the set $\mathcal{P}(\mathbf{X}^0)$ is compact.

Firstly, we will prove that the set $\mathcal{P}(\mathbf{X}^0)$ is closed. Let $\{\mathbf{X}^n\}_{i=1}^n$ be a sequence in $\mathcal{P}(\mathbf{X}^0)$, such that $\lim_{n\to\infty} \mathbf{X}^n = \hat{\mathbf{X}}$. Since for any $n = 1, 2, 3, \ldots$

$$\mathbf{A}^i \bullet \mathbf{X}^n = b_i,\ i = 1, \ldots, m, \quad \mathbf{X}^n \succ 0,$$

we have that

$$\mathbf{A}^i \bullet \hat{\mathbf{X}} = b_i,\ i = 1, \ldots, m, \quad \hat{\mathbf{X}} \succeq 0.$$

The continuity of $f_\mu^P(\mathbf{X})$ implies that

$$\lim_{n\to\infty} f_\mu^P(\mathbf{X}^n) = f_\mu^P(\hat{\mathbf{X}}) \leq f_\mu^P(\mathbf{X}^0). \tag{2.7}$$

Finally, if $\hat{\mathbf{X}} \succeq 0$ is singular, then $\det \hat{\mathbf{X}} = 0$ and hence $\ln\det \hat{\mathbf{X}} = -\infty$. But this contradicts (2.7) and therefore $\hat{\mathbf{X}}$ must be positive definite.

Now, we will prove that the set $\mathcal{P}(\mathbf{X}^0)$ is bounded. Suppose that $\mathcal{P}(\mathbf{X}^0)$ is unbounded. Observe that $\mathcal{P}(\mathbf{X}^0)$ is a convex set, since it is a sublevel set of a convex function. By applying Lemma 2.2.1 we obtain that there exist $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{P}(\mathbf{X}^0)$ satisfying

$$\mathbf{X}_1 - \mathbf{X}_2 \succeq 0, \quad \mathbf{X}_1 - \mathbf{X}_2 \neq 0$$

such that the ray

$$\{\mathbf{X}_t \mid \mathbf{X}_t = \mathbf{X}_1 + t(\mathbf{X}_1 - \mathbf{X}_2),\ t \geq 0\}$$

is included in $\mathcal{P}(\mathbf{X}^0)$. Therefore for all $t \geq 0$

$$f_\mu^P(\mathbf{X}_t) \leq f_\mu^P(\mathbf{X}^0). \tag{2.8}$$

Compute the limit

$$\lim_{t\to\infty} f_\mu^P(\mathbf{X}_t) = \lim_{t\to\infty} f_\mu^P(\mathbf{X}_1 + t(\mathbf{X}_1 - \mathbf{X}_2)) =$$

$$\lim_{t\to\infty}\Big[\mathbf{C}\bullet(\mathbf{X}_1+t(\mathbf{X}_1-\mathbf{X}_2))-\mu\ln\det(\mathbf{X}_1+t(\mathbf{X}_1-\mathbf{X}_2))\Big]=$$

$$\lim_{t\to\infty}\left[\mathbf{C}\bullet\mathbf{X}_1+t\Big(\mathbf{C}\bullet(\mathbf{X}_1-\mathbf{X}_2)-\frac{\mu\ln\det(\mathbf{X}_1+t(\mathbf{X}_1-\mathbf{X}_2))}{t}\Big)\right].$$

Denote

$$\mathbf{V}(t)=\mathbf{C}\bullet(\mathbf{X}_1-\mathbf{X}_2)-\frac{\mu\ln\det(\mathbf{X}_1+t(\mathbf{X}_1-\mathbf{X}_2))}{t}.$$

We will show that

$$0<\lim_{t\to\infty}\mathbf{V}(t)<\infty. \tag{2.9}$$

Denote $p_k(t)=\det(\mathbf{X}_1+t(\mathbf{X}_1-\mathbf{X}_2))$, which is a polynomial of a degree k. Using the L'Hospital's rule we obtain that

$$\lim_{t\to\infty}\frac{\mu\ln p_k(t)}{t}=\lim_{t\to\infty}\frac{\mu p_k'(t)}{p_k(t)}=0.$$

From Assumption (A2) it follows that there exists $(y^0,\mathbf{S}^0)\in\mathcal{D}^0$. The matrices $\mathbf{X}_1,\mathbf{X}_2$ are feasible and hence

$$\mathbf{C}\bullet(\mathbf{X}_1-\mathbf{X}_2)=\left(\sum_{i=1}^m\mathbf{A}^i y_i^0+\mathbf{S}^0\right)\bullet(\mathbf{X}_1-\mathbf{X}_2)=\mathbf{S}^0\bullet(\mathbf{X}_1-\mathbf{X}_2)>0,$$

where the inequality follows from Proposition A.2.5. We have shown that (2.9) holds and therefore

$$\lim_{t\to\infty}f_\mu^P(\mathbf{X}_t)=\infty.$$

But this contradicts to (2.8) and hence the theorem is proved.

$\square$

Denote $(\mathbf{X}(\mu),y(\mu),\mathbf{S}(\mu))$ the optimal solution of (2.4), (2.5). Theorem 2.2.1 implies that the central path can be well defined as follows:

Definition 2.2.1 *The set*

$$\{(\mathbf{X}(\mu),y(\mu),\mathbf{S}(\mu))\mid\mu>0\}$$

is called the central path.

Note. In the interior point theory the central path can be alternatively defined as the map

$$f_{cp} : R_{++} \to S^n_{++} \times R^m \times S^n_{++}, \quad f_{cp}(\mu) = (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu)).$$

The following theorem contains the necessary and sufficient conditions of optimality for the pair of barrier problems (2.4), (2.5).

Theorem 2.2.2 *Let $\mu > 0$. Then $(\mathbf{X}, y, \mathbf{S})$ is an optimal solution of (2.4), (2.5) if and only if it satisfies the system*

$$\begin{array}{rll} \mathbf{A}^i \bullet \mathbf{X} = b_i, & i = 1, \ldots, m, & \mathbf{X} \succ 0, \\ \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C}, & & \mathbf{S} \succ 0, \\ \mathbf{X}\mathbf{S} = \mu \mathbf{I}. & & \end{array} \tag{2.10}$$

Corollary 2.2.1 *For any $\mu > 0$ there exists a unique solution $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ of the system (2.10) and this solution lies on the central path.*

From the Corollary 2.2.1 it follows that the central path could be defined implicitly as the set of solutions of the system (2.10). However, the connection with the barrier problems (2.4), (2.5) is needed for proving the existence of the central path.

Note that the duality gap along the central path is equal to

$$\mathbf{C} \bullet \mathbf{X}(\mu) - b^T y(\mu) = \mathbf{X}(\mu) \bullet \mathbf{S}(\mu) = tr(\mathbf{X}(\mu)\mathbf{S}(\mu)) = tr(\mu \mathbf{I}) = n\mu.$$

Therefore, it is easy to see that if the central path has a limit point as $\mu \to 0$ then it is an optimal solution of (2.1), (2.2).

2.2.2 Analyticity and convergence of the central path

Theorem 2.2.3 *The function*

$$f_{cp} : \mu \mapsto (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$$

is an analytic function for $\mu > 0$.

The proof of this theorem is based on the fact, that the Jacobian of the map corresponding to (2.10) is nonsingular along the central path. Hence the following "analytic version" of the implicit function theorem (see also [16] Theorem 10.2.4, or [15] Theorem 3.2) can be applied. Since this theorem will be the main tool for many other results in this thesis, we provide here the full text of it.

Theorem 2.2.4 *(Implicit function theorem) Let $g : R^{r+s} \to R^s$ be an analytic function of $w \in R^r$ and $z \in R^s$ such that:*

1. *there exists $(\bar{w}, \bar{z}) \in R^{r+s}$ such that $g(\bar{w}, \bar{z}) = 0$,*
2. *The Jacobian of g with respect to z is nonsingular at $(\bar{w}, \bar{z})$.*

Then there exist (open) neighborhoods $\mathcal{O}(\bar{w}) \subset R^r$ and $\mathcal{O}(\bar{z}) \in R^s$ of $\bar{w}$ and $\bar{z}$ respectively, and an analytic function

$$f : \mathcal{O}(\bar{w}) \to \mathcal{O}(\bar{z})$$

such that $f(\bar{w}) = \bar{z}$ and

$$g(w, f(w)) = 0 \quad \forall\ w \in \mathcal{O}(\bar{w}).$$

The whole proof of Theorem 2.2.3 can be found e.g. in [15].

Theorem 2.2.5 *Let $\bar{\mu} > 0$. Then the set*

$$\{(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu)) | 0 < \mu \leq \bar{\mu}\}$$

is bounded.

From Theorem 2.2.5 it follows that the central path has a limit point as $\mu \to 0$ and hence there exists a sequence $\{\mu_k\}_{k=1}^{\infty} \in (0, \bar{\mu}\rangle$, such that $\lim_{k\to\infty} \mu_k = 0$ and

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

where $\mathbf{X}^* \in \mathcal{P}^*$ and $(y^*, \mathbf{S}) \in \mathcal{D}^*$.

Theorem 2.2.6 *The central path converges as $\mu \to 0$, i.e. it has a unique limit point, which is an optimal solution of* (2.1), (2.2).

The result stated in Theorem 2.2.6 was proved in the papers [29], [38], however the proof can be found also in [15]. In works [14], [20] it was shown that any limit point of the central path is a maximally complementary solution. Under the assumption of the existence of the strictly complementary optimal solution a characterization of the limit point of the central path as the so-called analytic center of the optimal solution set can be found in [43], [13], [20]. Partial characterization of the limit point of the central path as an analytic center of some convex subset of the optimal solution set was given in [28] and [63].

The following result was shown by Halická [25].

Theorem 2.2.7 *Assume that there exists a strictly complementary solution. Then the central path can be analytically extended to $\mu = 0$.*

The above result, however in the absence of strict complementarity, was proved in [11] for a certain type of degenerate SDP.

2.2.3 Symmetrization of the complementarity condition

It is well known that the product of two symmetric matrices is not necessary symmetric. This may cause problems in the interior point algorithms, which are based on solving the system (2.10). Therefore the matrix $\mathbf{XS}$ is replaced by some kind of symmetrization matrix $\Phi(\mathbf{X}, \mathbf{S}) \in S^n$ with the following property:

$$\text{If } \mathbf{X} \succeq 0, \mathbf{S} \succeq 0, \text{ then } \mathbf{XS} = 0 \text{ if and only if } \Phi(\mathbf{X}, \mathbf{S}) = 0.$$

In semidefinite programming and semidefinite complementarity problems the following symmetrization maps are discussed (see [4], [47], [49], [51], [50], [58], [59]):

$$\begin{array}{rcl}
\Phi_1(\mathbf{X}, \mathbf{S}) & = & (\mathbf{XS} + \mathbf{SX})/2 \\
\Phi_2(\mathbf{X}, \mathbf{S}) & = & \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}} \\
\Phi_3(\mathbf{X}, \mathbf{S}) & = & \mathbf{L}_{\mathbf{X}}^T\mathbf{S}\mathbf{L}_{\mathbf{X}} \\
\Phi_4(\mathbf{X}, \mathbf{S}) & = & (\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})/2 \\
\Phi_5(\mathbf{X}, \mathbf{S}) & = & (\mathbf{U}_{\mathbf{S}}^T\mathbf{L}_{\mathbf{X}} + \mathbf{L}_{\mathbf{X}}^T\mathbf{U}_{\mathbf{S}})/2
\end{array}$$

where $\mathbf{X}^{\frac{1}{2}}, \mathbf{S}^{\frac{1}{2}}$ are the square roots of the matrices $\mathbf{X}, \mathbf{S}$; $\mathbf{L}_{\mathbf{X}}$ is the lower Cholesky factor of the matrix $\mathbf{X}$ and $\mathbf{U}_{\mathbf{S}}$ is the upper Cholesky factor of the matrix $\mathbf{S}$. Remark that if $\mathbf{X} \succeq 0, \mathbf{S} \succeq 0$, then the matrices $\Phi_2(\mathbf{X}, \mathbf{S}), \Phi_3(\mathbf{X}, \mathbf{S})$ are positive semidefinite, however the other are not in general.

Chapter 3

Existence of weighted paths in SDP

3.1 Motivation

In linear programming, the concept of the central path can be easily extended to the concept of the weighted central path—by defining weighted logarithmic barrier functions. However, this technique can not be applied to semidefinite programming, in general. There were more approaches of how to define the weighted central path in SDP. One of them was developed by Monteiro et al. (see [47], [51]) originally for nonlinear semidefinite complementarity problems. Following this approach one can define the weighted central path for SDP as the set $\{ (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu)) \mid \mu > 0\}$ of the solutions of the parameterized systems

$$\begin{array}{l} \mathbf{A}^i \bullet \mathbf{X} = b_i + \mu \triangle b_i, \quad i = 1, \ldots, m, \quad \mathbf{X} \succ 0, \\ \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C} + \mu \triangle \mathbf{C}, \quad \mathbf{S} \succ 0, \\ \Phi_j(\mathbf{X}, \mathbf{S}) = \phi_j(\mu) \mathbf{W}, \end{array} \tag{3.1}$$

where $\triangle b \in R^m, \triangle \mathbf{C} \in S^n$ are fixed, $\mathbf{W} \succ 0$ is the weight, $\Phi_j(\mathbf{X}, \mathbf{S})$ is one of the symmetrization maps already discussed in Section 2.2.3:

$$\begin{array}{rcl} \Phi_1(\mathbf{X}, \mathbf{S}) & = & (\mathbf{XS} + \mathbf{SX})/2 \\ \Phi_2(\mathbf{X}, \mathbf{S}) & = & \mathbf{X}^{\frac{1}{2}} \mathbf{S} \mathbf{X}^{\frac{1}{2}} \\ \Phi_3(\mathbf{X}, \mathbf{S}) & = & \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}} \\ \Phi_4(\mathbf{X}, \mathbf{S}) & = & (\mathbf{X}^{\frac{1}{2}} \mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}} \mathbf{X}^{\frac{1}{2}})/2 \\ \Phi_5(\mathbf{X}, \mathbf{S}) & = & (\mathbf{U}_{\mathbf{S}}^T \mathbf{L}_{\mathbf{X}} + \mathbf{L}_{\mathbf{X}}^T \mathbf{U}_{\mathbf{S}})/2 \end{array} \tag{3.2}$$

and

$$\phi_j(\mu) = \mu, \quad j = 1, 2, 3; \qquad \phi_j(\mu) = \sqrt{\mu}, \quad j = 4, 5.$$

Therefore, according to the symmetrization map we will distinguish five types of weighted paths in semidefinite programming.

The authors Monteiro and Zanjacomo [51] have proved the existence of the weighted paths in nonlinear semidefinite complementarity problems using deep results from nonlinear analysis, based on the theory of the local homeomorphic maps. The another approach, used by Preiss and Stoer [58] was more elementary, it was essentially based on the implicit function theorem. However, latter authors proved the existence of the weighted path in linear complementarity problem associated only with the symmetrization $\Phi_1(\mathbf{X}, \mathbf{S})$. The same symmetrization and technique was used by the author in [70] for the existence of the weighted central path in SDP. In [71] the result [70] is extended to all five symmetrizations defined in (3.2).

In this chapter we present our results from [70], [71] in all details, that is, we prove that the weighted paths can be well defined (for appropriately chosen weights). To this aim we need to show that for fixed $\triangle b, \triangle \mathbf{C}$, properly chosen weight $\mathbf{W}$ and any $\mu > 0$ there exists a unique solution of the system (3.1). [1] Obviously, such weighted central paths do not lie in the interior of the feasible set in general, and hence they can be useful if an interior point does not exist or is unknown.

In the next we will consider Assumption (A1) (see Section 1.3) and instead of Assumption (A2) we will consider a weaker assumption:

Assumption (A3): The system (2.3) is solvable.

3.2 Nonsingularity of Fréchet derivatives

For fixed $\triangle b \in R^m, \triangle \mathbf{C} \in S^n$ consider the maps $F^j_{\mu,\mathbf{W}} : S^n \times R^m \times S^n \to R^m \times S^n \times S^n$ with parameters $\mu > 0$ and $\mathbf{W} \succ 0$ $(j = 1, .., 5)$:

$$F^j_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S}) = \begin{bmatrix} \mathcal{A}(\mathbf{X}) - b - \mu\triangle b \\ \mathcal{A}^*(y) + \mathbf{S} - \mathbf{C} - \mu\triangle \mathbf{C} \\ \Phi_j(\mathbf{X}, \mathbf{S}) - \phi_j(\nu)\mathbf{W} \end{bmatrix}. \tag{3.3}$$

[1] The proof of the existence of the solution of (3.1) can not be performed in the same way as in the case of the system (2.10)—it seems not to be possible to characterize the weighted central path in SDP using weighted logarithmic barrier problems.

Here the map $\mathcal{A}$ is defined as

$$\mathcal{A} : S^n \to R^m, \qquad \mathcal{A}(\mathbf{X}) = (\mathbf{A}^1 \bullet \mathbf{X}, \ldots, \mathbf{A}^m \bullet \mathbf{X}), \tag{3.4}$$

and its adjoint map is

$$\mathcal{A}^* : R^m \to S^n, \qquad \mathcal{A}^*(y) = \sum_{i=1}^{m} \mathbf{A}^i y_i.$$

Clearly, in this notation the system (3.1) is equivalent to

$$F^j_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S}) = 0, \qquad \mathbf{X} \succ 0,\ \mathbf{S} \succ 0.$$

The main tool we will use in the proof of the existence of the weighted paths is the analytic version of the implicit function theorem (Theorem 2.2.4, see also e.g. [16], [15]). In this context we will be interested in the Fréchet derivative of the maps $F^j_{\mu,\mathbf{W}}$. It can be derived that if $\mathbf{X} \succ 0, \mathbf{S} \succ 0$, the Fréchet derivative of $DF^j_{\mu,\mathbf{W}}$ at $(\mathbf{X}, y, \mathbf{S})$ is the linear map

$$DF^j_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] = \begin{bmatrix} \mathcal{A}(\triangle\mathbf{X}) \\ \mathcal{A}^*(\triangle y) + \triangle\mathbf{S} \\ D\Phi_j(\mathbf{X}, \mathbf{S})[\triangle\mathbf{X}, \triangle\mathbf{S}] \end{bmatrix}$$

with the variables $[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] \in S^n \times R^m \times S^n$ where $D\Phi_j(\mathbf{X}, \mathbf{S})$ are derived in Appendix B (see Corollary B.2.1).

In this section we will derive the sufficient conditions to ensure the nonsingularity of $DF^j_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$. These conditions will differ for particular $j \in \{1, \ldots, 5\}$.

3.2.1 Nonsingularity of $DF^1_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$

Recall that if $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$ then, according to Corollary B.2.1,

$$DF^1_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] = \begin{bmatrix} \mathcal{A}(\triangle\mathbf{X}) \\ \mathcal{A}^*(\triangle y) + \triangle\mathbf{S} \\ \frac{1}{2}(\triangle\mathbf{X}\mathbf{S} + \mathbf{S}\triangle\mathbf{X} + \triangle\mathbf{S}\mathbf{X} + \mathbf{X}\triangle\mathbf{S}) \end{bmatrix}$$

Define the matrix[2]

$$\mathbb{A} = \begin{pmatrix} svec(\mathbf{A}^1)^T \\ \vdots \\ svec(\mathbf{A}^m)^T \end{pmatrix} \in R^m \times R^{\bar{n}}. \tag{3.5}$$

[2]See Definition A.4.1 in Appendix A.4.

Then clearly

$$\mathcal{A}(\mathbf{X}) = \mathbb{A}svec(\mathbf{X}),$$
$$svec(\mathcal{A}^*(y)) = \mathbb{A}^T y.$$

Using the matrices defined in (3.5), the symmetric Kronecker product[3] and the *svec* operator we can write $DF^1_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ in a matrix-vector representation

$$\begin{pmatrix} \mathbb{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{A}^T & \mathbf{I} \\ \mathbf{S} \star \mathbf{I} & \mathbf{0} & \mathbf{X} \star \mathbf{I} \end{pmatrix} \begin{pmatrix} svec(\triangle \mathbf{X}) \\ \triangle y \\ svec(\triangle \mathbf{S}) \end{pmatrix},$$

which will be useful for proving the following proposition:

Proposition 3.2.1 *If* $\mathbf{XS} + \mathbf{SX} \succ 0$*, then* $DF^1_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ *is a nonsingular linear map.*

Proof. Assume

$$\begin{pmatrix} \mathbb{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{A}^T & \mathbf{I} \\ \mathbf{S} \star \mathbf{I} & \mathbf{0} & \mathbf{X} \star \mathbf{I} \end{pmatrix} \begin{pmatrix} svec(\triangle \mathbf{X}) \\ \triangle y \\ svec(\triangle \mathbf{S}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.6}$$

It suffices to show that the equations (3.6) imply

$$\begin{pmatrix} svec(\triangle \mathbf{X}) \\ \triangle y \\ svec(\triangle \mathbf{S}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

By Lemma 2.1.1 we have that

$$svec(\triangle \mathbf{X})^T svec(\triangle \mathbf{S}) = 0. \tag{3.7}$$

From the third equation of (3.6) it follows

$$svec(\triangle \mathbf{X}) = -(\mathbf{S} \star \mathbf{I})^{-1}(\mathbf{X} \star \mathbf{I}) svec(\triangle \mathbf{S}). \tag{3.8}$$

The equalities (3.7), (3.8) imply

$$svec(\triangle \mathbf{S})^T (\mathbf{S} \star \mathbf{I})^{-1}(\mathbf{X} \star \mathbf{I}) svec(\triangle \mathbf{S}) = 0$$

and hence, according to Corollary A.4.4, $svec(\triangle \mathbf{S}) = 0$. Finally, Assumption (A1) and (3.8) yield $\triangle y = 0$ and $svec(\triangle \mathbf{X}) = 0$, respectively.

□

[3] See Definition A.4.2 in Appendix A.4.

3.2.2 Nonsingularity of $DF^2_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$

Recall that if $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$, then, according to Corollary B.2.1,

$$DF^2_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] =$$

$$= \begin{bmatrix} \mathcal{A}(\triangle\mathbf{X}) \\ \mathcal{A}^*(\triangle y) + \triangle\mathbf{S} \\ \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}} \mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S} \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} \end{bmatrix}.$$

where $\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}$ is the unique solution $\mathbf{H}$ of the equation $\mathbf{X}^{\frac{1}{2}}\mathbf{H} + \mathbf{H}\mathbf{X}^{\frac{1}{2}} = \triangle\mathbf{X}$ (see Definition B.2.1 (a)). Using the matrices defined in (3.5), the symmetric Kronecker product and the *svec* operator, $DF^2_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ can be rewritten as:

$$\begin{pmatrix} \mathbb{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{A}^T & \mathbf{I} \\ (\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I})(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})^{-1} & \mathbf{0} & \mathbf{X}^{\frac{1}{2}} \star \mathbf{X}^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} svec(\triangle\mathbf{X}) \\ \triangle y \\ svec(\triangle\mathbf{S}) \end{pmatrix}$$

Proposition 3.2.2 *Let* $\mathbf{X} \succ 0, \mathbf{S} \succ 0$ *and* $\mathbf{X}^{\frac{1}{2}}\mathbf{S} + \mathbf{S}\mathbf{X}^{\frac{1}{2}} \succeq 0$. *Then* $DF^2_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ *is a nonsingular linear map.*

Proof. Assume $\mathbf{X} \succ 0, \mathbf{S} \succ 0$, $\mathbf{X}^{\frac{1}{2}}\mathbf{S} + \mathbf{S}\mathbf{X}^{\frac{1}{2}} \succeq 0$ and

$$\begin{pmatrix} \mathbb{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{A}^T & \mathbf{I} \\ (\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I})(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})^{-1} & \mathbf{0} & \mathbf{X}^{\frac{1}{2}} \star \mathbf{X}^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} svec(\triangle\mathbf{X}) \\ \triangle y \\ svec(\triangle\mathbf{S}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.9}$$

We will show that

$$\begin{pmatrix} svec(\triangle\mathbf{X}) \\ \triangle y \\ svec(\triangle\mathbf{S}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

By Lemma 2.1.1 we have that

$$svec(\triangle\mathbf{X})^T svec(\triangle\mathbf{S}) = 0. \tag{3.10}$$

From the third equation it follows

$$svec(\triangle\mathbf{S}) = -(\mathbf{X}^{\frac{1}{2}} \star \mathbf{X}^{\frac{1}{2}})^{-1}(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I})(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})^{-1} svec(\triangle\mathbf{X}). \tag{3.11}$$

The equalities (3.10), (3.11) imply

$$svec(\triangle\mathbf{X})^T(\mathbf{X}^{\frac{1}{2}} \star \mathbf{X}^{\frac{1}{2}})^{-1}(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I})(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})^{-1} svec(\triangle\mathbf{X}) = 0.$$

It remains to show that, under the given assumptions, the matrix

$$(\mathbf{X}^{\frac{1}{2}} \star \mathbf{X}^{\frac{1}{2}})^{-1}(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I})(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})^{-1}$$

is positive definite, or equivalently that the matrix

$$(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})(\mathbf{X}^{-\frac{1}{2}} \star \mathbf{X}^{-\frac{1}{2}})(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I})$$

is positive definite. Using the properties of the symmetric Kronecker product[4] we obtain

$$(\mathbf{X}^{\frac{1}{2}} \star \mathbf{I})(\mathbf{X}^{-\frac{1}{2}} \star \mathbf{X}^{-\frac{1}{2}})(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I}) = (\mathbf{I} \star \mathbf{X}^{-\frac{1}{2}})(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{I}) = \frac{1}{2}(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{X}^{-\frac{1}{2}} + \mathbf{I} \star \mathbf{S}).$$

The matrix $\mathbf{I} \star \mathbf{S}$ is positive definite and hence we only have to show that $\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{X}^{-\frac{1}{2}}$ is positive semidefinite. Let $\mathbf{V} \in S^n$ be arbitrary. Then

$$2svec(\mathbf{V})^T(\mathbf{X}^{\frac{1}{2}}\mathbf{S} \star \mathbf{X}^{-\frac{1}{2}})svec(\mathbf{V}) = \mathbf{V} \bullet [\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{V}\mathbf{X}^{-\frac{1}{2}} + \mathbf{X}^{-\frac{1}{2}}\mathbf{V}\mathbf{S}\mathbf{X}^{\frac{1}{2}}] =$$

$$= tr(\mathbf{V}\mathbf{X}^{-\frac{1}{2}}\mathbf{V}\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{V}\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{V}\mathbf{X}^{-\frac{1}{2}}) =$$

$$= tr(\mathbf{X}^{-\frac{1}{4}}\mathbf{V}\mathbf{S}\mathbf{X}^{\frac{1}{2}}\mathbf{V}\mathbf{X}^{-\frac{1}{4}} + \mathbf{X}^{-\frac{1}{4}}\mathbf{V}\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{V}\mathbf{X}^{-\frac{1}{4}}) = tr(\mathbf{X}^{-\frac{1}{4}}\mathbf{V}(\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S})\mathbf{V}\mathbf{X}^{-\frac{1}{4}}) \geq 0$$

and hence the proposition is proved.

□

Note. The proposition above defines a set of matrices $\mathbf{X}$ and $\mathbf{S}$ for which $DF^2_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ is a nonsingular map, however, it seems not to be possible to determine the set of suitable weights (which relates to the third equation in (3.1)) from the condition $\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S} \succeq 0$. Such a set can be found using the statement of the following lemma.

The following lemma is a consequence of Lemma 2.3 of [49]. Because of sake of the completeness of the explanation, the proof of this result is included in all details.

Lemma 3.2.1 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$ *and* $\gamma \in (0, \frac{1}{\sqrt{2}})$. *If there exists* $\mu > 0$ *such that* $\|\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}} - \mu\mathbf{I}\|_2 \leq \gamma\mu$, *then for* $\mathbf{U}, \triangle\mathbf{X}, \triangle\mathbf{S} \in S^n$ *the following implication holds:*

$$\left.\begin{array}{r} \triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0, \\ \mathbf{U}\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{U} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} = 0, \\ \mathbf{U}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{U} = \triangle\mathbf{X} \end{array}\right\} \Rightarrow \mathbf{U} = \triangle\mathbf{X} = \triangle\mathbf{S} = 0. \tag{3.12}$$

[4]see Appendix A.4

Proof. Assume $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$, $\gamma \in (0, \frac{1}{\sqrt{2}})$ and

$$\left.\begin{aligned} \triangle\mathbf{X} \bullet \triangle\mathbf{S} &= 0, \\ \mathbf{U}\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{U} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} &= 0, \\ \mathbf{U}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{U} &= \triangle\mathbf{X} \end{aligned}\right\} \tag{3.13}$$

If we multiply the third equation in (3.13) from the left and from the right by $\mathbf{X}^{-\frac{1}{2}}$, we obtain

$$\mathbf{X}^{-\frac{1}{2}}\mathbf{U} + \mathbf{U}\mathbf{X}^{-\frac{1}{2}} = \mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}$$

and hence, according to Proposition A.2.9,

$$\|\mathbf{U}\mathbf{X}^{-\frac{1}{2}}\|_F \leq \frac{\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F}{\sqrt{2}}. \tag{3.14}$$

Obviously, for any $\mu > 0$

$$\mu(\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} - \mathbf{X}^{-\frac{1}{2}}\mathbf{U} - \mathbf{U}\mathbf{X}^{-\frac{1}{2}}) = 0,$$

and therefore, using the second equality in (3.13), we obtain

$$\mu(\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} - \mathbf{X}^{-\frac{1}{2}}\mathbf{U} - \mathbf{U}\mathbf{X}^{-\frac{1}{2}}) + \mathbf{U}\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{U} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} = 0,$$

which can be rewritten as

$$\mu\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} = \mathbf{U}\mathbf{X}^{-\frac{1}{2}}(\mu\mathbf{I} - \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}) + (\mu\mathbf{I} - \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}})\mathbf{X}^{-\frac{1}{2}}\mathbf{U}. \tag{3.15}$$

Since

$$\mu\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} \bullet \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} = \mu\triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0,$$

we have that

$$\begin{aligned} \|\mu\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F^2 &\leq \|\mu\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F^2 + \|\mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_F^2 = \\ &= \|\mu\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_F^2. \end{aligned} \tag{3.16}$$

From (3.15), properties of matrix norm (see Proposition A.2.7) and (3.14) we obtain that

$$\begin{aligned} \|\mu\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_F &= \|\mathbf{U}\mathbf{X}^{-\frac{1}{2}}(\mu\mathbf{I} - \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}) + (\mu\mathbf{I} - \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}})\mathbf{X}^{-\frac{1}{2}}\mathbf{U}\|_F \leq \\ &\leq 2\|\mathbf{U}\mathbf{X}^{-\frac{1}{2}}(\mu\mathbf{I} - \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}})\|_F \leq 2\|\|\mathbf{U}\mathbf{X}^{-\frac{1}{2}}\|_F\|\mu\mathbf{I} - \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_2 \leq \end{aligned}$$

$$\sqrt{2}\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F\|\mu\mathbf{I}-\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_2.$$

This inequality, together with (3.16) implies

$$\mu\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F \le \sqrt{2}\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F\|\mu\mathbf{I}-\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_2$$

and hence

$$(\mu-\sqrt{2}\|\mu\mathbf{I}-\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_2)\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F \le 0.$$

From the assumptions of the lemma it follows that $\mu-\sqrt{2}\|\mu\mathbf{I}-\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}\|_2 > 0$ and therefore $\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F = 0$. Since $\mathbf{X}^{-\frac{1}{2}} \succ 0$, we have that $\triangle\mathbf{X} = 0$. Proposition A.4.8 and (3.13) imply $\mathbf{U} = 0$ and $\triangle\mathbf{S} = 0$.

□

Corollary 3.2.1 *Let $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. If there exists $\mu > 0$ such that $\|\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}-\mu\mathbf{I}\|_2 < \frac{\mu}{\sqrt{2}}$, then (3.12) holds.*

Proposition 3.2.3 *Let $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. If there exists $\mu > 0$ such that*

$$\|\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}-\mu\mathbf{I}\|_2 < \frac{\mu}{\sqrt{2}},$$

then $DF^2_{\mu,\mathbf{W}}(\mathbf{X},y,\mathbf{S})$ is a nonsingular linear map.

Proof. Assume

$$\begin{array}{l} \mathcal{A}(\triangle\mathbf{X}) = 0 \\ \mathcal{A}^*(\triangle y) + \triangle\mathbf{S} = 0 \\ \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}\mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}} + \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}} = 0 \end{array}. \tag{3.17}$$

We have to show that $\triangle\mathbf{X} = 0$, $\triangle\mathbf{S} = 0$ and $\triangle y = 0$. The first two equations in (3.17) imply $\triangle\mathbf{X}\bullet\triangle\mathbf{S} = 0$. If we denote $\mathbf{U} = \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}$, using Corollary 3.2.1 we immediately obtain $\triangle\mathbf{X} = 0$, $\triangle\mathbf{S} = 0$. Assumption (A1) implies $\triangle y = 0$.

□

3.2.3 Nonsingularity of $DF^3_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$

Recall that if $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$ then

$$DF^3_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] =$$

$$= \begin{bmatrix} \mathcal{A}(\triangle\mathbf{X}) \\ \mathcal{A}^*(\triangle y) + \triangle\mathbf{S} \\ [[\triangle\mathbf{X}]]^T_{\mathbf{L_X}}\mathbf{S}\mathbf{L_X} + \mathbf{L_X}{}^T\mathbf{S}[[\triangle\mathbf{X}]]_{\mathbf{L_X}} + \mathbf{L_X}{}^T\triangle\mathbf{S}\mathbf{L_X} \end{bmatrix}$$

where $[[\triangle\mathbf{X}]]_{\mathbf{L_X}}$ is the unique solution $\mathbf{H} \in L^n$ of the equation $\mathbf{L_X}\mathbf{H}^T + \mathbf{H}\mathbf{L_X}{}^T = \triangle\mathbf{X}$ (see Definition B.2.1 (b)).

Proposition 3.2.4 *If $\mathbf{X} \succ 0, \mathbf{S} \succ 0$ and $\mathbf{L}^T_{\mathbf{X}}\mathbf{S}\mathbf{L_X} \in D^n_{++}$, then for $\mathbf{U} \in L^n$ and $\triangle\mathbf{X}, \triangle\mathbf{S} \in S^n$ the following implication holds:*

$$\left.\begin{aligned} \triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0, \\ \mathbf{U}^T\mathbf{S}\mathbf{L_X} + \mathbf{L_X}{}^T\mathbf{S}\mathbf{U} + \mathbf{L_X}{}^T\triangle\mathbf{S}\mathbf{L_X} = 0, \\ \mathbf{L_X}\mathbf{U}^T + \mathbf{U}\mathbf{L}^T_{\mathbf{X}} = \triangle\mathbf{X} \end{aligned}\right\} \Rightarrow \mathbf{U} = \triangle\mathbf{X} = \triangle\mathbf{S} = 0. \tag{3.18}$$

Proof. Assume that $\mathbf{X} \succ 0, \mathbf{S} \succ 0$, $\mathbf{L}^T_{\mathbf{X}}\mathbf{S}\mathbf{L_X} \in D^n_{++}$ and

$$\begin{aligned} \triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0, \\ \mathbf{U}^T\mathbf{S}\mathbf{L_X} + \mathbf{L_X}{}^T\mathbf{S}\mathbf{U} + \mathbf{L_X}{}^T\triangle\mathbf{S}\mathbf{L_X} = 0, \\ \mathbf{L_X}\mathbf{U}^T + \mathbf{U}\mathbf{L}^T_{\mathbf{X}} = \triangle\mathbf{X} \end{aligned}$$

Since $\mathbf{L_X} \in L^n_{++}$, we can express

$$\triangle\mathbf{S} = -(\mathbf{L_X}{}^{-T}\mathbf{U}\mathbf{S} + \mathbf{S}\mathbf{U}\mathbf{L}^{-1}_{\mathbf{X}})$$

and obtain

$$0 = -\triangle\mathbf{X} \bullet \triangle\mathbf{S} = (\mathbf{L_X}\mathbf{U}^T + \mathbf{U}\mathbf{L}^T_{\mathbf{X}}) \bullet (\mathbf{L_X}{}^{-T}\mathbf{U}\mathbf{S} + \mathbf{S}\mathbf{U}\mathbf{L}^{-1}_{\mathbf{X}}) =$$

$$2tr(\mathbf{U}^T\mathbf{S}\mathbf{U}) + 2tr[(\mathbf{L}^T_{\mathbf{X}}\mathbf{S}\mathbf{L_X})(\mathbf{U}\mathbf{L}^{-1}_{\mathbf{X}})^2].$$

From the assumptions and properties of triangular matrices[5] we have that $tr(\mathbf{U}^T\mathbf{S}\mathbf{U}) = 0$ and therefore also $\mathbf{U} = 0, \triangle\mathbf{X} = 0, \triangle\mathbf{S} = 0$.

□

[5] see Appendix A.5

Note. It can be easily seen that the condition $\mathbf{L}_\mathbf{X}^T\mathbf{S}\mathbf{L}_\mathbf{X} \in D_{++}^n$ is equivalent to $\mathbf{XS} \in L_{++}^n$ (follows from Proposition A.5.3 (c),(j)).

The following lemma is a consequence of Lemma 2.4 of [50].

Lemma 3.2.2 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$ *and* $\gamma \in (0, \frac{1}{\sqrt{2}})$. *If there exists* $\mu > 0$ *such that* $\|\mathbf{L}_\mathbf{X}^T\mathbf{S}\mathbf{L}_\mathbf{X} - \mu\mathbf{I}\|_2 \le \gamma\mu$, *then* (3.18) *holds.*

Proof. Assume $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$, $\gamma \in (0, \frac{1}{\sqrt{2}})$ and

$$\left.\begin{array}{r}\triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0,\\ \mathbf{U}^T\mathbf{S}\mathbf{L}_\mathbf{X} + {\mathbf{L}_\mathbf{X}}^T\mathbf{S}\mathbf{U} + {\mathbf{L}_\mathbf{X}}^T\triangle\mathbf{S}\mathbf{L}_\mathbf{X} = 0,\\ \mathbf{L}_\mathbf{X}\mathbf{U}^T + \mathbf{U}\mathbf{L}_\mathbf{X}^T = \triangle\mathbf{X}\end{array}\right\} \tag{3.19}$$

If we multiply the third equation in (3.19) by $\mathbf{L}_\mathbf{X}^{-1}$ from the left by $\mathbf{L}_\mathbf{X}^{-T}$ from the right, we obtain

$$\mathbf{U}^T\mathbf{L}_\mathbf{X}^{-T} + \mathbf{L}_\mathbf{X}^{-1}\mathbf{U} = \mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T}$$

and hence, according to Proposition A.5.6

$$\|\mathbf{U}^T\mathbf{L}_\mathbf{X}^{-T}\|_F \le \frac{\|\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T}\|_F}{\sqrt{2}}. \tag{3.20}$$

Obviously, for any $\mu > 0$

$$\mu(\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T} - \mathbf{U}^T\mathbf{L}_\mathbf{X}^{-T} + \mathbf{L}_\mathbf{X}^{-1}\mathbf{U}) = 0,$$

and therefore, using the second equality in (3.19), we obtain

$$\mu(\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T} - \mathbf{U}^T\mathbf{L}_\mathbf{X}^{-T} + \mathbf{L}_\mathbf{X}^{-1}\mathbf{U}) + \mathbf{U}^T\mathbf{S}\mathbf{L}_\mathbf{X} + {\mathbf{L}_\mathbf{X}}^T\mathbf{S}\mathbf{U} + {\mathbf{L}_\mathbf{X}}^T\triangle\mathbf{S}\mathbf{L}_\mathbf{X} = 0$$

which can be rewritten as

$$\mu\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T} + {\mathbf{L}_\mathbf{X}}^T\triangle\mathbf{S}\mathbf{L}_\mathbf{X} = \mathbf{U}^T\mathbf{L}_\mathbf{X}^{-T}(\mu\mathbf{I} - \mathbf{L}_\mathbf{X}^T\mathbf{S}\mathbf{L}_\mathbf{X}) + (\mu\mathbf{I} - \mathbf{L}_\mathbf{X}^T\mathbf{S}\mathbf{L}_\mathbf{X})\mathbf{L}_\mathbf{X}^{-1}\mathbf{U}. \tag{3.21}$$

Since

$$\mu\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T} \bullet {\mathbf{L}_\mathbf{X}}^T\triangle\mathbf{S}\mathbf{L}_\mathbf{X} = \mu\triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0,$$

we have that

$$\|\mu\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T}\|_F^2 \le \|\mu\mathbf{L}_\mathbf{X}^{-1}\triangle\mathbf{X}\mathbf{L}_\mathbf{X}^{-T}\|_F^2 + \|{\mathbf{L}_\mathbf{X}}^T\triangle\mathbf{S}\mathbf{L}_\mathbf{X}\|_F^2 =$$

$$= \|\mu \mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T} + \mathbf{L}_{\mathbf{X}}{}^{T} \triangle \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_F^2. \tag{3.22}$$

From (3.21), properties of matrix norm (see Proposition A.2.7) and (3.20) we obtain that

$$\|\mu \mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T} + \mathbf{L}_{\mathbf{X}}{}^{T} \triangle \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_F = \|\mathbf{U}^T \mathbf{L}_{\mathbf{X}}^{-T} (\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}) + (\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}) \mathbf{L}_{\mathbf{X}}^{-1} \mathbf{U}\|_F \leq$$
$$\leq 2\|\mathbf{U}^T \mathbf{L}_{\mathbf{X}}^{-T} (\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}})\|_F \leq 2\|\mathbf{U}^T \mathbf{L}_{\mathbf{X}}^{-T}\|_F \|\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_2 \leq$$
$$\sqrt{2}\|\mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T}\|_F \|\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_2.$$

This inequality, together with (3.22), implies

$$\mu \|\mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T}\|_F \leq \sqrt{2}\|\mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T}\|_F \|\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_2$$

and hence

$$(\mu - \sqrt{2}\|\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_2) \|\mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T}\|_F \leq 0.$$

From the assumptions of the lemma it follows that $\mu - \sqrt{2}\|\mu \mathbf{I} - \mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}}\|_2 > 0$ and therefore $\|\mathbf{L}_{\mathbf{X}}^{-1} \triangle \mathbf{X} \mathbf{L}_{\mathbf{X}}^{-T}\|_F = 0$. Since $\mathbf{L}_{\mathbf{X}}^{-1} \in L_{++}^n$, we have that $\triangle \mathbf{X} = 0$. Proposition A.5.5 and (3.19) imply $\mathbf{U} = 0$ and $\triangle \mathbf{S} = 0$, respectively.

□

Corollary 3.2.2 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. *If there exists* $\mu > 0$ *such that* $\|\mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}} - \mu \mathbf{I}\|_2 < \frac{\mu}{\sqrt{2}}$, *then* (3.18) *holds.*

Proposition 3.2.5 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. *If*
(a) there exists $\mu > 0$ *such that* $\|\mathbf{L}_{\mathbf{X}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}} - \mu \mathbf{I}\|_2 < \frac{\mu}{\sqrt{2}}$,
or
(b) $\mathbf{L}_{\mathbf{X}}{}^{T} \mathbf{S} \mathbf{L}_{\mathbf{X}} \in D_{++}^n$,
then $DF_{\mu,\mathbf{W}}^3(\mathbf{X}, y, \mathbf{S})$ *is a nonsingular linear map.*

Proof. Assume

$$\begin{array}{c} \mathcal{A}(\triangle \mathbf{X}) = 0 \\ \mathcal{A}^*(\triangle y) + \triangle \mathbf{S} = 0 \\ [[\triangle \mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}}^T \mathbf{S} \mathbf{L}_{\mathbf{X}} + \mathbf{L}_{\mathbf{X}}{}^{T} \mathbf{S} [[\triangle \mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}} + \mathbf{L}_{\mathbf{X}}{}^{T} \triangle \mathbf{S} \mathbf{L}_{\mathbf{X}} = 0 \end{array} \tag{3.23}$$

We have to show that $\triangle \mathbf{X} = 0$, $\triangle \mathbf{S} = 0$ and $\triangle y = 0$. The first two equations in (3.23) imply $\triangle \mathbf{X} \bullet \triangle \mathbf{S} = 0$. If we denote $\mathbf{U} = [[\triangle \mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}}$, using Corollary 3.2.2 in the case (a), and using Proposition 3.2.4 in the case (b), we immediately obtain $\triangle \mathbf{X} = 0$, $\triangle \mathbf{S} = 0$. Assumption (A1) implies $\triangle y = 0$.

□

3.2.4 Nonsingularity of $DF^4_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$

Recall that if $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$, then

$$DF^4_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] =$$

$$= \begin{bmatrix} \mathcal{A}(\triangle\mathbf{X}) \\ \tilde{\mathcal{A}}(\triangle y) + \triangle\mathbf{S} \\ \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}} + \langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}} \end{bmatrix}.$$

where $\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}$ is the unique solution $\mathbf{H}$ of the equation $\mathbf{X}^{\frac{1}{2}}\mathbf{H} + \mathbf{H}\mathbf{X}^{\frac{1}{2}} = \triangle\mathbf{X}$ and $\langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}}$ is the unique solution $\mathbf{H}$ of the equation $\mathbf{S}^{\frac{1}{2}}\mathbf{H} + \mathbf{H}\mathbf{S}^{\frac{1}{2}} = \triangle\mathbf{S}$ (see Definition B.2.1 (a)).

The following lemma is a consequence of Proposition 4 of [51]. As in the case of Lemma 3.2.2 the whole proof is included.

Lemma 3.2.3 *Let $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$ and $\gamma \in (0, \frac{1}{3\sqrt{2}})$ be given. If there exists $\mu > 0$ such that $\|(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})/2 - \mu\mathbf{I}\|_F \leq \gamma\mu$, then for $\mathbf{U}, \mathbf{V}, \triangle\mathbf{X}, \triangle\mathbf{S} \in S^n$ the following implication holds:*

$$\left.\begin{array}{r} \triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0 \\ \mathbf{V}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{V} + \mathbf{S}\mathbf{X}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{U} = 0 \\ \mathbf{U}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{U} = \triangle\mathbf{X} \\ \mathbf{V}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{V} = \triangle\mathbf{S} \end{array}\right\} \Rightarrow \mathbf{U} = \mathbf{V} = \triangle\mathbf{X} = \triangle\mathbf{S} = 0. \tag{3.24}$$

Proof. If we multiply the third equation in (3.24) from the left and right by $\mathbf{X}^{-\frac{1}{2}}$ and the last equation in (3.24) from the left and right by $\mathbf{S}^{-\frac{1}{2}}$, we obtain

$$\mathbf{X}^{-\frac{1}{2}}\mathbf{U} + \mathbf{U}\mathbf{X}^{-\frac{1}{2}} = \mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}} \text{ and } \mathbf{S}^{-\frac{1}{2}}\mathbf{V} + \mathbf{V}\mathbf{S}^{-\frac{1}{2}} = \mathbf{S}^{-\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{-\frac{1}{2}}.$$

From Proposition A.2.9 we have that

$$\|\mathbf{X}^{-\frac{1}{2}}\mathbf{U}\|_F \leq \frac{\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F}{\sqrt{2}} \text{ and } \|\mathbf{S}^{-\frac{1}{2}}\mathbf{V}\|_F \leq \frac{\|\mathbf{S}^{-\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{-\frac{1}{2}}\|_F}{\sqrt{2}} \tag{3.25}$$

Denote

$$\overline{\triangle\mathbf{X}} = \mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{S}^{\frac{1}{2}} \text{ and } \overline{\triangle\mathbf{S}} = \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{-\frac{1}{2}}.$$

Obviously

$$\overline{\triangle\mathbf{X}} = 0 \Leftrightarrow \triangle\mathbf{X} = 0, \text{ and } \overline{\triangle\mathbf{S}} = 0 \Leftrightarrow \triangle\mathbf{S} = 0$$

and

$$\overline{\triangle\mathbf{X}} \bullet \overline{\triangle\mathbf{S}} = tr[(\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{S}^{\frac{1}{2}})(\mathbf{S}^{-\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{\frac{1}{2}})] =$$

$$= tr[\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\triangle\mathbf{S}\mathbf{S}^{\frac{1}{2}}] = tr[\triangle\mathbf{X}\triangle\mathbf{S}] = 0. \tag{3.26}$$

By multiplying the third equation in (3.24) by $\mathbf{X}^{-\frac{1}{2}}$ from the left and by $\mathbf{S}^{\frac{1}{2}}$ from the right and by multiplying the fourth equation in (3.24) by $\mathbf{S}^{-\frac{1}{2}}$ from the right and by $\mathbf{X}^{\frac{1}{2}}$ from the left we obtain that

$$\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{S}^{\frac{1}{2}} = \mathbf{X}^{-\frac{1}{2}}\mathbf{U}\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{U}\mathbf{S}^{\frac{1}{2}} \qquad \mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{-\frac{1}{2}} = \mathbf{X}^{\frac{1}{2}}\mathbf{V} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{V}\mathbf{S}^{-\frac{1}{2}}.$$

From this it follows

$$\overline{\triangle\mathbf{X}} + \overline{\triangle\mathbf{S}} = \mathbf{X}^{-\frac{1}{2}}\mathbf{U}\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{U}\mathbf{S}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{V} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{V}\mathbf{S}^{-\frac{1}{2}} = \tag{3.27}$$

$$= \mathbf{X}^{-\frac{1}{2}}\mathbf{U}\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mathbf{S}^{\frac{1}{2}}\mathbf{U} - \mathbf{V}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{V}\mathbf{S}^{-\frac{1}{2}} =$$

$$= \mathbf{X}^{-\frac{1}{2}}\mathbf{U}\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mathbf{S}^{\frac{1}{2}}(\mathbf{X}^{\frac{1}{2}}\mathbf{X}^{-\frac{1}{2}})\mathbf{U} - \mathbf{V}(\mathbf{S}^{-\frac{1}{2}}\mathbf{S}^{\frac{1}{2}})\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{V}\mathbf{S}^{-\frac{1}{2}} =$$

(we add zero in the form $\mu(\mathbf{X}^{-\frac{1}{2}}\mathbf{U} - \mathbf{X}^{-\frac{1}{2}}\mathbf{U} + \mathbf{V}\mathbf{S}^{-\frac{1}{2}} - \mathbf{V}\mathbf{S}^{-\frac{1}{2}})$, where $\mu > 0$)

$$= \mathbf{X}^{-\frac{1}{2}}\mathbf{U}\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{X}^{-\frac{1}{2}}\mathbf{U} + \mu\mathbf{X}^{-\frac{1}{2}}\mathbf{U} - \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}}\mathbf{X}^{-\frac{1}{2}}\mathbf{U} -$$

$$-\mathbf{V}\mathbf{S}^{-\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}} + \mu\mathbf{V}\mathbf{S}^{-\frac{1}{2}} - \mu\mathbf{V}\mathbf{S}^{-\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{V}\mathbf{S}^{-\frac{1}{2}}$$

$$= \mathbf{X}^{-\frac{1}{2}}\mathbf{U}(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{I}) + (\mu\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})\mathbf{X}^{-\frac{1}{2}}\mathbf{U} + \mathbf{V}\mathbf{S}^{-\frac{1}{2}}(\mu\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}}) + (\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{I})\mathbf{V}\mathbf{S}^{-\frac{1}{2}}.$$

Proposition A.2.12 implies that if $\|(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})/2 - \mu\mathbf{I}\|_F \leq \gamma\mu$, then

$$\|\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{I}\|_F \leq \sqrt{2}\gamma\mu \tag{3.28}$$

and

$$\|\mathbf{X}^{-\frac{1}{2}}\mathbf{S}^{-\frac{1}{2}}\|_2 \leq \frac{1}{(1-\sqrt{2}\gamma)\mu}. \tag{3.29}$$

From (3.26) it follows

$$\|\overline{\triangle\mathbf{X}} + \overline{\triangle\mathbf{S}}\|_F^2 = \|\overline{\triangle\mathbf{X}}\|_F^2 + 2\overline{\triangle\mathbf{X}} \bullet \overline{\triangle\mathbf{S}} + \|\overline{\triangle\mathbf{S}}\|_F^2 = \|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2$$

and hence from (3.27) we have

$$(\|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2)^{\frac{1}{2}} =$$

$$= \|\mathbf{X}^{-\frac{1}{2}}\mathbf{U}(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{I}) + (\mu\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})\mathbf{X}^{-\frac{1}{2}}\mathbf{U} + \mathbf{V}\mathbf{S}^{-\frac{1}{2}}(\mu\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}}) + (\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{I})\mathbf{V}\mathbf{S}^{-\frac{1}{2}}\|_F \leq$$

(from properties of matrix norm)

$$\leq 2\|\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} - \mu\mathbf{I}\|_F(\|\mathbf{V}\mathbf{S}^{-\frac{1}{2}}\|_F + \|\mathbf{X}^{-\frac{1}{2}}\mathbf{U}\|_F) \leq$$

(using (3.28))

$$\leq 2\sqrt{2}\gamma\mu(\|\mathbf{S}^{-\frac{1}{2}}\mathbf{V}\|_F + \|\mathbf{X}^{-\frac{1}{2}}\mathbf{U}\|_F) \leq$$

(using (3.25))

$$\leq 2\gamma\mu(\|\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{X}^{-\frac{1}{2}}\|_F + \|\mathbf{S}^{-\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{-\frac{1}{2}}\|_F) =$$

$$= 2\gamma\mu(\|(\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{S}^{\frac{1}{2}})(\mathbf{S}^{-\frac{1}{2}}\mathbf{X}^{-\frac{1}{2}})\|_F + \|(\mathbf{S}^{-\frac{1}{2}}\mathbf{X}^{-\frac{1}{2}})(\mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{-\frac{1}{2}})\|_F) \leq$$

(from statement (b) of Proposition A.2.7)

$$\leq 2\gamma\mu\|\mathbf{S}^{-\frac{1}{2}}\mathbf{X}^{-\frac{1}{2}}\|_2((\|\overline{\triangle\mathbf{X}}\|_F + \|\overline{\triangle\mathbf{S}}\|_F) \leq$$

(using (3.29) and the inequality: if $a \geq 0, b \geq 0$, then $a + b \leq \sqrt{2}(a^2 + b^2)^{\frac{1}{2}}$)

$$\leq \frac{2\sqrt{2}\gamma}{1 - \sqrt{2}\gamma}(\|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2)^{\frac{1}{2}}$$

Since $\frac{2\sqrt{2}\gamma}{1-\sqrt{2}\gamma} < 1$, it holds $(\|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2)^{\frac{1}{2}} = 0$ and therefore also $\triangle\mathbf{X} = \triangle\mathbf{S} = 0$.

□

Corollary 3.2.3 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. *If there exists* $\mu > 0$ *such that*

$$\|(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})/2 - \mu\mathbf{I}\|_F < \frac{1}{3\sqrt{2}}\mu,$$

then (3.24) *holds.*

Proposition 3.2.6 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. *If there exists* $\mu > 0$ *such that*

$$\|(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})/2 - \mu\mathbf{I}\|_F < \frac{1}{3\sqrt{2}}\mu,$$

then $DF^4_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ *is a nonsingular linear map.*

Proof. Assume

$$\begin{array}{c}\mathcal{A}(\triangle\mathbf{X}) = 0\\ \tilde{\mathcal{A}}(\triangle y) + \triangle\mathbf{S} = 0\\ \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}\mathbf{S}^{\frac{1}{2}} + \mathbf{S}^{\frac{1}{2}}\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}} + \langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}} = 0.\end{array} \tag{3.30}$$

We have to show that $\triangle\mathbf{X} = 0$, $\triangle\mathbf{S} = 0$ and $\triangle y = 0$. The first two equations in (3.30) imply $\triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0$. If we denote $\mathbf{U} = \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}$, and $\mathbf{V} = \langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}}$, using Corollary 3.2.3 we immediately obtain $\triangle\mathbf{X} = 0$, $\triangle\mathbf{S} = 0$. Assumption (A1) implies $\triangle y = 0$.

□

3.2.5 Nonsingularity of $DF^5_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$

Recall that if $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$, then

$$DF^5_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})[\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}] =$$

$$= \begin{bmatrix} \mathcal{A}(\triangle\mathbf{X}) \\ \tilde{\mathcal{A}}(\triangle y) + \triangle\mathbf{S} \\ [[\triangle\mathbf{X}]]^T_{\mathbf{L_X}}\mathbf{U_S} + \mathbf{U_S}^T[[\triangle\mathbf{X}]]_{\mathbf{L_X}} + [[\triangle\mathbf{S}]]^T_{\mathbf{U_S}}\mathbf{L_X} + \mathbf{L_X}^T[[\triangle\mathbf{S}]]_{\mathbf{U_S}} \end{bmatrix}.$$

where $[[\triangle\mathbf{X}]]_{\mathbf{L_X}}$ is the unique solution $\mathbf{H}$ of the equation $\mathbf{L_X}\mathbf{H}^T + \mathbf{H}\mathbf{L_X}^T = \triangle\mathbf{X}$ and $[[\triangle\mathbf{S}]]_{\mathbf{U_S}}$ is the unique solution $\mathbf{H}$ of the equation $\mathbf{U_S}\mathbf{H}^T + \mathbf{H}\mathbf{U_S}^T = \triangle\mathbf{S}$.

The following lemma is a consequence of Proposition 5 of [51].

Lemma 3.2.4 *Let $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$ and $\gamma \in (0, \frac{1}{3\sqrt{2}})$ be given. If there exists $\mu > 0$ such that $\|(\mathbf{U_S}^T\mathbf{L_X} + \mathbf{L_X}^T\mathbf{U_S})/2 - \mu\mathbf{I}\|_F \leq \gamma\mu$, then for $\mathbf{L} \in L^n, \mathbf{U} \in U^n$ and $\triangle\mathbf{X}, \triangle\mathbf{S} \in S^n$ the following implication holds:*

$$\left.\begin{array}{r} \triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0 \\ \mathbf{L}^T\mathbf{U_S} + \mathbf{U_S}^T\mathbf{L} + \mathbf{U}^T\mathbf{L_X} + \mathbf{L_X}^T\mathbf{U} = 0 \\ \mathbf{L_X}\mathbf{L}^T + \mathbf{L}\mathbf{L_X}^T = \triangle\mathbf{X} \\ \mathbf{U_S}\mathbf{U}^T + \mathbf{U}\mathbf{U_S}^T = \triangle\mathbf{S} \end{array}\right\} \Rightarrow \mathbf{U} = \mathbf{L} = \triangle\mathbf{X} = \triangle\mathbf{S} = 0. \tag{3.31}$$

Proof. If we multiply the third equation in (3.31) from the left by $\mathbf{L_X}^{-1}$ and from the right by $\mathbf{L_X}^{-T}$ and the last equation in (3.31) from the left $\mathbf{U_S}^{-1}$ and from the right by $\mathbf{U_S}^{-1}$, we obtain

$$\mathbf{L}^T\mathbf{L_X}^{-T} + \mathbf{L_X}^{-1}\mathbf{L} = \mathbf{L_X}^{-1}\triangle\mathbf{X}\mathbf{L_X}^{-T} \text{ and } \mathbf{U}^T\mathbf{U_S}^{-T} + \mathbf{U_S}^{-1}\mathbf{U} = \mathbf{U_S}^{-1}\triangle\mathbf{S}\mathbf{U_S}^{-T}.$$

From Proposition A.5.6 we have that

$$\left.\begin{array}{l} \|\mathbf{L}^T\mathbf{L_X}^{-T}\|_F = \|\mathbf{L_X}^{-1}\mathbf{L}\|_F \leq \frac{\|\mathbf{L_X}^{-1}\triangle\mathbf{X}\mathbf{L_X}^{-T}\|_F}{\sqrt{2}} \\ \|\mathbf{U}^T\mathbf{U_S}^{-T}\|_F = \|\mathbf{U_S}^{-1}\mathbf{U}\|_F \leq \frac{\|\mathbf{U_S}^{-1}\triangle\mathbf{S}\mathbf{U_S}^{-T}\|_F}{\sqrt{2}} \end{array}\right\}. \tag{3.32}$$

Denote

$$\overline{\triangle\mathbf{X}} = \mathbf{U_S}^T\triangle\mathbf{X}\mathbf{L_X}^{-T} \text{ and } \overline{\triangle\mathbf{S}} = \mathbf{U_S}^{-1}\triangle\mathbf{S}\mathbf{L_X}$$

Obviously,

$$\overline{\triangle\mathbf{X}} = 0 \Leftrightarrow \triangle\mathbf{X} = 0, \text{ and } \overline{\triangle\mathbf{S}} = 0 \Leftrightarrow \triangle\mathbf{S} = 0.$$

Hence

$$\overline{\triangle\mathbf{X}} \bullet \overline{\triangle\mathbf{S}} = tr[(\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\mathbf{S}^{\frac{1}{2}})(\mathbf{S}^{-\frac{1}{2}}\triangle\mathbf{S}\mathbf{S}^{\frac{1}{2}})] =$$

$$= tr[\mathbf{X}^{-\frac{1}{2}}\triangle\mathbf{X}\triangle\mathbf{S}\mathbf{S}^{\frac{1}{2}}] = tr[\triangle\mathbf{X}\triangle\mathbf{S}] = 0. \tag{3.33}$$

By multiplying the third equation in (3.31) by ${\mathbf{U_S}}^T$ from the left and by ${\mathbf{L_X}}^{-T}$ from the right and by multiplying the fourth equation in (3.31) by $\mathbf{L_X}$ a from the right and by ${\mathbf{U_S}}^{-1}$ from the left we obtain that

$$\overline{\triangle\mathbf{X}} = {\mathbf{U_S}}^T(\mathbf{L_X}\mathbf{L}^T + \mathbf{L}{\mathbf{L_X}}^T){\mathbf{L_X}}^{-T} = {\mathbf{U_S}}^T\mathbf{L_X}\mathbf{L}^T{\mathbf{L_X}}^{-T} + {\mathbf{U_S}}^T\mathbf{L},$$

$$\overline{\triangle\mathbf{S}} = {\mathbf{U_S}}^{-1}(\mathbf{U_S}\mathbf{U}^T + \mathbf{U}{\mathbf{U_S}}^T)\mathbf{L_X} = \mathbf{U}^T\mathbf{L_X} + {\mathbf{U_S}}^{-1}\mathbf{U}{\mathbf{U_S}}^T\mathbf{L_X}.$$

From this it follows

$$\overline{\triangle\mathbf{X}} + \overline{\triangle\mathbf{S}} = {\mathbf{U_S}}^T\mathbf{L_X}\mathbf{L}^T{\mathbf{L_X}}^{-T} + {\mathbf{U_S}}^T\mathbf{L} + \mathbf{U}^T\mathbf{L_X} + {\mathbf{U_S}}^{-1}\mathbf{U}{\mathbf{U_S}}^T\mathbf{L_X} = \tag{3.34}$$

$$= {\mathbf{U_S}}^T\mathbf{L_X}\mathbf{L}^T{\mathbf{L_X}}^{-T} - \mathbf{L}^T\mathbf{U_S} - {\mathbf{L_X}}^T\mathbf{U} + {\mathbf{U_S}}^{-1}\mathbf{U}{\mathbf{U_S}}^T\mathbf{L_X} =$$

$$= {\mathbf{U_S}}^T\mathbf{L_X}\mathbf{L}^T{\mathbf{L_X}}^{-T} - \mathbf{L}^T({\mathbf{L_X}}^{-T}{\mathbf{L_X}}^T)\mathbf{U_S} - {\mathbf{L_X}}^T(\mathbf{U_S}{\mathbf{U_S}}^{-1})\mathbf{U} + {\mathbf{U_S}}^{-1}\mathbf{U}{\mathbf{U_S}}^T\mathbf{L_X} =$$

(we add zero in the form $\mu(\mathbf{L}^T{\mathbf{L_X}}^{-T} - \mathbf{L}^T{\mathbf{L_X}}^{-T} + {\mathbf{U_S}}^{-1}\mathbf{U} - {\mathbf{U_S}}^{-1}\mathbf{U})$, where $\mu > 0$)

$$= {\mathbf{U_S}}^T\mathbf{L_X}\mathbf{L}^T{\mathbf{L_X}}^{-T} - \mu\mathbf{L}^T{\mathbf{L_X}}^{-T} + \mu\mathbf{L}^T{\mathbf{L_X}}^{-T} - \mathbf{L}^T{\mathbf{L_X}}^{-T}{\mathbf{L_X}}^T\mathbf{U_S} -$$

$$-{\mathbf{L_X}}^T\mathbf{U_S}{\mathbf{U_S}}^{-1}\mathbf{U} + \mu{\mathbf{U_S}}^{-1}\mathbf{U} - \mu{\mathbf{U_S}}^{-1}\mathbf{U} + {\mathbf{U_S}}^{-1}\mathbf{U}{\mathbf{U_S}}^T\mathbf{L_X}$$

$$= ({\mathbf{U_S}}^T\mathbf{L_X} - \mu\mathbf{I})\mathbf{L}^T{\mathbf{L_X}}^{-T} + \mathbf{L}^T{\mathbf{L_X}}^{-T}(\mu\mathbf{I} - {\mathbf{L_X}}^T\mathbf{U_S}) +$$

$$(\mu\mathbf{I} - {\mathbf{L_X}}^T\mathbf{U_S}){\mathbf{U_S}}^{-1}\mathbf{U} + {\mathbf{U_S}}^{-1}\mathbf{U}({\mathbf{U_S}}^T\mathbf{L_X} - \mu\mathbf{I}).$$

From (3.33) it follows

$$\|\overline{\triangle\mathbf{X}} + \overline{\triangle\mathbf{S}}\|_F^2 = \|\overline{\triangle\mathbf{X}}\|_F^2 + 2\overline{\triangle\mathbf{X}} \bullet \overline{\triangle\mathbf{S}} + \|\overline{\triangle\mathbf{S}}\|_F^2 = \|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2$$

and hence from (3.34) we have

$$(\|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2)^{\frac{1}{2}} =$$

$$= \|({\mathbf{U_S}}^T\mathbf{L_X} - \mu\mathbf{I})\mathbf{L}^T{\mathbf{L_X}}^{-T} + \mathbf{L}^T{\mathbf{L_X}}^{-T}(\mu\mathbf{I} - {\mathbf{L_X}}^T\mathbf{U_S}) +$$

$$(\mu\mathbf{I} - {\mathbf{L_X}}^T\mathbf{U_S}){\mathbf{U_S}}^{-1}\mathbf{U} + {\mathbf{U_S}}^{-1}\mathbf{U}({\mathbf{U_S}}^T\mathbf{L_X} - \mu\mathbf{I})\|_F \leq$$

(from properties of matrix norm)

$$\leq 2\|{\mathbf{U_S}}^T\mathbf{L_X} - \mu\mathbf{I}\|_F(\|\mathbf{L}^T{\mathbf{L_X}}^{-T}\|_F + \|{\mathbf{U_S}}^{-1}\mathbf{U}\|_F) \leq$$

(by applying the statement (a) of Proposition A.2.12)

$$\leq 2\sqrt{2}\gamma\mu(\|\mathbf{L}^T{\mathbf{L_X}}^{-T}\|_F + \|{\mathbf{U_S}}^{-1}\mathbf{U}\|_F) \leq$$

(by using (3.25))

$$\leq 2\gamma\mu(\|\mathbf{L_X}^{-1}\triangle\mathbf{X}\mathbf{L_X}^{-T}\|_F + \|\mathbf{U_S}^{-1}\triangle\mathbf{S}\mathbf{U_S}^{-T}\|_F) =$$

$$= 2\gamma\mu(\|(\mathbf{L_X}^{-1}\mathbf{U_S}^{-T})(\mathbf{U_S}^T\triangle\mathbf{X}\mathbf{L_X}^{-T})\|_F + \|(\mathbf{U_S}^{-1}\triangle\mathbf{S}\mathbf{L_X})(\mathbf{L_X}^{-1}\mathbf{U_S}^{-T})\|_F) \leq$$

(from statement (b) of Proposition A.2.7)

$$\leq 2\gamma\mu\|\mathbf{L_X}^{-1}\mathbf{U_S}^{-T}\|_2((\|\overline{\triangle\mathbf{X}}\|_F + \|\overline{\triangle\mathbf{S}}\|_F) \leq$$

(from statement (b) of Proposition A.2.12 and from the inequality: if $a \geq 0, b \geq 0$, then $a + b \leq \sqrt{2}(a^2 + b^2)^{\frac{1}{2}}$)

$$\leq \frac{2\sqrt{2}\gamma}{1-\sqrt{2}\gamma}(\|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2)^{\frac{1}{2}}$$

Since $\frac{2\sqrt{2}\gamma}{1-\sqrt{2}\gamma} < 1$, it holds $(\|\overline{\triangle\mathbf{X}}\|_F^2 + \|\overline{\triangle\mathbf{S}}\|_F^2)^{\frac{1}{2}} = 0$ and therefore also $\triangle\mathbf{X} = \triangle\mathbf{S} = 0$.

□

Corollary 3.2.4 *Let $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. If there exists $\mu > 0$ such that*

$$\|(\mathbf{U_S}^T\mathbf{L_X} + \mathbf{L_X}^T\mathbf{U_S})/2 - \mu\mathbf{I}\|_F \leq \frac{1}{3\sqrt{2}}\mu,$$

then (3.31) *holds.*

Proposition 3.2.7 *Let $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. If there exists $\mu > 0$ such that*

$$\|(\mathbf{U_S}^T\mathbf{L_X} + \mathbf{L_X}^T\mathbf{U_S})/2 - \mu\mathbf{I}\|_F \leq \frac{1}{3\sqrt{2}}\mu,$$

then $DF^5_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$ is a nonsingular linear map.

Proof. Assume

$$\begin{array}{c} \mathcal{A}(\triangle\mathbf{X}) = 0 \\ \tilde{\mathcal{A}}(\triangle y) + \triangle\mathbf{S} = 0 \\ [[\triangle\mathbf{X}]]_{\mathbf{L_X}}^T\mathbf{U_S} + \mathbf{U_S}^T[[\triangle\mathbf{X}]]_{\mathbf{L_X}} + [[\triangle\mathbf{S}]]_{\mathbf{U_S}}^T\mathbf{L_X} + \mathbf{L_X}^T[[\triangle\mathbf{S}]]_{\mathbf{U_S}} = 0. \end{array} \tag{3.35}$$

We have to show that $\triangle\mathbf{X} = 0$, $\triangle\mathbf{S} = 0$ and $\triangle y = 0$. The first two equations in (3.35) imply $\triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0$. If we denote $\mathbf{L} = [[\triangle\mathbf{X}]]_{\mathbf{L_X}}$, and $\mathbf{U} = [[\triangle\mathbf{S}]]_{\mathbf{U_S}}$, using Corollary 3.2.4 we immediately obtain $\triangle\mathbf{X} = 0$, $\triangle\mathbf{S} = 0$. Assumption (A1) implies $\triangle y = 0$.

□

3.3 Sets of suitable weights

In this section the results of the previous section will be used for a definition of sets of suitable weights for particular type of symmetrization. First define a set $\mathcal{M}_\varepsilon$ as follows.

For $\varepsilon > 0$ denote

$$\mathcal{M}_\varepsilon = \{ \ \mathbf{W} \in S^n_{++}; \exists \nu : \ \ \|\mathbf{W} - \nu\mathbf{I}\|_2 < \varepsilon\nu\}.$$

It can be easily seen that these sets are conic neighborhoods of the identity matrix $\mathbf{I}$. The following lemma provides a nice description of $\mathcal{M}_\varepsilon$.

Lemma 3.3.1 *The set $\mathcal{M}_\varepsilon$ is a convex cone. Moreover, if $\varepsilon \in (0,1)$ then*

$$\mathbf{W} \in \mathcal{M}_\varepsilon \ \Leftrightarrow \ \kappa(\mathbf{W}) = \frac{\lambda_{max}(\mathbf{W})}{\lambda_{min}(\mathbf{W})} < \frac{1+\varepsilon}{1-\varepsilon}.$$

(Here $\kappa(\mathbf{W})$ means the condition number of $\mathbf{W}$.)

Proof. The proof that $\mathcal{M}_\varepsilon$ is a convex cone is straightforward. We will prove the second part of the lemma. Denote

$$\lambda_{max}(\mathbf{W}) = \lambda_1(\mathbf{W}) \geq \lambda_2(\mathbf{W}) \geq \cdots \geq \lambda_n(\mathbf{W}) = \lambda_{min}(\mathbf{W})$$

the eigenvalues of $\mathbf{W}$. It holds that $\mathbf{W} \in \mathcal{M}_\varepsilon$ if and only if $\mathbf{W} \succ 0$ and there exists $\nu > 0$ such that

$$\|\mathbf{W} - \nu\mathbf{I}\|_2 = \max_i |\lambda_i(\mathbf{W}) - \nu| < \nu\varepsilon. \tag{3.36}$$

The inequality (3.36) is equivalent to

$$(1-\varepsilon)\nu < \lambda_{min}(\mathbf{W}) \leq \lambda_{max}(\mathbf{W}) < (1+\varepsilon)\nu. \tag{3.37}$$

From this follows that

$$\frac{\lambda_{max}(\mathbf{W})}{\lambda_{min}(\mathbf{W})} < \frac{1+\varepsilon}{1-\varepsilon}. \tag{3.38}$$

Now assume (3.38). Then

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

Therefore there exists $\nu > 0$ such that

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

or, equivalently, $\nu > 0$ such that (3.37) holds.

□

Now, for any $j = 1, \dots, 5$ we are ready to define the set $\mathcal{W}_j$ of suitable weights in the following way:

- $\mathcal{W}_1 = S^n_{++}$
- $\mathcal{W}_2 = \{ \ \mathbf{W} \in S^n_{++}; \exists \nu : \ \|\mathbf{W} - \nu \mathbf{I}\|_2 < \frac{\nu}{\sqrt{2}} \}$
- $\mathcal{W}_3 = \{ \ \mathbf{W} \in S^n_{++}; \exists \nu : \ \|\mathbf{W} - \nu \mathbf{I}\|_2 < \frac{\nu}{\sqrt{2}} \}$ *or* $\mathcal{W}_3 = D^n_{++}$
- $\mathcal{W}_4 = \{ \ \mathbf{W} \in S^n_{++}; \exists \nu : \ \|\mathbf{W} - \nu \mathbf{I}\|_F < \frac{\nu}{3\sqrt{2}} \}$
- $\mathcal{W}_5 = \{ \ \mathbf{W} \in S^n_{++}; \exists \nu : \ \|\mathbf{W} - \nu \mathbf{I}\|_F < \frac{\nu}{3\sqrt{2}} \}$

Note. Obviously, we can write $\mathcal{W}_2 = \mathcal{W}_3 = \mathcal{M}_{\frac{1}{\sqrt{2}}}$. However, Proposition A.2.7 (a) implies, that we can also set $\mathcal{W}_4 = \mathcal{W}_5 = \mathcal{M}_{\frac{1}{3\sqrt{2n}}}$.

In the previous section we have actually proved the following theorem:

Theorem 3.3.1 *Let* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. *Then for any* $j \in \{1, \dots, 5\}$

$$\Phi_j(\mathbf{X}, \mathbf{S}) \in \mathcal{W}_j \ \Rightarrow \ DF^j_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S}) \ \textit{is a nonsingular linear map.}$$

Proof. The statement of the theorem for $j =, 1, 2, 3$ follows from Proposition 3.2.1, Proposition 3.2.3 and Proposition 3.2.5. The statement for $j = 4, 5$ follows from Proposition 3.2.6 and Proposition 3.2.7.

□

Note. Let us remark that the nonsingularity of $DF^j_{\mu,\mathbf{W}}$, $(j = 2, 3, 4, 5)$ in the context of nonlinear semidefinite complementarity problems was shown by Monteiro and Zanjacomo [51] on the sets $\mathcal{W}_j$. The result for $j = 2, 3$ was improved by Tunçel and Wolkowicz [72] to the set $\mathcal{M}_{\sqrt{3}-1}$. Moreover, the result for $j = 3$ was extended by Chua and Tunçel [9] to the more general class of sets $\mathcal{M}_{\mathbf{D}}$, defined for any $\mathbf{D} = diag(d_1, \dots d_n) \in D^n_{++}$ as

$$\mathcal{M}_{\mathbf{D}} = \left\{ \ \mathbf{W} \succ 0; \exists \nu : \ \|\mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} - \nu \mathbf{I}\|_2 < \sqrt{\frac{d_{min}}{2 d_{max}}} \nu \right\}.$$

3.4 Properties of symmetrization maps

All the symmetrization maps in (3.2) have some properties useful for proving the existence of the weighted paths. These are stated in the following lemmas.

Lemma 3.4.1 *If* $\mathbf{X} \succeq 0, \mathbf{S} \succeq 0$, *then*
a) $\mathbf{X} \bullet \mathbf{S} = tr(\Phi_j(\mathbf{X}, \mathbf{S})),\ j = 1, 2, 3;$
b) $\mathbf{X} \bullet \mathbf{S} \leq 2\ tr(\Phi_j(\mathbf{X}, \mathbf{S})^2),\ j = 4, 5.$

Proof. The statement a) follows directly from the properties of the trace. We will prove the statement b). From the assumptions of the lemma it follows that the matrices $\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}$ and $\mathbf{U}_{\mathbf{S}}^T\mathbf{L}_{\mathbf{X}}$ have nonnegative eigenvalues. But for every square matrix $\mathbf{A}$ with nonnegative eigenvalues it holds that

$$tr\Big(\frac{\mathbf{A} + \mathbf{A}^T}{2}\Big)^2 = tr\Big(\frac{\mathbf{A}^2 + \mathbf{A}\mathbf{A}^T + \mathbf{A}^T\mathbf{A} + (\mathbf{A}^T)^2}{4}\Big) = tr\Big(\frac{\mathbf{A}^2 + \mathbf{A}\mathbf{A}^T}{2}\Big) \geq tr\Big(\frac{\mathbf{A}\mathbf{A}^T}{2}\Big).$$

The rest of the proof follows from the fact that

$$\mathbf{X} \bullet \mathbf{S} = tr(\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}}) = tr(\mathbf{U}_{\mathbf{S}}^T\mathbf{L}_{\mathbf{X}}\mathbf{L}_{\mathbf{X}}^T\mathbf{U}_{\mathbf{S}}^T).$$

□

Lemma 3.4.2 *Let* $j \in \{1, \ldots, 5\}$ *be arbitrary. If* $\mathbf{X} \succeq 0$, $\mathbf{S} \succeq 0$ *and* $\Phi_j(\mathbf{X}, \mathbf{S}) \succ 0$, *then* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$.

Proof. The statement for j=2,3 is obvious. If j=1, and $\mathbf{X} \succeq 0$ is singular, then $\mathbf{Q}^T\mathbf{X}\mathbf{Q} = \mathbf{D} = diag(d_1, \ldots, d_k, 0, \ldots, 0)$ for some orthogonal matrix $\mathbf{Q}$ and hence the matrix

$$\mathbf{Q}^T(\mathbf{X}\mathbf{S} + \mathbf{S}\mathbf{X})\mathbf{Q} = \mathbf{D}\mathbf{Q}^T\mathbf{S}\mathbf{Q} + \mathbf{Q}^T\mathbf{S}\mathbf{Q}\mathbf{D}$$

is singular. However this contradicts the assumption. If j=4, the proof is the same. The statement for j=5 follows from the fact that if $\mathbf{X}$ is singular, then there exists an index i such that $(\mathbf{L}_{\mathbf{X}})_{ii} = 0$. Then also $(\mathbf{U}_{\mathbf{S}}^T\mathbf{L}_{\mathbf{X}} + \mathbf{L}_{\mathbf{X}}^T\mathbf{U}_{\mathbf{S}})_{ii} = 0$, but this contradicts the assumption.

□

Lemma 3.4.3 *Let* $\nu > 0$ *and* $\mathbf{X} \succ 0$, $\mathbf{S} \succ 0$. *Then* $\mathbf{X}\mathbf{S} = \nu\mathbf{I}$ *if and only if*
a) $\Phi_j(\mathbf{X}, \mathbf{S}) = \nu\mathbf{I},\ j = 1, 2, 3$
b) $\Phi_j(\mathbf{X}, \mathbf{S}) = \sqrt{\nu}\mathbf{I},\ j = 4, 5.$

Proof. $\Rightarrow$ The statement for j=1,2,3 is obvious. Consider j=4. The matrices $\mathbf{X}, \mathbf{S}$ commute and therefore are simultaneously diagonalizable, that is, there exists an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{X} = \mathbf{Q}\mathbf{D}_\mathbf{X}\mathbf{Q}^T$ and $\mathbf{S} = \mathbf{Q}\mathbf{D}_\mathbf{S}\mathbf{Q}^T$. Therefore

$$(\mathbf{XS})^{\frac{1}{2}} = \mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}} = \sqrt{\nu}\mathbf{I} = \mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}} = (\mathbf{SX})^{\frac{1}{2}}$$

Let j=5. The matrix $\mathbf{L} = \mathbf{U}_\mathbf{S}^T\mathbf{L}_\mathbf{X}$ is lower triangular with positive diagonal entries. We have that $\mathbf{XS} = \nu\mathbf{I}$ if and only if $\mathbf{U}_\mathbf{S}^T\mathbf{X}\mathbf{U}_\mathbf{S} = \nu\mathbf{I}$. On the other hand $\mathbf{L}\mathbf{L}^T = \mathbf{U}_\mathbf{S}^T\mathbf{X}\mathbf{U}_\mathbf{S} = (\sqrt{\nu}\mathbf{I})(\sqrt{\nu}\mathbf{I})$ and therefore from the uniqueness of the Cholesky factor we obtain that $\mathbf{L} = \mathbf{L}^T = \sqrt{\nu}\mathbf{I}$.
$\Leftarrow$ The statement is obvious for j=2,3. For j=1 assume that $\mathbf{XS} + \mathbf{SX} = 2\nu\mathbf{I}$. For any symmetric matrix $\mathbf{A}$ and positive diagonal matrix $\mathbf{D}$ we have that

$$(\mathbf{AD} + \mathbf{DA})_{ij} = (\mathbf{D}_{ii} + \mathbf{D}_{jj})\mathbf{A}_{ij} = 0 \;\Leftrightarrow\; \mathbf{A}_{ij} = 0.$$

We have that $\mathbf{X} = \mathbf{QDQ}^T$ for some orthogonal matrix $\mathbf{Q}$ and positive diagonal matrix $\mathbf{D}$ and hence

$$\mathbf{XS} + \mathbf{SX} = \mathbf{QDQ}^T\mathbf{S} + \mathbf{SQDQ}^T = 2\nu\mathbf{I} \;\Leftrightarrow\; \mathbf{DQ}^T\mathbf{SQ} + \mathbf{Q}^T\mathbf{SQDQ}^T = 2\nu\mathbf{I}.$$

Therefore $\mathbf{Q}^T\mathbf{SQ}$ must be diagonal. We obtain that $\mathbf{X}$, $\mathbf{S}$ are simultaneously diagonalizable and so they commute. The proof for j=4 is similar. Finally assume that j=5 and $\mathbf{U}_\mathbf{S}^T\mathbf{L}_\mathbf{X} + \mathbf{L}_\mathbf{X}^T\mathbf{U}_\mathbf{S}^T = 2\nu\mathbf{I}$. Because the matrix $\mathbf{U}_\mathbf{S}^T\mathbf{L}_\mathbf{X}$ is lower triangular, $\mathbf{U}_\mathbf{S}^T\mathbf{L}_\mathbf{X} = \mathbf{L}_\mathbf{X}^T\mathbf{U}_\mathbf{S}^T = \sqrt{\nu}\mathbf{I}$. So we obtain $\mathbf{U}_\mathbf{S}^T\mathbf{L}_\mathbf{X}\mathbf{L}_\mathbf{X}^T\mathbf{U}_\mathbf{S}^T = \mathbf{U}_\mathbf{S}^T\mathbf{X}\mathbf{U}_\mathbf{S}^T = \nu\mathbf{I}$ that is equivalent to $\mathbf{XS} = \nu\mathbf{I}$.

□

3.5 Boundedness of weighted paths

For $\mu > 0$ and $\mathbf{W} \succ 0$ we will denote

$$(\mathbf{X}_{(\mu,\mathbf{W})}, y_{(\mu,\mathbf{W})}, \mathbf{S}_{(\mu,\mathbf{W})})$$

a solution of the system (3.1) (for some $j \in \{1, \ldots, 5\}$). Obviously, the solution needs not exist nor be unique. Nevertheless, we will prove in this section that the set of all solutions for some μ and $\mathbf{W}$ is bounded. To this aim, besides Assumption (A1) and Assumption (A3) will be needed the following:

Assumption (A4): For any $j \in \{1,\dots,5\}$ let $\triangle b, \triangle \mathbf{C}$ be such that there exists $\mathbf{W}^0 \in \mathcal{W}_j$ and $\mu_0 > 0$ such that the system (3.1) is solvable for $\mathbf{W} = \mathbf{W}^0$ and $\mu = \mu_0$.

In what follows, by $\mathbf{W}^0$ and μ_0 we will denote the weight $\mathbf{W}^0 \in \mathcal{W}_j$ and $\mu_0 > 0$ from Assumption (A4) for which the system (3.1) is solvable.

Let us remark that for $j \in \{1,\dots,5\}$ there always exist $\triangle b, \triangle \mathbf{C}$ such that they satisfy Assumption (A4). In fact, we can choose a weight $\mathbf{W}^0 \in \mathcal{W}_j$ and $\mu_0 > 0$ and pick up $(\mathbf{X}^0, y^0, \mathbf{S}^0) \in S^n_{++} \times R^m \times S^n_{++}$ such that

$$\Phi_j(\mathbf{X}^0, \mathbf{S}^0) = \phi_j(\mu^0)\mathbf{W}^0.$$

Then if we let

$$\triangle b = \frac{\mathcal{A}(\mathbf{X}^0) - b}{\phi_j(\mu^0)}, \qquad \triangle \mathbf{C} = \frac{\mathcal{A}^*(y^0) + \mathbf{S}^0 - \mathbf{C}}{\phi_j(\mu^0)},$$

then $(\mathbf{X}^0, y^0, \mathbf{S}^0)$ is a solution of the system (3.1) for $\mu = \mu_0$ and $\mathbf{W} = \mathbf{W}^0$. On the other hand, if Assumption (A2) holds, then $\triangle b = 0, \triangle \mathbf{C} = 0$ satisfy Assumption (A4) with $\mathbf{W}^0 = \mathbf{I}$ and any $\mu_0 > 0$, since the central path exists according to Corollary 2.2.1.

Lemma 3.5.1 *Let $\mathcal{O}(\mathbf{W}^0) \subset S^n_{++}$ be a bounded neighborhood of $\mathbf{W}^0$. Then the set*

$$\mathcal{M} = \{(\mathbf{X}_{(\mu,\mathbf{W})}, y_{(\mu,\mathbf{W})}, \mathbf{S}_{(\mu,\mathbf{W})}) \mid 0 < \mu \le \mu_0, \mathbf{W} \in \mathcal{O}(\mathbf{W}^0)\}$$

is bounded.

Proof. Let $(\mathbf{X}^0, y^0, \mathbf{S}^0)$ be the solution of (3.1) for $\mu = \mu_0$ and $\mathbf{W} = \mathbf{W}^0$. Let $0 < \mu \le \mu_0$ and $\mathbf{W} \in \mathcal{O}(\mathbf{W}^0)$ be arbitrary, such that there exist a solution $(\mathbf{X}_{(\mu,\mathbf{W})}, y_{(\mu,\mathbf{W})}, \mathbf{S}_{(\mu,\mathbf{W})})$ of the system (3.1). From Assumption (A3) we have that there exists $(\mathbf{X}^*, y^*, \mathbf{S}^*)$ such that:

$$\mathbf{A}^i \bullet \mathbf{X}^* = b_i, \quad \sum_{i=1}^m \mathbf{A}^i y_i^* + \mathbf{S}^* = \mathbf{C}, \quad \mathbf{X}^* \succeq 0, \quad \mathbf{S}^* \succeq 0, \quad \mathbf{X}^*\mathbf{S}^* = 0.$$

Define

$$\begin{pmatrix} \hat{\mathbf{X}} \\ \hat{y} \\ \hat{\mathbf{S}} \end{pmatrix} = \frac{\mu}{\mu_0} \begin{pmatrix} \mathbf{X}^0 \\ y^0 \\ \mathbf{S}^0 \end{pmatrix} + \left(1 - \frac{\mu}{\mu_0}\right) \begin{pmatrix} \mathbf{X}^* \\ y^* \\ \mathbf{S}^* \end{pmatrix}.$$

Clearly

$$\mathbf{A}^i \bullet \hat{\mathbf{X}} = \frac{\mu}{\mu_0}\mathbf{A}^i \bullet \mathbf{X}^0 + \left(1 - \frac{\mu}{\mu_0}\right)\mathbf{A}^i \bullet \mathbf{X}^* = b_i + \mu \triangle b_i, \quad \forall i = 1,\dots,m,$$

$$\sum_{i=1}^{m} \mathbf{A}^i \hat{y}_i + \hat{\mathbf{S}} = \frac{\mu}{\mu_0}\left(\sum_{i=1}^{m} \mathbf{A}^i y_i^0 + \mathbf{S}^0\right) + \left(1 - \frac{\mu}{\mu_0}\right)\left(\sum_{i=1}^{m} \mathbf{A}^i y_i^0 + \mathbf{S}^0\right) = \mathbf{C} + \mu\triangle\mathbf{C}$$

and hence

$$\mathbf{A}^i \bullet (\hat{\mathbf{X}} - \mathbf{X}_{(\mu,\mathbf{W})}) = 0, \quad \sum_{i=1}^{m} \mathbf{A}^i(\hat{y}_i - (y_{(\mu,\mathbf{W})})_i) + (\hat{\mathbf{S}} - \mathbf{S}_{(\mu,\mathbf{W})}) = 0.$$

Therefore

$$(\hat{\mathbf{X}} - \mathbf{X}_{\mu,\mathbf{W}}) \bullet (\hat{\mathbf{S}} - \mathbf{S}_{\mu,\mathbf{W}}) = 0.$$

This gives

$$\hat{\mathbf{X}} \bullet \mathbf{S}_{(\mu,\mathbf{W})} + \mathbf{X}_{(\mu,\mathbf{W})} \bullet \hat{\mathbf{S}} = \hat{\mathbf{X}} \bullet \hat{\mathbf{S}} + \mathbf{X}_{(\mu,\mathbf{W})} \bullet \mathbf{S}_{(\mu,\mathbf{W})}. \tag{3.39}$$

We first observe that

$$\hat{\mathbf{X}} \bullet \hat{\mathbf{S}} = \left(\frac{\mu}{\mu_0}\right)^2 \mathbf{X}^0 \bullet \mathbf{S}^0 + \left(1 - \frac{\mu}{\mu_0}\right)^2 \mathbf{X}^* \bullet \mathbf{S}^* + \frac{\mu}{\mu_0}\left(1 - \frac{\mu}{\mu_0}\right)(\mathbf{X}^0 \bullet \mathbf{S}^* + \mathbf{S}^0 \bullet \mathbf{X}^*) =$$

$$= \left(\frac{\mu}{\mu_0}\right)^2 \mathbf{X}^0 \bullet \mathbf{S}^0 + \frac{\mu}{\mu_0}\left(1 - \frac{\mu}{\mu_0}\right)(\mathbf{X}^0 \bullet \mathbf{S}^* + \mathbf{S}^0 \bullet \mathbf{X}^*)$$

$$\begin{cases} \leq \mu_0 \; tr(\mathbf{W}^0) + (\mathbf{X}^0 \bullet \mathbf{S}^* + \mathbf{S}^0 \bullet \mathbf{X}^*) & j = 1,2,3, \\ \leq 2\mu_0 tr((\mathbf{W}^0)^2) + (\mathbf{X}^0 \bullet \mathbf{S}^* + \mathbf{S}^0 \bullet \mathbf{X}^*) & j = 4,5, \end{cases} \tag{3.40}$$

where the inequalities follow from Lemma 3.4.1. According to the same lemma we have

$$\begin{array}{l} \mathbf{X}_{(\mu,\mathbf{W})} \bullet \mathbf{S}_{(\mu,\mathbf{W})} = \mu \; tr(\mathbf{W}) \leq \beta, \quad j = 1,2,3, \\ \mathbf{X}_{(\mu,\mathbf{W})} \bullet \mathbf{S}_{(\mu,\mathbf{W})} \leq 2tr(\mathbf{W}^2) \leq \beta, \quad j = 4,5 \end{array} \tag{3.41}$$

for some $\beta > 0$, since $0 < \mu < \mu_0$ a $\mathbf{W} \in \mathcal{O}(\mathbf{W}_0)$, which is bounded. Finally, from (3.39), (3.40), (3.41) we obtain that

$$\hat{\mathbf{X}} \bullet \mathbf{S}_{(\mu,\mathbf{W})} + \mathbf{X}_{(\mu,\mathbf{W})} \bullet \hat{\mathbf{S}} \leq \gamma$$

for some $\gamma > 0$ and hence the set

$$\mathcal{M}_1 = \left\{(\mathbf{X}_{(\mu,\mathbf{W})}, \mathbf{S}_{(\mu,\mathbf{W})}) \;\middle|\; \mu \in (0,\mu_0\rangle,\; \mathbf{W} \in \mathcal{O}(\mathbf{W}_0)\right\}$$

is included in the simplex

$$\left\{(\mathbf{X},\mathbf{S}) \mid \mathbf{X} \succeq 0,\; \mathbf{S} \succeq 0,\; \hat{\mathbf{X}} \bullet \mathbf{S} + \mathbf{X} \bullet \hat{\mathbf{S}} \leq \gamma\right\}$$

which is bounded, since $\hat{\mathbf{X}} \succ 0, \hat{\mathbf{S}} \succ 0$.
We will prove the boundedness of the set $\mathcal{M}$. (This part of the proof is adapted from [47].) Assume that there exists a sequence $\{(\mu_k, \mathbf{W}_k)\} \in (0, \mu_0\rangle \times \mathcal{O}(\mathbf{W}_0)$ and the associated solutions

$$(\mathbf{X}_k, y_k, \mathbf{S}_k) = (\mathbf{X}^j_{(\mu_k, \mathbf{W}_k)}, y^j_{(\mu_k, \mathbf{W}_k)} \mathbf{S}^j_{(\mu_k, \mathbf{W}_k)})$$

such that the sequence $\{(\mathbf{X}_k, \mathbf{S}_k)\}$ is bounded (this follows from the first part of the proof), however $\lim_{k\to\infty} \|y_k\| = \infty$. Obviously, the sequence $\frac{y_k}{\|y_k\|}$ is bounded and therefore it has a convergent subsequence. Assume (without loss of generality) that

$$\lim_{k\to\infty} \frac{y_k}{\|y_k\|} = \triangle y,$$

where obviously

$$\|\triangle y\| = 1. \tag{3.42}$$

We also have that

$$\lim_{k\to\infty} \frac{\tilde{\mathcal{A}}(y_k) + \mathbf{S}_k}{\|y_k\|} = \lim_{k\to\infty} \frac{\mathbf{C} + \mu_k \triangle \mathbf{C}}{\|y_k\|} = 0.$$

This implies

$$\lim_{k\to\infty} \frac{\tilde{\mathcal{A}}(y_k) + \mathbf{S}_k}{\|y_k\|} = \lim_{k\to\infty} \tilde{\mathcal{A}}(\frac{y_k}{\|y_k\|}) + \lim_{k\to\infty} \frac{\mathbf{S}_k}{\|y_k\|} = \mathcal{A}(\triangle y) = 0$$

and hence (from Assumption (A1)) $\triangle y = 0$, which contradicts to (3.42).

□

3.6 Existence of weighted path

In this section we prove (under Assumption (A1), Assumption (A3) and Assumption (A4)) that for any $\mu \in (0, \mu_0\rangle$ and $\mathbf{W} \in \mathcal{W}_j$ there exists a unique solution of (3.1) which is necessary for correct definition of weighted path. We first prove the existence.

Let $j \in \{1, \ldots, 5\}$. Consider the map

$$G_j : S^n \times R^m \times S^n \times R \times S^n \to R^m \times S^n \times S^n$$

such that

$$G_j(\mathbf{X}, y, \mathbf{S}, \mu, \mathbf{W}) = F^j_{\mu,\mathbf{W}}(\mathbf{X}, y, \mathbf{S})$$

Obviously $G_j(\mathbf{X}^0, y^0, \mathbf{S}^0, \mu_0, \mathbf{W}^0) = 0$. The following technique is called the analytic continuation and was used by Preiss and Stoer to prove the existence of the weighted path in linear complementarity problem associated with the symmetrization $(\mathbf{XS} + \mathbf{SX})/2$ (see Lemma 3.5, Lemma 3.6, Lemma 3.7 in [58]).

Lemma 3.6.1 *Let $j \in \{1, 2, \ldots, 5\}$. Assume $\mu_1 \in (0, \mu_0\rangle, \mathbf{W}^1 \in \mathcal{W}_j$ and let*

$$\psi : \langle 0, 1\rangle \to (0, \mu_0\rangle \times \mathcal{W}_j, \qquad \psi(t) = (\mu_t, \mathbf{W}^t)$$

be a continuous path from $\psi(0) = (\mu_0, \mathbf{W}^0)$ to $\psi(1) = (\mu_1, \mathbf{W}^1)$. Then for all $t \in \langle 0, 1\rangle$ the system

$$G_j(\mathbf{X}, y, \mathbf{S}, \mu_t, \mathbf{W}^t) = 0$$

has a locally unique solution $(\mathbf{X}^t, y^t, \mathbf{S}^t)$, where $\mathbf{X}^t \succ 0, \mathbf{S}^t \succ 0$. Moreover, there exists a function

$$g_j : R_{++} \times S^n_{++} \to S^n_{++} \times R^m \times S^n_{++},$$

which is defined and analytic on some neighborhood of $\psi(t)$, satisfies $g_j(\psi(t)) = (\mathbf{X}^t, y^t, \mathbf{S}^t)$ and

$$G_j(g_j(\psi(t)), \psi(t)) = 0.$$

Proof. For $t \in \langle 0, 1\rangle$ consider the system

$$G_j(\mathbf{X}, y, \mathbf{S}, \psi(t)) = 0, \quad \mathbf{X} \succ 0,\ \mathbf{S} \succ 0. \tag{3.43}$$

The point $(\mathbf{X}^0, y^0, \mathbf{S}^0)$ is the solution of this system for $t = 0$. From Theorem 3.3.1 it follows that the partial Fréchet derivative $DG_j(\mathbf{X}, y, \mathbf{S}, \phi(t))$ concerning the variables $(\mathbf{X}, y, \mathbf{S})$ is nonsingular in $(\mathbf{X}^0, y^0, \mathbf{S}^0, \phi(0))$. From the implicit function theorem we obtain that there exists an analytic function g_j defined on some neighborhood of $\psi(0) = (\mu_0, \mathbf{W}^0)$ such that

$$g_j(\psi(0)) = g_j(\mu_0, \mathbf{W}^0) = (\mathbf{X}^0, y^0, \mathbf{S}^0)$$

and

$$G_j(g_j(\psi(t)), \psi(t)) = G_j(g_j(\mu_t, \mathbf{W}^t), \mu_t, \mathbf{W}^t) = 0$$

on some neighborhood of $t = 0$. Actually, there is a maximal $\bar{t} \in (0, 1\rangle$ such that

$$g_j(\psi(t)) = g_j(\mu_t, \mathbf{W}^t) = (\mathbf{X}^t, y^t, \mathbf{S}^t), \quad \forall t \in \langle 0, \bar{t}).$$

That means $(\mathbf{X}^t, y^t, \mathbf{S}^t)$ is a locally unique solution of

$$F^j_{\mu_t, \mathbf{W}^t}(\mathbf{X}, y, \mathbf{S}) = 0, \qquad \forall\ t \in \langle 0, \bar{t}).$$

Moreover, from Lemma 3.4.2 we have that $\mathbf{X}^t \succ 0, \mathbf{S}^t \succ 0$. From the continuity of ψ it follows that $\psi(\langle 0, 1\rangle)$ is a compact subset of $(0, \mu_0\rangle \times \mathcal{W}_j$, therefore, according to Lemma 3.5.1 the set

$$\{g(\psi(t)) = (\mathbf{X}^t, y^t, \mathbf{S}^t),\ |\ t \in \langle 0, \bar{t})\}$$

is bounded. Let $t_k \in \langle 0, \bar{t})$ for $k = 1, 2, \ldots$ and $\lim_{k\to\infty} t_k = \bar{t}$. Then there exists a sequence $\{t_{k_j}\}_{j=1}^{\infty}$ chosen from $\{t_k\}_{k=1}^{\infty}$ such that

$$\lim_{j\to\infty} g_j(\psi(t_{k_j})) = (\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}}).$$

Because

$$\begin{aligned} \mathbf{A}^i \bullet \mathbf{X}^{t_{k_j}} &= b_i + \mu_{t_{k_j}} \triangle b_i, \quad i = 1, \ldots, m, \quad \mathbf{X}^{t_{k_j}} \succ 0, \\ \textstyle\sum_{i=1}^m \mathbf{A}^i y_i^{t_{k_j}} + \mathbf{S}^{t_{k_j}} &= \mathbf{C} + \mu_{t_{k_j}} \triangle \mathbf{C}, \quad \mathbf{S}^{t_{k_j}} \succ 0, \\ \Psi_j(\mathbf{X}^{t_{k_j}}, \mathbf{S}^{t_{k_j}}) &= \phi_j(\mu_{t_{k_j}}) \mathbf{W}^{t_{k_j}}, \end{aligned}$$

by taking limit $j \to \infty$ we obtain

$$\begin{aligned} \mathbf{A}^i \bullet \bar{\mathbf{X}} &= b_i + \mu_{\bar{t}} \triangle b_i, \quad i = 1, \ldots, m, \quad \bar{\mathbf{X}} \succeq 0, \\ \textstyle\sum_{i=1}^m \mathbf{A}^i \bar{y}_i + \bar{\mathbf{S}} &= \mathbf{C} + \mu_{\bar{t}} \triangle \mathbf{C}, \quad \bar{\mathbf{S}} \succeq 0 \\ \Psi_j(\bar{\mathbf{X}}, \bar{\mathbf{S}}) &= \phi_j(\mu_{\bar{t}}) \mathbf{W}^{\bar{t}}. \end{aligned}$$

Applying Lemma 3.4.2 again we have that $\bar{\mathbf{X}}, \bar{\mathbf{S}}$ are positive definite. Therefore $(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}})$ is the solution of the system (3.43) for $t = \bar{t}$. The partial Fréchet derivative

$$DG_j(\mathbf{X}, y, \mathbf{S}, \psi(t))$$

concerning the variables $(\mathbf{X}, y, \mathbf{S})$ is nonsingular in $(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}})$ and $(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}})$ is locally unique solution of the system

$$F^j_{\mu_{\bar{t}}, \mathbf{W}^{\bar{t}}}(\mathbf{X}, y, \mathbf{S}) = 0.$$

By applying the implicit function theorem again and from the maximality of $\bar{t}$ we obtain that $\bar{t} = 1$.

□

Corollary 3.6.1 *For any $\mu \in (0, \mu_0\rangle$ and $\mathbf{W} \in \mathcal{W}_j$ there exists a solution of* (3.1).

Proof. It suffices to prove that having $\mathbf{W} \in \mathcal{W}_j$ and $\mu \in (0, \mu_0\rangle$ one can find a continuous path from $(\mu_0, \mathbf{W}^0)$ to $(\mu, \mathbf{W})$. However, we can define

$$\psi(t) = (t\mu + (1-t)\mu_0, t\mathbf{W} + (1-t)\mathbf{W}^0).$$

Obviously $t\mu + (1-t)\mu_0 \in (0, \mu_0\rangle$ for all $t \in \langle 0, 1\rangle$ and since $\mathcal{W}_j$ is convex, $t\mathbf{W} + (1 - t)\mathbf{W}^0 \in \mathcal{W}_j$.

□

Having the existence result stated in Corollary 3.5.1 we turn our attention to the uniqueness of the solutions. As a consequence of Lemma 3.6.1 we obtain the following result that will be useful later.

Corollary 3.6.2 *For all $t \in \langle 0, 1\rangle$ the function $g(\psi(t))$ from Lemma 3.6.1 is uniquely determined by the path ψ and the starting value $g(\psi(0))$.*

First, we prove the uniqueness of (3.1) for a special choice of the weight matrix $\mathbf{W} = \mathbf{I}$. This result will be used then in the proof of Lemma 3.6.3.

Lemma 3.6.2 *Let $j \in \{1, 2, \ldots, 5\}$ be arbitrary. If the system*

$$\left.\begin{array}{rl} \mathcal{A}(\mathbf{X}) = b + \mu\triangle b, & \mathbf{X} \succ 0, \\ \mathcal{A}^*(y) + \mathbf{S} = \mathbf{C} + \mu\triangle C, & \mathbf{S} \succ 0, \\ \Phi_j(\mathbf{X}, \mathbf{S}) = \phi_j(\mu)\mathbf{I} & \end{array}\right\} \tag{3.44}$$

has a solution for some $\mu > 0$, then this solution is unique.

Proof. Suppose there are two solutions $(\mathbf{X}_1, y_1, \mathbf{S}_1)$, $(\mathbf{X}_2, y_2, \mathbf{S}_2)$ of the system (3.44). Let $(\triangle\mathbf{X}, \triangle y, \triangle\mathbf{S}) = (\mathbf{X}_1, y_1, \mathbf{S}_1) - (\mathbf{X}_2, y_2, \mathbf{S}_2)$. Then $\mathcal{A}(\triangle\mathbf{X}) = 0, \tilde{\mathcal{A}}(\triangle y) + \triangle\mathbf{S} = 0$ and hence $\triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0$. Lemma 3.4.3 states that

$$\Phi_j(\mathbf{X}_i, \mathbf{S}_i) = \phi_j(\mu)\mathbf{I} \;\Leftrightarrow\; \mathbf{X}_i\mathbf{S}_i = \mu\mathbf{I}, \qquad i = 1, 2.$$

Therefore

$$\mu\mathbf{I} = \mathbf{X}_1\mathbf{S}_1 = (\mathbf{X}_2 + \triangle\mathbf{X})(\mathbf{S}_2 + \triangle\mathbf{S}) = \mathbf{X}_2\mathbf{S}_2 + \mathbf{X}_2\triangle\mathbf{S} + \triangle\mathbf{X}\mathbf{S}_2 + \triangle\mathbf{X}\triangle\mathbf{S},$$

$$\mu\mathbf{I} = \mathbf{X}_2\mathbf{S}_2 = (\mathbf{X}_1 - \triangle\mathbf{X})(\mathbf{S}_1 - \triangle\mathbf{S}) = \mathbf{X}_1\mathbf{S}_1 - \mathbf{X}_1\triangle\mathbf{S} - \triangle\mathbf{X}\mathbf{S}_1 + \triangle\mathbf{X}\triangle\mathbf{S}$$

and by subtracting the equations above we obtain that

$$(\mathbf{X}_1 + \mathbf{X}_2)\triangle\mathbf{S} + \triangle\mathbf{X}(\mathbf{S}_1 + \mathbf{S}_2) = 0.$$

We can express $\triangle\mathbf{S}$ as

$$\triangle\mathbf{S} = -(\mathbf{X}_1 + \mathbf{X}_2)^{-1}\triangle\mathbf{X}(\mathbf{S}_1 + \mathbf{S}_2) \tag{3.45}$$

and hence

$$0 = \triangle\mathbf{X} \bullet \triangle\mathbf{S} = tr(\triangle\mathbf{X}\triangle\mathbf{S}) = -tr(\triangle\mathbf{X}(\mathbf{X}_1 + \mathbf{X}_2)^{-1}\triangle\mathbf{X}(\mathbf{S}_1 + \mathbf{S}_2)) =$$

$$= -tr((\mathbf{S}_1 + \mathbf{S}_2)^{\frac{1}{2}} \triangle\mathbf{X}(\mathbf{X}_1 + \mathbf{X}_2)^{-1} \triangle\mathbf{X}(\mathbf{S}_1 + \mathbf{S}_2)^{\frac{1}{2}}).$$

The trace of the positive semidefinite matrix is zero if and only if it is the zero matrix. That's why $\triangle\mathbf{X} = 0$ and from (3.45) also $\triangle\mathbf{S} = 0$. Finally, Assumption (A1) gives $\triangle y = 0$.

□

We now prove the uniqueness for the general positive definite weight matrix $\mathbf{W}$.

Lemma 3.6.3 *If the system*

$$\left.\begin{array}{rl} \mathcal{A}(\mathbf{X}) = b + \mu\triangle b, & \mathbf{X} \succ 0, \\ \mathcal{A}^*(y) + \mathbf{S} = \mathbf{C} + \mu\triangle\mathbf{C}, & \mathbf{S} \succ 0, \\ \Phi_j(\mathbf{X}, \mathbf{S}) = \phi_j(\mu)\mathbf{W} & \end{array}\right\} \tag{3.46}$$

has a solution for some $\mu > 0$, then this solution is unique.

Proof. Let $\mu > 0$ and suppose there are two solutions $(\mathbf{X}_1, y_1, \mathbf{S}_1), (\mathbf{X}_2, y_2, \mathbf{S}_2)$. Consider the line connecting $(\mu, \mathbf{W})$ with $(\mu, \mathbf{I})$, i.e.

$$\psi : \langle 0, 1\rangle \to R_{++} \times \mathcal{W}_j, \quad \psi(t) = (\mu, t\mathbf{I} + (1-t)\mathbf{W}).$$

Lemma 3.6.1 states that there exist analytic continuations from $(\mathbf{X}_1, y_1, \mathbf{S}_1)$ and $(\mathbf{X}_2, y_2, \mathbf{S}_2)$ along ψ to the solution of the system (3.44), which is unique (Lemma 3.6.2). Denote this solution $(\mathbf{X}_I, y_I, \mathbf{S}_I)$. The analytic continuation from $(\mathbf{X}_I, y_I, \mathbf{S}_I)$ along the inverse path $\phi^{-1}(t) = \phi(1-t)$ leads to both $(\mathbf{X}_1, y_1, \mathbf{S}_1)$ and $(\mathbf{X}_2, y_2, \mathbf{S}_2)$. The uniqueness of the analytic continuation (Corollary 3.6.2) implies $(\mathbf{X}_1, y_1, \mathbf{S}_1) = (\mathbf{X}_2, y_2, \mathbf{S}_2)$.

□

Now, we can formulate the main result of this chapter, which is a simple consequence of Corollary 3.6.1 and Lemma 3.6.3. Let us recall that it was proved under Assumptions (A1), (A3), (A4).

Theorem 3.6.1 *Let $j \in \{1, 2, \ldots, 5\}$. Then for any $\mu \in (0, \mu_0\rangle$ and $\mathbf{W} \in \mathcal{W}_j$ there exists unique solution of the system* (3.1)*.*

As it was mentioned above, under Assumption (A2) the choice $\triangle b = 0, \triangle\mathbf{C} = 0$ satisfies Assumption (A4). Hence we obtain the following corollary of Theorem 3.6.1.

Corollary 3.6.3 *Consider Assumptions (A1) and (A2) and let* $j \in \{1, 2, \ldots, 5\}$*. Then for any* $\mu > 0$ *and* $\mathbf{W} \in \mathcal{W}_j$ *there exists unique solution of the system*

$$\left.\begin{array}{rl} \mathcal{A}(\mathbf{X}) = b, & \mathbf{X} \succ 0, \\ \mathcal{A}^*(y) + \mathbf{S} = \mathbf{C}, & \mathbf{S} \succ 0, \\ \Phi_j(\mathbf{X}, \mathbf{S}) = \phi_j(\mu)\mathbf{W}. & \end{array}\right\} \tag{3.47}$$

Definition 3.6.1 *(a) Assume (A1), (A3), (A4). Let* $j \in \{1, 2, \ldots, 5\}$ *and* $\mathbf{W} \in \mathcal{W}_j$*. Then the infeasible weighted central path (with the weight* $\mathbf{W}$*) is defined as the set*

$$\{\ (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))\ |\mu \in (0, \mu_0\rangle\}$$

of the solutions of the system (3.1)*, or alternatively as the map*

$$f_{IWP} : (0, \mu_0\rangle \to S^n \times R^m \times S^n, \qquad \mu \mapsto (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu)).$$

(b) Assume (A1) and (A2). Let $j \in \{1, 2, \ldots, 5\}$ *and* $\mathbf{W} \in \mathcal{W}_j$*. Then the feasible weighted central path (with the weight* $\mathbf{W}$*) is defined as the set*

$$\{\ (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))\ |\mu > 0\}$$

of the solutions of the system (3.47)*, or alternatively as the map*

$$f_{FWP} : R_{++} \to S^n \times R^m \times S^n, \qquad \mu \mapsto (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu)).$$

The following proposition follows directly from Definition 3.6.1 and the analyticity of the systems (3.1) and (3.47) for $\mu > 0$.

Proposition 3.6.1 *The (infeasible/feasible) weighted central path is an analytic function for* $\mu > 0$*.*

Chapter 4

Limiting behavior of weighted paths

In the previous chapter the existence of the (infeasible) weighted paths was shown, that were associated with various symmetrization maps (see (3.2)) and defined as the sets

$$\{\ (\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))\ |\mu \in (0, \mu_0\rangle\} \tag{4.1}$$

of the solutions of the system (3.1) for some fixed weight $\mathbf{W} \in \mathcal{W}_j$, where μ_0 is given by Assumption (A4) (see Definition 3.6.1).
Recall, that the existence was shown under Assumptions (A1), (A3), (A4). In what follows the assumption (A3) will be replaced with a stronger assumption:

Assumption (A5): There exists a strictly complementary optimal solution of the system (2.3).

This assumption is restrictive, though it is necessary for an analysis of the limiting behavior of the paths (see e.g. [25], [41], [42], [59], [7], [8]). Therefore, from now we will suppose that Assumption (A1), Assumption (A4) and Assumption (A5) hold.

Let $(\mathbf{X}^*, y^*, \mathbf{S}^*)$ be a strictly complementary optimal solution (see Definition 2.1.1). Since $\mathbf{X}^*\mathbf{S}^* = 0$, the matrices $\mathbf{X}^*, \mathbf{S}^*$ commute and therefore there exists an orthogonal matrix $\mathbf{Q}$ such that the matrices $\mathbf{Q}\mathbf{X}^*\mathbf{Q}^T$, $\mathbf{Q}\mathbf{S}^*\mathbf{Q}^T$ are diagonal (see Theorem A.1.2). Therefore, without loss of generality (applying an orthogonal transformation on the data, if necessary), we may assume, that

$$\mathbf{X}^* = \begin{pmatrix} \Lambda_B^* & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{S}^* = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_N^* \end{pmatrix},$$

where $\Lambda_B^* = diag(\lambda_1^*, \ldots, \lambda_{|B|}^*) \succ 0$, $\Lambda_N^* = diag(\lambda_{|B|+1}^*, \ldots, \lambda_n^*) \succ 0$.
Let $(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})$ be an another (not necessary strictly complementary) optimal solution of the system (2.3). Assume that $(\hat{\mathbf{X}}, \hat{\mathbf{S}})$ is partitioned in the following way:

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

From the complementarity property it follows, that

$$0 = \mathbf{X}^*\hat{\mathbf{S}} = \begin{pmatrix} \Lambda_B^* & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{S}}_B & \hat{\mathbf{S}}_V \\ \hat{\mathbf{S}}_V^T & \hat{\mathbf{S}}_N \end{pmatrix} = \begin{pmatrix} \Lambda_B^*\hat{\mathbf{S}}_B & \Lambda_B^*\hat{\mathbf{S}}_V \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since $\Lambda_B^* \succ 0$, we have that $\hat{\mathbf{S}}_B = 0$, $\hat{\mathbf{S}}_V = 0$. In the similar way it can be shown, that $\hat{\mathbf{X}}_N = 0$, $\hat{\mathbf{X}}_V = 0$. Therefore any optimal solution pair $(\hat{\mathbf{X}}, \hat{\mathbf{S}})$ is in the form

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

where $\hat{\mathbf{X}}_B \succeq 0, \hat{\mathbf{S}}_N \succeq 0$.

In what follows, we will assume, that any square symmetric matrix $\mathbf{M} \in S^n$ has the partition

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_B & \mathbf{M}_V \\ \mathbf{M}_V^T & \mathbf{M}_N \end{pmatrix}. \tag{4.2}$$

Lemma 4.0.4 *Let* $\mathbf{M}, \mathbf{N} \in S^n$. *Then*

$$\mathbf{M} \bullet \mathbf{N} = \mathbf{M}_B \bullet \mathbf{N}_B + 2\mathbf{M}_V \bullet \mathbf{N}_V + \mathbf{M}_N \bullet \mathbf{N}_N.$$

4.1 Asymptotic behavior of weighted paths

In this section we will study the asymptotic behavior of the blocks $\mathbf{X}_B(\mu)$, $\mathbf{S}_B(\mu)$, $\mathbf{X}_V(\mu)$, $\mathbf{S}_V(\mu)$, $\mathbf{X}_N(\mu)$, $\mathbf{S}_N(\mu)$ of the matrices $\mathbf{X}(\mu), \mathbf{S}(\mu)$ for $\mu \to 0$. These results are obtained by extending the technique of Preiss and Stoer [59][1] to all types of weighted paths. For the definition of the $\mathcal{O}$, Θ and o notation see Appendix C.

[1]These authors studied the asymptotic behavior of the paths associated with the symmetrization $(\mathbf{XS} + \mathbf{SX})/2$. Their results are included here in order to obtain a complete view on the asymptotic behavior of all types of weighted paths.

4.1.1 Asymptotic properties in $\mathcal{O}$-notation

Proposition 4.1.1 *Let $j \in \{1, \ldots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\mathbf{X}(\mu) = \mathcal{O}(1), \quad y(\mu) = \mathcal{O}(1), \quad \mathbf{S}(\mu) = \mathcal{O}(1).$$

Proof. The statement follows from Lemma 3.5.1 (boundedness of the weighted path).

$\square$

Proposition 4.1.2 *Let $j \in \{1, \ldots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\begin{array}{ll} \mathbf{X}_B(\mu) = \mathcal{O}(1), & \mathbf{S}_B(\mu) = \mathcal{O}(\mu), \\ \mathbf{X}_V(\mu) = \mathcal{O}(\sqrt{\mu}), & \mathbf{S}_V(\mu) = \mathcal{O}(\sqrt{\mu}), \\ \mathbf{X}_N(\mu) = \mathcal{O}(\mu), & \mathbf{S}_N(\mu) = \mathcal{O}(1). \end{array}$$

Proof. Let $(\mathbf{X}^*, \mathbf{S}^*)$ be the strictly complementary optimal solution given by assumption (A5). Define

$$\bar{\mathbf{X}} = \frac{\mu}{\mu_0}\mathbf{X}(\mu_0) + \Big(1 - \frac{\mu}{\mu_0}\Big)\mathbf{X}^*, \tag{4.3}$$

$$\bar{\mathbf{S}} = \frac{\mu}{\mu_0}\mathbf{S}(\mu_0) + \Big(1 - \frac{\mu}{\mu_0}\Big)\mathbf{S}^*. \tag{4.4}$$

It can be easily seen, that

$$(\bar{\mathbf{X}} - \mathbf{X}(\mu)) \bullet (\bar{\mathbf{S}} - \mathbf{S}(\mu)) = 0$$

and hence

$$\bar{\mathbf{X}} \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \bar{\mathbf{S}} = \bar{\mathbf{X}} \bullet \bar{\mathbf{S}} + \mathbf{X}(\mu) \bullet \mathbf{S}(\mu). \tag{4.5}$$

By inserting (4.3), (4.4) into (4.5) we obtain

$$\Big(1 - \frac{\mu}{\mu_0}\Big)[\mathbf{X}^* \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}^*] + \frac{\mu}{\mu_0}[\mathbf{X}(\mu_0) \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}(\mu_0)] =$$

$$= \Big(\frac{\mu}{\mu_0}\Big)^2 \mathbf{X}(\mu_0) \bullet \mathbf{S}(\mu_0) + \frac{\mu}{\mu_0}\Big(1 - \frac{\mu}{\mu_0}\Big)[\mathbf{X}^* \bullet \mathbf{S}(\mu_0) + \mathbf{X}(\mu_0) \bullet \mathbf{S}^*] + \mathbf{X}(\mu) \bullet \mathbf{S}(\mu) \tag{4.6}$$

By multiplying the equation (4.6) by $\frac{\mu_0}{\mu}$, we obtain

$$\Big(\frac{\mu_0 - \mu}{\mu}\Big)[\mathbf{X}^* \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}^*] + [\mathbf{X}(\mu_0) \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}(\mu_0)] =$$

$$= \frac{\mu}{\mu_0}\mathbf{X}(\mu_0) \bullet \mathbf{S}(\mu_0) + \frac{\mu_0 - \mu}{\mu_0}[\mathbf{X}^* \bullet \mathbf{S}(\mu_0) + \mathbf{X}(\mu_0) \bullet \mathbf{S}^*] + \frac{\mu_0}{\mu}\mathbf{X}(\mu) \bullet \mathbf{S}(\mu). \tag{4.7}$$

Obviously $\frac{\mu}{\mu_0} \leq 1$, $\frac{\mu_0 - \mu}{\mu_0} \leq 1$ and from the equality

$$\Phi_j(\mathbf{X}(\mu), \mathbf{S}(\mu)) = \phi_j(\mu)\mathbf{W}$$

and Lemma 3.4.1 it follows, that

$$\frac{\mu_0}{\mu}\mathbf{X}(\mu) \bullet \mathbf{S}(\mu) = \frac{\mu_0}{\mu}tr(\mu\mathbf{W}) = \mu_0 tr\mathbf{W},$$

for $j = 1, 2, 3$, and

$$\frac{\mu_0}{\mu}\mathbf{X}(\mu) \bullet \mathbf{S}(\mu) \leq \frac{\mu_0}{\mu}2tr(\mu\mathbf{W}^2) = 2\mu_0 tr\mathbf{W^2}$$

for $j = 4, 5$. Therefore there exists a constant $C > 0$ such that

$$\Big(\frac{\mu_0 - \mu}{\mu}\Big)[\mathbf{X}^* \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}^*] + [\mathbf{X}(\mu_0) \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}(\mu_0)] \leq C.$$

The both addends on the left hand side are nonnegative and therefore

$$\Big(\frac{\mu_0 - \mu}{\mu}\Big)[\mathbf{X}^* \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}^*] \leq C.$$

Let $\varepsilon \in (0, \mu_0)$. Then for $\mu \in (0, \mu_0 - \varepsilon\rangle$ it holds

$$[\mathbf{X}^* \bullet \mathbf{S}(\mu) + \mathbf{X}(\mu) \bullet \mathbf{S}^*] \leq \frac{C\mu}{\mu_0 - \mu} \leq \frac{C\mu}{\varepsilon}$$

and hence, for sufficiently small μ

$$\sum_{i=1}^{|B|} \lambda_i^* \mathbf{S}_{ii}(\mu) + \sum_{i=|B|+1}^{n} \lambda_i^* \mathbf{X}_{ii}(\mu) = \mathcal{O}(\mu).$$

Because $\lambda_i^* > 0$ $\forall i = 1, \dots n$, we have

$$\mathbf{S}_{ii}(\mu) = \mathcal{O}(\mu),\ \forall i = 1, \dots |B|, \quad \mathbf{X}_{ii}(\mu) = \mathcal{O}(\mu),\ \forall i = |B| + 1, \dots n,$$

and therefore

$$tr(\mathbf{S}_B(\mu)) = \mathcal{O}(\mu) \quad tr(\mathbf{X}_N(\mu)) = \mathcal{O}(\mu).$$

Proposition C.2.2 implies $\mathbf{S}_B(\mu) = \mathcal{O}(\mu)$ and $\mathbf{X}_N(\mu) = \mathcal{O}(\mu)$. From Proposition A.3.3 it follows, that there exists $C_1 > 0$ such that

$$\|\mathbf{X}_V(\mu)\|_F^2 \leq tr(\mathbf{X}_B(\mu))tr(\mathbf{X}_N(\mu)) \leq C_1\mu.$$

Therefore $\mathbf{X}_V(\mu) = \mathcal{O}(\sqrt{\mu})$. It can be similarly shown that $\mathbf{S}_V(\mu) = \mathcal{O}(\sqrt{\mu})$.

□

Denote

$$\mathbf{Y}(\mu) := \mathbf{X}^{\frac{1}{2}}(\mu), \qquad \mathbf{Z}(\mu) := \mathbf{S}^{\frac{1}{2}}(\mu) \tag{4.8}$$

the square root of the matrix $\mathbf{X}(\mu)$ and $\mathbf{S}(\mu)$, which exist and are uniquely defined (see Appendix A.1). Obviously

$$\begin{array}{ll} \mathbf{X}_B(\mu) = \mathbf{Y}_B^2(\mu) + \mathbf{Y}_V(\mu)\mathbf{Y}_V^T(\mu), & \mathbf{S}_B(\mu) = \mathbf{Z}_B^2(\mu) + \mathbf{Z}_V(\mu)\mathbf{Z}_V^T(\mu), \\ \mathbf{X}_V(\mu) = \mathbf{Y}_B(\mu)\mathbf{Y}_V(\mu) + \mathbf{Y}_V(\mu)\mathbf{Y}_N(\mu), & \mathbf{S}_V(\mu) = \mathbf{Z}_B(\mu)\mathbf{Z}_V(\mu) + \mathbf{Z}_V(\mu)\mathbf{Z}_N(\mu), \\ \mathbf{X}_N(\mu) = \mathbf{Y}_N^2(\mu) + \mathbf{Y}_V^T(\mu)\mathbf{Y}_V(\mu), & \mathbf{S}_N(\mu) = \mathbf{Z}_N^2(\mu) + \mathbf{Y}_V^T(\mu)\mathbf{Y}_V(\mu). \end{array} \tag{4.9}$$

Proposition 4.1.3 *Let $j \in \{1, \dots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\begin{array}{ll} \mathbf{Y}_B(\mu) = \mathcal{O}(1), & \mathbf{Z}_B(\mu) = \mathcal{O}(\sqrt{\mu}), \\ \mathbf{Y}_V(\mu) = \mathcal{O}(\sqrt{\mu}), & \mathbf{Z}_V(\mu) = \mathcal{O}(\sqrt{\mu}), \\ \mathbf{Y}_N(\mu) = \mathcal{O}(\sqrt{\mu}), & \mathbf{Z}_N(\mu) = \mathcal{O}(1). \end{array}$$

Proof. Since by Proposition 4.1.2 $\mathbf{X}_B(\mu) = \mathcal{O}(1), \mathbf{S}_N(\mu) = \mathcal{O}(1)$, we have that $\mathbf{Y}_B(\mu) = \mathcal{O}(1), \mathbf{Z}_N(\mu) = \mathcal{O}(1)$. Moreover,

$$\max\{\|\mathbf{Y}_N(\mu)\|_F^2, \|\mathbf{Y}_V(\mu)\|_F^2\} \le \|\mathbf{Y}_N(\mu)\|_F^2 + \|\mathbf{Y}_V(\mu)\|_F^2 =$$

$$= tr(\mathbf{Y}_N^2(\mu)) + tr(\mathbf{Y}_V^T(\mu)\mathbf{Y}_V(\mu)) = tr(\mathbf{X}_N(\mu)) = \mathcal{O}(\mu),$$

which yield $\mathbf{Y}_N(\mu) = \mathcal{O}(\sqrt{\mu})$ and $\mathbf{Y}_V(\mu) = \mathcal{O}(\sqrt{\mu})$. Analogously,

$$\max\{\|\mathbf{Z}_B(\mu)\|_F^2, \|\mathbf{Z}_V(\mu)\|_F^2\} \le \|\mathbf{Z}_B(\mu)\|_F^2 + \|\mathbf{Z}_V(\mu)\|_F^2 =$$

$$= tr(\mathbf{Z}_B^2(\mu)) + tr(\mathbf{Z}_V(\mu)\mathbf{Z}_V^T(\mu)) = tr(\mathbf{S}_B(\mu)) = \mathcal{O}(\mu)$$

which yield $\mathbf{Z}_B(\mu) = \mathcal{O}(\sqrt{\mu})$ and $\mathbf{Z}_V(\mu) = \mathcal{O}(\sqrt{\mu})$.

□

Denote $\mathbf{L}(\mu) := \mathbf{L}_{\mathbf{X}(\mu)} \in L_{++}^n$ the lower Cholesky factor of the matrix $\mathbf{X}(\mu)$ and $\mathbf{U}(\mu) := \mathbf{U}_{\mathbf{S}(\mu)} \in U_{++}^n$ the upper Cholesky factor of the matrix $\mathbf{S}(\mu)$ (which exist and are uniquely determined—see Theorem A.1.3). It holds

$$\mathbf{X}(\mu) = \mathbf{L}(\mu)\mathbf{L}^T(\mu), \qquad \mathbf{S}(\mu) = \mathbf{U}(\mu)\mathbf{U}^T(\mu),$$

where we denote $\mathbf{L}^T(\mu) := (\mathbf{L}(\mu))^T$ and $\mathbf{U}^T(\mu) := (\mathbf{U}(\mu))^T$. Assume, that any lower triangular matrix $\mathbf{L}$ and upper triangular matrix $\mathbf{U}$ is partitioned in the following way:

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_B & 0 \\ \mathbf{L}_V^T & \mathbf{L}_N \end{pmatrix}, \qquad \mathbf{U} = \begin{pmatrix} \mathbf{U}_B & \mathbf{U}_V \\ 0 & \mathbf{U}_N \end{pmatrix}.$$

Then the associated blocks satisfy the following equalities

$$\begin{array}{ll} \mathbf{X}_B(\mu) = \mathbf{L}_B(\mu)\mathbf{L}_B^T(\mu), & \mathbf{S}_B(\mu) = \mathbf{U}_B(\mu)\mathbf{U}_B^T(\mu) + \mathbf{U}_V(\mu)\mathbf{U}_V^T(\mu), \\ \mathbf{X}_V(\mu) = \mathbf{L}_B(\mu)\mathbf{L}_V(\mu), & \mathbf{S}_V(\mu) = \mathbf{U}_V(\mu)\mathbf{U}_N^T(\mu), \\ \mathbf{X}_N(\mu) = \mathbf{L}_V^T(\mu)\mathbf{L}_V(\mu) + \mathbf{L}_N(\mu)\mathbf{L}_N^T(\mu), & \mathbf{S}_N(\mu) = \mathbf{U}_N(\mu)\mathbf{U}_N^T(\mu). \end{array} \tag{4.10}$$

Proposition 4.1.4 *Let $j \in \{1, \ldots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\begin{array}{ll} \mathbf{L}_B(\mu) = \mathcal{O}(1), & \mathbf{U}_B(\mu) = \mathcal{O}(\sqrt{\mu}), \\ \mathbf{L}_V(\mu) = \mathcal{O}(\sqrt{\mu}), & \mathbf{U}_V(\mu) = \mathcal{O}(\sqrt{\mu}), \\ \mathbf{L}_N(\mu) = \mathcal{O}(\sqrt{\mu}), & \mathbf{U}_N(\mu) = \mathcal{O}(1). \end{array}$$

Proof is similar to the proof of Proposition 4.1.3.

□

4.1.2 Asymptotic properties in Θ-notation

Lemma 4.1.1 *Let $j \in \{1, \ldots, 5\}$. Then there exists a constant $K > 0$ such that for any $\mu \in (0, \mu_0\rangle$*

$$K \leq \frac{1}{\mu^n} \det(\mathbf{X}(\mu) \det \mathbf{S}(\mu)).$$

Proof. The statement will be proved separately for each $j = 1, \ldots, 5$. Assume $j = 1$, i.e. the functions $\mathbf{X}(\mu), \mathbf{S}(\mu)$ satisfy $\mathbf{X}(\mu)\mathbf{S}(\mu) + \mathbf{S}(\mu)\mathbf{X}(\mu) = 2\mu\mathbf{W}$. Proposition A.1.8 then implies

$$0 < \det \mathbf{W} = \det \frac{\mathbf{X}(\mu)\mathbf{S}(\mu) + \mathbf{S}(\mu)\mathbf{X}(\mu)}{2\mu} \leq \frac{\det(\mathbf{X}(\mu)\mathbf{S}(\mu))}{\mu} = \frac{1}{\mu^n} \det \mathbf{X}(\mu) \det \mathbf{S}(\mu).$$

Let $j = 2$ and $\mathbf{X}(\mu)^{\frac{1}{2}}\mathbf{S}(\mu)\mathbf{X}(\mu)^{\frac{1}{2}} = \mu\mathbf{W}$. Obviously, it holds

$$0 < \det \mathbf{W} = \det \frac{\mathbf{X}^{\frac{1}{2}}(\mu)\mathbf{S}(\mu)\mathbf{X}^{\frac{1}{2}}(\mu)}{\mu} =$$

$$= \frac{1}{\mu^n} \det \mathbf{X}^{\frac{1}{2}}(\mu) \det \mathbf{S}(\mu) \det \mathbf{X}^{\frac{1}{2}}(\mu) = \frac{1}{\mu^n} \det \mathbf{X}(\mu) \det \mathbf{S}(\mu)$$

Suppose $j = 3$, i.e. $\mathbf{L}(\mu)^T \mathbf{S}(\mu) \mathbf{L}(\mu) = \mu \mathbf{W}$. Then

$$0 < \det \mathbf{W} = \det \frac{\mathbf{L}(\mu)^T \mathbf{S}(\mu) \mathbf{L}(\mu)}{\mu} = \frac{1}{\mu^n} \det(\mathbf{L}(\mu)\mathbf{L}(\mu)^T) \det \mathbf{S}(\mu) = \frac{1}{\mu^n} \det \mathbf{X}(\mu) \det \mathbf{S}(\mu).$$

Let $j = 4$, i.e. $(\mathbf{X}^{\frac{1}{2}}(\mu)\mathbf{S}^{\frac{1}{2}}(\mu) + \mathbf{S}^{\frac{1}{2}}(\mu)\mathbf{X}^{\frac{1}{2}}(\mu))/2 = \sqrt{\mu}\mathbf{W}$. From Proposition A.1.8 it follows, that

$$0 < \det \mathbf{W}^2 = \det \Big(\frac{\mathbf{X}^{\frac{1}{2}}(\mu)\mathbf{S}^{\frac{1}{2}}(\mu) + \mathbf{S}^{\frac{1}{2}}(\mu)\mathbf{X}^{\frac{1}{2}}(\mu)}{2\sqrt{\mu}} \Big)^2 \leq$$

$$\leq \det \Big(\frac{\mathbf{X}^{\frac{1}{2}}(\mu)\mathbf{S}^{\frac{1}{2}}(\mu)}{\sqrt{\mu}} \Big)^2 = \frac{1}{\mu^n} \det \mathbf{X}(\mu) \det \mathbf{S}(\mu).$$

Finally, let $j = 5$ and $(\mathbf{U}^T_{\mathbf{S}(\mu)}\mathbf{L}_{\mathbf{X}(\mu)} + \mathbf{L}^T_{\mathbf{X}(\mu)}\mathbf{U}_{\mathbf{S}}(\mu))/2 = \sqrt{\mu}\mathbf{W}$. Proposition A.1.8 implies

$$0 < \det \mathbf{W}^2 = \det \Big(\frac{\mathbf{U}(\mu)^T \mathbf{L}(\mu) + \mathbf{L}(\mu)^T \mathbf{U}(\mu)}{2\sqrt{\mu}} \Big)^2 \leq$$

$$\leq \det \Big(\frac{\mathbf{U}(\mu)^T \mathbf{L}(\mu)}{\sqrt{\mu}} \Big)^2 = \frac{1}{\mu^n} (\det \mathbf{U}(\mu)^T \det \mathbf{L}(\mu))^2 =$$

$$\frac{1}{\mu^n} \det(\mathbf{U}(\mu)\mathbf{U}(\mu)^T) \det(\mathbf{L}(\mu)\mathbf{L}(\mu)^T) = \frac{1}{\mu^n} \det \mathbf{X}(\mu) \det \mathbf{S}(\mu).$$

Hence, we can take $K := \det \mathbf{W}$ if $j \in \{1, 2, 3\}$, and $K := \det(\mathbf{W}^2)$ if $j \in \{4, 5\}$.

□

Proposition 4.1.5 *Let $j \in \{1, \ldots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\begin{array}{ll} \mathbf{X}_B(\mu) = \Theta(1), & \mathbf{S}_B(\mu) = \Theta(\mu), \\ \mathbf{X}_N(\mu) = \Theta(\mu), & \mathbf{S}_N(\mu) = \Theta(1). \end{array}$$

Proof. From Fischer inequality (Theorem A.1.5) it follows that

$$\frac{1}{\mu^n} \det \mathbf{X}(\mu) \det \mathbf{S}(\mu) \leq \frac{1}{\mu^n} \det \mathbf{X}_B(\mu) \det \mathbf{X}_N(\mu) \det \mathbf{S}_B(\mu) \det \mathbf{S}_N(\mu) =$$

$$= \det \mathbf{X}_B(\mu) \det \frac{\mathbf{X}_N(\mu)}{\mu} \det \frac{\mathbf{S}_B(\mu)}{\mu} \det \mathbf{S}_N(\mu). \tag{4.11}$$

From (4.11) and Lemma 4.1.1 we have that there exists a constant $K > 0$ such that

$$\ln K \leq \ln \det \mathbf{X}_B(\mu) + \ln \det \frac{\mathbf{X}_N(\mu)}{\mu} + \ln \det \frac{\mathbf{S}_B(\mu)}{\mu} + \ln \det \mathbf{S}_N(\mu) := V(\mu). \quad (4.12)$$

Proposition 4.1.2 implies, that all addends in $V(\mu)$ are bounded above for sufficiently small μ. We will show that $\ln \det \mathbf{X}_B(\mu)$, $\ln \det \frac{\mathbf{X}_N(\mu)}{\mu}$, $\ln \det \frac{\mathbf{S}_B(\mu)}{\mu}$, $\ln \det \mathbf{S}_N(\mu)$ are bounded below. Suppose that there exists a sequence $\{\mu_k\}_{k=1}^{\infty} \to 0$ and an addend in $V(\mu)$ (e.g. the first one) such that

$$\lim_{k\to\infty} \ln \det \mathbf{X}_B(\mu_k) = -\infty.$$

Since all addends in $V(\mu)$ are bounded above, we have that $\lim_{k\to\infty} V(\mu_k) = -\infty$, however this contradicts to (4.12). Therefore there exists a constant C such, that

$$C \leq \min\{\ln \det \mathbf{X}_B(\mu), \ln \det \frac{\mathbf{X}_N(\mu)}{\mu}, \ln \det \frac{\mathbf{S}_B(\mu)}{\mu}, \ln \det \mathbf{S}_N(\mu)\}.$$

The rest follows from Proposition C.2.3.

□

Proposition 4.1.6 *Let $j \in \{1, \ldots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\begin{array}{ll} \mathbf{Y}_B(\mu) = \Theta(1), & \mathbf{Z}_B(\mu) = \Theta(\sqrt{\mu}), \\ \mathbf{Y}_N(\mu) = \Theta(\sqrt{\mu}), & \mathbf{Z}_N(\mu) = \Theta(1). \end{array}$$

Proof. From Lemma 4.1.1 we have that there exists a constant $K > 0$ such that

$$K \leq \frac{1}{\mu^n} \det \mathbf{Y}^2(\mu) \det \mathbf{Z}^2(\mu) = \frac{1}{\mu^n} [\det \mathbf{Y}(\mu) \det \mathbf{Z}(\mu)]^2.$$

We can apply Fischer inequality to obtain

$$\sqrt{K} \leq \det \mathbf{Y}_B(\mu) \det \frac{\mathbf{Y}_N(\mu)}{\sqrt{\mu}} \det \frac{\mathbf{Z}_B(\mu)}{\sqrt{\mu}} \det \mathbf{Z}_N(\mu)$$

The rest of the proof follows from Proposition 4.1.3 and is analogous to the proof of Proposition 4.1.5.

□

Proposition 4.1.7 *Let $j \in \{1, \ldots, 5\}$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\begin{array}{ll} \mathbf{L}_B(\mu) = \Theta(1), & \mathbf{U}_B(\mu) = \Theta(\sqrt{\mu}), \\ \mathbf{L}_N(\mu) = \Theta(\sqrt{\mu}), & \mathbf{U}_N(\mu) = \Theta(1). \end{array}$$

Proof. From Lemma 4.1.1 we have that there exists a constant $K > 0$ such that

$$K \leq \frac{1}{\mu^n} \det[\mathbf{L}(\mu)\mathbf{L}^T(\mu)] \det[\mathbf{U}(\mu)\mathbf{U}^T(\mu)] = \frac{1}{\mu^n}[\det \mathbf{L}(\mu) \det \mathbf{U}(\mu)]^2.$$

Obviously

$$\sqrt{K} \leq \det \mathbf{L}_B(\mu) \det \frac{\mathbf{L}_N(\mu)}{\sqrt{\mu}} \det \frac{\mathbf{U}_B(\mu)}{\sqrt{\mu}} \det \mathbf{U}_N(\mu).$$

The rest of the proof follows from Proposition 4.1.4 and is analogous to the proof of Proposition 4.1.5.

□

4.1.3 Asymptotic properties in o-notation

Let $j \in \{1, \ldots, 5\}$ and consider the weighted path given in (4.1). From Proposition 4.1.2 it follows that the functions

$$\mathbf{X}_B(\mu), \frac{\mathbf{S}_B(\mu)}{\mu}, \frac{\mathbf{X}_V(\mu)}{\sqrt{\mu}}, \frac{\mathbf{S}_V(\mu)}{\sqrt{\mu}}, \frac{\mathbf{X}_N(\mu)}{\mu}, \mathbf{S}_N(\mu)$$

are bounded. Similarly, from Proposition 4.1.3 and Proposition 4.1.4 we have that the functions

$$\mathbf{Y}_B(\mu), \frac{\mathbf{Z}_B(\mu)}{\sqrt{\mu}}, \frac{\mathbf{Y}_V(\mu)}{\sqrt{\mu}}, \frac{\mathbf{Z}_V(\mu)}{\sqrt{\mu}}, \frac{\mathbf{Y}_N(\mu)}{\sqrt{\mu}}, \mathbf{Z}_N(\mu),$$

and

$$\mathbf{L}_B(\mu), \frac{\mathbf{U}_B(\mu)}{\sqrt{\mu}}, \frac{\mathbf{L}_V(\mu)}{\sqrt{\mu}}, \frac{\mathbf{U}_V(\mu)}{\sqrt{\mu}}, \frac{\mathbf{L}_N(\mu)}{\sqrt{\mu}}, \mathbf{U}_N(\mu)$$

are bounded.

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

$$\begin{array}{ll} \mathbf{X}_B(\mu) = \tilde{\mathbf{X}}_B(\rho), & \mathbf{S}_B(\mu) = \rho^2 \tilde{\mathbf{S}}_B(\rho), \\ \mathbf{X}_V(\mu) = \rho \tilde{\mathbf{X}}_V(\rho), & \mathbf{S}_V(\mu) = \rho \tilde{\mathbf{S}}_V(\rho), \\ \mathbf{X}_N(\mu) = \rho^2 \tilde{\mathbf{X}}_N(\rho), & \mathbf{S}_N(\mu) = \tilde{\mathbf{S}}_N(\rho). \end{array} \tag{4.13}$$

Similarly we can define the matrices $\tilde{\mathbf{Y}}(\rho), \tilde{\mathbf{Z}}(\rho)$ and $\tilde{\mathbf{L}}(\rho), \tilde{\mathbf{U}}(\rho)$ with the equalities

$$\begin{array}{ll} \mathbf{Y}_B(\mu) = \tilde{\mathbf{Y}}_B(\rho), & \mathbf{Z}_B(\mu) = \rho\tilde{\mathbf{Z}}_B(\rho), \\ \mathbf{Y}_V(\mu) = \rho\tilde{\mathbf{Y}}_V(\rho), & \mathbf{Z}_V(\mu) = \rho\tilde{\mathbf{Z}}_V(\rho), \\ \mathbf{Y}_N(\mu) = \rho\tilde{\mathbf{Y}}_N(\rho), & \mathbf{Z}_N(\mu) = \tilde{\mathbf{Z}}_N(\rho), \end{array} \tag{4.14}$$

and

$$\begin{array}{ll} \mathbf{L}_B(\mu) = \tilde{\mathbf{L}}_B(\rho), & \mathbf{U}_B(\mu) = \rho\tilde{\mathbf{U}}_B(\rho), \\ \mathbf{L}_V(\mu) = \rho\tilde{\mathbf{L}}_V(\rho), & \mathbf{U}_V(\mu) = \rho\tilde{\mathbf{U}}_V(\rho), \\ \mathbf{L}_N(\mu) = \rho\tilde{\mathbf{L}}_N(\rho), & \mathbf{U}_N(\mu) = \tilde{\mathbf{U}}_N(\rho), \end{array} \tag{4.15}$$

respectively.

The matrices $\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{S}}(\rho),\ \tilde{\mathbf{Y}}(\rho), \tilde{\mathbf{Z}}(\rho),\ \tilde{\mathbf{L}}(\rho), \tilde{\mathbf{U}}(\rho)$ are bounded and therefore there exists a sequence

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

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

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

Lemma 4.1.2 *We have*

$$a)\quad \tilde{\mathbf{X}}_B^* \succ 0,\ \ \tilde{\mathbf{X}}_N^* \succ 0,\ \ \tilde{\mathbf{S}}_B^* \succ 0,\ \ \tilde{\mathbf{S}}_N^* \succ 0,$$

$$b)\quad \tilde{\mathbf{Y}}_B^* \succ 0,\ \ \tilde{\mathbf{Y}}_N^* \succ 0,\ \ \tilde{\mathbf{Z}}_B^* \succ 0,\ \ \tilde{\mathbf{Z}}_N^* \succ 0,$$

$$c)\quad \tilde{\mathbf{L}}_B^* \in L_{++}^n,\ \ \tilde{\mathbf{L}}_N^* \in L_{++}^n,\ \ \tilde{\mathbf{U}}_B^* \in U_{++}^n, \tilde{\mathbf{U}}_N^* \in U_{++}^n.$$

Proof. The statement follows from Proposition 4.1.5, Proposition 4.1.6, Proposition 4.1.7.

□

Lemma 4.1.3 *We have*

$$a)\quad \begin{array}{lcl} \tilde{\mathbf{X}}_B^* &=& (\tilde{\mathbf{Y}}_B^*)^2, \\ \tilde{\mathbf{X}}_V^* &=& \tilde{\mathbf{Y}}_B^*\tilde{\mathbf{Y}}_V^*, \\ \tilde{\mathbf{X}}_N^* &=& (\tilde{\mathbf{Y}}_N^*)^2 + (\tilde{\mathbf{Y}}_V^*)^T\tilde{\mathbf{Y}}_V^*, \end{array} \qquad \begin{array}{lcl} \tilde{\mathbf{S}}_B^* &=& (\tilde{\mathbf{Z}}_B^*)^2 + \tilde{\mathbf{Z}}_V^*(\tilde{\mathbf{Z}}_V^*)^T, \\ \tilde{\mathbf{S}}_V^* &=& \tilde{\mathbf{Z}}_B^*\tilde{\mathbf{Z}}_V^*, \\ \tilde{\mathbf{S}}_N^* &=& (\tilde{\mathbf{Z}}_N^*)^2, \end{array}$$

$$b)\quad \begin{array}{lcl} \tilde{\mathbf{X}}_B^* &=& \tilde{\mathbf{L}}_B^*(\tilde{\mathbf{L}}_B^*)^T, \\ \tilde{\mathbf{X}}_V^* &=& \tilde{\mathbf{L}}_B^*\tilde{\mathbf{L}}_V^*, \\ \tilde{\mathbf{X}}_N^* &=& \tilde{\mathbf{L}}_N^*(\tilde{\mathbf{L}}_N^*)^T + (\tilde{\mathbf{L}}_V^*)^T\tilde{\mathbf{L}}_V^*, \end{array} \qquad \begin{array}{lcl} \tilde{\mathbf{S}}_B^* &=& \tilde{\mathbf{U}}_B^*(\tilde{\mathbf{U}}_B^*)^T + \tilde{\mathbf{U}}_V^*(\tilde{\mathbf{U}}_V^*)^T, \\ \tilde{\mathbf{S}}_V^* &=& \tilde{\mathbf{U}}_V^*(\tilde{\mathbf{U}}_N^*)^T, \\ \tilde{\mathbf{S}}_N^* &=& \tilde{\mathbf{U}}_N^*(\tilde{\mathbf{U}}_N^*)^T. \end{array}$$

Proof. From (4.9) and (4.14) it follows

$$\begin{aligned}
\tilde{\mathbf{X}}_B(\rho_k) &= \tilde{\mathbf{Y}}_B(\rho_k)^2 + \rho_k^2 \tilde{\mathbf{Y}}_V(\rho_k)\tilde{\mathbf{Y}}_V(\rho_k)^T, \\
\rho_k \tilde{\mathbf{X}}_V(\rho_k) &= \rho_k \tilde{\mathbf{Y}}_B(\rho_k)\tilde{\mathbf{Y}}_V(\rho_k) + \rho_k^2 \tilde{\mathbf{Y}}_V(\rho_k)\tilde{\mathbf{Y}}_N(\rho_k), \\
\rho_k^2 \tilde{\mathbf{X}}_N(\rho_k) &= \rho_k^2 \tilde{\mathbf{Y}}_N(\rho_k)^2 + \rho_k^2 \tilde{\mathbf{Y}}_V(\rho_k)^T \tilde{\mathbf{Y}}_V(\rho_k), \\
\rho_k^2 \tilde{\mathbf{S}}_B(\rho_k) &= \rho_k^2 \tilde{\mathbf{Z}}_B(\rho_k)^2 + \rho_k^2 \tilde{\mathbf{Z}}_V(\rho_k)\tilde{\mathbf{Z}}_V(\rho_k)^T, \\
\rho_k \tilde{\mathbf{S}}_V(\rho_k) &= \rho_k \tilde{\mathbf{Z}}_B(\rho_k)\tilde{\mathbf{Z}}_V(\rho_k) + \rho_k^2 \tilde{\mathbf{Z}}_V(\rho_k)\tilde{\mathbf{Z}}_N(\rho_k), \\
\tilde{\mathbf{S}}_N(\rho_k) &= \tilde{\mathbf{Z}}_N(\rho_k)^2 + \rho_k^2 \tilde{\mathbf{Z}}_V(\rho_k)^T \tilde{\mathbf{Z}}_V(\rho_k),
\end{aligned}$$

and from (4.10) and (4.15) it follows

$$\begin{aligned}
\tilde{\mathbf{X}}_B(\rho_k) &= \tilde{\mathbf{L}}_B(\rho_k)\tilde{\mathbf{L}}_B(\rho_k)^T, \\
\rho_k \tilde{\mathbf{X}}_V(\rho_k) &= \rho_k \tilde{\mathbf{L}}_B(\rho_k)\tilde{\mathbf{L}}_V(\rho_k), \\
\rho_k^2 \tilde{\mathbf{X}}_N(\rho_k) &= \rho_k^2 \tilde{\mathbf{L}}_N(\rho_k)\tilde{\mathbf{L}}_N(\rho_k)^T + \rho_k^2 \tilde{\mathbf{L}}_V(\rho_k)^T \tilde{\mathbf{L}}_V(\rho_k), \\
\rho_k^2 \tilde{\mathbf{S}}_B(\rho_k) &= \rho_k^2 \tilde{\mathbf{U}}_B(\rho_k)\tilde{\mathbf{U}}_B(\rho_k)^T + \rho_k^2 \tilde{\mathbf{U}}_V(\rho_k)\tilde{\mathbf{U}}_V(\rho_k)^T, \\
\rho_k \tilde{\mathbf{S}}_V(\rho_k) &= \rho_k \tilde{\mathbf{U}}_V(\rho_k)\tilde{\mathbf{U}}_N(\rho_k)^T, \\
\tilde{\mathbf{S}}_N(\rho_k) &= \tilde{\mathbf{U}}_N(\rho_k)\tilde{\mathbf{U}}_N(\rho_k)^T.
\end{aligned}$$

The statement of the lemma follows from the boundedness of the "tilde" matrices.

□

Lemma 4.1.4 *Let $j \in \{1, 2, 3\}$. Then*

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

Proof. Let $(\mathbf{X}^*, \mathbf{S}^*)$ be the strictly complementary optimal solution given by assumption (A5). Define

$$\bar{\mathbf{X}} = \frac{\mu_l}{\mu_k}\mathbf{X}(\mu_k) + \Big(1 - \frac{\mu_l}{\mu_k}\Big)\mathbf{X}^*, \tag{4.16}$$

$$\bar{\mathbf{S}} = \frac{\mu_l}{\mu_k}\mathbf{S}(\mu_k) + \Big(1 - \frac{\mu_l}{\mu_k}\Big)\mathbf{S}^*. \tag{4.17}$$

It can be easily seen, that

$$(\bar{\mathbf{X}} - \mathbf{X}(\mu_l)) \bullet (\bar{\mathbf{S}} - \mathbf{S}(\mu_l)) = 0$$

and hence

$$\bar{\mathbf{X}} \bullet \mathbf{S}(\mu_l) + \mathbf{X}(\mu_l) \bullet \bar{\mathbf{S}} = \bar{\mathbf{X}} \bullet \bar{\mathbf{S}} + \mathbf{X}(\mu_l) \bullet \mathbf{S}(\mu_l). \tag{4.18}$$

By inserting (4.16), (4.17) into (4.18) we obtain

$$\frac{\mu_k - \mu_l}{\mu_k}[\mathbf{X}^* \bullet \mathbf{S}(\mu_l) + \mathbf{X}(\mu_l) \bullet \mathbf{S}^*] + \frac{\mu_l}{\mu_k}[\mathbf{X}(\mu_k) \bullet \mathbf{S}(\mu_l) + \mathbf{X}(\mu_l) \bullet \mathbf{S}(\mu_k)] =$$

$$= \Big(\frac{\mu_l}{\mu_k}\Big)^2 \mathbf{X}(\mu_k)\bullet\mathbf{S}(\mu_k) + \frac{\mu_l}{\mu_k}\Big(\frac{\mu_k-\mu_l}{\mu_k}\Big)[\mathbf{X}^*\bullet\mathbf{S}(\mu_k)+\mathbf{X}(\mu_k)\bullet\mathbf{S}^*]+\mathbf{X}(\mu_l)\bullet\mathbf{S}(\mu_l). \quad (4.19)$$

By Lemma 3.4.1 we have

$$\begin{array}{rcl} \mathbf{X}(\mu_k)\bullet\mathbf{S}(\mu_k) & = & \mu_k tr(\mathbf{W}), \\ \mathbf{X}(\mu_l)\bullet\mathbf{S}(\mu_l) & = & \mu_l tr(\mathbf{W}). \end{array} \quad (4.20)$$

By inserting (4.20) into (4.19) and by multiplying by $\frac{\mu_k^2}{\mu_l}$ we have that

$$\frac{\mu_k(\mu_k-\mu_l)}{\mu_l}[\mathbf{X}^*\bullet\mathbf{S}(\mu_l)+\mathbf{X}(\mu_l)\bullet\mathbf{S}^*]+\mu_k[\mathbf{X}(\mu_k)\bullet\mathbf{S}(\mu_l)+\mathbf{X}(\mu_l)\bullet\mathbf{S}(\mu_k)] =$$

$$= (\mu_l\mu_k+\mu_k^2)tr(\mathbf{W}) + (\mu_k-\mu_l)[\mathbf{X}^*\bullet\mathbf{S}(\mu_k)+\mathbf{X}(\mu_k)\bullet\mathbf{S}^*]. \quad (4.21)$$

It holds

$$\frac{\mathbf{X}^*\bullet\mathbf{S}(\mu_l)}{\mu_l} = \frac{1}{\mu_l}\begin{pmatrix} \mathbf{X}_B^* & 0 \\ 0 & 0 \end{pmatrix}\bullet\begin{pmatrix} \mathbf{S}_B(\mu_l) & \mathbf{S}_V(\mu_l) \\ \mathbf{S}_V^T(\mu_l) & \mathbf{S}_N(\mu_l) \end{pmatrix} = \frac{1}{\mu_l}tr\begin{pmatrix} \mathbf{X}_B^*\mathbf{S}_B(\mu_l) & * \\ * & 0 \end{pmatrix} =$$

$$= \mathbf{X}_B^*\bullet\frac{\mathbf{S}_B(\mu_l)}{\mu_l} = \mathbf{X}_B^*\bullet\tilde{\mathbf{S}}_B(\rho_l)$$

and hence

$$\lim_{l\to\infty}\frac{\mathbf{X}^*\bullet\mathbf{S}(\mu_l)}{\mu_l} = \lim_{l\to\infty}\mathbf{X}_B^*\bullet\tilde{\mathbf{S}}_B(\rho_l) = \mathbf{X}_B^*\bullet\tilde{\mathbf{S}}_B^*.$$

It can be shown similarly that

$$\lim_{l\to\infty}\frac{\mathbf{X}(\mu_l)\bullet\mathbf{S}^*}{\mu_l} = \lim_{l\to\infty}\tilde{\mathbf{X}}_N(\rho_l)\bullet\mathbf{S}_N^* = \tilde{\mathbf{X}}_N^*\bullet\mathbf{S}_N^*.$$

Further, we have

$$\mathbf{X}(\mu_k)\bullet\mathbf{S}(\mu_l) = \mathbf{X}(\rho_k^2)\bullet\mathbf{S}(\rho_l^2) = \begin{pmatrix} \tilde{\mathbf{X}}_B(\rho_k) & \rho_k\tilde{\mathbf{X}}_V(\rho_k) \\ \rho_k\tilde{\mathbf{X}}_V^T(\rho_k) & \rho_k^2\tilde{\mathbf{X}}_N(\rho_k) \end{pmatrix}\bullet\begin{pmatrix} \rho_l^2\tilde{\mathbf{S}}_B(\rho_l) & \rho_l\tilde{\mathbf{S}}_V(\rho_l) \\ \rho_l\tilde{\mathbf{S}}_V^T(\rho_l) & \tilde{\mathbf{S}}_N(\rho_l) \end{pmatrix} =$$

$$= tr\begin{pmatrix} \rho_l^2\tilde{\mathbf{X}}_B(\rho_k)\tilde{\mathbf{S}}_B(\rho_l)+\rho_k\rho_l\tilde{\mathbf{X}}_V(\rho_k)\tilde{\mathbf{S}}_V^T(\rho_l) & * \\ * & \rho_k\rho_l\tilde{\mathbf{X}}_V(\rho_k)^T\tilde{\mathbf{S}}_V(\rho_l)+\rho_k^2\tilde{\mathbf{X}}_N(\rho_k)\tilde{\mathbf{S}}_N(\rho_l) \end{pmatrix}$$

and therefore

$$\lim_{l\to\infty}\mathbf{X}(\mu_k)\bullet\mathbf{S}(\mu_l) = \mu_k\tilde{\mathbf{X}}_N(\rho_k)\bullet\tilde{\mathbf{S}}_N^*$$

and similarly

$$\lim_{l\to\infty}\mathbf{X}(\mu_l)\bullet\mathbf{S}(\mu_k) = \mu_k\tilde{\mathbf{X}}_B^*\bullet\tilde{\mathbf{S}}_B(\rho_k).$$

It can be easily seen that

$$\mathbf{X}(\mu_k) \bullet \mathbf{S}^* = \mu_k \tilde{\mathbf{X}}_N(\rho_k) \bullet \mathbf{S}_N^*$$

and

$$\mathbf{X}^* \bullet \mathbf{S}(\mu_k) = \mu_k \mathbf{X}_B^* \bullet \tilde{\mathbf{S}}_B(\rho_k).$$

Compute the limit of the left hand side and right hand side of (4.21) as $l \to \infty$. By inserting the expressions above and after multiplying by $\frac{1}{\mu_k^2}$ we obtain

$$[\mathbf{X}_B^* \bullet \tilde{\mathbf{S}}_B^* + \tilde{\mathbf{X}}_N^* \bullet \mathbf{S}_N^*] + [\tilde{\mathbf{X}}_B^* \bullet \tilde{\mathbf{S}}_B(\rho_k) + \tilde{\mathbf{X}}_N(\rho_k) \bullet \tilde{\mathbf{S}}_N^*] = tr\mathbf{W} + [\mathbf{X}_B^* \bullet \tilde{\mathbf{S}}_B(\rho_k) + \tilde{\mathbf{X}}_N(\rho_k) \bullet \mathbf{S}_N^*].$$

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

$$[\mathbf{X}_B^* \bullet \tilde{\mathbf{S}}_B^* + \tilde{\mathbf{X}}_N^* \bullet \mathbf{S}_N^*] + [\tilde{\mathbf{X}}_B^* \bullet \tilde{\mathbf{S}}_B^* + \tilde{\mathbf{X}}_N^* \bullet \tilde{\mathbf{S}}_N^*] = tr\mathbf{W} + [\mathbf{X}_B^* \bullet \tilde{\mathbf{S}}_B^* + \tilde{\mathbf{X}}_N^* \bullet \mathbf{S}_N^*]$$

and hence

$$\tilde{\mathbf{X}}_B^* \bullet \tilde{\mathbf{S}}_B^* + \tilde{\mathbf{X}}_N^* \bullet \tilde{\mathbf{S}}_N^* = tr\mathbf{W}. \tag{4.22}$$

From Lemma 4.0.4 it follows that

$$\rho_k^2 tr\mathbf{W} = \rho_k^2 \tilde{\mathbf{X}}_B(\rho_k) \bullet \tilde{\mathbf{S}}_B(\rho_k) + 2\rho_k^2 \tilde{\mathbf{X}}_V(\rho_k) \bullet \tilde{\mathbf{S}}_V(\rho_k) + \rho_k^2 \tilde{\mathbf{X}}_N(\rho_k) \bullet \tilde{\mathbf{S}}_N(\rho_k).$$

If we multiply the equality above by $\frac{1}{\rho_k^2}$ and take the limit $k \to \infty$, we obtain

$$\tilde{\mathbf{X}}_B^* \bullet \tilde{\mathbf{S}}_B^* + 2\tilde{\mathbf{X}}_V^* \bullet \tilde{\mathbf{S}}_V^* + \tilde{\mathbf{X}}_N^* \bullet \tilde{\mathbf{S}}_N^* = tr\mathbf{W}$$

from which, using (4.22), it follows that $\tilde{\mathbf{X}}_V^* \bullet \tilde{\mathbf{S}}_V^* = 0$.

$\square$

Proposition 4.1.8 *Let $j = 1$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds*

$$\mathbf{X}_V(\mu) = o(\sqrt{\mu}), \qquad \mathbf{S}_V(\mu) = o(\sqrt{\mu}).$$

Proof. It sufficies to show that $\tilde{\mathbf{X}}_V = 0, \tilde{\mathbf{S}}_V^* = 0$. From the equality

$$\mathbf{X}(\mu)\mathbf{S}(\mu) + \mathbf{S}(\mu)\mathbf{X}(\mu) = 2\mu\mathbf{W},$$

we have that

$$\mathbf{X}_B(\mu)\mathbf{S}_V(\mu) + \mathbf{X}_V(\mu)\mathbf{S}_N(\mu) + \mathbf{S}_B(\mu)\mathbf{X}_V(\mu) + \mathbf{S}_V(\mu)\mathbf{X}_N(\mu) = 2\mu\mathbf{W}_V. \tag{4.23}$$

By dividing (4.23) by $\sqrt{\mu}$ and by replacing μ by ρ_k^2, we obtain

$$\tilde{\mathbf{X}}_B(\rho_k)\tilde{\mathbf{S}}_V(\rho_k) + \tilde{\mathbf{X}}_V(\rho_k)\tilde{\mathbf{S}}_N(\rho_k) + \rho_k^2[\tilde{\mathbf{S}}_B(\rho_k)\tilde{\mathbf{X}}_V(\rho_k) + \tilde{\mathbf{S}}_V(\rho_k)\tilde{\mathbf{X}}_N(\rho_k)] = 2\rho_k \mathbf{W}_V.$$

By taking limit $k \to \infty$ we obtain

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

Therefore

$$\tilde{\mathbf{S}}_V^* = -(\tilde{\mathbf{X}}_B^*)^{-1}\tilde{\mathbf{X}}_V^*\tilde{\mathbf{S}}_N^*. \tag{4.24}$$

From Lemma 4.1.4 it follows

$$0 = \tilde{\mathbf{X}}_V^* \bullet \tilde{\mathbf{S}}_V^* = tr[(\tilde{\mathbf{X}}_V^*)^T\tilde{\mathbf{S}}_V^*] = -tr[(\tilde{\mathbf{X}}_V^*)^T(\tilde{\mathbf{X}}_B^*)^{-1}\tilde{\mathbf{X}}_V^*\tilde{\mathbf{S}}_N^*] =$$

$$= -tr[(\tilde{\mathbf{S}}_N^*)^{\frac{1}{2}}(\tilde{\mathbf{X}}_V^*)^T(\tilde{\mathbf{X}}_B^*)^{-1}\tilde{\mathbf{X}}_V^*(\tilde{\mathbf{S}}_N^*)^{\frac{1}{2}}].$$

Since

$$(\tilde{\mathbf{S}}_N^*)^{\frac{1}{2}}(\tilde{\mathbf{X}}_V^*)^T(\tilde{\mathbf{X}}_B^*)^{-1}\tilde{\mathbf{X}}_V^*(\tilde{\mathbf{S}}_N^*)^{\frac{1}{2}} \succeq 0, \quad \tilde{\mathbf{S}}_N^* \succ 0, \quad \tilde{\mathbf{X}}_B^* \succ 0,$$

it holds $\tilde{\mathbf{X}}_V^* = 0$. This fact together with (4.24) implies $\tilde{\mathbf{S}}_V^* = 0$.

□

Proposition 4.1.9 *Let $j = 2$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds that $\mathbf{W}_V = 0$ if and only if $\mathbf{X}_V(\mu) = o(\sqrt{\mu})$, $\mathbf{Y}_V(\mu) = o(\sqrt{\mu})$, $\mathbf{S}_V(\mu) = o(\sqrt{\mu})$.*

Proof. The equality

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

implies

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

$$+ \mathbf{Y}_B(\mu)\mathbf{S}_V(\mu)\mathbf{Y}_N(\mu) + \mathbf{Y}_V(\mu)\mathbf{S}_N(\mu)\mathbf{Y}_N(\mu) = \mu\mathbf{W}_V. \tag{4.25}$$

If we divide (4.25) by μ 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

$$\tilde{\mathbf{Y}}_B^*\tilde{\mathbf{S}}_V^*\tilde{\mathbf{Y}}_N^* + \tilde{\mathbf{Y}}_V^*\tilde{\mathbf{S}}_N^*\tilde{\mathbf{Y}}_N^* = \mathbf{W}_V. \tag{4.26}$$

The implication ($\Leftarrow$) follows from (4.26) and the facts, that if $\mathbf{X}_V(\mu) = o(\sqrt{\mu})$ then $\tilde{\mathbf{Y}}_V^* = 0$ and if $\mathbf{S}_V(\mu) = o(\sqrt{\mu})$ then $\tilde{\mathbf{S}}_V^*$. To prove the reverse implication ($\Rightarrow$), assume $\mathbf{W}_V = 0$. Because $\tilde{\mathbf{Y}}_N^* \succ 0$ (see Lemma 4.1.2), from (4.26) it follows, that

$$\tilde{\mathbf{Y}}_B^* \tilde{\mathbf{S}}_V^* + \tilde{\mathbf{Y}}_V^* \tilde{\mathbf{S}}_N^* = 0. \tag{4.27}$$

From Lemma 4.1.3, Lemma 4.1.4 and (4.27) we obtain that

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

Lemma 4.1.2 states, that $\tilde{\mathbf{S}}_N^* \succ 0$ which gives $\tilde{\mathbf{Y}}_V^* = 0$ and therefore also

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

□

Proposition 4.1.10 *Let $j = 3$. Then for $\mu \in (0, \mu_0\rangle$ sufficiently small it holds, that $\mathbf{W}_V = 0$ if and only if $\mathbf{X}_V(\mu) = o(\sqrt{\mu})$, $\mathbf{L}_V(\mu) = o(\sqrt{\mu})$, $\mathbf{S}_V(\mu) = o(\sqrt{\mu})$.*

Proof. The equality

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

gives

$$\mathbf{L}_B^T(\mu)\mathbf{S}_V(\mu)\mathbf{L}_N(\mu) + \mathbf{L}_V(\mu)\mathbf{S}_N(\mu)\mathbf{L}_N(\mu) = \mu \mathbf{W}_V. \tag{4.28}$$

If we divide (4.28) by μ and put $\mu = \rho_k^2$, we obtain

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

From this, by taking the limit $k \to \infty$, we obtain

$$(\tilde{\mathbf{L}}_B^*)^T \tilde{\mathbf{S}}_V^* \tilde{\mathbf{L}}_N^* + \tilde{\mathbf{L}}_V^* \tilde{\mathbf{S}}_N^* \tilde{\mathbf{L}}_N^* = \mathbf{W}_V.$$

The implication ($\Leftarrow$) follows from the fact, that $\mathbf{X}_V(\mu) = o(\sqrt{\mu}), \mathbf{S}_V(\mu) = o(\sqrt{\mu})$ imply $\tilde{\mathbf{L}}_V^* = 0$ and $\tilde{\mathbf{S}}_V^* = 0$. We will prove the statement ($\Rightarrow$). Since $\tilde{\mathbf{L}}_N^* \in L_{++}^n$ (see Lemma 4.1.2) and $\mathbf{W}_V = 0$, it holds

$$(\tilde{\mathbf{L}}_B^*)^T \tilde{\mathbf{S}}_V^* + \tilde{\mathbf{L}}_V^* \tilde{\mathbf{S}}_N^* = 0. \tag{4.29}$$

From Lemma 4.1.3, Lemma 4.1.4 and (4.29) it follows

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

Lemma 4.1.3 states that $\tilde{\mathbf{S}}_N^* \succ 0$ and hence $\tilde{\mathbf{L}}_V^* = 0$. Therefore also $\tilde{\mathbf{X}}_V^* = 0$ and $\tilde{\mathbf{S}}_V^* = 0$.

□

4.1.4 Summarization of asymptotic behavior of weighted paths

Consider the five types of the weighted paths, associated with symmetrization maps $\Phi_j(\mathbf{X},\mathbf{S})$ for $j=1,\ldots,5$. In the previous section we have shown, that these paths posses two types of asymptotic behavior:

$$\mathbf{X}(\mu)=\begin{pmatrix}\Theta(1) & o(\sqrt{\mu})\\ o(\sqrt{\mu}) & \Theta(\mu)\end{pmatrix},\quad \mathbf{S}(\mu)=\begin{pmatrix}\Theta(\mu) & o(\sqrt{\mu})\\ o(\sqrt{\mu}) & \Theta(1)\end{pmatrix} \tag{4.30}$$

and/or

$$\mathbf{X}(\mu)=\begin{pmatrix}\Theta(1) & \mathcal{O}(\sqrt{\mu})\\ \mathcal{O}(\sqrt{\mu}) & \Theta(\mu)\end{pmatrix},\quad \mathbf{S}(\mu)=\begin{pmatrix}\Theta(\mu) & \mathcal{O}(\sqrt{\mu})\\ \mathcal{O}(\sqrt{\mu}) & \Theta(1)\end{pmatrix} \tag{4.31}$$

In the concrete,

- if $j=1$, the path functions have the property (4.30) (and therefore also (4.31))[2];
- if $j=2,3$, the behavior depends on whether the weight matrix $\mathbf{W}$ is block diagonal - if so, then the path functions have the property (4.30), else their asymptotic behavior is described by (4.31);
- if $j=4,5$, the path functions have the property (4.31).

Moreover, the asymptotic behavior of the square root and Cholesky factors of the path functions was studied. These can be interesting in the case of the last four types of weighted paths. Similarly, these functions posses also two kinds of asymptotic behavior:

$$\mathbf{Y}(\mu)=\mathbf{X}^{\frac{1}{2}}(\mu)=\begin{pmatrix}\Theta(1) & o(\sqrt{\mu})\\ o(\sqrt{\mu}) & \Theta(\mu)\end{pmatrix},\quad \mathbf{L}(\mu)=\mathbf{L}_{\mathbf{X}(\mu)}=\begin{pmatrix}\Theta(1) & 0\\ o(\sqrt{\mu}) & \Theta(\sqrt{\mu})\end{pmatrix} \tag{4.32}$$

and/or

$$\mathbf{Y}(\mu)=\mathbf{X}^{\frac{1}{2}}(\mu)=\begin{pmatrix}\Theta(1) & \mathcal{O}(\sqrt{\mu})\\ \mathcal{O}(\sqrt{\mu}) & \Theta(\sqrt{\mu})\end{pmatrix},\quad \mathbf{Z}(\mu)=\mathbf{S}^{\frac{1}{2}}(\mu)=\begin{pmatrix}\Theta(\sqrt{\mu}) & \mathcal{O}(\sqrt{\mu})\\ \mathcal{O}(\sqrt{\mu}) & \Theta(1)\end{pmatrix}$$

$$\mathbf{L}(\mu)=\mathbf{L}_{\mathbf{X}(\mu)}=\begin{pmatrix}\Theta(1) & 0\\ \mathcal{O}(\sqrt{\mu}) & \Theta(\sqrt{\mu})\end{pmatrix},\quad \mathbf{U}(\mu)=\mathbf{U}_{\mathbf{S}(\mu)}=\begin{pmatrix}\Theta(\sqrt{\mu}) & \mathcal{O}(\sqrt{\mu})\\ 0 & \Theta(1)\end{pmatrix} \tag{4.33}$$

Here

- if $j=2,3$, the behavior, again, depends on whether the weight matrix $\mathbf{W}$ is block diagonal - if so, then the appropriate functions have the property (4.32), else their asymptotic behavior is described by (4.33);
- if $j=4,5$, the functions have the property (4.33).

[2]Recall, that this result was obtained by Preiss and Stoer [59].

4.2 Analyticity of weighted paths at the boundary point

In this section the analyticity of the weighted paths at the boundary will be studied. To this aim, the results stated in the previous section will be very useful. In Section 4.1.4, we could see that the first three types of weighted paths (case $j = 1, 2, 3$) posses "better" behavior and for this reason we will be interested only in these three types of weighted paths.

Note, that the analyticity of these paths was already studied by several authors. The path associated with the symmetrization $(\mathbf{XS} + \mathbf{SX})/2$ (case $j = 1$) was studied by Preiss and Stoer [59], in the context of linear complementarity problems. We will use several ideas of these authors to obtain new interesting results for the another two symmetrizations. The result of [59] will be included also in our study in order to give the complete overview. The analyticity of the weighted path associated with the map $\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}$ (case $j = 2$) was studied by Lu and Monteiro [42]. The weighted path associated with the symmetrization $\mathbf{L_X}^T\mathbf{S}\mathbf{L_X}$ (case $j = 3$) was studied by Chua [8], however only diagonal weights were considered. We will analyze the analyticity of the path for $j = 3$ and we will consider nondiagonal positive definite weights in general. The result [8] follows from our results as a special case. Moreover, for the weighted path studied in [42] (case $j = 2$) we obtain more complete results.

4.2.1 Transformation of feasibility conditions

In order to separate the blocks of the path matrices that posses different types of asymptotic behavior, we need to transform the system of the equations in the feasibility conditions

$$\begin{array}{c} \mathbf{A}^i \bullet \mathbf{X} = b_i + \mu\triangle b_i, \quad i = 1, \dots, m \\ \sum_{i=1}^m \mathbf{A}^i y_i + \mathbf{S} = \mathbf{C} + \mu\triangle\mathbf{C}. \end{array}$$

Consider the partition of the matrices given in (4.2). Then the system above can be rewritten in the following way:

$$\begin{array}{c} \mathbf{A}^i_B \bullet \mathbf{X}_B + 2\mathbf{A}^i_V \bullet \mathbf{X}_V + \mathbf{A}^i_N \bullet \mathbf{X}_N = b_i + \mu\triangle b_i, \ i = 1, \dots, m, \\ \sum_{i=1}^m \mathbf{A}^i_B y_i + \mathbf{S}_B = \mathbf{C}_B + \mu\triangle\mathbf{C}_B, \\ \sum_{i=1}^m \mathbf{A}^i_V y_i + \mathbf{S}_V = \mathbf{C}_V + \mu\triangle\mathbf{C}_V, \\ \sum_{i=1}^m \mathbf{A}^i_N y_i + \mathbf{S}_N = \mathbf{C}_N + \mu\triangle\mathbf{C}_N. \end{array} \tag{4.34}$$

Define the matrices

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

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 (4.34) has the matrix-vector form

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

Rewrite the system above once more as

$$\mathbb{P}v + \mathbb{Q}w + \mathbb{R}z = d + \mu\triangle d, \tag{4.35}$$

where

$$\mathbb{P} = \left[\begin{array}{ccc} \mathbb{A}_B & 0 & 0 \\ 0 & \mathbb{A}_B^T & 0 \\ 0 & \mathbb{A}_V^T & 0 \\ 0 & \mathbb{A}_N^T & \mathbf{I}_{\bar{N}} \end{array}\right], \quad \mathbb{Q} = \left[\begin{array}{cc} 2\mathbb{A}_V & 0 \\ 0 & 0 \\ 0 & \mathbf{I}_{\bar{V}} \\ 0 & 0 \end{array}\right], \quad \mathbb{R} = \left[\begin{array}{cc} \mathbb{A}_N & 0 \\ 0 & \mathbf{I}_{\bar{B}} \\ 0 & 0 \\ 0 & 0 \end{array}\right];$$

$$v = \left[\begin{array}{c} svec(\mathbf{X}_B) \\ y \\ svec(\mathbf{S}_N) \end{array}\right], \quad w = \left[\begin{array}{c} vec(\mathbf{X}_V) \\ vec(\mathbf{S}_V) \end{array}\right], \quad z = \left[\begin{array}{c} svec(\mathbf{X}_N) \\ svec(\mathbf{S}_B) \end{array}\right]$$

and

$$d = \left[\begin{array}{c} b \\ svec(\mathbf{C}_B) \\ vec(\mathbf{C}_V) \\ svec(\mathbf{C}_N) \end{array}\right], \qquad \triangle d = \left[\begin{array}{c} \triangle b \\ svec(\triangle\mathbf{C}_B) \\ vec(\triangle\mathbf{C}_V) \\ svec(\triangle\mathbf{C}_N) \end{array}\right].$$

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

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

Lemma 4.2.1 *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* $rank(\mathbf{M}_1\mathbf{A}) = s$*.*

Proof. The existence of the matrix $\mathbf{M}$ follows from the Gaussian elimination. According to the well-known Sylvester theorem (Theorem 2.6 of [81]) we have

$$rank(\mathbf{M}_1\mathbf{A}) = rank(\mathbf{A}) - \dim(Im(\mathbf{A}) \cap Ker(\mathbf{M}_1)).$$

We will show that $Im(\mathbf{A}) \cap Ker(\mathbf{M}_1) = \{0\}$. Assume that there exists $y \neq 0$ such that $y \in Im(\mathbf{A}) \cap Ker(\mathbf{M}_1)$. Then $y = \mathbf{A}x$ for some $x \in R^m, x \neq 0$ and $\mathbf{M}_1 y = 0$. From this we have that $\mathbf{M}_1\mathbf{A}x = 0$ and therefore $\mathbf{MA}x = 0$. From the nonsingularity of $\mathbf{M}$ it follows that $\mathbf{A}x = 0 = y$.

□

Let

$$s := rank(\mathbb{P}) \leq \min\{k_1, m + \bar{n}\} = k_1. \tag{4.36}$$

Then from Lemma 4.2.1 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 (4.35) by $\mathbf{M}$ from the left we obtain an equivalent system

$$\begin{aligned} \mathbf{M}_1\mathbb{P}v + \mathbf{M}_1\mathbb{Q}w + \mathbf{M}_1\mathbb{R}z &= \mathbf{M}_1(d + \mu\triangle d), \\ \mathbf{M}_2\mathbb{Q}w + \mathbf{M}_2\mathbb{R}z &= \mathbf{M}_2(d + \mu\triangle d). \end{aligned} \tag{4.37}$$

Now let

$$t - s := rank(\mathbf{M}_2\mathbb{Q}) \leq \min\{m + \bar{n} - s, k_2\}. \tag{4.38}$$

Then again, Lemma 4.2.1 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 (4.37) is equivalent to

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

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

$$\begin{aligned}
\tilde{\mathbb{P}}_1v + \tilde{\mathbb{Q}}_1w + \tilde{\mathbb{R}}_1z &= \tilde{d}_1 + \mu\triangle\tilde{d}_1,\\
\tilde{\mathbb{Q}}_2w + \tilde{\mathbb{R}}_2z &= \tilde{d}_2 + \mu\triangle\tilde{d}_2,\\
\tilde{\mathbb{R}}_3z &= \tilde{d}_3 + \mu\triangle\tilde{d}_3.
\end{aligned} \tag{4.39}$$

The following lemma states a well-known linear algebra result, which will be useful in the next.

Lemma 4.2.2 *If $\mathbf{A}$ is $n\times m$ matrix and $n \le m$. Then the map $x \mapsto \mathbf{A}x$ is surjective if and only if $rank(\mathbf{A}) = n$.*

Lemma 4.2.3 *The linear maps*

$$v \mapsto \tilde{\mathbb{P}}_1v, \quad w \mapsto \tilde{\mathbb{Q}}_2w, \quad z \mapsto \tilde{\mathbb{R}}_3z$$

are surjective.

Proof. $\tilde{\mathbb{P}}_1 = \mathbf{M}_1\mathbb{P}$ is $s\times k_1$ matrix and $rank(\tilde{\mathbb{P}}_1) = s$. From (4.36) and Lemma 4.2.2 it follows that the map $v \mapsto \tilde{\mathbb{P}}_1v$ is surjective. Similarly $\tilde{\mathbb{Q}}_2 = \mathbf{N}_1\mathbf{M}_2\mathbb{Q}$ is $(t-s)\times k_2$ matrix and $rank(\tilde{\mathbb{Q}}_2) = t-s$. From (4.38) and Lemma 4.2.2. it follows that the map $w \mapsto \tilde{\mathbb{Q}}_2w$ is surjective. Finally, from Assumption (A1) and Lemma 2.1.2 we have that

$$\{[\mathcal{A}(\mathbf{X}), \mathcal{A}^*(y)+\mathbf{S}] \mid \mathbf{X}\in S^n, y\in R^m, \mathbf{S}\in S^n\} = R^m\times S^n$$

and hence also the map z $\mapsto \tilde{\mathbb{R}}_3z$ is surjective.

$\square$

Proposition 4.2.1 *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 4.2.3) we have that there exist matrices $\mathbf{V}_1, \mathbf{V}_2$ of dimension $|B| \times |N|$ such that

$$\tilde{\mathbb{Q}}_2 \begin{bmatrix} vec(\mathbf{V}_1) \\ vec(\mathbf{V}_2) \end{bmatrix} + \tilde{\mathbb{R}}_2 \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0. \tag{4.40}$$

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

$$\tilde{\mathbb{P}}_1 \begin{bmatrix} svec(\mathbf{U}_1) \\ u \\ svec(\mathbf{U}_2) \end{bmatrix} + \tilde{\mathbb{Q}}_1 \begin{bmatrix} vec(\mathbf{V}_1) \\ vec(\mathbf{V}_2) \end{bmatrix} + \tilde{\mathbb{R}}_1 \begin{bmatrix} svec(\triangle \mathbf{X}_N) \\ svec(\triangle \mathbf{S}_B) \end{bmatrix} = 0. \tag{4.41}$$

The equations (4.40), (4.41) 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

$$\mathcal{A}\left(\begin{bmatrix} \mathbf{U}_1 & \mathbf{V}_1 \\ \mathbf{V}_1^T & \triangle \mathbf{X}_N \end{bmatrix} \right) = 0, \quad \mathcal{A}^*(u) + \begin{bmatrix} \triangle \mathbf{S}_B & \mathbf{V}_2 \\ \mathbf{V}_2^T & \mathbf{U}_2 \end{bmatrix} = 0. \tag{4.42}$$

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

$$\mathcal{A}\left(\begin{bmatrix} \triangle\mathbf{X}_B & 0 \\ 0 & 0 \end{bmatrix}\right) = 0, \quad \mathcal{A}^*(\triangle y) + \begin{bmatrix} 0 & 0 \\ 0 & \triangle\mathbf{S}_N \end{bmatrix} = 0. \tag{4.43}$$

From (4.42) and (4.43) 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 Lemma 2.1.1. 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 4.2.3) 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

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

Lemma 2.1.1 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

$$\tilde{\mathbb{P}}_1 \begin{bmatrix} svec(\mathbf{W}_1) \\ w \\ svec(\mathbf{W}_2) \end{bmatrix} + \tilde{\mathbb{Q}}_1 \begin{bmatrix} vec(\triangle\mathbf{X}_V) \\ vec(\triangle\mathbf{S}_V) \end{bmatrix} + \tilde{\mathbb{R}}_1 \begin{bmatrix} svec(\triangle\mathbf{X}_N) \\ svec(\triangle\mathbf{S}_B) \end{bmatrix} = 0. \tag{4.44}$$

The equation (4.44) 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

$$\mathcal{A}\left(\begin{bmatrix} \mathbf{W}_1 & \triangle\mathbf{X}_V \\ \triangle\mathbf{X}_V^T & \triangle\mathbf{X}_N \end{bmatrix}\right) = 0, \quad \mathcal{A}^*(u) + \begin{bmatrix} \triangle\mathbf{S}_B & \triangle\mathbf{S}_V \\ \triangle\mathbf{S}_V^T & \mathbf{W}_2 \end{bmatrix} = 0. \tag{4.45}$$

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 $\triangle\mathbf{X}_N = 0$. Lemma 2.1.1, (4.45) and Lemma 4.0.4 yield

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

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

$\square$

4.2.2 Normalization of feasibility conditions

Consider the normalized matrices defined in (4.13). Recall, that

$$\tilde{\mathbf{X}}(\rho) := \begin{pmatrix} \mathbf{X}_B(\rho^2) & \mathbf{X}_V(\rho^2)/\rho \\ \mathbf{X}_V(\rho^2)^T/\rho & \mathbf{X}_N(\rho^2)/\rho^2 \end{pmatrix}, \; \tilde{\mathbf{S}}(\rho) := \begin{pmatrix} \mathbf{S}_B(\rho^2)/\rho^2 & \mathbf{S}_V(\rho^2)/\rho \\ \mathbf{S}_V(\rho^2)^T/\rho & \mathbf{S}_N(\rho^2) \end{pmatrix}, \tag{4.46}$$

where $\rho := \sqrt{\mu}$.

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 (4.39) we obtain

$$\begin{array}{rcl} \tilde{\mathbb{P}}_1\tilde{v}(\rho) + \rho\tilde{\mathbb{Q}}_1\tilde{w}(\rho) + \rho^2\tilde{\mathbb{R}}_1\tilde{z}(\rho) & = & \tilde{d}_1 + \rho^2\triangle\tilde{d}_1, \\ \rho\tilde{\mathbb{Q}}_2\tilde{w}(\rho) + \rho^2\tilde{\mathbb{R}}_2\tilde{z}(\rho) & = & \tilde{d}_2 + \rho^2\triangle\tilde{d}_2, \\ \rho^2\tilde{\mathbb{R}}_3\tilde{z}(\rho) & = & \tilde{d}_3 + \rho^2\triangle\tilde{d}_3. \end{array} \tag{4.47}$$

From the asymptotic behavior of the path functions (see Section 4.1.4 or Proposition 4.1.2 and Proposition 4.1.5) it follows that (for any $j \in \{1, 2, 3\}$)

$$\tilde{\mathbf{X}}(\rho) = \begin{pmatrix} \Theta(1) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}, \quad \tilde{\mathbf{S}}(\rho) = \begin{pmatrix} \Theta(1) & O(1) \\ O(1) & \Theta(1) \end{pmatrix}.$$

Therefore, for any sequence $\{\rho_k\} \to 0$, the matrices $\tilde{\mathbf{X}}(\rho_k)$, $\tilde{\mathbf{S}}(\rho_k)$ and the vector $\tilde{y}(\rho_k)$ are bounded, so we may assume that the limit $\lim_{k\to\infty}(\tilde{\mathbf{X}}(\rho_k), \tilde{y}(\rho_k), \tilde{\mathbf{S}}(\rho_k)) = (\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*)$ exists. Moreover, the matrices $\tilde{\mathbf{X}}_B^*, \tilde{\mathbf{X}}_N^*, \tilde{\mathbf{S}}_B^*, \tilde{\mathbf{S}}_N^*$ are positive definite (see Lemma 4.1.2). 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 (4.47) and letting $\rho_k \to 0$ we find that $\tilde{d}_2 = 0, \tilde{d}_3 = 0$.

Define the map Ψ in the following way:

$$\Psi(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = \begin{bmatrix} \tilde{\mathbb{P}}_1\tilde{v} + \rho\tilde{\mathbb{Q}}_1\tilde{w} + \rho^2\tilde{\mathbb{R}}_1\tilde{z} - \tilde{d}_1 + \rho^2\triangle\tilde{d}_1 \\ \tilde{\mathbb{Q}}_2\tilde{w} + \rho\tilde{\mathbb{R}}_2\tilde{z} - \rho\triangle\tilde{d}_2 \\ \tilde{\mathbb{R}}_3\tilde{z} - \triangle\tilde{d}_3 \end{bmatrix}, \tag{4.48}$$

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 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

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

From Proposition 4.2.1 it immediately follows that

$$D\Psi(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] = 0 \;\Rightarrow\; \triangle\tilde{\mathbf{X}} \bullet \triangle\tilde{\mathbf{S}} = 0. \tag{4.49}$$

4.2.3 Introduction of normalized system and nonsingularity of Fréchet derivative (I)

In this section, for each $j = 1, 2, 3$ we define a normalized map $\tilde{F}^j$, such, that for any $\rho \in (0, \rho_0\rangle$, where $\rho_0 := \sqrt{\mu_0}$, the triple $(\tilde{\mathbf{X}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ is the unique solution of the system

$$\tilde{F}^j(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = 0, \quad \tilde{\mathbf{X}} \succ 0, \tilde{\mathbf{S}} \succ 0.$$

Moreover, it will be shown that the Fréchet derivative of $\tilde{F}^j$ with respect to $(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}})$ is a nonsingular linear map at the point $(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$.

Symmetrization $(\mathbf{XS} + \mathbf{SX})/2$

Consider the last condition in the system (3.1):

$$\mathbf{XS} + \mathbf{SX} = 2\mu\mathbf{W}$$

and rewrite it in the block form:

$$\begin{array}{rcl} \mathbf{X}_B\mathbf{S}_B + \mathbf{S}_B\mathbf{X}_B + \mathbf{X}_V\mathbf{S}_V^T + \mathbf{S}_V\mathbf{X}_V^T & = & 2\mu\mathbf{W}_B \\ \mathbf{X}_B\mathbf{S}_V + \mathbf{S}_B\mathbf{X}_V + \mathbf{X}_V\mathbf{S}_N + \mathbf{S}_V\mathbf{X}_N & = & 2\mu\mathbf{W}_V \\ \mathbf{X}_N\mathbf{S}_N + \mathbf{S}_N\mathbf{X}_N + \mathbf{X}_V^T\mathbf{S}_V + \mathbf{S}_V^T\mathbf{X}_V & = & 2\mu\mathbf{W}_N \end{array} \tag{4.50}$$

Because $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ satisfies the system above for $\mu > 0$, we have that for any $\rho > 0$ the triple $(\tilde{\mathbf{X}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ satisfies

$$\begin{array}{rcl} \tilde{\mathbf{X}}_B(\rho)\tilde{\mathbf{S}}_B(\rho) + \tilde{\mathbf{S}}_B(\rho)\tilde{\mathbf{X}}_B(\rho) + \tilde{\mathbf{X}}_V(\rho)\tilde{\mathbf{S}}_V(\rho)^T + \tilde{\mathbf{S}}_V(\rho)\tilde{\mathbf{X}}_V(\rho)^T & = & 2\mathbf{W}_B \\ \tilde{\mathbf{X}}_B(\rho)\tilde{\mathbf{S}}_V(\rho) + \rho^2\tilde{\mathbf{S}}_B(\rho)\tilde{\mathbf{X}}_V(\rho) + \tilde{\mathbf{X}}_V(\rho)\tilde{\mathbf{S}}_N(\rho) + \rho^2\tilde{\mathbf{S}}_V(\rho)\tilde{\mathbf{X}}_N(\rho) & = & 2\rho\mathbf{W}_V \\ \tilde{\mathbf{X}}_N(\rho)\tilde{\mathbf{S}}_N(\rho) + \tilde{\mathbf{S}}_N(\rho)\tilde{\mathbf{X}}_N(\rho) + \tilde{\mathbf{X}}_V(\rho)^T\tilde{\mathbf{S}}_V(\rho) + \tilde{\mathbf{S}}_V(\rho)^T\tilde{\mathbf{X}}_V(\rho) & = & 2\mathbf{W}_N \end{array} \tag{4.51}$$

If we put $\rho := \rho_k$ and take the limit $k \to \infty$ in the equations above we find that

$$\begin{array}{c} \tilde{\mathbf{X}}_B^*\tilde{\mathbf{S}}_B^* + \tilde{\mathbf{S}}_B^*\tilde{\mathbf{X}}_B^* = 2\mathbf{W}_B \succ 0 \\ \tilde{\mathbf{X}}_N^*\tilde{\mathbf{S}}_N^* + \tilde{\mathbf{S}}_N^*\tilde{\mathbf{X}}_N^* = 2\mathbf{W}_N \succ 0 \end{array} \tag{4.52}$$

Moreover, because of (4.30), we have that $\tilde{\mathbf{X}}_V^* = 0$ and $\tilde{\mathbf{S}}_V^* = 0$. Define the map $\tilde{F}^1$ in the following way:

$$\tilde{F}^1(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = \begin{bmatrix} \Psi(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) \\ \tilde{\mathbf{X}}_B\tilde{\mathbf{S}}_B + \tilde{\mathbf{S}}_B\tilde{\mathbf{X}}_B + \tilde{\mathbf{X}}_V\tilde{\mathbf{S}}_V^T + \tilde{\mathbf{S}}_V\tilde{\mathbf{X}}_V^T - 2\mathbf{W}_B \\ \tilde{\mathbf{X}}_B\tilde{\mathbf{S}}_V + \rho^2\tilde{\mathbf{S}}_B\tilde{\mathbf{X}}_V + \tilde{\mathbf{X}}_V\tilde{\mathbf{S}}_N + \rho^2\tilde{\mathbf{S}}_V\tilde{\mathbf{X}}_N - 2\rho\mathbf{W}_V \\ \tilde{\mathbf{X}}_N\tilde{\mathbf{S}}_N + \tilde{\mathbf{S}}_N\tilde{\mathbf{X}}_N + \tilde{\mathbf{X}}_V^T\tilde{\mathbf{S}}_V + \tilde{\mathbf{S}}_V^T\tilde{\mathbf{X}}_V - 2\mathbf{W}_N \end{bmatrix},$$

where Ψ is defined by (4.48). Obviously, for $\rho > 0$ (sufficiently small)

$$\tilde{F}^1(\tilde{\mathbf{X}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho), \rho) = 0,$$

and also

$$\tilde{F}^1(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0) = 0.$$

The Fréchet derivative of $\tilde{F}^1$ with respect to $(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}})$ at the point $(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ is the linear map given by

$$D\tilde{F}^1(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] = \begin{bmatrix} D\Psi(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] \\ \triangle\tilde{\mathbf{X}}_B\tilde{\mathbf{S}}^*_B + \triangle\tilde{\mathbf{S}}_B\tilde{\mathbf{X}}^*_B + \tilde{\mathbf{X}}^*_B\triangle\tilde{\mathbf{S}}_B + \tilde{\mathbf{S}}^*_B\triangle\tilde{\mathbf{X}}_B \\ \tilde{\mathbf{X}}^*_B\triangle\tilde{\mathbf{S}}_V + \triangle\tilde{\mathbf{X}}_V\tilde{\mathbf{S}}^*_N \\ \triangle\tilde{\mathbf{X}}_N\tilde{\mathbf{S}}^*_N + \triangle\tilde{\mathbf{S}}_N\tilde{\mathbf{X}}^*_N + \tilde{\mathbf{X}}^*_N\triangle\tilde{\mathbf{S}}_N + \tilde{\mathbf{S}}^*_N\triangle\tilde{\mathbf{X}}_N \end{bmatrix}.$$

Lemma 4.2.4 $D\tilde{F}^1(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ *is a nonsingular linear map.*

Proof. Assume

$$D\tilde{F}^1(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] = 0 \tag{4.53}$$

This implies

$$\triangle\tilde{\mathbf{X}}_B\tilde{\mathbf{S}}^*_B + \triangle\tilde{\mathbf{S}}_B\tilde{\mathbf{X}}^*_B + \tilde{\mathbf{X}}^*_B\triangle\tilde{\mathbf{S}}_B + \tilde{\mathbf{S}}^*_B\triangle\tilde{\mathbf{X}}_B = 0,$$

which is equivalent to

$$(\mathbf{I} \star \tilde{\mathbf{S}}^*_B)svec(\triangle\tilde{\mathbf{X}}_B) + (\mathbf{I} \star \tilde{\mathbf{X}}^*_B)svec(\triangle\tilde{\mathbf{S}}_B) = 0,$$

(see $(A.2)$ in Appendix A.4). Since $\tilde{\mathbf{S}}^*_B \succ 0$ (Lemma 4.1.2), it holds $\mathbf{I} * \tilde{\mathbf{S}}^*_B \succ 0$. Therefore

$$svec(\triangle\tilde{\mathbf{X}}_B) = -(\mathbf{I} * \tilde{\mathbf{S}}^*_B)^{-1}(\mathbf{I} * \tilde{\mathbf{X}}^*_B)svec\triangle\tilde{\mathbf{S}}_B.$$

By using Proposition 4.2.1 we obtain

$$0 = (svec\triangle\tilde{\mathbf{X}}_B)^T(svec\triangle\tilde{\mathbf{S}}_B) = -(svec\triangle\tilde{\mathbf{S}}_B)^T(\mathbf{I} * \tilde{\mathbf{X}}^*_B)(\mathbf{I} * \tilde{\mathbf{S}}^*_B)^{-1}(svec\triangle\tilde{\mathbf{S}}_B)$$

From Lemma 4.1.2, (4.52) and Corollary A.4.4 it follows that

$$(\mathbf{I} * \tilde{\mathbf{X}}^*_B)(\mathbf{I} * \tilde{\mathbf{S}}^*_B)^{-1} \succ 0$$

and hence $\triangle\tilde{\mathbf{S}}_B = 0$ and $\triangle\tilde{\mathbf{X}}_B = 0$.

Proposition 4.2.1 states that $\triangle\tilde{\mathbf{X}}_N \bullet \triangle\tilde{\mathbf{S}}_N = 0$. Similarly one can obtain $\triangle\tilde{\mathbf{S}}_N = 0$ and $\triangle\tilde{\mathbf{X}}_N = 0$.

Finally, from (4.53) we have that

$$\tilde{\mathbf{X}}_B^* \triangle \tilde{\mathbf{S}}_V + \triangle \tilde{\mathbf{X}}_V \tilde{\mathbf{S}}_N^* = 0$$

and therefore

$$\triangle \tilde{\mathbf{S}}_V = -(\tilde{\mathbf{X}}_B^*)^{-1} \triangle \tilde{\mathbf{X}}_V \tilde{\mathbf{S}}_N^*.$$

By using Proposition 4.2.1 we obtain

$$0 = tr[(\triangle \tilde{\mathbf{X}}_V)^T \triangle \tilde{\mathbf{S}}_V] = -tr[(\triangle \tilde{\mathbf{X}}_V)^T (\tilde{\mathbf{X}}_B^*)^{-1} \triangle \tilde{\mathbf{X}}_V \tilde{\mathbf{S}}_N^*] =$$

$$= -tr[(\tilde{\mathbf{S}}_N^*)^{\frac{1}{2}} (\triangle \tilde{\mathbf{X}}_V)^T (\tilde{\mathbf{X}}_B^*)^{-1} \triangle \tilde{\mathbf{X}}_V (\tilde{\mathbf{S}}_N^*)^{\frac{1}{2}}]$$

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

□

Symmetrization $\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}$

Note that this part of the proof is adapted from [42], and is included in order to give a complete analysis of the problem.

Consider the condition

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

which can be equivalently rewritten as the pair

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

Consider the normalized matrices $\tilde{\mathbf{X}}(\rho)$, $\tilde{\mathbf{S}}(\rho)$ given in (4.46), or (4.13), associated with the weighted path functions $\mathbf{X}(\mu)$, $\mathbf{S}(\mu)$ for some weight $\mathbf{W} \in \mathcal{W}_2$.

Let $\mathcal{U}_{BN}^n$ be the vector space of all upper block triangular matrices with symmetric diagonal blocks of dimensions $|B| \times |B|$ and $|N| \times |N|$ and L be the linear map $\mathcal{U}_{BN}^n \to R^{n\times n}$ defined as

$$L\left(\begin{pmatrix} \mathbf{Y}_B & \mathbf{Y}_V \\ 0 & \mathbf{Y}_N \end{pmatrix}\right) = \begin{pmatrix} 0 & 0 \\ \mathbf{Y}_V^T & 0 \end{pmatrix}.$$

Define $\tilde{\mathbf{U}}(\rho) \in \mathcal{U}_{BN}^n$ as

$$\tilde{\mathbf{U}}(\rho) = \begin{pmatrix} \mathbf{Y}_B(\rho^2) & \mathbf{Y}_V(\rho^2)/\rho \\ 0 & \mathbf{Y}_N(\rho^2)/\rho \end{pmatrix}.$$

From Proposition 4.1.2 and Proposition 4.1.3 it follows that for any sequence $\{\rho_k\} \to 0$ the sequence

$$(\tilde{\mathbf{X}}(\rho_k), \tilde{\mathbf{U}}(\rho_k), \tilde{y}(\rho_k), \tilde{\mathbf{S}}(\rho_k))$$

is bounded and hence we may assume that the limit $\lim_{k\to\infty}(\tilde{\mathbf{X}}(\rho_k), \tilde{\mathbf{U}}(\rho_k), \tilde{y}(\rho_k), \tilde{\mathbf{S}}(\rho_k)) = (\tilde{\mathbf{X}}^*, \tilde{\mathbf{U}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*)$ exists. It can be easily seen that

$$\begin{array}{rcl} (\tilde{\mathbf{U}}(\rho) + \rho L(\tilde{\mathbf{U}}(\rho)))\tilde{\mathbf{S}}(\rho)(\tilde{\mathbf{U}}(\rho) + \rho L(\tilde{\mathbf{U}}(\rho)))^T & = & \mathbf{W} \\ (\tilde{\mathbf{U}}(\rho) + \rho L(\tilde{\mathbf{U}}(\rho)))^T(\tilde{\mathbf{U}}(\rho) + \rho L(\tilde{\mathbf{U}}(\rho))) & = & \tilde{\mathbf{X}}(\rho). \end{array} \tag{4.54}$$

Define the map $\tilde{F}^2$ in the following way:

$$\tilde{F}^2(\tilde{\mathbf{X}}, \tilde{\mathbf{U}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) = \left[\begin{array}{c} \Psi(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho) \\ (\tilde{\mathbf{U}} + \rho L(\tilde{\mathbf{U}}))\tilde{\mathbf{S}}(\tilde{\mathbf{U}} + \rho L(\tilde{\mathbf{U}}))^T - \mathbf{W} \\ (\tilde{\mathbf{U}} + \rho L(\tilde{\mathbf{U}}))^T(\tilde{\mathbf{U}} + \rho L(\tilde{\mathbf{U}})) - \tilde{\mathbf{X}} \end{array} \right],$$

where Ψ is defined by (4.48). Obviously for $\rho > 0$ (sufficiently small)

$$\tilde{F}^2(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{U}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho), \rho) = 0$$

and

$$\tilde{F}^2(\tilde{\mathbf{X}}^*, \tilde{\mathbf{U}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0) = 0.$$

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

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

$$= \left[\begin{array}{c} D\Psi(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)[\triangle\tilde{\mathbf{X}}, \triangle\tilde{y}, \triangle\tilde{\mathbf{S}}] \\ \triangle\tilde{\mathbf{U}}\tilde{\mathbf{S}}^*(\tilde{\mathbf{U}}^*)^T + \tilde{\mathbf{U}}^*\triangle\tilde{\mathbf{S}}(\tilde{\mathbf{U}}^*)^T + \tilde{\mathbf{U}}^*\tilde{\mathbf{S}}^*(\triangle\tilde{\mathbf{U}})^T \\ (\triangle\tilde{\mathbf{U}})^T\tilde{\mathbf{U}}^* + (\tilde{\mathbf{U}}^*)^T\triangle\tilde{\mathbf{U}} - \triangle\tilde{\mathbf{X}} \end{array} \right].$$

The following lemma is a consequence of Lemma 3.8 and Lemma 3.9 of [42] and can be proved using a similar technique like Lemma 3.2.1 or Lemma 3.2.2.

Lemma 4.2.5 *Let $\gamma \in \langle 0, \frac{1}{\sqrt{2}})$, and let $\mathbf{U} \in \mathcal{U}^n_{BN}$ and $\mathbf{S} \in S^n$ be such that $\mathbf{U}_B \succ 0$, $\mathbf{U}_N \succ 0$ and $\|\mathbf{U}\mathbf{S}\mathbf{U}^T - \nu\mathbf{I}\|_2 \leq \gamma\nu$ for some $\nu > 0$. Then for $\triangle\mathbf{U} \in \mathcal{U}^n_{BN}$ and $\triangle\mathbf{X}, \triangle\mathbf{S} \in S^n$ the following implication holds:*

$$\left. \begin{array}{r} \triangle\mathbf{U}\mathbf{S}\mathbf{U}^T + \mathbf{U}\triangle\mathbf{S}\mathbf{U}^T + \mathbf{U}\mathbf{S}\triangle\mathbf{U}^T = 0 \\ \triangle\mathbf{U}^T\mathbf{U} + \mathbf{U}^T\triangle\mathbf{U} = \triangle\mathbf{X} \\ \triangle\mathbf{X} \bullet \triangle\mathbf{S} = 0 \end{array} \right\} \Rightarrow \ \triangle\mathbf{U} = \triangle\mathbf{X} = \triangle\mathbf{S} = 0. \tag{4.55}$$

Corollary 4.2.1 *Let* $\mathbf{U} \in \mathcal{U}^n_{BN}$ *and* $\mathbf{S} \in S^n$ *be such that* $\mathbf{U}_B \succ 0$, $\mathbf{U}_N \succ 0$. *If there exists* $\nu > 0$ *such that* $\|\mathbf{U}\mathbf{S}\mathbf{U}^T - \nu\mathbf{I}\|_2 < \frac{\nu}{\sqrt{2}}$, *then* (4.55) *holds.*

The following lemma directly follows from (4.49) and Corollary 4.2.1.

Lemma 4.2.6 $D\tilde{F}^2(\tilde{\mathbf{X}}^*, \tilde{\mathbf{U}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ *is a nonsingular linear map.*

Proof. Assume

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

From (4.49) it follows that $\triangle\tilde{\mathbf{X}} \bullet \triangle\tilde{\mathbf{S}} = 0$. By inserting $\rho = \rho_k$ into the first equation of (4.54) and by taking the limit $k \to \infty$ we obtain that $\tilde{\mathbf{U}}^*\tilde{\mathbf{S}}^*(\tilde{\mathbf{U}}^*)^T = \mathbf{W}$. Since $\mathbf{W} \in \mathcal{W}_{\frac{1}{\sqrt{2}}}$ and $\tilde{\mathbf{U}}^*_B \succ 0$, $\tilde{\mathbf{U}}^*_N \succ 0$ (due to Lemma 4.1.2), the assumptions of the Corollary 4.2.1 are satisfied. Threfore $\triangle\tilde{\mathbf{U}} = 0, \triangle\tilde{\mathbf{X}} = 0, \triangle\tilde{\mathbf{S}} = 0$. Assumption (A1) yields $\triangle\tilde{y}$.

□

Symmetrization $\mathbf{L_X}^T\mathbf{S}\mathbf{L_X}$

Consider the condition

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

from (3.1), which can be rewritten equivalently as the pair

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

Consider the normalized matrices $\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{S}}(\rho)$ defined in (4.46), or (4.13), which are associated to the weighted path functions $\mathbf{X}(\mu), \mathbf{S}(\mu)$ with a fixed weight $\mathbf{W}$ (from $\mathcal{W}_{\frac{1}{\sqrt{2}}}$ or D^n_{++}). Define

$$\tilde{\mathbf{L}}(\rho) := \begin{pmatrix} \mathbf{L}_B(\rho^2) & 0 \\ \mathbf{L}_V(\rho^2)^T/\rho & \mathbf{L}_N(\rho^2)/\rho \end{pmatrix}$$

Lemma 4.2.7 *The systems*

$$\left.\begin{aligned} \mathbf{L}(\mu)^T\mathbf{S}(\mu)\mathbf{L}(\mu) &= \mu\mathbf{W} \\ \mathbf{L}(\mu)\mathbf{L}(\mu)^T &= \mathbf{X}(\mu) \end{aligned}\right\}$$

and

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

are equivalent (for $\rho = \sqrt{\mu} > 0$ *sufficiently small).*

Proof follows from simple computation: we can rewrite the equality

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

as

$$\mathbf{L}(\rho^2)^T \mathbf{S}(\rho^2) \mathbf{L}(\rho^2) = \rho^2 \mathbf{W}.$$

However, by definition, the left-hand side is equal to

$$\begin{bmatrix} \tilde{\mathbf{L}}_B(\rho)^T & \rho\tilde{\mathbf{L}}_V(\rho) \\ 0 & \rho\tilde{\mathbf{L}}_N(\rho)^T \end{bmatrix} \begin{bmatrix} \rho^2\tilde{\mathbf{S}}_B(\rho) & \rho\tilde{\mathbf{S}}_V(\rho) \\ \rho\tilde{\mathbf{S}}_V(\rho) & \tilde{\mathbf{S}}_N(\rho) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{L}}_B(\rho) & 0 \\ \rho\tilde{\mathbf{L}}_V(\rho)^T & \rho\tilde{\mathbf{L}}_N(\rho) \end{bmatrix} =$$

$$= \rho^2 \tilde{\mathbf{L}}(\rho)^T \tilde{\mathbf{S}}(\rho) \tilde{\mathbf{L}}(\rho).$$

Similarly, the equality

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

can be rewritten as

$$\mathbf{L}(\rho^2)\mathbf{L}(\rho^2)^T = \mathbf{X}(\rho^2),$$

or

$$\begin{bmatrix} \tilde{\mathbf{L}}_B(\rho) & 0 \\ \rho\tilde{\mathbf{L}}_V(\rho)^T & \rho\tilde{\mathbf{L}}_N(\rho) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{L}}_B(\rho)^T & \rho\tilde{\mathbf{L}}_V(\rho) \\ 0 & \rho\tilde{\mathbf{L}}_N(\rho)^T \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}_B(\rho) & \rho\tilde{\mathbf{X}}_V(\rho) \\ \rho\tilde{\mathbf{X}}_V^T(\rho) & \rho^2\tilde{\mathbf{X}}_N(\rho) \end{bmatrix},$$

which is equivalent to $\tilde{\mathbf{L}}(\rho)\tilde{\mathbf{L}}(\rho)^T = \tilde{\mathbf{X}}(\rho)$.

□

From Proposition 4.1.2 and Proposition 4.1.4 it follows that for any sequence $\rho_k \to 0$ is

$$(\tilde{\mathbf{X}}(\rho_k), \tilde{\mathbf{L}}(\rho_k), \tilde{y}(\rho_k), \tilde{\mathbf{S}}(\rho_k))$$

bounded, so we may assume that the limit

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

exists. By taking the limits $\rho_k \to 0$ in the equalities

$$\tilde{\mathbf{L}}(\rho_k)^T \tilde{\mathbf{S}}(\rho_k) \tilde{\mathbf{L}}(\rho_k) = \mathbf{W}$$

$$\tilde{\mathbf{L}}(\rho_k)^T \tilde{\mathbf{L}}(\rho_k) = \tilde{\mathbf{X}}(\rho_k)$$

we obtain that

$$\begin{array}{c} (\tilde{\mathbf{L}}^*)^T \tilde{\mathbf{S}}^* \tilde{\mathbf{L}}^* = \mathbf{W} \\ \tilde{\mathbf{L}}^* (\tilde{\mathbf{L}}^*)^T = \tilde{\mathbf{X}}^*. \end{array} \tag{4.56}$$

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

Define the map $\tilde{F}^3$ 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 Ψ is defined by (4.48). Obviously 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

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

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

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

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

Lemma 4.2.8 $D\tilde{F}^3(\tilde{\mathbf{X}}^*, \tilde{\mathbf{L}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ *is a nonsingular linear map.*

Proof. 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 (4.49) we have that $\triangle\tilde{\mathbf{X}} \bullet \triangle\tilde{\mathbf{S}} = 0$. Due to (4.56), the case $\mathbf{W} \in D^n_{++}$ follows from Proposition 3.2.4 and the case $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ follows from Corollary 3.2.2.

□

4.2.4 Analyticity of weighted path as a function of $\sqrt{\mu}$ at $\mu = 0$

The aim of this section is to prove the following proposition.

Proposition 4.2.2 *Let $j \in \{1, 2, 3\}$. 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. Assume j=1. Recall that $\tilde{F}^1 : S^n \times R^m \times S^n \times R \to R^m \times S^n \times S^n$ is an analytic function of $(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}}, \rho)$ such that

1. There exists $(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$ such that $\tilde{F}^1(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0) = 0$,
2. The Fréchet derivative of the map $\tilde{F}^1$ with respect to $(\tilde{\mathbf{X}}, \tilde{y}, \tilde{\mathbf{S}})$ is nonsingular in $(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*, 0)$—see Lemma 4.2.4.

Now we can apply the implicit function theorem to obtain that there exists a neighborhood $\mathcal{I}$ of $\rho = 0$, a neighborhood $\mathcal{U}$ of $(\tilde{\mathbf{X}}^*, \tilde{y}^*, \tilde{\mathbf{S}}^*)$ and an analytic function

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

such that

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

and

$$\tilde{F}^1((\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\rho), \rho) = 0 \quad \forall \rho \in \mathcal{I}. \tag{4.57}$$

There exists $\bar{k} > 0$ such that for all $k \geq \bar{k}$: $\rho_k \in \mathcal{I}$, $(\tilde{\mathbf{X}}(\rho_k), \tilde{y}(\rho_k), \tilde{\mathbf{S}}(\rho_k)) \in \mathcal{U}$ and therefore

$$(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\rho_k) = (\tilde{\mathbf{X}}(\rho_k), \tilde{y}(\rho_k), \tilde{\mathbf{S}}(\rho_k)) \qquad \forall k \geq \bar{k}.$$

However, sice $(\tilde{\mathbf{X}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ and $(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\rho)$ are solutions of (4.57) for $\rho > 0$, we have that

$$(\tilde{\mathbf{X}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho)) = (\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\rho) \qquad \forall \rho \in \mathcal{I} \cap (0, \infty)$$

by the uniqueness of positive definite solutions. Thus the function $(\tilde{\mathbf{X}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ is analytically extendable to $\rho = 0$ by prescription

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

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

Assume $j = 2$ and consider the map $\tilde{F}^2$. Using similar arguments as in the case $j = 1$ and Lemma 4.2.6 it can be shown that the function $(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{U}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ can be analytically extended to $\rho = 0$ by prescription

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

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

Let $j = 3$ and consider the map $\tilde{F}^3$. Similarly as in the case of the map $\tilde{F}^1$, using Lemma 4.2.8 it can be shown that the function $(\tilde{\mathbf{X}}(\rho), \tilde{\mathbf{L}}(\rho), \tilde{y}(\rho), \tilde{\mathbf{S}}(\rho))$ can be analytically extended to $\rho = 0$ by prescription

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

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

□

4.2.5 Introduction of new normalized matrices and transformation of feasibility conditions

Lemma 4.2.9 *Let $j \in \{1,2,3\}$ be arbitrary and assume $\mathbf{W}_V = 0$ if $j \neq 1$. 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 ρ 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

$$\mathbf{X}_V(\mu) = \rho \tilde{\mathbf{X}}_V(\rho) = \rho \mathcal{O}(\rho) = \mathcal{O}(\rho^2) = \mathcal{O}(\mu). \tag{4.58}$$

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 4.1.6 ($\mathbf{Y}_B(\mu) = \Theta(1)$) 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 (4.58) and Proposition 4.1.3.

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 $\mathbf{W}_V = 0$ whenever considering the case $j = 2$ or $j = 3$.

From Lemma 4.2.9 it follows that the path matrices posses the following asymptotic behavior:

$$\mathbf{X}(\mu) = \begin{pmatrix} \Theta(1) & \mathcal{O}(\mu) \\ \mathcal{O}(\mu) & \Theta(\mu) \end{pmatrix}, \quad \mathbf{S}(\mu) = \begin{pmatrix} \Theta(\mu) & \mathcal{O}(\mu) \\ \mathcal{O}(\mu) & \Theta(1) \end{pmatrix}. \tag{4.59}$$

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

$$\mathbf{L}(\mu) = \begin{pmatrix} \Theta(1) & 0 \\ \mathcal{O}(\mu) & \Theta(\sqrt{\mu}) \end{pmatrix}, \quad \mathbf{Y}(\mu) = \begin{pmatrix} \Theta(1) & \mathcal{O}(\mu) \\ \mathcal{O}(\mu) & \Theta(\sqrt{\mu}) \end{pmatrix}. \tag{4.60}$$

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

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

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 (4.39) using the new normalized matrices:

$$\begin{array}{rcl} \tilde{\mathbb{P}}_1\bar{v} + \mu\tilde{\mathbb{Q}}_1\bar{w} + \mu\tilde{\mathbb{R}}_1\bar{z} &=& \tilde{d}_1 + \mu\triangle\tilde{d}_1, \\ \mu\tilde{\mathbb{Q}}_2\bar{w} + \mu\tilde{\mathbb{R}}_2\bar{z} &=& \tilde{d}_2 + \mu\triangle\tilde{d}_2, \\ \mu\tilde{\mathbb{R}}_3\bar{z} &=& \tilde{d}_3 + \mu\triangle\tilde{d}_3. \end{array} \tag{4.61}$$

From the asymptotic behavior given in (4.59) 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 (4.61) and letting $k \to \infty$ we obtain that $\tilde{d}_2 = 0$ and $\tilde{d}_3 = 0$. Hence the system (4.61) has the form

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

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

$$\bar{\Psi}(\bar{\mathbf{X}},\bar{y},\bar{\mathbf{S}},\mu)=\begin{bmatrix}\tilde{\mathbb{P}}_1\bar{v}+\mu\tilde{\mathbb{Q}}_1\bar{w}+\mu\tilde{\mathbb{R}}_1\bar{z}-\tilde{d}_1+\mu\triangle\tilde{d}_1\\ \tilde{\mathbb{Q}}_2\bar{w}+\tilde{\mathbb{R}}_2\bar{z}-\triangle\tilde{d}_2\\ \tilde{\mathbb{R}}_3\bar{z}-\triangle\tilde{d}_3\end{bmatrix}. \tag{4.62}$$

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

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

4.2.6 Introduction of normalized system and nonsingularity of Fréchet derivatives (II)

Symmetrization $(\mathbf{XS}+\mathbf{SX})/2$

Consider the last condition $\mathbf{XS}+\mathbf{SX}=2\mu\mathbf{W}$ in the system (3.1), for which the block representation was given in (4.50). Because $(\mathbf{X}(\mu),y(\mu),\mathbf{S}(\mu))$ satisfies the system (4.50) for $\mu>0$ (sufficiently small), we have that the triple $(\bar{\mathbf{X}}(\mu),\bar{y}(\mu),\bar{\mathbf{S}}(\mu))$ satisfies

$$\begin{array}{rcl}\bar{\mathbf{X}}_B(\mu)\bar{\mathbf{S}}_B(\mu)+\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{X}}_B(\mu)+\mu\bar{\mathbf{X}}_V(\mu)\bar{\mathbf{S}}_V(\mu)^T+\mu\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{X}}_V(\mu)^T&=&2\mathbf{W}_B,\\ \bar{\mathbf{X}}_B(\mu)\bar{\mathbf{S}}_V(\mu)+\mu\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{X}}_V(\mu)+\bar{\mathbf{X}}_V(\mu)\bar{\mathbf{S}}_N(\mu)+\mu\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{X}}_N(\mu)&=&2\mathbf{W}_V,\\ \bar{\mathbf{X}}_N(\mu)\bar{\mathbf{S}}_N(\mu)+\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{X}}_N(\mu)+\mu\bar{\mathbf{X}}_V(\mu)^T\bar{\mathbf{S}}_V(\mu)+\mu\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{X}}_V(\mu)&=&2\mathbf{W}_N.\end{array} \tag{4.63}$$

If we put $\mu:=\mu_k$ and take the limit $k\to\infty$ in the equations above we find that

$$\begin{array}{l}\bar{\mathbf{X}}_B^*\bar{\mathbf{S}}_B^*+\bar{\mathbf{S}}_B^*\bar{\mathbf{X}}_B^*=2\mathbf{W}_B\succ 0\\ \bar{\mathbf{X}}_B^*\bar{\mathbf{S}}_V^*+\bar{\mathbf{X}}_V^*\bar{\mathbf{S}}_N^*=2\mathbf{W}_V\\ \bar{\mathbf{X}}_N^*\bar{\mathbf{S}}_N^*+\bar{\mathbf{S}}_N^*\bar{\mathbf{X}}_N^*=2\mathbf{W}_N\succ 0\end{array} \tag{4.64}$$

Define the map $\bar{F}^1$ in the following way:

$$\bar{F}^1(\bar{\mathbf{X}},\bar{y},\bar{\mathbf{S}},\mu)=\begin{bmatrix}\bar{\Psi}(\bar{\mathbf{X}},\bar{y},\bar{\mathbf{S}},\mu)\\ \bar{\mathbf{X}}_B\bar{\mathbf{S}}_B+\bar{\mathbf{S}}_B\bar{\mathbf{X}}_B+\mu\bar{\mathbf{X}}_V\bar{\mathbf{S}}_V^T+\mu\bar{\mathbf{S}}_V\bar{\mathbf{X}}_V^T-2\mathbf{W}_B\\ \bar{\mathbf{X}}_B\bar{\mathbf{S}}_V+\mu\bar{\mathbf{S}}_B\bar{\mathbf{X}}_V+\bar{\mathbf{X}}_V\bar{\mathbf{S}}_N+\mu\bar{\mathbf{S}}_V\bar{\mathbf{X}}_N-2\mathbf{W}_V\\ \bar{\mathbf{X}}_N\bar{\mathbf{S}}_N+\bar{\mathbf{S}}_N\bar{\mathbf{X}}_N+\mu\bar{\mathbf{X}}_V^T\bar{\mathbf{S}}_V+\mu\bar{\mathbf{S}}_V^T\bar{\mathbf{X}}_V-2\mathbf{W}_N\end{bmatrix},$$

where $\bar{\Psi}$ is defined (4.62). Obviously, for $\mu > 0$ (sufficiently small)

$$\bar{F}^1(\bar{\mathbf{X}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu), \mu) = 0, \tag{4.65}$$

and also

$$\bar{F}^1(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0) = 0. \tag{4.66}$$

The Fréchet derivative of $\bar{F}^1$ with respect to $(\bar{\mathbf{X}}, \bar{y}, \bar{\mathbf{S}})$ at the point $(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)$ is the linear map given by

$$D\bar{F}^1(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = \begin{bmatrix} D\bar{\Psi}(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] \\ \triangle\bar{\mathbf{X}}_B\bar{\mathbf{S}}^*_B + \triangle\bar{\mathbf{S}}_B\bar{\mathbf{X}}^*_B + \bar{\mathbf{X}}^*_B\triangle\bar{\mathbf{S}}_B + \bar{\mathbf{S}}^*_B\triangle\bar{\mathbf{X}}_B \\ \triangle\bar{\mathbf{X}}_B\bar{\mathbf{S}}^*_V + \triangle\bar{\mathbf{X}}_V\bar{\mathbf{S}}^*_N + \bar{\mathbf{X}}^*_B\triangle\bar{\mathbf{S}}_V + \bar{\mathbf{X}}^*_V\triangle\bar{\mathbf{S}}_N \\ \triangle\bar{\mathbf{X}}_N\bar{\mathbf{S}}^*_N + \triangle\bar{\mathbf{S}}_N\bar{\mathbf{X}}^*_N + \bar{\mathbf{X}}^*_N\triangle\bar{\mathbf{S}}_N + \bar{\mathbf{S}}^*_N\triangle\bar{\mathbf{X}}_N \end{bmatrix}.$$

Lemma 4.2.10 *$D\bar{F}^1(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)$ is a nonsingular linear map.*

Proof. Assume

$$D\bar{F}^1(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = 0. \tag{4.67}$$

From the first equation of (4.67) and Proposition 4.2.1 (b) it follows that $\triangle\bar{\mathbf{X}}_B \bullet \triangle\bar{\mathbf{S}}_B = 0$ and $\triangle\bar{\mathbf{X}}_N \bullet \triangle\bar{\mathbf{S}}_N = 0$. Using similar arguments as in the proof of Lemma 4.2.4 it can be shown that $\triangle\bar{\mathbf{X}}_B = 0$, $\triangle\bar{\mathbf{X}}_N = 0$, $\triangle\bar{\mathbf{S}}_B = 0$, $\triangle\bar{\mathbf{S}}_N = 0$. This fact together with the third equation in (4.67) yield

$$\triangle\bar{\mathbf{X}}_V\bar{\mathbf{S}}^*_N + \bar{\mathbf{X}}^*_B\triangle\bar{\mathbf{S}}_V = 0. \tag{4.68}$$

Moreover, from Proposition 4.2.1 (b) we now have that $\triangle\bar{\mathbf{X}}_V \bullet \triangle\bar{\mathbf{S}}_V = 0$. Similarly as in the proof of Lemma 4.2.4 it can be shown that this fact together with (4.68) imply $\triangle\bar{\mathbf{X}}_V = 0$, $\triangle\bar{\mathbf{S}}_V = 0$. Finally, Assumption (A1) yields $\triangle\bar{y} = 0$.

□

Symmetrization $\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}}$

Consider the condition

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

again, or equivalently

$$\mathbf{Y}\mathbf{S}\mathbf{Y} = \mu\mathbf{W},$$
$$\mathbf{Y}^2 = \mathbf{X},$$

which can be rewritten in the following block form:

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

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

$$\begin{gathered}
\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{Y}}_B(\mu)+\mu\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{Y}}_B(\mu)\\
+\mu\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{Y}}_V(\mu)^T+\mu\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{Y}}_V(\mu)^T=\mathbf{W}_B,\\
\sqrt{\mu}\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{Y}}_V(\mu)+\mu\sqrt{\mu}\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{Y}}_V(\mu)+\\
\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{Y}}_N(\mu)+\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{Y}}_N(\mu)=0,\\
\mu^2\bar{\mathbf{Y}}_V(\mu)^T\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{Y}}_V(\mu)+\mu\bar{\mathbf{Y}}_N(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{Y}}_V(\mu)+\\
\mu\bar{\mathbf{Y}}_V(\mu)^T\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{Y}}_N(\mu)+\bar{\mathbf{Y}}_N(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{Y}}_N(\mu)=\mathbf{W}_N,\\
\bar{\mathbf{Y}}_B(\mu)^2+\mu^2\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{Y}}_V(\mu)^T=\bar{\mathbf{X}}_B(\mu),\\
\bar{\mathbf{Y}}_B(\mu)\bar{\mathbf{Y}}_V(\mu)+\sqrt{\mu}\bar{\mathbf{Y}}_V(\mu)\bar{\mathbf{Y}}_N(\mu)=\bar{\mathbf{X}}_V(\mu),\\
\bar{\mathbf{Y}}_N(\mu)^2+\mu\bar{\mathbf{Y}}_V(\mu)^T\bar{\mathbf{Y}}_V(\mu)=\bar{\mathbf{X}}_N(\mu),
\end{gathered} \tag{4.69}$$

From (4.59) and (4.60) 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

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

where $\bar{\Psi}$ is defined in (4.62). For $\mu>0$ sufficiently small it holds that

$$\bar{F}^2(\bar{\mathbf{X}}(\mu),\bar{\mathbf{Y}}(\mu),\bar{y}(\mu),\bar{\mathbf{S}}(\mu),\mu)=0 \tag{4.70}$$

and

$$\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0) = 0. \tag{4.71}$$

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 4.2.11 $D\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{Y}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}]$ *is a nonsingular linear map.*

Proof. Assume

$$D\bar{F}^2(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{Y}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = 0. \tag{4.72}$$

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 (4.69), then by taking the limit $k \to \infty$ we obtain that

$$\bar{\mathbf{Y}}_B^*\bar{\mathbf{S}}_B^*\bar{\mathbf{Y}}_B^* = \mathbf{W}_B, \qquad \bar{\mathbf{Y}}_N^*\bar{\mathbf{S}}_N^*\bar{\mathbf{Y}}_N^* = \mathbf{W}_N. \tag{4.73}$$

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 (4.72) and Proposition 4.2.1 (b) we have that

$$\triangle\bar{\mathbf{X}}_B \bullet \triangle\bar{\mathbf{S}}_B = 0, \quad \triangle\bar{\mathbf{X}}_N \bullet \triangle\bar{\mathbf{S}}_N = 0. \tag{4.74}$$

We can apply Corollary 3.2.1 and obtain 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 (4.72) 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 4.2.1 (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 Proposition A.2.2, Proposition A.1.5 (c) and positive definitnes of $\bar{\mathbf{S}}_N^*$ it follows $\triangle\bar{\mathbf{Y}}_V = 0$. Therefore also $\triangle\bar{\mathbf{X}}_V = \triangle\bar{\mathbf{S}}_V = 0$. Assumption (A1) gives $\triangle\bar{y} = 0$.

□

Symmetrization $\mathbf{L_X S L_X}$

Consider the condition

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

again, or equivalently

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

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

$$\begin{gathered}\bar{\mathbf{L}}_B(\mu)^T\bar{\mathbf{S}}_B(\mu)\bar{\mathbf{L}}_B(\mu) + \mu\bar{\mathbf{L}}_V(\mu)\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{L}}_B(\mu) +\\ \mu\bar{\mathbf{L}}_B(\mu)^T\bar{\mathbf{S}}_V(\mu)^T\bar{\mathbf{L}}_V(\mu)^T + \mu\bar{\mathbf{L}}_V(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{L}}_V(\mu)^T = \mathbf{W}_B,\\ \bar{\mathbf{L}}_B(\mu)^T\bar{\mathbf{S}}_V(\mu)\bar{\mathbf{L}}_N(\mu) + \bar{\mathbf{L}}_V(\mu)\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{L}}_N(\mu) = 0,\\ \bar{\mathbf{L}}_N(\mu)^T\bar{\mathbf{S}}_N(\mu)\bar{\mathbf{L}}_N(\mu) = \mu\mathbf{W}_N,\\ \bar{\mathbf{L}}_B(\mu)\bar{\mathbf{L}}_B(\mu)^T = \mathbf{X}_B(\mu),\\ \bar{\mathbf{L}}_B(\mu)\bar{\mathbf{L}}_V(\mu) = \mathbf{X}_V(\mu),\\ \bar{\mathbf{L}}_N(\mu)\bar{\mathbf{L}}_N(\mu)^T + \mu\bar{\mathbf{L}}_V(\mu)^T\bar{\mathbf{L}}_V(\mu) = \mathbf{X}_N(\mu).\end{gathered} \tag{4.75}$$

From (4.59) and (4.59) 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 in Section 4.2.5. For $\mu > 0$ sufficiently small it holds, that

$$\bar{F}^3(\bar{\mathbf{X}}(\mu), \bar{\mathbf{L}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu), \mu) = 0 \tag{4.76}$$

and

$$\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0) = 0. \tag{4.77}$$

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 4.2.12 $D\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{L}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}]$ *is a nonsingular linear map.*

Proof. Assume

$$D\bar{F}^3(\bar{\mathbf{X}}^*, \bar{\mathbf{L}}^*, \bar{y}^*, \bar{\mathbf{S}}^*, 0)[\triangle\bar{\mathbf{X}}, \triangle\bar{\mathbf{L}}, \triangle\bar{y}, \triangle\bar{\mathbf{S}}] = 0. \tag{4.78}$$

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 (4.75), then by taking the limit $k \to \infty$ we obtain that

$$(\bar{\mathbf{L}}_B^*)^T \bar{\mathbf{S}}_B^* \bar{\mathbf{L}}_B = \mathbf{W}_B, \qquad (\bar{\mathbf{L}}_N^*)^T \bar{\mathbf{S}}_N^* \bar{\mathbf{L}}_N = \mathbf{W}_N. \tag{4.79}$$

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 (4.78) and Proposition 4.2.1 (b) we have that

$$\triangle\bar{\mathbf{X}}_B \bullet \triangle\bar{\mathbf{S}}_B = 0, \quad \triangle\bar{\mathbf{X}}_N \bullet \triangle\bar{\mathbf{S}}_N = 0. \tag{4.80}$$

By applying Corollary 3.2.2 in the case of $\mathbf{W} \in \mathcal{M}_{\frac{1}{\sqrt{2}}}$ or Proposition 3.2.4 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 (4.78) 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 4.2.1 (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 $\bar{\mathbf{S}}^*_N \succ 0$ and because of Proposition A.2.2 we have that $\triangle\bar{\mathbf{L}}_V = 0$. Therefore also $\triangle\bar{\mathbf{X}}_V = \triangle\bar{\mathbf{S}}_V = 0$. Assumption (A1) gives $\triangle\bar{y} = 0$.

□

4.2.7 Analyticity of weighted path as a function of μ at $\mu = 0$

Proposition 4.2.3 [3] *Let $j = 1$. Then the weighted path $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ is an analytic function of μ for all $\mu \geq 0$ (sufficiently small).*

Proof. From (4.65), (4.66) and Lemma 4.2.10 it follows that implicit function theorem can be applied: there exists a neighborhood $\mathcal{I}$ of $\mu = 0$, a neighborhood $\mathcal{U}$ of $(\bar{\mathbf{X}}^*, \bar{y}^*, \bar{\mathbf{S}}^*)$ and an analytic function $(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}}) : \mathcal{I} \to \mathcal{U}$ such that

$$(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(0) = (\bar{\mathbf{X}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$$

and

$$\bar{F}^1((\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\mu), \mu) = 0 \quad \forall \mu \in \mathcal{I}. \tag{4.81}$$

There exists $\bar{k} > 0$ such that for all $k \geq \bar{k}$: $\mu_k \in \mathcal{I}$, $(\bar{\mathbf{X}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k)) \in \mathcal{U}$ and therefore

$$(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\mu_k) = (\bar{\mathbf{X}}(\mu_k), \bar{y}(\mu_k), \bar{\mathbf{S}}(\mu_k)) \qquad \forall k \geq \bar{k}.$$

[3] This result was obtained by Preiss and Stoer [59] for weighted paths in linear complementarity problems.

However, sice $(\bar{\mathbf{X}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ and $(\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\mu)$ are solutions of (4.81) for $\mu > 0$, we have that

$$(\bar{\mathbf{X}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu)) = (\hat{\mathbf{X}}, \hat{y}, \hat{\mathbf{S}})(\mu) \qquad \forall \mu \in \mathcal{I} \cap (0, \infty)$$

by the uniqueness of positive definite solutions. Thus the function $(\bar{\mathbf{X}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ is analytically extendable to $\mu = 0$ by prescription

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

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

□

Lemma 4.2.13 *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

$$\frac{d\mathbf{X}_V(\mu)}{d\mu} = \frac{d[\rho\tilde{\mathbf{X}}_V(\rho)]}{d\mu} = \frac{d\rho}{d\mu}\tilde{\mathbf{X}}_V(\rho) + \rho\frac{d\tilde{\mathbf{X}}_V(\rho)}{d\mu} = \frac{1}{2\sqrt{\mu}}\left[\tilde{\mathbf{X}}_V(\rho) + \rho\frac{d\tilde{\mathbf{X}}_V(\rho)}{d\rho}\right]. \quad (4.82)$$

$d\tilde{\mathbf{X}}_V(\rho)$ is an analytic function of ρ at $\rho = 0$ (see the proof of Proposition 4.2.2). Therefore $\frac{d\tilde{\mathbf{X}}_V(\rho)}{d\rho}$ is bounded as $\rho \to 0$ and

$$\lim_{\rho \to 0} \rho \frac{d\tilde{\mathbf{X}}_V(\rho)}{d\rho} = 0.$$

Hence if $\frac{d\mathbf{X}_V(\mu)}{d\mu}$ was bounded (as $\mu \to 0$), then from (4.82) we have that

$$\lim_{\rho \to 0} \tilde{\mathbf{X}}_V(\rho) = 0.$$

But this implies $\mathbf{W}_V = 0$ (see Proposition 4.1.9 for $j = 2$ or Proposition 4.1.10 if $j = 3$), which contradicts to the assumption. The statement for $\frac{d\mathbf{S}_V(\mu)}{d\mu}$ can be proved similarly.

□

Proposition 4.2.4 *Let $j \in \{2, 3\}$. Then the associated weighted path $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ is an analytic function of μ for all $\mu \geq 0$ if and only if $\mathbf{W}_V = 0$.*

Proof. Assume $\mathbf{W}_V = 0$. Using (4.70), (4.71), Lemma 4.2.11 and similar arguments as in the case of the symmetrization Φ_1, it can be shown, that the function $(\bar{\mathbf{X}}(\mu), \bar{\mathbf{Y}}(\mu), \bar{y}(\mu), \bar{\mathbf{S}}(\mu))$ can be analytically extended to $\mu = 0$ by prescription

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

Therefore also the path function $(\mathbf{X}(\mu), y(\mu), \mathbf{S}(\mu))$ associated with the symmetrization Φ_2 is analytically extendable to $\mu = 0$. Similarly, using (4.76), (4.77), Lemma 4.2.12 and similar arguments as in the case of the symmetrization Φ_1, it can be shown, 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 Φ_3 is analytically extendable to $\mu = 0$. The reverse implication follows from Lemma 4.2.13 and properties of analytic functions.

$\square$

Chapter 5

Conclusion

In this thesis the weighted interior point paths in semidefinite programming were studied. In the concrete, we focused on the existence, the asymptotic behavior and the analyticity of the weighted paths at the boundary point. The main results included in this thesis could be summarized as follows.

We presented a new and relatively simple proof of the existence of weighted central paths associated with certain type of symmetrization map. These types of weighted paths were studied for the nonlinear complementarity problems [51] where the result was proved using the theory of local homeomorphic maps. In the work [58] the weighted path for linear complementarity problems was studied, however only the one, associated with the AHO-symmetrization. The weighted path in SDP associated with the Cholesky-type-symmetrization and positive diagonal weight was studied in the paper [7] and the existence was shown by defining weighted logarithmic barrier functions. The proof presented in the thesis is based on generalization of the result of [58] to all five symmetrization maps and the assumptions for the existence are formulated in terms of semidefinite programming. The main existence results are stated in Theorem 3.6.1 and Corollary 3.6.3. It seems that this technique can be also used for proving the existence of the weighted paths associated with the Cholesky symmetrization and weights included in the more general set of weights defined in [9].

The second part consists of the results concerning the asymptotic behavior of the weighted paths. These results were obtained under the assumption of the existence of the strict complementary optimal solution and can be considered as the generalization of the results of [59] (for the AHO-symmetrization) to all types of symmetrization maps. Moreover the square root and the Cholesky factors of the weighted paths were studied. The results concerning the $\mathcal{O}$-notation and Θ-notation are stated in Proposition 4.1.2, Proposition 4.1.3, Proposition 4.1.4, Proposition 4.1.5, Proposition 4.1.6 and Proposition

4.1.7. Also the asymptotic behavior in o-notation was studied and it was shown that this behavior depends not only on the symmetrization type but also on the type of the weight matrix. Related results are stated in Proposition 4.1.9 and Proposition 4.1.10. Summarization of the asymptotic behavior is given in Section 4.1.4. All properties included in this part were useful for the analysis of the interior point algorithms and analyticity of the weighted paths at the boundary point.

Finally, the analyticity of the weighted paths at the boundary point was analyzed in this thesis. We followed the known results from this area: in the papers [59], [41] it was shown that the weighted central path associated with AHO-symmetrization is an analytic function of μ at $\mu = 0$. The authors of [42] proved that the weighted path associated with the square-root-type symmetrization is analytic at $\mu = 0$ as a function of $\sqrt{\mu}$. Finally, in the work [8] it was shown that the weighted path associated with the Cholesky-type symmetrization and positive diagonal weight is an analytic function of μ at $\mu = 0$. In this thesis the weighted path associated with Cholesky-type symmetrization and a suitable symmetric positive definite weight was studied and it was proved that this path is analytic at $\mu = 0$ as a function of $\sqrt{\mu}$ (Propostion 4.2.2). As a consequence we obtained that if the weight matrix is block diagonal, then the off-diagonal blocks of the weighted path posses "better" asymptotic behavior in $\mathcal{O}$-notation and it was shown that the same result holds for the path associated with square-root-type symmetrization (Lemma 4.2.9). Moreover, it was shown that the weighted paths (associated with the both, the square-root-type and Cholesky-type symmetrization) are analytic functions of μ (at the boundary point) if and only if the weight matrix is block diagonal (Proposition 4.2.4). These results could be useful for the error bound and superlinear convergence analysis of the interior point algorithms for SDP.

Appendix A

Useful facts from matrix theory

This appendix includes the properties of matrices that were needed in the previous parts of this thesis. In the section A.1 we summarize the basic properties of real symmetric and positive semidefinite matrices. Most of the statements can be found e.g. in [32], [81], [18]. In the section A.2 we give some properties of the trace operator and two special types of matrix norms (the Frobenius and the spectral norm). We mainly focus on the properties concerning positive semidefinite matrices and include some special properties taken from [59], [51]. More about general properties of matrix norms can be found in [32]. The section A.3 includes basic properties of the Schur complement. More can be found e.g. in [18, 81]. In the section A.4 the definition and the review of the properties of the symmetric Kronecker product is given. For more, see e.g. [33], [81] and [34] (here one can find also an overview of its history and applications and complete proofs of the mentioned properties). In the section A.5 we give some simple, however useful properties of triangular matrices.

A.1 Symmetric and positive semidefinite matrices

We will denote S^n the vector space of all $n \times n$ real symmetric matrices. Obviously

$$\dim S^n = \frac{n(n+1)}{2}.$$

Theorem A.1.1 *(Spectral Decomposition) Let $\mathbf{A} \in S^n$. Then there exists an orthogonal matrix $\mathbf{Q} \in R^{n\times n}$ ($\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$) and a (real) diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$ and the diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$.*

The product of two symmetric matrices is not symmetric in general. For instance, take

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathbf{AB} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

Therefore, $\mathbf{AB} \neq \mathbf{BA}$ in general for $\mathbf{A}, \mathbf{B} \in S^n$. The following theorem is a consequence of Theorem 1.3.2 of [32] and gives a necessary and sufficient condition for $\mathbf{AB} = \mathbf{BA}$.

Theorem A.1.2 *Let $\mathbf{A}, \mathbf{B} \in S^n$. Then $\mathbf{AB} = \mathbf{BA}$ if and only if they are simultaneously diagonalizable, that is, there exists an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}^T\mathbf{AQ} = \mathbf{D}_A$ and $\mathbf{Q}^T\mathbf{BQ} = \mathbf{D}_B$.*

Proof. ($\Leftarrow$) The proof is straightforward.
($\Rightarrow$) Theorem A.1.1 implies that the matrices $\mathbf{A}, \mathbf{B}$ are orthogonally diagonalizable—there exist orthogonal matrices $\mathbf{U}, \mathbf{V}$ such that $\mathbf{A} = \mathbf{U}\mathbf{D}_\mathbf{A}\mathbf{U}^T$, $\mathbf{B} = \mathbf{V}\mathbf{D}_\mathbf{B}\mathbf{V}^T$. Therefore

$$\mathbf{AB} = \mathbf{BA} \;\Leftrightarrow\; \mathbf{U}\mathbf{D}_\mathbf{A}\mathbf{U}^T\mathbf{V}\mathbf{D}_\mathbf{B}\mathbf{V}^T = \mathbf{V}\mathbf{D}_\mathbf{B}\mathbf{V}^T\mathbf{U}\mathbf{D}_\mathbf{A}\mathbf{U}^T \;\Leftrightarrow$$

$$\Leftrightarrow\; \mathbf{D}_\mathbf{A}(\mathbf{U}^T\mathbf{V}\mathbf{D}_\mathbf{B}\mathbf{V}^T\mathbf{U}) = (\mathbf{U}^T\mathbf{V}\mathbf{D}_\mathbf{B}\mathbf{V}^T\mathbf{U})\mathbf{D}_\mathbf{A}$$

and, without loss of generality, we may assume that $\mathbf{A}$ is diagonal.

Let $\mathbf{A} = diag(A_{11}, \ldots, A_{nn})$ and $\mathbf{B} = [B_{ij}]$. Let $\alpha_1, \ldots, \alpha_k$ be such that $\alpha_1 < \ldots < \alpha_k$ and $\mathbf{A}$ can be written as

$$\mathbf{A} = \begin{pmatrix} \alpha_1\mathbf{I} & 0 & 0 & 0 \\ 0 & \alpha_2\mathbf{I} & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & \alpha_k\mathbf{I} \end{pmatrix}.$$

The equality $\mathbf{AB} = \mathbf{BA}$ then implies that $A_iB_{ij} = B_{ij}A_j$. Therefore $B_{ij} = 0$ whenever $A_{ii} \neq A_{jj}$ and the matrix $\mathbf{B}$ is a block diagonal:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 & 0 & 0 & 0 \\ 0 & \mathbf{B}_2 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & \mathbf{B}_k \end{pmatrix}.$$

Each block $\mathbf{B}_i$ corresponds to the block $\alpha_i\mathbf{I}$ and is a symmetric matrix. From Theorem A.1.1 we have that there exist orthogonal matrices $\mathbf{Q}_i$, such that $\mathbf{Q}_i^T\mathbf{B}_i\mathbf{Q}_i = \mathbf{D}_i$.

Finally, it is easy to see that the matrices $\mathbf{A}, \mathbf{B}$ are simultaneously diagonalizable by the matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1 & 0 & 0 & 0 \\ 0 & \mathbf{Q}_2 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & \mathbf{Q}_k \end{pmatrix}.$$

$\square$

Definition A.1.1 *The matrix $\mathbf{A} \in S^n$ is called positive definite ($\mathbf{A} \in S^n_{++}$ or $\mathbf{A} \succ 0$) if for all $x \in R^n$, $x \neq 0$ it holds*

$$x^T \mathbf{A} x > 0.$$

The matrix $\mathbf{A} \in S^n$ is called positive semidefinite ($\mathbf{A} \in S^n_+$ or $\mathbf{A} \succeq 0$) if for all $x \in R^n$ it holds

$$x^T \mathbf{A} x \geq 0.$$

In what follows we will list some well known properties of positive semidefinite matrices.

Proposition A.1.1 $\mathbf{A} \succ 0$ *if and only if* $\mathbf{A} \succeq 0$ *and* $\mathbf{A}$ *is nonsingular.*

Proposition A.1.2 *If* $\mathbf{A} \succ 0$, *then* $\mathbf{A}^{-1} \succ 0$.

Proposition A.1.3 *Let* $\mathbf{A} \succeq 0$ *and* $x^T \mathbf{A} x = 0$ *for some* $x \in R^n$. *Then* $\mathbf{A}x = 0$.

Proposition A.1.4 *Let* $\mathbf{A}_1, \ldots, \mathbf{A}_m \in S^n$ *and* $\alpha_1, \ldots, \alpha_m \in R$.

(a) *If* $\mathbf{A}_1 \succeq 0$ *and* $\mathbf{A}_2 \succ 0$ *then* $\mathbf{A}_1 + \mathbf{A}_2 \succ 0$.

(b) *If* $\mathbf{A}_i \succeq 0$ *and* $\alpha_i \geq 0$ *for all* $i = 1, 2, \ldots, m$, *then* $\alpha_1 \mathbf{A}_1 + \cdots + \alpha_m \mathbf{A}_m \succeq 0$.

(c) *If* $\mathbf{A}_i \succ 0$, $\alpha_i \geq 0$ *for all* $i = 1, 2, \ldots, m$ *and* $\alpha_j > 0$ *for some* $j \in \{1, \ldots, m\}$, *then* $\alpha_1 \mathbf{A}_1 + \cdots + \alpha_m \mathbf{A}_m \succ 0$.

From the last proposition it follows that S^n_+ is a convex cone.

Proposition A.1.5 *(a)* *Let* $\mathbf{A} \in S^n_{++}$ *and* $\mathbf{C} \in R^{n \times n}$ *is a nonsingular matrix, then* $\mathbf{C}^T \mathbf{A} \mathbf{C} \succ 0$.

(b) *Let* $\mathbf{A} \in S^n_+$ *and* $\mathbf{C} \in R^{n \times m}$, *then* $\mathbf{C}^T \mathbf{A} \mathbf{C} \succeq 0$.

(c) *Let* $\mathbf{A} \in S^n_{++}$ *and* $\mathbf{C} \in R^{n\times m}$, *then* $rank(\mathbf{C}^T\mathbf{AC}) = rank(\mathbf{C})$ *and hence* $\mathbf{C}^T\mathbf{AC} \succ 0$ *if and only if* $rank(\mathbf{C}) = m$.

Theorem A.1.3 *Let* $\mathbf{A} \in S^n$. *The following statements are equivalent:*

(a) $\mathbf{A} \in S^n_{++}$,

(b) *the eigenvalues* $\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A})$ *are positive,*

(c) *the determinant of every leading principal submatrix of* $\mathbf{A}$ *is positive,*

(d) *the determinant of every principal submatrix of* $\mathbf{A}$ *is positive,*

(e) *there exists a unique lower triangular matrix* $\mathbf{L}$ *with positive diagonal entries such that* $\mathbf{A} = \mathbf{LL}^T$,

(f) *there exists a unique upper triangular matrix* $\mathbf{U}$ *with positive diagonal entries such that* $\mathbf{A} = \mathbf{UU}^T$,

(g) *there exists a nonsingular matrix* $\mathbf{C} \in R^{n\times n}$ *such that* $\mathbf{A} = \mathbf{CC}^T$,

(h) *there exists an orthogonal matrix* $\mathbf{Q} \in R^{n\times n}$ $(\mathbf{QQ}^T = \mathbf{I})$ *and a positive diagonal matrix* $\mathbf{D}$ *such that* $\mathbf{A} = \mathbf{QDQ}^T$ *and the diagonal entries of* $\mathbf{D}$ *are the eigenvalues of* $\mathbf{A}$.

Theorem A.1.4 *Let* $\mathbf{A} \in S^n$. *The following statements are equivalent:*

(a) $\mathbf{A} \in S^n_{+}$,

(b) *the eigenvalues* $\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A})$ *are nonnegative,*

(c) *the determinant of every principal submatrix of* $\mathbf{A}$ *is nonnegative,*

(d) *there exists a lower triangular matrix* $\mathbf{L}$ *with nonnegative diagonal entries such that* $\mathbf{A} = \mathbf{LL}^T$,

(e) *there exists a upper triangular matrix* $\mathbf{U}$ *with nonnegative diagonal entries such that* $\mathbf{A} = \mathbf{UU}^T$,

(f) *there exists a matrix* $\mathbf{C} \in R^{n\times n}$ *such that* $\mathbf{A} = \mathbf{CC}^T$,

(g) *there exists an orthogonal matrix* $\mathbf{Q} \in R^{n\times n}$ $(\mathbf{QQ}^T = \mathbf{I})$ *and a nonnegative diagonal matrix* $\mathbf{D}$ *such that* $\mathbf{A} = \mathbf{QDQ}^T$ *and the diagonal entries of* $\mathbf{D}$ *are the eigenvalues of* $\mathbf{A}$, *(h)* $\mathbf{A} + \varepsilon\mathbf{I} \in S^n_{++}$ *for every* $\varepsilon > 0$.

Note. The matrix $\mathbf{L}$ from Theorem A.1.3 (e) and Theorem A.1.4 (d), respectively, is called the lower Cholesky factor of the positive (semi)definite matrix $\mathbf{A}$. Similarly, the matrix $\mathbf{U}$ from Theorem A.1.3 (f) and Theorem A.1.4 (e), respectively, is called the upper Cholesky factor of the positive (semi)definite matrix $\mathbf{A}$.

Note. For every positive semidefinite matrix $\mathbf{A}$ there exists a matrix $\mathbf{A}^{\frac{1}{2}}$ such that

$$\mathbf{A}^{\frac{1}{2}}\mathbf{A}^{\frac{1}{2}} = \mathbf{A}.$$

In fact, from the Theorem A.1.4 (g) it follows that
$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T = \mathbf{Q}diag(\lambda_1, \ldots, \lambda_n)\mathbf{Q}^T$, where $\lambda_i \geq 0$ for all $i = 1, \ldots, n$, and therefore one can define $\mathbf{A}^{\frac{1}{2}}$ as

$$\mathbf{A}^{\frac{1}{2}} = \mathbf{Q}diag(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})\mathbf{Q}^T.$$

Such matrix $\mathbf{A}^{\frac{1}{2}}$ is called the square root of $\mathbf{A}$. Moreover, it can be shown that the square root of $\mathbf{A}$ is uniquely determined for any $\mathbf{A} \succeq 0$ (see e.g. [81], Theorem 6.4).

The following properties of positive semidefinite matrices are less known. For this reason also the proofs or exact references are added.

Proposition A.1.6 *If* $\mathbf{A} \in S^n_{++}$ *and* $\mathbf{B} \in S^n$ *then there exists a nonsingular matrix* $\mathbf{P}$ *and a diagonal matrix* $\mathbf{D}$ *such that*

$$\mathbf{A} = \mathbf{P}\mathbf{P}^T \quad \textit{and} \quad \mathbf{B} = \mathbf{P}\mathbf{D}\mathbf{P}^T.$$

Moreover if $\mathbf{B} \in S^n_{++}$ *then* $\mathbf{D}$ *has positive diagonal entries.*

Proof. Denote $\mathbf{A}^{-\frac{1}{2}} = (\mathbf{A}^{\frac{1}{2}})^{-1}$. The matrix $\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{A}^{-\frac{1}{2}}$ is symmetric (and positive definite if $\mathbf{B}$ is positive definite) and hence there exists an orthogonal matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{A}^{-\frac{1}{2}} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$. One can then define $\mathbf{P} = \mathbf{A}^{\frac{1}{2}}\mathbf{Q}$. Then

$$\mathbf{A} = \mathbf{A}^{\frac{1}{2}}\mathbf{A}^{\frac{1}{2}} = \mathbf{A}^{\frac{1}{2}}\mathbf{Q}\mathbf{Q}^T\mathbf{A}^{\frac{1}{2}} = \mathbf{P}\mathbf{P}^T$$

and

$$\mathbf{B} = \mathbf{A}^{\frac{1}{2}}\mathbf{Q}\mathbf{D}\mathbf{Q}^T\mathbf{A}^{\frac{1}{2}} = \mathbf{P}\mathbf{D}\mathbf{P}^T.$$

□

Proposition A.1.7 *(a)* *If* $\mathbf{A} \succeq 0$, *then* $|\mathbf{A}_{ij}| \leq \sqrt{\mathbf{A}_{ii}\mathbf{A}_{jj}}$.

(b) If $\mathbf{A} \succ 0$ *and* $i \neq j$, *then* $|\mathbf{A}_{ij}| < \sqrt{\mathbf{A}_{ii}\mathbf{A}_{jj}}$.

Proof. Consider the principal submatrix

$$\begin{pmatrix} \mathbf{A}_{ii} & \mathbf{A}_{ij} \\ \mathbf{A}_{ij} & \mathbf{A}_{jj} \end{pmatrix}$$

of the matrix $\mathbf{A}$. The statemens (a) and (b) follow from Theorem A.1.4 (c) and Theorem A.1.3 (d), respectively.

□

Corollary A.1.1 *If* $\mathbf{A} \succeq 0$ *and* $\mathbf{A}_{ii} = 0$, *then* $\mathbf{A}_{ij} = \mathbf{A}_{ji} = 0$ *for all* $j = 1, \ldots, n$.

Example A.1.1 *(Sturm at al. [67]) Assume* $\mathbf{A} \succ 0, \mathbf{B} \succ 0$. *Then the matrix* $\mathbf{AB} + \mathbf{BA}$ *is symmetric but not necessary positive definite. In fact, let*

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}.$$

Then $\mathbf{AB} + \mathbf{BA} = \begin{pmatrix} -6 & 0 \\ 0 & 42 \end{pmatrix}$.

Theorem A.1.5 (Theorem 6.10 of [81]) *(Fischer inequality) Let* $\mathbf{A}$ *be a square (complex) matrix.*

$$\text{If} \quad \mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \succeq 0, \quad \text{then} \quad \det \mathbf{A} \leq \det A_{11} \det A_{22}.$$

Theorem A.1.6 (Theorem 6.11 of [81]) *(Hadamard inequality)*
If $\mathbf{A} \succeq 0$ *is a square (complex) matrix, then*

$$\det \mathbf{A} \leq \prod_{i=1}^{n} \mathbf{A}_{ii}$$

and the equality holds if and only if $\mathbf{A}_{ii} = 0$ *for some* i *or* $\mathbf{A}$ *is diagonal.*

Proposition A.1.8 (Lemma 2.3 of [59]) *If* $\mathbf{A} + \mathbf{A}^T \in S^n_{++}$, *then*

$$\det \left(\frac{\mathbf{A} + \mathbf{A}^T}{2} \right) \leq |\det \mathbf{A}|.$$

A.2 Trace and matrix norms

Let $\mathbf{A} \in R^{n\times n}$. The trace of the matrix $\mathbf{A}$ is defined as

$$tr(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}_{ii}.$$

The trace is a linear operator and has the following properties:

Proposition A.2.1 *Let $\mathbf{A}, \mathbf{B} \in R^{n\times n}$ and denote $\lambda_i(\mathbf{A}), i = 1, \ldots, n$ the eigenvalues of $\mathbf{A}$. Then*

(a) $tr(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i(\mathbf{A})$

(b) $tr(\mathbf{A}) = tr(\mathbf{A}^T)$

(c) $tr(\mathbf{AB}) = tr(\mathbf{BA})$

Using the trace one can define the inner product on $R^{n\times n}$ as follows:

$$\mathbf{A} \bullet \mathbf{B} = tr(\mathbf{A}^T\mathbf{B}) = tr(\mathbf{A}\mathbf{B}^T) = tr(\mathbf{B}^T\mathbf{A}) = tr(\mathbf{B}\mathbf{A}^T).$$

The trace operator and the associated inner product have the following properties on the set S^n_+.

Proposition A.2.2 *If $\mathbf{X} \succeq 0$ then $tr(\mathbf{A}) \geq 0$, and $tr(\mathbf{X}) = 0$ if and only if $\mathbf{X} = 0$.*

Proposition A.2.3 *If $\mathbf{X} \succeq 0$, $\mathbf{Y} \succeq 0$, then $\mathbf{X} \bullet \mathbf{Y} \geq 0$.*

Proposition A.2.4 *If $\mathbf{X} \succeq 0, \mathbf{Y} \succeq 0$. Then $\mathbf{XY} = 0 \Leftrightarrow \mathbf{X} \bullet \mathbf{Y} = 0$.*

Proposition A.2.5 *If $\mathbf{X} \succ 0, \mathbf{Y} \succeq 0$. Then $\mathbf{Y} = 0 \Leftrightarrow \mathbf{X} \bullet \mathbf{Y} = 0$.*

Definition A.2.1 *The function $\| \centerdot \| : R^{n\times n} \to R$ is called a* ***matrix norm*** *if for all $\mathbf{A}, \mathbf{B} \in R^{n\times n}$ it satisfies*

(a) $\|\mathbf{A}\| \geq 0$ *and* $\|\mathbf{A}\| = 0$ *if and only if* $\mathbf{A} = 0$,

(b) $\|c\mathbf{A}\| = |c|\|\mathbf{A}\|$,

(c) $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$,

(d) $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$.

The inner product above induces the so called **Frobenius (or Euclidean) matrix norm** on $R^{n\times n}$:

$$\|\mathbf{A}\|_F = \sqrt{\mathbf{A} \bullet \mathbf{A}} = \sqrt{tr(\mathbf{A}\mathbf{A}^T)}.$$

An another matrix norm on $R^{n\times n}$ is the so called **spectral norm**, defined as

$$\|\mathbf{A}\|_2 = \max_i \left\{ \sqrt{\lambda_i(\mathbf{A}^T\mathbf{A})} \right\}.$$

Proposition A.2.6 *(a) The spectral norm is the operator norm induced by the Euclidean vector norm $\|x\| = \sqrt{x^T x}$, that is,*

$$\|\mathbf{A}\|_2 = \max_{\|x\|=1} \|Ax\|$$

(b) If $\mathbf{A} \in S^n$, then $\|\mathbf{A}\|_2 = \max_i |\lambda_i(\mathbf{A})|$.

(c) If $\mathbf{A} \in S^n_+$, then $\|\mathbf{A}\|_2 = \lambda_{max}(\mathbf{A})$.

Proof. (a) The matrix $\mathbf{A}^T\mathbf{A}$ is positive semidefinite and hence by Theorem A.1.4 (g) there exists an orthogonal matrix $\mathbf{Q}$ and a nonnegative diagonal matrix $\mathbf{D} = diag(\lambda_1, \ldots, \lambda_n)$ such that $\mathbf{A}^T\mathbf{A} = \mathbf{Q}^T\mathbf{D}\mathbf{Q}$ and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}^T\mathbf{A}$. Therefore

$$\|\mathbf{A}\|_2 = \max_i \{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\} = \max_{\|x\|=1} \sqrt{\sum_{i=1}^n \lambda_i x_i^2} =$$

$$= \max_{\|x\|=1} \sqrt{x^T\mathbf{D}x} = \max_{\|x\|=1} \sqrt{x^T\mathbf{Q}^T\mathbf{D}\mathbf{Q}x} = \max_{\|x\|=1} \sqrt{x^T\mathbf{A}^T\mathbf{A}x} = \max_{\|x\|=1} \|\mathbf{A}x\|.$$

□

Proposition A.2.7 *If $\mathbf{A} \in S^n$, $\mathbf{B} \in R^{n\times n}$, then*

(a) $\|\mathbf{A}\|_2 \le \|\mathbf{A}\|_F \le \sqrt{n}\|\mathbf{A}\|_2$,

(b) $\|\mathbf{A}\mathbf{B}\|_F \le \|\mathbf{A}\|_2\|\mathbf{B}\|_F$.

Proof. (a) If $\mathbf{A} \in S^n$, then $\|\mathbf{A}\|_F^2 = tr(\mathbf{A}^2) = \sum_{i=1}^n \lambda_i^2(\mathbf{A})$. The statement follows from the inequalities

$$\|\mathbf{A}\|_2^2 = \big(\max_i |\lambda_i(\mathbf{A})|\big)^2 \le \sum_{i=1}^n \lambda_i^2(\mathbf{A}) \le n\big(\max_i |\lambda_i(\mathbf{A})|\big)^2 = n\|\mathbf{A}\|_2^2.$$

(b) If $\mathbf{A} \in S^n$, then $\mathbf{A}^2 \succeq 0$ and hence by Theorem A.1.4 (g) there exists an orthogonal matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{D} = diag(\lambda_1(\mathbf{A}^2), \ldots, \lambda_n(\mathbf{A}^2))$ such that $\mathbf{Q}\mathbf{A}^2\mathbf{Q}^T = \mathbf{D}$. Therefore

$$\|\mathbf{AB}\|_F^2 = tr(\mathbf{ABB}^T\mathbf{A}) = tr(\mathbf{A}^2\mathbf{BB}^T) =$$

$$tr(\mathbf{QA}^2\mathbf{Q}^T\mathbf{QBB}^T\mathbf{Q}^T) = tr(\mathbf{DQBB}^T\mathbf{Q}^T) =$$

$$= \sum_{i=1}^{n} \lambda_i(\mathbf{A}^2)(\mathbf{QBB}^T\mathbf{Q}^T)_{ii} \leq \lambda_{max}(\mathbf{A}^2) tr(\mathbf{QBB}^T\mathbf{Q}^T) = \|\mathbf{A}\|_2^2 \|\mathbf{B}\|_F^2.$$

□

Proposition A.2.8 *If* $\mathbf{A} \succeq 0$ *then*

(a) $\|\mathbf{A}\|_F \leq tr(\mathbf{A})$,

(b) $\|\mathbf{A}\|_2 \leq tr(\mathbf{A})$.

Proof. The statement (a) follows from

$$\|\mathbf{A}\|_F^2 = tr(\mathbf{A}^2) = \sum_{i=1}^{n} \lambda_i(\mathbf{A}^2) \leq \big(\sum_{i=1}^{n} \lambda_i(\mathbf{A})\big)^2.$$

The statement (b) is obvious.

□

Proposition A.2.9 *If the matrices* $\mathbf{B} \succ 0$ *and* $\mathbf{W}, \mathbf{H} \in S^n$ *satisfy* $\mathbf{BH} + \mathbf{HB} = \mathbf{W}$, *then*

$$\|\mathbf{BH}\|_F \leq \frac{\|\mathbf{W}\|_F}{\sqrt{2}}.$$

Proof.

$$\|\mathbf{W}\|_F^2 = \mathbf{W} \bullet \mathbf{W} = (\mathbf{BH} + \mathbf{HB}) \bullet (\mathbf{BH} + \mathbf{HB}) = 2tr(\mathbf{HB}^2\mathbf{H}) + 2tr(\mathbf{BHBH}) =$$

$$2\|\mathbf{BH}\|_F^2 + 2\|\mathbf{B}^{\frac{1}{2}}\mathbf{HB}^{\frac{1}{2}}\|_F^2 \geq 2\|\mathbf{BH}\|_F^2.$$

□

As an immediate consequence we obtain (see also Proposition A.4.8):

Corollary A.2.1 *Let* $\mathbf{B} \succ 0$ *be fixed and* $\mathbf{X} \in S^n$. *Then the equation*

$$\mathbf{BX} + \mathbf{XB} = 0$$

has the only solution $\mathbf{X} = 0$.

Proposition A.2.10 (Lemma 2.1.of [59]) *If* $\mathbf{A} \succ 0$ *and* $\mathbf{B} \succeq 0$ *then*

$$tr(\mathbf{A}^{-1}\mathbf{B}) \geq \frac{tr(\mathbf{B})}{\lambda_{max}(\mathbf{A})} \geq \frac{tr(\mathbf{B})}{tr(\mathbf{A})}$$

Proof. Since $\mathbf{A} \succ 0$, there exists an orthogonal matrix $\mathbf{Q}$ and a diagonal matrix

$$\mathbf{D} = diag(\lambda_1(\mathbf{A}), \ldots, \lambda_n(\mathbf{A}))$$

such that $\mathbf{QAQ}^T = \mathbf{D}$. Therefore $\mathbf{QA}^{-1}\mathbf{Q}^T = \mathbf{D}^{-1}$ and

$$tr(\mathbf{A}^{-1}\mathbf{B}) = tr(\mathbf{QA}^{-1}\mathbf{Q}^T\mathbf{QBQ}^T) = tr(\mathbf{D}^{-1}\mathbf{QBQ}^T) =$$
$$= \sum_{i=1}^{n} \lambda_i(\mathbf{A})^{-1}(\mathbf{QBQ}^T)_{ii} \geq \frac{\sum_{i=1}^{n}(\mathbf{QBQ}^T)_{ii}}{\lambda_{max}(\mathbf{A})} \geq \frac{tr(\mathbf{B})}{tr(\mathbf{A})}.$$

□

Proposition A.2.11 (Lemma 7 of [51]) *If* $\mathbf{A}$ *is a square matrix such that* $tr(\mathbf{A}^2) \geq 0$, *then*

$$\|\mathbf{A}\|_F \leq \sqrt{2}\left\|\frac{\mathbf{A}+\mathbf{A}^T}{2}\right\|_F \leq \sqrt{2}tr(\mathbf{A}).$$

Proof. The second inequality follows from Proposition A.2.7. We will prove the first inequality. Every square matrix is the sum of its symmetric and skewsymmetric part.

$$\mathbf{A} = \frac{\mathbf{A}+\mathbf{A}^T}{2} + \frac{\mathbf{A}-\mathbf{A}^T}{2}.$$

It holds $tr[(\mathbf{A}+\mathbf{A}^T)(\mathbf{A}-\mathbf{A}^T)] = tr(\mathbf{AA} - \mathbf{A}^T\mathbf{A}^T) = 0$ and therefore

$$tr(\mathbf{A}^2) = tr\left(\frac{\mathbf{A}+\mathbf{A}^T}{2}\right)^2 + tr\left(\frac{\mathbf{A}-\mathbf{A}^T}{2}\right)^2 = \left\|\frac{\mathbf{A}+\mathbf{A}^T}{2}\right\|_F^2 - \left\|\frac{\mathbf{A}-\mathbf{A}^T}{2}\right\|_F^2$$

and

$$\|\mathbf{A}\|_F^2 = tr\left(\frac{\mathbf{A}+\mathbf{A}^T}{2}\right)^2 - tr\left(\frac{\mathbf{A}-\mathbf{A}^T}{2}\right)^2 = \left\|\frac{\mathbf{A}+\mathbf{A}^T}{2}\right\|_F^2 + \left\|\frac{\mathbf{A}-\mathbf{A}^T}{2}\right\|_F^2.$$

By adding the above equalities and by the assumption $tr(\mathbf{A}^2) \geq 0$ we obtain

$$\|\mathbf{A}\|_F^2 \leq \|\mathbf{A}\|_F^2 + tr(\mathbf{A}^2) = 2\left\|\frac{\mathbf{A} + \mathbf{A}^T}{2}\right\|_F^2.$$

□

Theorem A.2.1 (Theorem 7.1.1 of [40]) *Let $\|\cdot\|$ be any matrix norm, such that $\|\mathbf{I}\| = 1$. If $\nu = \|\mathbf{M}\| < 1$, then the matrix $\mathbf{I} + \mathbf{M}$ is invertible, moreover it holds*

$$(\mathbf{I} + \mathbf{M})^{-1} = \mathbf{I} - \mathbf{M} + \mathbf{M}^2 - \ldots$$

and

$$\|(\mathbf{I} + \mathbf{M})^{-1}\| \leq \frac{1}{1 - \nu}.$$

Proposition A.2.12 (Lemma 8 of [51]) *Let $\mathbf{B}$ be a square matrix with real eigenvalues. Let $t \in (0, \frac{1}{\sqrt{2}})$ be given. Then if*

$$\left\|\frac{\mathbf{B} + \mathbf{B}^T}{2} - \mathbf{I}\right\|_F \leq t,$$

then

(a) $\|\mathbf{B} - \mathbf{I}\|_F \leq \sqrt{2}t$;

(b) $\|\mathbf{B}^{-1}\|_2 \leq \frac{1}{1 - \sqrt{2}t}$.

Proof. (a) Obviously the matrix $\mathbf{B} - \mathbf{I}$ has real eigenvalues and hence $tr(\mathbf{B} - \mathbf{I})^2 \geq 0$. Proposition A.2.11 then implies (let $\mathbf{A} = \mathbf{B} - \mathbf{I}$)

$$\|\mathbf{B} - \mathbf{I}\|_F \leq \sqrt{2}\left\|\frac{\mathbf{B} + \mathbf{B}^T}{2} - \mathbf{I}\right\|_F \leq \sqrt{2}t.$$

(b) From the assumption, the statement (a) and Proposition A.2.7 (a) it follows that $\|\mathbf{B} - \mathbf{I}\|_2 < 1$. Therefore we can use Theorem A.2.1 and obtain that

$$\|\mathbf{B}^{-1}\|_2 = \|(\mathbf{I} + (\mathbf{B} - \mathbf{I}))^{-1}\|_2 \leq \frac{1}{1 - \|\mathbf{B} - \mathbf{I}\|_2} \leq \frac{1}{1 - \sqrt{2}t}.$$

□

A.3 Schur complement

The concept of the Schur complement is a main tool in handling the matrices in semidefinite optimization. Many properties of the Schur complement are described e.g. in [81]. In this thesis we will need the properties presented in this section.

Definition A.3.1 *Let*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_{11}} & \mathbf{A_{12}} \\ \mathbf{A_{21}} & \mathbf{A_{22}} \end{pmatrix}$$

be a partitioned matrix (not square in general), where $\mathbf{A_{11}}$ *is a square submatrix. Then if* $\mathbf{A_{11}}$ *is nonsingular, then the matrix*

$$[\mathbf{A}/\mathbf{A_{11}}] = \mathbf{A_{22}} - \mathbf{A_{21}}\mathbf{A_{11}^{-1}}\mathbf{A_{12}}.$$

is called the ***Schur complement*** *of* $\mathbf{A_{11}}$ *in* $\mathbf{A}$.

Proposition A.3.1 *Let*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_{11}} & \mathbf{A_{12}} \\ \mathbf{A_{21}} & \mathbf{A_{22}} \end{pmatrix}$$

be a square matrix with a nonsingular square submatrix $\mathbf{A_{11}}$. *Then*

(a) $\det \mathbf{A} = \det \mathbf{A_{11}} \det[\mathbf{A}/\mathbf{A_{11}}]$,

(b) If $\mathbf{A} \in S^n$ *and* $\mathbf{A}_{11} \succ 0$, *then* $\mathbf{A} \succeq 0$ *if and only if* $[\mathbf{A}/\mathbf{A_{11}}] \succeq 0$,

(c) If $\mathbf{A} \in S^n$, *then* $\mathbf{A}_{11} \succ 0$ *if and only if* $\mathbf{A} \succ 0$ *and* $[\mathbf{A}/\mathbf{A_{11}}] \succ 0$.

Proof. All of the statements follow from the identity

$$\begin{pmatrix} \mathbf{I} & 0 \\ -\mathbf{A_{21}}\mathbf{A_{11}}^{-1} & \mathbf{I} \end{pmatrix} \mathbf{A} \begin{pmatrix} \mathbf{I} & -\mathbf{A_{11}}^{-1}\mathbf{A_{12}} \\ 0 & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{A_{11}} & 0 \\ 0 & [\mathbf{A}/\mathbf{A_{11}}] \end{pmatrix} \tag{A.1}$$

□

Corollary A.3.1 *Let* $\mathbf{A}$ *be a square matrix with a nonsingular square submatrix* $\mathbf{A_{11}}$. *Then* $\mathbf{A}$ *is nonsingular if and only if* $[\mathbf{A}/\mathbf{A_{11}}]$ *is nonsingular.*

Note. The statement (b) in Proposition A.3.1 says that if $\mathbf{A_{11}} \succ 0$, then $\mathbf{A} \succeq 0$ if and only if $\mathbf{A_{22}} \succeq \mathbf{A_{21}}\mathbf{A_{11}^{-1}}\mathbf{A_{12}}$ [1]. Therefore the equality $\mathbf{A_{22}} = \mathbf{A_{21}}\mathbf{A_{11}^{-1}}\mathbf{A_{12}}$ implies

$$\begin{pmatrix} \mathbf{A_{11}} & \mathbf{A_{12}} \\ \mathbf{A_{21}} & \mathbf{A_{21}}\mathbf{A_{11}^{-1}}\mathbf{A_{12}} \end{pmatrix} \succeq 0$$

[1] If $\mathbf{A}, \mathbf{B} \in S^n$ we write $\mathbf{A} \succeq \mathbf{B}$ if and only if $\mathbf{A} - \mathbf{B} \succeq 0$. The relation $\succeq$ is referred to as the Löwner partial ordering (see [81], Section 6.2).

and $\mathbf{A_{21}A_{11}^{-1}A_{12}}$ is "the smallest" matrix (in the Löwner partial ordering sense) to make the block matrix positive semidefinite.

Proposition A.3.2 *Let* $\mathbf{A}$ *be a nonsingular matrix with partition given in Definition A.3.1. If*

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{B_{11}} & \mathbf{B_{12}} \\ \mathbf{B_{21}} & \mathbf{B_{22}} \end{pmatrix}$$

then $\mathbf{B_{22}}^{-1} = [\mathbf{A}/\mathbf{A_{11}}]$.

Proof. By inverting $(A.1)$ we obtain

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{A_{11}}^{-1}\mathbf{A_{12}} \\ 0 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A_{11}}^{-1} & 0 \\ 0 & [\mathbf{A}/\mathbf{A_{11}}]^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & 0 \\ \mathbf{A_{21}}\mathbf{A_{11}}^{-1} & \mathbf{I} \end{pmatrix}.$$

The rest follows from a simple computation.

□

Proposition A.3.3 (Lemma 2.2 of [59]) *If* $\mathbf{A} \in S^n$ *and*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_{11}} & \mathbf{A_{12}} \\ \mathbf{A_{21}} & \mathbf{A_{22}} \end{pmatrix} \succeq 0$$

($\mathbf{A_{21}} = \mathbf{A_{12}}^T$*), then*

$$\|\mathbf{A}_{12}\|_F^2 \leq tr(\mathbf{A}_{11})tr(\mathbf{A}_{22}).$$

Proof. If $\mathbf{A} \succeq 0$, then $\mathbf{A} + \varepsilon\mathbf{I} \succ 0$ for any $\varepsilon > 0$. Denote

$$[\mathbf{A}/\mathbf{A_{11}}]_\varepsilon = \mathbf{A_{22}} + \varepsilon\mathbf{I} - \mathbf{A}_{21}(\mathbf{A_{11}} + \varepsilon\mathbf{I})^{-1}\mathbf{A}_{12}$$

the Schur complement of $\mathbf{A}_{11}+\varepsilon\mathbf{I}$ in $\mathbf{A}+\varepsilon\mathbf{I}$. Proposition A.3.1 (c) states that $[\mathbf{A}/\mathbf{A_{11}}]_\varepsilon \succ 0$ and hence

$$tr(\mathbf{A_{22}} + \varepsilon\mathbf{I}) > tr(\mathbf{A}_{21}(\mathbf{A_{11}} + \varepsilon\mathbf{I})^{-1}\mathbf{A}_{12}) = tr((\mathbf{A_{11}} + \varepsilon\mathbf{I})^{-1}\mathbf{A}_{12}\mathbf{A}_{12}^T) \geq \frac{tr(\mathbf{A}_{12}\mathbf{A}_{12}^T)}{tr(\mathbf{A_{11}} + \varepsilon\mathbf{I})},$$

where the last inequality follows from Proposition A.2.10. Hence for any $\varepsilon > 0$ we have $tr(\mathbf{A_{22}}+\varepsilon\mathbf{I})tr(\mathbf{A_{11}}+\varepsilon\mathbf{I}) > tr(\mathbf{A}_{12}\mathbf{A}_{12}^T)$. Taking the limit $\varepsilon \to 0$ leads to $tr(\mathbf{A_{22}})tr(\mathbf{A_{11}}) \geq tr(\mathbf{A}_{12}\mathbf{A}_{12}^T) = \|\mathbf{A}_{12}\|_F^2$.

□

A.4 Symmetric Kronecker product

The symmetric Kronecker product is a useful tool for expressing linear matrix maps. One can find its applications in the theory of interior point methods for semidefinite programming. In this section we will list the most important properties of the symmetric Kronecker product. More properties and proofs can be found e.g. in [34].

Definition A.4.1 *For any symmetric matrix*

$$\mathbf{X} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{pmatrix}$$

we define the vector $svec(\mathbf{X}) \in R^{\frac{n(n+1)}{2}}$ *as*

$$svec(\mathbf{X}) = (X_{11}, \sqrt{2}X_{12}, \dots, \sqrt{2}X_{1n}, X_{22}, \sqrt{2}X_{23}, \dots, \sqrt{2}X_{2n}, \dots, X_{nn})^T.$$

Note. The map $\mathbf{X} \mapsto svec(\mathbf{X})$ is an isomorphism $S^n \to R^{\frac{n(n+1)}{2}}$. Moreover, if $\mathbf{X}, \mathbf{Y} \in S^n$, then $\mathbf{X} \bullet \mathbf{Y} = svec(\mathbf{X})^T svec(\mathbf{Y})$.

Definition A.4.2 *Let* $\mathbf{M}, \mathbf{N}$ *are (not necessary symmetric) square matrices in* $R^{n\times n}$*. Then the map* $\mathbf{M} \star \mathbf{N} : S^n \to S^n$ *defined as*

$$\mathbf{X} \mapsto \frac{1}{2}(\mathbf{N}\mathbf{X}\mathbf{M}^T + \mathbf{M}\mathbf{X}\mathbf{N}^T)$$

is called the ***symmetric Kronecker product*** *of the matrices* $\mathbf{M}, \mathbf{N}$.

Obviously, the matrix $(\mathbf{M} \star \mathbf{N}) \in R^{\bar{n}\times\bar{n}}$ and

$$(\mathbf{M} \star \mathbf{N}) svec(\mathbf{X}) = \frac{1}{2} svec(\mathbf{N}\mathbf{X}\mathbf{M}^T + \mathbf{M}\mathbf{X}\mathbf{N}^T). \tag{A.2}$$

Let $\mathbf{M}, \mathbf{N} \in R^{n\times n}$. In what follows we list some basic properties of the symmetric Kronecker product $(\mathbf{M} \star \mathbf{N})$.

Proposition A.4.1 *(a) The symmetric Kronecker product is commutative:*

$$\mathbf{M} \star \mathbf{N} = \mathbf{N} \star \mathbf{M}.$$

(b) The symmetric Kronecker product is not associative in general:

$$\mathbf{K} \star (\mathbf{M} \star \mathbf{N}) \neq (\mathbf{K} \star \mathbf{M}) \star \mathbf{N}.$$

Proposition A.4.2

$$(\alpha\mathbf{M}) \star \mathbf{N} = \mathbf{M} \star (\alpha\mathbf{N}) = \alpha(\mathbf{M} \star \mathbf{N}) \qquad \forall\ \alpha \in R.$$

Proposition A.4.3

$$(\mathbf{M} \star \mathbf{N})^T = \mathbf{M}^T \star \mathbf{N}^T$$

and hence $\mathbf{M} \star \mathbf{I}$ *is symmetric if and only if* $\mathbf{M}$ *is symmetric.*

Proposition A.4.4

$$(a) \quad (\mathbf{K} + \mathbf{M}) \star \mathbf{N} = \mathbf{K} \star \mathbf{N} + \mathbf{M} \star \mathbf{N},$$

$$(b) \quad \mathbf{K} \star (\mathbf{M} + \mathbf{N}) = \mathbf{K} \star \mathbf{M} + \mathbf{K} \star \mathbf{N}.$$

Proposition A.4.5

$$(\mathbf{K} \star \mathbf{L})(\mathbf{M} \star \mathbf{N}) = \frac{1}{2}(\mathbf{KM} \star \mathbf{LN} + \mathbf{KN} \star \mathbf{LM}),$$

furthermore

$$(\mathbf{K} \star \mathbf{L})(\mathbf{M} \star \mathbf{M}) = \mathbf{KM} \star \mathbf{LM},$$

and

$$(\mathbf{K} \star \mathbf{K})(\mathbf{M} \star \mathbf{N}) = \mathbf{KM} \star \mathbf{KN}.$$

Proposition A.4.6 *Let* $\mathbf{M}, \mathbf{N} \in R^{n \times n}$ *be nonsingular. Then*

$$(a) \quad (\mathbf{M} \star \mathbf{M})^{-1} = \mathbf{M}^{-1} \star \mathbf{M}^{-1},$$

but

$$(b) \quad (\mathbf{M} \star \mathbf{N})^{-1} \neq \mathbf{M}^{-1} \star \mathbf{N}^{-1},$$

in general.

Proposition A.4.7 *If* $\mathbf{M} \succ 0, \mathbf{N} \succ 0$*, then* $\mathbf{M} \star \mathbf{N} \succ 0$.

Proof. Let $\mathbf{M} \succ 0, \mathbf{N} \succ 0$ and $\mathbf{U} \in S^n, \mathbf{U} \neq \mathbf{0}$. Then

$$svec(\mathbf{U})^T(\mathbf{M} \star \mathbf{N})svec(\mathbf{U}) = \frac{1}{2}svec(\mathbf{U})^T svec(\mathbf{NUM} + \mathbf{MUN}) =$$

$$= \frac{1}{2}[\mathbf{U} \bullet \mathbf{NUM} + \mathbf{U} \bullet \mathbf{MUN}] = \mathbf{UMU} \bullet \mathbf{N} > 0$$

□

Corollary A.4.1 *Let* $\mathbf{A} \in R^{n\times n}$, $\mathbf{B} \in S^n$ *and* $\mathbf{X} \in S^n$. *Then*

$$svec(\mathbf{AXA}^T) = (\mathbf{A} \star \mathbf{A})svec(\mathbf{X})$$

and

$$svec(\mathbf{BX} + \mathbf{XB}) = 2(\mathbf{B} \star \mathbf{I})svec(\mathbf{X}).$$

Proposition A.4.8 *Let* $\mathbf{B} \succ 0$ *be fixed and* $\mathbf{X} \in S^n$. *Then the equation*

$$\mathbf{BX} + \mathbf{XB} = 0$$

has the only solution $\mathbf{X} = 0$.

Proof. From the symmetric Kronecker product representation in Corollary A.4.1 we can see that $\mathbf{BX} + \mathbf{XB} = 0$ if and only if $(\mathbf{B} \star \mathbf{I})svec(\mathbf{X}) = 0$. Proposition A.4.7 implies that $(\mathbf{B} \star \mathbf{I})$ is positive definite and hence nonsingular. Therefore $\mathbf{X} = 0$.

□

Corollary A.4.2 *If* $\mathbf{B} \succ 0$, *then the map* $\mathbf{X} \to \mathbf{BX} + \mathbf{XB}$ *is a linear isomorphism* $S^n \to S^n$.

Corollary A.4.3 *Let* $\mathbf{B} \succ 0$ *and* $\mathbf{W} \in S^n$. *Then the equation* $\mathbf{BX} + \mathbf{XB} = \mathbf{W}$ *has a unique solution* $\mathbf{X} \in S^n$.

Proposition A.4.9 *Let* $\mathbf{X} \succ 0, \mathbf{S} \succ 0$ *and* $\mathbf{XS} + \mathbf{SX} \succ 0$. *Then* $(\mathbf{X}\star\mathbf{I})(\mathbf{S}\star\mathbf{I})$ *is positive definite (however not necessary symmetric).*

Proof. Let $\mathbf{V} \in S^n$ and denote $v = svec(\mathbf{V})$. It suffices to show that

$$v^T(\mathbf{X} \star \mathbf{I})(\mathbf{S} \star \mathbf{I})v > 0.$$

We have that

$$v^T(\mathbf{X} \star \mathbf{I})(\mathbf{S} \star \mathbf{I})v = \frac{1}{2}v^T(\mathbf{XS} \star \mathbf{I} + \mathbf{S} \star \mathbf{X})v > \frac{1}{2}v^T(\mathbf{XS} \star \mathbf{I})v =$$

$$= \frac{1}{2}(svec(\mathbf{V}))^T(\mathbf{XS} \star \mathbf{I})svec(\mathbf{V}) = \frac{1}{4}(svec(\mathbf{V}))^T svec(\mathbf{XSV} + \mathbf{VSX}) =$$

$$\frac{1}{4}[\mathbf{V} \bullet (\mathbf{XSV}) + \mathbf{V} \bullet (\mathbf{VXS})] = \frac{1}{4}tr(\mathbf{V}^2\mathbf{SX} + \mathbf{VXSV}) = \frac{1}{2}tr[\mathbf{V}(\mathbf{XS} + \mathbf{SX})\mathbf{V}] > 0.$$

□

From Proposition A.4.9, Proposition A.4.3 and Proposition A.1.5 it immediately follows:

Corollary A.4.4 *Let* $\mathbf{X} \succ 0, \mathbf{S} \succ 0$ *and* $\mathbf{XS} + \mathbf{SX} \succ 0$. *Then all the matrices (a)-(h) are positive definite (however not necessary symmetric):*

$$\begin{array}{llll} (a) & (\mathbf{S} \star \mathbf{I})^{-1}(\mathbf{X} \star \mathbf{I}) & (e) & (\mathbf{S} \star \mathbf{I})(\mathbf{X} \star \mathbf{I}) \\ (b) & (\mathbf{X} \star \mathbf{I})(\mathbf{S} \star \mathbf{I})^{-1} & (f) & (\mathbf{X} \star \mathbf{I})(\mathbf{S} \star \mathbf{I}) \\ (c) & (\mathbf{S} \star \mathbf{I})(\mathbf{X} \star \mathbf{I})^{-1} & (g) & (\mathbf{S} \star \mathbf{I})^{-1}(\mathbf{X} \star \mathbf{I})^{-1} \\ (d) & (\mathbf{X} \star \mathbf{I})^{-1}(\mathbf{S} \star \mathbf{I}) & (h) & (\mathbf{X} \star \mathbf{I})^{-1}(\mathbf{S} \star \mathbf{I})^{-1} \end{array}$$

A.5 Triangular matrices

We will denote L^n and U^n the vector space of all $n \times n$ lower and upper triangular matrices, respectively. Obviously

$$\dim L^n = \dim U^n = \frac{n(n+1)}{2}.$$

Define

$$L^n_+ = \{\mathbf{L} \in L^n,\ \mathbf{L}_{ii} \geq 0,\ \forall i = 1, \ldots, n\},$$
$$L^n_{++} = \{\mathbf{L} \in L^n,\ \mathbf{L}_{ii} > 0,\ \forall i = 1, \ldots, n\},$$

and similarly

$$U^n_+ = \{\mathbf{U} \in U^n,\ \mathbf{U}_{ii} \geq 0,\ \forall i = 1, \ldots, n\},$$
$$U^n_{++} = \{\mathbf{U} \in L^n,\ \mathbf{U}_{ii} > 0,\ \forall i = 1, \ldots, n\}.$$

Proposition A.5.1 *Let* $\mathbf{L}, \mathbf{L}_1, \mathbf{L}_2 \in L^n$ *and* $\mathbf{U}, \mathbf{U}_1, \mathbf{U}_2 \in U^n$.

(a) *If* $\mathbf{L} \in L^n_+$, *then* $\alpha\mathbf{L} \in L^n_+$ *for all* $\alpha \geq 0$,

(b) *If* $\mathbf{L} \in L^n_{++}$, *then* $\alpha\mathbf{L} \in L^n_{++}$ *for all* $\alpha > 0$,

(c) *If* $\mathbf{U} \in U^n_+$, *then* $\alpha\mathbf{U} \in U^n_+$ *for all* $\alpha \geq 0$,

(d) *If* $\mathbf{U} \in U^n_{++}$, *then* $\alpha\mathbf{U} \in U^n_{++}$ *for all* $\alpha \geq 0$,

(e) *If* $\mathbf{L}_1, \mathbf{L}_2 \in L^n_+$, *then* $\mathbf{L}_1 + \mathbf{L}_2 \in L^n_+$,

(f) *If* $\mathbf{L}_1, \mathbf{L}_2 \in L^n_{++}$, *then* $\mathbf{L}_1 + \mathbf{L}_2 \in L^n_{++}$,

(g) *If* $\mathbf{U}_1, \mathbf{U}_2 \in U^n_+$, *then* $\mathbf{U}_1 + \mathbf{U}_2 \in U^n_+$,

(h) *If* $\mathbf{U}_1, \mathbf{U}_2 \in U^n_{++}$, *then* $\mathbf{U}_1 + \mathbf{U}_2 \in U^n_{++}$.

Corollary A.5.1 *The sets* L^n_+, U^n_+ (L^n_{++}, U^n_{++}) *form a closed (open) convex cone.*

Proposition A.5.2 *(a)* $\mathbf{L} \in L^n_{++}$ *if and only if* $\mathbf{L} \in L^n_+$ *and* $\mathbf{L}$ *is nonsingular.*

(b) $\mathbf{U} \in U^n_{++}$ *if and only if* $\mathbf{U} \in U^n_+$ *and* $\mathbf{U}$ *is nonsingular.*

Proposition A.5.3 *Let* $\mathbf{L}, \mathbf{L}_1, \mathbf{L}_2 \in L^n$ *and* $\mathbf{U}, \mathbf{U}_1, \mathbf{U}_2 \in U^n$.

(a) *If* $\mathbf{L}_1, \mathbf{L}_2 \in L^n$, *then* $\mathbf{L}_1\mathbf{L}_2 \in L^n$,

(b) *If* $\mathbf{L}_1, \mathbf{L}_2 \in L^n_+$, *then* $\mathbf{L}_1\mathbf{L}_2 \in L^n_+$,

(c) *If* $\mathbf{L}_1, \mathbf{L}_2 \in L^n_{++}$, *then* $\mathbf{L}_1\mathbf{L}_2 \in L^n_{++}$,

(d) *If* $\mathbf{U}_1, \mathbf{U}_2 \in U^n$, *then* $\mathbf{U}_1\mathbf{U}_2 \in U^n$,

(e) *If* $\mathbf{U}_1, \mathbf{U}_2 \in U^n_+$, *then* $\mathbf{U}_1\mathbf{U}_2 \in U^n_+$,

(f) *If* $\mathbf{U}_1, \mathbf{U}_2 \in U^n_{++}$, *then* $\mathbf{U}_1\mathbf{U}_2 \in U^n_{++}$,

(g) *If* $\mathbf{L} \in L^n$, *then* $\mathbf{L}^2 \in L^n_+$, *moreover if* $\mathbf{L}$ *is nonsingular, then* $\mathbf{L}^2 \in L^n_{++}$,

(h) *If* $\mathbf{U} \in U^n$, *then* $\mathbf{U}^2 \in L^n_+$, *moreover if* $\mathbf{U}$ *is nonsingular, then* $\mathbf{U}^2 \in L^n_{++}$,

(i) *If* $\mathbf{L} \in L^n$ *is nonsingular, then* $\mathbf{L}^{-1} \in L^n$,

(j) *If* $\mathbf{L} \in L^n_{++}$, *then it is nonsingular and* $\mathbf{L}^{-1} \in L^n_{++}$,

(k) *If* $\mathbf{U} \in U^n$ *is nonsingular, then* $\mathbf{U}^{-1} \in U^n$,

(l) *If* $\mathbf{U} \in U^n_{++}$, *then it is nonsingular and* $\mathbf{U}^{-1} \in U^n_{++}$.

Proof. Denote $r_i(\mathbf{A})$ and $s_j(\mathbf{A})$ the vector which corresponds to the i-th row and the j-th column of the matrix $\mathbf{A}$, respectively. Let $\mathbf{L}_1, \mathbf{L}_2 \in L^n$ and $\mathbf{U}_1, \mathbf{U}_2 \in U^n$. The statements (a)-(h) then follow from

$$(\mathbf{L}_1\mathbf{L}_2)_{ii} = r_i(\mathbf{L}_1)^T s_j(\mathbf{L}_2) = \begin{cases} 0 & i > j \\ (\mathbf{L}_1)_{ii}(\mathbf{L}_2)_{ii} & i = j, \end{cases}$$

$$(\mathbf{U}_1\mathbf{U}_2)_{ii} = r_i(\mathbf{U}_1)^T s_j(\mathbf{U}_2) = \begin{cases} 0 & i < j \\ (\mathbf{U}_1)_{ii}(\mathbf{U}_2)_{ii} & i = j. \end{cases}$$

The statements (i)-(l) follow from the fact that the inverse of a nonsingular matrix $\mathbf{A}$ can be obtained from the adjoint of $\mathbf{A}$, more precisely

$$(\mathbf{A}^{-1})_{ij} = \frac{[adj(\mathbf{A})]_{ij}}{\det \mathbf{A}} = \frac{(-1)^{i+j} \det \mathbf{A}(j|i)}{\det \mathbf{A}},$$

where $\mathbf{A}(j|i)$ is the submatrix of $\mathbf{A}$ obtained by deleting the j-th row and the i-th column—and from the observation that if $\mathbf{L} \in L^n$, then $[adj(\mathbf{L})]_{ij} = 0$ for $i > j$ and if $\mathbf{U} \in U^n$, then $[adj(\mathbf{U})]_{ij} = 0$ for $i < j$.

□

Proposition A.5.4 *If* $\mathbf{L} \in L^n$ *(*$\mathbf{U} \in U^n$*), then the eigenvalues of* $\mathbf{L}$ *(*$\mathbf{U}$*) are the diagonal entries* $\mathbf{L}_{ii}$ *(*$\mathbf{U}_{ii}$*).*

Lemma A.5.1 *Let* $\mathbf{L} \in L^n$, $\mathbf{U} \in U^n$*. Then*

(a) $\mathbf{L} + \mathbf{L}^T = 0$ *if and only if* $\mathbf{L} = 0$.

(b) $\mathbf{U} + \mathbf{U}^T = 0$ *if and only if* $\mathbf{U} = 0$.

Proposition A.5.5 *(a) Let* $\mathbf{L} \in L^n_{++}$ *be fixed and* $\mathbf{X} \in L^n$*. Then the equation*

$$\mathbf{L}\mathbf{X}^T + \mathbf{X}\mathbf{L}^T = 0$$

has the only solution $\mathbf{X} = 0$.

(b) Let $\mathbf{U} \in U^n_{++}$ *be fixed and* $\mathbf{X} \in U^n$*. Then the equation*

$$\mathbf{U}\mathbf{X}^T + \mathbf{X}\mathbf{U}^T = 0$$

has the only solution $\mathbf{X} = 0$.

Proof. (a) The matrix $\mathbf{L}$ is nonsingular and hence $\mathbf{L}\mathbf{X}^T + \mathbf{X}\mathbf{L}^T = 0$ if and only if $\mathbf{L}^{-1}(\mathbf{L}\mathbf{X}^T + \mathbf{X}\mathbf{L}^T)\mathbf{L}^{-T} = \mathbf{X}^T\mathbf{L}^{-T} + \mathbf{L}^{-1}\mathbf{X} = 0$. From Proposition A.5.3 it follows $\mathbf{L}^{-1}\mathbf{X} \in L^n$ and moreover $(\mathbf{L}^{-1}\mathbf{X})^T = \mathbf{X}^T\mathbf{L}^{-T}$. Lemma A.5.1 implies $\mathbf{L}^{-1}\mathbf{X} = 0$ and hence from the nonsingularity of $\mathbf{L}$ we obtain $\mathbf{X} = 0$.
The statement (b) can be proved similarly.

□

Corollary A.5.2 *If* $\mathbf{L} \in L^n_{++}$ *(*$\mathbf{U} \in U^n_{++}$*), then the map* $\mathbf{X} \to \mathbf{L}\mathbf{X}^T + \mathbf{X}\mathbf{L}^T$ *(*$\mathbf{X} \to \mathbf{U}\mathbf{X}^T + \mathbf{X}\mathbf{U}^T$*) is a linear isomorphism* $L^n \to S^n$ *(*$U^n \to S^n$*).*

Corollary A.5.3 *Let* $\mathbf{L} \in L^n_{++}$ *(*$\mathbf{U} \in U^n_{++}$*) and* $\mathbf{W} \in S^n$*. Then the equation* $\mathbf{L}\mathbf{X}^T + \mathbf{X}\mathbf{L}^T = \mathbf{W}$ *(*$\mathbf{U}\mathbf{X}^T + \mathbf{X}\mathbf{U}^T = \mathbf{W}$*) has a unique solution* $\mathbf{X} \in L^n(U^n)$*.*

Proposition A.5.6 *If* $\mathbf{L} \in L^n$ *and* $\mathbf{L} + \mathbf{L}^T = \mathbf{W}$*, then*

$$\|\mathbf{L}\|_F \leq \frac{\|\mathbf{W}\|_F}{\sqrt{2}}.$$

Proof.

$$\|\mathbf{W}\|_F^2 = \mathbf{W} \bullet \mathbf{W} = 2tr(\mathbf{L}^2) + 2tr(\mathbf{L}\mathbf{L}^T) \geq 2\|\mathbf{L}\|_F^2.$$

□

Appendix B

Fréchet derivative of some matrix functions

In this Appendix we derive the Fréchet derivatives of some special matrix maps that are used in this thesis. In the first section we recall some basic definitions and properties (see also [39]). In the section B.2 we focus on concrete matrix maps defined on open subsets of S^n or L^n, respectively.

B.1 Basic definitions and properties

Definition B.1.1 *Let X, Y be normed vector spaces and $U \subset X$ be an open subset of X. A map $F : U \to Y$ is called Fréxhet differentiable at $x \in U$ if there exists a bounded linear operator $DF(x) : X \to Y$, such that*

$$F(x+h) - F(x) = DF(x)[h] + \omega(x,h)$$

and

$$\lim_{\|h\|\to 0} \frac{\|\omega(x,h)\|}{\|h\|} = 0$$

For all $h \in X$ it holds $DF(x)[h] \in Y$ and the operator $DF(x)$ is called the Fréchet derivative of F at x.

Proposition B.1.1 *(a) If $F(x) = c$, then $DF(x)[h] = 0$ for all $h \in X$.*

(b) If F is a linear map, then $DF(x) = F$.

(c) *Let F, G, H be continuous maps from X to Y. If the maps are Fréchet differentiable at $x \in X$, then $F + G, aF, FG, FGH$ are Fréchet differentiable at x and*

$$\begin{aligned} D(F+G)(x) &= DF(x) + DG(x) \\ D(aF)(x) &= aDF(x) \\ D(FG)(x) &= DF(x)G(x) + F(x)DG(x) \\ D(FGH)(x) &= DF(x)G(x)H(x) + F(x)DG(x)H(x) + F(x)G(x)DH(x) \end{aligned}$$

(d) *Let X, Y, Z be normed vector spaces, $O(x)$ be a neighbourhood of x, F be a continuous map from $O(x)$ to Y and $F(x) = y$. Let $O(y)$ be a neighbourhood of y and G be a continuous map from $O(y)$ to Z. If F is Fréchet differentiable at x and G Fréchet differentiable at y, then the composition $G \circ F$ (which is defined and continuous on some neighbourhood of x) is Fréchet differentiable at x and*

$$D(G \circ F)(x) = DG(y)DF(x).$$

Theorem B.1.1 *Let $F : X \to Y$ be continuously differentiable on an open subset V of X and $x \in V$. If $DF(x)$ is an isomorphism, then there exist an open neighbourhood $O(x) \subseteq V$ of x and a differentiable inverse function $F^{-1} : F(O(x)) \to O(x)$, where*

$$DF^{-1}(y) = [DF(x)]^{-1} \quad \text{and} \quad x = F^{-1}(y).$$

B.2 Fréchet derivatives of some matrix functions

In the following we will be interested in some types of matrix functions on normed vector spaces $(S^n, \|\cdot\|)$, $(L^n, \|\cdot\|)$ and $(U^n, \|\cdot\|)$. For two vector spaces V, W we will denote $L(V, W)$ the space of linear operators mapping $V \to W$.

Example B.2.1 *Let $\mathbf{A} \in R^{n \times n}$ and $\mathbf{B} \in S^n$. Consider the following linear maps:*

$$\begin{aligned} &\psi_1 : L^n \to U^n, \quad \psi_1(\mathbf{L}) = \mathbf{L}^T, \\ &\psi_2 : S^n \to S^n, \quad \psi_2(\mathbf{X}) = \mathbf{AXA}^T, \\ &\psi_3 : S^n \to S^n, \quad \psi_3(\mathbf{X}) = \mathbf{BX} + \mathbf{XB}. \end{aligned}$$

Obviously

$$D\psi_1 : L^n \to L(L^n, U^n), \quad D\psi_2 : S^n \to L(S^n, S^n), \quad D\psi_3 : S^n \to L(S^n, S^n)$$

and for fixed $\mathbf{L} \in L^n$ and $\mathbf{X} \in S^n$

$$\begin{aligned} &D\psi_1(\mathbf{L}) : L^n \to U^n, \quad D\psi_1(\mathbf{L})[\mathbf{H}] = \mathbf{H}^T \\ &D\psi_2(\mathbf{X}) : S^n \to S^n, \quad D\psi_2(\mathbf{X})[\mathbf{H}] = \mathbf{AHA}^T \\ &D\psi_3(\mathbf{X}) : S^n \to S^n, \quad D\psi_3(\mathbf{X})[\mathbf{H}] = \mathbf{BH} + \mathbf{HB} \end{aligned}$$

We turn our attention to the nonlinear maps:

$$\begin{array}{ll} \phi_1 : S^n \to S^n, & \phi_1(\mathbf{X}) = \mathbf{X}^2, \\ \phi_2 : L^n \to S^n, & \phi_2(\mathbf{L}) = \mathbf{L}\mathbf{L}^T, \\ \phi_3 : S^n_{++} \to S^n(S^n_{++}), & \phi_3(\mathbf{X}) = \mathbf{X}^{\frac{1}{2}}, \\ \phi_4 : S^n_{++} \to L^n(L^n_{++}), & \phi_4(\mathbf{X}) = \mathbf{L}_{\mathbf{X}}. \end{array}$$

These maps are important for an analysis of the symmetrization maps that are used in context of the weighted paths (see Chapter 2).

Proposition B.2.1 *Let $\phi_1 : S^n \to S^n,\ \phi_1(\mathbf{X}) = \mathbf{X}^2$. Then for any $\mathbf{X} \in S^n$ it holds*

$$D\phi_1(\mathbf{X}) : S^n \to S^n, \quad D\phi_1(\mathbf{X})[\mathbf{H}] = \mathbf{X}\mathbf{H} + \mathbf{H}\mathbf{X}.$$

Proof. Obviously, for any $\mathbf{H} \in S^n$

$$\phi_1(\mathbf{X} + \mathbf{H}) - \phi_1(\mathbf{X}) = (\mathbf{X} + \mathbf{H})^2 - \mathbf{X}^2 = \mathbf{X}\mathbf{H} + \mathbf{H}\mathbf{X} + \mathbf{H}^2.$$

The proposition now follows from the fact that the map $\mathbf{H} \to \mathbf{X}\mathbf{H} + \mathbf{H}\mathbf{X}$ is linear and

$$\lim_{\|\mathbf{H}\| \to 0} \frac{\|\mathbf{H}^2\|}{\|\mathbf{H}\|} = 0$$

since from the matrix norm submultiplicativity we have

$$0 \leq \frac{\|\mathbf{H}^2\|}{\|\mathbf{H}\|} \leq \|\mathbf{H}\|.$$

□

Proposition B.2.2 *Let $\phi_2 : L^n \to S^n,\ \phi_2(\mathbf{L}) = \mathbf{L}\mathbf{L}^T$. Then for any $\mathbf{L} \in L^n$ it holds*

$$D\phi_2(\mathbf{L}) : L^n \to S^n, \quad D\phi_2(\mathbf{L})[\mathbf{H}] = \mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T.$$

Proof. Let $\mathbf{H} \in L^n$. Then

$$\phi_2(\mathbf{L} + \mathbf{H}) - \phi_2(\mathbf{L}) = (\mathbf{L} + \mathbf{H})(\mathbf{L} + \mathbf{H})^T - \mathbf{L}\mathbf{L}^T = \mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T + \mathbf{H}\mathbf{H}^T.$$

Again, $\mathbf{H} \to \mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T$ is a linear map $L^n \to S^n$ and

$$\lim_{\|\mathbf{H}\| \to 0} \frac{\|\mathbf{H}\mathbf{H}^T\|}{\|\mathbf{H}\|} = 0$$

since

$$0 \leq \frac{\|\mathbf{H}\mathbf{H}^T\|}{\|\mathbf{H}\|} \leq \|\mathbf{H}^T\| = \|\mathbf{H}\|$$

□

Obviously, the pair of maps $\phi_1|_{S^n_{++}} : \mathbf{X} \mapsto \mathbf{X}^2$ and $\phi_3 : \mathbf{X} \mapsto \mathbf{X}^{\frac{1}{2}}$ as well as $\phi_2|_{L^n_{++}}$: $\mathbf{L} \mapsto \mathbf{L}\mathbf{L}^T$ and $\phi_4 : \mathbf{X} \mapsto \mathbf{L}_\mathbf{X}$ form a pair of inverse maps. Hence we can derive the Fréchet derivatives of ϕ_3, ϕ_4 using Theorem B.1.1. To this aim we will need the following two lemmas which are simple consequences of Proposition A.4.8 and Proposition A.5.5.

Lemma B.2.1 *If* $\mathbf{X} \succ 0$*, then* $D\phi_1(\mathbf{X}) : \mathbf{H} \mapsto \mathbf{XH} + \mathbf{HX}$ *is an isomorphism* $S^n \to S^n$*.*

Lemma B.2.2 *If* $\mathbf{L} \in L^n_{++}$*, then* $D\phi_2(\mathbf{L}) : \mathbf{H} \mapsto \mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T$ *is an isomorphism* $L^n \to S^n$*.*

From Lemma B.2.1 it follows that if $\mathbf{X} \succ 0$ is fixed, then for any $\mathbf{W} \in S^n$ there exists a unique solution $\mathbf{H} \in S^n$ of the equation $\mathbf{XH} + \mathbf{HX} = \mathbf{W}$. Similarly, from Lemma B.2.2 it follows that if $\mathbf{L} \in L^n_{++}$ is fixed, then for any $\mathbf{W} \in S^n$ there exists unique solution $\mathbf{H} \in L^n$ of the equation $\mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T = \mathbf{W}$.

Definition B.2.1 *(a) Let* $\mathbf{X} \succ 0$ *and* $\mathbf{W} \in S^n$ *be fixed. The unique solution* $\mathbf{H} \in S^n$ *of the equation* $\mathbf{XH} + \mathbf{HX} = \mathbf{W}$ *will be denoted as*

$$\mathbf{H} = \langle\langle \mathbf{W} \rangle\rangle_\mathbf{X}\,.$$

(b) Let $\mathbf{L} \in L^n_{++}$ *and* $\mathbf{W} \in S^n$ *be fixed. The unique solution* $\mathbf{H} \in L^n$ *of the equation* $\mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T = \mathbf{W}$ *will be denoted as*

$$\mathbf{H} = [[\mathbf{W}]]_\mathbf{L}.$$

Proposition B.2.3 *Let* $\phi_3 : S^n_{++} \to S^n_{++},\ \phi_3(\mathbf{X}) = \mathbf{X}^{\frac{1}{2}}$*. Then*

$$D\phi_3(\mathbf{X}) : S^n \to S^n, \quad D\phi_3(\mathbf{X})[\mathbf{W}] = \langle\langle \mathbf{W} \rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}\,.$$

Proof. Consider the map

$$\phi_1 : S^n \to S^n, \qquad \phi_1(\mathbf{X}) = \mathbf{X}^2.$$

The cone S^n_{++} is an open subset of S^n and $\phi_1(S^n_{++}) = S^n_{++}$. Let $\mathbf{X} \in S^n_{++}$. According to Lemma B.2.1 the map $D\phi_1(\mathbf{X})$ given by

$$D\phi_1(\mathbf{X}) : S^n \to S^n, \quad D\phi_1(\mathbf{X})[\mathbf{H}] = \mathbf{XH} + \mathbf{HX}.$$

is an isomorphism and hence from Theorem B.1.1 for the map

$$\phi_3 : S^n_{++} \to S^n_{++}, \quad \phi_3(\mathbf{X}) = \mathbf{X}^{\frac{1}{2}}$$

it holds

$$D\phi_3(\mathbf{X}) = [D\phi_1(\mathbf{X}^{\frac{1}{2}})]^{-1}.$$

Therefore $D\phi_3(\mathbf{X})[\mathbf{W}] = \langle\langle \mathbf{W} \rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}$.

□

Proposition B.2.4 *Let* $\phi_4 : S^n_{++} \to L^n_{++},\ \phi_4(\mathbf{X}) = \mathbf{L}_{\mathbf{X}}$. *Then*

$$D\phi_4(\mathbf{X}) : S^n \to L^n, \quad D\phi_4(\mathbf{X})[\mathbf{W}] = [[\mathbf{W}]]_{\mathbf{L}_{\mathbf{X}}}.$$

Proof. Consider the map

$$\phi_2 : L^n \to S^n, \qquad \phi_2(\mathbf{L}) = \mathbf{L}\mathbf{L}^T.$$

The cone L^n_{++} is an open set of L^n and $\phi_2(L^n_{++}) = S^n_{++}$. Let $\mathbf{L} \in L^n_{++}$. According to Lemma B.2.2 the map $D\phi_2(\mathbf{L})$ given by

$$D\phi_2(\mathbf{L}) : L^n \to S^n, \quad D\phi_2(\mathbf{L})[\mathbf{H}] = \mathbf{L}\mathbf{H}^T + \mathbf{H}\mathbf{L}^T.$$

is an isomorphism and hence from Theorem B.1.1 for the map $\phi_4 : S^n_{++} \to L^n_{++}, \phi_4(\mathbf{X}) = \mathbf{L}_{\mathbf{X}}$ it holds

$$D\phi_4(\mathbf{X}) = [D\phi_2(\mathbf{L}_{\mathbf{X}})]^{-1}$$

Therefore $D\phi_4(\mathbf{X})[\mathbf{W}] = [[\mathbf{W}]]_{\mathbf{L}_{\mathbf{X}}}$.

□

Proposition B.2.5 *Let* $\mathbf{A} \in S^n$. *The Fréchet derivative of the maps*

$$\psi_4 : S^n_{++} \to S^n \quad \psi_4(\mathbf{X}) = \mathbf{X}^{\frac{1}{2}}\mathbf{A}\mathbf{X}^{\frac{1}{2}}$$

and

$$\psi_5 : S^n_{++} \to S^n \quad \psi_5(\mathbf{X}) = \mathbf{L}^T_{\mathbf{X}}\mathbf{A}\mathbf{L}_{\mathbf{X}},$$

at $\mathbf{X}$ *is*

$$D\psi_4(\mathbf{X}) : S^n \to S^n, \quad D\psi_4(\mathbf{X})[\mathbf{W}] = \langle\langle \mathbf{W} \rangle\rangle_{\mathbf{X}^{\frac{1}{2}}} \mathbf{A}\mathbf{X}^{\frac{1}{2}} + \mathbf{X}^{\frac{1}{2}}\mathbf{A} \langle\langle \mathbf{W} \rangle\rangle_{\mathbf{X}^{\frac{1}{2}}},$$

$$D\psi_5(\mathbf{X}) : S^n \to S^n, \quad D\psi_5(\mathbf{X})[\mathbf{W}] = [[\mathbf{W}]]^T_{\mathbf{L}_{\mathbf{X}}}\mathbf{A}\mathbf{L}_{\mathbf{X}} + \mathbf{L}^T_{\mathbf{X}}\mathbf{A}[[\mathbf{W}]]_{\mathbf{L}_{\mathbf{X}}}.$$

Proof. Follows from Proposition B.1.1, Propositions B.2.3 and B.2.4.

□

As a consequence of the statements above we obtain the following corollary:

Corollary B.2.1 *Consider the maps*

$$\begin{aligned}
\Phi_1(\mathbf{X},\mathbf{S}) &= (\mathbf{XS}+\mathbf{SX})/2,\\
\Phi_2(\mathbf{X},\mathbf{S}) &= \mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}},\\
\Phi_3(\mathbf{X},\mathbf{S}) &= \mathbf{L}_{\mathbf{X}}^T\mathbf{S}\mathbf{L}_{\mathbf{X}},\\
\Phi_4(\mathbf{X},\mathbf{S}) &= (\mathbf{X}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}+\mathbf{S}^{\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})/2,\\
\Phi_5(\mathbf{X},\mathbf{S}) &= (\mathbf{U}_{\mathbf{S}}^T\mathbf{L}_{\mathbf{X}}+\mathbf{L}_{\mathbf{X}}^T\mathbf{U}_{\mathbf{S}})/2.
\end{aligned}$$

Then

$$D\Phi_1(\mathbf{X},\mathbf{S})[\triangle\mathbf{X},\triangle\mathbf{S}] = \tfrac{1}{2}(\triangle\mathbf{X}\mathbf{S}+\mathbf{S}\triangle\mathbf{X}+\triangle\mathbf{S}\mathbf{X}+\mathbf{X}\triangle\mathbf{S}),$$

$$D\Phi_2(\mathbf{X},\mathbf{S})[\triangle\mathbf{X},\triangle\mathbf{S}] = \langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}\,\mathbf{S}\mathbf{X}^{\frac{1}{2}}+\mathbf{X}^{\frac{1}{2}}\mathbf{S}\,\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}+\mathbf{X}^{\frac{1}{2}}\triangle\mathbf{S}\mathbf{X}^{\frac{1}{2}},$$

$$D\Phi_3(\mathbf{X},\mathbf{S})[\triangle\mathbf{X},\triangle\mathbf{S}] = [[\triangle\mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}}^T\mathbf{S}\mathbf{L}_{\mathbf{X}}+\mathbf{L}_{\mathbf{X}}{}^T\mathbf{S}[[\triangle\mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}}+\mathbf{L}_{\mathbf{X}}{}^T\triangle\mathbf{S}\mathbf{L}_{\mathbf{X}},$$

$$D\Phi_4(\mathbf{X},\mathbf{S})[\triangle\mathbf{X},\triangle\mathbf{S}] = \tfrac{1}{2}\big(\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}\,\mathbf{S}^{\frac{1}{2}}+\mathbf{S}^{\frac{1}{2}}\,\langle\langle\triangle\mathbf{X}\rangle\rangle_{\mathbf{X}^{\frac{1}{2}}}+\langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}}\,\mathbf{X}^{\frac{1}{2}}+\mathbf{X}^{\frac{1}{2}}\,\langle\langle\triangle\mathbf{S}\rangle\rangle_{\mathbf{S}^{\frac{1}{2}}}\big),$$

$$D\Phi_5(\mathbf{X},\mathbf{S})[\triangle\mathbf{X},\triangle\mathbf{S}] = \tfrac{1}{2}\big([[\triangle\mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}}^T\mathbf{U}_{\mathbf{S}}+\mathbf{U}_{\mathbf{S}}{}^T[[\triangle\mathbf{X}]]_{\mathbf{L}_{\mathbf{X}}}+[[\triangle\mathbf{S}]]_{\mathbf{U}_{\mathbf{S}}}^T\mathbf{L}_{\mathbf{X}}+\mathbf{L}_{\mathbf{X}}{}^T[[\triangle\mathbf{S}]]_{\mathbf{U}_{\mathbf{S}}}\big).$$

Appendix C

Asymptotic notation

In this Appendix we introduce basic definitions and properties concerning the asymptotic notation of the matrix functions.

C.1 Definitions

Before we start with defining the asymptotic notations of matrix functions, we recall the well known definitions of the $\mathcal{O}$-notation, Θ-notation and o-notation of real functions. For our purpose it will be sufficient only to consider the real functions $R_{++} \to R$.

Definition C.1.1 *Let $f : R_{++} \to R$, $g : R_{++} \to R$ be real functions. We will write*

$$g(\mu) = \mathcal{O}(f(\mu))$$

if and only if there exists $\mu_0 > 0$ and $\gamma > 0$ such that for all $\mu \in (0, \mu_0\rangle$ it holds

$$|g(\mu)| \leq \gamma |f(\mu)|.$$

Definition C.1.2 *Let $f : R_{++} \to R$, $g : R_{++} \to R$ be real functions. We will write*

$$g(\mu) = o(f(\mu))$$

if and only if

$$\lim_{\mu \to 0} \frac{|g(\mu)|}{|f(\mu)|} = 0.$$

In what follows, we will assume that

$$\mathbf{A} : R_{++} \to R^{k\times l}, \qquad \mathbf{A} : \mu \mapsto \mathbf{A}(\mu)$$

is a matrix function and

$$f : R_{++} \to R_{++}, \qquad f : \mu \mapsto f(\mu)$$

is a real function.

Definition C.1.3 *We will write*

$$\mathbf{A}(\mu) = \mathcal{O}(f(\mu))$$

if and only if $\|\mathbf{A}(\mu)\|_F = \mathcal{O}(f(\mu))$ *, that is, there exists* $\mu_0 > 0$ *and* $\gamma > 0$ *such that for all* $\mu \in (0, \mu_0\rangle$ *it holds*

$$\|\mathbf{A}(\mu)\|_F \leq \gamma f(\mu).$$

Definition C.1.4 *We will write*

$$\mathbf{A}(\mu) = o(f(\mu))$$

if and only if $\|\mathbf{A}(\mu)\|_F = o(f(\mu))$ *that is*

$$\lim_{\mu\to 0} \frac{\|\mathbf{A}(\mu)\|_F}{f(\mu)} = 0.$$

Definition C.1.5 *Assume* $\mathbf{A}(\mu) \in S^n$*. We will write* $\mathbf{A}(\mu) = \Theta(f(\mu))$ *if and only if there exists* $\mu_0 > 0$ *and* $\alpha > 0$ *such that for all* $\mu \in (0, \mu_0\rangle$ *it holds*

$$\frac{1}{\alpha}\mathbf{I} \preceq \frac{\mathbf{A}(\mu)}{f(\mu)} \preceq \alpha\mathbf{I}.$$

Note, that instead $\mathbf{X} \succeq \mathbf{Y}$ one can write $\mathbf{X} - \mathbf{Y} \in S^n_+$. By replacing the symmetric cone S^n_+ by the cone L^n_+ and U^n_+[1] we obtain the following definitions of Θ-notation for lower and upper triangular matrices, respectively.

Definition C.1.6 *Assume* $\mathbf{A}(\mu) \in L^n$*. We will write* $\mathbf{A}(\mu) = \Theta(f(\mu))$ *if and only if there exists* $\mu_0 > 0$ *and* $\alpha > 0$ *such that for all* $\mu \in (0, \mu_0\rangle$ *it holds*

$$\frac{\mathbf{A}(\mu)}{f(\mu)} - \frac{1}{\alpha}\mathbf{I} \in L^n_+ \quad \textit{and} \quad \alpha\mathbf{I} - \frac{\mathbf{A}(\mu)}{f(\mu)} \in L^n_+.$$

Definition C.1.7 *Assume* $\mathbf{A}(\mu) \in U^n$*. We will write* $\mathbf{A}(\mu) = \Theta(f(\mu))$ *if and only if there exists* $\mu_0 > 0$ *and* $\alpha > 0$ *such that for all* $\mu \in (0, \mu_0\rangle$ *it holds*

$$\frac{\mathbf{A}(\mu)}{f(\mu)} - \frac{1}{\alpha}\mathbf{I} \in U^n_+ \quad \textit{and} \quad \alpha\mathbf{I} - \frac{\mathbf{A}(\mu)}{f(\mu)} \in U^n_+.$$

[1]See Appendix A.5.

C.2 Basic properties

Proposition C.2.1 *Let* $\mathbf{A}(\mu) \in R^{n\times n}$ *and* $\mathbf{A}(\mu) = \mathcal{O}(f(\mu))$ *and* $\mathbf{A}_1(\mu)$ *is a square submatrix of* $\mathbf{A}(\mu)$. *Then*

(a) $\mathbf{A}_1(\mu) = \mathcal{O}(f(\mu))$,

(b) $\mathbf{A}_{ij}(\mu) = \mathcal{O}(f(\mu))$,

(c) $tr(\mathbf{A}(\mu)) = \mathcal{O}(f(\mu))$,

(d) If $\mathbf{A}(\mu) \in S_+^n(L_+^n, U_+^n)$, *then* $\lambda_i(\mathbf{A}(\mu)) = \mathcal{O}(f(\mu))$,

(e) If $\mathbf{A}(\mu) \in S_+^n(L_+^n, U_+^n)$, *then* $\det(\mathbf{A}(\mu)) = \mathcal{O}(f(\mu)^n)$.

Proof. The statement (a) follows from the inequality $\|\mathbf{A}_1(\mu)\|_F \le \|\mathbf{A}(\mu)\|_F$. The statement (b) is a special case of the statement (a). Using this statement, we obtain that there exist positive constants γ_i, $i = 1, \dots n$ such that for any i it holds $|\mathbf{A}(\mu)_{ii}| \le c_i f(\mu)$. Denote $\gamma = \sum_{i=1}^n \gamma_i > 0$. The statement (c) then follows from the inequality

$$|tr(\mathbf{A}(\mu))| = |\sum_{i=1}^n \mathbf{A}(\mu)_{ii}| \le \sum_{i=1}^n |\mathbf{A}(\mu)_{ii}| \le (\sum_{i=1}^n \gamma_i) f(\mu) = \gamma f(\mu).$$

Let us prove the statement (d). If $\mathbf{A}(\mu) \in S_+^n$, then its eigenvalues are nonnegative real numbers and therefore for any $i \in \{1, \dots, n\}$ it holds

$$0 \le \lambda_i(\mathbf{A}(\mu)) \le \sum_{j=1}^n \lambda_j(\mathbf{A}(\mu)) = tr(\mathbf{A}(\mu)) \le \gamma f(\mu),$$

according to the statement (c). Finally, the statement (d) implies the statement (e), since $\det(\mathbf{A}(\mu)) = \prod_{i=1}^n \lambda_i(\mathbf{A}(\mu))$.

□

Proposition C.2.2 *Assume* $\mathbf{A}(\mu) \in S_+^n$. *Then*

$$\mathbf{A}(\mu) = \mathcal{O}(f(\mu)) \;\Leftrightarrow\; tr(\mathbf{A}(\mu)) = \mathcal{O}(f(\mu)).$$

Proof. The implication from the left to the right is included in the statement (c) of the Proposition C.2.1. The reverse implication follows from the inequality $\|\mathbf{A}(\mu)\|_F \le tr(\mathbf{A}(\mu))$ (see Proposition A.2.8).

□

Proposition C.2.3 *Assume* $\mathbf{A}(\mu) \in S_+^n(L_+^n, U_+^n)$ *and* $\mathbf{A}(\mu) = \mathcal{O}(f(\mu))$. *Then for sufficiently small* μ *the following implication holds*

$$\exists c \in R: \ c \leq \ln\det \frac{\mathbf{A}(\mu)}{f(\mu)} \ \Rightarrow \ \mathbf{A}(\mu) = \Theta(f(\mu)),$$

according to Definition C.1.5 (Definition C.1.6, Definition C.1.7).

Proof. Let $\mathbf{A}(\mu) \in S_+^n$ and $\mathbf{A}(\mu) = \mathcal{O}(\mu)$. Denote $\tilde{\mathbf{A}}(\mu) = \frac{\mathbf{A}(\mu)}{f(\mu)}$. Obviously $\tilde{\mathbf{A}}(\mu) \in S_+^n$ and $\tilde{\mathbf{A}}(\mu) = \mathcal{O}(1)$. Assume that there exists $c \in R$ such that

$$c \leq \ln\det \tilde{\mathbf{A}}(\mu),$$

which is equivalent to

$$c_1 \leq \det \tilde{\mathbf{A}}(\mu) \tag{C.1}$$

for some positive real number c_1. From Proposition C.2.1 (d) we have that $\lambda_{max}(\tilde{\mathbf{A}}(\mu)) = \mathcal{O}(1)$, that is

$$\lambda_{max}(\tilde{\mathbf{A}}(\mu)) \leq a \tag{C.2}$$

for some $a > 0$. We will show, that there also exists $b > 0$ such that

$$b \leq \lambda_{min}(\tilde{\mathbf{A}}(\mu)).$$

If this is not true, then there exists a sequence $\{\mu_k\} \to 0$ such that

$$\lim_{k\to\infty} \lambda_{min}(\tilde{\mathbf{A}}(\mu_k)) = 0. \tag{C.3}$$

Because of $(C.2)$, the eigenvalues of $\tilde{\mathbf{A}}(\mu_k)$ are bounded above and hence $(C.3)$ yields it holds

$$\lim_{k\to\infty} \det \tilde{\mathbf{A}}(\mu_k) = 0,$$

which contradicts to $(C.1)$. Put $d = \max\{a, 1/b\} > 0$. Then for any $i \in \{1, \ldots, n\}$ it holds

$$\frac{1}{d} \leq \lambda_i(\tilde{\mathbf{A}}(\mu_k)) \leq d,$$

which is equivalent to

$$\frac{1}{d}\mathbf{I} \preceq diag(\lambda_1(\tilde{\mathbf{A}}(\mu_k)), \ldots, \lambda_n(\tilde{\mathbf{A}}(\mu_k))) \preceq d\mathbf{I}$$

or

$$\frac{1}{d}\mathbf{I} \preceq \tilde{\mathbf{A}}(\mu_k) \preceq d\mathbf{I}.$$

The proof for $\mathbf{A}(\mu) \in L_+^n$ or $\mathbf{A}(\mu) \in U_+^n$ can be carried out analogously.

□

Appendix D

Assumptions

In this appendix we review the assumptions needed in this thesis.

Assumption (A1): The matrices $\mathbf{A}_1, \ldots, \mathbf{A}_m$ are linearly independent.

Assumption (A2): $\mathcal{P}^0 \neq \emptyset, \mathcal{D}^0 \neq \emptyset$.

Assumption (A3): The system (2.3) is solvable.

Assumption (A4): For any $j \in \{1, \ldots, 5\}$ let $\triangle b, \triangle \mathbf{C}$ be such that there exists $\mathbf{W}^0 \in \mathcal{W}_j$ and $\mu_0 > 0$ such that the system (3.1) is solvable for $\mathbf{W} = \mathbf{W}^0$ and $\mu = \mu_0$.

Assumption (A5): There exists a strictly complementary optimal solution of the system (2.3).

Note. Recall that Assumption (A1) ensures the one-to-one correspondence between the dual variables y and $\mathbf{S}$. Assumption (A2) is necessary for the theory of interior point methods in SDP. The both assumptions together are equivalent to the fact that the primal and dual optimal solution sets are nonempty and bounded (See [69]). Therefore, Assumption (A3) is weaker then Assumption (A2). However it is sufficient for studying the infeasible weighted paths. Assumption (A4) ensures that the parameters $\triangle b, \triangle \mathbf{C}$ in the system (3.1) are chosen properly, in order to well-define the infeasible weighted central path (see also Definition 3.6.1). Finally, Assumption (A5) is necessary for the asymptotic analysis of the weighted paths.

Bibliography

[1] I. Adler and R. D. C. Monteiro, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Math. Program. **50** (1991), no. 1, 29–51.

[2] F. Alizadeh, *Optimization over positive semi-definite cone; interior-point methods and combinatorial applications*, Advances in Optimization and Parallel Computing (P.M. Pardalos, ed.), Holland, 1992, pp. 1–25.

[3] ———, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM Journal on Optimization **5** (1995), no. 1, 13–51.

[4] F. Alizadeh, J.-P. Haeberly, and M. L. Overton, *Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results*, SIAM Journal on Optimization **8** (1998), 746–768.

[5] R. Bellman and K. Fan, *On systems of linear inequalities in hermitian matrix variables*, Convexity, vol. 7 Proc. Symposia in Pure Mathematics, Amer. Math. Soc., Providence, RI, 1963, pp. 1–11.

[6] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, Studies in Applied Mathematics **15** (1994).

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

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

[9] C. B. Chua and L. Tunçel, *Invariance and efficiency of convex representations*, Math. Program. **111** (2007), no. 1, 113–140.

[10] B. Craven and B. Mond, *Linear programming with matrix variables*, Linear Algebra and Appl. **38** (1981), 73–80.

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

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

[13] ———, *Infeasible start semidefinite programming algorithms via self–dual embeddings*, Topics in Semidefinite and Interior–Point Methods (P. M. Pardalos and H. Wolkowicz, eds.), vol. 18, American Mathematical Society (AMS), Providence, Rhode Island, USA, 1998, pp. 215–236.

[14] ———, *Initialization in semidefinite programming via a self-dual skew-symmetric embedding*, Operations Research Letters **20** (1998), 213–221.

[15] E. deKlerk, *Aspects of semidefinite programming*, Kluwer Academic Publishers, 2002.

[16] J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York, 1960.

[17] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: Sequential unconstrained minimation techniques*, Classics in Applied Mathematics (reprinted in 1990) **4** (1968).

[18] M. Fiedler, *Speciální matice a jejich použitív numerické matematice*, STNL, Praha, 1981.

[19] R. Fletcher, *Semidefinite matrix constraints in optimization*, SIAM Journal on Control and Opt. **23** (1985), 493–513.

[20] D. Goldfarb and K. Scheinberg, *Interior point trajectories in semidefinite programming*, SIAM Journal on Optimization **8** (1998), 871–886.

[21] C. C. Gonzaga, *Path following methods for linear programming*, SIAM Review **34** (1992), 167–227.

[22] O. Güler, *Limiting behavior of weighted central paths in linear programming*, Mathematical Programming **95** (1994), 347–363.

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

[24] ______, *Two simple proofs for analyticity of the central path in linear programming*, Operations Research Letters **28** (2001), no. 1, 9–19.

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

[26] ______, *Semidefinite programming - lecture notes*, 2003.

[27] ______, *Dvadsať rokov moderných metód vnútorného bodu*, Pokroky matematiky a fyziky **49** (2004), 234–244.

[28] M. Halická, E. de Klerk, and C. Roos, *Limiting behavior of the central path in semidefinite optimization*, 2002, citeseer.ist.psu.edu/halicka02limiting.html.

[29] 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.

[30] M. Halická and M. Trnovská, *Weighted interior-point paths in semidefinite programming*, contribution, 22nd European Conference on Operational Research, Prague, 2007.

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

[32] R. Horn and C. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1995.

[33] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.

[34] K. Shäcke, *On the kronecker product*, http://orion.math.uwaterloo.ca/ hwolkowi /henry/reports/kronessaykatrin.ps.

[35] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica **4** (1984), 373–395.

[36] M. Kojima, S. Mizuno, and T. Noma, *Limiting behavior of trajectories generated by continuation method for monotone complementarity problems*, Mathematics of Operations Research **15** (1990), 662–675.

[37] M. Kojima, M. Shida, and S. Shindoh, *Local convergence of predictor-corrector infeasible-interior-point algorithms for sdps and sdlcps*, Mathematical Programming **80** (1998), 129–160.

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

[39] A. N. Kolmogorov and S. V. Fomin, *Základy teorie funkcií a funkcionální analýzy*, STNL, Praha, 1975.

[40] P. Lankaster, *Theory of matrices*, Academic Press, New York - London, 1969.

[41] Z. Lu and R. D. C. Monteiro, *Limiting behavior of the alizadeh-haeberly-overton weighted paths in semidefinite programming*, to appear in Optimization Methods and Software.

[42] ———, *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.

[43] 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.

[44] R. D. C. Monteiro, *Convergence and boundary behavior of the projective scaling trajectories for linear programming*, Mathematics of Operations Research **16** (1991), 842–858.

[45] ———, *Primal-dual path following algorithms for semidefinite programming*, SIAM J. Optim. **7** (1997), 663–678.

[46] R. D. C. Monteiro and J.-S. Pang, *Properties of an interior-point mapping for mixed complementarity problems*, Mathematics of Operations Research **21** (1996), 629–654.

[47] 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.

[48] R. D. C. Monteiro and T. Tsuchiya, *Limiting behavior of the derivatives of the certain trajectories associated with a monotone horizontal linear complementarity problem*, Mathematics of Operations Research **21** (1996), 793–814.

[49] ———, *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming*, SIAM Journal on Optimization **9(3)** (1999), 551–577.

[50] 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.

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

[52] Y. Nesterov and A. Nemirovski, *A general approach to polynomial-time algorithms design for convex programming*, tech. report, Centr. Econ. & Math. Inst., USSR Acad. Sci., Moscow (1988).

[53] ———, *Optimization over positive semidefinite matrices: Mathematical background and user's manual*, Centr. Econ. & Math. Inst., USSR Acad. Sci., Moscow (1990).

[54] ———, *Self-concordant functions and polynomial time methods for convex programming*, tech. report, Centr. Econ. & Math. Inst., USSR Acad. Sci., Moscow (1990).

[55] ———, *Interior point polynomial algorithms in convex programming*, SIAM Publications, Philadelphia, USA (1994).

[56] F. A. Potra and R. Sheng, *On homogenous interior-point algorithms for semidefinite programming*, Optimization Methods and Software **9** (1998), 161–184.

[57] ———, *A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming*, SIAM J. Optim. **8** (1998), 1007–1028.

[58] 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.

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

[60] C. Roos, T. Terlaky, and J.-Ph. Vial, *Theory and algorithms for linear optimization: An interior point approach*, John Wiley & Sons, New York, 1997.

[61] A. Shapiro, *Extremal problems on the set of nonnegative definite matrices*, Linear algebra and Appl. **67** (1985), 7–18.

[62] M. Shida, S. Shindoh, and M. Kojima, *Existence of search directions in interior-point algorithms for the sdp and the monotone sdlcp*, SIAM J. Optim **8** (1998), 387–396.

[63] 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.

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

[65] ______, *On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems*, Numer. Math. **81** (1999), 631–645.

[66] J. F. Sturm and S. Zhang, *Symmetric primal-dual path following algorithms for semidefinite programming*, Applied Numerical Mathematics **29** (1999), 301–315.

[67] ______, *On weighted centers for semidefinite programming*, European Journal of Operations Research **126** (2000), 391–407, citeseer.ist.psu.edu/sturm96weighted.html.

[68] M. J. Todd, *Semidefinite optimization*, Acta Numerica **10** (2001).

[69] M. Trnovská, *Strong duality conditions in semidefinite programming*, Journal of Electrical Engineering **56** (2005), 87–89.

[70] ______, *Weighted central path in semidefinite programming associated with the symmetrization map* $(xs + sx)/2$, Journal of Electrical Engineering **57** (2006), 57–60.

[71] ______, *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.

[72] L. Tunçel and H. Wolkowicz, *Strengthened existence and uniqueness conditions for search directions in semidefinite programming*, Linear algebra and its applications **400** (2005), 31–60.

[73] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Review **38** (1996), no. 1, 49–95.

[74] ______, *Applications of semidefinite programming*, Applied Numerical Mathematics: Transactions of IMACS **29** (1999), no. 3, 283–299.

[75] M. Wechs, *The analyticity of the interior-point-paths at strictly complementary solutions of linear programs*, Optimization Methods and Software **9** (1998), 209–234.

[76] C. Witzgal, P. T. Boggs, and P. D. Domich, *On the convergence behavior of trajectories for linear programming*, Contemporary Mathematics **114** (1990), 161–187.

[77] H. Wolkowicz, *Some applications of optimization in matrix theory*, Linear algebra and Appl. **40** (1981), 101–118.

[78] H. Wolkowicz, R. Saigal, and L. Vandenberghe (eds.), *Handbook of semidefinite programming: Theory, algorithms and applications*, Kluwer Academic Publisher, 2000.

[79] M. H. Wright, *The interior-point revolution in constrained optimization*, High-Performance Algorithms and Software in Nonlinear Optimization (R. DeLeone, A. Murli, P. M. Pardalos, and G. Toraldo, eds.), Kluwer Academic Publishers, 1998, pp. 359–381.

[80] S. J. Wright, *Primal-dual interior point methods*, SIAM, Philadelphia, 1997.

[81] F. Zhang, *Matrix theory: Basic results and techniques*, Springer-Verlag New York Inc., 1999.