2019

Conic Optimization: Optimal Partition, Parametric, and Stability Analysis

Ali Mohammad-Nezhad
Lehigh University, alm413@lehigh.edu

Conic Optimization: Optimal Partition, Parametric, and Stability Analysis

by

Ali Mohammad-Nezhad

A Dissertation
Presented to the Graduate Committee
of Lehigh University
in Candidacy for the Degree of
Doctor of Philosophy
in
Industrial and Systems Engineering

Lehigh University
January 2019

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Ali Mohammad-Nezhad

Conic Optimization: Optimal Partition, Parametric, and Stability Analysis

Date

(Committee Chair), Dissertation Director, Chair

(Must Sign with Blue Ink)

Accepted Date

Committee Members

Dr. Garth Isaak

Dr. Katya Scheinberg

Dr. Tamás Terlaky, Committee chair

Dr. Henry Wolkowicz

Dr. Luis Zuluaga

Acknowledgements

It is my pleasure to thank my PhD adviser, Professor Tamás Terlaky, who was incredibly kind, considerate, and supportive throughout my PhD studies. I wish to extend my gratitude to him for believing in my capabilities and his word of encouragement. I am strongly inspired by his enthusiasm, positive attitude, and his remarkable tenacity in research, discussions and tackling open mathematical questions. I always benefited from his valuable comments and suggestions during our collaborations and conversations.

During my PhD studies, I was exhilarated to collaborate with great faculty members specializing in theoretical and computational aspects of mathematical optimization. I would like to thank Professor Frank Curtis, Professor Gábor Pataki, Professor Robert Freund, Professor Rainer Sinn, and Professor James Renegar for several productive conversations and discussions. I am sincerely indebted to my PhD committee members, Professor Katya Scheinberg, Professor Henry Wolkowicz, Professor Garth Isaak, and Professor Zuluaga for reading through my PhD thesis and making insightful comments. Last but not least, I would like to thank Professor Greg Blekherman and Professor Cynthia Vinzant for inviting me to the summer school on Hyperbolic Polynomials and Optimization and the workshop on Real Algebraic Geometry and Optimization.

My special thanks go to my beloved wife, Sahar Tahernejad, without her support and help I could not have made it through the hard moments of my PhD studies. I am sincerely grateful to my parents, my in-laws, and my brother and sisters Mohammad, Bahareh, and Banafsheh for their love and warm support.

At the end, I would like to acknowledge that our work was supported by Air force Office of Scientific Research (AFOSR) Grant # FA9550-15-1-0222.

Contents

List of Tables

List of Figures

Notation and Symbols

Linear conic optimization

(B, N) The optimal partition of a linear optimization problem

(B, T, N) The optimal partition of a linear complementarity problem

(x^*, y^*, s^*) A maximally complementary optimal solution for linear conic optimization problem

ϵ Perturbation parameter

$\mathbb{V}$ A finite dimensional real vector space

$\mathcal{A}$ A linear transformation by which the equality constraints of a linear conic optimization problem is defined

$\mathcal{D}^*_{\mathrm{LCO}}$ The dual optimal set of a linear conic optimization problem

$\mathcal{D}_{\mathrm{LCO}}$ The dual feasible set of a linear conic optimization problem

$\mathcal{E}$ The domain of the optimal value function

$\mathcal{K}$ A closed convex cone

$\mathcal{P}^*_{\mathrm{LCO}}$ The primal optimal set of a linear conic optimization problem

$\mathcal{P}_{\mathrm{LCO}}$ The primal feasible set of a linear conic optimization problem

$\pi_{\mathrm{LCO}}(.)$ The optimal partition of a parametric linear conic optimization problem

$\varphi(.)$ The optimal value function of a parametric linear conic optimization problem

$L(.)$ A symmetric matrix which characterizes the Jordan product for a symmetric conic optimization problem

P_w The quadratic representation of w for a symmetric conic optimization problem

LCO Linear conic optimization

LCQO Linearly constrained quadratic optimization

LO Linear optimization

NT Nesterov-Todd

General notations

Ker(.) The null space of a matrix or the kernel of a linear operator

$\mathcal{C}(.)$ The cone of critical directions for a nonlinear optimization problem

$\mathcal{L}(.)$ The Lagrangian function of a nonlinear optimization problem

$\mathcal{R}(.)$ The column space of a matrix or the range of a linear operator

$\sigma_{\min}(.)$ The minimum singular value of a matrix

$\varphi(.)$ The optimal value function of a parametric nonlinear optimization problem

$B_r(.)$ An open ball of radius r

IPM Interior point method

LCP Linear complementarity problem

LICQ Linear independence constraint qualification

NLO Nonlinear optimization

Second-order conic optimization

$(\mathcal{B}, \mathcal{N}, \mathcal{R}, \mathcal{T})$ The optimal partition of a second-order conic optimization problem

$(x^*; y^*; s^*)$ A maximally complementary optimal solution for second-order conic optimization problem

$(x^\mu; y^\mu; s^\mu)$ A central solution of second-order conic optimization problem

$\bar{n}$ The total number of variables in a second-order conic optimization problem

$\mathbb{L}_+^{n_i}$ A second-order cone of dimension n_i for $i = 1, \ldots, p$

$\mathcal{D}^*_{\mathrm{SOCO}}$ The dual optimal set of a second-order conic optimization problem

$\mathcal{D}^*_{\mathrm{SOCO}}(.)$ The dual optimal set mapping of a second-order conic optimization problem

$\mathcal{D}_{\mathrm{SOCO}}$ The dual feasible set of a second-order conic optimization problem

$\mathcal{D}_{\mathrm{SOCO}}(.)$ The dual feasible set mapping of a second-order conic optimization problem

$\mathcal{L}_+^{\bar{n}}$ The Cartesian product of p second-order cones

$\mathcal{P}^*_{\mathrm{SOCO}}$ The primal optimal set of a second-order conic optimization problem

$\mathcal{P}^*_{\mathrm{SOCO}}(.)$ The primal optimal set mapping of a second-order conic optimization problem

$\mathcal{P}_{\text{SOCO}}$ The primal feasible set of a second-order conic optimization problem

$\mathcal{P}_{\text{SOCO}}(.)$ The primal feasible set mapping of a second-order conic optimization problem

$\pi_{\text{SOCO}}(.)$ The optimal partition of a parametric second-order conic optimization problem

$\sigma_1, \sigma_2, \sigma_3$ Condition numbers defined for a second-order conic optimization problem

$A := (A_1, \ldots, A_p)$ The coefficient matrix of a second-order conic optimization problem

$c := (c^1; \ldots; c^p)$ The objective vector of a second-order conic optimization problem

n_i The number of variables in the i^{th} second-order cone

p The number of second-order cones in a second-order conic optimization problem

P_i The matrix of the eigenvectors of $L(x^i)$ for the i^{th} second-order cone of a second-order conic optimization problem

SOCO Second-order conic optimization

Semidefinite optimization

$(\mathcal{B}, \mathcal{T}, \mathcal{N})$ The optimal partition of a semidefinite optimization problem

(X^*, y^*, S^*) A maximally complementary optimal solution for semidefinite optimization problem

(X^μ, y^μ, S^μ) A central solution of a semidefinite optimization problem

$\bar{C}$ The fixed perturbation direction for a parametric semidefinite optimization

γ	The exponent of Hölderian error bounds for the distance of an interior solution from the optimal set
κ	The positive condition number of Hölderian error bounds
$\Lambda(.)$	The diagonal matrix of the eigenvalues of a symmetric matrix
$\lambda_{[i]}(.)$	The i^{th} eigenvalue of a symmetric matrix
$\mathbb{S}^n$	The vector space of $n \times n$ symmetric matrices
$\mathbb{S}^n_+$	The cone of $n \times n$ positive semidefinite matrices
$\mathcal{D}^*_{\text{SDO}}$	The dual optimal set of a semidefinite optimization problem
$\mathcal{D}^*_{\text{SDO}}(.)$	The dual optimal set mapping of a semidefinite optimization problem
$\mathcal{D}_{\text{SDO}}$	The dual feasible set of a semidefinite optimization problem
$\mathcal{D}_{\text{SDO}}(.)$	The dual feasible set mapping of a semidefinite optimization problem
$\mathcal{P}^*_{\text{SDO}(.)}$	The primal optimal set mapping of a semidefinite optimization problem
$\mathcal{P}^*_{\text{SDO}}$	The primal optimal set of a semidefinite optimization problem
$\mathcal{P}_{\text{SDO}}$	The primal feasible set of a semidefinite optimization problem
$\mathcal{P}_{\text{SDO}}(.)$	The primal feasible set mapping of a semidefinite optimization problem
$\pi_{\text{SDO}}(.)$	The optimal partition of a parametric semidefinite optimization problem
σ	A condition number defined for a semidefinite optimization problem
A^i	The symmetric coefficient matrices of a semidefinite optimization problem for $i = 1, \ldots, m$

b The right hand side vector of a linear conic optimization problem

C The coefficient matrix of the objective function for a semidefinite optimization problem

$Q := (Q_{\mathcal{B}}, Q_{\mathcal{T}}, Q_{\mathcal{N}})$ An orthonormal basis of the optimal partition for semidefinite optimization

LMI Linear matrix inequality

SDO Semidefinite optimization

Abstract

A linear conic optimization problem consists of the minimization of a linear objective function over the intersection of an affine space and a closed convex cone. In recent years, linear conic optimization has received significant attention, partly due to the fact that we can take advantage of linear conic optimization to reformulate and approximate intractable optimization problems. Steady advances in computational optimization have enabled us to approximately solve a wide variety of linear conic optimization problems in polynomial time. Nevertheless, preprocessing methods, rounding procedures and sensitivity analysis tools are still the missing parts of conic optimization solvers. Given the output of a conic optimization solver, we need methodologies to generate approximate complementary solutions or to speed up the convergence to an exact optimal solution. A preprocessing method reduces the size of a problem by finding the minimal face of the cone which contains the set of feasible solutions. However, such a preprocessing method assumes the knowledge of an exact solution. More importantly, we need robust sensitivity and post-optimal analysis tools for an optimal solution of a linear conic optimization problem. Motivated by the vital importance of linear conic optimization, we take active steps to fill this gap.

This thesis is concerned with several aspects of a linear conic optimization problem, from algorithm through solution identification, to parametric analysis, which have

not been fully addressed in the literature. We specifically focus on three special classes of linear conic optimization problems, namely semidefinite and second-order conic optimization, and their common generalization, symmetric conic optimization. We propose a polynomial time algorithm for symmetric conic optimization problems. We show how to approximate/identify the optimal partition of semidefinite optimization and second-order conic optimization, a concept which has its origin in linear optimization. Further, we use the optimal partition information to either generate an approximate optimal solution or to speed up the convergence of a solution identification process to the unique optimal solution of the problem. Finally, we study the parametric analysis of semidefinite and second-order conic optimization problems. We investigate the behavior of the optimal partition and the optimal set mapping under perturbation of the objective function vector.

Chapter 1

Introduction

In this chapter, we lay the groundwork for our contributions. We provide the preliminary concepts for linear conic optimization (LCO), duality, optimality, nondegeneracy and second-order sufficient conditions. We then narrow down our attention to three special cases of LCO, i.e., symmetric conic optimization (SCO), semidefinite optimization (SDO), and second-order conic optimization (SOCO). We introduce the concept of the optimal partition for LCO, SDO and SOCO, and highlight its application in rounding procedures. Finally, we review some classical results for the sensitivity and the stability of nonlinear optimization (NLO) problems.

1.1 Definitions and notation

For LCO and facial description of a closed convex cone, we use the terminology from [25] and [131]. Throughout this thesis, int(.) and ri(.) stand for the interior and relative interior of a set, respectively, span(.) denotes the linear span of a set, i.e., the minimal subspace of $\mathbb{V}$ which contains the set, cl(.) is the closure of a set, bd(.) denotes the boundary of a set, and Ker(.) and $\mathcal{R}(.)$ serve as the kernel and the range of a linear transformation.

General notations: Let $\mathbb{V}$ be a finite dimensional real vector space endowed with an inner product $\langle ., . \rangle$. We define a cone as a set $\mathcal{D}$ so that if $x \in \mathcal{D}$, then $\lambda x \in \mathcal{D}$ for all $\lambda \geq 0$. Let $\mathcal{K} \subseteq \mathbb{V}$ be a closed convex cone. Then, the *dual* cone of $\mathcal{K}$ is defined as

$$\mathcal{K}^* := \{s : \langle x, s \rangle \geq 0, \quad \text{for all } x \in \mathcal{K}\}.$$

The cone $\mathcal{K}$ is called *self-dual* if $\mathcal{K} = \mathcal{K}^*$. The cone $\mathcal{K}$ is called proper if it is pointed, i.e., $x, -x \in \mathcal{K}$ implies $x = 0$, and if $\text{int}(\mathcal{K}) \neq \emptyset$. The convex cone $\mathcal{K}$ is referred to as a *homogeneous* cone if for every $x, s \in \text{int}(\mathcal{K})$, there exists an invertible linear map $\mathcal{A}$ so that $\mathcal{A}(x) = s$ and $\mathcal{A}(\mathcal{K}) = \mathcal{K}$. The cone $\mathcal{K}$ is *symmetric* if it is both self-dual and homogeneous.

- Let $\mathbb{V} = \mathbb{S}^n$ be the vector space of symmetric matrices endowed with the inner product

  $$\langle X, S \rangle := \text{Trace}(XS), \qquad \forall X, S \in \mathbb{S}^n,$$

  in which

  $$X := \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{12} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{pmatrix}$$

  is a symmetric matrix with real entries, and S is defined analogously. The *positive semidefinite cone* is defined as

  $$\mathbb{S}^n_+ := \left\{X \in \mathbb{S}^n \mid \lambda_{\min}(X) \geq 0\right\},$$

  where $\lambda_{\min}(X)$ denotes the minimum eigenvalue of X.

- Let $\mathbb{V} = \mathbb{R}^n$ endowed with the inner product

  $$\langle x, s \rangle := x^T s, \qquad \forall x, s \in \mathbb{R}^n,$$

  and the Euclidean norm $\|.\|_2 := \sqrt{\langle ., . \rangle}$. The *second-order cone*, or Lorentz cone, is defined as

  $$\mathbb{L}_+^n := \left\{ x := (x_1, \ldots, x_n)^T \in \mathbb{R}^n \mid x_1 \geq \|x_{2:n}\|_2 \right\}.$$

Both the positive semidefinite cone and second-order cone are symmetric.

Within the realm of Euclidean Jordan algebra, see Section A.1, the Frobenius norm of $x \in \mathbb{V}$, induced by the inner product, is denoted by $\|x\|_F$:

$$\|x\|_F := \sqrt{\langle x, x \rangle}.$$

In case that $\mathbb{V} = \mathbb{R}^n$, then $\|.\|_F$ reduces to the l_2 norm. If $\mathbb{V} = \mathbb{R}^{m \times n}$, then $\|.\|_F$ indicates the Frobenius norm of a matrix. For a matrix $A \in \mathbb{R}^{m \times n}$, $\|.\|_2$ stand for the induced 2-norm (spectral norm):

$$\|A\|_2 := \max_{\substack{x \in \mathbb{R}^n \\ \|x\|_2 = 1}} \|Ax\|_2.$$

Additionally, $\sigma_{\min}(A)$ denotes the smallest singular value of A:

$$\sigma_{\min}(A) := \begin{cases} \sqrt{\lambda_{\min}(AA^T)}, & m \leq n, \\ \sqrt{\lambda_{\min}(A^T A)}, & m > n. \end{cases}$$

For a set $\mathcal{D} \subseteq \mathbb{V}$, $\mathcal{D}^\perp$ denotes the orthogonal complement of the linear span of $\mathcal{D}$. Let $\mathcal{D}$ be a nonempty convex set. A face $\mathcal{F}$ of $\mathcal{D}$, denoted by $\mathcal{F} \trianglelefteq \mathcal{D}$, is defined as a nonempty convex subset of $\mathcal{D}$ such that if $x_1, x_2 \in \mathcal{D}$ and $\alpha x_1 + (1 - \alpha) x_2 \in \mathcal{F}$ for some $0 < \alpha < 1$, then we have $x_1, x_2 \in \mathcal{F}$, see e.g., Section 18 in [147]. Let $\mathcal{F}$

be a face of a closed convex cone $\mathcal{K}$. Then the conjugate face of $\mathcal{F}$ is defined as $\mathcal{F}^{\triangle} := \mathcal{F}^{\perp} \cap \mathcal{K}^{*} = \{\hat{x}\}^{\perp} \cap \mathcal{K}^{*}$ for some $\hat{x} \in \mathrm{ri}(\mathcal{F})$, i.e.,

$$\mathcal{F}^{\triangle} := \left\{ s \in \mathcal{K}^{*} \mid \langle \hat{x}, s \rangle = 0 \right\}, \tag{1.1}$$

where $\mathcal{F}^{\triangle} \trianglelefteq \mathcal{K}^{*}$. A face $\mathcal{F}$ of $\mathcal{K}$ is called exposed if $\mathcal{F} = \mathcal{K} \cap \{s\}^{\perp}$ for some $s \in \mathcal{K}^{*}$. If every face of $\mathcal{K}$ is exposed, then $\mathcal{K}$ is referred to as a facially exposed cone, see Section 2.2 in [25]. For a convex set $\mathcal{D} \subseteq \mathcal{K}$, the minimal face of $\mathcal{K}$ containing $\mathcal{D}$ is denoted by $\mathcal{F}_{\mathcal{D}}$, and it is defined as

$$\mathcal{F}_{\mathcal{D}} := \bigcap \{ \mathcal{F} \mid \mathcal{F} \trianglelefteq \mathcal{K},\ \mathcal{D} \subseteq \mathcal{F} \}.$$

For a convex set $\mathcal{D}$, the set of feasible directions at a given point $x \in \mathcal{D}$ is defined as

$$\mathrm{dir}(x, \mathcal{D}) := \{ d \in \mathbb{V} \mid x + td \in \mathcal{D},\ \text{for some } t > 0 \}.$$

Using the set of feasible directions we can define the tangent space as

$$\tan(x, \mathcal{D}) := \mathrm{cl}(\mathrm{dir}(x, \mathcal{D})) \cap -\mathrm{cl}(\mathrm{dir}(x, \mathcal{D})). \tag{1.2}$$

Notations for SDO: For a symmetric matrix $X \in \mathbb{S}^n$, $\lambda_{[i]}(X)$ denotes the i^{th} largest eigenvalue of X so that

$$\lambda_{[1]}(X) \geq \lambda_{[2]}(X) \geq \ldots \geq \lambda_{[n]}(X).$$

Thus, $\lambda_{\max}(X) := \lambda_{[1]}(X)$, $\lambda_{\min}(X) := \lambda_{[n]}(X)$, and $\Lambda(X)$ denotes the diagonal matrix of the eigenvalues of X. For X, $\mathrm{svec} : \mathbb{S}^n \to \mathbb{R}^{n(n+1)/2}$ is a linear transformation which multiplies the off-diagonal entries of a symmetric matrix by $\sqrt{2}$ and stacks the upper triangular part into a vector, i.e.,

$$\mathrm{svec}(X) := \left(X_{11}, \sqrt{2}X_{12}, \ldots, \sqrt{2}X_{1n}, X_{22}, \sqrt{2}X_{23}, \ldots, \sqrt{2}X_{2n}, \ldots, X_{nn} \right)^T.$$

We also define $\text{vec} : \mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$ as the concatenation of the columns of a matrix. The symmetric Kronecker product of any two square matrices K_1 and K_2 is defined as a mapping

$$(K_1 \otimes_s K_2)\,\text{svec}(X) := \frac{1}{2}\,\text{svec}\left(K_2 X K_1^T + K_1 X K_2^T\right).$$

See e.g., [32] for more details. By $\text{dist}(\mathcal{S}_1, \mathcal{S}_2)$ we mean the distance between two subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathbb{R}^n$ having the same dimension, which is defined as

$$\text{dist}(\mathcal{S}_1, \mathcal{S}_2) := \left\| \text{Proj}_{\mathcal{S}_1} - \text{Proj}_{\mathcal{S}_2} \right\|_2,$$

where $\text{Proj}_{\mathcal{S}_1}$ and $\text{Proj}_{\mathcal{S}_2}$ are the orthogonal projections onto the subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, see Section 2.5.3 in [62].

Notations for SOCO: We adopt the notation $(.;.;\ldots;.)$ and $(.,.,\ldots,.)$ to indicate the concatenation and side by side arrangement of column vectors, respectively. For any solution $x \in \mathbb{R}^n$ we define the minimum eigenvalue as

$$\lambda_{\min}(x) := x_1 - \|x_{2:n}\|_2.$$

Let $A_i \in \mathbb{R}^{m\times n_i}$ for $i = 1, \ldots, p$ and $\mathcal{I} \subseteq \{1, \ldots, p\}$ be an index set. Then $|\mathcal{I}|$ denotes the cardinality of $\mathcal{I}$, and both $(A_i)_{i\in\mathcal{I}}$ and $A_{\mathcal{I}}$ represent the matrix composed by matrices A_i for $i \in \mathcal{I}$.

1.2 Linear conic optimization (LCO)

Following the generalization of interior point methods (IPMs) to convex optimization, LCO, as a special case, has received special attention. An LCO problem optimizes a linear objective function over a closed convex cone $\mathcal{K} \subseteq \mathbb{V}$ intersected with affine constraints. Mathematically speaking, a pair of primal-dual LCO problems is written as

$$
\begin{aligned}
&(\mathrm{P_{LCO}}) \qquad z^*_{\mathrm{P_{LCO}}} := \min_x \{\langle c, x\rangle \mid \mathcal{A}x = b,\ x \in \mathcal{K}\},\\
&(\mathrm{D_{LCO}}) \qquad z^*_{\mathrm{D_{LCO}}} := \max_{y,s} \left\{b^T y \mid \mathcal{A}^* y + s = c,\ s \in \mathcal{K}^*\right\},
\end{aligned}
$$

where $c \in \mathbb{V}$, $b \in \mathbb{R}^m$, $\mathcal{A} : \mathbb{V} \to \mathbb{R}^m$ is a linear transformation and $\mathcal{A}^*$ denotes its adjoint. A primal-dual optimal solution of LCO, if there exists any, is denoted by $(\tilde{x}, \tilde{y}, \tilde{s})$.

Assumption 1.2.1. *The linear transformation $\mathcal{A}$ is surjective.*

The duality gap is defined as the difference between $z^*_{\mathrm{P_{LCO}}}$ and $z^*_{\mathrm{D_{LCO}}}$. It directly follows from the primal and dual formulations of LCO that, for any primal-dual feasible pair (x, y, s) the objective value of $(\mathrm{P_{LCO}})$ is greater than or equal to the objective value of $(\mathrm{D_{LCO}})$, since

$$\langle c, x\rangle - b^T y = \langle \mathcal{A}^* y + s, x\rangle - b^T y = \langle x, s\rangle \geq 0.$$

This is so called the *weak duality* property.

Theorem 1.2.1 (Weak Duality Theorem)**.** *For any primal-dual feasible solution (x, y, s) we have $\langle c, x\rangle \geq b^T y$. In particular, $z^*_{\mathrm{P_{LCO}}} \geq z^*_{\mathrm{D_{LCO}}}$. If $\langle c, x\rangle = b^T y$, then (x, y, s) is a primal-dual optimal solution.*

In contrast to linear optimization (LO), there might be a positive duality gap at optimality for LCO, and/or the primal/dual optimal value may not be attained. *Strong duality* holds for $(\mathrm{P_{LCO}})$ and $(\mathrm{D_{LCO}})$ if $z^*_{\mathrm{P_{LCO}}} = z^*_{\mathrm{D_{LCO}}}$, and the optimal value of $(\mathrm{P_{LCO}})$ and $(\mathrm{D_{LCO}})$ are attained. The following lemma is in order.

Lemma 1.2.1. *For the LCO problems* $(\mathrm{P_{LCO}})$ *and* $(\mathrm{D_{LCO}})$*,* (x, y, s) *satisfies*

$$
\begin{aligned}
\mathcal{A}x = b, &\quad x \in \mathcal{K},\\
\mathcal{A}^* y + s = c, &\quad s \in \mathcal{K}^*, \qquad (1.3)\\
\langle x, s\rangle = 0. &
\end{aligned}
$$

if and only if x is an optimal solution for $(\mathrm{P_{LCO}})$ *and* (y, s) *is an optimal solution for* $(\mathrm{D_{LCO}})$*, and* $z^*_{\mathrm{P_{LCO}}} = z^*_{\mathrm{D_{LCO}}}$.

Proof. The proof is immediate from the Weak Duality Theorem and (1.3). $\square$

Let $\mathcal{P}_{\mathrm{LCO}}$ and $\mathcal{D}_{\mathrm{LCO}}$ be the primal and dual feasible sets of LCO as defined below

$$\begin{aligned}
\mathcal{P}_{\mathrm{LCO}} &:= \{x \mid \mathcal{A}x = b, \quad x \in \mathcal{K}\},\\
\mathcal{D}_{\mathrm{LCO}} &:= \{(y, s) \mid \mathcal{A}^* y + s = c, \;\; s \in \mathcal{K}^*\}.
\end{aligned}$$

Furthermore, let $\mathcal{P}^*_{\mathrm{LCO}}$ and $\mathcal{D}^*_{\mathrm{LCO}}$ denote the primal and dual optimal sets, respectively. The next theorem provides sufficient conditions under which strong duality holds and $\mathcal{P}^*_{\mathrm{LCO}}, \mathcal{D}^*_{\mathrm{LCO}} \neq \emptyset$.

Definition 1.2.1. *Interior point condition is said to hold if there exists a primal-dual feasible solution* $(x^\circ, y^\circ, s^\circ)$ *so that* $x \in \mathrm{int}(\mathcal{K})$ *and* $s \in \mathrm{int}(\mathcal{K}^*)$.

The interior point condition is standard in the literature of IPMs for linear and conic optimization [151].

Theorem 1.2.2 (Theorem 2.4.1 in [14], Theorem 5.81 in [23])**.** *Suppose that the interior point condition holds. Then strong duality holds, i.e, the duality gap is zero, and both primal and dual optimal solutions are attained. Furthermore, both* $\mathcal{P}^*_{\mathrm{LCO}}$ *and* $\mathcal{D}^*_{\mathrm{LCO}}$ *are compact sets.*

If $\mathcal{K}$ is a proper cone, then the interior point condition for both primal and dual problems can be checked using a theorem of the alternative, as stated in the following theorem, see also Corollary 2 in [105].

Theorem 1.2.3 (Theorems 3.3.10 and 3.3.11 in [25])**.** *Suppose that $\mathcal{K}$ is a proper cone and* $(\mathrm{D_{LCO}})$ *is feasible. Then* $(\mathrm{D_{LCO}})$ *fails the interior point condition if and* *only if the system*

$$\mathcal{A}x = 0,$$
$$\langle c, x \rangle = 0,$$
$$x \in \mathcal{K},$$

has a nonzero solution. In a similar fashion, assume that $(\mathrm{P_{LCO}})$ *is feasible. Then* $(\mathrm{P_{LCO}})$ *fails the interior point condition if and only if the system*

$$\mathcal{A}^* y \in \mathcal{K}^*,$$
$$\mathcal{A}^* y \neq 0,$$
$$b^T y = 0,$$

has a solution.

In order to guarantee zero duality gap and attainment of the optimal values, we make the following assumption throughout this paper:

Assumption 1.2.2. *The interior point condition holds for both* $(\mathrm{P_{LCO}})$ *and* $(\mathrm{D_{LCO}})$*.*

As a result of Assumption 1.2.2, all $(x, y, s) \in \mathcal{P}^*_{\mathrm{LCO}} \times \mathcal{D}^*_{\mathrm{LCO}}$ satisfy (1.3). The condition $\langle x, s \rangle = 0$ in (1.3) is the optimality condition.

Definition 1.2.2. *A primal-dual optimal solution* $(x^*, y^*, s^*) \in \mathcal{P}^*_{\mathrm{LCO}} \times \mathcal{D}^*_{\mathrm{LCO}}$ *is called maximally complementary if*

$$x^* \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{LCO}}), \quad \textit{and} \quad (y^*, s^*) \in \mathrm{ri}(\mathcal{D}^*_{\mathrm{LCO}}).$$

A maximally complementary optimal solution (x^*, y^*, s^*) *is called strictly complementary if*

$$x^* \in \mathrm{ri}\left(\{s^*\}^\perp \cap \mathcal{K}\right), \quad \textit{or} \quad s^* \in \mathrm{ri}\left(\{x^*\}^\perp \cap \mathcal{K}^*\right).$$

This is equivalent to

$$\mathcal{F}_{s^*}^{\triangle} = \mathcal{F}_{x^*}, \quad or \quad \mathcal{F}_{x^*}^{\triangle} = \mathcal{F}_{s^*}.$$

In this thesis, any maximally complementary optimal solution is indicated by superscript *.

1.2.1 Symmetric conic optimization (SCO)

When $\mathcal{K}$ is a symmetric cone, then LCO problem reduces to a SCO problem. Note that $\mathcal{K} = \mathcal{K}^*$ by the definition of a symmetric cone. In fact, SCO includes LO, SOCO, and SDO along with their complex variants, see Theorem A.1.2.

We define a bilinear map $x \circ s$ as

$$x \circ s := L(x)s. \tag{1.4}$$

which is called Jordan product, and it is characterized by a symmetric matrix $L(x)$, see Definition A.1.1. Then, by Theorem 1.2.2 and [44], (x, y, s) is an optimal solution to the primal and dual SCO problems if and only if (x, y, s) satisfies

$$\begin{aligned} \mathcal{A}x = b, &\quad x \in \mathcal{K}, \\ \mathcal{A}^*y + s = c, &\quad s \in \mathcal{K}, \\ x \circ s = 0, & \end{aligned} \tag{1.5}$$

where $x \circ s = 0$ denotes the complementarity condition. The primal and dual SCO problems are referred to as ($\mathrm{P}_{\mathrm{SCO}}$) and ($\mathrm{D}_{\mathrm{SCO}}$), respectively.

1.2.2 Semidefinite optimization (SDO)

SDO is known as a generalization of LO, where the nonnegative orthant is substituted by the cone of symmetric positive semidefinite matrices. In SDO, one minimizes/maximizes the linear objective function

$$\langle C, X\rangle := \mathrm{Trace}(CX),$$

where C and X are $n \times n$ symmetric matrices, over the intersection of the positive semidefinite cone and a set of affine constraints. Mathematically, an SDO problem is written as

$$(\mathrm{P}_{\mathrm{SDO}}) \qquad \min\left\{\langle C, X\rangle \mid \langle A^i, X\rangle = b_i, \quad i = 1, \ldots, m, \ X \succeq 0\right\},$$

where A^i for $i = 1, \ldots, m$ are $n \times n$ symmetric matrices, $b \in \mathbb{R}^m$, and $X \in \mathbb{S}^n_+$. Alternatively, $X \succeq 0$ indicates that X is positive semidefinite. The dual SDO problem is given by

$$(\mathrm{D}_{\mathrm{SDO}}) \qquad \max\left\{b^T y \mid \sum_{i=1}^{m} y_i A^i + S = C, \ S \succeq 0, \ y \in \mathbb{R}^m\right\}.$$

Let $\mathcal{P}_{\mathrm{SDO}}$ and $\mathcal{D}_{\mathrm{SDO}}$ denote the primal and dual feasible sets, respectively. In light of this notation, the primal and dual optimal sets are denoted by $\mathcal{P}^*_{\mathrm{SDO}}$ and $\mathcal{D}^*_{\mathrm{SDO}}$, respectively. Note that $(\mathrm{P}_{\mathrm{SDO}})$ and $(\mathrm{D}_{\mathrm{SDO}})$ can be represented in an LCO format if we define

$$\mathcal{A}^s := \left(\mathrm{svec}(A^1), \ldots, \mathrm{svec}(A^m)\right)^T.$$

Then the primal and dual problems can be rephrased as

$$\begin{aligned}
&\min\left\{\mathrm{svec}(C)^T \mathrm{svec}(X) \mid \mathcal{A}^s \,\mathrm{svec}(X) = b, \ X \succeq 0\right\},\\
&\max\left\{b^T y \mid (\mathcal{A}^s)^T y + \mathrm{svec}(S) = \mathrm{svec}(C), \ S \succeq 0, \ y \in \mathbb{R}^m\right\}.
\end{aligned}$$

An optimal solution of SDO is denoted by $(\tilde{X}, \tilde{y}, \tilde{S})$. The surjective assumption for the linear map $\mathcal{A}$ reduces to the linear independence of the matrices A^i for $i = 1, \ldots, m$. Analogous to LCO case, it is assumed that the interior point condition holds, i.e., there exists $(X^\circ, y^\circ, S^\circ) \in \mathcal{P}_{\mathrm{SDO}} \times \mathcal{D}_{\mathrm{SDO}}$ with $X^\circ, S^\circ \succ 0$, where

$\succ 0$ means positive definite. The linear independence of A^i guarantees that y is uniquely determined for a given dual solution S, and the interior point condition ensures that the strong duality holds and that both the primal and dual optimal sets are compact. The interior point condition may be assumed w.l.o.g., since any SDO problem can be cast into a self-dual embedding format, for which the interior point condition always holds, see [34] for details.

SDO problems are frequently used in many applications, e.g., control theory, structural optimization, statistics, robust optimization, eigenvalue optimization, pattern recognition, and combinatorial optimization, see [4, 77, 172, 177] for a detailed description of the problems which can be represented as an SDO problem.

Since the interior point condition holds, the system of optimality conditions for $\mathrm{P_{SDO}}$ and $\mathrm{D_{SDO}}$ is a special case of (1.5) which is given by

$$\begin{aligned} \langle A^i, X\rangle &= b_i, \quad i = 1, \ldots, m, \\ \sum_{i=1}^m A^i y_i + S &= C, \\ XS = 0, &\quad X, S \succeq 0, \end{aligned} \tag{1.6}$$

where $XS = 0$ is referred to as the complementarity condition. A solution (X, y, S) which satisfies $XS = 0$ is called complementary. In a similar fashion, the definitions of strict and maximal complementarity can be specialized for $\mathrm{P_{SDO}}$ and $\mathrm{D_{SDO}}$ from Definition 1.2.2.

Definition 1.2.3 (Definition 2.7 in [32])**.** *A primal-dual optimal solution* $(X^*, y^*, S^*) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$ *is called maximally complementary if* $X^* \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}})$ *and* $(y^*, S^*) \in \mathrm{ri}(\mathcal{D}^*_{\mathrm{SDO}})$*. A maximally complementary optimal solution* (X^*, y^*, S^*) *is called strictly complementary if* $X^* + S^* \succ 0$*.*

The strict complementarity condition holds for $\mathrm{P_{SDO}}$ and $\mathrm{D_{SDO}}$ if there exists a strictly complementary optimal solution. Note that strict complementarity may fail in SDO, i.e., an SDO problem might have no strictly complementary optimal solution.

Example 1.2.1. *Consider the following SDO problem from [6]:*

$$A^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad b = (1,\ 0,\ 0)^T.$$

The problem has the unique optimal solution

$$X^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad y^* = (0,\ 0,\ 0)^T, \quad S^* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

which is not strictly complementary.

In light of Lemma 4 in [11], a maximally complementary optimal solution can be equivalently defined as a primal-dual solution for which $\operatorname{rank}(X^* + S^*)$ is maximal over the optimal set. As a result, all $X^* \in \operatorname{ri}(\mathcal{P}^*_{\text{SDO}})$ have the same range space. Analogously, all S^* have identical range spaces, where $(y^*, S^*) \in \operatorname{ri}(\mathcal{D}^*_{\text{SDO}})$, see e.g., Lemma 2.3 in [32] or Lemma 3.1 in [58].

1.2.3 Second-order conic optimization (SOCO)

SOCO problems minimize a linear objective function over the intersection of an affine space and Cartesian product of p second-order cones of dimension n_i, i.e.,

$$\mathcal{L}^{\bar{n}}_+ := \mathbb{L}^{n_1}_+ \times \ldots \times \mathbb{L}^{n_p}_+, \qquad \bar{n} := \sum_{i=1}^{p} n_i,$$

where

$$\mathbb{L}_+^{n_i} = \left\{x^i := (x_1^i, \dots, x_{n_i}^i)^T \in \mathbb{R}^{n_i} \mid x_1^i \geq \|x_{2:n_i}^i\|_2\right\}, \quad i = 1, \dots, p. \tag{1.7}$$

The primal and dual SOCO problems in standard form are represented as

$$\begin{aligned} &(\mathrm{P_{SOCO}}) \qquad \min\left\{c^T x \mid Ax = b,\ x \in \mathcal{L}_+^{\bar{n}}\right\}, \\ &(\mathrm{D_{SOCO}}) \qquad \max\left\{b^T y \mid A^T y + s = c,\ s \in \mathcal{L}_+^{\bar{n}}\right\}, \end{aligned}$$

where $b \in \mathbb{R}^m$, $A := (A_1, \dots, A_p)$[1], $x := (x^1; \dots; x^p)$, $s := (s^1; \dots; s^p)$, and $c := (c^1; \dots; c^p)$, in which $A_i \in \mathbb{R}^{m \times n_i}$, $s^i \in \mathbb{R}^{n_i}$, and $c^i \in \mathbb{R}^{n_i}$ for $i = 1, \dots, p$. Notice that x, s, and c are concatenation of the column vectors x^i, s^i, and c^i, respectively. An optimal solution of SOCO, if there exists any, is denoted by $(\tilde{x}; \tilde{y}; \tilde{s})$.

A wide range of applications in engineering, control, robust optimization, and combinatorial optimization can be modeled as SOCO problems, see e.g., [5, 100] for the applications of SOCO.

From an algebraic point of view, SOCO can be embedded in an SDO problem using the following equivalence between a second-order cone and a positive semidefinite cone:

$$L(x^i) := \begin{pmatrix} x_1^i & (x_{2:n_i}^i)^T \\ x_{2:n_i}^i & x_1^i I_{n_i - 1} \end{pmatrix} \succeq 0 \ \Leftrightarrow\ x^i \in \mathbb{L}_+^{n_i} \ \Leftrightarrow\ (x^i)^T R_i x^i \geq 0, \quad x_1^i \geq 0, \tag{1.8}$$

where R_i is an $n_i \times n_i$ diagonal matrix given by

$$R_i := \operatorname{diag}\big(1,\ -1, \dots, -1\big). \tag{1.9}$$

The surjective assumption for the linear map $\mathcal{A}$ for LCO is equivalent to a full row rank A matrix. Further, by the interior point condition, there exists a feasible

[1] The reader should differentiate the rectangular matrix A in $(\mathrm{P_{SOCO}})$ from the linear transformation $\mathcal{A}$ for LCO and the symmetric matrix A^i defined for SDO.

solution $(x^\circ; y^\circ; s^\circ)$ such that for all $i = 1, \dots, p$ we have $(x^\circ)^i, (s^\circ)^i \in \mathrm{int}(\mathbb{L}_+^{n_i})$, where

$$\mathrm{int}(\mathbb{L}_+^{n_i}) := \left\{x^i \in \mathbb{R}^{n_i} \mid x_1^i > \|x_{2:n_i}^i\|_2\right\}, \quad i = 1, \dots, p.$$

As a result, at optimality the duality gap is zero, and the optimal value of $(\mathrm{P}_{\mathrm{SOCO}})$ as well as that of $(\mathrm{D}_{\mathrm{SOCO}})$ is attained. Since strong duality holds, the optimal set for $(\mathrm{P}_{\mathrm{SOCO}})$ and $(\mathrm{D}_{\mathrm{SOCO}})$ can be represented as

$$\begin{aligned} Ax = b, &\quad x \in \mathcal{L}_+^{\bar{n}}, \\ A^T y + s = c, &\quad s \in \mathcal{L}_+^{\bar{n}}, \\ x \circ s = 0, & \end{aligned} \tag{1.10}$$

in which $x \circ s = 0$ denotes the complementarity condition, where

$$x \circ s := (x^1 \circ s^1; \dots; x^p \circ s^p),$$

and the Jordan product is defined as

$$x^i \circ s^i = L(x^i)s^i = L(s^i)x^i = \begin{pmatrix} (x^i)^T s^i \\ x_1^i s_{2:n_i}^i + s_1^i x_{2:n_i}^i \end{pmatrix}, \qquad i = 1, \dots, p, \tag{1.11}$$

as demonstrated in Example A.1.1. Any solution $(x; y; s)$ satisfying $x \circ s = 0$ is called complementary. Let $\mathcal{P}_{\mathrm{SOCO}}^*$ and $\mathcal{D}_{\mathrm{SOCO}}^*$ denote the primal and dual optimal sets, respectively. By the interior point condition, both $\mathcal{P}_{\mathrm{SOCO}}^*$ and $\mathcal{D}_{\mathrm{SOCO}}^*$ are nonempty and compact.

In light of Lemma 7 in [21] and Definition 1.2.2, the concepts of strict and maximal complementarity can be specialized for SOCO.

Definition 1.2.4 (Definition 23 in [5], Definition 5 in [21]). *Let $(x^*; y^*; s^*) \in \mathcal{P}_{\mathrm{SOCO}}^* \times \mathcal{D}_{\mathrm{SOCO}}^*$. Then $(x^*; y^*; s^*)$ is called strictly complementary if*

$$x^* + s^* \in \mathrm{int}(\mathcal{L}_+^{\bar{n}}).$$

An optimal solution $(x^*; y^*; s^*)$ *is called maximally complementary if* $x^* \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SOCO}})$ *and* $(y^*; s^*) \in \mathrm{ri}(\mathcal{D}^*_{\mathrm{SOCO}})$.

Equivalently, a primal-dual optimal solution $(x^*; y^*; s^*)$ is maximally complementary if $x^* + s^*$ has maximal number of second-order cones i for which $(x^*)^i + (s^*)^i \in \mathrm{int}(\mathbb{L}^{n_i}_+)$. Under the interior point condition, a primal-dual maximally complementary optimal solution always exists for $(\mathrm{P}_{\mathrm{SOCO}})$ and $(\mathrm{D}_{\mathrm{SOCO}})$, but a strictly complementary optimal solution may not exist.

1.3 Nondegeneracy conditions for LCO

In this section, we briefly review the nondegeneracy conditions for LCO from [131]. The nondegeneracy conditions for SDO, SOCO, and SCO have been studied in [6, 8, 45]. The nondegeneracy conditions for nonlinear SDO have been worked out in [23].

Using the facial description of the feasible set, we can define the primal and dual nondegeneracy conditions for LCO. Recall that $\mathcal{F}_x$ denotes the minimal face of the corresponding cone which contains $\{x\}$. The same is analogously defined for s.

Definition 1.3.1. *Let* (x, y, s) *be a primal-dual feasible solution for* $(\mathrm{P}_{\mathrm{LCO}})$ *and* $(\mathrm{D}_{\mathrm{LCO}})$. *Then* x *satisfies primal nondegeneracy condition if* $\mathrm{span}(\mathcal{F}_x^{\triangle}) \cap \mathcal{R}(\mathcal{A}^*) = \{0\}$. *Furthermore,* (y, s) *satisfies the dual nondegeneracy condition if* $\mathrm{span}(\mathcal{F}_s^{\triangle}) \cap \mathrm{Ker}(\mathcal{A}) = \{0\}$.

It is easy to verify that the primal nondegeneracy condition is equivalent to

$$(\mathcal{F}_x^{\triangle})^{\perp} + \mathrm{Ker}(\mathcal{A}) = \mathbb{V}. \tag{1.12}$$

The concept of primal nondegeneracy simply means that the tangent spaces at a solution x to the primal feasible set and the cone $\mathcal{K}$ span $\mathbb{V}$. In a similar manner, we can show that the dual nondegeneracy condition can be equivalently written as

$$(\mathcal{F}_s^{\triangle})^{\perp} + \mathcal{R}(\mathcal{A}^*) = \mathbb{V}. \tag{1.13}$$

Assume that the interior point condition holds for ($\mathrm{P_{LCO}}$) and ($\mathrm{D_{LCO}}$). The next theorem states the relationship between nondegeneracy conditions and the uniqueness of the optimal solution.

Theorem 1.3.1. *Let $(\bar{x}, \bar{y}, \bar{s})$ be a primal-dual optimal solution. If $\bar{x}$ is primal nondegenerate, then $(\bar{y}, \bar{s})$ is a unique dual optimal solution. Similarly, if $(\bar{y}, \bar{s})$ is a nondegenerate dual optimal solution, then $\bar{x}$ is a unique primal optimal solution. If $(\bar{x}, \bar{y}, \bar{s})$ is a strictly complementary optimal solution, then the converse of both statements are true.*

The primal or dual nondegeneracy condition is not necessary for the existence of a unique dual or primal optimal solution, respectively, see Example 5.92 in [23].

1.3.1 Nondegeneracy conditions for SDO

The expression (1.12) can be specialized for ($\mathrm{P_{SDO}}$) by noting that $\big(\mathrm{span}(\mathcal{F}_X^{\triangle})\big)^{\perp} = (\mathcal{F}_X^{\triangle})^{\perp}$, and that $\mathbb{S}_+^n$ is a *nice* cone, see (1.2) and Lemma 3.2.1 in [131]. Then we get

$$(\mathcal{F}_X^{\triangle})^{\perp} = \tan(X, \mathbb{S}_+^n).$$

Therefore, $X \in \mathcal{P}_{\mathrm{SDO}}$ is primal nondegenerate if

$$\tan(X, \mathbb{S}_+^n) + \mathrm{Ker}(\mathcal{A}) = \mathbb{S}^n.$$

Similarly, $(y, S) \in \mathcal{D}_{\mathrm{SDO}}$ is dual nondegenerate if

$$\tan(S, \mathbb{S}_+^n) + \mathcal{R}(\mathcal{A}^*) = \mathbb{S}^n,$$

where $\mathcal{A}$ is a linear transformation analogously defined for SDO.

The tangent spaces of X and S can be characterized using their spectral decomposition. Let $(X, y, S) \in \mathcal{P}_{\text{SDO}} \times \mathcal{D}_{\text{SDO}}$ be a primal-dual feasible solution as

$$X := M\Lambda(X)M^T, \qquad S := N\Lambda(S)N^T,$$

where $n_X := \operatorname{rank}(X)$ and $n_S := \operatorname{rank}(S)$, and $M \in \mathbb{R}^{n\times n}$ and $N \in \mathbb{R}^{n\times n}$ are orthogonal matrices. Note that M and N might be non-unique up to the sign of their columns. Let $M := (M_1,\ M_2)$ and $N := (N_1,\ N_2)$ where M_1 corresponds to the n_X positive eigenvalues of X and N_2 is associated with the n_S positive eigenvalues of S. Then the tangent space to the cone of positive semidefinite matrices at X is represented by

$$\tan(X, \mathbb{S}^n_+) := \left\{ M \begin{pmatrix} U & V \\ V^T & 0 \end{pmatrix} M^T : U \in \mathbb{S}^{n_X},\ V \in \mathbb{R}^{n_X \times (n-n_X)} \right\}.$$

Analogously, the tangent space to the cone of positive semidefinite matrices at (y, S) is given by

$$\tan(S, \mathbb{S}^n_+) := \left\{ N \begin{pmatrix} 0 & V \\ V^T & W \end{pmatrix} N^T : W \in \mathbb{S}^{n_S},\ V \in \mathbb{R}^{(n-n_S)\times n_S} \right\}.$$

Then the algebraic definition of a tangent space to the cone of $\mathbb{S}^n$ can be employed to characterize more convenient conditions for primal and dual nondegeneracy of a feasible solution.

Lemma 1.3.1 (Theorems 6 and 9 in [6])**.** *Let $(X, y, S) \in \mathcal{P}_{\text{SDO}} \times \mathcal{D}_{\text{SDO}}$ be a primal-dual feasible solution. Then X is primal nondegenerate if and only if the matrices*

$$\begin{pmatrix} M_1^T A^i M_1 & M_1^T A^i M_2 \\ M_2^T A^i M_1 & 0 \end{pmatrix}, \qquad i = 1, \ldots, m,$$

are linearly independent in $\mathbb{S}^n$. A dual feasible solution (y, S) is dual nondegenerate if and only if the matrices

$$N_1^T A^i N_1, \qquad i = 1, \dots, m$$

span $\mathbb{S}^{n-\mathrm{rank}(S)}$.

In case that M or N is non-unique, then any orthogonal matrix of X or S can be used to check the nondegeneracy of the feasible solution.

1.3.2 Nondegeneracy conditions for SOCO

Using the definition (1.12) and the niceness property of the second-order cone, the primal and dual nondegeneracy conditions can be specialized for ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$), as presented in [5, 8]. A primal feasible solution x is nondegenerate if

$$\tan(x^1, \mathbb{L}_+^{n_1}) \times \ldots \times \tan(x^p, \mathbb{L}_+^{n_p}) + \mathrm{Ker}(A) = \mathbb{R}^{\bar{n}}. \tag{1.14}$$

Analogously, a dual feasible solution (y, s) is nondegenerate if

$$\tan(s^1, \mathbb{L}_+^{n_1}) \times \ldots \times \tan(s^p, \mathbb{L}_+^{n_p}) + \mathcal{R}(A^T) = \mathbb{R}^{\bar{n}}. \tag{1.15}$$

Depending on x^i, the tangent space $\tan(x^i, \mathbb{L}_+^{n_i})$ has different characterization. If $x^i = 0$, then $\tan(x^i, \mathbb{L}_+^{n_i}) = \{0\}$, and for $x^i \in \mathrm{int}(\mathbb{L}_+^{n_i})$, we have $\tan(x^i, \mathbb{L}_+^{n_i}) = \mathbb{R}^{n_i}$. Finally, for $x^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}$ the tangent space is the orthogonal complement of the minimal face of $\mathbb{L}_+^{n_i}$ which contains $R_i x^i$, where R_i is defined in (1.9). From the spectral decomposition (A.1), we can realize that $x^i := \lambda_2^i p_2^i$, where p_2^i corresponds to λ_2^i the positive eigenvalue of x^i and $\{p_1^i, p_2^i\}$ denotes a Jordan frame. Hence, for any $x^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}$ we have

$$\tan(x^i, \mathbb{L}_+^{n_i}) = \big\{d \mid (p_1^i)^T d = 0\big\}.$$

The tangent space for s^i is defined analogously.

The existence of a primal (dual) nondegenerate optimal solution implies the uniqueness of the dual (primal) optimal solution. If the strict complementarity condition holds, then the converse is true as well, see Theorem 22 in [5]. This is a special case of Theorem 1.3.1.

1.4 IPMs and central path for SCO

In recent years, IPMs have been effectively tailored to solve SCO problems. The first idea of IPMs goes back to the work of Frisch [51] who suggested using logarithmic barrier functions in LO. Then, IPMs were extensively studied for NLO problems in the 1960's by Fiacco and McCormick [47]. Karmarkar revived the interest in IPMs by his polynomial-time algorithm for LO [92]. The extension of IPMs from LO to SDO was done independently by Alizadeh [4] and Nesterov and Nemirovskii [123]. Nesterov and Nemirovskii [123] proved that the theoretical efficiency of IPMs is maintained when a so-called self-scaled cone (which is identical to a symmetric cone [69]) replaces the nonnegative orthant, see also [140]. Many variants of IPMs have been introduced for SDO based on how the search direction and the neighborhood of the central path is defined. To name a few, we can mention the AHO [7], HRVW/KSH/M [78, 95, 120] and the Nesterov-Todd (NT) search directions [124, 125]. A search direction using a least squares solution of an overdetermined system was proposed by Kruk et al. [98]. de Klerk et al. [33] showed that a scaling scheme can be used to prove the polynomial time convergence of Gauss-Newton method.

The study of primal-dual IPMs for SCO problems was introduced by Nesterov and

Todd [124, 125] for LO problems over self-scaled cones. Faybusovich [44, 46] invoked Euclidean Jordan algebras to analyze a variety of search directions for SCO. Sturm [168, 169] established the underlying theory of his SeDuMi software in the context of Euclidean Jordan algebras. Schmieta [153] and Schmieta and Alizadeh [154–156] used the Euclidean Jordan algebraic framework to extend the analysis of the Monteiro-Zhang family [186] to all symmetric cones. Rangarajan [139] and Gu et al. [67] applied the Euclidean Jordan algebras in their analysis of infeasible IPMs. Tsuchiya [173, 174] studied various search directions of IPMs for SOCO using a Jordan algebra, see also [1, 8, 119]. For an excellent survey of IPMs for LO, SDO and SOCO, see [118, 121, 134, 136]. Further details on the application of Newton-type methods in IPMs can be found in Gonzaga [63] and Hertog [37].

Primal-dual path-following IPMs deal with a relaxation of (1.5) by replacing the complementarity condition by $x \circ s = \mu e$ with $\mu > 0$ as given below

$$\begin{aligned} \mathcal{A}x = b, &\quad x \in \operatorname{int}(\mathcal{K}), \\ \mathcal{A}^* y + s = c, &\quad s \in \operatorname{int}(\mathcal{K}), \\ x \circ s = \mu e, & \end{aligned} \tag{1.16}$$

where $x \circ s = \mu e$ is called the centrality condition, and e denotes the identity element, see the beginning part of Appendix A.1. Assuming that the interior point condition holds and $\mathcal{A}$ is surjective, for all $\mu > 0$ this system of equations has a unique solution (x^μ, y^μ, s^μ), which is called a central solution. The trajectory of the central solutions is known as the central path of an LCO problem [44]. Notice that at a central solution we have

$$\langle x^\mu, s^\mu \rangle := \operatorname{Trace}(x^\mu \circ s^\mu) = \operatorname{Trace}(\mu e) = r\mu,$$

where r denotes the order of the symmetric cone $\mathcal{K}$, see Theorem A.1.1. For SDO and SOCO, r is equal to n and $2p$, respectively. Applying the Newton method to the system (1.16) gives

$$
\begin{aligned}
\mathcal{A}\Delta x &= 0, \\
\mathcal{A}^*\Delta y + \Delta s &= 0, \\
x \circ \Delta s + s \circ \Delta x &= \mu e - x \circ s.
\end{aligned} \tag{1.17}
$$

Note that even if $x \in \operatorname{int}(\mathcal{K})$ and $s \in \operatorname{int}(\mathcal{K})$, the system (1.17) is not necessarily well-defined. For instance, the coefficient matrix in SOCO case might be singular [134]. An effective way to get around this problem is to scale ($\mathrm{P_{SCO}}$) and ($\mathrm{D_{SCO}}$) to project x and s on the same point. This is known as the NT scaling scheme [124, 125] which is as follows. For strictly feasible solutions x and s, there exists a unique $w \in \operatorname{int}(\mathcal{K})$ so that [46]

$$
v := P_w^{-\frac{1}{2}} x = P_w^{\frac{1}{2}} s, \tag{1.18}
$$

where $P_w := 2L(w^2) - L(w)^2$ is called the quadratic representation of w, see the beginning part of the Appendix A.1, and w itself is defined as

$$
w := \left[P_{s^{-\frac{1}{2}}} (P_{s^{\frac{1}{2}}} x)^{\frac{1}{2}} \right]^{-\frac{1}{2}} = \left[P_{x^{\frac{1}{2}}} (P_{x^{\frac{1}{2}}} s)^{-\frac{1}{2}} \right]^{-\frac{1}{2}}. \tag{1.19}
$$

It follows from (1.19) and Lemma A.1.5 that $x, s \in \operatorname{int}(\mathcal{K})$ implies $w \in \operatorname{int}(\mathcal{K})$, and the latter implies non-singularity of P_w. Thus, Lemma A.1.5 implies that $v \in \operatorname{int}(\mathcal{K})$. Now, by Part 2 of Lemma A.1.1 and using simple algebraic manipulations, we can verify that the scaled Newton system is given by

$$
\begin{aligned}
\bar{\mathcal{A}} d_x &= 0, \\
\bar{\mathcal{A}}^* \Delta y + d_s &= 0, \\
d_x + d_s &= \mu v^{-1} - v,
\end{aligned} \tag{1.20}
$$

where

$$
\begin{aligned}
d_x &:= P_w^{-\frac{1}{2}} \Delta x, \\
d_s &:= P_w^{\frac{1}{2}} \Delta s, \\
\bar{\mathcal{A}} &:= \mathcal{A} P_w^{\frac{1}{2}}.
\end{aligned} \tag{1.21}
$$

Note that d_x and d_s belong to the null space and row space of $\bar{\mathcal{A}}$, respectively. All this implies that the Newton system (1.20) uniquely determines d_x and d_s as the orthogonal components of $\mu v^{-1} - v$. The duality gap for the scaled problem is given by

$$\langle x, s\rangle = \operatorname{Trace}(x \circ s) = \operatorname{Trace}(P_w^{\frac{1}{2}} v \circ P_w^{-\frac{1}{2}} v) = \operatorname{Trace}(v^2) = \|v\|_F^2.$$

1.4.1 The central path for SDO

Analogous to LO, SDO problems can be solved in polynomial time using IPMs, though they require significantly more computational effort per iteration. From (1.16), the central path for SDO is simplified to the set of solutions of

$$\begin{aligned} \langle A^i, X\rangle &= b_i, \qquad i = 1, \dots, m, \\ \sum_{i=1}^{m} A^i y_i + S &= C, \\ XS &= \mu I_n, \\ X, S &\succeq 0, \end{aligned} \tag{1.22}$$

where $XS = \mu I_n$ is called the centrality condition, and I_n denotes the identity matrix of size n. For any given $\mu > 0$, the central solution (X^μ, y^μ, S^μ) to this system exists, and it is uniquely defined under the interior point condition and the linear independence of A^i for $i = 1, \dots, m$, see Theorem 3.1 in [32]. It readily follows from the centrality condition that X^μ and S^μ commute, and thus they have a common eigenvector basis. For $0 \leq \mu \leq \bar{\mu}$, where $\bar{\mu} > 0$, the set of solutions of (1.22) is bounded, see Lemma 3.2 in [32], and the trajectory of the central solutions has accumulation points in the relative interior of the optimal set, see e.g., Theorem 3.4 in [32]. A proof was given by [76] for the fact that the central path converges to a maximally complementary optimal solution.

The analyticity and limiting behavior of the central path for SDO have been extensively studied in the literature, see [2, 68, 74, 162, 163] for the analyticity results of LO and LCP. Luo et al. [106] established the superlinear convergence of an IPM for SDO under the strict complementarity condition and a condition for the size of the neighborhood of the central path. The convergence of the central path to the so called analytic center of the optimal set was established by Luo et al. [106] and de Klerk et al. [34] (see also Theorem 3.5 and Example 3.1 in [32]) under the strict complementarity condition. Halická et al. [76] showed that the convergence of the central path to the analytic center of the optimal set is not guaranteed when the strict complementarity condition fails. Here is the counterexample.

Example 1.4.1 (Section 2 in [76])**.** *Consider the following SDO problem:*

$$A^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A^4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad b = (1,\ 0,\ 0,\ 0)^T,$$

for which the primal optimal set can be represented as

$$X^* = \begin{pmatrix} 1 - x_{22} & x_{12} & 0 & 0 \\ x_{12} & x_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \succeq 0.$$

As $\mu \to 0$ *the central path converges to*

$$X^{**} = \begin{pmatrix} 2/5 & 0 & 0 & 0 \\ 0 & 3/5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

while the analytic center of the primal optimal set is given by

$$X^{a} = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Goldfarb and Scheinberg [58] showed, under the strict complementarity and primal-dual nondegeneracy conditions, that the first order derivatives of the central path converge as $\mu \to 0$. However, the first order derivatives may be unbounded if strict complementarity fails to hold. Using the strict complementarity condition only, Halická [75] showed the extension of the analyticity of the central path to $\mu = 0$.

1.4.2 The central path for SOCO

By (1.8), SOCO problems are polynomially solvable using an IPM for SDO. However, a direct implementation of IPMs for SOCO is proven to be more efficient in terms of computational complexity than IPMs applied to the equivalent SDO formulation [123].

Let $e_i := (1; \mathbf{0})$ denote the identity vector for the i^{th} second-order cone, and $e :=$

$(e_1; \dots; e_p)$. Then for $\mu > 0$ the central path can be phrased as

$$\begin{aligned} Ax = b, &\quad x \in \operatorname{int}(\mathcal{L}_+^{\bar{n}}), \\ A^T y + s = c, &\quad s \in \operatorname{int}(\mathcal{L}_+^{\bar{n}}), \\ x \circ s = \mu e. & \end{aligned} \tag{1.23}$$

Under the rank and the interior point conditions, for all $\mu > 0$ system (1.23) has a unique solution $(x^\mu; y^\mu; s^\mu)$, where $x^\mu, s^\mu \in \operatorname{int}(\mathcal{L}_+^{\bar{n}})$. For $\mu > 0$ the set of central solutions forms a smooth analytical curve which converges to a maximally complementary optimal solution, see [123] or Corollary 3.5 in [171]. Unless the strict complementarity condition holds, the central path may not converge to the analytic center of the optimal set. A SOCO counterexample has been provided in [76].

1.5 The optimal partition for LCO

The concept of the optimal partition was originally introduced for LO and linear complementarity problems (LCPs), where $\mathbb{V} = \mathbb{R}^n$ and $\mathcal{K} = \mathbb{R}^n_+$. For LO, the Goldman-Tucker theorem [60] proves the existence of strictly complementary solutions, and hence the partition of the index set of the variables to two disjoint complementary sets. From the complementarity condition for LO, which reduces to $x_j s_j = 0$ for $j = 1, \dots, n$, it can be seen that for every optimal solution $(\tilde{x}; \tilde{y}; \tilde{s})$, either of $\tilde{x}_j$ or $\tilde{s}_j$ should be zero. Further, there always exists a strictly complementary optimal solution, i.e., an optimal solution with $\tilde{x}_j + \tilde{s}_j > 0$ for every $j = 1, \dots, n$. Then the optimal partition is defined as the two disjoint sets B and N:

$$\begin{aligned} B &:= \big\{ j \in \{1, \dots, n\} \mid \tilde{x}_j > 0, \text{ for some primal optimal solution } \tilde{x} \big\}, \\ N &:= \big\{ j \in \{1, \dots, n\} \mid \tilde{s}_j > 0, \text{ for some dual optimal solution } (\tilde{y}; \tilde{s}) \big\}, \end{aligned}$$

where $B \cup N = \{1, \dots, n\}$. We can use the optimal partition to characterize the primal and dual optimal sets as

$$\mathcal{P}^*_{\text{LO}} := \big\{x \mid Ax = b,\ x_j \geq 0,\ \forall j \in B,\ \ x_j = 0,\ \ \forall j \in N\big\},$$
$$\mathcal{D}^*_{\text{LO}} := \big\{(y; s) \mid A^T y + s = c,\ s_j \geq 0,\ \forall j \in N,\ \ s_j = 0,\ \ \forall j \in B\big\}.$$

For LCPs [85], only maximally complementary solutions exist. In fact, there might exist a third set T defined as

$$T := \{1, \dots, n\} \setminus B \cup N.$$

We have both $\tilde{x}_j = 0$ and $\tilde{s}_j = 0$ for $j \in T$ for every optimal solution of the LCP. Goldfarb and Scheinberg [58] extended the concept of the optimal partition to SDO, and Yildirim [184] made a generalization for LCO with self-dual cones. Bonnans and Ramírez [21] established another algebraic definition of the optimal partition for SOCO. Peña and Roshchina [133] extended the idea of the complementarity partition for a linear system to a homogeneous convex conic system comprising of regular closed convex cones.

The concept of the optimal partition is well-defined only when strong duality[2] holds. Yildirim [184] derived a facial description of the optimal partition for (P_{LCO}) and (D_{LCO}). In this extended definition, the optimal partition is defined using a face of the convex cone $\mathcal{K}$ and its conjugate.

Lemma 1.5.1 (Proposition 3.2.4 in [25])**.** *Let* $(x^*, y^*, s^*) \in \operatorname{ri}(\mathcal{P}^*_{\text{LCO}} \times \mathcal{D}^*_{\text{LCO}})$*. Then we have*

$$\mathcal{F}_{x^*} \subseteq \mathcal{F}^{\triangle}_{s^*}, \qquad \mathcal{F}_{s^*} \subseteq \mathcal{F}^{\triangle}_{x^*}.$$

[2]Recall that strong duality here means that both the primal and dual problems admit optimal solutions with equal objective values.

Lemma 1.5.1 also implies that

$$\mathcal{F}_{\mathcal{P}^*_{\text{LCO}}} \subseteq \mathcal{F}^{\triangle}_{\mathcal{D}^*_{\text{LCO}}}, \qquad \mathcal{F}_{\mathcal{D}^*_{\text{LCO}}} \subseteq \mathcal{F}^{\triangle}_{\mathcal{P}^*_{\text{LCO}}}.$$

Since $\mathcal{F}^{\triangle}_{x^*}$ and $\mathcal{F}_{s^*}$ are both faces of $\mathcal{K}^*$, then $\mathcal{F}_{x^*}$ is a (proper) face of $\mathcal{F}^{\triangle}_{s^*}$, see Proposition 2.2.2 in [25]. The following lemma plays a central role in the characterization of the optimal partition for LCO, which depends on the self-duality of the cone $\mathcal{K}$. See also [175] for the complementary partition of $\mathcal{K}$ and $\mathcal{K}^*$.

Lemma 1.5.2 (Lemma 2.4 in [184])**.** *For any face $\mathcal{F}$ of the self-dual convex cone $\mathcal{K}$ we have $\mathcal{K} = \mathcal{F} \vee \mathcal{F}^{\triangle}$, where $\mathcal{F} \vee \mathcal{F}^{\triangle}$ denotes the minimal face of $\mathcal{K}$ containing $\mathcal{F}$ and $\mathcal{F}^{\triangle}$.*

From Lemma 1.5.2 the following definition is in order.

Definition 1.5.1. *For a maximally complementary solution (x^*, y^*, s^*) we define the optimal partition as $(\mathcal{F}_{x^*}, \mathcal{F}_{s^*}, \mathcal{G})$, where*

$$\mathcal{G} := (\mathcal{F}_{x^*} \vee \mathcal{F}_{s^*})^{\triangle}.$$

In other words, the self-dual cone $\mathcal{K}$ can be represented as

$$\mathcal{K} = \mathcal{F}_{x^*} \vee \mathcal{F}_{s^*} \vee \mathcal{G}, \tag{1.24}$$

which implies $\mathcal{G} = \{0\}$ if and only if (x^*, y^*, s^*) is a strictly complementary optimal solution for LCO.

In the rest of this section, we show how the optimal partition $(\mathcal{F}_{x^*}, \mathcal{F}_{s^*}, \mathcal{G})$ can be specialized for SDO and SOCO.

1.5.1 The optimal partition for SDO

Let $(X^*, y^*, S^*) \in \operatorname{ri}\left(\mathcal{P}^*_{\text{SDO}} \times \mathcal{D}^*_{\text{SDO}}\right)$ be a maximally complementary optimal solution. Since X^* and S^* commute by the complementarity condition, see (1.6), they are simultaneously diagonalizable, i.e., there exists an orthogonal matrix Q^* so that $X^* = Q^*\Lambda(X^*)(Q^*)^T$ and $S^* = Q^*\Lambda(S^*)(Q^*)^T$, where $\Lambda(X^*)$ and $\Lambda(S^*)$ are diagonal matrices of the eigenvalues. Then $\mathcal{R}(X^*) = \mathcal{R}(Q^*\Lambda(X^*))$ and $\mathcal{R}(S^*) = \mathcal{R}(Q^*\Lambda(S^*))$ indicate that the subspaces $\mathcal{R}(X^*)$ and $\mathcal{R}(S^*)$ are orthogonal, and they are spanned by the eigenvectors associated with the positive eigenvalues of X^* and S^*, respectively. Let us define $\mathcal{B} := \mathcal{R}(X^*)$, $\mathcal{N} := \mathcal{R}(S^*)$, and $\mathcal{T} := \big(\mathcal{R}(X^*) + \mathcal{R}(S^*)\big)^\perp$.

Definition 1.5.2. *The partition ($\mathcal{B}$, $\mathcal{N}$,$\mathcal{T}$) of $\mathbb{R}^n$ is called the optimal partition of an SDO problem.*

It follows from Definition 1.2.3 that $\mathcal{R}(\tilde{X}) \subseteq \mathcal{B}$ and $\mathcal{R}(\tilde{S}) \subseteq \mathcal{N}$ for all $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$. Since the subspaces $\mathcal{B}$ and $\mathcal{N}$ are orthogonal, it is immediate that $\mathcal{T} = \{0\}$ if and only if a strictly complementary solution exists.

We consider $Q := (Q_{\mathcal{B}}, Q_{\mathcal{T}}, Q_{\mathcal{N}})$ as an orthonormal basis partitioned according to the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$. For instance, the columns of Q^* corresponding to the positive eigenvalues of X^* can be chosen as an orthonormal basis for $\mathcal{B}$. In fact, any matrix with orthonormal columns which span $\mathcal{B}$ would be an orthonormal basis for $\mathcal{B}$. Analogously, we can choose the columns of Q^* corresponding to the positive eigenvalues of S^* as an orthonormal basis for $\mathcal{N}$. Since $(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}})$, the optimal partition is invariant with respect to the choice of (X^*, y^*, S^*).

Remark 1.5.1. *If the interior point condition fails for either* $(\mathrm{P_{SDO}})$ *or* $(\mathrm{D_{SDO}})$*, but a primal-dual optimal solution exists, and the duality gap is 0, then the optimal partition of* $(\mathrm{P_{SDO}})$ *and* $(\mathrm{D_{SDO}})$ *can be recovered from the optimal partition of the problem in self-dual embedding format, see [35].*

Theorem 1.5.1 characterizes $\mathcal{P}^*_{\mathrm{SDO}}$ and $\mathcal{D}^*_{\mathrm{SDO}}$. For brevity, we define $n_{\mathcal{B}} := \dim(\mathcal{B})$, $n_{\mathcal{T}} := \dim(\mathcal{T})$, and $n_{\mathcal{N}} := \dim(\mathcal{N})$.

Theorem 1.5.1 (Theorem 2.7 in [32]). *For every primal-dual optimal solution* $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$ *we can represent* $\tilde{X}$ *and* $\tilde{S}$ *as*

$$\tilde{X} = Q_{\mathcal{B}} U_{\tilde{X}} Q_{\mathcal{B}}^T, \qquad \tilde{S} = Q_{\mathcal{N}} U_{\tilde{S}} Q_{\mathcal{N}}^T,$$

where $U_{\tilde{X}} \in \mathbb{S}^{n_{\mathcal{B}}}_+$ *and* $U_{\tilde{S}} \in \mathbb{S}^{n_{\mathcal{N}}}_+$. *If* $n_{\mathcal{B}} > 0$ *and* $\tilde{X} \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}})$, *then there exists* $U_{\tilde{X}} \succ 0$. *Similarly, if* $n_{\mathcal{N}} > 0$ *and* $(\tilde{y}, \tilde{S}) \in \mathrm{ri}(\mathcal{D}^*_{\mathrm{SDO}})$, *then there exists* $U_{\tilde{S}} \succ 0$.

Notice the necessity of the condition $n_{\mathcal{B}} > 0$ or $n_{\mathcal{N}} > 0$ in Theorem 1.5.1. For instance, if $n_{\mathcal{B}} = 0$, then we have $\mathcal{P}^*_{\mathrm{SDO}} = \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}}) = \{0\}$, which implies $U_{\tilde{X}} = 0$.

Remark 1.5.2. *Note that* $Q_{\mathcal{B}} \mathbb{S}^{n_{\mathcal{B}}}_+ Q_{\mathcal{B}}^T$ *is a face of* $\mathbb{S}^n_+$, *see Proposition 2.2.14 in [25], such that* $\mathcal{P}^*_{\mathrm{SDO}} \subseteq Q_{\mathcal{B}} \mathbb{S}^{n_{\mathcal{B}}}_+ Q_{\mathcal{B}}^T$ *by Theorem 1.5.1. We can show that* $Q_{\mathcal{B}} \mathbb{S}^{n_{\mathcal{B}}}_+ Q_{\mathcal{B}}^T$ *is indeed the minimal face. To that end, observe that* $(\mathcal{P}^*_{\mathrm{SDO}})^\perp \cap \mathbb{S}^n_+ = \{X^*\}^\perp \cap \mathbb{S}^n_+$ *is a face of* $\mathbb{S}^n_+$ *containing* $\mathcal{P}^*_{\mathrm{SDO}}$ *for every* $X^* \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}})$, *which is equivalent to* $Q_{\mathcal{T} \cup \mathcal{N}} \mathbb{S}^{n_{\mathcal{T}} + n_{\mathcal{N}}}_+ Q_{\mathcal{T} \cup \mathcal{N}}^T$. *Since the positive semidefinite cone is facially exposed, then it follows from Corollary 2.2.10 in [25] that*

$$\mathcal{F}_{X^*} = (\{X^*\}^\perp \cap \mathbb{S}^n_+)^\triangle = Q_{\mathcal{B}} \mathbb{S}^{n_{\mathcal{B}}}_+ Q_{\mathcal{B}}^T.$$

In a similar manner, we can show that $Q_{\mathcal{N}} \mathbb{S}^{n_{\mathcal{N}}}_+ Q_{\mathcal{N}}^T$ *is the minimal face of* $\mathbb{S}^n_+$ *which contains* $\mathcal{D}^*_{\mathrm{SDO}}$.

Remark 1.5.3. *By the interior point condition, at least one of* $n_{\mathcal{B}}$ *or* $n_{\mathcal{N}}$ *has to be positive. In fact, if* $X^* = 0$ *is the unique primal optimal solution of* $(\mathrm{P}_{\mathrm{SDO}})$, *then any dual feasible solution is also dual optimal. Therefore, by the interior point condition, there exists a dual optimal solution* (y^*, S^*) *where* S^* *is positive definite. Similarly, for a unique dual optimal solution* (y^*, S^*) *with* $S^* = 0$ *there exists a primal optimal solution* X^* *which is positive definite. Consequently, when either* $n_{\mathcal{B}} = 0$ *or* $n_{\mathcal{N}} = 0$ *holds, then there exists an optimal solution which is strictly complementary.*

An orthogonal transformation of $(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}})$ with respect to Q reveals the optimal partition as

$$Q^T X^* Q = \begin{pmatrix} U_{X^*} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q^T S^* Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & U_{S^*} \end{pmatrix},$$

where $U_{X^*} \succ 0$ and $U_{S^*} \succ 0$ if $n_{\mathcal{B}}, n_{\mathcal{N}} > 0$. As a result of Theorem 1.5.1 we have

$$\begin{aligned} Q^T_{\mathcal{T}\cup\mathcal{N}} \tilde{X} Q_{\mathcal{T}\cup\mathcal{N}} &= 0, & \forall\ \tilde{X} &\in \mathcal{P}^*_{\mathrm{SDO}}, \\ Q^T_{\mathcal{B}\cup\mathcal{T}} \tilde{S} Q_{\mathcal{B}\cup\mathcal{T}} &= 0, & \forall\ (\tilde{y}, \tilde{S}) &\in \mathcal{D}^*_{\mathrm{SDO}}, \end{aligned}$$

where $Q_{\mathcal{T}\cup\mathcal{N}} := (Q_{\mathcal{T}}\ Q_{\mathcal{N}})$, and $Q_{\mathcal{B}\cup\mathcal{T}} := (Q_{\mathcal{B}}\ Q_{\mathcal{T}})$.

Let $\Gamma_{\mathcal{B}}$ and $\Gamma_{\mathcal{N}}$ denote the set of all orthonormal bases for $\mathcal{B}$ and $\mathcal{N}$, respectively. The following lemma is in order.

Lemma 1.5.3. *The sets $\Gamma_{\mathcal{B}}$ and $\Gamma_{\mathcal{N}}$ are compact.*

Proof. If $\mathcal{B} = \{0\}$, then the lemma holds trivially. Hence, we can assume that $\mathcal{B} \neq \{0\}$. Then it is known that for a given subspace $\mathcal{B}$, any two orthonormal bases $Q_{\mathcal{B}}$ and $\bar{Q}_{\mathcal{B}}$ are related by $Q_{\mathcal{B}} U = \bar{Q}_{\mathcal{B}}$ for some orthogonal matrix $U \in \mathbb{R}^{n_{\mathcal{B}} \times n_{\mathcal{B}}}$, see e.g., Lemma 2.4 in [32]. The result follows by noting that the set of orthogonal matrices is compact. The compactness of $\Gamma_{\mathcal{N}}$ follows analogously. $\square$

1.5.2 The optimal partition for SOCO

The notion of the optimal partition of LO can be extended to SOCO. Even though a SOCO problem can be embedded in SDO, the optimal partition in SOCO may be more nuanced when it is defined and analyzed directly in the SOCO setting. In SOCO, the index set $\{1, \dots, p\}$ of the second-order cones is partitioned into four subsets $\mathcal{B}, \mathcal{N}, \mathcal{R}$, and $\mathcal{T} := (\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$ as defined in [21]:

$$\begin{aligned}
\mathcal{B} &:= \left\{ i \mid \tilde{x}_1^i > \|\tilde{x}_{2:n_i}^i\|_2, \text{ for some } \tilde{x} \in \mathcal{P}^*_{\text{SOCO}} \right\},\\
\mathcal{N} &:= \left\{ i \mid \tilde{s}_1^i > \|\tilde{s}_{2:n_i}^i\|_2, \text{ for some } \tilde{s} \in \mathcal{D}^*_{\text{SOCO}} \right\},\\
\mathcal{R} &:= \left\{ i \mid \tilde{x}_1^i = \|\tilde{x}_{2:n_i}^i\|_2 > 0,\ \tilde{s}_1^i = \|\tilde{s}_{2:n_i}^i\|_2 > 0, \text{ for some } (\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}} \right\},\\
\mathcal{T}_1 &:= \left\{ i \mid \tilde{x}^i = \tilde{s}^i = 0, \text{ for all } (\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}} \right\},\\
\mathcal{T}_2 &:= \left\{ i \mid \tilde{s}^i = 0, \text{ for all } (\tilde{y};\tilde{s}) \in \mathcal{D}^*_{\text{SOCO}},\ \tilde{x}_1^i = \|\tilde{x}_{2:n_i}^i\|_2 > 0, \text{ for some } \tilde{x} \in \mathcal{P}^*_{\text{SOCO}} \right\},\\
\mathcal{T}_3 &:= \left\{ i \mid \tilde{x}^i = 0, \text{ for all } \tilde{x} \in \mathcal{P}^*_{\text{SOCO}},\ \tilde{s}_1^i = \|\tilde{s}_{2:n_i}^i\|_2 > 0, \text{ for some } (\tilde{y};\tilde{s}) \in \mathcal{D}^*_{\text{SOCO}} \right\}.
\end{aligned}$$

The convexity of the optimal set implies that $\mathcal{B}, \mathcal{N}, \mathcal{R}$, and $\mathcal{T}$ are mutually disjoint and their union is the index set $\{1, \ldots, p\}$. Therefore, it follows from the complementarity condition that for all $(\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}}$, $\tilde{x}^i = 0$ for all $i \in \mathcal{N}$, and $\tilde{s}^i = 0$ for all $i \in \mathcal{B}$, see e.g., Lemma 3.1 in [171]. Additionally, it follows from (1.11) and the complementarity condition that for all $i \in \mathcal{R}$

$$\tilde{x}^i = \tilde{\alpha}^i \begin{pmatrix} 1 \\ \frac{\tilde{x}_{2:n_i}^i}{\|\tilde{x}_{2:n_i}^i\|_2} \end{pmatrix}, \quad \tilde{s}^i = \tilde{\beta}^i \begin{pmatrix} 1 \\ \frac{\tilde{s}_{2:n_i}^i}{\|\tilde{s}_{2:n_i}^i\|_2} \end{pmatrix}, \qquad \frac{\tilde{x}_{2:n_i}^i}{\|\tilde{x}_{2:n_i}^i\|_2} = -\frac{\tilde{s}_{2:n_i}^i}{\|\tilde{s}_{2:n_i}^i\|_2}, \tag{1.25}$$

where $\tilde{\alpha}^i = \tilde{x}_1^i \geq 0$, $\tilde{\beta}^i = \tilde{s}_1^i \geq 0$, and for at least one $(\tilde{x};\tilde{y};\tilde{s})$ we have both $\tilde{\alpha}^i, \tilde{\beta}^i > 0$. Let $(\breve{x};\breve{y};\breve{s}) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}}$ be a primal-dual optimal solution. From the complementarity condition it can be interpreted that if $\breve{x}^i \in \text{int}(\mathbb{L}_+^{n_i})$ holds, then $\tilde{s}^i = 0$ for all $\tilde{s} \in \mathcal{D}^*_{\text{SOCO}}$. Analogously, if $\breve{s}^i \in \text{int}(\mathbb{L}_+^{n_i})$, then $\tilde{x}^i = 0$ for all $\tilde{x} \in \mathcal{P}^*_{\text{SOCO}}$.

Using the concept of the optimal partition, we can conclude that there exists a strictly complementary optimal solution $(x^*; y^*; s^*)$ if and only if $\mathcal{T} = \emptyset$. Otherwise, one has only a maximally complementary optimal solution. Furthermore, an optimal solution $(x^*; y^*; s^*) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}}$ is maximally complementary if

$$\begin{aligned}
(x^*)^i &\in \text{int}(\mathbb{L}_+^{n_i}), & \forall i \in \mathcal{B},\\
(s^*)^i &\in \text{int}(\mathbb{L}_+^{n_i}), & \forall i \in \mathcal{N},\\
(x^*)_1^i, (s^*)_1^i &\neq 0, & \forall i \in \mathcal{R}.
\end{aligned} \tag{1.26}$$

1.6 Second-order sufficient condition

In [21], Bonnans and Ramírez studied a second-order sufficient condition for nonlinear SOCO problems. A general nonlinear SOCO problem can be phrased as

$$
\begin{aligned}
(\mathrm{P}_{\mathrm{NSOCO}}) \quad & \min_{y \in \mathbb{R}^m} \ f(y) \\
& \text{s.t.} \ \ g_i(y) = s^i \in \mathbb{L}_+^{n_i}, \qquad i = 1, \ldots, p,
\end{aligned}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^m \to \mathbb{R}^{n_i}$ are twice continuously differentiable functions. For $(\mathrm{P}_{\mathrm{NSOCO}})$ the first-order optimality conditions are written as

$$
\begin{aligned}
\nabla \mathcal{L}((y;x)) = \nabla f(y) - \sum_{i=1}^{p} \nabla g_i^T(y) x^i = 0,& \\
g_i(y) = s^i \in \mathbb{L}_+^{n_i}, \qquad & i = 1, \ldots, p, \\
x^i \in \mathbb{L}_+^{n_i}, \qquad & i = 1, \ldots, p, \\
s^i \circ x^i = 0, \qquad & i = 1, \ldots, p,
\end{aligned} \tag{1.27}
$$

where $x^i \in \mathbb{R}^{n_i}$ is the Lagrange multiplier associated with $g_i(y)$,

$$
\mathcal{L}((y;x)) := f(y) - \sum_{i=1}^{p} g_i^T(y) x^i
$$

denotes the Lagrangian function of $(\mathrm{P}_{\mathrm{NSOCO}})$, and the bilinear form $\circ : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$ is defined as in (1.11). Any $(\bar{y}; \bar{x})$ satisfying (1.27) is called a stationary solution of $(\mathrm{P}_{\mathrm{NSOCO}})$.

A second-order sufficient condition of $(\mathrm{P}_{\mathrm{NSOCO}})$ has been investigated by Bonnans and Ramírez [21] which relies on the concepts of the tangent cone and the cone of critical directions. The tangent cone to $\mathbb{L}_+^{n_i}$ at x^i is defined as

$$
T_{\mathbb{L}_+^{n_i}}(s^i) = \begin{cases} \mathbb{R}^{n_i} & s^i \in \operatorname{int}(\mathbb{L}_+^{n_i}), \\ \mathbb{L}_+^{n_i}, & s^i = 0, \\ d \in \mathbb{R}^{n_i} : d_{2:n_i}^T s_{2:n_i}^i - d_1 s_1^i \le 0, & s^i \in \operatorname{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}. \end{cases}
$$

Let y be a stationary solution of ($\mathrm{P}_{\mathrm{NSOCO}}$) and $\Upsilon(y)$ be the set of Lagrange multipliers corresponding to y, and assume that $\Upsilon(y) \neq \emptyset$. Then a second-order sufficient condition is satisfied at y if

$$\sup_{x\in\Upsilon(y)} h^T\nabla^2\mathcal{L}((y;x))h + h^T H(y,x)h > 0, \quad \forall h \in \mathcal{C}(y) \setminus \{0\}, \tag{1.28}$$

where

$$H(y,x) = \sum_{i=1}^{p} H^i(y,x),$$

$$H^i(y,x) = \begin{cases} -\frac{x_1^i}{s_1^i}\nabla g_i^T(y)R_i\nabla g_i(y), & s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ \mathbf{0}_{m\times m}, & \text{otherwise}, \end{cases}$$

and $\mathcal{C}(y)$ denotes the cone of critical directions:

$$\mathcal{C}(y) = \begin{cases} h \in \mathbb{R}^m, & \\ \nabla g_i(y)h \in T_{\mathbb{L}_+^{n_i}}(s^i), & x^i = 0, \\ \nabla g_i(y)h = 0, & x^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), \\ (x^i)^T\nabla g_i(y)h = 0, & x^i, s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ \nabla g_i(y)h \in \mathbb{R}_+(x_1^i; -x_{2:n_i}^i) & x^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\},\ s^i = 0, \end{cases}$$

where $\mathbb{R}_+(x_1^i; -x_{2:n_i}^i)$ denotes the set of all nonnegative multiples of $R_i x^i$. A special form of (1.28) is discussed in Section 4.2.2.1.

1.7 Sensitivity and stability analysis

Sensitivity and stability analysis investigates the behavior of optimal solutions and optimal objective value under perturbation in objective function or constraints. Classical results about semicontinuity of the optimal set and the optimal value function date back to 1960's using the set-valued mapping theory [15, 82], see Sections 1.7.3 and 1.7.4. Dantzig et al. [31] as well as Robinson and Day [146] provided sufficient conditions for the continuity of the optimal set mapping. Evans and Gould [41] studied the stability of the feasible set and continuity of the optimal value function. Zlobec et al. [13, 187] identified the region of stability for perturbed convex optimization problems. Continuity and Lipschitz continuity of the optimal value function were established by Hogan [83] and Stern and Topkis [159] under convexity conditions. See also [52, 53, 56] for more continuity results of the optimal value function.

There has been a comprehensive study on the differential stability of the optimal value function and optimal solutions for NLO problems. Danskin [30] provided sufficient conditions for the existence of directional derivatives for the optimal value function of an abstract NLO problem, see Section 1.7.2. Using convexity assumptions and explicit forms for inequality constraints, Hogan [81] derived more specialized results for the existence of directional derivatives of the optimal value function, see also [61]. Gauvin and Tolle [56] derived strong bounds on *Dini*[3] upper and lower derivatives of the optimal value function. Gauvin and Janin [55] proved Hölder and Lipschitz continuity results for the solution of an NLO problem with perturbation along fixed directions. The sensitivity of KKT points was studied by Fiacco [49] and Fiacco and McCormick [47] using the implicit function theorem, see Section 1.7.1. Their analysis was based on linear independence constraint qualification, second-order sufficient condition, and strict complementarity condition. Furthermore, Fiacco [49] showed how to compute/approximate the partial derivatives of a locally optimal solution, see also [18, 48, 88, 89]. Robinson [145] released the strict complementarity condition but imposed a stronger second-order sufficient

[3] See Page 29 in [50] for the definition of Dini upper and lower derivatives.

condition. Kojima [94] removed the dependence on the strict complementarity condition by invoking the degree theory of a continuous map, see e.g., [129]. For convex optimization problems, Dempe [36] established the directional differentiability of optimal solutions under the Slater condition and a strong second-order sufficient condition. The stability of solutions for linear and nonlinear systems of inequalities were studied by Robinson [141, 142]. Robinson [143] derived an implicit function theorem for a *generalized equation*. Further, Robinson [144] introduced the concept of strong regularity to establish the existence and Lipschitz continuity of solutions for generalized equations. He then applied the results to NLO problems. Interestingly, KKT systems and variational inequalities can be represented as generalized equations, see e.g. [42, 149]. We refer the reader to [10, 22, 50] for surveys of classical results.

In the following sections, we concisely review the application of implicit function theorem in the stability analysis of locally optimal solutions. Furthermore, we present classical results for the directional stability of the optimal value function and semicontinuity of feasible and optimal solutions.

1.7.1 Sensitivity of locally optimal solutions

An NLO problem can be formulated as

$$
\begin{aligned}
(\mathrm{P}_{\mathrm{NLO}}) \quad \min \quad & f(x) \\
\text{s.t.} \quad & g_i(x) \leq 0, && i = 1, \ldots, m_1, \\
& h_i(x) = 0, && i = 1, \ldots, m_2, \\
& x \in \mathcal{X},
\end{aligned}
$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is a nonempty open set, and $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m_1$, and $h_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m_2$ are differentiable functions. For a given solution x, we define $\mathcal{I}(x)$ as the index set of active constraints at x, i.e.,

$$\mathcal{I}(x) := \{i \in \{1, \ldots, m_1\} \mid g_i(x) = 0\}.$$

The Jacobian of active constraints at a given solution x is defined as

$$J(x) := \Big([\nabla g_i(x)]_{i \in \mathcal{I}(x)} \quad [\nabla h_i(x)]_{i \in \{1,\ldots,m_2\}} \Big)^T.$$

Definition 1.7.1. *Linear independence constraint qualification (LICQ) holds at x if $J(x)$ is of full row rank.*

For ($\mathrm{P_{NLO}}$) the Lagrangian function is defined as

$$\mathcal{L}((x; u; v)) := f(x) + \sum_{i=1}^{m_1} u_i g_i(x) + \sum_{i=1}^{m_2} v_i h_i(x),$$

where $u \in \mathbb{R}^{m_1}$ and $v \in \mathbb{R}^{m_2}$ denote the Lagrange multipliers. The following theorem states the first-order optimality conditions for ($\mathrm{P_{NLO}}$).

Theorem 1.7.1 (Theorem 4.3.7 in [12])**.** *Let $\bar{x}$ be a feasible solution for* ($\mathrm{P_{NLO}}$)*, and assume that h_i for $i = 1, \ldots, m_2$ are continuously differentiable at $\bar{x}$ and LICQ holds at $\bar{x}$. If $\bar{x}$ is a locally optimal solution for* ($\mathrm{P_{NLO}}$)*, then there exist unique Lagrange multipliers $\bar{u}_i \geq 0$ for $i \in \mathcal{I}(\bar{x})$ and $\bar{v}_i$ for $i = 1, \ldots, m_2$ such that*

$$\nabla \mathcal{L}((\bar{x}; \bar{u}; \bar{v})) = \nabla f(\bar{x}) + \sum_{i \in \mathcal{I}(\bar{x})} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{i=1}^{m_2} \bar{v}_i \nabla h_i(\bar{x}) = 0.$$

From Theorem 1.7.1, the first-order optimality conditions can be equivalently written as

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^{m_1} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{i=1}^{m_2} \bar{v}_i \nabla h_i(\bar{x}) = 0&, \\ \bar{u}_i g_i(\bar{x}) = 0,& \quad i = 1, \ldots, m_1 \\ \bar{u}_i \geq 0,& \quad i = 1, \ldots, m_1, \\ g_i(\bar{x}) \leq 0,& \quad i = 1, \ldots, m_1, \\ h_i(\bar{x}) = 0,& \quad i = 1, \ldots, m_2, \end{aligned} \tag{1.29}$$

which we call KKT conditions. If LICQ holds at $\bar{x}$, then the KKT conditions will be necessary. A solution $(\bar{x}; \bar{u}; \bar{v})$ satisfying (1.29) is called a KKT solution. Furthermore, $(\bar{x}; \bar{u}; \bar{v})$ satisfies the strict complementarity condition if

$$g_i(\bar{x}) + \bar{u}_i > 0, \qquad \forall i = 1, \dots, m. \tag{1.30}$$

For a given KKT solution $(\bar{x}; \bar{u}; \bar{v})$, let $\mathcal{I}_+(\bar{x})$ denote the set of active constraints with $\bar{u}_i > 0$, i.e.,

$$\mathcal{I}_+(\bar{x}) := \{i \in \{1, \dots, m_1\} \mid g_i(\bar{x}) = 0, \text{ and } \bar{u}_i > 0\}.$$

A second-order sufficient condition holds at $\bar{x}$ if there exists Lagrange multipliers $\bar{u}$ and $\bar{v}$ so that (1.29) holds, and

$$z^T \nabla^2 \mathcal{L}((\bar{x}; \bar{u}; \bar{v})) z > 0, \qquad \forall z \in \mathcal{C}((\bar{x}; \bar{u}; \bar{v})), \tag{1.31}$$

where $\nabla^2 \mathcal{L}((\bar{x}; \bar{u}; \bar{v}))$ is called the Hessian of the Lagrangian function and

$$\mathcal{C}((\bar{x}; \bar{u}; \bar{v})) = \Big\{ z \mid \nabla g_i(\bar{x}) z \geq 0, \quad \forall\, i \in \mathcal{I}(\bar{x}) \setminus \mathcal{I}_+(\bar{x}), \quad \nabla g_i(\bar{x}) z = 0, \quad \forall i \in \mathcal{I}_+(\bar{x}), \\ \nabla h_i(\bar{x}) z = 0, \quad i = 1, \dots, m_2 \Big\}$$

is called the cone of critical directions. The second-order sufficient condition (1.31) plays a central role in the local convergence of NLO algorithms [127]. The following theorem provides sufficient conditions for the existence of a locally optimal solution.

Theorem 1.7.2 (Lemma 3.2.1 in [50])**.** *Assume that the functions f, g_i and h_i are twice differentiable in a neighborhood of $\bar{x}$, and there exist Lagrange multipliers $\bar{u}$ and $\bar{v}$ so that* (1.29)*and* (1.31) *hold. Then $\bar{x}$ is a strict locally optimal solution.*

A classical result for the sensitivity of locally optimal solutions can be obtained using the application of the implicit function theorem, see Theorem A.4.3. Indeed, under

LICQ, the second-order sufficient condition (1.31), and the strict complementarity condition (1.30), the implicit function theorem can be applied to the KKT system to characterize the behavior of locally optimal solutions with respect to the simultaneous perturbations in the objective function and the constraints. See Chapter 3 in [50].

Consider the perturbed NLO problem

$$
\begin{aligned}
(\mathrm{P}^{\omega}_{\mathrm{NLO}}) \quad \min \;\; & f(x,\omega), \\
\text{s.t.} \;\; & g(x,\omega) \le 0, \\
& h(x,\omega) = 0,
\end{aligned}
\tag{1.32}
$$

where $\omega \in \mathbb{R}^k$, $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{m_1}$, and $h : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{m_2}$. By the application of the implicit function theorem to the first-order optimality conditions of (1.32), the existence of a continuously differentiable mapping $(\bar{x}(.); \bar{u}(.); \bar{v}(.))$ on a sufficiently small neighborhood of 0 can be proven. For the sake of completeness we provide a proof here.

Theorem 1.7.3 (Theorem 3.2.2 in [50])**.** *Let $\bar{x}$ be a locally optimal solution of $(\mathrm{P}^{0}_{\mathrm{NLO}})$, and assume that f, g, and h are twice continuously differentiable in (x,ω) in a neighborhood of $(\bar{x},0)$. Furthermore, assume that LICQ holds at $\bar{x}$, the second-order sufficient condition (1.31) holds at $(\bar{x};\bar{u};\bar{v})$, and the strict complementarity condition holds. Then,*

- *$\bar{x}$ is an isolated locally optimal solution of $(\mathrm{P}^{0}_{\mathrm{NLO}})$.*
- *For small values of ω there exists a unique solution $\big(\bar{x}(\omega);\bar{u}(\omega);\bar{v}(\omega)\big)$ which satisfies the second-order sufficient condition. Further, $(\bar{x}(.);\bar{u}(.);\bar{v}(.))$ is a continuously differentiable mapping on a sufficiently small neighborhood of 0.*
- *For small values of ω, $(\bar{x}(\omega);\bar{u}(\omega);\bar{v}(\omega))$ is an isolated locally optimal solution of $(\mathrm{P}^{\omega}_{\mathrm{NLO}})$.*

- *For small values of ω, LICQ holds at $\bar{x}(\omega)$, and the strict complementarity condition holds.*

Proof. From the second-order sufficient condition and LICQ it follows that $\bar{x}$ is a strict locally optimal solution which satisfies the first-order optimality conditions of (1.32), and the Lagrange multipliers $(\bar{u}; \bar{v})$ are unique. For the remaining parts we apply the implicit function theorem to the mapping

$$F_{\mathrm{NLO}} : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^k \to \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

which is defined as

$$F_{\mathrm{NLO}}((x; u; v), \omega) := \begin{pmatrix} \nabla_x f(x, \omega) + \sum_{i=1}^{m_1} u_i \nabla_x g_i(x, \omega) + \sum_{i=1}^{m_2} v_i \nabla_x h_i(x, \omega) \\ ug(x, \omega) \\ h(x, \omega) \end{pmatrix},$$

where $ug(x, \omega)$ denotes the coordinatewise product of u and $g(x)$. By the assumptions, F_{NLO} is a continuously differentiable mapping. Furthermore, it follows from LICQ and the second-order sufficient condition that ∇F_{NLO} is nonsingular at $(\bar{x}; \bar{u}; \bar{v})$, see Theorem 14 in [47]. Then by the implicit function theorem, there exists $\varsigma > 0$ and a unique continuously differentiable mapping $(\bar{x}(.); \bar{u}(.); \bar{v}(.))$ on $B_\varsigma(0)$ which satisfies $F_{\mathrm{NLO}}((\bar{x}(\omega); \bar{u}(\omega); \bar{v}(\omega)), \omega) = 0$ for all $\omega \in B_\varsigma(0)$ and $(\bar{x}; \bar{u}; \bar{v}) = (\bar{x}(0); \bar{u}(0); \bar{v}(0))$. By the strict complementarity condition and the continuity of the mapping $\bar{u}(.)$ we have

$$\bar{u}_i(\omega) + g_i(\bar{x}(\omega), \omega) > 0, \qquad i = 1, \ldots, m_1,$$

for all ω sufficiently close to 0. Then together with the complementarity condition $u_i g_i(x, \omega) = 0$ all this means that $(\bar{x}(\omega); \bar{u}(\omega); \bar{v}(\omega))$ is a KKT solution at every ω in a small neighborhood of 0.

Notice that for sufficiently small $\|\omega\|_2$ the active set remains unchanged. In fact, from the complementarity condition and the continuity of both $g_i(x,\omega)$ and $\bar{u}(.)$ we can conclude that

$$\begin{array}{rcccccl} g_i(\bar{x},0)=0 & \Longrightarrow & \bar{u}_i>0 & \Longrightarrow & \bar{u}_i(\omega)>0 & \Longrightarrow & g_i(\bar{x}(\omega),\omega)=0, \\ \bar{u}_i=0 & \Longrightarrow & g_i(\bar{x},0)>0 & \Longrightarrow & g_i(\bar{x}(\omega),\omega)>0 & \Longrightarrow & \bar{u}_i(\omega)=0 \end{array}$$

hold for all $i=1,\dots,m_1$, implying the stability of active set for ω near 0. On the other hand, Since $J(\bar{x})$ has full row rank, the rank of $J(\bar{x}(.))$ stays constant for small ω. Consequently, both the strict complementarity condition and LICQ are valid in a neighborhood of $\omega=0$.

Finally, we show that the second-order sufficient condition (1.31) is stable in a neighborhood of $\omega=0$, i.e.,

$$z(\epsilon)^T\nabla^2\mathcal{L}\big((\bar{x}(\omega);\bar{u}(\omega);\bar{v}(\omega))\big)z(\omega)>0, \qquad \forall z(\omega)\in\mathcal{C}\big((\bar{x}(\omega);\bar{u}(\omega);\bar{v}(\omega))\big),$$

where

$$\begin{aligned}\mathcal{C}\big((\bar{x}(\omega);\bar{u}(\omega);\bar{v}(\omega))\big)=\Big\{z\ |\nabla g_i(\bar{x}(\omega),\omega)z\ge 0, &\qquad \forall i\in\mathcal{I}(\bar{x}(\omega))\setminus\mathcal{I}_+(\bar{x}(\omega)),\\ \nabla g_i(\bar{x}(\omega),\omega)z=0, &\qquad \forall i\in\mathcal{I}_+(\bar{x}(\omega)),\\ \nabla h_i(\bar{x}(\omega),\omega)z=0, &\qquad i=1,\dots,m_2\Big\}.\end{aligned}$$

Suppose that there exist $\omega_k\to 0$ and $z_k\in\mathcal{C}\big((\bar{x}(\omega_k);\bar{u}(\omega_k);\bar{v}(\omega_k))\big)$ so that

$$z_k^T\nabla^2\mathcal{L}\big((\bar{x}(\omega_k);\bar{u}(\omega_k);\bar{v}(\omega_k))\big)z_k\le 0. \tag{1.33}$$

We can assume w.l.o.g. that $\|z_k\|_2=1$, i.e., z_k has an accumulation point $\hat{z}$. Then taking a subsequence which converges to $\hat{z}$ and letting $k\to\infty$ we get

$$\hat{z}^T\nabla^2\mathcal{L}((\bar{x};\bar{u};\bar{v}))\hat{z}\le 0,$$

where $\hat{z}\in\mathcal{C}((\bar{x};\bar{u};\bar{v}))$. However, this contradicts the assumption of second-order sufficient condition. This completes the proof. $\square$

The result of Theorem 1.7.3 can be specialized for an unconstrained NLO problem. More precisely, if $f(x,\omega)$ is a twice continuously differentiable function in x, $\nabla_x f(x,\omega)$ is continuously differentiable in ω and in a neighborhood of $\bar{x}$, and if the Hessian of $f(x,\omega)$ is positive definite at $(\bar{x},0)$, then $\bar{x}(\omega)$ is a locally optimal solution of $f(x,\omega)$ with $\nabla^2 f(\bar{x}(\omega),\omega) \succ 0$ for all ω sufficiently close to 0.

1.7.2 Differential stability of the optimal value function

This section delves into the behavior of the optimal value of $(\mathrm{P}^{\omega}_{\mathrm{NLO}})$ as a function of ω, and its differential properties. Consider an abstract NLO problem where only the objective function is affected by the perturbation, i.e., we have

$$
\begin{aligned}
(\mathrm{P}^{\omega}_{\mathrm{NLO}}) \quad \min \;\; & f(x,\omega) \\
\text{s.t.} \;\; & x \in \Phi,
\end{aligned}
$$

where $\Phi \subseteq \mathbb{R}^n$. Then the optimal value function is defined as

$$
\varphi(\omega) = \min\{f(x,\epsilon) \mid x \in \Phi\},
$$

and

$$
\Psi(\omega) := \{x \in \Phi \mid f(x,\omega) = \varphi(\omega)\} \tag{1.34}
$$

is called the optimal set mapping. A basic stability result was proven by Danskin [30] providing the conditions for the existence of the directional derivatives of $\varphi(.)$. A more general result can be found in Theorem 4.13 in [23]. The continuity of f and its partial derivatives are presumed in the following theorem.

Theorem 1.7.4 (Theorem 1 in [30], Theorem 1.29 in [70])**.** *Consider* $(\mathrm{P}^{\omega}_{\mathrm{NLO}})$*, and assume that* Φ *is a nonempty and compact set. Further, assume that* f *and* $\frac{\partial f}{\partial \omega_i}$ *are* *continuous. Then, at a given ω, the directional derivative of $\varphi(.)$ exists along any direction d, and the directional derivative is given by*

$$D_d\varphi(\omega) := \min_{x\in\Psi(\omega)} \nabla_\omega f(x,\omega)^T d.$$

Theorem 1.7.4 stays valid even if Φ belongs to an abstract space. The result of Danskin's theorem can be applied to the dual of a convex program with right hand side perturbation. A more general result with right hand side perturbation was studied by Hogan [81], where the objective function and constraints are convex and continuously differentiable.

1.7.3 Set-valued analysis

In this section, we briefly review the notions of set convergence, set-valued mapping, and semicontinuity of set-valued mappings from Chapters 4 and 5 in [150] and Chapter 3 in [149]. See also Sections 2.3 and 4.1 in [23], Section 1.5 in [93], Chapter VI in [15], and [82] for further reading.

Let $\mathbb{N}$ be the set of natural numbers, $\mathcal{J}$ be the collection of subsets $J \subset \mathbb{N}$ so that $\mathbb{N}\setminus J$ is finite, and $\mathcal{J}_\infty$ denote the collection of all infinite subsets of $\mathbb{N}$. Furthermore, let $\{\mathcal{D}_k\}_{k=1}^\infty$ be a sequence of subsets of $\mathbb{R}^n$. Then the outer limit of $\{\mathcal{D}_k\}_{k=1}^\infty$ is defined as

$$\limsup_{k\to\infty} \mathcal{D}_k := \{x \mid \exists\ J \in \mathcal{J}_\infty \text{ and } x_k \in \mathcal{D}_k \text{ for } k \in J \text{ such that } \lim_{k\in J} x_k = x\},$$

where $\lim_{k\in J} x_k$ indicates the limit of x_k as $k \to \infty$ and $k \in J$. The inner limit of $\{\mathcal{D}_k\}_{k=1}^\infty$ is defined as

$$\liminf_{k\to\infty} \mathcal{D}_k := \{x \mid \exists\ J \in \mathcal{J} \text{ and } x_k \in \mathcal{D}_k \text{ for } k \in J \text{ such that } \lim_{k\in J} x_k = x\}.$$

If the inner and outer limits coincide, the limit of $\{\mathcal{D}_k\}_{k=1}^{\infty}$ exists and converges to $\mathcal{D}$ in the sense of Painlevé-Kuratowski, i.e.,

$$\lim_{k\to\infty} \mathcal{D}_k := \limsup_{k\to\infty} \mathcal{D}_k = \liminf_{k\to\infty} \mathcal{D}_k = \mathcal{D}.$$

In simple words, when $\mathcal{D}_k \neq \emptyset$, $\limsup\limits_{k\to\infty} \mathcal{D}_k$ denotes the collection of all accumulation points of $\{x_k\}_{k=1}^{\infty}$ such that $x_k \in \mathcal{D}_k$, while $\liminf\limits_{k\to\infty} \mathcal{D}_k$ represents the collection of all limit points of $\{x_k\}_{k=1}^{\infty}$. Alternatively, lim sup and lim inf of $\{\mathcal{D}_k\}_{k=1}^{\infty}$ can be defined in terms of the Euclidean distance:

$$\begin{aligned}
\limsup_{k\to\infty} \mathcal{D}_k &= \{x \mid \liminf_{k\to\infty}\ \mathrm{dist}(x, \mathcal{D}_k) = 0\}, \\
\liminf_{k\to\infty} \mathcal{D}_k &= \{x \mid \limsup_{k\to\infty}\ \mathrm{dist}(x, \mathcal{D}_k) = 0\}.
\end{aligned}$$

It directly follows from the definition that $\liminf_k \mathcal{D}_k \subseteq \limsup_k \mathcal{D}_k$. Furthermore, it turns out that lim sup and lim inf of $\{\mathcal{D}_k\}_{k=1}^{\infty}$ are closed sets, see Section 3.1 in [149].

In what follows, we formally define a set-valued mapping and its semicontinuity properties. Note that a set-valued mapping is referred to as a point-to-set mapping in [50] and a multifunction in [23].

Definition 1.7.2. *A set-valued mapping $\Phi : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ assigns a subset of $\mathbb{R}^n$ to each element of $\omega \in \mathbb{R}^m$.*

The domain of the set-valued mapping Φ is defined as

$$\mathrm{dom}(\Phi) := \{\omega \in \mathbb{R}^m \mid \Phi(\omega) \neq \emptyset\},$$

and its range space is given by

$$\mathcal{R}(\Phi) := \{x \in \mathbb{R}^n \mid x \in \Phi(\omega),\ \text{for some}\ \omega \in \mathbb{R}^m\}.$$

There are various forms of continuity for a set-valued mapping. Here, we adopt the definition from [149, 150] which is in accordance with [82]. For the other definitions of regularity and continuity, see Section 2.2 in [50] and Section 2.2 in [10].

In this thesis, Painlevé-Kuratowski set convergence forms the basis of the semicontinuity/continuity for a set-valued mapping. Let us define

$$\limsup_{\omega \to \bar{\omega}} \Phi(\omega) := \bigcup_{\omega \to \bar{\omega}} \limsup_{k \to \infty} \Phi(\omega_k) = \{x \mid \exists\ \omega_k \to \bar{\omega} \text{ s.t. } \liminf_{k \to \infty} \ \operatorname{dist}(x, \Phi(\omega_k)) = 0\},$$

$$\liminf_{\omega \to \bar{\omega}} \Phi(\omega) := \bigcap_{\omega \to \bar{\omega}} \liminf_{k \to \infty} \Phi(\omega_k) = \{x \mid \forall\ \omega_k \to \bar{\omega} \text{ s.t. } \limsup_{k \to \infty} \ \operatorname{dist}(x, \Phi(\omega_k)) = 0\},$$

or equivalently,

$$\limsup_{\omega \to \bar{\omega}} \Phi(\omega) := \bigcup_{\omega \to \bar{\omega}} \limsup_{k \to \infty} \Phi(\omega_k) = \big\{x \mid \exists\ \omega_k \to \bar{\omega}, \exists\ x_k \to x, \text{with } x_k \in \Phi(\omega_k)\big\},$$

$$\begin{aligned} \liminf_{\omega \to \bar{\omega}} \Phi(\omega) &:= \bigcap_{\omega \to \bar{\omega}} \liminf_{k \to \infty} \Phi(\omega_k) \\ &= \big\{x \mid \forall\ \omega_k \to \bar{\omega},\ \exists\ J \in \mathcal{J},\ \text{s.t. } x_k \to x \text{ with } x_k \in \Phi(\omega_k)\big\}. \end{aligned}$$

Then $\Phi(.)$ is called outer semicontinuous at $\bar{\omega}$ if

$$\limsup_{\omega \to \bar{\omega}} \Phi(\omega) \subseteq \Phi(\bar{\omega}),$$

and inner semicontinuous at $\bar{\omega}$ if

$$\liminf_{\omega \to \bar{\omega}} \Phi(\omega) \supseteq \Phi(\bar{\omega})$$

holds. The set-valued mapping $\Phi(.)$ is Painlevé-Kuratowski continuous at $\bar{\omega}$ if it is both outer and inner semicontinuous at $\bar{\omega}$. We refer the reader to [149, 150] for the proofs.

1.7.4 Continuity of the objective and solution set mapping

In this section, we discuss the stability of the feasible set, optimal set, and the optimal value function for an abstract NLO problem. It is worth investigating the conditions under which the solution set and the optimal objective value of an NLO problem do not change drastically after a slight perturbation. The following instance from [41] exemplifies an unstable behavior of a solution set mapping:

$$\Phi(\omega) = \{x \in \mathbb{R}^n \mid g(x) \leq \omega\},$$

where

$$g(x) = \begin{cases} x^3, & x \leq 0, \\ 0, & 0 < x \leq 1, \\ (x-1)^3, & x > 1, \end{cases}$$

and $\omega \in \mathbb{R}$. One can observe that $\Phi(0) = (-\infty, 1]$. However, for every $\omega < 0$ the solution set changes to $\Phi(\epsilon) = (-\infty, \sqrt[3]{\omega}]$, which does not include the interval $[0, 1]$ anymore.

Consider the following abstract NLO problem

$$\begin{aligned} (\mathrm{P}^{\omega}_{\mathrm{NLO}}) \quad \min \quad & f(x, \omega) \\ \text{s.t.} \quad & x \in \Phi(\omega), \end{aligned}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{-\infty, +\infty\}$, and $\Phi : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is called the feasible set mapping. The optimal set mapping $\Psi(.)$ is defined analogously as in (1.34). The continuity of the optimal set mapping and the optimal value function is dependent on the inner semicontinuity of the feasible set mapping $\Phi(.)$ and the continuity of $f(x, .)$.

Definition 1.7.3. *Let $\varphi : \mathbb{R}^m \to \mathbb{R} \cup \{-\infty, +\infty\}$. Then, φ is called*

- *lower semicontinuous at $\bar{\omega}$ if* $\liminf_{\omega \to \bar{\omega}} \varphi(\omega) \geq \varphi(\bar{\omega})$,
- *upper semicontinuous at $\bar{\omega}$ if* $\limsup_{\omega \to \bar{\omega}} \varphi(\omega) \leq \varphi(\bar{\omega})$,

- *continuous at $\bar{\omega}$ if it is both lower and upper semicontinuous at $\bar{\omega}$.*

A classical result for the semicontinuity of the optimal value function is given by the following theorem.

Theorem 1.7.5 (Theorem 6 in [82]). *Suppose that $\Phi(.)$ is inner semicontinuous at $\bar{\omega}$ and f is lower semicontinuous on $\Phi(\bar{\omega}) \times \bar{\omega}$. Then the optimal value function $\varphi(.)$ is lower semicontinuous at $\bar{\omega}$.*

The set-valued mapping $\Phi(.)$ is called uniformly bounded near $\bar{\omega}$ if there exists $\varsigma > 0$ and a compact set $\mathcal{D} \subset \mathbb{R}^n$ such that

$$\bigcup_{\omega \in B_\varsigma(\bar{\omega})} \Phi(\omega) \subseteq \mathcal{D}. \tag{1.35}$$

The following result is in order.

Theorem 1.7.6 (Theorem 5 in [82]). *Assume that $\Phi(.)$ is outer semicontinuous at $\bar{\omega}$ and uniformly bounded near $\bar{\omega}$, and that f is upper semicontinuous on $\Phi(\bar{\omega}) \times \bar{\omega}$. Then $\varphi(.)$ is outer semicontinuous at $\bar{\omega}$.*

The uniform boundedness is needed to ensure the outer semicontinuity of the optimal value function. For instance, consider the following example from [54], where the feasible set mapping is not uniformly bounded near $\bar{\omega} = 0$:

$$\begin{aligned} \max \quad & x \\ \text{s.t.} \quad & g(x) - \omega \leq 0, \end{aligned}$$

where

$$g(x) = \begin{cases} -(x + \frac{1}{2})^2 + \frac{5}{4}, & x < 0, \\ e^{-x}, & x \geq 0. \end{cases}$$

Then the optimal value function is given by

$$\varphi(\omega) = \begin{cases} g^{-1}(y), & y \leq 0, \\ +\infty, & y > 0, \end{cases}$$

which is not upper semicontinuous at $\bar{\omega} = 0$.

Now, we can resort to the results of Theorems 1.7.5 and 1.7.6 to conclude the continuity of the optimal value function and the outer semicontinuity of the optimal set mapping.

Theorem 1.7.7 (Theorems 7 and 8 in [82])**.** *Suppose that $\Phi(.)$ is continuous at $\bar{\omega}$ and f is continuous on $\Phi(\bar{\omega}) \times \bar{\omega}$. Then $\Psi(.)$ is outer semicontinuous at $\bar{\omega}$. Furthermore, if $\Phi(.)$ is uniformly bounded near $\bar{\omega}$, then $\varphi(.)$ is continuous at $\bar{\omega}$.*

A sufficient condition can be provided to ensure the continuity of the optimal set mapping.

Theorem 1.7.8 (Corollary 8.1 in [82])**.** *Assume that $\Phi(.)$ is continuous at $\bar{\omega}$, f is continuous on $\Phi(\bar{\omega}) \times \bar{\omega}$, $\Phi(.)$ is uniformly bounded near $\bar{\omega}$, and $\Psi(.)$ is single-valued at $\bar{\omega}$. Then $\Psi(.)$ is continuous at $\bar{\omega}$.*

Lipschitz continuity of the optimal set mapping, see Chapter 9 in [150], can be established by imposing stronger regularity conditions, see Section 4 in [23] for a detailed discussion. We investigate the semicontinuity of the optimal set mapping for parametric SDO and SOCO problems in Chapter 5.

1.8 Outline of the thesis

Thus far, we have reviewed the preliminary concepts in LCO, nondegeneracy, optimal partition, and sensitivity analysis. In the rest of this thesis, we present our main contributions, in order, as follows:

In Chapter 2, we generalize the primal-dual Dikin-type affine scaling method from LO to SCO using Euclidean Jordan algebraic tools. A Euclidean Jordan algebra is a commutative algebra over the field of real numbers which is not necessarily associative. A Euclidean Jordan algebra indeed provides the machinery for defining characteristic polynomial and eigenvalues for a symmetric cone. These tools enabled us to generalize Dikin-type search directions and the Dikin ellipsoid to SCO. This generalization has an $\mathcal{O}(\xi r L)$ iteration complexity, where ξ and r denote the measure of proximity and the order of symmetric cone, respectively, and L is the input length. Furthermore, the method's iteration complexity bound is analogous to the SDO case. The Dikin-type algorithm was tested against the SeDuMi, MOSEK and SDPT3 solvers on benchmark SOCO problems.

In the first part of Chapter 3, we investigate the identification of the optimal partition of SDO, for which we provide an approximation from a bounded sequence of solutions on, or in a neighborhood of the central path. We show how the complexity of approximating the optimal partition depends on condition numbers of the problem. Using bounds on the magnitude of the eigenvalues we identify the subsets of eigenvectors of the interior solutions whose accumulation points are orthonormal bases for the subspaces of the optimal partition. The magnitude of the eigenvalues of an interior solution is quantified using a condition number and an upper bound on the distance of an interior solution to the optimal set. We provide a measure of proximity of the approximation obtained from the central solutions to the true optimal partition of the problem.

In the second part of Chapter 3, we revisit the identification of the optimal partition for SOCO from [171]. We reproduce the bounds for the identification of the optimal partition using an error bound result for linear conic systems.

In the first part of Chapter 4, we investigate solution identification for SDO. We use an approximation of the optimal partition in a rounding procedure to generate an approximate maximally complementary solution. The procedure generates a rounded primal-dual solution from an interior solution, sufficiently close to the optimal set, which has approximate primal-dual feasibility and zero duality gap.

In the second part of Chapter 4, we investigate solution identification for SOCO. We establish quadratic convergence of Newton's method to the unique optimal solution of SOCO under both the primal and dual nondegeneracy conditions. Our local convergence result depends on the optimal partition of the problem, which can be identified from a bounded sequence of interior solutions. We provide a theoretical complexity bound for identifying the quadratic convergence region of Newton's method from the trajectory of central solutions. By way of experimentation, we illustrate quadratic convergence of Newton's method on some SOCO problems which fail the strict complementarity condition. At the end, we propose a rounding procedure for an approximate maximally complementary solution of SOCO.

In Chapter 5, we study parametric analysis of SDO and SOCO problems, where the objective function is perturbed along a fixed direction. We introduce the notions of nonlinearity interval and transition point for the optimal partition of the problem. Further, we investigate the continuity of optimal solutions and the behavior of the optimal partition in a nonlinearity interval. For SDO we investigate the sensitivity of the approximation of the optimal partition with respect to the perturbation of the objective vector, and we derive an upper bound on the distance between the invariant subspaces spanned by the approximation of the optimal partition. For SOCO we show how to compute a subinterval of a nonlinearity interval under strict complementarity condition. Additionally, we show how to identify a transition point from the higher-order derivatives of the unique optimal solution, when strict complementarity fails. At the end, we partially extend our derivations and continuity results for LCO.

1.8.1 Technical reports and publications

The generalization of the Dikin-type primal-dual affine scaling algorithm for SCO has been published in the journal Computational Optimization and Applications [113]:

- Chapter 2: A polynomial primal-dual affine scaling algorithm for symmetric conic optimization. *Computational Optimization and Applications* (2017) 66:577-600.

The results for the identification of the optimal partition for SDO and SOCO have been submitted for publication to three peer reviewed journals [112, 114, 115]:

- Chapter 3: On the identification of the optimal partition for semidefinite optimization: Under second round review in *INFOR: Information Systems and Operational Research*.

- Chapter 4: A rounding procedure for semidefinite optimization: To appear in *Operations Research Letters*.

- Chapter 4: Quadratic convergence to the optimal solution of second-order conic optimization without strict complementarity. *Optimization Methods and Software* (2018), DOI: 10.1080/10556788.2018.1528249.

The outcome of the work on the parametric analysis of SDO and SOCO problems are submitted [116], and to be submitted under the following titles:

- Chapter 5: Parametric analysis of semidefinite optimization: Submitted to *Optimization*.

- Chapter 5: On the nonlinearity interval of second-order conic optimization: To be submitted to *SIAM Journal on Optimization*.

Chapter 2

Numerical algorithms for SCO

Dikin's affine scaling method is originally a primal (or dual) method, where each step aims for minimizing the objective function over an ellipsoid inscribed in the primal feasible region. The notion of affine scaling methods were extended to the primal-dual space by Monteiro et al. [117] with worst-case iteration complexity $\mathcal{O}(nL^2)$. In 1996, Jansen et al. [86] derived a primal-dual Dikin-type affine scaling method which at each iteration minimizes the duality gap over the so-called Dikin ellipsoid in the primal-dual space. This method not only has an improved $\mathcal{O}(nL)$ polynomial complexity but it also features both centering and reduction of the duality gap in contrast to the method of Monteiro et al. [117]. de Klerk et al. [35] generalized the methods of Monteiro et al. [117] and Jansen et al. [86] to SDO. Nevertheless, the extension of affine scaling methods from LO to SCO, which includes SOCO, is not as straightforward as from LO to SDO, because SCO relies on a rather different algebra. As indicated in the book [134], the extension of IPMs with self-regular barrier functions from LO to SOCO does not follow in the same direction as in the extension from LO to SDO.

In this chapter, we review the derivation of the Dikin-type affine scaling method for

SDO from Chapter 6 in [35]. We then generalize the primal-dual Dikin-type affine scaling method of Jansen et al. [86] and de Klerk et al. [35] to SCO. We present the extension from [113].

2.1 Dikin-type affine scaling algorithm for SDO

We only present the derivation of the Dikin ellipsoid and Dikin-type search directions for SDO. The feasibility of a Dikin-step and iteration complexity for SDO are analogous to the case of SCO.

Recall the central path equations from Section 1.4.1. Applying the Newton method to the system (1.22) leads to

$$\begin{aligned} \langle A^i, \Delta X\rangle &= 0, \\ \sum_{i=1}^{m} A^i \Delta y_i + \Delta S &= 0, \\ X\Delta S + S\Delta X &= \mu e - XS. \end{aligned} \tag{2.1}$$

The symmetry of ΔS follows from the symmetry of A^i for $i = 1, \ldots, m$. However, ΔX may not be symmetric; hence the system (2.1) is not necessarily well-defined. To resolve the symmetry issue, we can employ the NT scaling scheme which projects X and S onto the same point V. Let D be a scaling matrix defined as

$$D := S^{-\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{\frac{1}{2}}S^{-\frac{1}{2}}, \tag{2.2}$$

from which one can easily see that $D^{-1}X = SD$. Then V is defined as

$$V := D^{-\frac{1}{2}}XD^{-\frac{1}{2}} = D^{\frac{1}{2}}SD^{\frac{1}{2}}. \tag{2.3}$$

Note that $V^2 = D^{-\frac{1}{2}}XS^{\frac{1}{2}}$, which implies that V^2 is symmetric and is similar to XS, i.e., the eigenvalues of V^2 and XS are identical. Therefore, for a given primal-dual solution (X, y, S) the complementarity gap is given by

$$\mathrm{Trace}(XS) = \mathrm{Trace}(V^2).$$

In a similar fashion, the search directions can be scaled as

$$\begin{aligned} D_X &:= D^{-\frac{1}{2}} \Delta X D^{-\frac{1}{2}}, \\ D_S &:= D^{\frac{1}{2}} \Delta S D^{\frac{1}{2}}. \end{aligned}$$

From the orthogonality of ΔX and ΔS, the orthogonality of D_X and D_S follows. Hence, the Newton step in the scaled space is given by

$$D_V = D_X + D_S.$$

The complementarity gap of the new iterate is then equal to

$$\begin{aligned} \mathrm{Trace}\left((X + \Delta X)(S + \Delta S)\right) &= \mathrm{Trace}\left((V + D_X)(V + D_S)\right) \\ &= \mathrm{Trace}(V^2 + V D_V). \end{aligned} \tag{2.4}$$

The proximity of (X, y, S) to the central path can be measured by

$$\mathrm{prox}(XS) := \frac{\lambda_{\max}(XS)}{\lambda_{\min}(XS)}, \qquad (X, y, S) \in \mathrm{ri}(\mathcal{P}_{\mathrm{SDO}} \times \mathcal{D}_{\mathrm{SDO}}). \tag{2.5}$$

Notice that XS has the same eigenvalues as $X^{\frac{1}{2}} S X^{\frac{1}{2}}$, i.e., XS has real positive eigenvalues even though it is not necessarily symmetric. Further, it follows from (2.5) that $\mathrm{prox}(XS) \geq 1$, and the equality holds only when (X, y, S) is on the central path. Therefore, a neighborhood of the central path can be defined as

$$\mathcal{N}_{\mathrm{prox}}(\xi) := \left\{ (X, y, S) \in \mathrm{ri}(\mathcal{P}_{\mathrm{SDO}} \times \mathcal{D}_{\mathrm{SDO}}) \mid \mathrm{prox}(XS) \leq \xi \right\}, \tag{2.6}$$

where $\xi > 1$.

2.1.1 The Dikin-type search directions

The extension of the Dikin ellipsoid from LO is given by

$$\mathcal{E}ll(V) := \left\{ D_V \mid \|V^{-\frac{1}{2}} D_V V^{-\frac{1}{2}}\|_F \leq 1 \right\}, \tag{2.7}$$

where V is positive definite. In the Dikin-type affine scaling method, the goal is to minimize the complementarity gap (2.4) over the Dikin ellipsoid (2.7). Note that $\mathrm{Trace}(V^2)$ is a constant term in (2.4). Hence, Dikin-type search directions are derived by solving

$$D_V^* := \mathrm{argmin}\left\{ \mathrm{Trace}(V D_V) \mid \|V^{-\frac{1}{2}} D_V V^{-\frac{1}{2}}\|_F \leq 1 \right\}. \tag{2.8}$$

The optimal solution of (2.8) can be obtained analytically, which is given by

$$D_V^* = D_X^* + D_S^* = -\frac{V^3}{\|V^2\|_F}.$$

By scaling the search directions back into the original space we get

$$\Delta X + D \Delta S D = -\frac{XSX}{\left(\mathrm{Trace}\left((XS)^2\right)\right)^{\frac{1}{2}}}.$$

Consequently, Dikin-type search directions in the original space are obtained by solving the following system of equations

$$\begin{aligned} \langle A^i, \Delta X \rangle &= 0, \\ \sum_{i=1}^{m} A^i \Delta y_i + \Delta S &= 0, \\ \Delta X + D \Delta S D &= -\frac{XSX}{\left(\mathrm{Trace}\left((XS)^2\right)\right)^{\frac{1}{2}}}, \end{aligned} \tag{2.9}$$

which is analogous to (2.1), except that it gives symmetric ΔX and ΔS.

The Dikin-type affine scaling method iteratively solves (2.9) to generate a new iterate and employs the neighborhood (2.6) to stay in close proximity to the central path.

2.2 Dikin-type affine scaling algorithm for SCO

Recall from (1.20) and (1.21) that d_x and d_s are orthogonal. Then a feasible primal-dual step along the search directions arrives at the duality gap

$$\mathrm{Trace}((v + d_x) \circ (v + d_s)) = \mathrm{Trace}(v^2 + v \circ d_v), \tag{2.10}$$

where $d_v = d_x + d_s$ stands for the Newton step in the scaled primal-dual space, referred to as the v-space. The Dikin-type algorithm aims for minimizing the duality gap (2.10) over a suitable ellipsoid in the v-space which is given by

$$\|v^{-1} \circ d_v\|_F \leq 1. \tag{2.11}$$

Instead of following the central path, the Dikin-type algorithm chooses the scaled primal-dual solutions from the ellipsoid (2.11). Ellipsoid (2.11) in the original space can be given as

$$\|P_w^{\frac{1}{2}} x^{-1} \circ P_w^{-\frac{1}{2}} \Delta x + P_w^{-\frac{1}{2}} s^{-1} \circ P_w^{\frac{1}{2}} \Delta s\|_F \leq 1.$$

It can be easily verified that this ellipsoid is indeed a generalization of the suitable ellipsoid introduced for LO in [86]. It is worth mentioning that a word-for-word generalization of the Dikin ellipsoid is written as

$$\mathcal{E}ll(x, s) := \{(\Delta x, \Delta s) \mid \|x^{-1} \circ \Delta x + s^{-1} \circ \Delta s\|_F \leq 1\}.$$

Notice that $\mathcal{E}ll(x, s)$ intersected with

$$\mathcal{A}\Delta x = 0, \quad \mathcal{A}^*\Delta y + \Delta s = 0$$

is not necessarily bounded because the ellipsoid $\mathcal{E}ll(x, s)$ contains the affine space $x^{-1} \circ \Delta x + s^{-1} \circ \Delta s = 0$. Thus, the system

$$\begin{aligned} \mathcal{A}\Delta x &= 0, \\ \mathcal{A}^*\Delta y + \Delta s &= 0, \\ x^{-1} \circ \Delta x + s^{-1}\Delta s &= 0, \end{aligned}$$

does not necessarily have a unique solution as $L^{-1}(x^{-1})$ and $L(s^{-1})$ do not commute in general.

2.2.1 Minimizing the duality gap over the ellipsoid

The Dikin-type search directions are derived by minimizing the duality gap (2.10) over the ellipsoid (2.11)

$$\begin{aligned} \min \quad & \mathrm{Trace}(v^2 + v \circ d_v) \\ \text{s.t.} \quad & \|v^{-1} \circ d_v\|_F \leq 1. \end{aligned} \tag{2.12}$$

Recall that $x, s \succ_{\mathcal{K}} 0$ implies $v \succ_{\mathcal{K}} 0$, where $\succ_{\mathcal{K}} 0$ simply denotes the cone inclusion. Then, we can realize that $\mathrm{Trace}(v \circ d_v) = \mathrm{Trace}(v^2 \circ (v^{-1} \circ d_v))$, where $v^2 := v \circ v$. Now, letting $\hat{v} := v^{-1} \circ d_v$, optimization problem (2.12) can be written as

$$\begin{aligned} \min \quad & \mathrm{Trace}(v^2 + v^2 \circ \hat{v}) \\ \text{s.t.} \quad & \|\hat{v}\|_F \leq 1. \end{aligned} \tag{2.13}$$

It is easy to show that the optimal solution of (2.13) is given by $\hat{v}^* = -\frac{v^2}{\|v^2\|_F}$, and the optimal objective value is $\|v\|_F^2 - \|v^2\|_F$. Therefore, we have

$$d_v^* = L(v^{-1})^{-1}\hat{v}^* = -\frac{v^3}{\|v^2\|_F}, \tag{2.14}$$

where we have used the fact that v and v^{-1} have the same Jordan frame, and $L(v^{-1})$ is invertible, see Theorem A.1.1 and Lemma A.1.5. Consequently, the Dikin-type search directions are obtained by solving

$$\begin{aligned} \bar{\mathcal{A}} d_x &= 0, \\ \bar{\mathcal{A}}^* \Delta y + d_s &= 0, \\ d_x + d_s &= -\frac{v^3}{\|v^2\|_F}. \end{aligned} \tag{2.15}$$

Akin to the Newton system (1.20), this system of equations has a unique solution. Taking a Dikin step along the search directions d_x and d_s, the new iterate in the v-space is obtained as

$$\begin{aligned} v_x^\alpha &:= v + \alpha d_x, \\ v_s^\alpha &:= v + \alpha d_s, \end{aligned} \tag{2.16}$$

and in the original space as

$$\begin{aligned} x^\alpha &:= x + \alpha \Delta x = x + \alpha P_w^{\frac{1}{2}} d_x, \\ s^\alpha &:= s + \alpha \Delta s = x + \alpha P_w^{-\frac{1}{2}} d_s. \end{aligned}$$

2.2.2 Proximity to the central path and feasibility

While reducing the duality gap, the Dikin-type algorithm keeps the iterates in a predefined neighborhood of the central path. The proximity measure given in [86] is generalized to

$$\operatorname{prox}(v^2) := \frac{\lambda_{\max}(v^2)}{\lambda_{\min}(v^2)}, \tag{2.17}$$

where $\lambda_{\min}(v^2)$ and $\lambda_{\max}(v^2)$ denote the smallest and largest eigenvalues of v^2, respectively. Note that $\operatorname{prox}(v^2) \geq 1$, and equality holds only when $x \circ s = \mu e$, see Lemma 28 in [156]. Further

$$x \circ s = \mu e \iff v^2 = \mu e.$$

2.2.3 The Dikin-type algorithm

The outline of the Dikin-type algorithm is described in Algorithm 1. The Dikin-type algorithm starts with a strictly feasible primal-dual solution (x^0, y^0, s^0) which is close enough to the central path in terms of the proximity measure prox(.). The algorithm uses the default steplength $\frac{1}{\xi\sqrt{r}}$ which, after each Dikin step, maintains feasibility and proximity to the central path. Dikin steps are taken until the duality gap decreases below the accuracy parameter ε.

Algorithm 1 Dikin-type algorithm

Input

A strictly feasible solution (x^0, y^0, s^0)

Parameters

Proximity measure $\xi > 1$ so that $\text{prox}(x^0 \circ s^0) \leq \xi$

Steplength α with default value $\frac{1}{\xi\sqrt{r}}$

Accuracy parameter ε

$x := x^0, \;\; s := s^0$

repeat

Obtain $(\Delta x, \Delta s)$ by solving (2.15) and then using (1.21)

Set $x := x + \alpha\Delta x$

Set $s := s + \alpha\Delta s$

until $\text{Trace}(x \circ s) \leq \varepsilon$

2.2.4 Complexity analysis of the Dikin-type algorithm

In this section, we provide technical results to show that the default steplength $\alpha = \frac{1}{\xi\sqrt{r}}$ leads to a strictly feasible primal-dual solution which also stays in a close proximity to the central path. We also prove that the Dikin-type algorithm arrives at a strictly feasible ε-optimal solution in $\mathcal{O}\left(\xi r \log\left(\frac{\text{Trace}(x^0 \circ s^0)}{\varepsilon}\right)\right)$ iterations.

The next lemma provides a sufficient condition for the steplength α, by which the Dikin step gives a strictly feasible primal-dual solution.

Lemma 2.2.1. *Let $\alpha \geq 0$, and assume that $v \succ_{\mathcal{K}} 0$. Then, the steplength $\bar{\alpha}$ is* *feasible if*

$$v_x^\alpha \circ v_s^\alpha \succ_{\mathcal{K}} 0, \qquad \forall\ 0 \le \alpha \le \bar{\alpha},$$

where $v_x^0 := v$, $v_s^0 := v$, *and* v_x^α *and* v_s^α *are defined by* (2.16).

Proof. By Lemma A.1.3, $v_x^\alpha \circ v_s^\alpha \succ_{\mathcal{K}} 0$ implies that $\det(v_x^\alpha) \neq 0$ and $\det(v_s^\alpha) \neq 0$ for $0 \le \alpha \le \bar{\alpha}$. Since the eigenvalues of v_x^α and v_s^α are continuous functions of α and $v \succ_{\mathcal{K}} 0$ holds, then the eigenvalues of v_x^α and v_s^α do not vanish and remain positive on $[0, \bar{\alpha}]$. $\square$

Lemma 2.2.2. *Let* $v^\alpha := P_{w^\alpha}^{-\frac{1}{2}} x^\alpha = P_{w^\alpha}^{\frac{1}{2}} s^\alpha$, *where* w^α *denotes the scaling point of* x^α *and* s^α, *where*

$$\begin{aligned} x^\alpha &= x + \alpha \Delta x, \\ s^\alpha &= s + \alpha \Delta s. \end{aligned}$$

Then, we have

$$\begin{aligned} \operatorname{prox}((v^\alpha)^2) &\le \operatorname{prox}(v_x^\alpha \circ v_s^\alpha), \\ \operatorname{prox}((v^\alpha)^2) &\le \operatorname{prox}(x^\alpha \circ s^\alpha). \end{aligned}$$

Proof. Since $w \succ_{\mathcal{K}} 0$ and $s \succ_{\mathcal{K}} 0$, it follows from part 1 of Lemma A.1.1 and part 2 of Lemma A.1.6 that

$$\begin{aligned} v^\alpha = P_{w^\alpha}^{\frac{1}{2}} s^\alpha \sim P_{(s^\alpha)^{\frac{1}{2}}} w^\alpha &= P_{(s^\alpha)^{\frac{1}{2}}} P_{(s^\alpha)^{-\frac{1}{2}}} \left[P_{(s^\alpha)^{\frac{1}{2}}} x^\alpha \right]^{\frac{1}{2}} \\ &= (P_{(s^\alpha)^{\frac{1}{2}}} x^\alpha)^{\frac{1}{2}} \sim (P_{(x^\alpha)^{\frac{1}{2}}} s^\alpha)^{\frac{1}{2}}, \end{aligned}$$

where $\sim$ denotes the similarity of the eigenvalues. Thus, according to Theorem A.1.1, we get

$$(v^\alpha)^2 \sim P_{(x^\alpha)^{\frac{1}{2}}} s^\alpha,$$

where

$$\begin{aligned} P_{(x^\alpha)^{\frac{1}{2}}} s^\alpha &= P^{\frac{1}{2}}_{P^{\frac{1}{2}}_{w^\alpha}(v+\alpha d_x)} (P^{-\frac{1}{2}}_{w^\alpha}(v+\alpha d_s)) \\ &\sim P^{\frac{1}{2}}_{v+\alpha d_x}(v+\alpha d_s). \end{aligned}$$

Now, considering Lemma A.1.7, we can conclude that

$$\begin{aligned} \operatorname{prox}((v^\alpha)^2) &= \operatorname{prox}(P^{\frac{1}{2}}_{v+\alpha d_x}(v+\alpha d_s)) \leq \operatorname{prox}(v_x^\alpha \circ v_s^\alpha), \\ \operatorname{prox}((v^\alpha)^2) &= \operatorname{prox}(P_{(x^\alpha)^{\frac{1}{2}}} s^\alpha) \leq \operatorname{prox}(x^\alpha \circ s^\alpha), \end{aligned}$$

which completes the proof. □

Assume that a strictly feasible solution (x, y, s) is given satisfying $\operatorname{prox}(v^2) \leq \xi$, where $\xi \geq 1$. Then, it follows from (3.29) that there exists $\xi_1, \xi_2 > 0$ with $\xi_2 = \xi\xi_1$ so that

$$\xi_2 e \succeq_{\mathcal{K}} v^2 \succeq_{\mathcal{K}} \xi_1 e. \tag{2.18}$$

Now, Lemma 2.2.3 establishes a bound on the steplength α which guarantees feasibility and proximity to the central path after a Dikin step.

Lemma 2.2.3. *The steps x^α and s^α are strictly feasible and* $\operatorname{prox}((v^\alpha)^2) \leq \xi$ *if*

$$\alpha \leq \min\left\{ \frac{\|v^2\|_F}{2\xi_2}, \frac{4\xi_1}{\|v^2\|_F} \right\}.$$

Proof. Recall from (1.18) that $P_w^{\frac{1}{2}}$ and $P_w^{-\frac{1}{2}}$ are invertible maps from $\operatorname{int}(\mathcal{K})$ to $\operatorname{int}(\mathcal{K})$. Therefore, $x^\alpha, s^\alpha \succ_{\mathcal{K}} 0$ if and only if $v_x^\alpha, v_s^\alpha \succ_{\mathcal{K}} 0$. Hence, by considering Lemma 2.2.1, we only need to show that $v_x^\alpha \circ v_s^\alpha \succ_{\mathcal{K}} 0$, where

$$v_x^\alpha \circ v_s^\alpha = (v + \alpha d_x) \circ (v + \alpha d_s) = v^2 - \alpha \frac{v^4}{\|v^2\|_F} + \alpha^2 d_x \circ d_s.$$

Note that v^2 and v^4 share the same Jordan frame by Theorem A.1.1. Hence, $\lambda_k(v^2)-\alpha\frac{\lambda_k(v^2)^2}{\|v^2\|_F}$ serves as the eigenvalue of $\varphi(v^2)$, where

$$\varphi(t) := t - \alpha\frac{t^2}{\|v^2\|_F},$$

and $\lambda_k(v^2)$ denotes the eigenvalue of v^2 for $k = 1, \dots, r$. For $\alpha \leq \frac{\|v^2\|_F}{2\xi_2}$, function $\varphi(t)$ is monotonically increasing on $[0, \xi_2]$. All this means that

$$\varphi(\xi_2)e \succeq_{\mathcal{K}} v^2 - \alpha\frac{v^4}{\|v^2\|_F} \succeq_{\mathcal{K}} \varphi(\xi_1)e,$$

and thus

$$\varphi(\xi_2)e + \alpha^2 d_x \circ d_s \succeq_{\mathcal{K}} v_x^\alpha \circ v_s^\alpha \succeq_{\mathcal{K}} \varphi(\xi_1)e + \alpha^2 d_x \circ d_s.$$

As long as the Dikin step is feasible, i.e., $\varphi(\xi_1)e + \alpha^2 d_x \circ d_s \succeq_{\mathcal{K}} 0$, we will have $\operatorname{prox}(v_x^\alpha \circ v_s^\alpha) \leq \xi$ if

$$\xi(\varphi(\xi_1)e + \alpha^2 d_x \circ d_s) \succ_{\mathcal{K}} \varphi(\xi_2)e + \alpha^2 d_x \circ d_s,$$

which can be further simplified to

$$\frac{\xi_1\xi_2}{\|v^2\|_F}e + \alpha d_x \circ d_s \succ_{\mathcal{K}} 0. \tag{2.19}$$

By Lemma A.1.2, $\frac{1}{4}\|d_x + d_s\|_F^2$ gives an upper bound on the eigenvalues of $d_x \circ d_s$. Thus, by Lemma A.1.4 and (2.18), we get

$$\frac{1}{4}\xi_2 e \succeq_{\mathcal{K}} \frac{1}{4}\lambda_{\max}(v^2)e \succeq_{\mathcal{K}} \frac{1}{4}\frac{\operatorname{Trace}(v^6)}{\|v^2\|_F^2}e = \frac{1}{4}\|\frac{v^3}{\|v^2\|_F}\|_F^2 e = \frac{1}{4}\|d_x + d_s\|_F^2 e \succeq_{\mathcal{K}} d_x \circ d_s.$$

Consequently, condition (2.19) is satisfied if

$$\left(\frac{\xi_1\xi_2}{\|v^2\|_F} - \frac{1}{4}\alpha\xi_2\right)e \succ_{\mathcal{K}} 0,$$

which in turn implies that $\alpha < \frac{4\xi_1}{\|v^2\|_F}$. Thus, considering Lemma 3.30, the result follows. $\square$

The next lemma shows that after a feasible Dikin step, the duality gap is reduced by at least a factor of $\left(1 - \frac{\alpha}{\sqrt{r}}\right)$.

Lemma 2.2.4. *Let (x, y, s) be a feasible primal-dual solution. Then, after a feasible Dikin step, we get*

$$\mathrm{Trace}(x^\alpha \circ s^\alpha) \leq \left(1 - \frac{\alpha}{\sqrt{r}}\right) \mathrm{Trace}(x \circ s).$$

Proof. Since $\mathrm{Trace}(d_x \circ d_s) = 0$, it follows that

$$\mathrm{Trace}((v + \alpha d_x) \circ (v + \alpha d_s)) = \mathrm{Trace}\left(v^2 - \alpha \frac{v^4}{\|v^2\|_F}\right) = \|v\|_F^2 - \alpha\|v^2\|_F.$$

Using the Cauchy-Schwarz inequality, a lower bound of $\|v\|_F^2$ is given by

$$\|v\|_F^2 = \mathrm{Trace}(v^2) = \mathrm{Trace}(v^2 \circ e) \leq \|v^2\|_F \|e\|_F = \sqrt{r}\|v^2\|_F.$$

Thus, we can conclude that

$$\mathrm{Trace}(x^\alpha \circ s^\alpha) = \mathrm{Trace}((v + \alpha d_x) \circ (v + \alpha d_s)) \leq \left(1 - \frac{\alpha}{\sqrt{r}}\right) \mathrm{Trace}(v^2),$$

which completes the proof. □

Theorem 2.2.1. *Let $\varepsilon > 0$, $\alpha = \frac{1}{\xi\sqrt{r}}$ and $\xi > 1$ so that $\mathrm{prox}(x^0 \circ s^0) \leq \xi$. Then, the Dikin-type algorithm terminates after at most $\lceil \xi r \log \frac{\mathrm{Trace}(x^0 \circ s^0)}{\varepsilon} \rceil$ iterations yielding a feasible solution (x, y, s) such that $\mathrm{prox}(v^2) \leq \xi$ and $\mathrm{Trace}(x \circ s) \leq \varepsilon$.*

Proof. By the left hand side inequality in (2.18), we have $\|v^2\|_F \geq \|\xi_1 e\|_F$ and thus

$$\alpha = \frac{1}{\xi\sqrt{r}} = \frac{\xi_1}{\xi_2\sqrt{r}} \leq \frac{\xi_1\sqrt{r}}{2\xi_2} = \frac{\|\xi_1 e\|_F}{2\xi_2} \leq \frac{\|v^2\|_F}{2\xi_2}.$$

By the right hand side inequality in (2.18) we have $\|v^2\|_F \leq \xi_2\sqrt{r}$, and thus

$$\frac{4\xi_1}{\|v^2\|_F} \geq \frac{4\xi_1}{\xi_2\sqrt{r}} = \frac{4}{\xi\sqrt{r}} > \alpha.$$

Therefore, the default value of α satisfies the conditions in Lemma 2.2.3.

As Lemma 2.2.4 proves, each Dikin step with the default value of α reduces the duality gap by a factor of $\left(1 - \frac{1}{\xi r}\right)$. Consequently, the duality gap reduces below ε after k iterations if

$$\left(1 - \frac{1}{\xi r}\right)^k \mathrm{Trace}(x^0 \circ s^0) \leq \varepsilon.$$

Taking the logarithm of both sides gives

$$k \log\left(1 - \frac{1}{\xi r}\right) + \log(\mathrm{Trace}(x^0 \circ s^0)) \leq \log(\varepsilon).$$

This inequality is satisfied if

$$\frac{k}{\xi r} \geq \log(\mathrm{Trace}(x^0 \circ s^0)) - \log(\varepsilon) = \log\left(\frac{\mathrm{Trace}(x^0 \circ s^0)}{\varepsilon}\right),$$

where we have used the fact that $-\log\left(1-\frac{1}{\xi r}\right) \geq \frac{1}{\xi r}$. This completes the proof. $\square$

2.2.5 Numerical results

As stated in Theorem A.1.2, SCO only includes LO, SOCO, and SDO along with their complex variants or a combination of them. The extensions of the Dikin-type algorithm for LO and SDO problems have been already investigated in [86] and [35], respectively. Thus, in this section, we investigate the performance of the Dikin-type algorithm only for SOCO problems. Toward this end, a set of 13 SOCO test problems are chosen from the DIMACS library as listed in Table 2.1. The test problems are of minimization type, and their optimal solutions are provided by Mittelmann in [110].

We adopt SeDuMi 1.3 [166], SDPT3-4.0 [172, 176], and MOSEK 7.1[1] as competing methods for comparison purposes. SeDuMi, and MOSEK apply feasible IPMs on a

[1] https://www.mosek.com/

Table 2.1: The specifications of the SOCO problems.

Name	#Rows	#Lorentz Cones	#Linear Variables	Optimal Value
nql30new	3680	[900; 900x3]	3602	-0.946028
nql30old	3601	[900; 900x3]	5560	0.946028
nql60new	14560	[3600;3600x3]	14402	-0.935423
nql60old	14401	[3600; 3600x3]	21920	0.935423
nql180new	130080	[32400;32400x3]	129602	-0.927717
nb	123	[793; 793x 3]	4	-0.050703
nb-L1	915	[793; 793x 3]	797	-13.01227
nb-L2	123	[839; 1x1677,838x3]	4	-1.628972
nb-L2-Bessel	123	[839; 1x 123, 838x 3]	4	-0.102571
qssp30new	3691	[1891; 1891x 4]	2	-6.496675
qssp30old	5674	[1891; 1891x 4]	3600	6.496675
qssp60new	14581	[7381; 7381x 4]	2	-6.562696
qssp60old	22144	[7381; 7381x 4]	14400	6.562696

self-dual embedding format (see [134] for details) while SDPT3 employs an infeasible IPM. The Dikin-type method is implemented in SeDuMi's framework, where it uses its own search direction, neighborhood and rule of steplength. All the methods are run in MATLAB 8.1 on a Core i7 @3.4GHz CPU with 8 GB RAM. We leave the default parameters unchanged and let the competing methods terminate according to their own stopping criterion. However, the threshold for the primal and dual feasibility and duality gap, and the maximum number of iterations have been set to 10^{-8} and 150, respectively, for all the competing methods. The Dikin-type algorithm complies with SeDuMi's default settings so that it terminates if the primal and dual infeasibility along with the duality gap drops below the SeDuMi's default threshold (10^{-8}). Further, the Dikin-type algorithm terminates if it gains less than 0.1% improvement in duality gap.

A safeguard procedure is considered in the Dikin-type algorithm for the case when the Dikin step provides no significant improvement in the duality gap (i.e., the relative improvement is less than 5%). In this situation, we skip the Dikin step and take a centering step which is obtained by solving

$$\begin{aligned} \bar{\mathcal{A}} d_x &= 0, \\ \bar{\mathcal{A}}^* \Delta y + d_s &= 0, \\ d_x + d_s &= \mu v^{-1} - v, \end{aligned}$$

where $\mu := \frac{\mathrm{Trace}(v^2)}{r}$. Doing so, we indeed improve the centrality but keep the duality gap constant.

In practice, the theoretical steplength $\alpha = \frac{1}{\xi\sqrt{r}}$ is nearly zero for large values of r, which prevents taking a long Dikin step. To remove this drawback, we apply a decreasing sequence of steplengths in $(0, \alpha_{\max})$ and take the largest value of α which satisfies

$$\mathrm{prox}(v_x^\alpha \circ v_s^\alpha) \leq \xi.$$

In our experiments, ξ is fixed at 4.2, and $\alpha_{\max}$ is set to the threshold value for the boundary of the cone. In fact $\xi = 4.2$ obtained the best results in our initial experiments. Smaller values of ξ lead to relatively fewer centering steps but with no significant improvement in the duality gap. For large values of ξ, we have significant improvement in the duality gap during the initial iterations, but the method gets very close to the boundary of the cone, resulting in almost no significant improvement afterwards.

Tables 2.2 to 2.5 illustrate the results of the Dikin-type algorithm and the competing methods, in which the best primal objective (Primal), relative duality gap (rgap), relative primal infeasibility (rpinf), relative dual infeasibility (rdinf), computational time (CPU), the number of iterations (#Iter), relative minimum eigenvalue of x (releig), and the number of centering steps (#cent.) are provided. The relative primal and dual infeasibility, relative duality gap, and relative minimum eigenvalue are defined, respectively, as

$$\begin{aligned}
\text{rpinf} &:= \frac{\|Ax - b\|_2}{1 + \|b\|_2}, \\
\text{rdinf} &:= \frac{\lambda_{\max}(\mathcal{A}^* y - c)}{1 + \|c\|_2}, \\
\text{rgap} &:= \frac{c^T x - b^T y}{1 + |c^T x| + |b^T y|}, \\
\text{releig} &:= \frac{\lambda_{\min}(x)}{1 + \|b\|_2}.
\end{aligned}$$

As illustrated by Table 2.2, both the Dikin-type algorithm and SeDuMi perform equally well on the test instances in terms of the objective value even though SeDuMi performs quite faster. In 10 out of 13 instances, the Dikin-type algorithm arrives at as good solutions as SeDuMi with 10^{-5} precision, and in 3 cases, the Dikin-type algorithm performs better. On "nql60old", "qssp30old", and "qssp60old", SeDuMi stops at non-optimal solutions. Further, the initial experiments showed that SeDuMi is not consistently stable on different platforms. For instance, when run on a machine under Windows operating system, SeDuMi fails in "qssp60old" after 3 iterations.

The Dikin-type algorithm obtains the average relative duality gap 1.24E-03 within the average of 67 iterations while SeDuMi ends up with 6.18E-03 in 18 iterations. The Dikin-type algorithm also outperforms SeDuMi in terms of primal and dual infeasibility. As demonstrated by the entries, the average of rpinf and rdinf over the test instances are 4.78E-04 and 8.46E-08, respectively for the Dikin-type algorithm and 2.07E-03 and 5.07E-07, respectively for SeDuMi.

It is worth mentioning that the Dikin-type algorithm takes only a few centering steps to get the iterates back to the vicinity of the central path. To be more precise, in 9 out of 13 instances, the Dikin-type algorithm uses less than 10% of the iterations for centering and more than 90% of iterations for reducing the duality gap. Further, no centering was used for 4 test problems. The worst case belongs to the test instances“qssp60new” and “qssp60old“, on which 19.3% and 19.7% of the iterations are spent on the centering.

The Dikin-type algorithm has been compared with MOSEK and SDPT3 in Tables 2.3 and 2.4, respectively. In 9 out of 13 instances, the Dikin-type algorithm obtains as good solutions as MOSEK and SDPT3. Nevertheless, MOSEK and SDPT3 outperform the Dikin-type algorithm in terms of solution quality. For all the test instances, MOSEK and SDPT3 have arrived at the best solutions. Here, SDPT3 performs better than MOSEK. On ”qssp60old”, MOSEK has gained a solution with relative primal infeasibility 0.0013 while this value is 8.83E-09 for SDPT3. In terms of accuracy and speed, MOSEK and SDPT3 are the winner. The average of relative primal and dual infeasibility are 1.01E-04 and 4.16E-09 for MOSEK and 2.99E-09 and 6.33E-13 for SDPT3, respectively, while these values are 4.78E-03 and 8.46E-08 for the Dikin-type algorithm.

The numerical experiments show that in many instances, the Dikin type algorithm provides solutions which are of higher accuracy than SeDuMi or almost as accurate as MOSEK and SDPT3 in terms of relative duality gap, and relative primal and dual infeasibility. Nevertheless, the numerical experiments also confirm that the Dikin-type search directions are less effective in decreasing the duality gap than the predictor-corrector search directions.

Table 2.2: Comparison of the Dikin-type algorithm and SeDuMi in terms of quality.

Instance	**Dikin-type affine scaling**				**SeDuMi**			
	Primal	rgap	rpinf	rdinf	Primal	rgap	rpinf	rdinf
nql30new	-0.94602800	7.33E-10	1.90E-09	7.96E-10	-0.94602788	1.11E-07	8.72E-10	1.02E-09
nql30old	0.94603380	1.62E-06	4.65E-08	5.92E-11	0.94604805	4.37E-09	8.21E-10	8.48E-10
nql60new	-0.93505133	1.01E-07	2.32E-09	9.70E-10	-0.93505119	9.35E-07	5.85E-10	1.57E-09
nql60old	0.93508946	1.07E-05	6.04E-07	4.89E-11	0.93516151	5.64E-09	2.92E-10	5.85E-10
nql180new	-0.92772591	1.70E-05	2.91E-09	1.08E-09	-0.92772365	3.42E-06	1.03E-09	1.30E-09
nb	-0.05070309	2.95E-11	1.26E-08	6.73E-12	-0.05070309	1.16E-12	3.74E-12	1.71E-12
nb_L1	-13.01226986	1.22E-08	2.52E-10	9.59E-11	-13.01227003	1.10E-10	1.56E-10	1.07E-10
nb_L2	-1.62896764	1.81E-06	1.50E-05	1.93E-10	-1.62897196	1.55E-10	3.74E-11	5.76E-10
nb_L2_Bessel	-0.10256951	2.03E-09	5.27E-12	1.85E-12	-0.10256950	1.10E-08	2.81E-11	2.44E-12
qssp30new	-6.49667496	2.85E-09	7.03E-10	3.22E-10	-6.49666910	7.66E-10	7.18E-10	2.96E-09
qssp30old	6.52338286	6.20E-03	0.0022	1.66E-12	6.70807589	3.28E-02	0.0149	5.75E-06
qssp60new	-6.56269019	1.98E-07	1.28E-07	1.10E-06	-6.56269721	8.23E-10	2.93E-10	1.05E-09
qssp60old	6.60855184	9.90E-03	0.004	5.40E-12	6.80741982	4.76E-02	0.012	8.24E-07
Average		1.24E-03	4.78E-04	8.46E-08		6.18E-03	2.07E-03	5.07E-07

Table 2.3: Comparison of the Dikin-type algorithm and MOSEK in terms of quality.

Instance	**Dikin-type affine scaling**				**MOSEK**			
	Primal	rgap	rpinf	rdinf	Primal	rgap	rpinf	rdinf
nql30new	-0.94602800	7.33E-10	1.90E-09	7.96E-10	-0.94602406	3.98E-09	5.03E-11	9.98E-09
nql30old	0.94603380	1.62E-06	4.65E-08	5.92E-11	0.94602932	3.39E-07	6.50E-06	5.12E-12
nql60new	-0.93505133	1.01E-07	2.32E-09	9.70E-10	-0.93504191	8.31E-10	1.13E-11	9.62E-09
nql60old	0.93508946	1.07E-05	6.04E-07	4.89E-11	0.93505940	2.47E-06	1.47E-07	1.62E-12
nql180new	-0.92772591	1.70E-05	2.91E-09	1.08E-09	-0.92764299	3.07E-09	2.75E-10	1.14E-08
nb	-0.05070309	2.95E-11	1.26E-08	6.73E-12	-0.05070309	1.41E-11	9.52E-09	1.24E-11
nb_L1	-13.01226986	1.22E-08	2.52E-10	9.59E-11	-13.01226111	1.65E-07	1.05E-09	2.45E-09
nb_L2	-1.62896764	1.81E-06	1.50E-05	1.93E-10	-1.62897188	2.41E-08	2.84E-07	1.48E-10
nb_L2_Bessel	-0.10256951	2.03E-09	5.27E-12	1.85E-12	-0.10256946	4.28E-08	1.92E-10	9.96E-11
qssp30new	-6.49667496	2.85E-09	7.03E-10	3.22E-10	-6.49665883	9.06E-12	8.93E-16	1.65E-08
qssp30old	6.52338286	6.20E-03	0.0022	1.66E-12	6.49667620	5.40E-08	1.20E-07	1.68E-12
qssp60new	-6.56269019	1.98E-07	1.28E-07	1.10E-06	-6.56269012	5.60E-13	2.12E-15	3.90E-09
qssp60old	6.60855184	9.90E-03	0.004	5.40E-12	6.56271306	1.04E-06	0.0013	8.66E-13
Average		1.24E-03	4.78E-04	8.46E-08		3.19E-07	1.01E-04	4.16E-09

Table 2.4: Comparison of the Dikin-type algorithm and SDPT3 in terms of quality.

Instance	**Dikin-type affine scaling**				**SDPT3**			
	Primal	rgap	rpinf	rdinf	Primal	rgap	rpinf	rdinf
nql30new	-0.94602800	7.33E-10	1.90E-09	7.96E-10	-0.94602850	8.49E-09	8.23E-14	2.48E-12
nql30old	0.94603380	1.62E-06	4.65E-08	5.92E-11	0.94602848	5.86E-09	1.62E-08	1.81E-15
nql60new	-0.93505133	1.01E-07	2.32E-09	9.70E-10	-0.93505295	7.37E-09	3.40E-13	5.62E-13
nql60old	0.93508946	1.07E-05	6.04E-07	4.89E-11	0.93505306	4.12E-08	9.13E-09	8.50E-15
nql180new	-0.92772591	1.70E-05	2.91E-09	1.08E-09	-0.92772862	6.85E-09	4.18E-12	5.21E-14
nb	-0.05070309	2.95E-11	1.26E-08	6.73E-12	-0.05070309	8.49E-09	2.26E-12	9.20E-17
nb_L1	-13.01226986	1.22E-08	2.52E-10	9.59E-11	-13.01227060	6.15E-09	2.92E-11	2.66E-12
nb_L2	-1.62896764	1.81E-06	1.50E-05	1.93E-10	-1.62897195	8.41E-09	1.91E-10	1.04E-13
nb_L2_Bessel	-0.10256951	2.03E-09	5.27E-12	1.85E-12	-0.10256950	6.16E-09	1.42E-12	5.25E-17
qssp30new	-6.49667496	2.85E-09	7.03E-10	3.22E-10	-6.49667573	1.80E-09	1.83E-11	8.08E-13
qssp30old	6.52338286	6.20E-03	0.0022	1.66E-12	6.49667571	6.26E-10	4.36E-09	9.47E-13
qssp60new	-6.56269019	1.98E-07	1.28E-07	1.10E-06	-6.56270646	3.93E-09	9.39E-11	6.03E-13
qssp60old	6.60855184	9.90E-03	0.004	5.40E-12	6.56270644	1.38E-09	8.83E-09	2.50E-16
Average		1.24E-03	4.78E-04	8.46E-08		8.21E-09	2.99E-09	6.33E-13

Table 2.5: Comparison in terms of solution time and the number of iterations.

Instance	**Dikin-type affine scaling**				**SeDuMi**			**MOSEK**			**SDPT3**		
	CPU	#Iter	#cent.	releig	CPU	#Iter	releig	CPU	#Iter	releig	CPU	#Iter	releig
nql30new	1.3	33	0	1.96E-09	0.3	15	2.41E-09	0.2	13	0	0.9	26	0
nql30old	3.8	41	1	3.89E-10	1.6	18	1.79E-09	0.6	20	0	1.6	30	0
nql60new	5.2	34	0	2.45E-09	1.0	14	2.48E-09	0.7	14	0	3.9	27	0
nql60old	20.4	66	4	8.32E-10	7.1	22	2.48E-09	1.8	20	0	7.3	29	0
nql180new	113.8	74	2	9.14E-10	13.2	16	1.51E-09	7.5	15	0	32.4	33	0
nb	2.1	46	2	6.61E-13	0.6	20	5.05E-15	0.3	20	0	0.4	22	0
nb_L1	1.9	31	0	3.03E-11	0.9	18	2.38E-11	0.3	16	0	3.2	30	0
nb_L2	3.4	53	2	2.34E-10	0.8	16	2.84E-12	0.3	13	0	0.5	15	0
nb_L2_Bessel	1.8	34	0	2.54E-12	0.5	16	5.53E-12	0.2	10	0	0.4	20	0
qssp30new	5.9	93	12	5.02E-12	0.5	20	4.78E-11	0.3	17	0	0.6	20	0
qssp30old	21.8	85	12	8.44E-11	2.5	12	2.94E-05	1.0	22	0	3.7	18	0
qssp60new	45.7	150	29	1.25E-10	2.8	27	9.60E-12	0.8	19	0	2.8	23	0
qssp60old	171.0	127	25	5.96E-11	19.6	18	6.87E-06	6.7	26	0	33.1	19	0
Average	30.6	67	7	5.45E-10	4.0	18	2.79E-06	1.6	17	0	7.0	24	0

Chapter 3

Identification of the optimal partition

In this chapter, we investigate the identification of the optimal partition for SDO and SOCO using the sequence of interior solutions on or in the vicinity of the central path. Section 3.1 investigates the identification of the optimal partition for SDO and Section 3.2 reviews the identification of the optimal partition for SOCO from [171]. Note that we use the same terminology for the optimal partitions of SDO and SOCO.

3.1 Identification of the optimal partition for SDO

In case of degeneracy, even for LO, the condition number of the Newton system of search directions goes to infinity, leading to ill-posed systems, during the final iterations of IPMs [71]. It would be helpful, like in LO and LCP [85, 151], if we could avoid this ill-conditioning, by switching over to a rounding procedure, when μ is sufficiently small. This motivates us to study the identification of the optimal partition for SDO. The optimal partition provides unique information about the optimal set of an SDO problem, regardless of nondegeneracy and strict complementarity conditions. As indicated in Section 1.5.1, the optimal partition is uniquely defined for SDO problems with strong duality.

We consider the identification of the optimal partition for SDO. The rationale behind the identification of the optimal partition is closely interconnected with the limiting behavior of the central path and the existence of a maximally complementary solution. Our goal is to approximate the optimal partition of an SDO problem using the limiting behavior of the central path and a bounded sequence of interior solutions in a neighborhood of the central path. We show how the complexity of approximating the optimal partition depends on condition numbers of the problem. Using bounds on the magnitude of the eigenvalues we identify the subsets of the eigenvectors of the interior solutions whose accumulation points form orthonormal bases for the subspaces of the optimal partition. The magnitude of the eigenvalues of an interior solution is quantified by using a condition number and an upper bound on the distance of an interior solution to the optimal set. In contrast to LO, there are certain instances of SDO for which the condition number is doubly exponentially small. We show that even approximation of the optimal partition is notably more expensive than the identification of the optimal partition for LO.

3.1.1 Identification along the central path

We provide a characterization of the optimal partition using the eigenvectors of a central solution, when μ is sufficiently close to 0. In Section 3.1.1.1, we define a condition number and employ an error bound result for linear matrix inequalities (LMIs), see Section A.2.1, to derive an upper bound on the distance of a central solution to the optimal set. In Section 3.1.1.2, we proceed with the approximation of the optimal partition using the condition number and the error bound result specified in Section 3.1.1.1. In Section 3.1.1.3, we measure the accuracy of the approximation of the optimal partition.

Remark 3.1.1. *The concept of the optimal partition is well-defined only when strong duality holds, and without the interior point condition the central path does not exist. It is known that the interior point condition can be made w.l.o.g., by using the self-dual embedding model, see e.g., [34]. Note that in this case, the embedding model is always well-posed (in terms of the interior point condition) even if the original problem is not.*

3.1.1.1 Condition number and error bound

Recall that the central path for ($\mathrm{P_{SDO}}$) and ($\mathrm{D_{SDO}}$) is defined by (1.22), and assume that $Q_{\mathcal{B}}$ and $Q_{\mathcal{N}}$ are known. To derive bounds on the magnitudes of the eigenvalues of X^μ and S^μ on the central path as $\mu \to 0$, we define a condition number σ as

$$\sigma := \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}, \tag{3.1}$$

where

$$\sigma_{\mathcal{B}} := \begin{cases} \displaystyle\max_{\tilde{X} \in \mathcal{P}^*_{\mathrm{SDO}}} \lambda_{\min}(Q_{\mathcal{B}}^T \tilde{X} Q_{\mathcal{B}}) \\ \qquad\qquad = \displaystyle\max_{\bar{Q}_{\mathcal{B}} \in \Gamma_{\mathcal{B}}} \max_{\tilde{X} \in \mathcal{P}^*_{\mathrm{SDO}}} \lambda_{\min}(\bar{Q}_{\mathcal{B}}^T \tilde{X} \bar{Q}_{\mathcal{B}}), & \text{if } n_{\mathcal{B}} > 0, \\ \infty, & \text{if } n_{\mathcal{B}} = 0, \end{cases} \tag{3.2}$$

$$\sigma_{\mathcal{N}} := \begin{cases} \displaystyle\max_{(\tilde{y},\tilde{S}) \in \mathcal{D}^*_{\mathrm{SDO}}} \lambda_{\min}(Q_{\mathcal{N}}^T \tilde{S} Q_{\mathcal{N}}) \\ \qquad\qquad = \displaystyle\max_{\bar{Q}_{\mathcal{N}} \in \Gamma_{\mathcal{N}}} \max_{(\tilde{y},\tilde{S}) \in \mathcal{D}^*_{\mathrm{SDO}}} \lambda_{\min}(\bar{Q}_{\mathcal{N}}^T \tilde{S} \bar{Q}_{\mathcal{N}}), & \text{if } \ n_{\mathcal{N}} > 0, \\ \infty, & \text{if } \ n_{\mathcal{N}} = 0. \end{cases} \tag{3.3}$$

The condition number σ is indeed a generalization of the analogous condition number from LO, as introduced by Ye [181].

Lemma 3.1.1. *The condition number σ is positive.*

Proof. By the interior point condition, $\mathcal{P}^*_{\text{SDO}} \times \mathcal{D}^*_{\text{SDO}}$ is nonempty and compact. Thus, σ is well-defined by Remark 1.5.3. Assume that $n_{\mathcal{B}} > 0$. Then there exists $\check{X} \in \mathcal{P}^*_{\text{SDO}}$ so that $\lambda_{\min}(Q_{\mathcal{B}}^T \check{X} Q_{\mathcal{B}}) > 0$. By the compactness of $\mathcal{P}^*_{\text{SDO}}$ and the continuity of the eigenvalues, there exists $\bar{X} \in \mathcal{P}^*_{\text{SDO}}$ so that

$$\max_{\tilde{X} \in \mathcal{P}^*_{\text{SDO}}} \lambda_{\min}(Q_{\mathcal{B}}^T \tilde{X} Q_{\mathcal{B}}) = \lambda_{\min}(Q_{\mathcal{B}}^T \bar{X} Q_{\mathcal{B}}) \geq \lambda_{\min}(Q_{\mathcal{B}}^T \check{X} Q_{\mathcal{B}}) > 0,$$

which implies that $\sigma_{\mathcal{B}} > 0$. A similar argument can be made to show that $\sigma_{\mathcal{N}} > 0$ if $n_{\mathcal{N}} > 0$. Consequently, it holds that $\sigma > 0$. $\square$

Remark 3.1.2. *In Appendix A.3, we provide a positive lower bound on the condition number σ as*

$$\sigma \geq \min\left\{\frac{1}{r_{\mathcal{P}^*_{\text{SDO}}} \sum_{i=1}^m \|A^i\|_F}, \frac{1}{r_{\mathcal{D}^*_{\text{SDO}}}}\right\}, \tag{3.4}$$

where

$$\log_2(r_{\mathcal{P}^*_{\text{SDO}}}) = (L+2)\Big(\max\{n,3\}(6n^2+2n+m)\Big)^{5n^2+2m},$$
$$\log_2(r_{\mathcal{D}^*_{\text{SDO}}}) = (L+2)\Big(\max\{n,3\}(7n^2+2n+2m)\Big)^{6n^2+m},$$

in which L is the binary length of the largest absolute value of the input data, when the problem is given by integers. See Lemma A.3.2 for the proof.

For LO, the condition number σ may be in the order of 2^{-L}. However, there are instances of SDO for which σ is doubly exponentially small, as the following example illustrates.

Example 3.1.1. *Consider Khachiyan's example which is adopted from [138]:*

$$\begin{aligned}
\max \quad & y_1 \\
\text{s.t.} \quad & G_i(y) := \begin{pmatrix} y_1 & 2y_i \\ 2y_i & y_{i+1} \end{pmatrix} \succeq 0, \qquad i = 1, \ldots, \bar{m}, \\
& y_{\bar{m}+1} \leq 1.
\end{aligned}$$

This problem can be represented in dual form ($\mathrm{D_{SDO}}$) *if we define*

$$\begin{aligned}
A^1 &= \begin{pmatrix} -1 & -2 \\ -2 & 0 \end{pmatrix} \oplus \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \oplus \ldots \oplus \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \oplus 0, \\
A^{i+1} &= \mathbf{0}_{2(i-1)\times 2(i-1)} \oplus \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \oplus \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix} \oplus \mathbf{0}_{(2(\bar{m}-i)-1)\times(2(\bar{m}-i)-1)}, \\
A^{\bar{m}+1} &= \mathbf{0}_{2(\bar{m}-1)\times 2(\bar{m}-1)} \oplus \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \oplus 1, \\
C &= \mathbf{0}_{2\bar{m}\times 2\bar{m}} \oplus 1, \\
b &= (1, \mathbf{0})^T,
\end{aligned}$$

where $i = 1, \ldots, \bar{m}-1$, $m = \bar{m}+1$, $n = 2\bar{m}+1$, *and the direct sum* $\oplus$ *forms a block diagonal matrix, i.e.,*

$$X \oplus S := \begin{pmatrix} X & 0 \\ 0 & S \end{pmatrix}.$$

From the LMIs we can observe that the volume of the feasible set is doubly exponentially small, since we have $4^{2^i-1}y_1 \leq y_{i+1}$ *and* $y_{i+1} \leq 1$ *for all* $i = 1, \ldots, \bar{m}$. *The optimal solution is unique, and it is given by* $y^*_{i+1} = 4^{2^i-2^{\bar{m}}}$ *for* $i = 0, \ldots, \bar{m}$.

Since $y^*_1 y^*_{i+1} = 4(y^*_i)^2$ *and* $G_i(.)$ *is a* 2×2 *matrix, we get*

$$\lambda_{\max}(G_i(y^*)) = \mathrm{Trace}(G_i(y^*)) = y^*_1 + y^*_{i+1} = 4^{1-2^{\bar{m}}} + 4^{2^i-2^{\bar{m}}}, \quad i = 1, \ldots, \bar{m}.$$

Therefore, an upper bound on the condition number σ is given by

$$\begin{aligned}\sigma \le \sigma_{\mathcal{N}} &= \lambda_{\min}\Big(Q_{\mathcal{N}}^T\Big(G_1(y^*) \oplus \ldots \oplus G_{\bar{m}}(y^*) \oplus (1 - y^*_{\bar{m}+1})\Big)Q_{\mathcal{N}}\Big) \\ &= \min_{i \in \{1,\ldots,\bar{m}\}} \{\lambda_{\max}(G_i(y^*))\} \\ &= 4^{1-2^{\bar{m}}} + 4^{2-2^{\bar{m}}} = 20 \times 4^{-2^{\bar{m}}}.\end{aligned}$$

Even though the lower bound (3.4) is doubly exponentially small, it is not too far from the actual value of σ for some instances of SDO. In fact, all this only indicates that an SDO problem is, in general, harder to solve exactly than an LO problem.

Example 3.1.2. *From (3.4) we get a doubly exponentially small lower bound on σ. Consider the SDO problem in Example 3.1.1 for which we have $\sigma \le 20 \times 4^{-2^{\bar{m}}}$. Given $n_{\mathcal{B}} \le 2\bar{m}+1$, $n_{\mathcal{N}} \le 2\bar{m}+1$, $\|A^1\|_F = \sqrt{\bar{m}+8}$, $\|A^{i+1}\|_F = 3$ for $i = 1,\ldots,\bar{m}-1$, $\|A^{\bar{m}+1}\|_F = \sqrt{2}$, and $L = \ell(2) = 1 + \lceil \log_2(3) \rceil = 3$, see (A.10), we can compute the lower bound (3.4). To do so, we have*

$$\bar{t}_p \le 6(2\bar{m}+1)^2 + 2(2\bar{m}+1) + \bar{m} + 1, \qquad \bar{t}_d \le 7(2\bar{m}+1)^2 + 2(2\bar{m}+1) + 2\bar{m} + 2,$$

$$\bar{s}_p \le 5(2\bar{m}+1)^2 + 2\bar{m} + 2, \qquad \bar{s}_d \le 6(2\bar{m}+1)^2 + \bar{m} + 1,$$

$$\bar{d}_p = \bar{d}_d \le 2\bar{m}+1,$$

$$\sum_{i=1}^{m} \|A^i\|_F = \sqrt{\bar{m}+8} + 3(\bar{m}-1) + \sqrt{2}.$$

Therefore, we get

$$\begin{aligned}\log(r_{\mathcal{P}^*_{\mathrm{SDO}}}) &= 5 \times (48\bar{m}^3 + 82\bar{m}^2 + 47\bar{m} + 9)^{20\bar{m}^2+22\bar{m}+7}, \\ \log(r_{\mathcal{D}^*_{\mathrm{SDO}}}) &= 5 \times (56\bar{m}^3 + 96\bar{m}^2 + 56\bar{m} + 11)^{24\bar{m}^2+25\bar{m}+7}.\end{aligned}$$

Consequently,

$$\begin{aligned}\sigma \ge \min\Big\{ &\big(\sqrt{\bar{m}+8} + 3(\bar{m}-1) + \sqrt{2}\big)2^{-5\times(48\bar{m}^3+82\bar{m}^2+47\bar{m}+9)^{20\bar{m}^2+22\bar{m}+7}}, \\ &2^{-5\times(56\bar{m}^3+96\bar{m}^2+56\bar{m}+11)^{24\bar{m}^2+25\bar{m}+7}}\Big\}.\end{aligned}$$

Remark 3.1.3. *One should be cautioned that computing an exact solution of an LO problem might be difficult too. More precisely, the condition number σ might be so small for an LO problem that very high accuracy is needed for the computation of an exact solution, far beyond the double precision arithmetic commonly used today. For instance, it might be extremely hard to exactly solve an LO problem with a Hilbert matrix of size larger than 20, regardless of the algorithm used.*

In what follows, we resort to a Hölderian error bound result for an LMI system from Theorem A.2.4. The Hölderian bound depends on the degree of singularity of the LMI system, see Section A.2.1. In Lemma 3.1.2, we employ the error bound result to specify an upper bound on the distance of a central solution to the optimal set. In Section 3.1.1.2, we use this upper bound along with the condition number σ to derive bounds on the magnitude of the eigenvalues of the central solutions.

Let $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$ be a primal-dual optimal solution. Then the primal and dual optimal sets can be equivalently written as

$$\begin{cases} \mathcal{A}^s \operatorname{svec}(X) = b, \\ \operatorname{svec}(\tilde{S})^T \operatorname{svec}(X) = 0, \\ X \succeq 0, \end{cases} \qquad \begin{cases} (\mathcal{A}^s)^T y + \operatorname{svec}(S) = \operatorname{svec}(C), \\ \operatorname{svec}(\tilde{X})^T \operatorname{svec}(S) = 0, \\ S \succeq 0, \end{cases} \tag{3.5}$$

see also Section 4 in [167]. Then the minimal subspaces containing the primal and dual optimal sets are given by

$$\begin{aligned} \bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SDO}}} &:= (\operatorname{Ker}(\mathcal{A}^s) \cap (\mathbb{R} \operatorname{svec}(\tilde{S}))^{\perp}) + \mathbb{R} \operatorname{svec}(\tilde{X}), \\ \bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SDO}}} &:= (\mathcal{R}((\mathcal{A}^s)^T) \cap (\mathbb{R} \operatorname{svec}(\tilde{X}))^{\perp}) + \mathbb{R} \operatorname{svec}(\tilde{S}), \end{aligned}$$

where $\mathbb{R} \operatorname{svec}(\tilde{X})$ and $\mathbb{R} \operatorname{svec}(\tilde{S})$ denote the set of all multiples of $\operatorname{svec}(\tilde{X})$ and $\operatorname{svec}(\tilde{S})$, respectively. From the primal-dual feasibility constraints we have

$$\langle X^{\mu} - \tilde{X}, S^{\mu} - \tilde{S} \rangle = 0,$$

which by (1.22) and the optimality of $\tilde{X}$ and $\tilde{S}$ gives

$$\langle X^{\mu}, \tilde{S}\rangle + \langle \tilde{X}, S^{\mu}\rangle = n\mu.$$

All this implies that $0 \le \operatorname{svec}(\tilde{S})^T \operatorname{svec}\big(X^{\mu}\big) \le n\mu$ and $0 \le \operatorname{svec}(\tilde{X})^T \operatorname{svec}\big(S^{\mu}\big) \le n\mu$. Therefore, by the orthogonal projection of $\operatorname{svec}\big(X^{\mu}\big)$ and $\operatorname{svec}\big(S^{\mu}\big)$ onto the affine subspaces

$$\begin{aligned}
&\big\{x \in \mathbb{R}^{n(n+1)/2} \mid x \in \operatorname{svec}(\tilde{X}) + \operatorname{Ker}(\mathcal{A}^s),\ \operatorname{svec}(\tilde{S})^T x = 0\big\},\\
&\big\{s \in \mathbb{R}^{n(n+1)/2} \mid s \in \operatorname{svec}(\tilde{S}) + \mathcal{R}\big((\mathcal{A}^s)^T\big),\ \operatorname{svec}(\tilde{X})^T s = 0\big\},
\end{aligned}$$

we get

$$\begin{aligned}
\operatorname{dist}\big(\operatorname{svec}\big(X^{\mu}\big),\ \{x \mid x \in \operatorname{svec}(\tilde{X}) + \operatorname{Ker}(\mathcal{A}^s),\ \operatorname{svec}(\tilde{S})^T x = 0\}\big) &\le \varpi_p n\mu,\\
\operatorname{dist}\big(\operatorname{svec}\big(S^{\mu}\big),\ \{s \mid s \in \operatorname{svec}(\tilde{S}) + \mathcal{R}\big((\mathcal{A}^s)^T\big),\ \operatorname{svec}(\tilde{X})^T s = 0\}\big) &\le \varpi_d n\mu,
\end{aligned}$$

in which ϖ_p and ϖ_d depend on $\mathcal{A}^s$ and $\tilde{S}$, and $\mathcal{A}^s$ and $\tilde{X}$, respectively. Interestingly, ϖ_p and ϖ_d can be interpreted as Hoffman condition numbers, see Theorem A.2.1. To see this, note that the application of the Hoffman error bound gives

$$\begin{aligned}
&\operatorname{dist}\big(\operatorname{svec}\big(X^{\mu}\big),\ \{x \mid x \in \operatorname{svec}(\tilde{X}) + \operatorname{Ker}(\mathcal{A}^s),\ \operatorname{svec}(\tilde{S})^T x = 0\}\big)\\
&\quad= \operatorname{dist}\big(\operatorname{svec}\big(X^{\mu}\big),\ \big\{x \mid \mathcal{A}^s x = b,\ \operatorname{svec}(\tilde{S})^T x = 0\big\}\big)\\
&\quad\le \varpi_p\big(\|\mathcal{A}^s \operatorname{svec}\big(X^{\mu}\big) - b\|_2 + \operatorname{svec}(\tilde{S})^T \operatorname{svec}\big(X^{\mu}\big)\big)\\
&\quad= \varpi_p \operatorname{svec}(\tilde{S})^T \operatorname{svec}\big(X^{\mu}\big) \le \varpi_p n\mu.
\end{aligned} \tag{3.6}$$

Analogously, we can derive

$$\begin{aligned}
&\operatorname{dist}\big(\operatorname{svec}\big(S^{\mu}\big),\ \{s \mid s \in \operatorname{svec}(\tilde{S}) + \mathcal{R}\big((\mathcal{A}^s)^T\big),\ \operatorname{svec}(\tilde{X})^T s = 0\}\big)\\
&\quad= \operatorname{dist}\big(\operatorname{svec}\big(S^{\mu}\big),\ \big\{s \mid \exists\, y \in \mathbb{R}^m,\ (\mathcal{A}^s)^T y + s = \operatorname{svec}(C),\ \operatorname{svec}(\tilde{X})^T s = 0\big\}\big)\\
&\quad\le \varpi_d\big(\|(\mathcal{A}^s)^T y^{\mu} + \operatorname{svec}\big(S^{\mu}\big) - \operatorname{svec}(C)\|_2 + \operatorname{svec}(\tilde{X})^T \operatorname{svec}\big(S^{\mu}\big)\big)\\
&\quad= \varpi_d \operatorname{svec}(\tilde{X})^T \operatorname{svec}\big(S^{\mu}\big) \le \varpi_d n\mu.
\end{aligned} \tag{3.7}$$

As a consequence, if

$$\mu \leq \hat{\mu} := \frac{1}{n} \min\left\{\varpi_p^{-1}, \varpi_d^{-1}\right\}, \tag{3.8}$$

then it holds that

$$\operatorname{dist}\Big(\operatorname{svec}\big(X^\mu\big),\ \bar{\mathcal{L}}_{\mathcal{P}}\Big) \leq 1, \qquad \operatorname{dist}\Big(\operatorname{svec}(S^\mu),\ \bar{\mathcal{L}}_{\mathcal{D}}\Big) \leq 1.$$

Now, we present the following lemma, as planned.

Lemma 3.1.2. *Let $\big(X^\mu, y^\mu, S^\mu\big)$ be a central solution with $\mu \leq \hat{\mu}$. Then there exist $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$, a positive condition number κ independent of μ, and an exponent $\gamma > 0$ so that*

$$\big\|X^\mu - X_\mu\big\|_F \leq \kappa(n\mu)^\gamma, \qquad \big\|S^\mu - S_\mu\big\|_F \leq \kappa(n\mu)^\gamma, \tag{3.9}$$

where γ depends on the degree of singularity of $\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SDO}}}$ and $\bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SDO}}}$.

Proof. The bounds in (3.9) can be established easily by applying the error bound result, as stated in Theorem A.2.4, to the LMIs in (3.5). As defined by (1.22), the set of central solutions (X^μ, y^μ, S^μ) for $0 < \mu \leq \hat{\mu}$ is bounded, see e.g., Lemma 3.2 in [32]. Therefore, from Theorem A.2.4 and the compactness of the optimal set it follows the existence of $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$, positive condition numbers κ_1 and κ_2 both independent of μ, and positive exponents γ_1 and γ_2 so that

$$\big\|X^\mu - X_\mu\big\|_F \leq \kappa_1(n\mu)^{\gamma_1}, \qquad \big\|S^\mu - S_\mu\big\|_F \leq \kappa_2(n\mu)^{\gamma_2},$$

where

$$\gamma_1 = 2^{-d(\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SDO}}}, \mathbb{S}^n_+)}, \qquad \gamma_2 = 2^{-d(\bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SDO}}}, \mathbb{S}^n_+)},$$

in which $d(\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SDO}}}, \mathbb{S}^n_+)$ and $d(\bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SDO}}}, \mathbb{S}^n_+)$ denote the degree of singularity of the subspaces $\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SDO}}}$ and $\bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SDO}}}$, respectively. Setting $\gamma := \min\{\gamma_1, \gamma_2\}$ and $\kappa := \max\{\kappa_1, \kappa_2\}$ we get the result as desired. $\square$

Remark 3.1.4. *From Theorem A.2.5 we can get a nontrivial upper bound $n-1$ on the degree of singularity. Therefore, we have $\gamma \geq 2^{1-n}$ for $n \geq 2$. However, we are not aware of any method to compute an upper bound on the condition number κ.*

Remark 3.1.5. *For the special case $\mathcal{T} = \{0\}$, the degree of singularity is at most 1 [167]. For instance, this special case happens when we embed an LO problem in SDO. Then Lemma 3.1.2 gives an upper bound $\mathcal{O}(\sqrt{n\mu})$ on the distance of a central solution to the optimal set. However, a direct application of the Hoffman error bound to the linear system of the optimality conditions results in the upper bound $\mathcal{O}(n\mu)$.*

3.1.1.2 Approximation of the optimal partition

Consider the orthogonal transformation of X^μ with respect to Q denoted by

$$\hat{X}^\mu := \begin{pmatrix} \hat{X}^\mu_{\mathcal{B}} & \hat{X}^\mu_{\mathcal{B}\mathcal{T}} & \hat{X}^\mu_{\mathcal{B}\mathcal{N}} \\ \hat{X}^\mu_{\mathcal{T}\mathcal{B}} & \hat{X}^\mu_{\mathcal{T}} & \hat{X}^\mu_{\mathcal{T}\mathcal{N}} \\ \hat{X}^\mu_{\mathcal{N}\mathcal{B}} & \hat{X}^\mu_{\mathcal{N}\mathcal{T}} & \hat{X}^\mu_{\mathcal{N}} \end{pmatrix}, \tag{3.10}$$

where $\hat{X}^\mu := Q^T X^\mu Q$. The orthogonal transformation of S^μ is defined analogously. Since the central path converges to a maximally complementary optimal solution, from the orthogonal transformation in (3.10) we have

$$\lim_{\mu \to 0} \hat{X}^\mu_{\mathcal{B}} = U_{X^{**}}, \qquad \text{and} \qquad \lim_{\mu \to 0} \hat{S}^\mu_{\mathcal{N}} = U_{S^{**}},$$

and

$$\lim\nolimits_{\mu \to 0} Q^T_{\mathcal{T}\cup\mathcal{N}} X^\mu Q_{\mathcal{T}\cup\mathcal{N}} = 0, \qquad \lim\nolimits_{\mu \to 0} Q^T_{\mathcal{B}\cup\mathcal{T}} S^\mu Q_{\mathcal{B}\cup\mathcal{T}} = 0,$$

where $\hat{X}^\mu_{\mathcal{B}} = Q^T_{\mathcal{B}} X^\mu Q_{\mathcal{B}}$ and $\hat{S}^\mu_{\mathcal{N}} = Q^T_{\mathcal{N}} S^\mu Q_{\mathcal{N}}$, and (X^{**}, y^{**}, S^{**}) denotes the limit point of the central path. The following lemma establishes upper bounds on the vanishing blocks of $\hat{X}^\mu$ and $\hat{S}^\mu$.

Lemma 3.1.3. *Let* $\left(X^\mu, y^\mu, S^\mu\right)$ *be a central solution with* $\mu \le \hat{\mu}$. *Then we have*

$$\operatorname{Trace}(\hat{X}^\mu_{\mathcal{N}}) \le \frac{n\mu}{\sigma}, \qquad \operatorname{Trace}(\hat{S}^\mu_{\mathcal{B}}) \le \frac{n\mu}{\sigma},$$
$$\left\|Q^T_{\mathcal{T}\cup\mathcal{N}} X^\mu Q_{\mathcal{T}\cup\mathcal{N}}\right\|_F \le \kappa(n\mu)^\gamma, \qquad \left\|Q^T_{\mathcal{B}\cup\mathcal{T}} S^\mu Q_{\mathcal{B}\cup\mathcal{T}}\right\|_F \le \kappa(n\mu)^\gamma.$$

Proof. By the compactness of $\mathcal{P}^*_{\text{SDO}}$, and the continuity of the eigenvalues, there exists $\bar{X} \in \mathcal{P}^*_{\text{SDO}}$ so that $\sigma_{\mathcal{B}} = \lambda_{\min}(Q^T_{\mathcal{B}}\bar{X}Q_{\mathcal{B}})$ as defined in (3.2). Analogously, it follows from (3.3) that $\sigma_{\mathcal{N}} = \lambda_{\min}(Q^T_{\mathcal{N}}\bar{S}Q_{\mathcal{N}})$ for some $(\bar{y}, \bar{S}) \in \mathcal{D}^*_{\text{SDO}}$. Since $\sigma = \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}$, then there exists $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^*_{\text{SDO}} \times \mathcal{D}^*_{\text{SDO}}$ so that

$$\lambda_{\min}(U_{\bar{X}}) \ge \sigma, \qquad \lambda_{\min}(U_{\bar{S}}) \ge \sigma, \tag{3.11}$$

where $U_{\bar{X}} = Q^T_{\mathcal{B}}\bar{X}Q_{\mathcal{B}}$ and $U_{\bar{S}} = Q^T_{\mathcal{N}}\bar{S}Q_{\mathcal{N}}$. Recall from the feasibility constraints that

$$\langle X^\mu - \bar{X}, S^\mu - \bar{S}\rangle = 0,$$

which by (1.22) and optimality of $\bar{X}$ and $\bar{S}$ gives

$$\langle X^\mu, \bar{S}\rangle + \langle \bar{X}, S^\mu\rangle = n\mu.$$

Since the inner product is invariant with respect to an orthogonal transformation, we get

$$\langle X^\mu, \bar{S}\rangle + \langle \bar{X}, S^\mu\rangle = \langle \hat{X}^\mu_{\mathcal{N}}, U_{\bar{S}}\rangle + \langle U_{\bar{X}}, \hat{S}^\mu_{\mathcal{B}}\rangle = n\mu,$$

where $\hat{S}^\mu_{\mathcal{B}} = Q^T_{\mathcal{B}} S^\mu Q_{\mathcal{B}}$ and $\hat{X}^\mu_{\mathcal{N}} = Q^T_{\mathcal{N}} X^\mu Q_{\mathcal{N}}$. Therefore, the positive definiteness of $\hat{X}^\mu_{\mathcal{N}}$ gives rise to $\langle \hat{X}^\mu_{\mathcal{N}}, U_{\bar{S}}\rangle \le n\mu$. Furthermore, from the inequality $\lambda_{\min}(U_{\bar{S}}) \operatorname{Trace}(\hat{X}^\mu_{\mathcal{N}}) \le \langle \hat{X}^\mu_{\mathcal{N}}, U_{\bar{S}}\rangle$, it immediately follows that

$$\lambda_{\min}(U_{\bar{S}}) \operatorname{Trace}(\hat{X}^\mu_{\mathcal{N}}) \le n\mu,$$

which by the lower bounds (3.11) gives

$$\operatorname{Trace}(\hat{X}^{\mu}_{\mathcal{N}}) \leq \frac{n\mu}{\sigma}.$$

In a similar manner, it follows from $\hat{S}^{\mu}_{\mathcal{B}} \succ 0$ that

$$\operatorname{Trace}(\hat{S}^{\mu}_{\mathcal{B}}) \leq \frac{n\mu}{\sigma}.$$

From Lemma 3.1.2 there exists $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^*_{\text{SDO}} \times \mathcal{D}^*_{\text{SDO}}$ so that (3.9) holds. Recall from Theorem 1.5.1 that X_μ can be represented as $Q_{\mathcal{B}} U_{X_\mu} Q_{\mathcal{B}}^T$ where $U_{X_\mu} \succeq 0$. Thus, we have

$$\left\| Q^T_{\mathcal{T}\cup\mathcal{N}} X^{\mu} Q_{\mathcal{T}\cup\mathcal{N}} \right\|_F = \left\| \begin{pmatrix} \hat{X}^{\mu}_{\mathcal{T}} & \hat{X}^{\mu}_{\mathcal{T}\mathcal{N}} \\ \hat{X}^{\mu}_{\mathcal{N}\mathcal{T}} & \hat{X}^{\mu}_{\mathcal{N}} \end{pmatrix} \right\|_F \leq \left\| X^{\mu} - X_{\mu} \right\|_F \leq \kappa (n\mu)^{\gamma},$$

and

$$\left\| Q^T_{\mathcal{B}\cup\mathcal{T}} S^{\mu} Q_{\mathcal{B}\cup\mathcal{T}} \right\|_F = \left\| \begin{pmatrix} \hat{S}^{\mu}_{\mathcal{B}} & \hat{S}^{\mu}_{\mathcal{B}\mathcal{T}} \\ \hat{S}^{\mu}_{\mathcal{T}\mathcal{B}} & \hat{S}^{\mu}_{\mathcal{T}} \end{pmatrix} \right\|_F \leq \left\| S^{\mu} - S_{\mu} \right\|_F \leq \kappa (n\mu)^{\gamma},$$

which completes the proof. □

Let $X^{\mu} = Q^{\mu}\Lambda(X^{\mu})(Q^{\mu})^T$ and $S^{\mu} = Q^{\mu}\Lambda(S^{\mu})(Q^{\mu})^T$ be the eigenvalue decompositions of X^{μ} and S^{μ}, where Q^{μ} denotes a common eigenvector basis. We show in Theorems 3.1.1 and 3.1.2 that it is possible to identify the subsets of columns of Q^{μ} whose accumulation points are orthonormal bases for the subspaces $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$, when μ is sufficiently small.

Lemma 3.1.4 (Theorem 4.5 in [161])**.** *Let $X \in \mathbb{S}^n$ and $Y \in \mathbb{R}^{n\times k}$. Then we have*

$$\begin{aligned} \lambda_{[n-k+1]}(X) + \ldots + \lambda_{[n]}(X) = \min_{Y} \quad & \operatorname{Trace}(Y^T X Y), \\ \text{s.t.} \quad & Y^T Y = I_k. \end{aligned}$$

Theorem 3.1.1. *For a central solution* (X^μ, y^μ, S^μ) *with* $\mu \leq \hat{\mu}$*, where* $\hat{\mu}$ *is given by* (3.8)*, it holds that:*

$$\lambda_{[n-i+1]}\big(S^\mu\big) \leq \frac{n\mu}{\sigma}, \qquad \lambda_{[i]}\big(X^\mu\big) \geq \frac{\sigma}{n}, \qquad i = 1, \ldots, n_{\mathcal{B}}, \tag{3.12}$$

$$\lambda_{[n-i+1]}\big(X^\mu\big) \leq \frac{n\mu}{\sigma}, \qquad \lambda_{[i]}\big(S^\mu\big) \geq \frac{\sigma}{n}, \qquad i = 1, \ldots, n_{\mathcal{N}}, \tag{3.13}$$

$$\frac{\mu}{c\sqrt{n}(n\mu)^\gamma} \leq \lambda_{[i]}\big(X^\mu\big),\ \lambda_{[n-i+1]}\big(S^\mu\big) \leq c\sqrt{n}(n\mu)^\gamma, \quad i = n_{\mathcal{B}}+1, \ldots, n_{\mathcal{B}}+n_{\mathcal{T}}. \tag{3.14}$$

If $n_{\mathcal{T}} > 0$*, then we have*

$$\kappa \geq \frac{\big(\min\{\varpi_p^{-1}, \varpi_d^{-1}\}\big)^{\frac{1}{2}-\gamma}}{n}, \qquad \frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}.$$

Proof. Recall that $\hat{S}^\mu_{\mathcal{B}} = Q_{\mathcal{B}}^T S^\mu Q_{\mathcal{B}}$ and $\hat{X}^\mu_{\mathcal{N}} = Q_{\mathcal{N}}^T X^\mu Q_{\mathcal{N}}$ as defined in (3.10). Then it follows from Lemma 3.1.4 that

$$\lambda_{[n-n_{\mathcal{B}}+1]}(S^\mu) + \ldots + \lambda_{[n]}(S^\mu) \leq \mathrm{Trace}(\hat{S}^\mu_{\mathcal{B}}) \leq \frac{n\mu}{\sigma},$$

$$\lambda_{[n-n_{\mathcal{N}}+1]}(X^\mu) + \ldots + \lambda_{[n]}(X^\mu) \leq \mathrm{Trace}(\hat{X}^\mu_{\mathcal{N}}) \leq \frac{n\mu}{\sigma}.$$

Therefore, noting that $\lambda_{\min}(X^\mu), \lambda_{\min}(S^\mu) > 0$, we get

$$\lambda_{[n-i+1]}(S^\mu) \leq \frac{n\mu}{\sigma}, \qquad i = 1, \ldots, n_{\mathcal{B}},$$

$$\lambda_{[n-i+1]}(X^\mu) \leq \frac{n\mu}{\sigma}, \qquad i = 1, \ldots, n_{\mathcal{N}}.$$

Further, from the centrality condition $\Lambda(X^\mu)\Lambda(S^\mu) = \mu I_n$, we can observe that the i^{th} largest eigenvalue of X^μ and the i^{th} smallest eigenvalue of S^μ have the same eigenvector, which implies $\lambda_{[i]}(X^\mu)\lambda_{[n-i+1]}(S^\mu) = \mu$. Hence, we can derive

$$\lambda_{[i]}(X^\mu) \geq \frac{\sigma}{n}, \qquad i = 1, \ldots, n_{\mathcal{B}},$$

$$\lambda_{[i]}(S^\mu) \geq \frac{\sigma}{n}, \qquad i = 1, \ldots, n_{\mathcal{N}}.$$

It follows from Lemmas 3.1.3 and 3.1.4 and $\mathrm{Trace}(X) \leq \sqrt{n}\|X\|_F$ that

$$\frac{1}{\sqrt{n}}\Big(\lambda_{[n-n_{\mathcal{N}}-n_{\mathcal{T}}+1]}(X^\mu) + \ldots + \lambda_{[n]}(X^\mu)\Big) \leq \left\|Q_{\mathcal{T}\cup\mathcal{N}}^T X^\mu Q_{\mathcal{T}\cup\mathcal{N}}\right\|_F \leq \kappa(n\mu)^\gamma,$$
$$\frac{1}{\sqrt{n}}\Big(\lambda_{[n-n_{\mathcal{B}}-n_{\mathcal{T}}+1]}(S^\mu) + \ldots + \lambda_{[n]}(S^\mu)\Big) \leq \left\|Q_{\mathcal{B}\cup\mathcal{T}}^T S^\mu Q_{\mathcal{B}\cup\mathcal{T}}\right\|_F \leq \kappa(n\mu)^\gamma,$$

which by the centrality condition gives (3.14).

By Theorem 3.1.1, if $n_{\mathcal{T}} > 0$, there exist $n_{\mathcal{T}}$ eigenvalues of X^μ and $n_{\mathcal{T}}$ eigenvalues of S^μ which stay within the interval $\left[\dfrac{\mu}{\kappa\sqrt{n}(n\mu)^\gamma},\ \kappa\sqrt{n}(n\mu)^\gamma\right]$, and thus both converge to 0 as $\mu \to 0$. Then it holds that

$$\kappa\sqrt{n}(n\mu)^\gamma \geq \frac{\mu}{\kappa\sqrt{n}(n\mu)^\gamma} \quad \Rightarrow \quad \kappa^2 n^2 \geq (n\mu)^{1-2\gamma}, \qquad \forall\ 0 < \mu \leq \hat{\mu},$$

which by the definition of $\hat{\mu}$ implies

$$\kappa \geq \frac{\left(\min\{\varpi_p^{-1}, \varpi_d^{-1}\}\right)^{\frac{1}{2}-\gamma}}{n}, \qquad \gamma \leq \frac{1}{2}.$$

The lower bound $\gamma \geq \frac{1}{2^{n-1}}$ follows from Lemma A.2.5. This completes the proof. $\square$

Since the central path is an analytic curve, the eigenvalues of X^μ and S^μ are continuous functions of μ, and the eigenvalues of central solutions converge to the eigenvalues of the limit point of the central path. Hence, one can observe from Theorem 3.1.1 that the eigenvalues of a central solution (X^μ, y^μ, S^μ) can be categorized into three subsets of eigenvalues as follows

1. $\lambda_i(X^\mu)$ converges to a positive value and $\lambda_i(S^\mu)$ converges to 0;

2. $\lambda_i(S^\mu)$ converges to a positive value and $\lambda_i(X^\mu)$ converges to 0;

3. both $\lambda_i(X^\mu)$ and $\lambda_i(S^\mu)$ converge to 0,

where $\lambda_i(X^\mu)$ and $\lambda_i(S^\mu)$ correspond to the i^{th} column of Q^μ. Let $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$ denote the subsets of columns of Q^μ corresponding to the above subsets of eigenvalues, respectively. Since the central path converges to a maximally complementary optimal solution, the accumulation points of $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$, when $\mu \to 0$, form orthonormal bases for the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$, respectively. Section 3.3 in [32] elucidates the details. The following theorem specifies an upper bound on μ which allows for the identification of $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$.

Theorem 3.1.2. *If μ satisfies*

$$\mu < \tilde{\mu} := \min\left\{\frac{1}{n}\left(\frac{\sigma}{\kappa n^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2},\ \hat{\mu}\right\}, \tag{3.15}$$

then we can identify $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$ from Q^μ.

Proof. From inequalities (3.12) and (3.13), we can deduce that the $n_{\mathcal{B}}$ largest eigenvalues of X^μ stay positive while the $n_{\mathcal{B}}$ smallest eigenvalues of S^μ will converge to 0. Similarly, the $n_{\mathcal{N}}$ largest eigenvalues of S^μ will remain positive while the last $n_{\mathcal{N}}$ eigenvalues of X^μ converge to 0 as $\mu \to 0$. Inequalities (3.14) also hint that, if $n_{\mathcal{T}} > 0$, then there should exist a set of $n_{\mathcal{T}}$ eigenvalues of X^μ and S^μ which stay within the interval $[\mu/\kappa\sqrt{n}(n\mu)^\gamma,\ \kappa\sqrt{n}(n\mu)^\gamma]$. Recall that the i^{th} largest eigenvalue of X^μ and the i^{th} smallest eigenvalue of S^μ have the same eigenvector. Thus, we can identify $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$ when $\mu \le \hat{\mu}$ and the intervals $[\mu/\kappa\sqrt{n}(n\mu)^\gamma,\ \kappa\sqrt{n}(n\mu)^\gamma]$, $(0, n\mu/\sigma]$, and $[\sigma/n, \infty)$ are disjoint, i.e., when

$$\frac{n\mu}{\sigma} < \frac{\mu}{\kappa\sqrt{n}(n\mu)^\gamma}, \qquad \kappa\sqrt{n}(n\mu)^\gamma < \frac{\sigma}{n}, \qquad \frac{n\mu}{\sigma} < \frac{\sigma}{n}, \tag{3.16}$$

which, by $\dfrac{\mu}{\kappa\sqrt{n}(n\mu)^\gamma} \le \kappa\sqrt{n}(n\mu)^\gamma$, is equivalent to

$$\mu < \frac{1}{n}\left(\frac{\sigma}{\kappa n^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}}. \tag{3.17}$$

Furthermore, in case that $\mathcal{T} = \{0\}$, μ needs to satisfy

$$\mu \leq \frac{\sigma^2}{n^2}.$$

Finally, $\mu \leq \hat{\mu}$ must hold as well in order to retain the validity of the bounds in (3.9). This completes the proof. $\square$

Remark 3.1.6. *In general, we do not know in advance if the strict complementarity condition holds for a given instance of SDO. Note that* (3.16) *and* (3.17) *imply that if* $n_{\mathcal{T}} > 0$*, then we have*

$$\frac{1}{n}\left(\frac{\sigma}{\kappa n^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}} \leq \frac{\sigma^2}{n^2}.$$

If $n_{\mathcal{T}} = 0$*, then we can make improvement on the bound* (3.15)*. In fact, the bounds* (3.14) *may provide no further information compared to* (3.12) *and* (3.13) *for small values of* μ*. Hence, in order to identify* $Q^{\mu}_{\mathcal{B}}$ *and* $Q^{\mu}_{\mathcal{N}}$ *it is enough to have*

$$\frac{n\mu}{\sigma} < \frac{\sigma}{n},$$

which reduces the bound (3.15) *to* $\mu < \frac{\sigma^2}{n^2}$*. This bound matches the one for LO, see Section 3.3.3 in [151].*

Remark 3.1.7. *Theorem 3.1.1 is used in Section 4.1.1 to derive bounds for the feasibility of a rounded primal-dual optimal solution. In Section 4.1.1,* $\mathcal{R}(Q^{\mu}_{\mathcal{B}})$*,* $\mathcal{R}(Q^{\mu}_{\mathcal{T}})$*, and* $\mathcal{R}(Q^{\mu}_{\mathcal{N}})$ *with* $\mu < \tilde{\mu}$ *are referred to as approximations of the subspaces* $\mathcal{B}$*,* $\mathcal{T}$*, and* $\mathcal{N}$*, respectively. Furthermore, we use Theorem 3.1.1 in Section 5.1.3 to investigate the sensitivity of the approximation of the optimal partition with respect to the perturbation of the objective vector.*

3.1.1.3 Proximity to the optimal partition

We can provide more information about the optimal partition of the problem by measuring the proximity of $\mathcal{R}(Q^\mu_\mathcal{B})$ and $\mathcal{R}(Q^\mu_\mathcal{N})$ to the subspaces $\mathcal{B}$ and $\mathcal{N}$, respectively, for $\mu < \tilde{\mu}$. To that end, we use the approach in [26] which measures the distance between a primal optimal solution $\tilde{X} \in \mathcal{P}^*_{\text{SDO}}$ and its projection onto $Q^\mu_{\mathcal{B}\cup\mathcal{T}}\mathbb{S}_+^{n_\mathcal{B}+n_\mathcal{T}}(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T$, which is a face of the positive semidefinite cone, see Proposition 2.2.14 in [25] for its proof. In fact, $\mathcal{P}^*_{\text{SDO}}$ is contained in the minimal face $Q_\mathcal{B}\mathbb{S}_+^{n_\mathcal{B}}Q_\mathcal{B}^T$ which itself is a face of $Q_{\mathcal{B}\cup\mathcal{T}}\mathbb{S}_+^{n_\mathcal{B}+n_\mathcal{T}}Q^T_{\mathcal{B}\cup\mathcal{T}}$. Analogously, we measure the distance between $\tilde{S}$, where $(\tilde{y},\tilde{S}) \in \mathcal{D}^*_{\text{SDO}}$, and its projection onto $Q^\mu_{\mathcal{T}\cup\mathcal{N}}\mathbb{S}_+^{n_\mathcal{T}+n_\mathcal{N}}(Q^\mu_{\mathcal{T}\cup\mathcal{N}})^T$.

The following technical lemma is in order.

Lemma 3.1.5. *Let (X^μ, y^μ, S^μ) be given so that $\mu \leq \hat{\mu}$. Then we have*

$$\sup_{\tilde{X}\in\mathcal{P}^*_{\text{SDO}}\setminus\{0\}} \frac{\langle S^\mu, \tilde{X}\rangle}{\|\tilde{X}\|_F} \leq \kappa(n\mu)^\gamma,$$

$$\sup_{(\tilde{y},\tilde{S})\in\mathcal{D}^*_{\text{SDO}},\ \tilde{S}\neq 0} \frac{\langle X^\mu, \tilde{S}\rangle}{\|\tilde{S}\|_F} \leq \kappa(n\mu)^\gamma.$$

Proof. Assume that $0 \neq \tilde{X} \in \mathcal{P}^*_{\text{SDO}}$ is given. Then for all $(\tilde{y},\tilde{S}) \in \mathcal{D}^*_{\text{SDO}}$ we have

$$\frac{\langle S^\mu, \tilde{X}\rangle}{\|\tilde{X}\|_F} = \frac{\langle S^\mu - \tilde{S} + \tilde{S}, \tilde{X}\rangle}{\|\tilde{X}\|_F} = \frac{\langle S^\mu - \tilde{S}, \tilde{X}\rangle}{\|\tilde{X}\|_F} \leq \|S^\mu - \tilde{S}\|_F.$$

Therefore, we get

$$\sup_{\tilde{X}\in\mathcal{P}^*_{\text{SDO}}\setminus\{0\}} \frac{\langle S^\mu, \tilde{X}\rangle}{\|\tilde{X}\|_F} \leq \min_{(\tilde{y},\tilde{S})\in\mathcal{D}^*_{\text{SDO}}} \|S^\mu - \tilde{S}\|_F \leq \kappa(n\mu)^\gamma,$$

where the last inequality follows from Lemma 3.1.2. The proof for the second part follows analogously. □

Theorem 3.1.3. *Let (X^μ, y^μ, S^μ) be given so that $\mu < \tilde{\mu}$. Then for all $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}}$ we have*

$$\|\tilde{X} - \tilde{X}_{\mathcal{F}_{\mathcal{BT}}}\|_F \leq \sqrt{2}\|\tilde{X}\|_F \sqrt{\frac{\kappa n (n\mu)^\gamma}{\sigma}}, \tag{3.18}$$

$$\|\tilde{S} - \tilde{S}_{\mathcal{F}_{\mathcal{TN}}}\|_F \leq \sqrt{2}\|\tilde{S}\|_F \sqrt{\frac{\kappa n (n\mu)^\gamma}{\sigma}}, \tag{3.19}$$

where $\tilde{X}_{\mathcal{F}_{\mathcal{BT}}}$ and $\tilde{S}_{\mathcal{F}_{\mathcal{TN}}}$ denote the projection of $\tilde{X}$ and $\tilde{S}$ onto the faces $\mathcal{F}_{\mathcal{BT}}$ and $\mathcal{F}_{\mathcal{TN}}$, respectively, in which

$$\begin{aligned}
\mathcal{F}_{\mathcal{BT}} &:= Q^\mu_{\mathcal{B}\cup\mathcal{T}} \mathbb{S}^{n_{\mathcal{B}}+n_{\mathcal{T}}}_+ (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T,\\
\mathcal{F}_{\mathcal{TN}} &:= Q^\mu_{\mathcal{T}\cup\mathcal{N}} \mathbb{S}^{n_{\mathcal{T}}+n_{\mathcal{N}}}_+ (Q^\mu_{\mathcal{T}\cup\mathcal{N}})^T.
\end{aligned}$$

Proof. If $\tilde{X} = 0$ or $\tilde{S} = 0$, then $\tilde{X}_{\mathcal{F}_{\mathcal{BT}}} = 0$ or $\tilde{S}_{\mathcal{F}_{\mathcal{TN}}} = 0$, and thus the inequalities (3.18) and (3.19) trivially hold. Note that the projection of $\tilde{X}$ onto the face $\mathcal{F}_{\mathcal{BT}}$ is the optimal solution to

$$\begin{aligned}
\tilde{X}_{\mathcal{F}_{\mathcal{BT}}} :&= \operatorname*{argmin}_{U \succeq 0} \left\| \tilde{X} - Q^\mu_{\mathcal{B}\cup\mathcal{T}} U (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \right\|_F \\
&= \operatorname*{argmin}_{U \succeq 0} \left\| \begin{pmatrix} (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}} - U & (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{N}} \\ (Q^\mu_{\mathcal{N}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}} & (Q^\mu_{\mathcal{N}})^T \tilde{X} Q^\mu_{\mathcal{N}} \end{pmatrix} \right\|_F,
\end{aligned}$$

which is given by $U^* = (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}}$. Then we get

$$\begin{aligned}
\|\tilde{X} - \tilde{X}_{\mathcal{F}_{\mathcal{BT}}}\|_F &= \left\| \tilde{X} - Q^\mu_{\mathcal{B}\cup\mathcal{T}} U^* (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \right\|_F \\
&= \left\| \tilde{X} - Q^\mu_{\mathcal{B}\cup\mathcal{T}} (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}} (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \right\|_F \\
&= \sqrt{\|\tilde{X}\|_F^2 - \left\| (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}} \right\|_F^2} \\
&\leq \|\tilde{X}\|_F \sqrt{1 - \frac{\left\| (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}} \right\|_F^2}{\|\tilde{X}\|_F^2}}.
\end{aligned}$$

Thus, it only remains to derive a lower bound on

$$\frac{\left\| (Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}} \right\|_F}{\|\tilde{X}\|_F}. \tag{3.20}$$

Let us define

$$\Lambda(S^\mu) =: \begin{pmatrix} \Lambda_{\mathcal{B}\cup\mathcal{T}}(S^\mu) & 0 \\ 0 & \Lambda_{\mathcal{N}}(S^\mu) \end{pmatrix}.$$

Then from Lemma 3.1.5 we get

$$\begin{aligned} \langle Q^\mu_{\mathcal{N}}\Lambda_{\mathcal{N}}(S^\mu)(Q^\mu_{\mathcal{N}})^T, \tilde{X}\rangle &\le \langle Q^\mu_{\mathcal{B}\cup\mathcal{T}}\Lambda_{\mathcal{B}\cup\mathcal{T}}(S^\mu)(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T, \tilde{X}\rangle \\ &\quad + \langle Q^\mu_{\mathcal{N}}\Lambda_{\mathcal{N}}(S^\mu)(Q^\mu_{\mathcal{N}})^T, \tilde{X}\rangle \\ &= \langle S^\mu, \tilde{X}\rangle \\ &\le \kappa(n\mu)^\gamma \|\tilde{X}\|_F. \end{aligned}$$

All this implies that

$$\begin{aligned} \min \quad & \left\|(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T X Q^\mu_{\mathcal{B}\cup\mathcal{T}}\right\|_F \\ \text{s.t.} \quad & \langle Q^\mu_{\mathcal{N}}\Lambda_{\mathcal{N}}(S^\mu)(Q^\mu_{\mathcal{N}})^T, X\rangle \le \kappa(n\mu)^\gamma, \\ & \|X\|_F = 1, \\ & X \succeq 0, \end{aligned} \tag{3.21}$$

gives a lower bound on (3.20). Let $\check{X} := (Q^\mu)^T X Q^\mu$, where

$$\check{X} := \begin{pmatrix} \check{X}_{\mathcal{B}\cup\mathcal{T}} & \check{X}_{(\mathcal{B}\cup\mathcal{T})\mathcal{N}} \\ \check{X}_{\mathcal{N}(\mathcal{B}\cup\mathcal{T})} & \check{X}_{\mathcal{N}} \end{pmatrix}.$$

Then auxiliary problem (3.21) is equivalent to

$$\begin{aligned} \min \quad & \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\|_F \\ \text{s.t.} \quad & \langle \Lambda_{\mathcal{N}}(S^\mu), \check{X}_{\mathcal{N}}\rangle \le \kappa(n\mu)^\gamma, \\ & \|\check{X}_{\mathcal{B}\cup\mathcal{T}}\|_F^2 + \|\check{X}_{\mathcal{N}}\|_F^2 + 2\|\check{X}_{(\mathcal{B}\cup\mathcal{T})\mathcal{N}}\|_F^2 = 1, \\ & \check{X} \succeq 0. \end{aligned} \tag{3.22}$$

Since $\check{X} \succeq 0$, we can use the inequality[1] $\left\|\check{X}_{(\mathcal{B}\cup\mathcal{T})\mathcal{N}}\right\|_F^2 \leq \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\|_F \left\|\check{X}_{\mathcal{N}}\right\|_F$ to derive a relaxation of (3.22) as

$$\begin{aligned} \min \quad & \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\|_F \\ \text{s.t.} \quad & \langle \Lambda_{\mathcal{N}}(S^\mu), \check{X}_{\mathcal{N}} \rangle \leq \kappa(n\mu)^\gamma, \\ & \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\|_F + \left\|\check{X}_{\mathcal{N}}\right\|_F \geq 1, \\ & \check{X}_{\mathcal{B}\cup\mathcal{T}} \succeq 0, \\ & \check{X}_{\mathcal{N}} \succeq 0. \end{aligned} \tag{3.23}$$

Finally, from the constraints in (3.23) we get

$$\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\|_F \geq 1 - \|\check{X}_{\mathcal{N}}\|_F \geq 1 - \frac{\kappa(n\mu)^\gamma}{\lambda_{[n_{\mathcal{N}}]}(S^\mu)} \geq 1 - \frac{\kappa n(n\mu)^\gamma}{\sigma} \tag{3.24}$$

$$> 1 - \frac{1}{\sqrt{n}} > 0, \tag{3.25}$$

in which (3.24) follows from (3.13) as well as

$$\lambda_{\min}\big(\Lambda_{\mathcal{N}}(S^\mu)\big)\|\check{X}_{\mathcal{N}}\|_F \leq \langle \Lambda_{\mathcal{N}}(S^\mu), \check{X}_{\mathcal{N}} \rangle \leq \kappa(n\mu)^\gamma,$$

and (3.25) results from $\mu < \tilde{\mu}$. In a similar way as in [26], it can be shown that $1 - \frac{\kappa(n\mu)^\gamma}{\lambda_{[n_{\mathcal{N}}]}(S^\mu)}$ is indeed the optimal value of (3.21). Consequently, we can conclude that

$$\begin{aligned} \left\|\tilde{X} - \tilde{X}_{\mathcal{F}_{\mathcal{B}\mathcal{T}}}\right\|_F &\leq \|\tilde{X}\|_F \sqrt{1 - \frac{\|(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}}\|_F^2}{\|\tilde{X}\|_F^2}} \\ &\leq \|\tilde{X}\|_F \sqrt{2\Big(\frac{\kappa n(n\mu)^\gamma}{\sigma}\Big) - \Big(\frac{\kappa n(n\mu)^\gamma}{\sigma}\Big)^2} \\ &\leq \sqrt{2}\|\tilde{X}\|_F \sqrt{\frac{\kappa n(n\mu)^\gamma}{\sigma}}. \end{aligned}$$

---

[1] The validity of this inequality can be verified by squaring both sides of $\left\|\begin{pmatrix} A & X \\ X^T & B \end{pmatrix}\right\|_F \leq \|A\|_F + \|B\|_F$, which is valid for all positive semidefinite $\begin{pmatrix} A & X \\ X^T & B \end{pmatrix}$. See Theorem 2.1 and Remark 2.3 in [99] for more general results.

Analogously, we can prove that

$$\left\|\tilde{S} - \tilde{S}_{\mathcal{F}_{\mathcal{T}\mathcal{N}}}\right\|_F \leq \|\tilde{S}\|_F \sqrt{1 - \frac{\|(Q^{\mu}_{\mathcal{T}\cup\mathcal{N}})^T \tilde{S} Q^{\mu}_{\mathcal{T}\cup\mathcal{N}}\|_F^2}{\|\tilde{S}\|_F^2}} \leq \sqrt{2}\|\tilde{S}\|_F \sqrt{\frac{\kappa n (n\mu)^{\gamma}}{\sigma}}.$$

$\square$

Under the assumption of primal-dual uniqueness, we can provide an upper bound on the distance between the subspaces $\mathcal{B}$ and $\mathcal{R}(Q^{\mu}_{\mathcal{B}})$, which are of the same dimension if $\mu < \tilde{\mu}$.

Theorem 3.1.4. *Assume that a central solution $(X^{\mu}, y^{\mu}, S^{\mu})$ is given with $\mu < \tilde{\mu}$. Further, let Q be an orthonormal basis partitioned according to $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$, as defined in Section 1.5.1. Then there exist $\rho > 0$ and $\upsilon > 0$ such that*

$$\begin{aligned}
&\operatorname{dist}(\mathcal{B}, \mathcal{R}(Q^{\mu}_{\mathcal{B}})) \leq \min\{2\rho(\sqrt{n}\mu)^{\upsilon}, 1\},\\
&\operatorname{dist}(\mathcal{T}, \mathcal{R}(Q^{\mu}_{\mathcal{T}})) \leq \min\{2\rho(\sqrt{n}\mu)^{\upsilon}, 1\},\\
&\operatorname{dist}(\mathcal{N}, \mathcal{R}(Q^{\mu}_{\mathcal{N}})) \leq \min\{2\rho(\sqrt{n}\mu)^{\upsilon}, 1\}.
\end{aligned}$$

Proof. An orthogonal projection matrix of the subspace $\mathcal{B}$ is given by

$$Q_{\mathcal{B}}(Q_{\mathcal{B}}^T Q_{\mathcal{B}})^{-1} Q_{\mathcal{B}}^T = Q_{\mathcal{B}} Q_{\mathcal{B}}^T.$$

Note that this projection matrix is invariant with respect to any choice of an orthonormal basis for $\mathcal{B}$, see e.g., Section 2.5.1 in [62]. Then we get

$$\begin{aligned}
\operatorname{dist}\left(\mathcal{B}, \mathcal{R}(Q^{\mu}_{\mathcal{B}})\right) &= \left\|Q^{\mu}_{\mathcal{B}}(Q^{\mu}_{\mathcal{B}})^T - Q_{\mathcal{B}}Q_{\mathcal{B}}^T\right\|_2\\
&= \left\|Q^{\mu}_{\mathcal{B}}(Q^{\mu}_{\mathcal{B}})^T - Q_{\mathcal{B}}Q_{\mathcal{B}}^T - Q^{\mu}_{\mathcal{B}}Q_{\mathcal{B}}^T + Q^{\mu}_{\mathcal{B}}Q_{\mathcal{B}}^T\right\|_2\\
&= \left\|Q^{\mu}_{\mathcal{B}}\left((Q^{\mu}_{\mathcal{B}})^T - Q_{\mathcal{B}}^T\right) + \left(Q^{\mu}_{\mathcal{B}} - Q_{\mathcal{B}}\right)Q_{\mathcal{B}}^T\right\|_2\\
&\leq \left\|Q^{\mu}_{\mathcal{B}}\right\|_2\left\|Q^{\mu}_{\mathcal{B}} - Q_{\mathcal{B}}\right\|_2 + \left\|Q^{\mu}_{\mathcal{B}} - Q_{\mathcal{B}}\right\|_2\left\|Q_{\mathcal{B}}\right\|_2\\
&\leq 2\left\|Q^{\mu}_{\mathcal{B}} - Q_{\mathcal{B}}\right\|_2,
\end{aligned}$$

which implies

$$\operatorname{dist}\big(\mathcal{B}, \mathcal{R}(Q^\mu_{\mathcal{B}})\big) \le 2 \min_{\bar{Q}_{\mathcal{B}} \in \Gamma_{\mathcal{B}}} \big\|Q^\mu_{\mathcal{B}} - \bar{Q}_{\mathcal{B}}\big\|_2. \tag{3.26}$$

The set of primal-dual optimal solutions can be represented as a system of polynomial equations and inequalities [7]. Let $\mathcal{C}$ be the set of all solutions

$$(\operatorname{vec}(\tilde{Q}); \operatorname{diag}(\Lambda(\tilde{X})); \tilde{y}; \operatorname{diag}(\Lambda(\tilde{S})))$$

of this system so that $\tilde{Q}\Lambda(\tilde{X})\tilde{Q}^T \in \mathcal{P}^*_{\mathrm{SDO}}$ and $(\tilde{y}, \tilde{Q}\Lambda(\tilde{S})\tilde{Q}^T) \in \mathcal{D}^*_{\mathrm{SDO}}$, where $\operatorname{diag}(.)$ denotes the vector of diagonal entries of a square matrix. By the centrality condition, a central solution (X^μ, y^μ, S^μ) violates the constraints $\Lambda_{ii}(\tilde{X})\Lambda_{ii}(\tilde{S}) = 0$ by μ for $i = 1, \ldots, n$. Since the set of central solutions (X^μ, y^μ, S^μ) with $\mu < \tilde{\mu}$ is contained in a compact set, it follows from (3.26) and Theorem A.2.8 that there exist $\upsilon > 0$ and $\rho > 0$ such that

$$\begin{aligned}
\operatorname{dist}\big(\mathcal{B}, \mathcal{R}(Q^\mu_{\mathcal{B}})\big) &\le 2 \inf_{(Q^*\Lambda(X^*)(Q^*)^T, y^*, Q^*\Lambda(S^*)(Q^*)^T) \in \operatorname{ri}(\mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}})} \Big\| \Big(\operatorname{vec}\big(Q^\mu - Q^*\big); \\
&\qquad \operatorname{diag}\big(\Lambda(X^\mu) - \Lambda(X^*)\big); y^\mu - y^*; \operatorname{diag}\big(\Lambda(S^\mu) - \Lambda(S^*)\big)\Big)\Big\|_2 \\
&= 2 \operatorname{dist}\Big(\Big(\operatorname{vec}\big(Q^\mu\big); \operatorname{diag}\big(\Lambda(X^\mu)\big); y^\mu; \operatorname{diag}\big(\Lambda(S^\mu)\big)\Big),\ \mathcal{C}\Big) \qquad (3.27) \\
&\le 2\rho(\sqrt{n}\mu)^\upsilon,
\end{aligned}$$

where the equality (3.27) follows from the uniqueness of the optimal solution. The proofs for the subspaces $\mathcal{T}$ and $\mathcal{N}$ are analogous. □

Remark 3.1.8. *Since $\mathcal{C}$ is a compact set, there exists a solution*

$$(\operatorname{vec}(\tilde{Q}_\mu); \operatorname{diag}(\Lambda(\tilde{X}_\mu)); \tilde{y}_\mu; \operatorname{diag}(\Lambda(\tilde{S}_\mu)))$$

of $\mathcal{C}$ whose distance from $\big(\operatorname{vec}(Q^\mu); \operatorname{diag}(\Lambda(X^\mu)); y^\mu; \operatorname{diag}(\Lambda(S^\mu))\big)$ is minimal. The assumption of uniqueness in Theorem 3.1.4 can be released if there exists a sequence of common eigenvector bases of maximally complementary solutions converging to $\tilde{Q}_\mu$. More precisely, assume that $Q_k \to \tilde{Q}_\mu$ for a convergent sequence

$(X_k, y_k, S_k) \to (\tilde{X}_\mu, \tilde{y}_\mu, \tilde{S}_\mu)$ *such that* $(X_k, y_k, S_k) \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}})$ *for all* k*, and* $X_k = Q_k \Lambda(X_k) Q_k^T$ *and* $S_k = Q_k \Lambda(S_k) Q_k^T$ *are eigenvalue decompositions. Then equality* (3.27) *holds, and thus the upper bounds in Theorem 3.1.4 are valid regardless of the uniqueness assumption. In particular, this condition holds if* $(\tilde{X}_\mu, \tilde{y}_\mu, \tilde{S}_\mu)$ *is a maximally complementary optimal solution, or if there exists a unique common eigenvector basis for* $(\tilde{X}_\mu, \tilde{y}_\mu, \tilde{S}_\mu)$*. For instance, consider the following SDO problem from [59]:*

$$A^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; A^3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix},$$

$$b = (1,\ 0,\ 0)^T.$$

The primal optimal set can be described as

$$X^*(\delta) = \begin{pmatrix} 1 & 2(\delta-1) & 2(\delta-1) \\ 2(\delta-1) & 4(1-\delta) & 4(1-\delta) \\ 2(\delta-1) & 4(1-\delta) & 4(1-\delta) \end{pmatrix}, \qquad 0 \le \delta \le 1,$$

and the unique dual optimal solution is

$$S^* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \qquad y^* = (1,\ 1,\ 1)^T.$$

One can verify that the eigenvalues of $X^*(\delta)$ *for* $0 \le \delta \le 1$ *are given by*

$$\begin{aligned}
\lambda_{[1]}(X^*(\delta)) &= \frac{1}{2}\sqrt{96\delta^2 - 176\delta + 81} - 4\delta + \frac{9}{2}, \\
\lambda_{[2]}(X^*(\delta)) &= -\frac{1}{2}\sqrt{96\delta^2 - 176\delta + 81} - 4\delta + \frac{9}{2}, \\
\lambda_{[3]}(X^*(\delta)) &= 0.
\end{aligned}$$

Observe that for all $0 < \delta < 1$, $(X^*(\delta), y^*, S^*)$ is strictly complementary, and that the multiplicity of the positive eigenvalues of $X^*(\delta)$ and S^* are 1. Hence, for all $0 < \delta < 1$, the eigenvalue decompositions of $X^*(\delta)$ and S^* are unique up to the sign of columns of the orthogonal matrices.

Suppose that $(\tilde{X}_\mu, \tilde{y}_\mu, \tilde{S}_\mu) = (X_1^*, y^*, S^*)$, which is not a strictly complementary optimal solution. Nevertheless, $\tilde{X}_\mu$ and $\tilde{S}_\mu$ have the unique common eigenvector basis

$$\tilde{Q}_\mu = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Therefore, there exists a sequence of unique common eigenvector bases Q_k, corresponding to $(X^*(\delta), y^*, S^*)$ for $0 < \delta < 1$, which converges to $\tilde{Q}_\mu$.

3.1.2 Identification in a neighborhood of the central path

Thus far, we assumed that the solution given by IPMs is exactly on the central path. In general, however, path-following IPMs operate in a specified vicinity of the central path by computing approximate solutions of (1.22). In this section, we consider a sequence of solutions in the relative interior of the primal-dual feasible set, which has accumulation points in the relative interior of the optimal set.

Consider a solution $(X^\circ, y^\circ, S^\circ) \in \mathrm{ri}(\mathcal{P}_{\mathrm{SDO}} \times \mathcal{D}_{\mathrm{SDO}})$ given by a primal-dual path-following IPM, where $X^\circ = M\Lambda(X^\circ)M^T$ and $S^\circ = N\Lambda(S^\circ)N^T$ are eigenvalue decompositions of X° and S°, respectively, and M and N are orthogonal matrices. In contrast to the result of Theorem 3.1.2, the accumulation points of the subsets of eigenvectors are not identical for X° and S°. The reason lies in the fact that X° and S° do not necessarily commute. For instance, consider the Nesterov-Todd scaling method, where X° and S° are projected onto the same point V defined as

$$V := D^{-\frac{1}{2}} X^\circ D^{-\frac{1}{2}} = D^{\frac{1}{2}} S^\circ D^{\frac{1}{2}},$$

which implies $X^\circ = DS^\circ D$, where $D \succ 0$ denotes the scaling matrix, see [124] for the definition of D. Then the eigenvalue decomposition of S° yields

$$X^\circ = DP\Lambda(S^\circ)P^T D. \tag{3.28}$$

Since $\Lambda(X^\circ)$ and $\Lambda(S^\circ)$ have nonzero diagonal entries, we may assume that $\Lambda(S^\circ) =: \Sigma^{\frac{1}{2}}\Lambda(X^\circ)\Sigma^{\frac{1}{2}}$ where $\Sigma^{\frac{1}{2}}$ is a positive definite diagonal matrix. Hence, from (3.28) we get

$$\Sigma^{-\frac{1}{2}} P^T D^{-1} X^\circ D^{-1} P \Sigma^{-\frac{1}{2}} = \Lambda(X^\circ).$$

Note that $D^{-1}P\Sigma^{-\frac{1}{2}}$ is an $n \times n$ invertible matrix but not necessarily equal to the orthogonal matrix M. Therefore, there exists an invertible matrix $N' \in \mathbb{R}^{n\times n}$ so that

$$M = D^{-1} N \Sigma^{-\frac{1}{2}} N',$$

implying that X° and S° do not necessarily share an eigenvector basis.

The proximity of $(X^\circ, y^\circ, S^\circ)$ to the central path, as defined in (2.6), can be measured by

$$\mathrm{prox}(X^\circ S^\circ) := \frac{\lambda_{\max}(X^\circ S^\circ)}{\lambda_{\min}(X^\circ S^\circ)}, \qquad (X^\circ, y^\circ, S^\circ) \in \mathrm{ri}(\mathcal{P}_{\mathrm{SDO}} \times \mathcal{D}_{\mathrm{SDO}}). \tag{3.29}$$

Notice that $X^\circ S^\circ$ has the same eigenvalues as $(X^\circ)^{\frac{1}{2}} S^\circ (X^\circ)^{\frac{1}{2}}$, i.e., $X^\circ S^\circ$ has real positive eigenvalues even though it is not necessarily symmetric. Further, it follows from (3.29) that $\mathrm{prox}(X^\circ S^\circ) \geq 1$, and the equality holds only when $(X^\circ, y^\circ, S^\circ)$ is on the central path. A neighborhood of the central path is defined by

$$\mathcal{N}_{\mathrm{prox}}(\xi) := \Big\{(X^\circ, y^\circ, S^\circ) \in \mathrm{ri}(\mathcal{P}_{\mathrm{SDO}} \times \mathcal{D}_{\mathrm{SDO}}) \mid \mathrm{prox}(X^\circ S^\circ) \leq \xi\Big\}, \tag{3.30}$$

where $\xi > 1$. Then for $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi)$ we have

$$\lambda_{\min}(X^\circ S^\circ) \leq \lambda_{[i]}(X^\circ S^\circ) \leq \xi \lambda_{\min}(X^\circ S^\circ), \qquad i = 1, \ldots, n. \tag{3.31}$$

Here, we use the application of Weyl theorem[2] in [101] to provide an upper bound on $\lambda_{\min}(X^\circ S^\circ)$.

Lemma 3.1.6 (Corollary 2.3 in [101])**.** *Let X and S be two $n \times n$ symmetric positive semidefinite matrices. Then for $j \leq \min\{\operatorname{rank}(X),\ \operatorname{rank}(S)\}$ we have*

$$\min_{1 \leq i \leq j} \{\lambda_{[i]}(X)\lambda_{[j-i+1]}(S)\} \geq \lambda_{[j]}(XS) \geq \max_{j \leq i \leq n} \{\lambda_{[i]}(X)\lambda_{[n+j-i]}(S)\}. \tag{3.32}$$

Lemma 3.1.7. *Let $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi)$. Then we have*

$$\lambda_{[i]}(X^\circ)\lambda_{[n-i+1]}(S^\circ) \geq \lambda_{\min}(X^\circ S^\circ), \qquad i = 1, \ldots, n. \tag{3.33}$$

Proof. The proof is straightforward from the first inequality in (3.32) and the positive definiteness of X° and S°. In fact, for the special case $k = n$ there holds that

$$\min\left\{\lambda_{[1]}(X^\circ)\lambda_{[n]}(S^\circ),\ \lambda_{[2]}(X^\circ)\lambda_{[n-1]}(S^\circ), \ldots, \lambda_{[n]}(X^\circ)\lambda_{[1]}(S^\circ)\right\} \geq \lambda_{\min}(X^\circ S^\circ),$$

which completes the proof. □

The following theorem generalizes the bounds derived in Theorem 3.1.1 to an approximate solution $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi)$. Analogous to the case of central solutions, we let $M = (M_{\mathcal{B}}, M_{\mathcal{T}}, M_{\mathcal{N}})$ and $N = (N_{\mathcal{B}}, N_{\mathcal{T}}, N_{\mathcal{N}})$ be the subsets of columns of M and N, respectively, associated with the eigenvalues of X° and S° whose accumulation points are positive and zero.

[2] See Theorem 4.3.7 in [84].

Theorem 3.1.5. *Let $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi)$ and $\mu := \frac{\langle X^\circ, S^\circ\rangle}{n}$. Then there exist a positive condition number κ' independent of μ and a positive exponent γ so that*

$$\begin{array}{lll} \lambda_{[n-i+1]}(S^\circ) \leq \dfrac{n\mu}{\sigma}, & \lambda_{[i]}(X^\circ) \geq \dfrac{\sigma}{n\xi}, & i = 1, \ldots, n_{\mathcal{B}}, \\ \lambda_{[n-i+1]}(X^\circ) \leq \dfrac{n\mu}{\sigma}, & \lambda_{[i]}(S^\circ) \geq \dfrac{\sigma}{n\xi}, & i = 1, \ldots, n_{\mathcal{N}}, \\ \dfrac{\mu}{\kappa'\sqrt{n}\xi(n\mu)^\gamma} \leq \lambda_{[i]}(X^\circ), & \lambda_{[n-i+1]}(S^\circ) \leq \kappa'\sqrt{n}(n\mu)^\gamma, & i = n_{\mathcal{B}}+1, \ldots, n_{\mathcal{B}}+n_{\mathcal{T}}. \end{array}$$

If $n_{\mathcal{T}} > 0$, then we have

$$\frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}.$$

If μ satisfies

$$\mu < \min\left\{\frac{1}{n}\left(\frac{\sigma}{\kappa' n^{\frac{3}{2}}\xi}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2\xi},\ \hat{\mu}\right\}, \tag{3.34}$$

then we can identify $M_{\mathcal{B}}$, $M_{\mathcal{T}}$, and $M_{\mathcal{N}}$ from X°, and $N_{\mathcal{B}}$, $N_{\mathcal{T}}$, and $N_{\mathcal{N}}$ from S°.

Proof. The proof technique can be traced back to Theorem 3.1.1 fairly easily. Let $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^*_{\text{SDO}} \times \mathcal{D}^*_{\text{SDO}}$ which satisfies (3.11) and $(\hat{X}^\circ, \hat{S}^\circ)$ denote the orthogonal transformation of (X°, S°) with respect to Q. Then it follows from the orthogonality between $(X^\circ - \bar{X})$ and $(S^\circ - \bar{S})$ that

$$\langle X^\circ, \bar{S}\rangle + \langle \bar{X}, S^\circ\rangle = \langle \hat{X}^\circ_{\mathcal{N}}, U_{\bar{S}}\rangle + \langle U_{\bar{X}}, \hat{S}^\circ_{\mathcal{B}}\rangle = \langle X^\circ, S^\circ\rangle,$$

where $\hat{S}^\circ_{\mathcal{B}} = Q^T_{\mathcal{B}} S^\circ Q_{\mathcal{B}}$ and $\hat{X}^\circ_{\mathcal{N}} = Q^T_{\mathcal{N}} X^\circ Q_{\mathcal{N}}$. Using the inequality $\lambda_{\min}(U_{\bar{S}}) \operatorname{Trace}(\hat{X}^\circ_{\mathcal{N}}) \leq \langle \hat{X}^\circ_{\mathcal{N}}, U_{\bar{S}}\rangle$ and the positive definiteness of X° and S° we have

$$\begin{aligned} \lambda_{\min}(U_{\bar{X}}) \operatorname{Trace}(\hat{S}^\circ_{\mathcal{B}}) &\leq \langle X^\circ, S^\circ\rangle \quad \Rightarrow \quad \operatorname{Trace}(\hat{S}^\circ_{\mathcal{B}}) &&\leq \frac{n\mu}{\sigma}, \\ \lambda_{\min}(U_{\bar{S}}) \operatorname{Trace}(\hat{X}^\circ_{\mathcal{N}}) &\leq \langle X^\circ, S^\circ\rangle \quad \Rightarrow \quad \operatorname{Trace}(\hat{X}^\circ_{\mathcal{N}}) &&\leq \frac{n\mu}{\sigma}, \end{aligned}$$

where the latter inequalities follow from (3.11). Now, Lemma 3.1.4 can be applied to get

$$\begin{aligned}
\lambda_{[n-n_{\mathcal{B}}+1]}(S^\circ) + \ldots + \lambda_{[n]}(S^\circ) \quad &\leq \mathrm{Trace}(\hat{S}^\circ_{\mathcal{B}}) \quad \leq \frac{n\mu}{\sigma}, \\
\lambda_{[n-n_{\mathcal{N}}+1]}(X^\circ) + \ldots + \lambda_{[n]}(X^\circ) \quad &\leq \mathrm{Trace}(\hat{X}^\circ_{\mathcal{N}}) \quad \leq \frac{n\mu}{\sigma},
\end{aligned}$$

which by $X^\circ, S^\circ \succ 0$ implies

$$\begin{aligned}
\lambda_{[n-i+1]}(S^\circ) &\leq \frac{n\mu}{\sigma}, \qquad i = 1, \ldots, n_{\mathcal{B}}, \\
\lambda_{[n-i+1]}(X^\circ) &\leq \frac{n\mu}{\sigma}, \qquad i = 1, \ldots, n_{\mathcal{N}}.
\end{aligned} \tag{3.35}$$

Recall from (3.31) that

$$n\mu = \langle X^\circ, S^\circ \rangle \leq n\xi\lambda_{\min}(X^\circ S^\circ),$$

which yields

$$\frac{\lambda_{\min}(X^\circ S^\circ)}{\mu} \geq \frac{1}{\xi}. \tag{3.36}$$

Then (3.33) and (3.36) can be applied to (3.35) to derive lower bounds on the eigenvalues of X° and S°:

$$\begin{aligned}
\lambda_{[i]}(X^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(S^\circ)} \geq \frac{\sigma\lambda_{\min}(X^\circ S^\circ)}{n\mu} \geq \frac{\sigma}{n\xi}, \qquad i = 1, \ldots, n_{\mathcal{B}}, \\
\lambda_{[i]}(S^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(X^\circ)} \geq \frac{\sigma\lambda_{\min}(X^\circ S^\circ)}{n\mu} \geq \frac{\sigma}{n\xi}, \qquad i = 1, \ldots, n_{\mathcal{N}}.
\end{aligned}$$

For the subspace $\mathcal{T}$ we should note that

$$\left\{(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi) \mid \langle X^\circ, S^\circ \rangle \leq \min\left\{\varpi_p^{-1}, \varpi_d^{-1}\right\}\right\},$$

where ϖ_p and ϖ_d are defined as in (3.6) and (3.7), is a bounded set by the interior point condition and the linear independence of A^i for $i = 1, \ldots, m$, see e.g., Lemma

3.1 in [32]. Furthermore, the amount of constraint violation with respect to the LMI system (3.5) for $(X^\circ, y^\circ, S^\circ)$ is equal to $n\mu$. Hence, the result of Theorem A.2.4 is still valid, i.e., for $0 < \mu \leq \hat{\mu}$ there exist $(X_{\mu^\circ}, y_{\mu^\circ}, S_{\mu^\circ}) \in \mathcal{P}^*_{\text{SDO}} \times \mathcal{D}^*_{\text{SDO}}$, a positive condition number κ' independent of μ, and a positive exponent γ so that

$$\left\|X^\circ - X_{\mu^\circ}\right\|_F \leq \kappa'(n\mu)^\gamma, \qquad \left\|S^\circ - S_{\mu^\circ}\right\|_F \leq \kappa'(n\mu)^\gamma, \tag{3.37}$$

where κ' and γ are defined as in Lemma 3.1.2. Analogous to the proof of Theorem 3.1.1, we can observe, using the orthogonal transformation Q, that

$$\begin{aligned} \left\|Q^T_{\mathcal{T}\cup\mathcal{N}} X^\circ Q_{\mathcal{T}\cup\mathcal{N}}\right\|_F &= \left\| \begin{pmatrix} \hat{X}^\circ_{\mathcal{T}} & \hat{X}^\circ_{\mathcal{T}\mathcal{N}} \\ \hat{X}^\circ_{\mathcal{N}\mathcal{T}} & \hat{X}^\circ_{\mathcal{N}} \end{pmatrix} \right\|_F \leq \left\|X^\circ - X_{\mu^\circ}\right\|_F \leq \kappa'(n\mu)^\gamma, \\ \left\|Q^T_{\mathcal{B}\cup\mathcal{T}} S^\circ Q_{\mathcal{B}\cup\mathcal{T}}\right\|_F &= \left\| \begin{pmatrix} \hat{S}^\circ_{\mathcal{B}} & \hat{S}^\circ_{\mathcal{B}\mathcal{T}} \\ \hat{S}^\circ_{\mathcal{T}\mathcal{B}} & \hat{S}^\circ_{\mathcal{T}} \end{pmatrix} \right\|_F \leq \left\|S^\circ - S_{\mu^\circ}\right\|_F \leq \kappa'(n\mu)^\gamma. \end{aligned} \tag{3.38}$$

Then it follows from Lemma 3.1.4 and (3.38) that

$$\begin{aligned} \lambda_{[n-n_{\mathcal{N}}-n_{\mathcal{T}}+1]}(X^\circ) + \ldots + \lambda_{[n]}(X^\circ) &\leq \kappa'\sqrt{n}(n\mu)^\gamma, \\ \lambda_{[n-n_{\mathcal{B}}-n_{\mathcal{T}}+1]}(S^\circ) + \ldots + \lambda_{[n]}(S^\circ) &\leq \kappa'\sqrt{n}(n\mu)^\gamma, \end{aligned}$$

and consequently,

$$\begin{aligned} \lambda_{[n-i+1]}(X^\circ) &\leq \kappa'\sqrt{n}(n\mu)^\gamma, & i &= 1, \ldots, n_{\mathcal{N}} + n_{\mathcal{T}}, \\ \lambda_{[n-i+1]}(S^\circ) &\leq \kappa'\sqrt{n}(n\mu)^\gamma, & i &= 1, \ldots, n_{\mathcal{B}} + n_{\mathcal{T}}. \end{aligned}$$

Using the bounds in (3.33) and (3.36) we can derive

$$\begin{aligned} \lambda_{[i]}(X^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(S^\circ)} \geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\kappa'\sqrt{n}(n\mu)^\gamma} \geq \frac{\mu}{\kappa'\sqrt{n}\xi(n\mu)^\gamma}, & i &= 1, \ldots, n_{\mathcal{B}} + n_{\mathcal{T}}, \\ \lambda_{[i]}(S^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(X^\circ)} \geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\kappa'\sqrt{n}(n\mu)^\gamma} \geq \frac{\mu}{\kappa'\sqrt{n}\xi(n\mu)^\gamma}, & i &= 1, \ldots, n_{\mathcal{N}} + n_{\mathcal{T}}. \end{aligned}$$

In the sequel, using the same argument as in Theorem 3.1.2, we can identify the subsets of columns of M and N whose accumulation points form orthonormal bases for $\mathcal{B}$, $\mathcal{T}$ and $\mathcal{N}$ if

$$\frac{n\mu}{\sigma} < \frac{\mu}{\kappa'\sqrt{n}\xi(n\mu)^{\gamma}}, \qquad \kappa'\sqrt{n}(n\mu)^{\gamma} < \frac{\sigma}{n\xi}. \tag{3.39}$$

Considering the case $\mathcal{T} = \{0\}$, we can represent (3.39) as

$$\mu < \min\left\{\frac{1}{n}\left(\frac{\sigma}{\kappa' n^{\frac{3}{2}}\xi}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2\xi}\right\}.$$

Including the condition $\mu \leq \hat{\mu}$ gives the result as desired. Further, if $n_{\mathcal{T}} > 0$, from $\dfrac{\mu}{\kappa'\sqrt{n}\xi(n\mu)^{\gamma}} \leq \kappa'\sqrt{n}(n\mu)^{\gamma}$ we get

$$(\kappa')^2 n^2 \xi \geq (n\mu)^{1-2\gamma}, \qquad \forall\ 0 < \mu \leq \hat{\mu},$$

which implies

$$\frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}.$$

This completes the proof. $\square$

Corollary 3.1.1. *Let* $(X^{(0)}, y^{(0)}, S^{(0)}) \in \mathcal{N}_{\text{prox}}(\xi)$ *be an initial solution,* $\mu^{(0)} := \frac{\langle X^{(0)}, S^{(0)}\rangle}{n}$*, and* $\log(.)$ *denote the natural logarithm. Then the Dikin-type primal-dual affine scaling method with steplength* $\alpha = \frac{1}{\xi\sqrt{n}}$ *and the neighborhood* (3.30)*, see Section 2.1, needs at most*

$$\left\lceil \xi n \log\left(\mu^{(0)}\left(\min\left\{\frac{1}{n}\left(\frac{\sigma}{\kappa' n^{\frac{3}{2}}\xi}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2\xi},\ \hat{\mu}\right\}\right)^{-1}\right)\right\rceil$$

iterations to get an $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi)$ *which allows to identify* $(M_{\mathcal{B}}, M_{\mathcal{T}}, M_{\mathcal{N}})$ *and* $(N_{\mathcal{B}}, N_{\mathcal{T}}, N_{\mathcal{N}})$*.*

Proof. The proof easily follows from the iteration complexity result for the Dikin-type primal-dual affine scaling method with steplength $\alpha = \frac{1}{\xi\sqrt{n}}$, see [32]. Then the complementarity gap drops below a threshold ε after

$$\left\lceil \xi n \log\left(\frac{n\mu^{(0)}}{\varepsilon}\right)\right\rceil$$

iterations. The result follows if we replace ε by the right hand side of (3.34) multiplied by n. $\square$

Remark 3.1.9. *In* (3.37)*, we employed the same exponent* γ *as in* (3.9) *but a different condition number* κ'*. In fact, the primal and dual systems in* (3.5) *are used for both Theorems 3.1.1 and 3.1.5. However, it is not known whether* κ *and* κ' *are identical or of the same order.*

3.2 Identification of the optimal partition for SOCO

To identify the optimal partition from a central solution, Terlaky and Wang [171] defined two condition numbers σ_1 and σ_2 as

$$\begin{aligned}
&\sigma_{\mathcal{B}} := \min_{i\in\mathcal{B}} \max_{\tilde{x}\in\mathcal{P}^*_{\text{SOCO}}} \left\{\tilde{x}^i_1 - \|\tilde{x}^i_{2:n_i}\|_2\right\}, \qquad \sigma_{\mathcal{N}} := \min_{i\in\mathcal{N}} \max_{(\tilde{y};\tilde{s})\in\mathcal{D}^*_{\text{SOCO}}} \left\{\tilde{s}^i_1 - \|\tilde{s}^i_{2:n_i}\|_2\right\},\\
&\sigma_1 := \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\},\\
&\sigma_2 := \min_{i\in\mathcal{R}} \max_{(\tilde{x};\tilde{y};\tilde{s})\in\mathcal{P}^*_{\text{SOCO}}\times\mathcal{D}^*_{\text{SOCO}}} \left\{\tilde{x}^i_1 + \tilde{s}^i_1 - \|\tilde{x}^i_{2:n_i} + \tilde{s}^i_{2:n_i}\|_2\right\}.
\end{aligned} \tag{3.40}$$

We also define

$$\sigma_3 := \max_{(\tilde{x};\tilde{y};\tilde{s})\in\mathcal{P}^*_{\text{SOCO}}\times\mathcal{D}^*_{\text{SOCO}}} \left\{\|(\tilde{x};\tilde{y};\tilde{s})\|_2\right\} \tag{3.41}$$

as the radius of the outer ball which circumscribes the optimal set. Note that the condition numbers σ_1, σ_2, and σ_3 are finite positive values by the interior point condition, see Lemma 3.1.1 and Lemma 3.3 in [171]. The condition numbers σ_1 and σ_2 could be doubly exponentially small for some instances of SOCO, see Example 3.1.1.

Remark 3.2.1. *A positive lower bound can be computed for the condition numbers σ_1 and σ_2 using the method given in Section A.3. Under the uniqueness condition, an upper bound can be computed for the condition number σ_3, by representing the set of primal-dual optimal solutions of SOCO as a semi-algebraic set, see Lemma A.3.1.*

3.2.1 Identification along the central path

Theorem 3.2.1 presents the magnitude of x^μ and s^μ, as given in Theorem 3.4 in [171]. In [171], the authors employed the error bound in Theorem A.2.3 to derive bounds on the magnitude of $(x^\mu)^i$ and $(s^\mu)^i$ for $i \in \mathcal{T}$. Here, we rely on an error bound result for a linear conic system as stated in Theorem A.2.6. We use Theorem A.2.6 to specify an upper bound on the distance of a central solution from the optimal set.

Let $(\tilde{x}; \tilde{y}; \tilde{s})$ be a primal-dual optimal solution of $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$. In a similar manner as in Section 3.1.1.1, the primal and dual optimal sets can be equivalently written as the following linear conic systems

$$\begin{cases} x \in \tilde{x} + \operatorname{Ker}(A), \\ \tilde{s}^T x = 0, \\ x \in \mathcal{L}^{\bar{n}}_+, \end{cases} \qquad\qquad \begin{cases} s \in \tilde{s} + \mathcal{R}(A^T), \\ \tilde{x}^T s = 0, \\ s \in \mathcal{L}^{\bar{n}}_+. \end{cases} \tag{3.42}$$

In this case, the linear subspace $\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SOCO}}}$, which contains $\mathcal{P}^*_{\mathrm{SOCO}}$, is defined as

$$\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SOCO}}} := \big(\operatorname{Ker}(A) \cap (\mathbb{R}\tilde{s})^\perp\big) + \mathbb{R}\tilde{x},$$

Analogously, the linear subspace $\bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SOCO}}}$, which is the orthogonal complement of $\bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SOCO}}}$, is defined as

$$\bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SOCO}}} := \big(\mathcal{R}(A^T) \cap (\mathbb{R}\tilde{x})^\perp\big) + \mathbb{R}\tilde{s}.$$

From the orthogonality of $(x^\mu - \tilde{x})$ and $(s^\mu - \tilde{s})$ we have

$$\tilde{x}^T s^\mu + \tilde{s}^T x^\mu = p\mu,$$

which implies $0 \le \tilde{s}^T x^\mu \le p\mu$ and $0 \le \tilde{x}^T s^\mu \le p\mu$. Then the application of the Hoffman error bound gives

$$\begin{aligned}
\operatorname{dist}\left(x^\mu,\ \bar{\mathcal{S}}_{\mathcal{P}^*_{\mathrm{SOCO}}}\right) &\le \operatorname{dist}\left(x^\mu,\ \left\{x \in \tilde{x} + \operatorname{Ker}(A) \mid \tilde{s}^T x = 0\right\}\right)\\
&= \operatorname{dist}\left(x^\mu,\ \left\{x \mid Ax = b,\ \tilde{s}^T x = 0\right\}\right)\\
&\le \varpi_p\left(\|Ax^\mu - b\|_2 + \tilde{s}^T x^\mu\right) = \varpi_p \tilde{s}^T x^\mu \le \varpi_p p\mu,
\end{aligned}$$

where $\varpi_p > 0$ denotes the Hoffman condition number. Analogously, we derive

$$\operatorname{dist}\left(s^\mu,\ \bar{\mathcal{S}}_{\mathcal{D}^*_{\mathrm{SOCO}}}\right) \le \operatorname{dist}\left(s^\mu,\ \left\{s \in \tilde{s} + \mathcal{R}(A^T) \mid \tilde{x}^T s = 0\right\}\right) \le \varpi_d p\mu,$$

where $\varpi_d > 0$ is the Hoffman condition number. Note that the condition numbers ϖ_p and ϖ_d depend on A and $\tilde{s}$, and A and $\tilde{x}$, respectively.

Lemma 3.2.1. *Let $(x^\mu; y^\mu; s^\mu)$ be a central solution with*

$$\mu \le \hat{\mu} := \frac{1}{p}\min\left\{\varpi_p^{-1}, \varpi_d^{-1}\right\}. \tag{3.43}$$

Then there exist $(x_\mu; y_\mu; s_\mu) \in \mathcal{P}^*_{\mathrm{SOCO}} \times \mathcal{D}^*_{\mathrm{SOCO}}$, a positive condition number κ independent of μ, and $\gamma > 0$ so that

$$\begin{aligned}
\|x^\mu - x_\mu\|_2 &\le \kappa(p\mu)^\gamma,\\
\|y^\mu - y_\mu\|_2 &\le \kappa(p\mu)^\gamma,\\
\|s^\mu - s_\mu\|_2 &\le \kappa(p\mu)^\gamma.
\end{aligned} \tag{3.44}$$

Proof. Conditions (A.8) hold if $\mu \le \hat{\mu}$. Moreover, the rank and the interior point conditions imply that the set

$$\left\{(x^\mu; y^\mu; s^\mu) \mid 0 < \mu \le \hat{\mu}\right\}$$

is contained in a compact set. Therefore, the result of Theorem A.2.6 is applicable to the linear conic systems in (3.42). Furthermore, the compactness of $\mathcal{P}^*_{\text{SOCO}}$ and $\mathcal{D}^*_{\text{SOCO}}$ implies the existence of $(x_\mu; y_\mu; s_\mu) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}}$ so that

$$\|x^\mu - x_\mu\|_2 \leq \kappa_1(p\mu)^{\gamma_1}, \quad \|s^\mu - s_\mu\|_2 \leq \kappa_2(p\mu)^{\gamma_2},$$

where $\gamma_1 := 2^{-d(\bar{\mathcal{S}}_{\mathcal{P}^*_{\text{SOCO}}}, \mathcal{L}^{\bar{n}}_+)}$ and $\gamma_2 := 2^{-d(\bar{\mathcal{S}}_{\mathcal{D}^*_{\text{SOCO}}}, \mathcal{L}^{\bar{n}}_+)}$. Since the rows of A are assumed to be linearly independent, system $A^T(y^\mu - y_\mu) = s^\mu - s_\mu$ has a unique solution. Therefore, using Theorem A.2.6 again, we get

$$\|y^\mu - y_\mu\|_2 \leq \left\|(A^T)^\dagger\right\|_2 \|s^\mu - s_\mu\|_2 \leq \kappa_3(p\mu)^{\gamma_2},$$

where $(A^T)^\dagger := (AA^T)^{-1}A$ stands for the pseudo-inverse of A^T [161]. Then, taking $\kappa := \max\{\kappa_1, \kappa_2, \kappa_3\}$ and $\gamma := \min\{\gamma_1, \gamma_2\}$, we get the results as desired. □

Remark 3.2.2. *We can obtain a nontrivial lower bound $\gamma \geq 2^{-p}$ from (A.9). However, we are not aware of any upper bound on the condition number κ.*

Now, the bounds on the magnitude of the central solutions are summarized in Theorem 3.2.1.

Theorem 3.2.1. *Let $(x^\mu; y^\mu; s^\mu)$ be a central solution such that $\mu \leq \hat{\mu}$, where $\hat{\mu}$ is* *defined in (3.43). Then we have*

$$
\begin{aligned}
&(x^\mu)^i_1 \geq (x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 > \frac{\sigma_1}{2p}, \quad \textit{and} \quad (s^\mu)^i_1 \leq \frac{p\mu}{\sigma_1}, && i \in \mathcal{B},\\
&(s^\mu)^i_1 \geq (s^\mu)^i_1 - \|(s^\mu)^i_{2:n_i}\|_2 > \frac{\sigma_1}{2p}, \quad \textit{and} \quad (x^\mu)^i_1 \leq \frac{p\mu}{\sigma_1}, && i \in \mathcal{N},\\
&(x^\mu)^i_1 > \frac{\sigma_2}{4p}, \quad \textit{and} \quad (s^\mu)^i_1 > \frac{\sigma_2}{4p}, && i \in \mathcal{R},\\
&\big((x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2\big) + \big((s^\mu)^i_1 - \|(s^\mu)^i_{2:n_i}\|_2\big) \leq \frac{2p\mu}{\sigma_2}, && i \in \mathcal{R},\\
&(x^\mu)^i_1 + (s^\mu)^i_1 - \|(x^\mu)^i_{2:n_i} + (s^\mu)^i_{2:n_i}\|_2 > \frac{\sigma_1}{2p}, && i \in \mathcal{B} \cup \mathcal{N},\\
&(x^\mu)^i_1 + (s^\mu)^i_1 - \|(x^\mu)^i_{2:n_i} + (s^\mu)^i_{2:n_i}\|_2 > \frac{\sigma_2}{2p}, && i \in \mathcal{R},\\
&\frac{\mu}{2\kappa(p\mu)^\gamma} \leq (x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 \leq (x^\mu)^i_1 \leq \kappa(p\mu)^\gamma, && i \in \mathcal{T}_1,\\
&\frac{\mu}{2\kappa(p\mu)^\gamma} \leq (s^\mu)^i_1 - \|(s^\mu)^i_{2:n_i}\|_2 \leq (s^\mu)^i_1 \leq \kappa(p\mu)^\gamma, && i \in \mathcal{T}_1,\\
&\frac{\mu}{2\kappa(p\mu)^\gamma} \leq (x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 \leq \sqrt{2}\kappa(p\mu)^\gamma, && i \in \mathcal{T}_2,\\
&\frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma} \leq (s^\mu)^i_1 \leq \kappa(p\mu)^\gamma, && i \in \mathcal{T}_2,\\
&\frac{\mu}{2\kappa(p\mu)^\gamma} \leq (s^\mu)^i_1 - \|(s^\mu)^i_{2:n_i}\|_2 \leq \sqrt{2}\kappa(p\mu)^\gamma, && i \in \mathcal{T}_3,\\
&\frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma} \leq (x^\mu)^i_1 \leq \kappa(p\mu)^\gamma, && i \in \mathcal{T}_3,\\
&\frac{\mu}{2\kappa(p\mu)^\gamma} \leq (x^\mu)^i_1 + (s^\mu)^i_1 - \|(x^\mu)^i_{2:n_i} + (s^\mu)^i_{2:n_i}\|_2 \leq 4\kappa(p\mu)^\gamma, && i \in \mathcal{T},
\end{aligned}
$$

where κ and γ are defined in Lemma 3.2.1.

Proof. The sketch of the proof for the subsets $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{R}$ is similar to Theorem 3.4 in [171]. Let $i \in \mathcal{T}$ denote a block $((x^\mu)^i; (y^\mu)^i; (s^\mu)^i)$ of the central solution with $\mu \leq \hat{\mu}$. From (1.11) and the central path equation $(x^\mu)^i \circ (s^\mu)^i = \mu e_i$ we get

$$(s^\mu)^i = \frac{\mu\big((x^\mu)^i_1; -(x^\mu)^i_{2:n_i}\big)}{((x^\mu)^i_1)^2 - \|(x^\mu)^i_{2:n_i}\|_2^2}.$$

Therefore, from $(x^\mu)^i_1 > \|(x^\mu)^i_{2:n_i}\|_2$ we have

$$\begin{aligned}(s^\mu)^i_1 &= \frac{\mu(x^\mu)^i_1}{((x^\mu)^i_1)^2 - \|(x^\mu)^i_{2:n_i}\|_2^2} = \frac{\mu}{(x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2}\frac{(x^\mu)^i_1}{(x^\mu)^i_1 + \|(x^\mu)^i_{2:n_i}\|_2} \\ &\geq \frac{\mu}{2((x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2)}.\end{aligned} \tag{3.45}$$

Analogously, we can derive

$$(x^\mu)^i_1 \geq \frac{\mu}{2((s^\mu)^i_1 - \|(s^\mu)^i_{2:n_i}\|_2)}. \tag{3.46}$$

- $i \in \mathcal{T}_1$: In this case, we have $\tilde{x}^i = \tilde{s}^i = 0$ for all $(\tilde{x}; \tilde{y}; \tilde{s}) \in \mathcal{P}^*_{\text{SOCO}} \times \mathcal{D}^*_{\text{SOCO}}$, and thus the bounds in (3.44) reduce to

$$\|(x^\mu)^i\|_2 \leq \kappa(p\mu)^\gamma, \quad \|(s^\mu)^i\|_2 \leq \kappa(p\mu)^\gamma. \tag{3.47}$$

Consequently, it can be deducted from (3.45) and (3.47) that

$$\begin{aligned}(x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 &\geq \frac{\mu}{2(s^\mu)^i_1} \geq \frac{\mu}{2\|(s^\mu)^i\|_2} \geq \frac{\mu}{2\kappa(p\mu)^\gamma}, \\ (x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 &\leq (x^\mu)^i_1 \leq \|(x^\mu)^i\|_2 \leq \kappa(p\mu)^\gamma.\end{aligned}$$

In a similar manner, using (3.46) we can show that

$$\frac{\mu}{2\kappa(p\mu)^\gamma} \leq (s^\mu)^i_1 - \|(s^\mu)^i_{2:n_i}\|_2 \leq (s^\mu)^i_1 \leq \kappa(p\mu)^\gamma,$$

which completes the first part of the proof.

- $i \in \mathcal{T}_2$: In this case, the bound in (3.44) reduces to $\|(s^\mu)^i\|_2 \leq \kappa(p\mu)^\gamma$. Thus, we can conclude from (3.45) that

$$(x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 \geq \frac{\mu}{2(s^\mu)^i_1} \geq \frac{\mu}{2\|(s^\mu)^i\|_2} \geq \frac{\mu}{2\kappa(p\mu)^\gamma}.$$

Furthermore, it follows from $(x_\mu)^i_1 = \|(x_\mu)^i_{2:n_i}\|_2$ that

$$\begin{aligned}(x^\mu)^i_1 - \|(x^\mu)^i_{2:n_i}\|_2 &= ((x^\mu)^i_1 - (x_\mu)^i_1) + \left(\|(x_\mu)^i_{2:n_i}\|_2 - \|(x^\mu)^i_{2:n_i}\|_2\right) \\ &\leq |(x^\mu)^i_1 - (x_\mu)^i_1| + \|(x_\mu)^i_{2:n_i} - (x^\mu)^i_{2:n_i}\|_2 \\ &\leq \sqrt{2}\|(x^\mu)^i - x^i_\mu\|_2 \leq \sqrt{2}\kappa(p\mu)^\gamma.\end{aligned} \tag{3.48}$$

Therefore, using (3.45) and (3.48) we get

$$\kappa(p\mu)^\gamma \geq (s^\mu)_1^i \geq \frac{\mu}{2\big((x^\mu)_1^i - \|(x^\mu)_{2:n_i}^i\|_2\big)} \geq \frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma},$$

which completes the proof for the second part.

- $i \in \mathcal{T}_3$: It immediately follows after reversing the roles of $(x^\mu)^i$ and $(s^\mu)^i$.

The rest of the theorem follows by applying the results from the previous parts as in Theorem 3.8 in [171]. □

From the bounds of Theorem 3.2.1 one can observe that a complete separation of the variables to the partition $\mathcal{B}$, $\mathcal{N}$, $\mathcal{R}$ and $\mathcal{T}$ can be made if

$$\frac{p\mu}{\sigma_1} < \min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{4p}\right\}, \quad \max\left\{\frac{p\mu}{\sigma_1}, \frac{2p\mu}{\sigma_2}\right\} < \frac{\sigma_1}{2p}, \quad 4\kappa(p\mu)^\gamma < \min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{2p}\right\},$$

which can be simplified to

$$\mu < \tilde{\mu} := \min\left\{\frac{\sigma_1^2}{2p^2}, \frac{\sigma_1\sigma_2}{4p^2}, \ \frac{1}{p}\left(\frac{1}{4\kappa}\min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{2p}\right\}\right)^{\frac{1}{\gamma}}, \ \hat{\mu}\right\}. \tag{3.49}$$

However, we do not have enough information for a further separation of $\mathcal{T}$ into $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$. To that end, we need positive lower bounds on $(x^\mu)_1^i$ and $(s^\mu)_1^i$ in $\mathcal{T}_2$ and $\mathcal{T}_3$, respectively, which cannot be directly obtained from Theorem 3.2.1. Nevertheless, we assume in the convergence analysis of Section 4.2.3 that $\mu < \tilde{\mu}$ is small enough for a complete identification of $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$.

Remark 3.2.3. *Theorem 3.2.1 specifies the bounds on the magnitude of the central solutions. The results can be extended to the case where IPMs generate an approximate solution $(x^\circ; y^\circ; s^\circ)$ in a neighborhood of the central path, see Section 3.1.2 and Section 4 in [171] for a detailed discussion.*

Chapter 4

Identification of optimal solutions

The optimal partition information can be used in a so called rounding procedure to generate either a maximally or strictly complementary optimal solution. For LO and sufficient LCPs, the optimal partition and a maximally complementary optimal solution can be identified in strongly polynomial time. Ye [180] proposed a finite termination strategy for IPMs which generates a strictly complementary optimal solution from a primal-dual solution sufficiently close to the optimal set. Under the interior point condition as well as the integrality of the data, Roos et al. [151] presented a rounding procedure which uses the optimal partition information to identify a strictly complementary optimal solution. Under the same conditions, Illés et al. [85] considered the identification of the optimal partition for sufficient LCPs and proposed a strongly polynomial rounding procedure to a maximally complementary optimal solution.

In chapter 3, we investigated the identification of the optimal partition for SDO and SOCO. The identification of the optimal partition was obtained from a bounded sequence of solutions on, or in a neighborhood of the central path. In this chapter, we use the (approximation of the) optimal partition to either generate an approximate maximally complementary solution or speed up the convergence to the unique optimal solution.

In Section 4.1, we use the approximation of the optimal partition in a rounding procedure to generate an approximate maximally complementary solution for SDO. In Section 4.2, we employ the optimal partition of SOCO to identify the quadratic convergence region of Newton's method. Furthermore, we use the optimal partition to generate an approximate maximally complementary optimal solution for SOCO.

4.1 Identification of optimal solutions for SDO

In this section, we use the approximation of the optimal partition from Section 3.1 to generate a primal-dual solution which approximately satisfies primal-dual feasibility constraints and has zero duality gap, so called an approximate maximally complementary solution. This is in contrast to a rounding procedure for LO, e.g., [151], which gives an exact strictly complementary solution. Hence, our procedure can be considered as an extension of the rounding procedure in [151] except that we use the approximation of the optimal partition. To the best of our knowledge, there is no theoretical/computational procedure for the identification of the optimal partition and optimal solutions for SDO.

4.1.1 A rounding procedure for central solutions

Even though only approximations of $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$ are available, from a central solution with sufficiently small μ we can make a projection onto the boundary of the positive semidefinite cone to generate a complementary solution with approximate primal-dual feasibility. We choose a central solution (X^μ, y^μ, S^μ) with $\mu < \tilde{\mu}$, see (3.15), and compute the eigenvectors $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$. We then generate a primal-dual solution with approximate primal-dual feasibility and zero duality gap. We prove that if μ is sufficiently small, then the rounded primal-dual solution satisfies the cone constraints.

Suppose that a central solution (X^μ, y^μ, S^μ) is given, where μ satisfies (3.15). The columns of a common eigenvectors basis Q^μ can be rearranged so that

$$Q^\mu := \left(Q^\mu_{\mathcal{B}}, Q^\mu_{\mathcal{T}}, Q^\mu_{\mathcal{N}}\right),$$

e.g., $Q^\mu_{\mathcal{B}}$ denotes the columns of Q^μ whose accumulation point is an orthonormal basis for $\mathcal{B}$.

Remark 4.1.1. *In order to correctly identify $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$, the knowledge of the condition number σ is needed in our rounding procedure. See Lemma A.3.2 for a positive lower bound on σ.*

Let (X^*, y^*, S^*) be a maximally complementary optimal solution and define $\hat{X}^* := (Q^\mu)^T X^* Q^\mu$, and $\hat{S}^* := (Q^\mu)^T S^* Q^\mu$, i.e.,

$$\hat{X}^* := \begin{pmatrix} \hat{X}^*_{\mathcal{B}} & \hat{X}^*_{\mathcal{B}\mathcal{T}} & \hat{X}^*_{\mathcal{B}\mathcal{N}} \\ \hat{X}^*_{\mathcal{T}\mathcal{B}} & \hat{X}^*_{\mathcal{T}} & \hat{X}^*_{\mathcal{T}\mathcal{N}} \\ \hat{X}^*_{\mathcal{N}\mathcal{B}} & \hat{X}^*_{\mathcal{N}\mathcal{T}} & \hat{X}^*_{\mathcal{N}} \end{pmatrix}, \quad \hat{S}^* := \begin{pmatrix} \hat{S}^*_{\mathcal{B}} & \hat{S}^*_{\mathcal{B}\mathcal{T}} & \hat{S}^*_{\mathcal{B}\mathcal{N}} \\ \hat{S}^*_{\mathcal{T}\mathcal{B}} & \hat{S}^*_{\mathcal{T}} & \hat{S}^*_{\mathcal{T}\mathcal{N}} \\ \hat{S}^*_{\mathcal{N}\mathcal{B}} & \hat{S}^*_{\mathcal{N}\mathcal{T}} & \hat{S}^*_{\mathcal{N}} \end{pmatrix}.$$

Further, let $\hat{A}^i := (Q^\mu)^T A^i Q^\mu$, and $\Lambda(X^\mu) := (Q^\mu)^T X^\mu Q^\mu$ be a diagonal matrix, in which

$$\hat{A}^i := \begin{pmatrix} \hat{A}^i_{\mathcal{B}} & \hat{A}^i_{\mathcal{B}\mathcal{T}} & \hat{A}^i_{\mathcal{B}\mathcal{N}} \\ \hat{A}^i_{\mathcal{T}\mathcal{B}} & \hat{A}^i_{\mathcal{T}} & \bar{A}^i_{\mathcal{T}\mathcal{N}} \\ \hat{A}^i_{\mathcal{N}\mathcal{B}} & \hat{A}^i_{\mathcal{N}\mathcal{T}} & \hat{A}^i_{\mathcal{N}} \end{pmatrix}, \quad \Lambda(X^\mu) := \begin{pmatrix} \Lambda_{\mathcal{B}}(X^\mu) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(X^\mu) & 0 \\ 0 & 0 & \Lambda_{\mathcal{N}}(X^\mu) \end{pmatrix}.$$

Then from the primal feasibility constraints we have

$$\langle \hat{A}^i, \hat{X}^* \rangle = b_i, \qquad i = 1, \ldots, m, \tag{4.1}$$

$$\langle \hat{A}^i, \Lambda(X^\mu) \rangle = b_i, \qquad i = 1, \ldots, m. \tag{4.2}$$

By subtracting (4.2) from (4.1) for each i, for $\Delta X^\mu_{\mathcal{B}} = \hat{X}^*_{\mathcal{B}} - \Lambda_{\mathcal{B}}(X^\mu)$ we get

$$\langle \hat{A}^i_{\mathcal{B}}, \Delta X^\mu_{\mathcal{B}} \rangle = \langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X^\mu) \rangle + \langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X^\mu) \rangle + (\xi_p)_i, \tag{4.3}$$

where the residual term $(\xi_p)_i$ is

$$(\xi_p)_i = -\langle \hat{A}^i_{\mathcal{N}}, \hat{X}^*_{\mathcal{N}} \rangle - \langle \hat{A}^i_{\mathcal{T}}, \hat{X}^*_{\mathcal{T}} \rangle - 2\Big(\langle \hat{A}^i_{\mathcal{BT}}, \hat{X}^*_{\mathcal{BT}} \rangle + \langle \hat{A}^i_{\mathcal{BN}}, \hat{X}^*_{\mathcal{BN}} \rangle + \langle \hat{A}^i_{\mathcal{TN}}, \hat{X}^*_{\mathcal{TN}} \rangle \Big).$$

Analogously, let $\hat{C} := (Q^\mu)^T C Q^\mu$, and

$$\Lambda(S^\mu) := \begin{pmatrix} \Lambda_{\mathcal{B}}(S^\mu) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S^\mu) & 0 \\ 0 & 0 & \Lambda_{\mathcal{N}}(S^\mu) \end{pmatrix}.$$

Then for the dual constraints we get

$$\sum_{i=1}^m y^*_i \hat{A}^i + \hat{S}^* = \hat{C}, \tag{4.4}$$

$$\sum_{i=1}^m y^\mu_i \hat{A}^i + \Lambda(S^\mu) = \hat{C}. \tag{4.5}$$

By subtracting (4.5) from (4.4) for $\Delta y^\mu_i = y^*_i - y^\mu_i$ and $\Delta S^\mu_{\mathcal{N}} = \hat{S}^*_{\mathcal{N}} - \Lambda_{\mathcal{N}}(S^\mu)$ we get

$$\sum_{i=1}^m \Delta y^\mu_i \hat{A}^i + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S^\mu_{\mathcal{N}} \end{pmatrix} = \begin{pmatrix} \Lambda_{\mathcal{B}}(S^\mu) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S^\mu) & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} \hat{S}^*_{\mathcal{B}} & \hat{S}^*_{\mathcal{BT}} & \hat{S}^*_{\mathcal{BN}} \\ \hat{S}^*_{\mathcal{TB}} & \hat{S}^*_{\mathcal{T}} & \hat{S}^*_{\mathcal{TN}} \\ \hat{S}^*_{\mathcal{NB}} & \hat{S}^*_{\mathcal{NT}} & 0 \end{pmatrix}. \tag{4.6}$$

Both the right hand sides in (4.3) and (4.6) depend on the chosen maximally complementary optimal solution. Therefore, the system of equations in (4.3) and (4.6) may not be solvable if we drop the unknown terms. Instead, we can solve two least square problems to obtain search directions towards primal and dual solutions.

Remark 4.1.2. *For an LO problem embedded in SDO, X^μ and S^μ are diagonal matrices. When the optimal partition is known, the coordinates of variables can be rearranged to get $Q^\mu = I_n$. All this implies that the unknown terms in (4.3) and (4.6) are just zero for the special case of LO.*

4.1.1.1 Primal least square problem

For the primal problem we solve

$$\begin{array}{ll} \min & \|\Delta X\|_F^2 + \|\varepsilon_p\|_2^2 \\ \text{s.t.} & \langle \hat{A}^i_{\mathcal{B}}, \Delta X\rangle - (\varepsilon_p)_i = \langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X^\mu)\rangle + \langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X^\mu)\rangle, \quad i = 1, \ldots, m. \end{array} \tag{4.7}$$

We may assume that $\hat{A}^i_{\mathcal{B}} \neq 0$ for some i. Otherwise, the optimal solution of (4.7) would give $\Delta X^* = 0$, and thus the effect of the vanishing terms is absorbed in primal infeasibility. For LO we have $\hat{A}^i_{\mathcal{B}} \neq 0$ for some i, since otherwise the primal optimal solution would be trivial.

The optimal solution $(\Delta X^*, \varepsilon_p^*)$ to the auxiliary problem (4.7) yields

$$\tilde{\tilde{X}}_{\mathcal{B}} := \Lambda_{\mathcal{B}}(X^\mu) + \Delta X^*$$

so that

$$\langle \hat{A}^i_{\mathcal{B}}, \tilde{\tilde{X}}_{\mathcal{B}}\rangle = b_i + (\varepsilon_p^*)_i, \qquad i = 1, \ldots, m.$$

Thus, $\tilde{\tilde{X}}_{\mathcal{B}}$ has $\|\varepsilon_p^*\|_2$ infeasibility for the primal constraints.

Let $r(n) := n(n+1)/2$ and define

$$\hat{\mathcal{A}}^s_{\mathcal{B}} := \left(\operatorname{svec}(\hat{A}^1_{\mathcal{B}}), \ldots, \operatorname{svec}(\hat{A}^m_{\mathcal{B}})\right)^T.$$

Note that $\hat{\mathcal{A}}^s_{\mathcal{B}}$ might be rank deficient. Then auxiliary problem (4.7) reduces to

$$\min \quad \|\Delta x\|_2^2 + \|\hat{\mathcal{A}}^s_{\mathcal{B}}\Delta x - \eta\|_2^2, \tag{4.8}$$

where $\Delta x = \operatorname{svec}(\Delta X)$, and $\eta_i = \langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X^\mu)\rangle + \langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X^\mu)\rangle$ denotes the vanishing term for $i = 1, \ldots, m$, which should be zero for all optimal solutions. Lemma 4.1.1 establishes upper bounds on $\|\Delta X^*\|_F$ and $\|\varepsilon_p^*\|_2$. The bounds depend on the constant

$$\pi_p := \prod_{k=1}^{r(n_{\mathcal{B}})} \left\| \left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})} \right)_{.k} \right\|_2,$$

where $\big((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)_{.k}$ denotes the k^{th} column of $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$. Using the upper bounds in Lemma 4.1.1 , we show in Theorem 4.1.1 that $\tilde{\tilde{X}}_{\mathcal{B}} \succ 0$ for sufficiently small μ.

Lemma 4.1.1. *Let $(\Delta X^*, \varepsilon_p^*)$ be the unique optimal solution of (4.7). Then we have*

$$\|\Delta X^*\|_F \leq 2\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\},$$
$$\|\varepsilon_p^*\|_2 \leq 2\|\mathcal{A}^s\|_F\left(\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^s\|_F^2 + 1\right)\max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}.$$

Proof. The optimality conditions for (4.8) are given by

$$\big((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)\Delta x = (\hat{\mathcal{A}}^s_{\mathcal{B}})^T\eta, \tag{4.9}$$

where $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})} \succ 0$. The unique solution Δx^* can be computed using Cramer's rule [84]:

$$\Delta x_j^* = \frac{\det\left(\big((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)^{(j)}\right)}{\det\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)}, \qquad j = 1, \ldots, r(n_{\mathcal{B}}),$$

in which the matrix $((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)^{(j)}$ in the numerator is obtained by substituting the j^{th} column of $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$ by $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\eta$. Noting that[1] $\det((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}) \geq 1$, we can deduce from Hadamard's inequality [84] that for $j = 1, \ldots, r(n_{\mathcal{B}})$

$$|\Delta x_j^*| \leq \left|\det\left(\big((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)^{(j)}\right)\right| \leq \big\|(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\eta\big\|_2 \prod_{\substack{k=1\\k\neq j}}^{r(n_{\mathcal{B}})} \big\|((\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})})_{.k}\big\|_2$$

hold. Since the diagonal entries of $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}}+I_{r(n_{\mathcal{B}})}$ are greater than or equal to 1, the norm of each column is at least 1, and thus a uniform bound for all $j = 1, \ldots, r(n_{\mathcal{B}})$

[1] This is true regardless of data type, since the eigenvalues of $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T\hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$ are at least 1.

can be derived as

$$|\Delta x_j^*| \leq \left\|(\hat{\mathcal{A}}_{\mathcal{B}}^s)^T\eta\right\|_2 \prod_{\substack{k=1\\k\neq j}}^{r(n_{\mathcal{B}})} \left\|((\hat{\mathcal{A}}_{\mathcal{B}}^s)^T\hat{\mathcal{A}}_{\mathcal{B}}^s + I_{r(n_{\mathcal{B}})})_{.k}\right\|_2 \leq \pi_p\|(\hat{\mathcal{A}}_{\mathcal{B}}^s)^T\eta\|_2. \tag{4.10}$$

Noting that $\|\hat{A}_{\mathcal{N}}^i\|_F \leq \|\hat{A}^i\|_F = \|A^i\|_F$ and $\|\hat{A}_{\mathcal{T}}^i\|_F \leq \|\hat{A}^i\|_F = \|A^i\|_F$, we can conclude from (3.13) and (3.14) that

$$\begin{aligned}
|\langle \hat{A}_{\mathcal{N}}^i, \Lambda_{\mathcal{N}}(X^\mu)\rangle| &\leq \frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}\|A^i\|_F, && i = 1,\ldots,m,\\
|\langle \hat{A}_{\mathcal{T}}^i, \Lambda_{\mathcal{T}}(X^\mu)\rangle| &\leq \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\|A^i\|_F, && i = 1,\ldots,m,
\end{aligned}$$

which yields the upper bound

$$|\eta_i| \leq 2\|A^i\|_F \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}, \qquad i = 1,\ldots,m. \tag{4.11}$$

Consequently, from (4.10) and (4.11) it follows that

$$|\Delta x_j^*| \leq \pi_p\|(\hat{\mathcal{A}}_{\mathcal{B}}^s)^T\eta\|_2 \leq 2\pi_p\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}, \quad j = 1,\ldots,r(n_{\mathcal{B}}),$$

where we have used $\|\mathcal{A}^s\|_F^2 = \sum_{i=1}^m \|A^i\|_F^2$, and the inequality $\|\hat{\mathcal{A}}_{\mathcal{B}}^s\|_F \leq \|\mathcal{A}^s\|_F$. As a result, we get

$$\begin{aligned}
\|\varepsilon_p^*\|_2 &= \|\hat{\mathcal{A}}_{\mathcal{B}}^s\Delta x^* - \eta\|_2 \leq \|\hat{\mathcal{A}}_{\mathcal{B}}^s\|_F\|\Delta x^*\|_2 + \|\eta\|_2\\
&\leq 2\|\mathcal{A}^s\|_F\left(\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^s\|_F^2 + 1\right)\max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}.
\end{aligned}$$

This completes the proof. □

4.1.1.2 Dual least square problem

Let E denote a slack matrix as

$$E := \begin{pmatrix} E_{\mathcal{B}} & E_{\mathcal{B}\mathcal{T}} & E_{\mathcal{B}\mathcal{N}}\\ E_{\mathcal{T}\mathcal{B}} & E_{\mathcal{T}} & E_{\mathcal{T}\mathcal{N}}\\ E_{\mathcal{N}\mathcal{B}} & E_{\mathcal{N}\mathcal{T}} & 0 \end{pmatrix}, \tag{4.12}$$

which is defined in accordance with the unknown right hand side matrix in (4.6). Then the auxiliary problem for an approximate dual solution is formulated as

$$\begin{aligned} \min \quad & \|\Delta y\|_2^2 + \|\Delta S\|_F^2 + \|E\|_F^2 \\ \text{s.t.} \quad & \sum_{i=1}^{m} \Delta y_i \hat{A}^i + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S \end{pmatrix} - E = \begin{pmatrix} \Lambda_{\mathcal{B}}(S^\mu) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S^\mu) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{4.13}$$

The optimal solution $(\Delta y^*, \Delta S^*, E^*)$ gives $\tilde{\tilde{y}}_i := y_i^\mu + \Delta y_i^*$ for $i = 1, \ldots, m$ and $\tilde{\tilde{S}}_{\mathcal{N}} := \Lambda_{\mathcal{N}}(S^\mu) + \Delta S^*$ with ε_d^* infeasibility for the dual constraints, where

$$\varepsilon_d^* := \|E^*\|_F. \tag{4.14}$$

For the sake of clarity, in what follows, auxiliary problem (4.13) is represented in vector form. To do so, analogous to the definition of $\hat{\mathcal{A}}_{\mathcal{B}}^s$, we apply the mapping svec(.) to each diagonal block of $\hat{A}_i$ to form $\hat{\mathcal{A}}_{\mathcal{N}}^s$ and $\hat{\mathcal{A}}_{\mathcal{T}}^s$. Further we apply the mapping vec(.) to form $\hat{\mathcal{A}}_{\mathcal{BT}}^v$, $\hat{\mathcal{A}}_{\mathcal{BN}}^v$, and $\hat{\mathcal{A}}_{\mathcal{TN}}^v$. Thus, auxiliary problem (4.13) reduces to the least square problem

$$\begin{aligned} \min \ & \|\Delta y\|_2^2 + \|(\hat{\mathcal{A}}_{\mathcal{B}}^s)^T \Delta y - \zeta_{\mathcal{B}}\|_2^2 + \|(\hat{\mathcal{A}}_{\mathcal{T}}^s)^T \Delta y - \zeta_{\mathcal{T}}\|_2^2 + \|(\hat{\mathcal{A}}_{\mathcal{N}}^s)^T \Delta y\|_2^2 \\ & + 2\|(\hat{\mathcal{A}}_{\mathcal{BT}}^v)^T \Delta y\|_2^2 + 2\|(\hat{\mathcal{A}}_{\mathcal{BN}}^v)^T \Delta y\|_2^2 + 2\|(\hat{\mathcal{A}}_{\mathcal{TN}}^v)^T \Delta y\|_2^2, \end{aligned} \tag{4.15}$$

where $\zeta_{\mathcal{B}} = \mathrm{svec}(\Lambda_{\mathcal{B}}(S^\mu))$ and $\zeta_{\mathcal{T}} = \mathrm{svec}(\Lambda_{\mathcal{T}}(S^\mu))$. Lemma 4.1.2 establishes upper bounds on ε_d^* and $\|\Delta S^*\|_F$. For the upper bounds we define the positive definite matrix

$$\begin{aligned} \mathcal{H} := {} & \hat{\mathcal{A}}_{\mathcal{B}}^s(\hat{\mathcal{A}}_{\mathcal{B}}^s)^T + \hat{\mathcal{A}}_{\mathcal{T}}^s(\hat{\mathcal{A}}_{\mathcal{T}}^s)^T + \hat{\mathcal{A}}_{\mathcal{N}}^s(\hat{\mathcal{A}}_{\mathcal{N}}^s)^T + 2\hat{\mathcal{A}}_{\mathcal{BT}}^v(\hat{\mathcal{A}}_{\mathcal{BT}}^v)^T + 2\hat{\mathcal{A}}_{\mathcal{BN}}^v(\hat{\mathcal{A}}_{\mathcal{BN}}^v)^T \\ & + 2\hat{\mathcal{A}}_{\mathcal{TN}}^v(\hat{\mathcal{A}}_{\mathcal{TN}}^v)^T + I_m \end{aligned}$$

and constant

$$\pi_d := \prod_{k=1}^{m} \|\mathcal{H}_{.k}\|_2,$$

where $\mathcal{H}_{.k}$ denotes the k^{th} column of $\mathcal{H}$. Theorem 4.1.1 proves that for sufficiently small μ we have $\tilde{\tilde{S}}_{\mathcal{N}} \succ 0$.

Lemma 4.1.2. *Problem (4.13) has a unique optimal solution $(\Delta y^*, \Delta S^*, E^*)$, which satisfies*

$$\|\Delta S^*\|_F \leq 2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\},$$

$$\varepsilon_d^* = \|E^*\|_F \leq \sqrt{2}(4\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 + 1)\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}.$$

Proof. The optimality conditions for (4.15) can be written as

$$\mathcal{H}\Delta y = \hat{\mathcal{A}}^s_{\mathcal{B}}\zeta_{\mathcal{B}} + \hat{\mathcal{A}}^s_{\mathcal{T}}\zeta_{\mathcal{T}},$$

where $\mathcal{H} \succ 0$. The unique solution of this system can be computed by using Cramer's rule as follows

$$\Delta y_i^* = \frac{\det(\mathcal{H}^{(i)})}{\det(\mathcal{H})}, \qquad i = 1, \ldots, m,$$

where matrix $\mathcal{H}^{(i)}$ in the numerator is obtained by substituting the i^{th} column of $\mathcal{H}$ by $\hat{\mathcal{A}}^s_{\mathcal{B}}\zeta_{\mathcal{B}} + \hat{\mathcal{A}}^s_{\mathcal{T}}\zeta_{\mathcal{T}}$. Note that $\lambda_{\min}(\mathcal{H}) \geq 1$, which implies $\det(\mathcal{H}) \geq 1$. Therefore, for $i = 1, \ldots, m$ we get

$$|\Delta y_i^*| \leq |\det(\mathcal{H}^{(i)})| \leq \|\hat{\mathcal{A}}^s_{\mathcal{B}}\zeta_{\mathcal{B}} + \hat{\mathcal{A}}^s_{\mathcal{T}}\zeta_{\mathcal{T}}\|_2 \prod_{\substack{k=1\\k\neq i}}^{m} \|\mathcal{H}_{.k}\|_2 \leq \pi_d\|\hat{\mathcal{A}}^s_{\mathcal{B}}\zeta_{\mathcal{B}} + \hat{\mathcal{A}}^s_{\mathcal{T}}\zeta_{\mathcal{T}}\|_2,$$

where the second inequality follows from Hadamard's inequality. Note that

$$\prod_{k=1,k\neq i}^{m} \|\mathcal{H}_{.k}\|_2 \leq \pi_d,$$

since the diagonal entries of $\mathcal{H}$ are at least 1. Furthermore, from Theorem 3.1.1 we get

$$\|\hat{\mathcal{A}}^s_{\mathcal{B}}\zeta_{\mathcal{B}} + \hat{\mathcal{A}}^s_{\mathcal{T}}\zeta_{\mathcal{T}}\|_2 \leq \|\hat{\mathcal{A}}^s_{\mathcal{B}}\zeta_{\mathcal{B}}\|_2 + \|\hat{\mathcal{A}}^s_{\mathcal{T}}\zeta_{\mathcal{T}}\|_2 \leq 2\|\mathcal{A}^s\|_F \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\},$$

which leads to

$$|\Delta y_i^*| \leq 2\pi_d \|\mathcal{A}^s\|_F \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}, \quad i = 1, \ldots, m. \tag{4.16}$$

Consequently, from (4.16) we have

$$\|\Delta S^*\|_F = \|(\hat{\mathcal{A}}_\mathcal{N}^s)^T \Delta y^*\|_2 \leq 2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma}, \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}.$$

Note that $\|QEQ^T\|_F = \|E\|_F$. Then we derive bounds on the components of E^* as follows

$$\begin{aligned}
\|E_\mathcal{B}^*\|_F &= \|(\hat{\mathcal{A}}_\mathcal{B}^s)^T \Delta y^* - \zeta_\mathcal{B}\|_2 \leq 2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\} \\
&\quad + \frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma} \leq (2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 + 1) \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}, \\
\|E_\mathcal{T}^*\|_F &= \|(\hat{\mathcal{A}}_\mathcal{T}^s)^T \Delta y^* - \zeta_\mathcal{T}\|_2 \leq 2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\} \\
&\quad + \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma \leq (2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 + 1) \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}, \\
\|E_{\mathcal{B}\mathcal{T}}^*\|_F, \|E_{\mathcal{B}\mathcal{N}}^*\|_F, \|E_{\mathcal{T}\mathcal{N}}^*\|_F &\leq 2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}.
\end{aligned}$$

Then we get

$$\begin{aligned}
\|E^*\|_F^2 &\leq \left(2\Big(2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 + 1\Big)^2 + 6\Big(2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2\Big)^2\right) \\
&\qquad \times \left(\max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}\right)^2 \\
&\leq 2\Big(4\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 + 1\Big)^2 \left(\max\left\{\frac{n\sqrt{n_\mathcal{B}}\mu}{\sigma},\ \kappa\sqrt{nn_\mathcal{T}}(n\mu)^\gamma\right\}\right)^2,
\end{aligned}$$

which gives the upper bound on ε_d^*. □

4.1.1.3 Cone feasibility

As specified by Lemmas 4.1.1 and 4.1.2, $\big(Q_\mathcal{B}^\mu \tilde{\tilde{X}}_\mathcal{B}(Q_\mathcal{B}^\mu)^T, \tilde{\tilde{y}}, Q_\mathcal{N}^\mu \tilde{\tilde{S}}_\mathcal{N}(Q_\mathcal{N}^\mu)^T\big)$ yields a complementary solution for ($\mathrm{P_{SDO}}$) and ($\mathrm{D_{SDO}}$). This primal-dual pair has $\varepsilon^* :=$

$\max\{\varepsilon_p^*, \varepsilon_d^*\}$ infeasibility with respect to the linear constraints. Theorem 4.1.1 shows that for a sufficiently small μ, the rounding procedure yields a primal-dual solution with $\tilde{\tilde{X}}_{\mathcal{B}}, \tilde{\tilde{S}}_{\mathcal{N}} \succ 0$.

Theorem 4.1.1. *Let* $\vartheta_1 := 2n^2\|\mathcal{A}^s\|_F^2$, $\vartheta_2 := 2\kappa n^{\frac{3}{2}}\sqrt{n_{\mathcal{T}}}\|\mathcal{A}^s\|_F^2$, *and*

$$\mu^r := \min\left\{\frac{\sigma^2}{\vartheta_1 \max\{\pi_p\sqrt{r(n_{\mathcal{B}})n_{\mathcal{N}}}, \pi_d\sqrt{mn_{\mathcal{B}}}\}}, \frac{1}{n}\left(\frac{\sigma}{\vartheta_2 \max\{\pi_p\sqrt{r(n_{\mathcal{B}})},\ \pi_d\sqrt{m}\}}\right)^{\frac{1}{\gamma}}, \tilde{\mu}\right\}.$$

If $\mu < \mu^r$, *then we have* $\tilde{\tilde{X}}_{\mathcal{B}}, \tilde{\tilde{S}}_{\mathcal{N}} \succ 0$.

Proof. We only need to show that for $\mu < \mu^r$ the rounding procedure results in $\tilde{\tilde{X}}_{\mathcal{B}}, \tilde{\tilde{S}}_{\mathcal{N}} \succ 0$. Noting that

$$|\lambda_{\min}(\Delta X^*)| \leq \|\Delta X^*\|_F, \qquad |\lambda_{\min}(\Delta S^*)| \leq \|\Delta S^*\|_F,$$

we can conclude from (3.12) and (3.13), and Lemmas 4.1.1 and 4.1.2 that

$$\begin{aligned}
\lambda_{\min}(\tilde{\tilde{X}}_{\mathcal{B}}) &\geq \lambda_{\min}(\Lambda_{\mathcal{B}}(X^\mu)) + \lambda_{\min}(\Delta X^*) \\
&\geq \frac{\sigma}{n} - 2\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}, \\
\lambda_{\min}(\tilde{\tilde{S}}_{\mathcal{N}}) &\geq \lambda_{\min}(\Lambda_{\mathcal{N}}(S^\mu)) + \lambda_{\min}(\Delta S^*) \\
&\geq \frac{\sigma}{n} - 2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}.
\end{aligned}$$

Consequently, $\tilde{\tilde{X}}_{\mathcal{B}}, \tilde{\tilde{S}}_{\mathcal{N}} \succ 0$ holds if

$$\begin{aligned}
2\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\} &< \frac{\sigma}{n}, \\
2\pi_d\sqrt{m}\|\mathcal{A}^s\|_F^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, \kappa\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\} &< \frac{\sigma}{n}.
\end{aligned}$$

These inequalities hold if $\mu < \mu^r$. The proof is complete. □

4.1.1.4 Outline of the procedure

Now, we can outline a simple procedure which yields an approximate maximally complementary solution.

Algorithm 2 Rounding procedure for SDO

Input (X^μ, y^μ, S^μ), where $\mu < \mu^r$ is fixed.

$(Q^\mu_\mathcal{B}, Q^\mu_\mathcal{T}, Q^\mu_\mathcal{N})$.

Do Solve least square problem (4.7) to get $\tilde{\tilde{X}}_\mathcal{B}$.

Solve least square problem (4.13) to get $(\tilde{\tilde{y}}, \tilde{\tilde{S}}_\mathcal{N})$.

Return $(Q^\mu_\mathcal{B}\tilde{\tilde{X}}_\mathcal{B}(Q^\mu_\mathcal{B})^T, \tilde{\tilde{y}}, Q^\mu_\mathcal{N}\tilde{\tilde{S}}_\mathcal{N}(Q^\mu_\mathcal{N}))^T$.

Even though Algorithm 2 produces an approximate maximally complementary solution and needs the eigenvalue decomposition of a central solution, it relies on solving two linear systems of equations with better conditioned coefficient matrices than the Jacobian of the Newton system.

Remark 4.1.3. *For a fixed μ the orthogonal matrix Q^μ, due to multiplicity of the eigenvalues, may not be unique. Then the solution given by Algorithm 2 varies with the choice of Q^μ. Obviously, this cannot happen for LO.*

Remark 4.1.4. *As indicated at the beginning of Section 4.1, Algorithm 2 can be inferred as an extension of the method in [151]. Computing an ε^*-feasible maximally complementary solution requires $\mathcal{O}(\max\{n^6_\mathcal{B}, m^3\})$ arithmetic operations. In fact, this is equivalent to solving two linear systems of equations, using the Gauss elimination method, with $r(n_\mathcal{B})$ and m variables, respectively.*

4.1.2 A rounding procedure for approximate solutions

We recall from Section 3.1.2 that the identification results can be extended for approximate solutions, which is best suited for primal-dual IPMs, since they generate a sequence of interior solutions in a neighborhood of the central path. This motivates the extension of the rounding procedure for solutions in a neighborhood of the central path.

Let $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_{\text{prox}}(\xi)$, and consider the eigenvalue decompositions

$$X^\circ = M\Lambda(X^\circ)M^T, \qquad S^\circ = N\Lambda(S^\circ)N^T,$$

where M and N are orthogonal matrices. Further, let $M := (M_{\mathcal{B}}, M_{\mathcal{T}}, M_{\mathcal{N}})$ and $N := (N_{\mathcal{B}}, N_{\mathcal{T}}, N_{\mathcal{N}})$, where the subsets of columns of M and N correspond to the optimal partition. When $\langle X^\circ, S^\circ\rangle$ is sufficiently small, we can identify $M_{\mathcal{B}}$, $M_{\mathcal{T}}$, and $M_{\mathcal{N}}$ from X°, and $N_{\mathcal{B}}$, $N_{\mathcal{T}}$, and $N_{\mathcal{N}}$ from S°. To extend the rounding procedure to solutions in the neighborhood of the central path, we need to choose the eigenvectors either from M or N, because X° and S° do not commute, see Section 3.1.2. To do so, we can solve the primal least square problem (4.7), where X^μ and Q^μ are replaced by X° and M, respectively, in the definition of $\hat{A}^i$, $\hat{C}$, and the right hand side in (4.7). We then solve the following least square problem to compute a dual solution:

$$\begin{aligned}
\min \quad & \|\Delta y\|_2^2 + \|\Delta S\|_F^2 + \|E\|_F^2 \\
\text{s.t.} \quad & \sum_{i=1}^m \Delta y_i \hat{A}^i + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S \end{pmatrix} - E = \begin{pmatrix} M_{\mathcal{B}}^T S^\circ M_{\mathcal{B}} & M_{\mathcal{B}}^T S^\circ M_{\mathcal{T}} & M_{\mathcal{B}}^T S^\circ M_{\mathcal{N}} \\ M_{\mathcal{T}}^T S^\circ M_{\mathcal{B}} & M_{\mathcal{T}}^T S^\circ M_{\mathcal{T}} & M_{\mathcal{T}}^T S^\circ M_{\mathcal{N}} \\ M_{\mathcal{N}}^T S^\circ M_{\mathcal{B}} & M_{\mathcal{N}}^T S^\circ M_{\mathcal{T}} & 0 \end{pmatrix},
\end{aligned} \tag{4.17}$$

where E is defined as in (4.12). Let $\left(M_{\mathcal{B}}\tilde{\tilde{X}}_{\mathcal{B}}M_{\mathcal{B}}^T, \tilde{\tilde{y}}, M_{\mathcal{N}}\tilde{\tilde{S}}_{\mathcal{N}}M_{\mathcal{N}}^T\right)$ be the updated primal-dual solution after applying the search directions from (4.7) and (4.17), where

$$\begin{aligned}
\tilde{\tilde{X}}_{\mathcal{B}} &= \Lambda_{\mathcal{B}}(X^\circ) + \Delta X^*, \\
\tilde{\tilde{S}}_{\mathcal{N}} &= M_{\mathcal{N}}^T S^\circ M_{\mathcal{N}} + \Delta S^*, \\
\tilde{\tilde{y}}_i &= y_i^\circ + \Delta y_i^*, \qquad\qquad i = 1, \ldots, m.
\end{aligned}$$

We can show, in a similar manner as in Section 4.1.1, that $\left(M_{\mathcal{B}}\tilde{\tilde{X}}_{\mathcal{B}}M_{\mathcal{B}}^T, \tilde{\tilde{y}}, M_{\mathcal{N}}\tilde{\tilde{S}}_{\mathcal{N}}M_{\mathcal{N}}^T\right)$ becomes an approximate complementary solution if the complementarity gap $\langle X^\circ, S^\circ\rangle$ is sufficiently small.

Alternatively, we may fix the basis at N and solve (4.13) to compute a dual solution, where (y^μ, S^μ) and Q^μ are replaced by (y°, S°) and N, respectively, in the definition of $\hat{A}^i$ and the right hand side in (4.13). Afterwards, we solve

$$\begin{aligned} \min \quad & \|\Delta X\|_F^2 + \|\varepsilon_p\|_2^2 \\ \text{s.t.} \quad & \langle \hat{A}^i_{\mathcal{B}}, \Delta X\rangle - (\varepsilon_p)_i = \bar{b}_i, \quad i = 1, \ldots, m, \end{aligned} \tag{4.18}$$

where

$$\begin{aligned} \bar{b}_i = {} & \langle \hat{A}^i_{\mathcal{T}}, N_{\mathcal{T}}^T X^\circ N_{\mathcal{T}}\rangle + \langle \hat{A}^i_{\mathcal{N}}, N_{\mathcal{N}}^T X^\circ N_{\mathcal{N}}\rangle + 2\langle \hat{A}^i_{\mathcal{B}\mathcal{T}}, N_{\mathcal{B}}^T X^\circ N_{\mathcal{T}}\rangle + 2\langle \hat{A}^i_{\mathcal{B}\mathcal{N}}, N_{\mathcal{B}}^T X^\circ N_{\mathcal{N}}\rangle \\ & + 2\langle \hat{A}^i_{\mathcal{T}\mathcal{N}}, N_{\mathcal{T}}^T X^\circ N_{\mathcal{N}}\rangle. \end{aligned}$$

Let $\left(N_{\mathcal{B}}\tilde{\tilde{X}}_{\mathcal{B}}N_{\mathcal{B}}^T, \tilde{\tilde{y}}, N_{\mathcal{N}}\tilde{\tilde{S}}_{\mathcal{N}}N_{\mathcal{N}}^T\right)$ be the new primal-dual solution after applying the search directions from (4.13) and (4.18), where

$$\begin{aligned} \tilde{\tilde{X}}_{\mathcal{B}} &= N_{\mathcal{B}}^T X^\circ N_{\mathcal{B}} + \Delta X^*, \\ \tilde{\tilde{S}}_{\mathcal{N}} &= \Lambda_{\mathcal{N}}(S^\circ) + \Delta S^*, \\ \tilde{\tilde{y}}_i &= y_i^\circ + \Delta y_i^*, \qquad i = 1, \ldots, m. \end{aligned}$$

Then for sufficiently small complementarity gap $\langle X^\circ, S^\circ\rangle$, we can show that

$$\left(N_{\mathcal{B}}\tilde{\tilde{X}}_{\mathcal{B}}N_{\mathcal{B}}^T, \tilde{\tilde{y}}, N_{\mathcal{N}}\tilde{\tilde{S}}_{\mathcal{N}}N_{\mathcal{N}}^T\right)$$

is an approximate maximally complementary solution. The approach to derive the feasibility bounds is analogous to Section 4.1.1.

4.2 Identification of optimal solutions for SOCO

Quadratic convergence of a primal-dual IPM for SOCO follows from Theorem 28 in [5] under nondegeneracy and strict complementarity conditions. Under the same conditions, the application of Newton's method to the optimality conditions of SDO enjoys a quadratic convergence, when the initial point is sufficiently close to the optimal set, see Corollary 3.2 in [7]. If strict complementarity condition fails, then this local convergence result is no longer maintained. See [106, 111, 135, 164, 165, 182, 183] for superlinear and quadratic convergence of IPMs for LO, LCP, and SDO.

To the best of our knowledge, only very few remedies are available to resolve the issue of strict complementarity for SDO and SOCO. Those mostly are based on nonsmooth analysis of the optimality conditions, see e.g., [24, 90, 91, 96, 97]. In such cases, the complementarity condition is replaced by $x^i - \Pi_{\mathbb{L}_+^{n_i}}(x^i - s^i) = 0$ for $i = 1, \dots, p$, where $\Pi_{\mathbb{L}_+^{n_i}}(.)$ denotes the Euclidean projection on $\mathbb{L}_+^{n_i}$. Due to non-differentiability of $\Pi_{\mathbb{L}_+^{n_i}}(.)$ at some points, smoothing functions have been proposed to reformulate the complementarity condition. Under primal and dual nondegeneracy conditions, Chan and Sun [24] and Kong [96] established the quadratic convergence of a smoothing Newton's method to the unique optimal solution of SDO and symmetric conic optimization, respectively. Extending a smoothing function from nonlinear complementarity problems, Chi and Liu [28] proposed a non-interior continuation method for SOCO with superlinear convergence rate in the absence of strict complementarity. See also [29] for another derivation of smoothing function and smoothing Newton's method for SOCO with quadratic convergence rate.

In this section, our goal is to establish quadratic convergence to the unique optimal solution of SOCO using the optimal partition of the problem. Analogous to e.g., [24, 96], we only assume primal and dual nondegeneracy conditions. However, in contrast to [24, 96], we do not consider the metric projection operator to reformulate the complementarity condition. In our case, quadratic convergence is established only when the solution is sufficiently centered and sufficiently close to the optimal set, which allows for the identification of the optimal partition. The optimal partition can be identified using a sequence of interior solutions which has accumulation points in the relative interior of the optimal set, as presented in Section 3.2. Given the optimal partition identified from a sequence of central solutions, we reformulate the SOCO problem as a reduced NLO problem and then apply Newton's method to the first-order optimality conditions of the reduced problem. We show that if the primal and dual nondegeneracy conditions hold, then the Jacobian of the equations in KKT system of the reduced NLO problem is nonsingular at the unique globally optimal solution of the NLO problem. As a result, starting from a solution sufficiently close to the optimal set, Newton's method converges quadratically to the unique optimal solution. We support the theory by some numerical results.

4.2.1 Nondegeneracy conditions for SOCO

Since we assume both the primal and dual nondegeneracy almost everywhere in this paper, using the optimal partition we characterize the primal and dual nondegeneracy conditions for the unique optimal solutions of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). In Section 1.3.2, primal-dual nondegeneracy is defined for any primal-dual feasible solution.

To that end, the matrix of the eigenvectors of $L(x^i)$ is denoted by $P_i := (\sqrt{2}p_1^i, \sqrt{2}p_2^i,\ \hat{P}_i)$, where

$$p_1^i := \frac{1}{2}\left(1; \frac{-x_{2:n_i}^i}{\|x_{2:n_i}^i\|_2}\right), \qquad p_2^i := \frac{1}{2}\left(1; \frac{x_{2:n_i}^i}{\|x_{2:n_i}^i\|_2}\right),$$

and $\hat{P}_i \in \mathbb{R}^{(n_i \times n_i - 2)}$ is a matrix with orthogonal columns. The eigenvectors of $L((x^*)^i)$ are indicated by superscript *.

Theorem 4.2.1 (Theorems 20 and 21 in [5], Proposition 19 in [21])**.** *Let* $(x^*; y^*; s^*)$ *be the unique optimal solution of* $(\mathrm{P_{SOCO}})$ *and* $(\mathrm{D_{SOCO}})$*. Then* x^* *is primal nondegenerate if and only if the matrix*

$$\big((A_i \bar{P}_i^*)_{i\in\mathcal{R}\cup\mathcal{T}_2},\ A_{\mathcal{B}}\big) \tag{4.19}$$

has full row rank, where $\bar{P}_i^* := (\sqrt{2}(p^*)_2^i,\ \hat{P}_i^*)$*. Furthermore,* $(y^*; s^*)$ *is dual nondegenerate if and only if the matrix*

$$\big((A_i R_i (s^*)^i)_{i\in\mathcal{R}\cup\mathcal{T}_3},\ A_{\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2}\big) \tag{4.20}$$

has full column rank, where R_i *is defined in* (1.9)*. If* $(x^*; y^*; s^*)$ *is both primal and dual nondegenerate, then we have*

$$\sum_{i\in\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2} n_i + |\mathcal{R}\cup\mathcal{T}_3| \le m \le \sum_{i\in\mathcal{B}\cup\mathcal{R}\cup\mathcal{T}_2} n_i - |\mathcal{R}\cup\mathcal{T}_2|. \tag{4.21}$$

For the sake of convenience, given the unique optimal solution $(x^*; y^*; s^*)$, the primal nondegeneracy of x^* and the dual nondegeneracy of $(y^*; s^*)$ are simply called the primal and dual nondegeneracy conditions, respectively.

4.2.2 Second-order sufficient condition for SOCO

We highlight the connection between the second-order sufficient condition of Bonnans and Ramírez [21] and the primal nondegeneracy condition. In Section 4.2.3, we use the second-order sufficient condition to show the nonsingularity of the Jacobian of the equations in the KKT conditions for a reduced NLO problem.

4.2.2.1 Second-order sufficient condition for the dual problem

Let $h \in \mathbb{R}^m$, and assume that $\mathcal{R}$ is nonempty so that there exists $(\check{y}; \check{s}) \in \mathcal{D}^*_{\mathrm{SOCO}}$ with $\check{s}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\}$ for some $i \in \mathcal{R}$. Then the specialization of the second-order sufficient condition for $(\mathrm{D}_{\mathrm{SOCO}})$, where the objective is to minimize $-b^T y$, is given by

$$\sup_{\tilde{x}\in\mathcal{P}^*_{\mathrm{SOCO}}} h^T H_{\mathrm{D}}(\check{y},\tilde{x})h > 0, \quad \forall h \in \mathcal{C}_{\mathrm{D}}(\check{y}) \setminus \{0\}, \tag{4.22}$$

where

$$\begin{aligned} H_{\mathrm{D}}(\check{y},\tilde{x}) &:= \sum_{i=1}^{p} H^i_{\mathrm{D}}(\check{y},\tilde{x}),\\ H^i_{\mathrm{D}}(\check{y},\tilde{x}) &:= \begin{cases} -\frac{\tilde{x}^i_1}{\check{s}^i_1} A_i R_i A_i^T, & \check{s}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\},\\ \mathbf{0}_{m\times m}, & \text{otherwise}, \end{cases} \qquad i = 1,\dots,p, \end{aligned}$$

and $\mathcal{C}_{\mathrm{D}}(\check{y})$ is the cone of critical directions for $(\mathrm{D}_{\mathrm{SOCO}})$ which is defined as follows

$$\begin{cases} h \in \mathbb{R}^m, & \\ -A_i^T h \in \mathbb{L}^{n_i}_+, & \tilde{x}^i, \check{s}^i = 0,\\ -A_i^T h \in \left\{ d \mid d^T_{2:n_i}\check{s}^i_{2:n_i} - d_1 \check{s}^i_1 \le 0 \right\}, & \tilde{x}^i = 0,\ \check{s}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\},\\ A_i^T h = 0, & \tilde{x}^i \in \mathrm{int}(\mathbb{L}^{n_i}_+),\\ (\tilde{x}^i)^T A_i^T h = 0, & \tilde{x}^i, \check{s}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\},\\ -A_i^T h \in \mathbb{R}_+(\tilde{x}^i_1; -\tilde{x}^i_{2:n_i}), & \tilde{x}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\},\ \check{s}^i = 0. \end{cases} \tag{4.23}$$

Then, by $\tilde{x}^i = 0$ for $i \in \mathcal{T}_3$, we have

$$h^T H_{\mathrm{D}}(\check{y},\tilde{x})h = \sum_{i\in\mathcal{R}} -\frac{\tilde{x}^i_1}{\check{s}^i_1} h^T A_i R_i A_i^T h.$$

It follows from the definition, see e.g., (3.109) in [23], that $h^T H_{\mathrm{D}}(\tilde{y},\tilde{x})h \ge 0$ for all $(\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^*_{\mathrm{SOCO}} \times \mathcal{D}^*_{\mathrm{SOCO}}$ and $h \in \mathcal{C}_{\mathrm{D}}(\tilde{y})$. To see this, let

$$A_i^T h =: \left((A_i^T h)_1;\ (A_i^T h)_{2:n_i}\right), \quad \forall i.$$

Then for all $h \in \mathcal{C}_{\mathrm{D}}(\tilde{y}) \setminus \{0\}$ we have

$$\begin{aligned}0 = (\tilde{x}^i)^T A_i^T h &= \tilde{x}_1^i (A_i^T h)_1 + (\tilde{x}_{2:n_i}^i)^T (A_i^T h)_{2:n_i} \\ &\geq \tilde{x}_1^i (A_i^T h)_1 - \big\|\tilde{x}_{2:n_i}^i\big\|_2 \big\|(A_i^T h)_{2:n_i}\big\|_2 \\ &= \tilde{x}_1^i \Big((A_i^T h)_1 - \big\|(A_i^T h)_{2:n_i}\big\|_2 \Big), \qquad \forall i \in \mathcal{R},\end{aligned}$$

where the last equality follows from $\tilde{x}_1^i = \|\tilde{x}_{2:n_i}^i\|_2$. Since $\tilde{x}_1^i > 0$, we can conclude that

$$(A_i^T h)_1 - \big\|(A_i^T h)_{2:n_i}\big\|_2 \leq 0, \qquad \forall i \in \mathcal{R}.$$

Analogously, we can derive

$$0 = (\tilde{x}^i)^T A_i^T h = \tilde{x}_1^i (A_i^T h)_1 + (\tilde{x}_{2:n_i}^i)^T (A_i^T h)_{2:n_i} \leq \tilde{x}_1^i \Big((A_i^T h)_1 + \big\|(A_i^T h)_{2:n_i}\big\|_2 \Big),$$

which implies

$$(A_i^T h)_1 + \big\|(A_i^T h)_{2:n_i}\big\|_2 \geq 0, \qquad \forall i \in \mathcal{R}.$$

Consequently,

$$h^T A_i R_i A_i^T h = \Big((A_i^T h)_1 - \big\|(A_i^T h)_{2:n_i}\big\|_2 \Big) \Big((A_i^T h)_1 + \big\|(A_i^T h)_{2:n_i}\big\|_2 \Big) \leq 0, \tag{4.24}$$

which implies $h^T H_{\mathrm{D}}(\tilde{y}, \tilde{x}) h \geq 0$.

The connection between the primal nondegeneracy condition and the second-order sufficient condition (4.22) is stated in the following lemma.

Lemma 4.2.1 (Proposition 3.2 in [97])**.** *Let $\mathcal{R} \neq \emptyset$ and $(x^*; y^*; s^*)$ be the unique optimal solution of* ($\mathrm{P_{SOCO}}$) *and* ($\mathrm{D_{SOCO}}$)*. Then, under the primal nondegeneracy condition, the second-order sufficient condition* (4.22) *holds at $(y^*; s^*)$.*

We can observe from the proof of Lemma 4.2.1 that under the primal nondegeneracy condition we have $h^T A_i R_i A_i^T h < 0$ for some $i \in \mathcal{R}$. More precisely, notice from $((x^*)^i)^T A_i^T h = 0$ that $A_i^T h \notin \mathrm{int}(\mathbb{L}_+^{n_i})$ and $-A_i^T h \notin \mathrm{int}(\mathbb{L}_+^{n_i})$ for every

$h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}$. Then from the characterization of the primal nondegeneracy condition in Theorem 4.2.1 we have that

$$\begin{aligned} A_i^T \eta = 0, &\quad i \in \mathcal{B}, \\ ((x^*)^i)^T A_i^T \eta = 0, &\quad i \in \mathcal{R} \cup \mathcal{T}_2, \\ (\hat{P}_i^*)^T A_i^T \eta = 0, &\quad i \in \mathcal{R} \cup \mathcal{T}_2 \end{aligned} \tag{4.25}$$

has only a trivial solution $\eta = 0$, where $\hat{P}_i^*$ is defined in (4.19), and $\eta \in \mathbb{R}^m$. From (4.23) we can observe that a critical direction $h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}$ satisfies

$$\begin{aligned} A_i^T h = 0, &\qquad i \in \mathcal{B}, \\ ((x^*)^i)^T A_i^T h = 0, &\qquad i \in \mathcal{R} \cup \mathcal{T}_2, \\ (\hat{P}_i^*)^T A_i^T h = 0, &\qquad i \in \mathcal{T}_2, \end{aligned}$$

where the last two equalities hold, because $-A_i^T h = \rho R_i (x^*)^i$ for some $\rho \geq 0$, and the columns of $\hat{P}_i^*$ are orthogonal to both $(x^*)^i$ and $R_i(x^*)^i$ for $i \in \mathcal{T}_2$. Therefore, we have $(\hat{P}_i^*)^T A_i^T h \neq 0$ for some $i \in \mathcal{R}$, since otherwise we would get a nontrivial solution η for (4.25). Consequently, from $(\hat{P}_i^*)^T A_i^T h \neq 0$ and $((x^*)^i)^T A_i^T h = 0$ it can be deducted that $A_i^T h \notin \mathbb{L}_+^{n_i}$ and $-A_i^T h \notin \mathbb{L}_+^{n_i}$ for all $h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}$.

Scheinberg in Section 4.2 in [152], and Bonnans and Shapiro in Theorem 5.91 in [23] assume the strict complementarity condition to establish a mutual relationship between the second-order sufficient condition, see Section 1.6), and the primal nondegeneracy condition for SDO. Here, we only need the primal nondegeneracy condition in Lemma 4.2.1 to ensure that the strong second-order sufficiency condition holds.

Example 4.2.1. *Consider the following SOCO problem from [5]:*

$$
\begin{aligned}
\min \quad & -x_2^1 \\
\text{s.t.} \quad & x_1^1 = 1, \\
& 2x_2^1 + x_3^1 - x_2^2 = 0, \\
& 2x_3^1 - x_3^2 = 0, \\
& x_1^2 = 2, \\
& x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2}, \\
& x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}.
\end{aligned} \tag{4.26}
$$

The SOCO problem (4.26) *satisfies the interior point condition, and it has the unique primal-dual optimal solution*

$$x^* = (1, 1, 0, 2, 2, 0)^T, \quad y^* = (-1, 0, 0, 0)^T, \quad s^* = (1, -1, 0, 0, 0, 0)^T,$$

which fails strict complementarity. The optimal partition is given by

$$\mathcal{R} = \{1\}, \quad \mathcal{T}_2 = \{2\}, \quad \mathcal{B} = \mathcal{N} = \mathcal{T}_1 = \mathcal{T}_3 = \emptyset.$$

We can check that both the primal and dual nondegeneracy conditions hold. Note that

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \bar{P}_1^* = \bar{P}_2^* = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1 \end{pmatrix},$$

which gives

$$(A_1 \bar{P}_1^*, A_2 \bar{P}_2^*) = \begin{pmatrix} 1/\sqrt{2} & 0 & 0 & 0 \\ \sqrt{2} & 1 & -1/\sqrt{2} & 0 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1/\sqrt{2} & 0 \end{pmatrix}, \; (A_1 R_1 (s^*)^1, A_2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Since both of these matrices are nonsingular, the unique optimal solution is primal and dual nondegenerate. For the dual problem the cone of critical directions is given by

$$\mathcal{C}_{\mathrm{D}}(y^*) = \begin{cases} h \in \mathbb{R}^4, \\ ((x^*)^1)^T A_1^T h = 0, \\ -A_2^T h \in \mathbb{R}_+ \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}, \end{cases}$$

which is equivalent to $\mathcal{C}_{\mathrm{D}}(y^*) = \left\{h \in \mathbb{R}^4 \mid h_1 \geq 0,\ h_3 = 0,\ h_2 = h_4 = -\frac{1}{2}h_1 \leq 0\right\}$. *Therefore, we get*

$$-\frac{(x^*)_1^1}{(s^*)_1^1} A_1 R_1 A_1^T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

which implies that

$$h^T H_{\mathrm{D}}(y^*, x^*) h = h_2^2 > 0, \qquad \forall h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}.$$

Thus, the second-order sufficient condition holds at $(y^*; s^*)$.

4.2.2.2 Second-order sufficient condition for the primal problem

It is straightforward to derive the cone of critical directions and the second-order sufficient condition for $(\mathrm{P_{SOCO}})$. To that end, note that $(\mathrm{P_{SOCO}})$ can be equivalently written as

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax - b \in \{0\}, \\ & x^i \in \mathbb{L}^{n_i}_+, \qquad i = 1, \ldots, p. \end{aligned}$$

In a similar manner, we can show that under the dual nondegeneracy condition, the second-order sufficient condition holds at the unique optimal solution of ($\mathrm{P}_{\mathrm{SOCO}}$). Let $\check{x} \in \mathcal{P}^*_{\mathrm{SOCO}}$ and assume that $\mathcal{R} \neq \emptyset$ with $\check{x}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\}$ for some $i \in \mathcal{R}$. We redefine $h := (h^1; \ldots; h^p)$, where $h^i \in \mathbb{R}^{n_i}$ for $i = 1, \ldots, p$, to refer to a critical directions[2] belonging to $\mathcal{C}_{\mathrm{P}}(\check{x})$ as defined as

$$\begin{cases} Ah = 0, & \\ h^i = 0, & \tilde{s}^i \in \mathrm{int}(\mathbb{L}^{n_i}_+), \\ h^i \in \mathbb{L}^{n_i}_+, & \check{x}^i, \tilde{s}^i = 0, \\ h^i \in \left\{ d \mid d^T_{2:n_i} \check{x}^i_{2:n_i} - d_1 \check{x}^i_1 \leq 0 \right\}, & \check{x}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\},\ \tilde{s}^i = 0, \\ h^i \in \mathbb{R}^{n_i}, & \check{x}^i \in \mathrm{int}(\mathbb{L}^{n_i}_+), \\ (\tilde{s}^i)^T h^i = 0, & \check{x}^i, \tilde{s}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\}, \\ h^i \in \mathbb{R}_+(\tilde{s}^i_1; -\tilde{s}^i_{2:n_i}), & \check{x}^i = 0,\ \tilde{s}^i \in \mathrm{bd}(\mathbb{L}^{n_i}_+) \setminus \{0\}. \end{cases} \tag{4.27}$$

Then the second-order sufficient condition for ($\mathrm{P}_{\mathrm{SOCO}}$) is given by

$$\sup_{(\tilde{y};\tilde{s}) \in \mathcal{D}^*_{\mathrm{SOCO}}} h^T H_{\mathrm{P}}(\check{x}, \tilde{s}) h > 0, \quad \forall h \in \mathcal{C}_{\mathrm{P}}(\check{x}) \setminus \{0\}, \tag{4.28}$$

[2]The proof is straightforward, and it easily follows from the complementarity of $\check{x}$ and $\tilde{s}$ and Proposition 3.10 in [23].

where

$$H_{\mathrm{P}}(\check{x}, \tilde{s}) := \sum_{i=1}^{p} H_{\mathrm{P}}^{i}(\check{x}, \tilde{s}),$$

$$H_{\mathrm{P}}^{i}(\check{x}, \tilde{s}) := \begin{cases} -\frac{\tilde{s}_1^i}{\check{x}_1^i}\mathrm{diag}(\mathbf{0}, R_i, \mathbf{0}), & \check{x}^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ \mathbf{0}_{\bar{n}\times\bar{n}}, & \text{otherwise}, \end{cases} \qquad i = 1, \ldots, p,$$

in which $\mathrm{diag}(\mathbf{0}, R_i, \mathbf{0})$ is a block diagonal matrix whose i^{th} block is R_i and 0 elsewhere. Then by $\tilde{s}^i = 0$ for $i \in \mathcal{T}_2$, we have

$$h^T H_{\mathrm{P}}(\check{x}, \tilde{s})h = \sum_{i\in\mathcal{R}} -\frac{\tilde{s}_1^i}{\check{x}_1^i}((h_1^i)^2 - \|h_{2:n_i}^i\|_2^2).$$

Analogously, we can show that $h^T H_{\mathrm{P}}(\tilde{x}, \tilde{s})h \geq 0$ for all $(\tilde{x}; \tilde{y}; \tilde{s}) \in \mathcal{P}^*_{\mathrm{SOCO}} \times \mathcal{D}^*_{\mathrm{SOCO}}$ and $h \in \mathcal{C}_{\mathrm{P}}(\tilde{x})$. To that end, we have

$$\tilde{s}_1^i\Big(h_1^i - \|h_{2:n_i}^i\|_2\Big) \leq 0 = (\tilde{s}^i)^T h^i \leq \tilde{s}_1^i\Big(h_1^i + \|h_{2:n_i}^i\|_2\Big), \qquad \forall i \in \mathcal{R},$$

which follows from $\tilde{s}_1^i = \|\tilde{s}_{2:n_i}^i\|_2$. Since $\tilde{s}_1^i > 0$, we have

$$(h_1^i)^2 - \left\|h_{2:n_i}^i\right\|_2^2 \leq 0, \qquad \forall i \in \mathcal{R}. \tag{4.29}$$

The next lemma shows that under the dual nondegeneracy condition (4.29) holds with strict inequality for some $i \in \mathcal{R}$. For the sake of completeness, we provide an illustrative proof.

Lemma 4.2.2. *Let $(x^*; y^*; s^*)$ be the unique optimal solution of* ($\mathrm{P_{SOCO}}$) *and* ($\mathrm{D_{SOCO}}$) *so that $\mathcal{R} \neq \emptyset$. Then, under the dual nondegeneracy condition the second-order sufficient condition* (4.28) *holds at x^*.*

Proof. Let $h \in \mathcal{C}_{\mathrm{P}}(x^*) \setminus \{0\}$, and assume that $(h_1^i)^2 - \|h_{2:n_i}^i\|_2^2 = 0$ for every $i \in \mathcal{R}$, which implies either $h^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})$ or $-h^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})$. Then, from $((s^*)^i)^T h^i = 0$, it follows that

$$h^i \in \mathbb{R}\big((s^*)_1^i; -(s^*)_{2:n_i}^i\big), \qquad i \in \mathcal{R}.$$

Noting that $h^i = 0$ for $i \in \mathcal{N}$ and $Ah = 0$, we get

$$\sum_{i\in\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2} A_i h^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_3} A_i h^i = \sum_{i\in\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2} A_i h^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_3} \alpha_i A_i R_i (s^*)^i = 0,$$

where $\alpha_i \geq 0$ for $i \in \mathcal{T}_3$. All this implies that

$$\Big((A_i R_i (s^*)^i)_{i\in\mathcal{R}\cup\mathcal{T}_3},\ A_{\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2} \Big)$$

has linearly dependent columns, which contradicts the dual nondegeneracy condition. This completes the proof. □

4.2.3 Quadratic convergence under failure of strict complementarity

In this section, under the primal and dual nondegeneracy conditions, we establish quadratic convergence of Newton's method to the unique optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). To that end, we need the optimal partition $(\mathcal{B}, \mathcal{N}, \mathcal{R}, \mathcal{T})$ to be known and $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$ to be correctly identified. Hence, it is assumed that $\mu < \tilde{\mu}$ allows for a complete identification of $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$.

Lemma 4.2.3. *Assume that the primal and dual nondegeneracy conditions hold. Then $\mathcal{R} = \emptyset$ implies $\mathcal{T} = \emptyset$.*

Proof. Suppose that $\mathcal{R} = \emptyset$ and $\mathcal{T} \neq \emptyset$, and $(x^*; y^*; s^*)$ is the unique[3] optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). Then, for every possible case in which $\mathcal{T}_1$, $\mathcal{T}_2$, or $\mathcal{T}_3$ is nonempty, the number of columns in (4.20) is strictly greater than the number of columns in (4.19), i.e., (4.19) and (4.20) have $\sum_{i\in\mathcal{B}\cup\mathcal{T}_2} n_i - |\mathcal{T}_2|$ and $\sum_{i\in\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2} n_i + |\mathcal{T}_3|$ columns, respectively. Thus, (4.19) and (4.20) cannot be simultaneously of full row rank and full column rank. □

[3] Otherwise, the nondegeneracy conditions fail.

As a result of Lemma 4.2.3, if $\mathcal{R} = \emptyset$, then $A_\mathcal{B}$ is a nonsingular matrix by the primal and dual nondegeneracy conditions. Therefore, the unique optimal solutions of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) can be obtained by solving two linear systems of equations. Hence, in the sequel we assume that $\mathcal{R} \neq \emptyset$.

Let $(x^*; y^*; s^*)$ be the unique optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) which satisfies the primal and dual nondegeneracy conditions. Further, let us assume that $\mathcal{T}_1, \mathcal{T}_3 \neq \emptyset$. If we drop the dual constraints $c^i - A_i^T y \in \mathbb{L}_+^{n_i}$ for $i \in \mathcal{T}_1 \cup \mathcal{T}_3$, then we obtain a relaxation of ($\mathrm{D_{SOCO}}$) as

$$(\mathrm{D'_{SOCO}}) \quad \max \big\{ b^T y \mid A_i^T y + s^i = c^i, \quad s^i \in \mathbb{L}_+^{n_i}, \quad i \in \{1, \ldots, p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\} \big\},$$

and its dual is written as

$$\begin{aligned}(\mathrm{P'_{SOCO}}) \quad \min \big\{ & \sum_{i \in \{1,\ldots,p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}} (c^i)^T x^i \mid \\ & \sum_{i \in \{1,\ldots,p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}} A_i x^i = b, \quad x^i \in \mathbb{L}_+^{n_i}, \quad i \in \{1, \ldots, p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\} \big\}.\end{aligned}$$

Since $(x^*)^i = 0$ for $i \in \mathcal{T}_1 \cup \mathcal{T}_3$, it follows from the optimality conditions, (4.19), and (4.20) that $((x^*)^i; y^*; (s^*)^i)$ for $i \in \{1, \ldots, p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}$ is a primal-dual optimal solution for ($\mathrm{P'_{SOCO}}$) and ($\mathrm{D'_{SOCO}}$), and it satisfies the primal and dual nondegeneracy conditions. To see this, the primal nondegeneracy condition is the same as the one for x^*, and the dual nondegeneracy condition needs

$$\big((A_i R_i (s^*)^i)_{i \in \mathcal{R}}, \ A_{\mathcal{B} \cup \mathcal{T}_2} \big)$$

to have linearly independent columns, which is true by the dual nondegeneracy of $(y^*; s^*)$. As a result, if we remove the columns of $\mathcal{T}_1$ and $\mathcal{T}_3$ from A and c, we can recover the unique optimal solutions of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) by solving ($\mathrm{P'_{SOCO}}$) and ($\mathrm{D'_{SOCO}}$). At the risk of causing confusion, we refer to $(\bar{x}; \bar{y}; \bar{s})$ as the unique optimal solution of ($\mathrm{P'_{SOCO}}$) and ($\mathrm{D'_{SOCO}}$).

The algebraic definition (1.7) can be used to reformulate ($\mathrm{D}'_{\mathrm{SOCO}}$) as a nonconvex NLO problem. Then inspired by the optimal partition information and the characteristics of a maximally complementary optimal solution, one can realize that the unique dual optimal solution $(\bar{y}; \bar{s})$ can be obtained by solving the NLO reformulation of ($\mathrm{D}'_{\mathrm{SOCO}}$) as

$$
\begin{aligned}
(\mathrm{D}_{\mathrm{NLO}}) \quad \min \quad & -b^T w \\
\text{s.t.} \quad A_i^T w &= c^i, && i \in \mathcal{B} \cup \mathcal{T}_2, \\
A_i^T w + z^i &= c^i, && i \in \mathcal{R} \cup \mathcal{N}, \\
(z^i)^T R_i z^i &= 0, && i \in \mathcal{R}, \\
z &\in \mathcal{W},
\end{aligned}
$$

where $w \in \mathbb{R}^m$, $z^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{R} \cup \mathcal{N}$, and $\mathcal{W}$ is a nonempty open convex cone defined as

$$
\mathcal{W} := \left\{ z \mid z_1^i > 0, \ i \in \mathcal{R}, \quad z^i \in \operatorname{int}(\mathbb{L}_+^{n_i}), \ i \in \mathcal{N} \right\}.
$$

Let z denote the concatenation of the column vectors z^i for $i \in \mathcal{R} \cup \mathcal{N}$. It then follows that ($\mathrm{D}_{\mathrm{NLO}}$) has the unique globally optimal solution $(\bar{w}; \bar{z})$, since otherwise the optimality or the uniqueness of $(\bar{y}; \bar{s})$ is contradicted. The unique globally optimal solution is given by

$$
\bar{w} := \bar{y}, \qquad \bar{z}^i := \bar{s}^i, \quad i \in \mathcal{R} \cup \mathcal{N}. \tag{4.30}
$$

In a similar manner, the unique optimal solution $\bar{x}$ can be computed by solving

$$
\begin{aligned}
(\mathrm{P}_{\mathrm{NLO}}) \quad \min \quad & \sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} (c^i)^T \nu^i \\
\text{s.t.} \quad & \sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} A_i \nu^i = b, \\
& (\nu^i)^T R_i \nu^i = 0, && i \in \mathcal{R} \cup \mathcal{T}_2, \\
& \nu \in \mathcal{V},
\end{aligned}
$$

where $\nu^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2$, and $\mathcal{V}$ is an open convex cone defined as

$$\mathcal{V} := \left\{ \nu \mid \nu_1^i > 0,\ i \in \mathcal{R} \cup \mathcal{T}_2, \quad \nu^i \in \operatorname{int}(\mathbb{L}_+^{n_i}),\ i \in \mathcal{B} \right\}.$$

For the sake of convenience, we only consider ($\mathrm{D_{NLO}}$). Analogous results can be derived for problem ($\mathrm{P_{NLO}}$).

Let $u^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}$ and $v \in \mathbb{R}^{|\mathcal{R}|}$ be the Lagrange multipliers associated with the constraints in ($\mathrm{D_{NLO}}$). The first-order optimality conditions, see (1.29), are given by

$$\begin{aligned}
-\sum_{i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}} A_i u^i &= b, \\
-u^i - 2v_i R_i z^i &= 0, \quad && i \in \mathcal{R}, \\
-u^i &= 0, \quad && i \in \mathcal{N}, \\
A_i^T w &= c^i, \quad && i \in \mathcal{B} \cup \mathcal{T}_2, \\
A_i^T w + z^i &= c^i, \quad && i \in \mathcal{R} \cup \mathcal{N}, \\
(z^i)^T R_i z^i &= 0, \quad && i \in \mathcal{R}, \\
z &\in \mathcal{W},
\end{aligned} \tag{4.31}$$

which bears a striking resemblance to the optimality conditions (1.10). Let u be the concatenation of the column vectors u^i for $i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}$. Then we can observe that for $\bar{z} \in \mathcal{W}$ there exist Lagrange multipliers $\bar{u}$ and $\bar{v}$ so that $(\bar{w}; \bar{z}; \bar{u}; \bar{v})$ satisfies the first-order optimality conditions (4.31). Such a solution can be obtained by setting

$$\begin{aligned}
\bar{u}^i &:= -\bar{x}^i, \quad && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R}, \\
\bar{u}^i &:= 0, && i \in \mathcal{N}, \\
\bar{v}_i &:= \frac{1}{2} \frac{\bar{x}_1^i}{\bar{s}_1^i}, && i \in \mathcal{R}.
\end{aligned} \tag{4.32}$$

We show in Lemma 4.2.4 that, under the dual nondegeneracy condition, the Lagrange multipliers $(\bar{u};\bar{v})$ are unique. Let $J((w;z))$ denote the Jacobian of the equality constraints in $(\mathrm{D}_{\mathrm{NLO}})$ as follows

$$J((w;z)) := \begin{pmatrix} A_{\mathcal{B}}^T & 0 & 0 \\ A_{\mathcal{T}_2}^T & 0 & 0 \\ A_{\mathcal{R}}^T & I & 0 \\ A_{\mathcal{N}}^T & 0 & I \\ 0 & Z_{\mathcal{R}} & 0 \end{pmatrix}, \tag{4.33}$$

where $Z_{\mathcal{R}}$ is given by

$$Z_{\mathcal{R}} := \begin{pmatrix} 2(z_1^1; -z_{2:n_1}^1)^T & 0 & 0 & 0 & \\ & 2(z_1^2; -z_{2:n_2}^2)^T & & & \\ 0 & 0 & \ddots & 0 & \\ & & & 2(z_1^i; -z_{2:n_i}^i)^T & \\ 0 & 0 & 0 & \ddots & \end{pmatrix},$$

in which $i \in \mathcal{R}$. Note that $Z_{\mathcal{R}}$ has full row rank since $(z^i)^T R_i \neq 0$ for every $i \in \mathcal{R}$.

Lemma 4.2.4. *Let $(\bar{w};\bar{z})$ be the unique globally optimal solution of* $(\mathrm{D}_{\mathrm{NLO}})$*. Then, under the dual nondegeneracy condition, $J((\bar{w};\bar{z}))$ has full row rank.*

Proof. We show that $J((\bar{w};\bar{z}))^T\eta = 0$ has only the trivial solution $\eta = 0$, where $\eta := (\eta^1;\ldots;\eta^5)$ is a vector of appropriate size. Then from $J((\bar{w};\bar{z}))^T\eta = 0$ we have

$$\begin{aligned} A_{\mathcal{B}}\eta^1 + A_{\mathcal{T}_2}\eta^2 + A_{\mathcal{R}}\eta^3 + A_{\mathcal{N}}\eta^4 &= 0, \\ \eta^3 + \bar{Z}_{\mathcal{R}}^T\eta^5 &= 0, \\ \eta^4 &= 0, \end{aligned}$$

which implies

$$A_{\mathcal{B}}\eta^1 + A_{\mathcal{T}_2}\eta^2 - A_{\mathcal{R}}\bar{Z}_{\mathcal{R}}^T\eta^5 = 0,$$

where $A_{\mathcal{R}}\bar{Z}_{\mathcal{R}}^T = (2A_1R_1\bar{z}^1, \ldots, 2A_iR_i\bar{z}^i, \ldots)$ for $i \in \mathcal{R}$. Since $(\bar{y}; \bar{s})$ is the unique dual nondegenerate optimal solution of ($\mathrm{D}'_{\mathrm{SOCO}}$), it follows from (4.20) that $(A_{\mathcal{R}}\bar{Z}_{\mathcal{R}}^T,\ A_{\mathcal{B}\cup\mathcal{T}_2})$ has full column rank, and thus $\eta = 0$ is the unique solution of $J((\bar{w}; \bar{z}))^T\eta = 0$. □

Under the full rank result of Lemma 4.2.4, LICQ holds at $(\bar{w}; \bar{z})$. This regularity condition guarantees that the set of Lagrange multipliers associated with $(\bar{w}; \bar{z})$ is a singleton.

For the sake of simplicity let $\vartheta := (w; z; u; v)$. The Lagrangian function of ($\mathrm{D}_{\mathrm{NLO}}$) is defined as

$$\begin{aligned}&\mathcal{L}(\vartheta) := \\ &\quad - b^Tw - \sum_{i\in\mathcal{B}\cup\mathcal{T}_2}(u^i)^T(A_i^Tw - c^i) - \sum_{i\in\mathcal{R}\cup\mathcal{N}}(u^i)^T(A_i^Tw + z^i - c^i) - \sum_{i\in\mathcal{R}} v_i(z^i)^TR_iz^i,\end{aligned}$$

and the Hessian of $\mathcal{L}(\vartheta)$ is given by

$$\nabla^2\mathcal{L}(\vartheta) := \begin{pmatrix} 0 & 0 & 0 \\ 0 & V_{\mathcal{R}} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where

$$V_{\mathcal{R}} := -2\mathrm{diag}(v_1R_1, v_2R_2, \ldots, v_iR_i, \ldots)$$

is a block diagonal matrix, in which $i \in \mathcal{R}$. Let $h = (h^1; h^2; h^3) \in \mathrm{Ker}(J((\bar{w}; \bar{z})))$, where $h^1 \in \mathbb{R}^m$ and h^2 as well as h^3 is the concatenation of the vectors $(h^2)^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{R}$ and $(h^3)^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{N}$, respectively. In Lemma 4.2.5, we show that under the primal nondegeneracy condition, the second-order sufficient condition for ($\mathrm{D}_{\mathrm{NLO}}$) holds at $(\bar{w}; \bar{z})$, i.e.,

$$h^T\nabla^2\mathcal{L}(\bar{\vartheta})h > 0, \quad \forall h \in \mathrm{Ker}(J((\bar{w}; \bar{z}))) \setminus \{0\}, \tag{4.34}$$

in which $\bar{\vartheta} := (\bar{w}; \bar{z}; \bar{u}; \bar{v})$.

Lemma 4.2.5. *Let $(\bar{w};\bar{z})$ be the unique globally optimal solution of* $(\mathrm{D_{NLO}})$*. Then, under the primal nondegeneracy condition, the second-order sufficient condition* (4.34) *holds at $(\bar{w};\bar{z})$.*

Proof. Note that $\mathrm{Ker}(J((\bar{w};\bar{z})))$ can be equivalently written as the solution set of

$$\begin{aligned} A_i^T h^1 &= 0, \quad i \in \mathcal{B} \cup \mathcal{T}_2, \\ A_i^T h^1 + (h^2)^i &= 0, \quad i \in \mathcal{R}, \\ A_i^T h^1 + (h^3)^i &= 0, \quad i \in \mathcal{N}, \\ (\bar{z}^i)^T R_i (h^2)^i &= 0, \quad i \in \mathcal{R}. \end{aligned} \tag{4.35}$$

Then we get

$$h^T \nabla^2 \mathcal{L}(\bar{\vartheta}) h = -2 \sum_{i \in \mathcal{R}} \bar{v}_i ((h^2)^i)^T R_i (h^2)^i = -2 \sum_{i \in \mathcal{R}} \bar{v}_i (h^1)^T A_i R_i A_i^T h^1.$$

By the primal nondegeneracy condition and the argument after Lemma 4.2.1, for the unique primal optimal solution $\bar{x}$ system (4.25) has only a trivial solution. Thus, we have $(\hat{Q}_i^*)^T A_i^T h^1 \neq 0$ for some $i \in \mathcal{R}$, where $\hat{P}_i^*$ is defined as in (4.19). Hence, it follows from (4.32) and (4.35) that $\bar{v}_i (h^1)^T A_i R_i A_i^T h^1 < 0$ for Lagrange multipliers $(\bar{u};\bar{v})$ for all $h^1 \neq 0$ satisfying (4.35). $\square$

Remark 4.2.1. *Since LICQ holds at $(\bar{w};\bar{z})$, and the second-order sufficient condition* (4.34) *holds at $\bar{\vartheta}$, it follows from Lemma 3.2.2 in [50], that $(\bar{w};\bar{z})$ is an isolated locally optimal solution, i.e., $(\bar{w};\bar{z})$ is the only locally optimal solution in some of its neighborhoods.*

Example 4.2.2. *Problem* (4.26) *in the nonlinear format* $(\mathrm{P_{NLO}})$ *has four isolated* *locally optimal solutions*

$$\begin{aligned}
\nu_{(1)} &= (1,\ -0.2425,\ 0.9701,\ 2,\ 0.4851,\ 1.9403)^T,\\
\nu_{(2)} &= (1,\ -1,\ 0,\ 2,\ -2,\ 0)^T,\\
\nu_{(3)} &= (1,\ 0.2425,\ -0.9701,\ 2,\ -0.4851,\ -1.9403)^T,\\
\nu_{(4)} &= (1,\ 1,\ 0,\ 2,\ 2,\ 0)^T,
\end{aligned}$$

where the objective values are 0.2425, 1, -0.2425, and -1, respectively. The second-order constraint $x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}$ *is weakly inactive at* $\nu_{(4)}$*, i.e., its removal does not affect the optimality of* $\nu_{(4)}$*. Removing the weakly inactive constraint reduces the set of locally optimal solutions to* $\{\nu_{(2)}, \nu_{(4)}\}$ *but leaves the set of globally optimal solutions* $\{\nu_{(4)}\}$ *unchanged. Note that the Jacobian matrix* (4.33) *is nonsingular at the unique globally optimal solution* $(\bar{w}; \bar{z})$*. Therefore, the second-order sufficient condition* (4.34) *trivially holds at* $(\bar{w}; \bar{z})$*.*

4.2.3.1 Quadratic convergence of Newton's method

We apply Newton's method to the first-order optimality conditions of ($\mathrm{D}_{\mathrm{NLO}}$). The idea is to start from a central solution, for which μ satisfies (3.49), and take Newton steps to converge to $\bar{\vartheta}$. The first-order optimality conditions (4.31) can be written as $G(\vartheta) = 0$ and $z \in \mathcal{W}$, where the mapping $G : \mathbb{R}^{\bar{n}_c} \to \mathbb{R}^{\bar{n}_c}$ is defined as

$$G(\vartheta) := \begin{pmatrix} -\sum_{i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}} A_i u^i - b & \\ -u^i - 2v_i R_i z^i & i \in \mathcal{R} \\ -u^i & i \in \mathcal{N} \\ A_i^T w - c^i & i \in \mathcal{B} \cup \mathcal{T}_2 \\ A_i^T w + z^i - c^i & i \in \mathcal{R} \cup \mathcal{N} \\ (z^i)^T R_i z^i & i \in \mathcal{R} \end{pmatrix}, \tag{4.36}$$

in which

$$\bar{n}_c := \sum_{i\in\mathcal{B}\cup\mathcal{T}_2\cup\mathcal{R}\cup\mathcal{N}} n_i + \sum_{i\in\mathcal{R}\cup\mathcal{N}} n_i + |\mathcal{R}| + m.$$

For ease of exposition, the equations of (4.31) are indexed in mapping G. The Jacobian of G is given by

$$\nabla G(\vartheta) := \begin{pmatrix} \nabla^2\mathcal{L}(\vartheta) & -J((w;z))^T \\ J((w;z)) & 0 \end{pmatrix}.$$

Letting $\vartheta^{(k)}$ be the k^{th} iterate, a Newton step is taken by computing

$$\vartheta^{(k+1)} := \vartheta^{(k)} + d\vartheta^{(k)}, \qquad d\vartheta^{(k)} := \left(dw^{(k)}; dz^{(k)}; du^{(k)}; dv^{(k)}\right), \tag{4.37}$$

where the search direction $d\vartheta^{(k)}$ is obtained by solving

$$\nabla G(\vartheta^{(k)})d\vartheta^{(k)} = -G(\vartheta^{(k)}). \tag{4.38}$$

Remark 4.2.2. *The underlying idea of the iterative procedure differs from a primal-dual IPM in that the iterative procedure applies the Newton's method to $G(.)$, while IPMs apply the Newton's method to the central path equations* (1.23)*. When $\mu = 0$, the primal-dual system* (1.23) *is equivalent to the optimal set of* ($\mathrm{P}_{\mathrm{SOCO}}$) *and* ($\mathrm{D}_{\mathrm{SOCO}}$) *while* (4.31) *includes locally optimal solutions. More importantly, the Jacobian of* (1.23) *at $\mu = 0$ is nonsingular under the strict complementarity and nondegeneracy conditions while only the latter is needed in our case.*

Lemma 4.2.4 shows that $J((\bar{w};\bar{z}))$ is of full row rank, and by Lemma 4.2.5 it holds that $\mathcal{L}(\bar{\vartheta})$ has a positive curvature in the null space of $J((\bar{w};\bar{z}))$. Now, we show that $\nabla G(\bar{\vartheta})$ is nonsingular.

Lemma 4.2.6. *Assume that the primal and dual nondegeneracy conditions hold. Then $\nabla G(\bar{\vartheta})$ is nonsingular.*

Proof. Let $\eta := (\eta^1; \eta^2)$ be a vector of appropriate size and consider the linear system $\nabla G(\bar{\vartheta})\eta = 0$. Then we have

$$\begin{aligned} \nabla^2 \mathcal{L}(\bar{\vartheta})\eta^1 - J((\bar{w};\bar{z}))^T \eta^2 = 0, \\ J((\bar{w};\bar{z}))\eta^1 = 0. \end{aligned}$$

From the first equation we have $(\eta^1)^T \nabla^2 \mathcal{L}(\bar{\vartheta})\eta^1 = 0$, which implies $\eta^1 = 0$ by Lemma 4.2.5. Setting $\eta^1 = 0$, the first equation gives $J((\bar{w};\bar{z}))^T \eta^2$, which implies $\eta^2 = 0$ by Lemma 4.2.4. □

The next lemma shows that ∇G is Lipschitz continuous, regardless of any regularity condition.

Lemma 4.2.7. *The Jacobian ∇G is Lipschitz continuous with global Lipschitz constant $\tau_1 := 2\sqrt{2}$.*

Proof. Let $\xi := (\xi^1; \ldots; \xi^8)$ be a vector of appropriate size. Then we have

$$\begin{aligned} \|\nabla G(\vartheta) - \nabla G(\vartheta')\|_2 &\leq \max_{\|\xi\|_2=1} \|(V_{\mathcal{R}} - V'_{\mathcal{R}})\xi^2\|_2 \\ &\quad + \max_{\|\xi\|_2=1} \|((Z'_{\mathcal{R}})^T - Z^T_{\mathcal{R}})\xi^8\|_2 + \max_{\|\xi\|_2=1} \|(Z_{\mathcal{R}} - Z'_{\mathcal{R}})\xi^2\|_2 \\ &\leq \max_{\|\xi^2\|_2=1} \|(V_{\mathcal{R}} - V'_{\mathcal{R}})\xi^2\|_2 + 2 \max_{\|\xi^2\|_2=1} \|(Z_{\mathcal{R}} - Z'_{\mathcal{R}})\xi^2\|_2 \\ &= \|V_{\mathcal{R}} - V'_{\mathcal{R}}\|_2 + 2\|Z_{\mathcal{R}} - Z'_{\mathcal{R}}\|_2. \end{aligned}$$

Then from the properties of the spectral norm we get

$$\begin{aligned} \|V_{\mathcal{R}} - V'_{\mathcal{R}}\|_2 &\leq 2 \max_{i\in\mathcal{R}} |v_i - v'_i| \leq 2\|v - v'\|_2, \\ \|Z_{\mathcal{R}} - Z'_{\mathcal{R}}\|_2 &\leq \sqrt{\max_{i\in\mathcal{R}} \|z^i - (z')^i\|_2^2} \leq \|z - z'\|_2. \end{aligned}$$

All this gives

$$\|\nabla G(\vartheta) - \nabla G(\vartheta')\|_2 \leq 2(\|v - v'\|_2 + \|z - z'\|_2) \leq 2\sqrt{2}\|\vartheta - \vartheta'\|_2,$$

for all ϑ and ϑ'. □

The following lemma will be useful for establishing the quadratic convergence of Newton's method.

Lemma 4.2.8. *Let* $(x^\mu; y^\mu; s^\mu)$ *be a central solution with* $\mu \le \hat{\mu}$*, where* $\hat{\mu}$ *is defined by* (3.43)*,* $(\bar{x}; \bar{y}; \bar{s})$ *be the unique optimal solution of* $(\mathrm{P}'_{\mathrm{SOCO}})$ *and* $(\mathrm{D}'_{\mathrm{SOCO}})$*, and* $(x^*; y^*; s^*)$ *be the unique optimal solution of* $(\mathrm{P}_{\mathrm{SOCO}})$ *and* $(\mathrm{D}_{\mathrm{SOCO}})$*. Then, under the primal and dual nondegeneracy conditions, we have*

$$\sqrt{\sum_{i\in\mathcal{R}} \left(\frac{(x^\mu)^i_1}{(s^\mu)^i_1} - \frac{\bar{x}^i_1}{\bar{s}^i_1} \right)^2} \le \frac{4p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right). \tag{4.39}$$

Proof. Note that for every $i \in \mathcal{R}$ we have

$$\frac{\bar{x}^i_{2:n_i}}{\|\bar{x}^i_{2:n_i}\|_2} = -\frac{\bar{s}^i_{2:n_i}}{\|\bar{s}^i_{2:n_i}\|_2}. \tag{4.40}$$

Since $\bar{x}^i = (x^*)^i$ and $\bar{s}^i = (s^*)^i$ for $i \in \mathcal{R}$, it follows from (3.40) and (4.40) that for every $i \in \mathcal{R}$

$$\begin{aligned} \sigma_2 \le \bar{x}^i_1 + \bar{s}^i_1 - \|\bar{x}^i_{2:n_i} + \bar{s}^i_{2:n_i}\|_2 &= \bar{x}^i_1 + \bar{s}^i_1 - |\bar{x}^i_1 - \bar{s}^i_1| \\ &= 2\min\{\bar{x}^i_1, \bar{s}^i_1\}. \end{aligned} \tag{4.41}$$

Furthermore, it holds that

$$\begin{aligned} \left| \frac{(x^\mu)^i_1}{(s^\mu)^i_1} - \frac{\bar{x}^i_1}{\bar{s}^i_1} \right| &= \left| \left(\frac{(x^\mu)^i_1}{(s^\mu)^i_1} - \frac{\bar{x}^i_1}{(s^\mu)^i_1} \right) + \left(\frac{\bar{x}^i_1}{(s^\mu)^i_1} - \frac{\bar{x}^i_1}{\bar{s}^i_1} \right) \right| \\ &\le \frac{1}{(s^\mu)^i_1} |(x^\mu)^i_1 - \bar{x}^i_1| + \bar{x}^i_1 \left| \frac{\bar{s}^i_1 - (s^\mu)^i_1}{(s^\mu)^i_1 \bar{s}^i_1} \right| \\ &\le \frac{1}{(s^\mu)^i_1} \|(x^\mu)^i - \bar{x}^i\|_2 + \frac{\bar{x}^i_1}{(s^\mu)^i_1 \bar{s}^i_1} \|(s^\mu)^i - \bar{s}^i\|_2 \le \frac{\kappa(p\mu)^\gamma}{(s^\mu)^i_1}\left(1 + \frac{\bar{x}^i_1}{\bar{s}^i_1}\right), \end{aligned}$$

where the last inequality follows from (3.44). Now using (3.41), (4.41), and Theorem 3.2.1 we get

$$\left| \frac{(x^\mu)^i_1}{(s^\mu)^i_1} - \frac{\bar{x}^i_1}{\bar{s}^i_1} \right| \le \frac{4p\kappa(p\mu)^\gamma}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right),$$

which completes the proof. $\square$

Let Newton's method be initiated with a given interior solution

$$\begin{aligned} w^{(0)} &:= y^\mu, \\ (z^i)^{(0)} &:= (s^\mu)^i, \qquad i \in \mathcal{R} \cup \mathcal{N}, \\ (u^i)^{(0)} &:= -(x^\mu)^i, \quad i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}, \\ v_i^{(0)} &:= \frac{1}{2}\frac{(x^\mu)_1^i}{(s^\mu)_1^i}, \quad i \in \mathcal{R}. \end{aligned} \tag{4.42}$$

Then a search direction is computed by using (4.38), and the new iterate is obtained by (4.37). Theorem 4.2.2 shows that if μ is sufficiently small, then Newton's method converges quadratically to the unique optimal solution $(\bar{x}; \bar{y}; \bar{s})$. To that end, we adopt the quadratic convergence result of Newton's method from Theorem A.4.1.

Theorem 4.2.2. *Assume that the primal and dual nondegeneracy conditions hold. Let*

$$\mu < \min\left\{ p^{-1}\left(4\sqrt{2}\theta_1\kappa\left(\sqrt{3} + \frac{2p\sqrt{|\mathcal{R}|}}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right)\right)\right)^{-\frac{1}{\gamma}},\ \tilde{\mu}\right\}, \tag{4.43}$$

in which θ_1 denotes an upper bound on $\|\nabla G(\bar{\vartheta})^{-1}\|_2$, and $\tilde{\mu}$ is defined in (3.49). Then, initiated as given in (4.42), Newton's method converges to $\bar{\vartheta}$ with quadratic rate. In particular, the convergence to the unique optimal solution $(\bar{x}; \bar{y}; \bar{s})$ is quadratic.

Proof. By Lemmas 4.2.6 and 4.2.7, the conditions of Theorem A.4.1 hold, and we get

$$r_n := \frac{1}{4\sqrt{2}\theta_1}.$$

Therefore, the Newton steps are well-defined in the neighborhood $B_{r_n}(\bar{\vartheta})$, and the convergence of Newton's method to $\bar{\vartheta}$ is quadratic if $\vartheta^{(0)} \in B_{r_n}(\bar{\vartheta})$. The quadratic convergence to $(\bar{x}; \bar{y}; \bar{s})$ follows from (4.30) and (4.32). Using the bounds in Theorem 3.2.1 and (4.39) we get

$$\|v^{(0)} - \bar{v}\|_2 = \sqrt{\frac{1}{4}\sum_{i \in \mathcal{R}}\left(\frac{(x^\mu)_1^i}{(s^\mu)_1^i} - \frac{\bar{x}_1^i}{\bar{s}_1^i}\right)^2} \leq \frac{2p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right).$$

Then, considering the error bounds given in (3.44), we obtain

$$\begin{aligned}\|\vartheta^{(0)} - \bar{\vartheta}\|_2 &\le \|(w^{(0)} - \bar{w}; z^{(0)} - \bar{z}; u^{(0)} - \bar{u})\|_2 + \|v^{(0)} - \bar{v}\|_2 \\ &\le \|(x^\mu - x^*; y^\mu - y^*; s^\mu - s^*)\|_2 + \|v^{(0)} - \bar{v}\|_2 \\ &\le \sqrt{3}\kappa(p\mu)^\gamma + \frac{2p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\Big(1 + \frac{2\sigma_3}{\sigma_2}\Big),\end{aligned}$$

where $(x^*; y^*; s^*)$ is the unique optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). The result of the theorem follows if we satisfy

$$\sqrt{3}\kappa(p\mu)^\gamma + \frac{2p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\Big(1 + \frac{2\sigma_3}{\sigma_2}\Big) < r_n,$$

or equivalently,

$$(p\mu)^\gamma < \frac{r_n}{\kappa\Big(\sqrt{3} + \frac{2p\sqrt{|\mathcal{R}|}}{\sigma_2}\big(1 + \frac{2\sigma_3}{\sigma_2}\big)\Big)}.$$

This completes the proof. □

Recall that $(\bar{x}; \bar{y}; \bar{s})$ is the unique optimal solution for ($\mathrm{P'_{SOCO}}$) and ($\mathrm{D'_{SOCO}}$). If $\mathcal{T}_1, \mathcal{T}_3 \neq \emptyset$, then we can recover the unique optimal solutions of the original problems ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) by appending $\mathcal{T}_1$ and $\mathcal{T}_3$ so that

$$\begin{aligned}(s^*)^i &:= c^i - A_i^T \bar{y}, && i \in \mathcal{T}_3, \\ (s^*)^i &:= 0, && i \in \mathcal{T}_1, \\ (x^*)^i &:= 0, && i \in \mathcal{T}_1 \cup \mathcal{T}_3.\end{aligned}$$

Remark 4.2.3. *Recall that Theorem 3.2.1 can be extended to the case when IPMs generate approximate solutions in a neighborhood of the central path. Hence, an upper bound analogous to* (4.43) *can be derived for approximate solutions in a neighborhood of the central path.*

4.2.4 Quadratic convergence under strict complementarity

When the strict complementarity condition holds in addition to the primal and dual nondegeneracy conditions, the quadratic convergence of Newton's method follows from Theorem 28 in [5]. In this case, a stronger complexity bound can be obtained in order to identify the quadratic convergence region. Note that the optimality conditions (1.10) can be written as $F_{\rm SO}((x;y;s)) = 0$ and $x, s \in \mathcal{L}_+^{\bar{n}}$, where the mapping

$$F_{\rm SO} : \mathbb{R}^{\bar{n}} \times \mathbb{R}^m \times \mathbb{R}^{\bar{n}} \rightarrow \mathbb{R}^m \times \mathbb{R}^{\bar{n}} \times \mathbb{R}^{\bar{n}}$$

is given by

$$F_{\rm SO}((x;y;s)) := \begin{pmatrix} Ax - b \\ A^T y + s - c \\ x \circ s \end{pmatrix}. \tag{4.44}$$

The Jacobian of $F_{\rm SO}$ is given by

$$\nabla F_{\rm SO}((x;y;s)) := \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L(s) & 0 & L(x) \end{pmatrix}, \tag{4.45}$$

where

$$L(x) := \operatorname{diag}(L(x^1), \ldots, L(x^p)),$$
$$L(s) := \operatorname{diag}(L(s^1), \ldots, L(s^p)).$$

The following technical lemma is in order.

Lemma 4.2.9 (Theorem 3.1 in [5] and [73]). *The Jacobian $\nabla F_{\rm SO}((x^*;y^*;s^*))$ is nonsingular if and only if the optimal solution $(x^*;y^*;s^*)$ satisfies strict complementarity, primal nondegeneracy, and dual nondegeneracy conditions.*

By Lemma 4.2.9, $\nabla F_{\mathrm{SO}}((x^*;y^*;s^*))$ is nonsingular, where $(x^*;y^*;s^*)$ is the unique optimal solution. Furthermore, analogous to Lemma 4.2.7, we can show that ∇F_{SO} is Lipschitz continuous with global Lipschitz constant $\tau_2 := 2$.

Lemma 4.2.10. *The Jacobian ∇F_{SO} is Lipschitz continuous with global Lipschitz constant $\tau_2 := 2$.*

Proof. Let $\xi := (\xi^1;\xi^2;\xi^3) \in \mathbb{R}^{\bar{n}} \times \mathbb{R}^m \times \mathbb{R}^{\bar{n}}$. Then from (1.8) and (4.45) we have

$$\begin{aligned}
&\|\nabla F_{\mathrm{SO}}((x;y;s)) - \nabla F_{\mathrm{SO}}((x';y';s'))\|_2 \\
&\qquad = \max_{\|\xi\|_2=1} \|\big(\nabla F_{\mathrm{SO}}((x;y;s)) - \nabla F_{\mathrm{SO}}((x';y';s'))\big)\xi\|_2 \\
&\qquad = \max_{\|\xi\|_2=1} \|L(s-s')\xi^1 + L(x-x')\xi^3\|_2 \\
&\qquad \le \max_{\|\xi\|_2=1} \|L(s-s')\xi^1\|_2 + \max_{\|\xi\|_2=1} \|L(x-x')\xi^3\|_2 \\
&\qquad \le \max_{\|\xi^1\|_2=1} \|L(s-s')\xi^1\|_2 + \max_{\|\xi^3\|_2=1} \|L(x-x')\xi^3\|_2 \\
&\qquad \le \|L(s-s')\|_2 + \|L(x-x')\|_2,
\end{aligned}$$

where $L(s-s')$ and $L(x-x')$ are block diagonal symmetric matrices. Then from Theorem 3 in [5] and the definition of the spectral norm we get

$$\begin{aligned}
\|L(s-s')\|_2 &= \max_{i=1,\dots,p} \max_{j=1,\dots,n} |\lambda_j(L(s^i-(s')^i))| \\
&\le \max_{i=1,\dots,p} \left|s_1^i - (s')_1^i\right| + \|s^i_{2:n_i} - (s')^i_{2:n_i}\|_2 \\
&\le \max_{i=1,\dots,p} \sqrt{2}\|s^i - (s')^i\|_2 \le \sqrt{2}\|s-s'\|_2.
\end{aligned}$$

The case for $\|L(x-x')\|_2$ is similar. Consequently, we get

$$\begin{aligned}
\|\nabla F_{\mathrm{SO}}((x;y;s)) - \nabla F_{\mathrm{SO}}((x';y';s'))\|_2 &\le \sqrt{2}\|s-s'\|_2 + \sqrt{2}\|x-x'\|_2 \\
&\le 2\|(x-x';y-y';s-s')\|_2,
\end{aligned}$$

which completes the proof. $\square$

Then the following result is immediate from Theorem A.4.1.

Theorem 4.2.3. *Assume that there exists $\theta_2 > 0$ so that*

$$\|\nabla F_{\mathrm{SO}}((x^*; y^*; s^*))^{-1}\|_2 \leq \theta_2.$$

Let $\hat{\mu}$ be defined by (3.43) and a central solution $(x^\mu; y^\mu; s^\mu)$ with

$$\mu < \min\left\{p^{-1}\left(4\sqrt{3}\theta_2\kappa\right)^{-\frac{1}{\gamma}},\ \hat{\mu}\right\} \tag{4.46}$$

be given, where κ and γ are defined as in (3.44). Then starting from a central solution $(x^\mu; y^\mu; s^\mu)$, Newton's method is quadratically convergent to $(x^*; y^*; s^*)$.

Proof. Since F is continuously differentiable, the result of Theorem A.4.1 is valid. Hence, Newton steps are well-defined in a neighborhood of $(x^*; y^*; s^*)$. Additionally, from Lemma 3.2.1 there exist positive κ and γ so that

$$\|(x^\mu - x^*; y^\mu - y^*; s^\mu - s^*)\|_2 \leq \sqrt{3}\kappa(p\mu)^\gamma.$$

Then it is immediate from (A.25) that $(x^\mu; y^\mu; s^\mu)$ is in the quadratic convergence region of Newton's method if

$$\sqrt{3}\kappa(p\mu)^\gamma < \frac{1}{4\theta_2},$$

which yields the result. $\square$

Remark 4.2.4. *Bound (4.43), relying on the condition numbers σ_1, σ_2, σ_3, and κ, is significantly more complicated than (4.46). In fact, the intricacy of bound (4.43) indicates that quadratic convergence is harder to achieve in the absence of strict complementarity. To that end, μ has to be small enough so that the optimal partition can be identified.*

4.2.5 Numerical results

We demonstrate quadratic convergence of Newton's method, applied to the first-order optimality conditions of $(\mathrm{D_{NLO}})$, on some instances of SOCO problems. For the first part of numerical experiments we solve (4.26) and the following SOCO problem

$$
\begin{array}{rr}
\min & -\frac{1}{2}x_2^1 - \frac{1}{2}x_3^1 \\
\text{s.t.} & x_1^1 = 1, \\
 & x_1^2 - x_4^1 = 1, \\
 & x_2^2 - x_2^1 = 0, \\
 & x_3^2 - x_3^1 = 0, \\
 & x_3^1 - x_4^1 - x_1^3 = -1, \\
 & x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2 + (x_4^1)^2}, \\
 & x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}, \\
 & x_1^3 \geq 0.
\end{array}
\tag{4.47}
$$

The SOCO problem (4.47) has the unique primal-dual optimal solution

$$
\begin{aligned}
x^* &= (1,\ 1/\sqrt{2},\ 1/\sqrt{2},\ 0,\ 1,\ 1/\sqrt{2},\ 1/\sqrt{2},\ 1+1/\sqrt{2})^T, \\
y^* &= (-1/\sqrt{2},\ 0,\ 0,\ 0,\ 0)^T, \\
s^* &= (1/\sqrt{2},\ -1/2,\ -1/2,\ 0,\ 0,\ 0,\ 0,\ 0)^T,
\end{aligned}
$$

and its optimal partition is given by

$$
\mathcal{B} = \{3\}, \quad \mathcal{R} = \{1\}, \quad \mathcal{T}_2 = \{2\}, \quad \mathcal{N} = \mathcal{T}_1 = \mathcal{T}_3 = \emptyset.
$$

For the second part of this section we generate a set of 10 random SOCO problems which fail the strict complementarity condition, but satisfy both the primal and dual nondegeneracy conditions, see [179] for degenerate random SDO problems. Specifically, we generate random problems with a unique primal-dual optimal solution. We choose m and the optimal partition $(\mathcal{B}, \mathcal{N}, \mathcal{R}, \mathcal{T})$, see Table 4.1, in such a way that the necessary conditions (4.21) are satisfied. Then the interior point condition

Table 4.1: The optimal partition and dimension of random problems.

Prob.	m	p	$(n_1,\ldots,n_p)$	$\mathcal{B}$	$\mathcal{N}$	$\mathcal{R}$	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_3$
1	7	4	$(3,6,2,2)$	$\{1\}$	$\emptyset$	$\{2\}$	$\{3\}$	$\emptyset$	$\{4\}$
2	6	2	$(3,5)$	$\emptyset$	$\emptyset$	$\{2\}$	$\emptyset$	$\{1\}$	$\emptyset$
3	6	3	$(3,6,2)$	$\{1\}$	$\emptyset$	$\{2\}$	$\emptyset$	$\{3\}$	$\emptyset$
4	4	2	$(5,3)$	$\emptyset$	$\emptyset$	$\{1\}$	$\{2\}$	$\emptyset$	$\emptyset$
5	9	5	$(5,6,4,2,3)$	$\{3\}$	$\{2\}$	$\{1,5\}$	$\emptyset$	$\{4\}$	$\emptyset$
6	11	6	$(5,6,5,2,3,2)$	$\{1\}$	$\{3\}$	$\{2,6\}$	$\{5\}$	$\emptyset$	$\{4\}$
7	7	5	$(3,9,3,3,4)$	$\emptyset$	$\{1\}$	$\{2,5\}$	$\emptyset$	$\{4\}$	$\{3\}$
8	18	6	$(10,5,7,8,2,8)$	$\{6\}$	$\emptyset$	$\{1,3,4\}$	$\emptyset$	$\{2,5\}$	$\emptyset$
9	22	7	$(10,5,7,8,2,8,5)$	$\emptyset$	$\emptyset$	$\{1,3,4\}$	$\{6\}$	$\{2,5\}$	$\{7\}$
10	35	7	$(8,8,8,8,8,8,8)$	$\{1,4\}$	$\emptyset$	$\{2,5\}$	$\{7\}$	$\{6\}$	$\{3\}$

automatically holds by the uniqueness of the optimal solution, see Theorem 5.81 in [23].

For the random problems we generate the unique optimal solution $(x^*;y^*;s^*)$ as follows

$$
\begin{aligned}
(x^*)^i_1 &\sim U\big(\|(x^*)^i_{2:n_i}\|_2+0.1,\ \|(x^*)^i_{2:n_i}\|_2+100.1\big), && i\in\mathcal{B},\\
(s^*)^i_1 &\sim U\big(\|(s^*)^i_{2:n_i}\|_2+0.1,\ \|(s^*)^i_{2:n_i}\|_2+100.1\big), && i\in\mathcal{N},\\
(x^*)^i &\sim U\big(0.1,100.1\big)\times\big(1;(\varrho^*)^i/\|(\varrho^*)^i\|_2\big), && i\in\mathcal{R}\cup\mathcal{T}_2,\\
(s^*)^i &\sim U\big(0.1,100.1\big)\times\big(1;-(\varrho^*)^i/\|(\varrho^*)^i\|_2\big), && i\in\mathcal{R}\cup\mathcal{T}_3,\\
y^*_i &\sim U(-100,100) && i=1,\ldots,m,
\end{aligned}
$$

where

$$
\begin{aligned}
(x^*)^i_j &\sim U(-100,100), && j=2,\ldots,n_i, && i\in\mathcal{B},\\
(s^*)^i_j &\sim U(-100,100), && j=2,\ldots,n_i, && i\in\mathcal{N},\\
(\varrho^*)^i_j &\sim U(-100,100), && j=1,\ldots,n_i-1, && i\in\mathcal{R}\cup\mathcal{T}_2\cup\mathcal{T}_3,
\end{aligned}
$$

in which $U(.,.)$ denotes the uniform distribution. For the rest of the variables we have $(x^*)^i=0$ for $i\in\mathcal{N}\cup\mathcal{T}_1\cup\mathcal{T}_3$ and $(s^*)^i=0$ for $i\in\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2$ by definition.

We then normalize the vectors by

$$x^* := \frac{x^*}{\|x^*\|_2}, \quad y^* := \frac{y^*}{\|y^*\|_2}, \quad s^* := \frac{s^*}{\|s^*\|_2}.$$

Analogous to the optimal solution, the entries of A are uniformly distributed in $(-100, 100)$. However, we keep generating random A until the rank condition and the nondegeneracy conditions in Theorem 4.2.1 hold[4]. We then compute $A := A/\|A\|_2$ and generate the right hand side and objective vectors by $b := Ax^*$ and $c := A^T y^* + s^*$.

We solve all the SOCO problems by using SeDuMi 1.3 included in the CVX optimization package [64, 65] and applying Newton's method to (4.36) and (4.44) throughout this section. The codes are run in MATLAB 9.2 environment on a MacBook Pro with Intel Core i5 CPU @ 2.3 GHz and 8GB of RAM. The Newton based approaches are referred to as NLO-Newton and SOCO-Newton, respectively. To solve (4.26) and (4.47) by Newton based approaches, we choose a central solution with $\mu = 10^{-2}$ as the initial point. However, we choose smaller values of μ for the random problems, as specified in Table 4.8. For the Newton based approaches $\|F_{\mathrm{SO}}((x^{(k)}; y^{(k)}; s^{(k)}))\|_2 \leq 10^{-14}$ is set as the terminating condition, and the optimality tolerance for SeDuMi is fixed at 10^{-15}. For NLO-Newton we remove the rows and columns associated with $\mathcal{T}_1$ and $\mathcal{T}_3$ from both the primal and dual problems. Additionally, we assign $(x^\mu; y^\mu; s^\mu)$ to $\vartheta^{(0)}$ according to (4.42), and we form the solution $(x^{(k)}; y^{(k)}; s^{(k)})$ by setting

$$\begin{aligned}
(x^i)^{(k)} &= -(u^i)^{(k)}, && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N},\\
y^{(k)} &= w^{(k)}, && \\
(s^i)^{(k)} &= (z^i)^{(k)}, && i \in \mathcal{R} \cup \mathcal{N},\\
(s^i)^{(k)} &= 0, && i \in \mathcal{B} \cup \mathcal{T}_2,
\end{aligned}$$

[4] Since the full row rank condition and the nondegeneracy conditions hold generically, see e.g., [6], the expected number of iterations to get the desired coefficient matrix is one.

Table 4.2: The numerical results of NLO-Newton on SOCO problem (4.26).

k	$\lVert Ax^{(k)} - b \rVert_2$	$c^T x^{(k)}$	$\lVert x^{(k)} \circ s^{(k)} \rVert_2$	$\lVert G(.) \rVert_2$	$\lVert F(.) \rVert_2$
0	5.551115E-17	-9.922138E-01	1.000000E-02	9.420176E-02	9.432417E-02
1	5.551115E-17	-1.004244E+00	1.132375E-02	2.952235E-02	1.132375E-02
2	0.000000E+00	-1.000041E+00	5.817148E-05	4.551179E-05	5.817148E-05
3	0.000000E+00	-1.000000E+00	2.149505E-10	1.600968E-10	2.149505E-10
4	0.000000E+00	-1.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00

Table 4.3: The numerical results of SOCO-Newton on SOCO problem (4.26).

k	$\lVert Ax^{(k)} - b \rVert_2$	$c^T x^{(k)}$	$\lVert x^{(k)} \circ s^{(k)} \rVert_2$	$\sigma_{\min}(\nabla F(.))$	$\lVert F(.) \rVert_2$
15	0.000000E+00	-1.000000E+00	1.540356E-11	1.270995E-06	1.540356E-11
16	0.000000E+00	-1.000000E+00	3.850891E-12	6.354972E-07	3.850891E-12
17	2.220446E-16	-1.000000E+00	9.627227E-13	3.177485E-07	9.627227E-13
18	2.220446E-16	-1.000000E+00	2.406807E-13	1.588743E-07	2.406808E-13
19	0.000000E+00	-1.000000E+00	6.017017E-14	7.943713E-08	6.017017E-14
20	0.000000E+00	-1.000000E+00	1.504254E-14	3.971856E-08	1.504254E-14
21	2.220446E-16	-1.000000E+00	3.760636E-15	1.985928E-08	3.768821E-15

since there is no s^i corresponding to $\mathcal{B}$ and $\mathcal{T}_2$ in (4.36).

Tables 4.2 to 4.4 illustrate the numerical results of the Newton based approaches on SOCO problem (4.26). For NLO-Newton we report both the Newton residuals $\|G(\vartheta^{(k)})\|_2$ and $\|F_{\mathrm{SO}}((x^{(k)}; y^{(k)}; s^{(k)}))\|_2$. NLO-Newton meets the stopping condition in only 4 iterations while this number is 21 for SOCO-Newton. As can be observed from Tables 4.2 and 4.3, the convergence of NLO-Newton to the unique optimal solution of (4.26) is quadratic while the convergence for SOCO-Newton is no better than linear. Additionally, SeDuMi arrives at the Newton residual 1.445657×10^{-12}, and in that sense it is less accurate than NLO-Newton and SOCO-Newton.

The numerical results of Newton based approaches on SOCO problem (4.47) are summarized in Tables 4.5 to 4.7. From Tables 4.5 and 4.6 we can observe the

Table 4.4: The k^{th} iterate of NLO-Newton and SOCO-Newton on SOCO problem (4.26).

NLO-Newton			SOCO-Newton		
k	$(x_3^1)^{(k)}$	$(x_3^2)^{(k)}$	k	$(x_3^1)^{(k)}$	$(x_3^2)^{(k)}$
0	-6.410764E-02	-1.282153E-01	17	-5.387688E-07	-1.077538E-06
1	8.783536E-03	1.756707E-02	18	-2.693844E-07	-5.387688E-07
2	0.000000E+00	0.000000E+00	19	-1.346922E-07	-2.693844E-07
3	0.000000E+00	0.000000E+00	20	-6.734610E-08	-1.346922E-07
4	0.000000E+00	0.000000E+00	21	-3.367305E-08	-6.734610E-08

Table 4.5: The numerical results of NLO-Newton on SOCO problem (4.47).

k	$\left\|Ax^{(k)} - b\right\|_2$	$c^T x^{(k)}$	$\left\|x^{(k)} \circ s^{(k)}\right\|_2$	$\|G(.)\|_2$	$\|F(.)\|_2$
0	1.110223E-16	-6.995520E-01	1.000000E-02	7.979631E-02	8.014507E-02
1	2.220446E-16	-7.102731E-01	6.634877E-03	8.185201E-03	6.634877E-03
2	3.330669E-16	-7.071020E-01	9.464771E-06	6.556691E-06	9.464771E-06
3	0.000000E+00	-7.071068E-01	6.959588E-11	4.962837E-11	6.959588E-11
4	0.000000E+00	-7.071068E-01	1.110223E-16	1.110223E-16	1.110223E-16

quadratic convergence of NLO-Newton, versus linear convergence of SOCO-Newton. Furthermore, Table 4.7 confirms that NLO-Newton evolves faster toward the unique optimal solution of (4.47) than SOCO-Newton. NLO-Newton arrives at the Newton residual 1.110223×10^{-16} in only 4 iterations, while SeDuMi ends up with the Newton residual 1.777953×10^{-9}.

We draw a sample of 100 instances for each random SOCO problem and report the average results in Tables 4.8 to 4.10. To ensure convergence of NLO-Newton to the unique optimal solution, we choose initial solutions with sufficiently small μ when solving the random problems. Table 4.8 reports the values of μ as well as the distance of the initial solution $\omega^{(0)} := (x^{(0)}; y^{(0)}; s^{(0)})$ and the SeDuMi's solution from the unique optimal solution $\omega^* := (x^*; y^*; s^*)$, where ω^*_{se}, ω^*_{nn}, and ω^*_{sn} stand for the solution output by SeDuMi, NLO-Newton, and SOCO-Newton. For NLO-Newton

Table 4.6: The numerical results of SOCO-Newton on SOCO problem (4.47).

k	$\left\|Ax^{(k)} - b\right\|_2$	$c^T x^{(k)}$	$\left\|x^{(k)} \circ s^{(k)}\right\|_2$	$\sigma_{\min}(\nabla F(.))$	$\|F(.)\|_2$
15	1.110223E-16	-7.071068E-01	9.289517E-12	1.937402E-06	9.289517E-12
16	1.110223E-16	-7.071068E-01	2.322376E-12	9.687002E-07	2.322376E-12
17	1.110223E-16	-7.071068E-01	5.805962E-13	4.843499E-07	5.805962E-13
18	0.000000E+00	-7.071068E-01	1.452016E-13	2.421749E-07	1.452016E-13
19	2.482534E-16	-7.071068E-01	3.629592E-14	1.210874E-07	3.629677E-14
20	0.000000E+00	-7.071068E-01	9.118149E-15	6.054371E-08	9.118149E-15

Table 4.7: The k^{th} iterate of NLO-Newton and SOCO-Newton on SOCO problem (4.47).

NLO-Newton			SOCO-Newton		
k	$(x_4^1)^{(k)}$	$(x_1^3)^{(k)}$	k	$(x_4^1)^{(k)}$	$(x_1^3)^{(k)}$
0	7.86081643E-02	1.62522185E+00	16	1.28146875E-06	1.70710550E+00
1	-4.67865604E-03	1.71469713E+00	17	6.40734373E-07	1.70710614E+00
2	0.00000000E+00	1.70710195E+00	18	3.20367187E-07	1.70710646E+00
3	0.00000000E+00	1.70710678E+00	19	1.60183593E-07	1.70710662E+00
4	0.00000000E+00	1.70710678E+00	20	8.00917970E-08	1.70710670E+00

the solution ω^*_{nn} is formed by $(x^*_{nn}; y^*_{nn}; s^*_{nn})$, where

$$\begin{aligned}
(x^*_{nn})^i &= -(u^i)^{(\bar{k})}, && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}, \\
(x^*_{nn})^i &= 0, && i \in \mathcal{T}_1 \cup \mathcal{T}_3, \\
y^*_{nn} &= w^{(\bar{k})}, && \\
(s^*_{nn})^i &= (z^i)^{(\bar{k})}, && i \in \mathcal{R} \cup \mathcal{N}, \\
(s^*_{nn})^i &= c^i - A_i^T w^{(\bar{k})}, && i \in \mathcal{T}_3, \\
(s^*_{nn})^i &= 0, && i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2,
\end{aligned}$$

where $\bar{k}$ denotes the index of final iterate.

Tables 4.9 and 4.10 demonstrate the numerical results of Newton based approaches on the random SOCO problems. The values of $\|\omega^*_{nn} - \omega^*\|_2$, $\|\omega^*_{sn} - \omega^*\|_2$, and "Iter" indicate the accuracy and fast convergence of NLO-Newton in comparison to

Table 4.8: The initial and optimal solutions for random SOCO problems.

Prob.	μ	$\lVert\omega^{(0)} - \omega^*\rVert_2$	$\lVert\omega^*_{se} - \omega^*\rVert_2$
1	1.00E-05	1.912776E-02	1.592721E-06
2	1.00E-05	9.125473E-02	6.738418E-06
3	1.00E-05	1.643906E-02	2.662254E-06
4	1.00E-05	1.795054E-02	1.607865E-06
5	1.00E-05	2.629707E-02	2.755846E-07
6	1.00E-05	4.050471E-02	3.003945E-05
7	1.00E-05	8.766428E-03	4.772793E-07
8	1.00E-06	1.555120E-02	3.128175E-06
9	1.00E-06	1.049691E-02	6.437721E-07
10	1.00E-06	2.267326E-02	1.741151E-06

Table 4.9: The numerical results of NLO-Newton on random SOCO problems.

Prob.	Iter	$\lVert F_{\mathrm{SO}}(\omega^{(0)})\rVert_2$	$\lVert F_{\mathrm{SO}}(\omega^*_{nn})\rVert_2$	$\lVert\omega^*_{nn} - \omega^*_{se}\rVert_2$	$\lVert\omega^*_{nn} - \omega^*\rVert_2$	$\lVert\omega^*_{nn} - \omega^*_{sn}\rVert_2$
1	3	1.958945E-03	2.059106E-16	1.592721E-06	2.021790E-13	7.452722E-07
2	3	1.888355E-02	2.589828E-16	6.738418E-06	1.390557E-14	1.056285E-06
3	3	8.310096E-04	1.908819E-16	2.662245E-06	8.634383E-12	5.002596E-07
4	3	7.542060E-04	2.555563E-16	6.078650E-07	1.555339E-13	4.794085E-07
5	4	2.494958E-03	5.557528E-16	2.755846E-07	1.975277E-14	6.050262E-07
6	4	1.962349E-03	6.526095E-16	3.003945E-05	2.171982E-12	2.831458E-04
7	4	1.149507E-03	6.627321E-16	4.772793E-07	1.107431E-14	4.519775E-07
8	3	1.088135E-04	6.271617E-16	3.128175E-06	7.187394E-13	9.655715E-07
9	4	1.720883E-03	3.192262E-16	6.437721E-07	7.164252E-14	4.071836E-07
10	4	1.960026E-03	5.365570E-16	1.741151E-06	8.266871E-13	8.583896E-07

SOCO-Newton.

4.2.6 Special case: a strongly polynomial rounding procedure

In a special case when the sets $\mathcal{R}$ and $\mathcal{T}$ are empty, a strictly complementary optimal solution can be obtained as easily as in LO [109, 180], regardless of the nondegeneracy conditions. More precisely, an interior solution (x^μ, y^μ, s^μ), with sufficiently small μ, can be rounded to an exact strictly complementary optimal solution in

Table 4.10: The numerical results of SOCO-Newton on random SOCO problems.

Prob.	Iter	$\|F_{\mathrm{SO}}(\omega^{(0)})\|_2$	$\|F_{\mathrm{SO}}(\omega^*_{sn})\|_2$	$\|\nabla F_{\mathrm{SO}}(\omega^*_{sn})\|_2$	$\|\omega^*_{sn}-\omega_{se}\|_2$	$\|\omega^*_{sn}-\omega^*\|_2$
1	15	1.996282E-06	6.229178E-15	9.947566E-09	1.226415E-06	7.452723E-07
2	16	1.414214E-05	5.999936E-15	1.413926E-08	5.774589E-06	1.056285E-06
3	14	1.732051E-06	4.582760E-15	1.458812E-08	2.463535E-06	5.002683E-07
4	14	1.414214E-06	6.101894E-15	1.575818E-08	1.226030E-06	4.794086E-07
5	15	2.236068E-05	6.644257E-15	1.941088E-08	3.546504E-07	6.050262E-07
6	14	2.449490E-06	6.379693E-15	8.065773E-09	3.108220E-04	2.831458E-04
7	14	2.236068E-06	6.201575E-15	1.642138E-08	5.217249E-07	4.519775E-07
8	13	2.449491E-07	5.319422E-15	4.816730E-09	2.633089E-06	9.655719E-07
9	15	2.645751E-06	4.950853E-15	9.779408E-09	3.668787E-07	4.071836E-07
10	15	2.645751E-06	5.222010E-15	4.778974E-09	1.179792E-06	8.583897E-07

strongly polynomial time through solving two least squares problems.

Let $(x^*; y^*; s^*) \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SOCO}} \times \mathcal{D}^*_{\mathrm{SOCO}})$ be a strictly complementary optimal solution of $(\mathrm{P}_{\mathrm{SOCO}})$ and $(\mathrm{D}_{\mathrm{SOCO}})$. Then the primal-dual feasibility constraints imply

$$\sum_{i\in\mathcal{B}} A_i (x^i)^* = b, \qquad \sum_{i\in\mathcal{B}\cup\mathcal{N}} A_i (x^\mu)^i = b,$$

and

$$\begin{aligned} A_i^T y^* &= c_i, \qquad & A_i^T y^\mu + (s^\mu)^i &= c_i, \qquad & i &\in \mathcal{B},\\ A_i^T y^* + (s^i)^* &= c_i, \qquad & A_i^T y^\mu + (s^\mu)^i &= c_i, \qquad & i &\in \mathcal{N}. \end{aligned}$$

Subtracting the right hand side equations from the left hand side ones we get

$$\begin{aligned} \sum_{i\in\mathcal{B}} A_i (\Delta x^\mu)^i &= \sum_{i\in\mathcal{N}} A_i (x^\mu)^i, & &\\ A_i^T \Delta y^\mu &= (s^\mu)^i, & i &\in \mathcal{B},\\ A_i^T \Delta y^\mu + (\Delta s^\mu)^i &= 0, & i &\in \mathcal{N}, \end{aligned}$$

where $\Delta y^\mu := y^* - y^\mu$, $(\Delta x^\mu)^i := (x^i)^* - (x^\mu)^i$ for $i \in \mathcal{B}$, and $(\Delta s^\mu)^i := (s^i)^* - (s^\mu)^i$ for $i \in \mathcal{N}$. Thus, we get a primal-dual solution with zero complementarity gap by solving the least squares problem

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{i\in\mathcal{B}} \|\Delta x^i\|_2^2 \\ \text{s.t.} \quad & \sum_{i\in\mathcal{B}} A_i \Delta x^i = \sum_{i\in\mathcal{N}} A_i (x^\mu)^i, \end{aligned} \tag{4.48}$$

for the primal solution, and the following least square problem

$$\begin{aligned} \min \quad & \frac{1}{2}\|\Delta y\|_2^2 + \frac{1}{2}\sum_{i\in\mathcal{N}} \|\Delta s^i\|_2^2 \\ \text{s.t.} \quad & A_i^T \Delta y = (s^\mu)^i, \quad i \in \mathcal{B}, \\ & A_i^T \Delta y + \Delta s^i = 0, \quad i \in \mathcal{N}, \end{aligned}$$

for the dual solution, which is equivalent to

$$\begin{aligned} \min \quad & \frac{1}{2}\|\Delta y\|_2^2 + \frac{1}{2}\sum_{i\in\mathcal{N}} \|A_i^T \Delta y\|_2^2 \\ \text{s.t.} \quad & A_i^T \Delta y = (s^\mu)^i, \quad i \in \mathcal{B}. \end{aligned} \tag{4.49}$$

The least squares problems (4.48) and (4.49) yield the complementary primal-dual pair $(\tilde{\tilde{x}}; \tilde{\tilde{y}}; \tilde{\tilde{s}})$, where

$$\begin{aligned} \tilde{\tilde{x}}^i &:= (x^\mu)^i + (\Delta x^*)^i, && i \in \mathcal{B}, \\ \tilde{\tilde{y}} &:= y^\mu + \Delta y^*, \\ \tilde{\tilde{s}}^i &:= (s^\mu)^i + (\Delta s^*)^i, && i \in \mathcal{N}. \end{aligned}$$

It can be easily shown that $(\tilde{\tilde{x}}; \tilde{\tilde{y}}; \tilde{\tilde{s}})$ is feasible with respect to the primal and dual affine constraints. Further, the feasibility of $(\tilde{\tilde{x}}; \tilde{\tilde{y}}; \tilde{\tilde{s}})$ with respect to the second-order cones can be established when μ is sufficiently small, see e.g., [109].

4.2.7 Approximate maximally complementary solutions

An idea similar to the one in Section 4.1.1 can be used to generate an approximate maximally complementary solution. Generally speaking, if $\mathcal{R} \neq \emptyset$, an exact solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) cannot be obtained even with rational data. However, given a central solution (x^μ, y^μ, s^μ) with sufficiently small μ, we can make a projection onto the boundary of the second-order cone to generate a solution with zero complementary gap. Toward this end, we can fix the Jordan frame $\{(p^\mu)_1^i, (p^\mu)_2^i\}$ for $i \in \mathcal{R} \cup \mathcal{T}_2 \cup \mathcal{T}_3$, and solve for their corresponding eigenvalues. Let $(x^*; y^*; s^*) \in \mathrm{ri}(\mathcal{P}^*_{\mathrm{SOCO}} \times \mathcal{D}^*_{\mathrm{SOCO}})$ be a primal-dual maximally complementary optimal solution. Then by using primal-dual feasibility constraints we have

$$\begin{aligned}
\sum_{i\in\mathcal{B}} A_i(x^*)^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} (\lambda^*)_2^i A_i(p^*)_2^i &= b,\\
\sum_{i\in\mathcal{B}\cup\mathcal{N}\cup\mathcal{T}_1\cup\mathcal{T}_3} A_i(x^\mu)^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} ((\lambda^\mu)_1^i A_i(p^\mu)_1^i + (\lambda^\mu)_2^i A_i(p^\mu)_2^i) &= b,
\end{aligned}$$

where $\{(p^\mu)_1^i, (p^\mu)_2^i\}$ denotes the Jordan frame and $(\lambda^\mu)_1^i$ and $(\lambda^\mu)_2^i$ are their corresponding eigenvalues, see (A.1). Subtracting the second equation from the first one, we get

$$\begin{aligned}
&\sum_{i\in\mathcal{B}} A_i(\Delta x^\mu)^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} (\Delta\lambda^\mu)^i A_i(p^\mu)_2^i = \\
&\qquad\qquad \sum_{i\in\mathcal{N}\cup\mathcal{T}_1\cup\mathcal{T}_3} A_i(x^\mu)^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} (\lambda^\mu)_1^i A_i(p^\mu)_1^i + \xi_p,
\end{aligned}$$

where

$$\begin{aligned}
(\Delta x^\mu)^i &= (x^*)^i - (x^\mu)^i, && i \in \mathcal{B},\\
(\Delta\lambda^\mu)^i &= (\lambda^*)_2^i - (\lambda^\mu)_2^i, && i \in \mathcal{R}\cup\mathcal{T}_2,\\
\xi_p &= -\sum_{i\in\mathcal{R}\cup\mathcal{T}_2} (\lambda^*)_2^i A_i((p^*)_2^i - (p^\mu)_2^i).
\end{aligned}$$

The right hand side vector depends on the solutions belonging to $\mathcal{N}$, $\mathcal{T}_1$, and $\mathcal{T}_3$ and the eigenvalues of solutions in $\mathcal{R}$, $\mathcal{T}_3$, which converge to 0. Hence, the right hand side vector can be made arbitrary small as $\mu \to 0$. However, the residual term ξ_p is not known since it depends on the optimal solution x^*. We can drop the residual term ξ_p and solve the least square problem

$$\begin{aligned}
\min_{\Delta x, \Delta\lambda, e_p} \quad & \sum_{i\in\mathcal{B}} \|\Delta x^i\|_2^2 + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} \|\Delta\lambda^i\|_2^2 + \|\varepsilon_p\|_2^2 \\
\text{s.t.} \quad & \sum_{i\in\mathcal{B}} A_i \Delta x^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} \Delta\lambda^i A_i (p^\mu)_2^i = \\
& \qquad \sum_{i\in\mathcal{N}\cup\mathcal{T}_1\cup\mathcal{T}_3} A_i (x^\mu)^i + \sum_{i\in\mathcal{R}\cup\mathcal{T}_2} (\lambda^\mu)_1^i A_i (p^\mu)_1^i - \varepsilon_p,
\end{aligned} \tag{4.50}$$

which, due to disregarding ξ_p, gives the direction $(\Delta x^*, \Delta\lambda^*)$ towards a nearly primal feasible solution, i.e.,

$$\begin{aligned}
\tilde{\tilde{x}}^i &:= (x^\mu)^i + (\Delta x^*)^i, && i \in \mathcal{B}, \\
\tilde{\tilde{\lambda}}_2^i &:= (\lambda^\mu)_2^i + (\Delta\lambda^*)^i, && i \in \mathcal{R}.
\end{aligned}$$

Given the dual optimal solution $(y^*; s^*)$, the dual feasibility constraints imply

$$\begin{aligned}
A_i^T \Delta y^\mu &= (s^\mu)^i, && i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2, \\
A_i^T \Delta y^\mu + (\Delta\delta^\mu)^i (p^\mu)_1^i &= (\delta^\mu)_2^i (p^\mu)_2^i + \xi_d, && i \in \mathcal{R} \cup \mathcal{T}_3, \\
A_i^T \Delta y^\mu + (\Delta s^\mu)^i &= 0, && i \in N,
\end{aligned}$$

where

$$\begin{aligned}
\Delta y^\mu &= y^* - y^\mu, && \\
(\Delta s^\mu)^i &= (s^*)^i - (s^\mu)^i, && i \in \mathcal{N}, \\
(\Delta\delta^\mu)^i &= (\delta^*)_1^i - (\delta^\mu)_1^i, && i \in \mathcal{R} \cup \mathcal{T}_3, \\
\xi_d &= -(\delta^*)_1^i \big((p^*)_1^i - (p^\mu)_1^i\big). &&
\end{aligned}$$

Analogous to the primal case, ξ_d depends on the optimal solution $(y^*; s^*)$. Hence we can drop ξ_d and solve the following least square problem

$$
\begin{aligned}
\min_{\Delta y, \Delta s, \Delta\delta, \varepsilon_d} \quad & \sum_{i\in\mathcal{N}} \|\Delta s^i\|_2^2 + \sum_{i\in\mathcal{R}\cup\mathcal{T}_3} \|\Delta\delta^i\|_2^2 + \|\Delta y\|_2^2 + \|\varepsilon_d\|_2^2 \\
\text{s.t.} \quad & A_i^T \Delta y = (s^\mu)^i, && i \in \mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2, \\
& A_i^T \Delta y + \Delta\delta^i (p^\mu)_1^i = (\delta^\mu)_2^i (p^\mu)_2^i - \varepsilon_d, && i \in \mathcal{R}\cup\mathcal{T}_3, \\
& A_i^T \Delta y + \Delta s^i = 0, && i \in \mathcal{N},
\end{aligned}
\tag{4.51}
$$

which gives the direction $(\Delta y^*; \Delta s^*; \Delta\delta^*)$ toward a nearly dual feasible solution, i.e.,

$$
\begin{aligned}
\tilde{\tilde{y}} &:= y^\mu + \Delta y^*, && \\
\tilde{\tilde{s}}^i &:= (s^\mu)^i + (\Delta s^*)^i, && i \in \mathcal{N}, \\
\tilde{\tilde{\delta}}_1^i &:= (\delta^\mu)_1^i + (\Delta\delta^*)^i, && i \in \mathcal{R}\cup\mathcal{T}_3.
\end{aligned}
$$

It can be easily shown that the primal-dual pair $(\tilde{\tilde{x}}; \tilde{\tilde{y}}; \tilde{\tilde{s}})$ has zero complementary gap, and it is ε^*-feasible with respect to the primal-dual affine constraints, where

$$
\varepsilon^* = \max\{\varepsilon_p^*, \varepsilon_d^*\}. \tag{4.52}
$$

Analogous to Section 4.2.6, feasibility with respect to the second-order cones can be achieved if μ is sufficiently small, i.e., there exists a positive $\tilde{\mu}$ such that for $\mu < \tilde{\mu}$ we have

$$
\begin{aligned}
\tilde{\tilde{x}}^i &\in \operatorname{int}(\mathbb{L}_+^{n_i}), && i \in \mathcal{B}, \\
\tilde{\tilde{\lambda}}_2^i &> 0, && i \in \mathcal{R}\cup\mathcal{T}_2, \\
\tilde{\tilde{s}}^i &\in \operatorname{int}(\mathbb{L}_+^{n_i}), && i \in \mathcal{N}, \\
\tilde{\tilde{\delta}}_1^i &> 0, && i \in \mathcal{R}\cup\mathcal{T}_3,
\end{aligned}
$$

see [112] for details. Doing so, we get an approximate maximally complementary solution $(\tilde{\tilde{x}}; \tilde{\tilde{y}}; \tilde{\tilde{s}})$.

Chapter 5

Sensitivity and stability analysis

Steady advances in computational optimization have enabled us to solve a wide variety of LCO problems in polynomial time. Nevertheless, sensitivity analysis tools are still the missing parts of LCO solvers, e.g., IPMs in SeDuMi, SDPT3, and MOSEK. The sensitivity and stability analysis of convex optimization problems have been worked out in [147, 148]. Shapiro [158] established the differentiability of the optimal solution for a nonlinear SDO problem using the standard implicit function theorem. Bonnans et al. [20] proposed a second-order regularity condition for finite dimensional optimization problems and established Hölder and Lipschitz continuity results for approximate optimal solutions. They also applied the theory to SDO and semi-infinite optimization problems. Under linear perturbations in objective vector, coefficient matrix, and right hand side, Nayakkankuppam and Overton [122] derived the region of stability around an optimal solution of SDO which satisfies strict complementarity and nondegeneracy conditions. The sensitivity of central solutions for SDO was treated in [128, 170]. Yildirim [185] proposed a sensitivity analysis approach based on IPM for LO and SDO. Bonnans and Ramírez [21] characterized strongly regular KKT solutions for nonlinear SOCO problems, see also [40].

Recently, Cheung and Wolkowicz [27] and Sekiguchi and Waki [157] studied the continuity of optimal value function for SDO problems which fail interior point condition. A comprehensive study of directional and differential stability of NLO problems in abstract spaces was given by Bonnans and Shapiro [23]. The results are mostly exploring a small neighborhood of a locally optimal solution, and they depend on strict complementarity and strong second-order sufficiency conditions. We also refer the reader to [93] for extensive results for the regularity of optimal solutions for parametric optimization problems.

Adler and Monteiro [3] studied the parametric analysis of LO problems using the concept of the optimal partition. Another treatment of sensitivity analysis for LO based on the optimal partition approach was given by Jansen et al. [87] and Greenberg [66]. Berkelaar et al. [16] extended the optimal partition approach to linearly constrained quadratic optimization (LCQO) with perturbation in the right hand side vector and showed that the optimal value function is convex and piecewise quadratic. There have been some studies on optimal partition and parametric analysis of conic optimization problems. Goldfarb and Scheinberg [59] considered a parametric SDO problem, where the objective is perturbed along a fixed direction. They derived auxiliary problems to compute the directional derivatives of the optimal value function and the so called *linearity interval* of the optimal partition, see also [152]. The optimal value function for SDO has been shown to be piecewise algebraic [126], i.e., for each piece there exists a polynomial function $\psi(.,.)$ so that $\psi(\varphi(\epsilon),\epsilon) = 0$, see also Section 5.3 in [19]. Yildirim [184] extended the concept of the optimal partition and the auxiliary problems in [59] for LCO problems.

In this chapter, we study the parametric analysis of SDO and SOCO problems and then extend the results for a parametric LCO problem. We introduce the concepts of nonlinearity interval and transition point for the optimal partition and provide sufficient conditions to get a subinterval of a nonlinearity interval. We study the relationship between the continuity of the optimal set mapping and a nonlinearity interval. Furthermore, we investigate the sensitivity of the approximation of the optimal partition for SDO. For SOCO we define auxiliary problems to compute a subinterval of a nonlinearity interval. Further, we show how to use derivative information to identify a transition point.

5.1 Parametric analysis of SDO

We consider a parametric SDO problem, which can be phrased as

$$\begin{aligned}
&(\mathrm{P}^{\epsilon}_{\mathrm{SDO}}) \qquad \min\left\{\langle C+\epsilon\bar{C},X\rangle \mid \langle A^i,X\rangle=b_i,\quad i=1,\ldots,m,\ X\succeq 0\right\},\\
&(\mathrm{D}^{\epsilon}_{\mathrm{SDO}}) \qquad \max\left\{b^Ty \mid \sum_{i=1}^m y_iA^i+S=C+\epsilon\bar{C},\ S\succeq 0,\ y\in\mathbb{R}^m\right\}.
\end{aligned}$$

where $\bar{C}\in\mathbb{S}^n$ is a fixed direction. The primal and dual feasible set mappings are defined as

$$\begin{aligned}
\mathcal{P}_{\mathrm{SDO}}(\epsilon) &:= \left\{X \mid \langle A^i,X\rangle=b_i,\quad i=1,\ldots,m,\ X\succeq 0\right\},\\
\mathcal{D}_{\mathrm{SDO}}(\epsilon) &:= \left\{(y,S) \mid \sum_{i=1}^m y_iA^i+S=C+\epsilon\bar{C}\right\}.
\end{aligned}$$

Since $\mathcal{P}_{\mathrm{SDO}}(.)$ is independent of ϵ, primal feasibility either holds or fails for all ϵ. Thus, we assume that $(\mathrm{P}^{\epsilon}_{\mathrm{SDO}})$ is solvable for all ϵ.

The optimal value of $(\mathrm{P}^{\epsilon}_{\mathrm{SDO}})$ yields the optimal value function $\varphi:\mathbb{R}\to\mathbb{R}\cup\{-\infty,+\infty\}$. Let $\mathcal{E}\subseteq\mathbb{R}$ be the set of all ϵ for which $\varphi(\epsilon)>-\infty$, and assume that $\mathcal{E}\neq\emptyset$. Then it is proven that $\varphi(.)$ is a proper concave function [148], and that it is continuous on $\mathrm{int}(\mathcal{E})$. The continuity of the optimal value function at ϵ follows from the concavity of $\varphi(.)$, as a special case of the following lemma.

Lemma 5.1.1 (Corollary 2.109 in [23]). *The concave function $\varphi(.)$ is continuous at $\bar{\epsilon}$ if $\bar{\epsilon} \in \operatorname{int}(\mathcal{E})$.*

Furthermore, $\mathcal{E}$ is a closed, possibly unbounded, interval, see e.g., Lemma 2.2 in [16].

In order to guarantee strong duality and the existence of the optimal partition, the following assumption is made throughout this section:

Assumption 5.1.1. *The interior point condition holds for both* $(\mathrm{P}^{\epsilon'}_{\mathrm{SDO}})$ *and* $(\mathrm{D}^{\epsilon'}_{\mathrm{SDO}})$ *for all* $\epsilon' \in \operatorname{int}(\mathcal{E})$*, i.e., there exists a feasible solution* $\big(X^\circ(\epsilon'), y^\circ(\epsilon'), S^\circ(\epsilon')\big)$ *such that* $X^\circ(\epsilon'), S^\circ(\epsilon') \succ 0$.

Assumption 5.1.1 is actually needed for the existence of the central path and uniformly boundedness of the primal and dual optimal set mappings which are given by

$$
\begin{aligned}
\mathcal{P}^*_{\mathrm{SDO}}(\epsilon) &:= \big\{X \mid \langle C + \epsilon\bar{C}, X\rangle = \varphi(\epsilon),\ X \in \mathcal{P}(\epsilon)\big\},\\
\mathcal{D}^*_{\mathrm{SDO}}(\epsilon) &:= \big\{(y, S) \mid b^T y = \varphi(\epsilon),\ (y, S) \in \mathcal{D}(\epsilon)\big\}.
\end{aligned}
$$

Consequently, both $\mathcal{P}^*_{\mathrm{SDO}}(\epsilon')$ and $\mathcal{D}^*_{\mathrm{SDO}}(\epsilon')$ are nonempty and compact for all $\epsilon' \in \operatorname{int}(\mathcal{E})$. Furthermore, for every $\epsilon' \in \operatorname{int}(\mathcal{E})$ there exists a maximally complementary optimal solution $\big(X^*(\epsilon'), y^*(\epsilon'), S^*(\epsilon')\big)$.

5.1.1 Continuity of optimal solutions for SDO

In this section, we investigate the continuity of the primal and dual feasible set mappings and the outer semicontinuity of the primal and dual optimal set mappings for SDO. For the sake of completeness, we include short proofs for the main results. The proofs can be found in [10] and [82], see also Section 1.7.4.

Since the primal feasible set is invariant with respect to the perturbation, $\mathcal{P}_{\mathrm{SDO}}(.)$ is continuous at any ϵ. For the dual problem $(\mathrm{D}^{\epsilon}_{\mathrm{SDO}})$ we define the set-valued mapping

$\mathcal{D}^y_{\text{SDO}} : \mathbb{R} \to \mathbb{R}^m$ as

$$\mathcal{D}^y_{\text{SDO}}(\epsilon) := \left\{ y \in \mathbb{R}^m \mid C + \epsilon\bar{C} - \sum_{i=1}^m y_i A^i \succeq 0 \right\},$$

Since $\mathbb{S}^n_+$ is a closed convex cone, it can be shown that $\mathcal{D}^y_{\text{SDO}}(.)$ is outer semicontinuous at every $\bar{\epsilon} \in \text{int}(\mathcal{E})$.

Lemma 5.1.2. *The set-valued mapping $\mathcal{D}^y_{\text{SDO}}(.)$ is outer semicontinuous at $\bar{\epsilon}$.*

Proof. Let $\{\epsilon_k\} \subset \mathbb{R}$, $\epsilon_k \to \bar{\epsilon}$, $\{y_k\} \subset \mathbb{R}^m$, $y_k \in \mathcal{D}^y_{\text{SDO}}(\epsilon_k)$, and $y_k \to y$. Since $y_k \in \mathcal{D}^y_{\text{SDO}}(\epsilon_k)$ we have

$$C + \epsilon_k\bar{C} - \sum_{i=1}^m (y_k)_i A^i \succeq 0.$$

From the continuity of the affine constraint and the closedness of $\mathbb{S}^n_+$ we have

$$C + \epsilon_k\bar{C} - \sum_{i=1}^m (y_k)_i A^i \to C + \bar{\epsilon}\bar{C} - \sum_{i=1}^m y_i A^i \succeq 0,$$

which proves the outer semicontinuity of $\mathcal{D}^y_{\text{SDO}}(.)$ at $\bar{\epsilon}$. $\square$

Furthermore, we can show that $\mathcal{D}^y_{\text{SDO}}(.)$ and thus $\mathcal{D}_{\text{SDO}}(.)$ is inner semicontinuous at every $\bar{\epsilon} \in \text{int}(\mathcal{E})$.

Lemma 5.1.3. *Let $\bar{y} \in \mathcal{D}^y_{\text{SDO}}(\bar{\epsilon})$ such that $C + \bar{\epsilon}\bar{C} - \sum_{i=1}^m \bar{y}_i A^i \succ 0$. Then the set-valued mapping $\mathcal{D}^y_{\text{SDO}}(.)$ is inner semicontinuous at every $\epsilon \in \text{int}(\mathcal{E})$.*

Proof. Let $\epsilon_k \to \bar{\epsilon}$, $\hat{y} \in \mathcal{D}^y_{\text{SDO}}(\bar{\epsilon})$, and assume that $\hat{S} := C + \bar{\epsilon}\bar{C} - \sum_{i=1}^m \hat{y}_i A^i$ has at least one zero eigenvalue. We can construct a convergent sequence $\{y_k\}$ with $y_k \to \hat{y}$ such that $C + \epsilon_k\bar{C} - \sum_{i=1}^m (y_k)_i A^i \succeq 0$ for large values of k.

Consider the sequence $y_k = (1 - \alpha_k)\hat{y} + \alpha_k\bar{y}$. Note that

$$C + \epsilon_k\bar{C} - \sum_{i=1}^m (y_k)_i A^i = (1 - \alpha_k)\Big(C + \epsilon_k\bar{C} - \sum_{i=1}^m \hat{y}_i A^i\Big) + \alpha_k\Big(C + \epsilon_k\bar{C} - \sum_{i=1}^m \bar{y}_i A^i\Big).$$

Then $C+\epsilon_k\bar{C}-\sum_{i=1}^m (y_k)_i A^i \succeq 0$ holds if

$$(1-\alpha_k)\Big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \hat{y}_i A^i\Big)+\alpha_k\Big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \bar{y}_i A^i\Big)\succeq 0. \tag{5.1}$$

If $0\le\alpha_k\le 1$, then (5.1) holds if

$$(1-\alpha_k)\lambda_{\min}\Big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \hat{y}_i A^i\Big)+\alpha_k\lambda_{\min}\Big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \bar{y}_i A^i\Big)\ge 0,$$

which is equivalent to

$$\alpha_k \ge \frac{-\lambda_{\min}\big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \hat{y}_i A^i\big)}{\lambda_{\min}\big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \bar{y}_i A^i\big)-\lambda_{\min}\big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \hat{y}_i A^i\big)}$$

for large k, since the denominator has to be positive. Letting $\alpha_k:=\max\{\rho_k,0\}$, where

$$\rho_k := \frac{-\lambda_{\min}\big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \hat{y}_i A^i\big)}{\lambda_{\min}\big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \bar{y}_i A^i\big)-\lambda_{\min}\big(C+\epsilon_k\bar{C}-\sum_{i=1}^m \hat{y}_i A^i\big)},$$

we can realize that $\rho_k\to 0$ and $y_k\in\mathcal{D}^y_{\mathrm{SDO}}(\epsilon_k)$ for large k, since $0\le\alpha_k\le 1$ and $\alpha_k\to 0$. This completes the proof. $\square$

By the interior point condition, we can prove the outer semicontinuity of the optimal set-valued mapping at any $\bar{\epsilon}\in\operatorname{int}(\mathcal{E})$.

Lemma 5.1.4. *The set-valued mapping $\mathcal{P}^*_{\mathrm{SDO}}(.)\times\mathcal{D}^*_{\mathrm{SDO}}(.)$ is outer semicontinuous at every $\bar{\epsilon}\in\operatorname{int}(\mathcal{E})$.*

Proof. The outer semicontinuity follows from the closedness of the positive semidefinite cone and continuity of the functions in (1.6), i.e., given the sequences $\epsilon_k\to\bar{\epsilon}$ and $(X(\epsilon_k),y(\epsilon_k),S(\epsilon_k))\in\mathcal{P}^*_{\mathrm{SDO}}(\epsilon_k)\times\mathcal{D}^*_{\mathrm{SDO}}(\epsilon_k)$ such that $(X(\epsilon_k),y(\epsilon_k),S(\epsilon_k))\to(\bar{X},\bar{y},\bar{S})$, the limit point $(\bar{X},\bar{y},\bar{S})$ satisfies (1.6) at $\bar{\epsilon}$. $\square$

Note that the optimal set mapping is not necessarily inner semicontinuous. We can provide a sufficient condition for the continuity of $\mathcal{P}^*_{\mathrm{SDO}}(.)$ and $\mathcal{D}^*_{\mathrm{SDO}}(.)$ even when both strict complementarity and nondegeneracy conditions fail. It can be shown that the optimal set is uniformly bounded near any $\bar{\epsilon} \in \operatorname{int}(\mathcal{E})$, see (1.35). The proof is adopted from [157].

Lemma 5.1.5. *The optimal set $\mathcal{P}^*_{\mathrm{SDO}}(.) \times \mathcal{D}^*_{\mathrm{SDO}}(.)$ is uniformly bounded near $\bar{\epsilon} \in \operatorname{int}(\mathcal{E})$.*

Proof. By the continuity of the set mappings $\mathcal{P}_{\mathrm{SDO}}(.)$ and $\mathcal{D}_{\mathrm{SDO}}(.)$ at $\bar{\epsilon}$ there exists a sequence of interior solutions $(X^\circ(\epsilon_k), y^\circ(\epsilon_k), S^\circ(\epsilon_k))$ converging to $(X^\circ(\bar{\epsilon}), y^\circ(\bar{\epsilon}), S^\circ(\bar{\epsilon}))$. Let $(X(\epsilon_k), y(\epsilon_k), S(\epsilon_k))$ be a primal-dual optimal solution of $(\mathrm{P}^{\epsilon_k}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon_k}_{\mathrm{SDO}})$. Then from the primal and dual constraints we have

$$\langle X(\epsilon_k) - X^\circ(\epsilon_k), S(\epsilon_k) - S^\circ(\epsilon_k)\rangle = 0,$$

which gives

$$\langle X(\epsilon_k), S^\circ(\epsilon_k)\rangle + \langle S(\epsilon_k), X^\circ(\epsilon_k)\rangle = \langle X^\circ(\epsilon_k), S^\circ(\epsilon_k)\rangle.$$

Then from $X(\epsilon_k) \succeq 0$ and $S^\circ(\epsilon_k) \succ 0$ we can derive

$$\|X(\epsilon_k)\|_F \leq \frac{\langle X^\circ(\epsilon_k), S^\circ(\epsilon_k)\rangle}{\lambda_{\min}(S^\circ(\epsilon_k))}.$$

Since $(X^\circ(\epsilon_k), y^\circ(\epsilon_k), S^\circ(\epsilon_k)) \to (X^\circ(\bar{\epsilon}), y^\circ(\bar{\epsilon}), S^\circ(\bar{\epsilon}))$, there exists $\varsigma > 0$ such that

$$\|X(\epsilon_k)\|_F \leq \frac{\langle X^\circ(\epsilon_k), S^\circ(\epsilon_k)\rangle}{\lambda_{\min}(S^\circ(\epsilon_k))} \leq \frac{\langle X^\circ(\bar{\epsilon}), S^\circ(\bar{\epsilon})\rangle + \varsigma}{\lambda_{\min}(S^\circ(\bar{\epsilon})) - \varsigma}.$$

In a similar manner, we can show that

$$\|S(\epsilon_k)\|_F \leq \frac{\langle X^\circ(\epsilon_k), S^\circ(\epsilon_k)\rangle}{\lambda_{\min}(X^\circ(\epsilon_k))} \leq \frac{\langle X^\circ(\bar{\epsilon}), S^\circ(\bar{\epsilon})\rangle + \varsigma}{\lambda_{\min}(X^\circ(\bar{\epsilon})) - \varsigma},$$

which completes the proof. □

Lemma 5.1.6. *Assume that $\mathcal{P}^*_{\mathrm{SDO}}(.) \times \mathcal{D}^*_{\mathrm{SDO}}(.)$ is single-valued at $\bar{\epsilon}$, then $\mathcal{P}^*_{\mathrm{SDO}}(.) \times \mathcal{D}^*_{\mathrm{SDO}}(.)$ is continuous at $\bar{\epsilon}$.*

Proof. Let $\mathcal{P}^*_{\text{SDO}}(\bar\epsilon) \times \mathcal{D}^*_{\text{SDO}}(\bar\epsilon) = \big\{\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)\big\}$. By the outer semicontinuity of the set-valued mapping we have

$$\liminf_{\epsilon\to\bar\epsilon} \mathcal{P}^*_{\text{SDO}}(\epsilon) \times \mathcal{D}^*_{\text{SDO}}(\epsilon) \subseteq \limsup_{\epsilon\to\bar\epsilon} \mathcal{P}^*_{\text{SDO}}(\epsilon) \times \mathcal{D}^*_{\text{SDO}}(\epsilon) = \big\{\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)\big\}. \tag{5.2}$$

We know from Lemma 5.1.5 that $\mathcal{P}^*_{\text{SDO}}(.) \times \mathcal{D}^*_{\text{SDO}}(.)$ is uniformly bounded near $\bar\epsilon$. Then from the error bound result for an LMI system, see Theorem A.2.4, we get

$$\operatorname{dist}\Big(\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big),\ \mathcal{P}^*_{\text{SDO}}(\epsilon) \times \mathcal{D}^*_{\text{SDO}}(\epsilon)\Big) \leq \big(|\epsilon-\bar\epsilon|\|\bar C\|_F\big)^\gamma,$$

when $|\epsilon-\bar\epsilon|\|\bar C\|_F \leq 1$. Then for any sequence $\epsilon_k \to \bar\epsilon$, there exists a sequence $\big(X(\epsilon_k), y(\epsilon_k), S(\epsilon_k)\big)$ converging to $\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)$, which by (5.2) proves the continuity of $\mathcal{P}^*_{\text{SDO}}(.) \times \mathcal{D}^*_{\text{SDO}}(.)$ at $\bar\epsilon$. $\square$

Remark 5.1.1. *Another proof can be given for the continuity of both $\mathcal{P}^*_{\text{SDO}}(.)$ and $\mathcal{D}^*_{\text{SDO}}(.)$. By the interior point condition, the set-valued mapping $\mathcal{D}^y_{\text{SDO}}(.)$ is continuous at $\bar\epsilon$, and $\mathcal{D}^*_{\text{SDO}}(.)$ is uniformly bounded near $\bar\epsilon$. Then the continuity of $\mathcal{D}^*_{\text{SDO}}(.)$ at $\bar\epsilon$ is immediate from Theorem 1.7.8.*

If the strict complementarity and primal-dual nondegeneracy conditions hold at $\bar\epsilon$, then the optimal solution is unique and strictly complementary. Hence, the continuity of $\mathcal{D}^*_{\text{SDO}}(.)$ at $\bar\epsilon$ follows from Lemma 5.1.6. However, we can provide stronger results by applying the implicit function theorem. Note that the optimality conditions for ($\text{P}^\epsilon_{\text{SDO}}$) and ($\text{D}^\epsilon_{\text{SDO}}$) can be written as

$$F_{\text{SD}}\big(X, y, S, \epsilon\big) := \begin{pmatrix} \mathcal{A}^s \operatorname{svec}(X) - b \\ (\mathcal{A}^s)^T y + \operatorname{svec}(S) - \operatorname{svec}(C) - \epsilon \operatorname{svec}(\bar C) \\ \frac{1}{2}\operatorname{svec}(XS+SX) \end{pmatrix} = 0, \tag{5.3}$$
$$X, S \succeq 0.$$

Then the Jacobian of the linear equations in (5.3) is given by

$$\nabla F_{\rm SD}(X,y,S) := \begin{pmatrix} \mathcal{A}^s & 0 & 0 \\ 0 & (\mathcal{A}^s)^T & I_{n(n+1)/2} \\ S\otimes_s I_n & 0 & X\otimes_s I_n \end{pmatrix},$$

where $\otimes_s$ denotes the *symmetric Kronecker product.* The following technical lemma is in order.

Lemma 5.1.7 (Theorem 3.1 in [7] and [73])**.** *The Jacobian $\nabla F_{\rm SD}\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)$ is nonsingular if and only if the optimal solution $\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)$ satisfies strict complementarity, and both the primal and dual nondegeneracy conditions.* □

Now, we have the following result.

Theorem 5.1.1. *Let $(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon))$ be a strictly complementary solution. Then there exists $\varsigma > 0$ and a unique continuously differentiable mapping $(X^*(.), y^*(.), S^*(.))$ on $(\bar\epsilon-\varsigma, \bar\epsilon+\varsigma)$ such that $(X^*(\epsilon''), y^*(\epsilon''), S^*(\epsilon''))$ is the unique strictly complementary optimal solution for $(\mathrm{P}^{\epsilon''}_{\rm SDO})$ and $(\mathrm{D}^{\epsilon''}_{\rm SDO})$ for all $\epsilon'' \in (\bar\epsilon-\varsigma, \bar\epsilon+\varsigma)$.*

Proof. We can observe from Lemma 5.1.7 that the conditions of the implicit function theorem hold, see Theorem A.4.3. Then there exists $\varsigma' > 0$ and a unique continuously differentiable mapping $(X'(.), y'(.), S'(.))$ on $(\bar\epsilon-\varsigma', \bar\epsilon+\varsigma')$ such that $(X'(\epsilon'), y'(\epsilon'), S'(\epsilon'))$ satisfies the equations in (5.3) for all $\epsilon' \in (\bar\epsilon-\varsigma', \bar\epsilon+\varsigma')$ and $(X'(\bar\epsilon), y'(\bar\epsilon), S'(\bar\epsilon)) = (X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon))$. Since $(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon))$ is strictly complementary and $(X'(.), y'(.), S'(.))$ is continuously differentiable on $(\bar\epsilon-\varsigma', \bar\epsilon+\varsigma')$, we have

$$X'(\epsilon) + S'(\epsilon) \succ 0,$$

for all ϵ sufficiently close to $\bar\epsilon$. Hence, from the complementarity condition $XS + SX = 0$ and $X'(\epsilon) + S'(\epsilon) \succ 0$ we can conclude that $X'(\epsilon) \succeq 0$ and $S'(\epsilon) \succeq 0$

and $X'(\epsilon)S'(\epsilon) = 0$, i.e., $(X'(\epsilon), y'(\epsilon), S'(\epsilon))$ is a strictly complementary optimal solution for $(\mathrm{P}^{\epsilon}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon}_{\mathrm{SDO}})$. Furthermore, by the continuity arguments and the nonsingularity of the Jacobian at $(X^*(\bar{\epsilon}), y^*(\bar{\epsilon}), S^*(\bar{\epsilon}))$, $\nabla F_{\mathrm{SD}}\big(X'(.), y'(.), S'(.)\big)$ is nonsingular in a sufficiently small neighborhood of $\bar{\epsilon}$. Consequently, by using Lemma 5.1.7 again, we can find a neighborhood of $\bar{\epsilon}$ such that the optimal solutions are strictly complementary and primal-dual nondegenerate. $\square$

5.1.2 Sensitivity of the optimal partition for SDO

In this section, we investigate the behavior of the optimal partition and the optimal set mapping under perturbation of the objective vector. From now on,

$$\pi_{\mathrm{SDO}}(\epsilon) := \big(\mathcal{B}(\epsilon), \mathcal{T}(\epsilon), \mathcal{N}(\epsilon)\big)$$

denotes the optimal partition of $(\mathrm{P}^{\epsilon}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon}_{\mathrm{SDO}})$ for a given ϵ. We introduce and characterize the subintervals of $\mathrm{int}(\mathcal{E})$ on which the optimal partition or the dimension of both $\mathcal{B}(\epsilon)$ and $\mathcal{N}(\epsilon)$ is stable. The discussion is motived by minimizing a parametric objective function on the 3-elliptope [19], see Figure 5.1:

$$\left\{(x,y,z) \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0 \right\}.$$

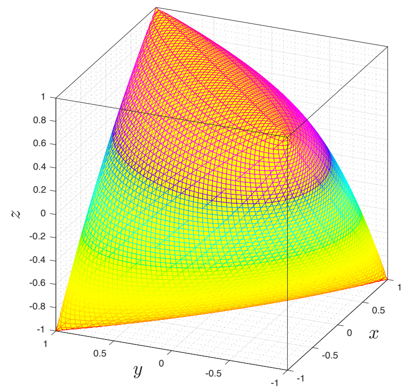

Figure 5.1: The illustration of a 3-elliptope.

Example 5.1.1. *Consider the following SDO problem:*

$$A^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}, \qquad \bar{C} = \begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}, \qquad b = (1,\ 1,\ 1)^T.$$

For all $\epsilon \in [0,1]$ *a maximally complementary solution is given by*

$$X^*(\epsilon) = \begin{pmatrix} 1 & \frac{1}{2} - \epsilon & \epsilon - \frac{1}{2} \\ \frac{1}{2} - \epsilon & 1 & 1 - 2(\epsilon - \frac{1}{2})^2 \\ \epsilon - \frac{1}{2} & 1 - 2(\epsilon - \frac{1}{2})^2 & 1 \end{pmatrix},$$

$$y^*(\epsilon) = (-(2\epsilon - 1)^2,\ -1,\ -1)^T,$$

$$S^*(\epsilon) = \begin{pmatrix} (2\epsilon - 1)^2 & 2\epsilon - 1 & 1 - 2\epsilon \\ 2\epsilon - 1 & 1 & -1 \\ 1 - 2\epsilon & -1 & 1 \end{pmatrix},$$

where the eigenvalues of $X^*(\epsilon)$ *and* $S^*(\epsilon)$ *are given by*

$$\begin{array}{lll}\lambda_{[1]}\big(X^*(\epsilon)\big) = -2\epsilon^2 + 2\epsilon + \dfrac{3}{2}, & \lambda_{[2]}\big(X^*(\epsilon)\big) = 2\epsilon^2 - 2\epsilon + \dfrac{3}{2}, & \lambda_{[3]}\big(X^*(\epsilon)\big) = 0,\\ \lambda_{[1]}\big(S^*(\epsilon)\big) = 4\epsilon^2 - 4\epsilon + 3, & \lambda_{[2]}\big(S^*(\epsilon)\big) = 0, & \lambda_{[3]}\big(S^*(\epsilon)\big) = 0.\end{array}$$

The optimal partition at $\epsilon = \frac{1}{2}$ *is given by*

$$\mathcal{B}(\epsilon) = \mathcal{R}\left(\begin{pmatrix} 0 & 1\\ 1/\sqrt{2} & 0\\ 1/\sqrt{2} & 0\end{pmatrix}\right), \qquad \mathcal{T}(\epsilon) = \{0\}, \qquad \mathcal{N}(\epsilon) = \mathcal{R}\left(\begin{pmatrix} 0\\ -1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}\right),$$

while for all $\epsilon \in [0,1] \setminus \{\frac{1}{2}\}$ *we have*

$$\begin{aligned}\mathcal{B}(\epsilon) &= \mathcal{R}\left(\begin{pmatrix} 0 & 2\operatorname{sgn}(2\epsilon-1)/\sqrt{2(2\epsilon-1)^2+4}\\ 1/\sqrt{2} & -|2\epsilon-1|/\sqrt{2(2\epsilon-1)^2+4}\\ 1/\sqrt{2} & |2\epsilon-1|/\sqrt{2(2\epsilon-1)^2+4}\end{pmatrix}\right),\\ \mathcal{T}(\epsilon) &= \{0\},\\ \mathcal{N}(\epsilon) &= \mathcal{R}\left(\begin{pmatrix} (1-2\epsilon)/\sqrt{(2\epsilon-1)^2+2}\\ -1/\sqrt{(2\epsilon-1)^2+2}\\ 1/\sqrt{(2\epsilon-1)^2+2}\end{pmatrix}\right),\end{aligned}$$

where $\operatorname{sgn}(.)$ *denotes the signum function. We can observe that* $\big(X^*(\epsilon), y^*(\epsilon), S^*(\epsilon)\big)$ *is strictly complementary for all* $\epsilon \in [0,1]$*, and both* $\operatorname{rank}\big(X^*(\epsilon)\big)$ *and* $\operatorname{rank}\big(S^*(\epsilon)\big)$ *are constant on* $[0,1]$*. It can be further investigated that the primal and dual nondegeneracy conditions hold at all* $\epsilon \in [0,1] \setminus \{\frac{1}{2}\}$*, and at* $\epsilon = \frac{1}{2}$ *the dual nondegeneracy condition fails. For instance, for all* $\epsilon \in [0,1] \setminus \{\frac{1}{2}\}$ *a common orthonormal eigenvector basis of* $X^*(\epsilon)$ *and* $S^*(\epsilon)$ *is given by*

$$Q^*(\epsilon) = \begin{pmatrix} 0 & 2\operatorname{sgn}(2\epsilon-1)/\sqrt{2(2\epsilon-1)^2+4} & (1-2\epsilon)/\sqrt{(2\epsilon-1)^2+2}\\ 1/\sqrt{2} & -|2\epsilon-1|/\sqrt{2(2\epsilon-1)^2+4} & -1/\sqrt{(2\epsilon-1)^2+2}\\ 1/\sqrt{2} & |2\epsilon-1|/\sqrt{2(2\epsilon-1)^2+4} & 1/\sqrt{(2\epsilon-1)^2+2}\end{pmatrix}.$$

Then, using the conditions in Lemma 1.3.1, one can check that the matrices

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & (2\mathrm{sgn}(2\epsilon-1)/\nu_1)^2 & 2\mathrm{sgn}(2\epsilon-1)(1-2\epsilon)/(\nu_1\nu_2) \\ 0 & 2\mathrm{sgn}(2\epsilon-1)(1-2\epsilon)/(\nu_1\nu_2) & 0 \end{pmatrix},$$

$$\begin{pmatrix} 1/2 & -|2\epsilon-1|/(\sqrt{2}\nu_1) & -1/(\sqrt{2}\nu_2) \\ -|2\epsilon-1|/(\sqrt{2}\nu_1) & ((2\epsilon-1)/\nu_1)^2 & |2\epsilon-1|/(\nu_1\nu_2) \\ -1/(\sqrt{2}\nu_2) & |2\epsilon-1|/(\nu_1\nu_2) & 0 \end{pmatrix},$$

$$\begin{pmatrix} 1/2 & |2\epsilon-1|/(\sqrt{2}\nu_1) & 1/(\sqrt{2}\nu_2) \\ |2\epsilon-1|/(\sqrt{2}\nu_1) & ((2\epsilon-1)/\nu_1)^2 & |2\epsilon-1|/(\nu_1\nu_2) \\ 1/(\sqrt{2}\nu_2) & |2\epsilon-1|/(\nu_1\nu_2) & 0 \end{pmatrix},$$

where

$$\nu_1 := \sqrt{2(2\epsilon-1)^2+4},$$
$$\nu_2 := \sqrt{(2\epsilon-1)^2+2},$$

are linearly independent. Furthermore, we can observe that the following matrices span $\mathbb{S}^2$:

$$\begin{pmatrix} 0 & 0 \\ 0 & (2\mathrm{sgn}(2\epsilon-1)/\nu_1)^2 \end{pmatrix}, \quad \begin{pmatrix} 1/2 & -|2\epsilon-1|/(\sqrt{2}\nu_1) \\ -|2\epsilon-1|/(\sqrt{2}\nu_1) & ((2\epsilon-1)/\nu_1)^2 \end{pmatrix},$$
$$\begin{pmatrix} 1/2 & |2\epsilon-1|/(\sqrt{2}\nu_1) \\ |2\epsilon-1|/(\sqrt{2}\nu_1) & ((2\epsilon-1)/\nu_1)^2 \end{pmatrix},$$

which implies dual nondegeneracy of the optimal solution.

As indicated in [59] and also demonstrated by Example 5.1.1, the optimal partition might vary with ϵ in a subinterval of $\mathrm{int}(\mathcal{E})$. However, the dimension of $\mathcal{B}(.)$ and $\mathcal{N}(.)$, or equivalently $\mathrm{rank}\big(X^*(.)\big)$ and $\mathrm{rank}\big(S^*(.)\big)$, might be stable in certain subintervals. This is in contrast to LO, where the interval $\mathcal{E}$ is divided into subintervals each with a unique optimal partition, see e.g., Jansen et al. [87].

Motivated by this observation, we make the following definitions:

Definition 5.1.1. *The two optimal partitions* $\pi_{\mathrm{SDO}}(\epsilon')$ *and* $\pi_{\mathrm{SDO}}(\epsilon'')$ *are called identical if* $\pi_{\mathrm{SDO}}(\epsilon') = \pi_{\mathrm{SDO}}(\epsilon'')$*, i.e.,*

$$\mathcal{B}(\epsilon') = \mathcal{B}(\epsilon''), \quad \textit{and} \quad \mathcal{N}(\epsilon') = \mathcal{N}(\epsilon'').$$

Otherwise, if

$$\dim\big(\mathcal{B}(\epsilon')\big) = \dim\big(\mathcal{B}(\epsilon'')\big), \quad \textit{and} \quad \dim\big(\mathcal{N}(\epsilon')\big) = \dim\big(\mathcal{N}(\epsilon'')\big),$$

then the two optimal partitions $\pi_{\mathrm{SDO}}(\epsilon')$ *and* $\pi_{\mathrm{SDO}}(\epsilon'')$ *are called weakly identical, and it is denoted by* $\pi_{\mathrm{SDO}}(\epsilon') \overset{w}{=} \pi_{\mathrm{SDO}}(\epsilon'')$*.*

It is immediate from the definition that if the partitions are not weakly identical, then they are not identical either.

Using the continuity arguments, we can estimate the behavior of the optimal partition $\pi_{\mathrm{SDO}}(\epsilon)$ in a neighborhood of ϵ.

Lemma 5.1.8. *If* $\mathcal{P}^*_{\mathrm{SDO}}(.)$ *is continuous at* $\bar{\epsilon}$*, then* $n_{\mathcal{B}(\bar{\epsilon})} \leq n_{\mathcal{B}(\epsilon')}$ *for all* ϵ' *sufficiently close to* $\bar{\epsilon}$*. If* $\mathcal{D}^*_{\mathrm{SDO}}(.)$ *is continuous at* $\bar{\epsilon}$*, then* $n_{\mathcal{N}(\bar{\epsilon})} \leq n_{\mathcal{N}(\epsilon')}$ *for all* ϵ' *in a small neighborhood of* $\bar{\epsilon}$*.*

Proof. Let $\big(X^*(\bar{\epsilon}), y^*(\bar{\epsilon}), S^*(\bar{\epsilon})\big)$ be a maximally complementary optimal solution. By the inner semicontinuity of the primal and dual optimal set mappings at $\bar{\epsilon}$, for ϵ' sufficiently close to $\bar{\epsilon}$ there exist primal optimal solution $X(\epsilon')$ and dual optimal solution $(y(\epsilon'), S(\epsilon'))$ close to $\big(X^*(\bar{\epsilon}), y^*(\bar{\epsilon}), S^*(\bar{\epsilon})\big)$ so that

$$\begin{aligned} &\lambda_{[i]}(X(\epsilon')) > 0, && i = 1, \ldots, n_{\mathcal{B}(\bar{\epsilon})}, \\ &\lambda_{[i]}(S(\epsilon')) > 0, && i = 1, \ldots, n_{\mathcal{N}(\bar{\epsilon})} \end{aligned}$$

holds. All this implies that $n_{\mathcal{B}(\bar{\epsilon})} \leq n_{\mathcal{B}(\epsilon')}$ and $n_{\mathcal{N}(\bar{\epsilon})} \leq n_{\mathcal{N}(\epsilon')}$. This completes the proof. $\square$

Remark 5.1.2. *More generally, the inclusion properties in Lemma 5.1.8 hold if there exists $\epsilon_k \to \bar{\epsilon}$ such that*

$$\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{SDO}}(\epsilon_k) \cap \mathrm{ri}\left(\mathcal{P}^*_{\mathrm{SDO}}(\bar{\epsilon})\right) \neq \emptyset, \qquad \liminf_{k\to\infty} \mathcal{D}^*_{\mathrm{SDO}}(\epsilon_k) \cap \mathrm{ri}\left(\mathcal{D}^*_{\mathrm{SDO}}(\bar{\epsilon})\right) \neq \emptyset.$$

In the following sections, we review linearity intervals from [59] and then introduce nonlinearity intervals and transition points of the optimal partition for $(\mathrm{P}^{\epsilon}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon}_{\mathrm{SDO}})$.

5.1.2.1 Linearity intervals

For a parametric LO and LCQO, the interval $\mathcal{E}$ can be partitioned into subintervals, so called linearity intervals, where each subinterval is associated with a unique optimal partition and a unique primal optimal set, e.g., for LO the index sets in $\mathcal{B}(\epsilon')$ and $\mathcal{N}(\epsilon')$ are invariant with respect to ϵ' in a linearity interval.

Let $\mathcal{I}_{\mathrm{lin}}$ be a subset of $\mathrm{int}(\mathcal{E})$. Then $\mathcal{I}_{\mathrm{lin}}$ is called a linearity interval if $\pi_{\mathrm{SDO}}(\epsilon') = \pi_{\mathrm{SDO}}(\epsilon'')$ for all $\epsilon', \epsilon'' \in \mathcal{I}_{\mathrm{lin}}$. The following result is an extension from LCQO [16, 17].

Lemma 5.1.9. *Let $\epsilon', \epsilon'' \in \mathrm{int}(\mathcal{E})$. If $\pi_{\mathrm{SDO}}(\epsilon') = \pi_{\mathrm{SDO}}(\epsilon'')$ and $\epsilon_\rho := \rho\epsilon' + (1-\rho)\epsilon''$ for every $0 \le \rho \le 1$, then $\pi_{\mathrm{SDO}}(\epsilon') = \pi_{\mathrm{SDO}}(\epsilon'') = \pi(\epsilon_\rho)$. Moreover,*

$$\begin{aligned} X^*(\epsilon_\rho) &:= \rho X^*(\epsilon') + (1-\rho)X^*(\epsilon''),\\ y^*(\epsilon_\rho) &:= \rho y^*(\epsilon') + (1-\rho)y^*(\epsilon''),\\ S^*(\epsilon_\rho) &:= \rho S^*(\epsilon') + (1-\rho)S^*(\epsilon'') \end{aligned} \tag{5.4}$$

is a maximally complementary solution of $(\mathrm{P}^{\epsilon_\rho}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon_\rho}_{\mathrm{SDO}})$.

Proof. Since $\mathcal{B}(\epsilon') = \mathcal{B}(\epsilon'')$ and $\mathcal{N}(\epsilon') = \mathcal{N}(\epsilon'')$, it is easy to see from Theorem 1.5.1 that $\left(X^*(\epsilon_\rho), y^*(\epsilon_\rho), S^*(\epsilon_\rho)\right)$ is a primal-dual optimal solution of $(\mathrm{P}^{\epsilon_\rho}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon_\rho}_{\mathrm{SDO}})$.

Furthermore, from (5.4) we get

$$X^*(\epsilon_\rho) = Q_{\mathcal{B}(\epsilon')}\big(\rho U_{X^*(\epsilon')} + (1-\rho)U_{X^*(\epsilon'')}\big)Q^T_{\mathcal{B}(\epsilon')}, \quad \rho U_{X^*(\epsilon')} + (1-\rho)U_{X^*(\epsilon'')} \succ 0,$$
$$S^*(\epsilon_\rho) = Q_{\mathcal{N}(\epsilon')}\big(\rho U_{S^*(\epsilon')} + (1-\rho)U_{S^*(\epsilon'')}\big)Q^T_{\mathcal{N}(\epsilon')}, \quad \rho U_{S^*(\epsilon')} + (1-\rho)U_{S^*(\epsilon'')} \succ 0,$$

which implies

$$\mathcal{B}(\epsilon') = \mathcal{R}(X^*(\epsilon_\rho)) \subseteq \mathcal{B}(\epsilon_\rho), \qquad \mathcal{N}(\epsilon') = \mathcal{R}(S^*(\epsilon_\rho)) \subseteq \mathcal{N}(\epsilon_\rho),$$

where the inclusions follow from the definition of a maximally complementary solution. Using the same argument, we can choose a sufficiently small κ to generate

$$X\big((1+\kappa)\epsilon'' - \kappa\epsilon'\big) = Q_{\mathcal{B}(\epsilon')}\big((1+\kappa)U_{X^*(\epsilon'')} - \kappa U_{X^*(\epsilon')}\big)Q^T_{\mathcal{B}(\epsilon')},$$
$$S\big((1+\kappa)\epsilon'' - \kappa\epsilon'\big) = Q_{\mathcal{N}(\epsilon')}\big((1+\kappa)U_{S^*(\epsilon'')} - \kappa U_{S^*(\epsilon')}\big)Q^T_{\mathcal{N}(\epsilon')},$$

which is an optimal solution for $(\mathrm{P}_{\mathrm{SDO}}^{(1+\kappa)\epsilon''-\kappa\epsilon'})$ and $(\mathrm{D}_{\mathrm{SDO}}^{(1+\kappa)\epsilon''-\kappa\epsilon'})$. Note that κ can be made so small that $(1+\kappa)\epsilon'' - \kappa\epsilon' \in \mathrm{int}(\mathcal{E})$. Now, if $\mathcal{T}(\epsilon') \supsetneq \mathcal{T}(\epsilon_\rho)$, then there would exist $0 \neq q \in \mathcal{R}\big(Q_{\mathcal{T}(\epsilon')}\big)$ so that

$$q^T\big(X^*(\epsilon_\rho) + S^*(\epsilon_\rho)\big)q > 0. \tag{5.5}$$

However, this would contradict the optimal partition at ϵ' and ϵ''. To see this, we can check that

$$\epsilon'' = \frac{\kappa}{\kappa+\rho}\epsilon_\rho + \frac{\rho}{\kappa+\rho}\epsilon'''$$

where $\epsilon''' := (1+\kappa)\epsilon'' - \kappa\epsilon'$. Then

$$X(\epsilon'') = \frac{\kappa}{\kappa+\rho}X^*(\epsilon_\rho) + \frac{\rho}{\kappa+\rho}X^*(\epsilon'''),$$
$$y(\epsilon'') = \frac{\kappa}{\kappa+\rho}y^*(\epsilon_\rho) + \frac{\rho}{\kappa+\rho}y^*(\epsilon'''),$$
$$S(\epsilon'') = \frac{\kappa}{\kappa+\rho}S^*(\epsilon_\rho) + \frac{\rho}{\kappa+\rho}S^*(\epsilon''')$$

gives an optimal solution for $(\mathrm{P}_{\mathrm{SDO}}^{\epsilon''})$ and $(\mathrm{D}_{\mathrm{SDO}}^{\epsilon''})$, since both $\big(X^*(\epsilon_\rho), y^*(\epsilon_\rho), S^*(\epsilon_\rho)\big)$ and $\big(X^*(\epsilon'''), y^*(\epsilon'''), S^*(\epsilon''')\big)$ can be represented using $Q_{\mathcal{B}(\epsilon')}$ and $Q_{\mathcal{N}(\epsilon')}$. However, we have from (5.5) that

$$q^T\big(X(\epsilon'') + S(\epsilon'')\big)q > 0,$$

which is a contradiction, since q is a common eigenvector of $X(\epsilon'')$ and $S(\epsilon'')$. Therefore we should have $\mathcal{T}(\epsilon') = \mathcal{T}(\epsilon_\rho)$, which induces $\mathcal{B}(\epsilon') = \mathcal{B}(\epsilon_\rho)$ and $\mathcal{N}(\epsilon') = \mathcal{N}(\epsilon_\rho)$. The second part of the proof is immediate. $\square$

Example 5.1.2. *Consider the following SDO problem from [59]:*

$$A^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$C = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}, \qquad \bar{C} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \qquad b = (1,\ 0,\ 0)^T,$$

where for all $\epsilon \in (-1, 2)$, both the primal and the dual problems satisfy the interior point condition. On the interval $(0, 2)$ the unique optimal solution is strictly complementary, and the optimal partition remains constant:

$$\mathcal{B}(\epsilon) = \mathcal{R}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{T}(\epsilon) = \{0\}, \quad \mathcal{N}(\epsilon) = \mathcal{R}\left(\begin{pmatrix} 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}\right), \qquad \forall \epsilon \in (0, 2).$$

Strict complementarity fails at $\epsilon = 0$, and the optimal partition changes to

$$\mathcal{B}(0) = \mathcal{R}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{T}(0) = \mathcal{R}\begin{pmatrix} 1 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \quad \mathcal{N}(0) = \mathcal{R}\begin{pmatrix} 1 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

We can observe that the optimal partition at $\epsilon = 0$ is different from the partitions in $(-1,0)$ and $(0,2)$. Furthermore, the optimal partition in $(-1,0)$ continuously varies with ϵ.

Let $\bar{\epsilon} \in \mathcal{I}_{\text{lin}}$. By the definition of a linearity interval, $\mathcal{I}_{\text{lin}}$ is the set of all $\epsilon' \in \operatorname{int}(\mathcal{E})$ for which the system

$$\begin{aligned} \langle A^i, Q_{\mathcal{B}(\bar{\epsilon})} U_X Q_{\mathcal{B}(\bar{\epsilon})}^T \rangle &= b_i, \qquad i = 1, \ldots, m, && U_X \succ 0, \\ \sum_{i=1}^m A^i y_i + Q_{\mathcal{N}(\bar{\epsilon})} U_S Q_{\mathcal{N}(\bar{\epsilon})}^T &= C + \epsilon \bar{C}, && U_S \succ 0 \end{aligned}$$

remain feasible. Therefore, from Lemma 5.1.9 it is immediate that $\mathcal{I}_{\text{lin}}$ is either a singleton or an open, possibly unbounded, interval.

Remark 5.1.3. *It follows from Lemma 5.1.9 that*

$$\langle C + \epsilon_\rho \bar{C}, X^*(\epsilon_\rho) \rangle = \rho \langle C + \epsilon' \bar{C}, X^*(\epsilon') \rangle + (1-\rho) \langle C + \epsilon'' \bar{C}, X^*(\epsilon'') \rangle,$$

i.e., the optimal value function is indeed linear on a linearity interval. Furthermore, either there exists a unique primal optimal solution, or a unique primal optimal set associated with a linearity interval. This is an extension of Corollary 2 in [87].

A linearity interval can be computed by solving a pair of auxiliary SDO problems.

Lemma 5.1.10 (Lemma 4.1 in [59])**.** *Assume that $\bar{\epsilon}$ belongs to a bounded linearity interval $\mathcal{I}_{\text{lin}}$. Then the extreme points of $\mathcal{I}_{\text{lin}}$ can be obtained by solving*

$$\begin{aligned} \alpha_{\text{lin}}(\beta_{\text{lin}}) := \inf(\sup) \quad & \epsilon \\ \text{s.t.} \quad & \sum_{i=1}^m y_i A^i + Q_{\mathcal{N}(\bar{\epsilon})} U_S Q_{\mathcal{N}(\bar{\epsilon})}^T = C + \epsilon \bar{C}, \\ & U_S \succ 0. \end{aligned} \tag{5.6}$$

If $\mathcal{I}_{\text{lin}}$ is unbounded, then we have either $\alpha_{\text{lin}} = -\infty$, $\beta_{\text{lin}} = \infty$, or both.

Remark 5.1.4. *The stability of strict complementarity in a linearity interval follows from the definition, i.e., it either holds or fails at every $\epsilon' \in \mathcal{I}_{\text{lin}}$. Stability holds as well for both the primal and dual nondegeneracy conditions. The case for primal nondegeneracy is obvious, since the primal optimal set mapping is constant in a linearity interval. Let $\bar{\epsilon} \in \mathcal{I}_{\text{lin}}$ and assume that $\big(y^*(\bar{\epsilon}), S^*(\bar{\epsilon})\big)$ is dual nondegenerate. Then the matrices*

$$Q^T_{\mathcal{B}(\bar{\epsilon})\cup\mathcal{T}(\bar{\epsilon})} A_i Q_{\mathcal{B}(\bar{\epsilon})\cup\mathcal{T}(\bar{\epsilon})}, \qquad i = 1, \dots, m$$

span $\mathbb{S}^{n-n_{\mathcal{N}(\bar{\epsilon})}}$ by the condition given in Section 1.3. Since the orthonormal basis $Q_{\mathcal{B}(.)\cup\mathcal{T}(.)}$ is constant on $\mathcal{I}_{\text{lin}}$, the dual nondegeneracy condition follows for all ϵ' in $\mathcal{I}_{\text{lin}}$.

5.1.2.2 Transition point and nonlinearity interval

As a result of Lemma 5.1.10, if $\alpha_{\text{lin}} < \beta_{\text{lin}}$, then $\alpha_{\text{lin}} < \bar{\epsilon} < \beta_{\text{lin}}$ belongs to the linearity interval $(\alpha_{\text{lin}}, \beta_{\text{lin}})$. Otherwise, $\alpha_{\text{lin}} = \bar{\epsilon} = \beta_{\text{lin}}$ indicates that the optimal partition changes in every neighborhood of $\bar{\epsilon}$. In other words, the optimal partitions around $\bar{\epsilon}$ are either nonidentical or weakly identical with $\pi_{\text{SDO}}(\bar{\epsilon})$. In the latter case, $\bar{\epsilon}$ belongs to a subinterval of $\operatorname{int}(\mathcal{E})$, where $\operatorname{rank}\big(X^*(.)\big)$ and $\operatorname{rank}\big(S^*(.)\big)$ remain constant. In Example 5.1.1, $[0, 1]$ is such a subinterval.

Definition 5.1.2. *The point $\bar{\epsilon} \in \operatorname{int}(\mathcal{E})$ is called a transition point if for every $\varsigma > 0$ there exists $\epsilon' \in (\bar{\epsilon} - \varsigma, \bar{\epsilon} + \varsigma) \subseteq \operatorname{int}(\mathcal{E})$ such that $\pi_{\text{SDO}}(\bar{\epsilon}) \neq \pi_{\text{SDO}}(\epsilon')$. A nonlinearity interval is defined as a (possibly unbounded) subinterval of maximal length $\mathcal{I}_{\text{non}} \subseteq \operatorname{int}(\mathcal{E})$ such that $\pi_{\text{SDO}}(\epsilon') \overset{w}{=} \pi_{\text{SDO}}(\epsilon'')$ for all $\epsilon', \epsilon'' \in \mathcal{I}_{\text{non}}$.*

In parametric LO and LCQO, two consecutive subintervals are adjoined at transition point which has a different optimal partition with respect to its neighboring subintervals, see e.g., Jansen et al. [87].

In contrast to the definition of a linearity interval, it does not follow from Definition 5.1.2 whether a nonlinearity interval is open. However, a sufficient condition can be given for the openness of a nonlinearity interval.

Lemma 5.1.11. *Let $\mathcal{I}_{\text{non}}$ be a nonlinearity interval. If for each $\epsilon' \in \mathcal{I}_{\text{non}}$ and for every $\{\epsilon_k\} \to \epsilon'$ there exists a sequence of maximally complementary solutions $\big(X^*(\epsilon_k), y^*(\epsilon_k), S^*(\epsilon_k)\big)$ converging to a maximally complementary solution $\big(X^*(\epsilon'), y^*(\epsilon'), S^*(\epsilon')\big)$, then $\mathcal{I}_{\text{non}}$ is an open interval.*

Proof. The proof follows from the constancy of rank in a nonlinearity interval, and the fact that for every $\epsilon' \in \mathcal{I}_{\text{non}}$ the eigenvalues of $X^*(.)$ and $S^*(.)$ vary continuously in a small neighborhood of ϵ'. □

Corollary 5.1.1. *Assume that both $\mathcal{P}^*_{\text{SDO}}(.)$ and $\mathcal{D}^*_{\text{SDO}}(.)$ are continuous at every $\epsilon' \in \mathcal{I}_{\text{non}}$. Then $\mathcal{I}_{\text{non}}$ is an open interval. In particular, $\mathcal{I}_{\text{non}}$ is open if both $\mathcal{P}^*_{\text{SDO}}(.)$ and $\mathcal{D}^*_{\text{SDO}}(.)$ are single-valued on $\mathcal{I}_{\text{non}}$.*

Proof. The proof follows from a more general result for inner semicontinuous set-valued mappings. See Corollary 1.3 in [82]. □

Using continuity arguments and the strict complementarity condition, we can provide sufficient conditions which guarantee that $\bar{\epsilon} \in \text{int}(\mathcal{E})$ belongs to the interior of a nonlinearity interval.

Lemma 5.1.12. *Assume that strict complementarity holds at $\bar{\epsilon} \in \text{int}(\mathcal{E})$. If for every $\{\epsilon_k\} \to \bar{\epsilon}$ there exists a sequence of optimal solutions $\big(X(\epsilon_k), y(\epsilon_k), S(\epsilon_k)\big)$ converging to a strictly complementary solution $\big(X^*(\bar{\epsilon}), y^*(\bar{\epsilon}), S^*(\bar{\epsilon})\big)$, then $\bar{\epsilon}$ belongs to the interior of a nonlinearity interval.*

Proof. Since $\big(X^*(\bar{\epsilon}), y^*(\bar{\epsilon}), S^*(\bar{\epsilon})\big)$ is strictly complementary, we have

$$\text{rank}\,\big(X^*(\bar{\epsilon})\big) + \text{rank}\,\big(S^*(\bar{\epsilon})\big) = n. \tag{5.7}$$

Then by the assumption and the lower semicontinuity of the rank function, see [79], there exists an optimal solution $\big(X(\epsilon_k), y(\epsilon_k), S(\epsilon_k)\big)$ such that

$$\begin{aligned}\operatorname{rank}\big(X(\epsilon_k)\big) &\geq \operatorname{rank}\big(X^*(\bar\epsilon)\big),\\ \operatorname{rank}\big(S(\epsilon_k)\big) &\geq \operatorname{rank}\big(S^*(\bar\epsilon)\big)\end{aligned}$$

for sufficiently large k, i.e., $\big(X(\epsilon_k), y(\epsilon_k), S(\epsilon_k)\big)$ is a strictly complementary solution for sufficiently large k. Therefore, it follows from (5.7) that the ranks of $X(.)$ and $S(.)$ stay constant in a small neighborhood of $\bar\epsilon$. $\square$

A special case of Lemma 5.1.12 happens when $\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)$ is unique and strictly complementary. Then $\nabla F_{\text{SD}}(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)$ is nonsingular by Lemma 5.1.7. Furthermore, by the implicit function theorem, see Theorem 5.1.1, there exists $\varsigma > 0$ so that $\big(X^*(.), y^*(.), S^*(.)\big)$ is unique and continuously differentiable on $(\bar\epsilon - \varsigma, \bar\epsilon + \varsigma)$. All this means that $\bar\epsilon$ belongs to the interior of a nonlinearity interval, as a result of Lemma 5.1.12.

Lemma 5.1.13. *Let $\big(X^*(\bar\epsilon), y^*(\bar\epsilon), S^*(\bar\epsilon)\big)$ be unique and strictly complementary. Furthermore, let $\mathcal{I}_{\text{inv}}$ be the open interval of maximal length containing $\bar\epsilon$ on which $\nabla F_{\text{SD}}(X^*(.), y^*(.), S^*(.)\big)$ is nonsingular. Then $\mathcal{I}_{\text{inv}}$ is a subinterval of the nonlinearity interval which contains $\bar\epsilon$. Further, if $\hat\epsilon$ is an extreme point[1] of $\mathcal{I}_{\text{inv}}$ and the strict complementarity condition fails at $\hat\epsilon$, then $\hat\epsilon$ is an extreme point of the nonlinearity interval which contains $\bar\epsilon$.*

Proof. The first part of proof follows from Lemmas 5.1.12 and 5.1.7, and the implicit function theorem. For the second part, note that

$$\begin{pmatrix} S^*(.) \otimes_s I_n & 0 & X^*(.) \otimes_s I_n \end{pmatrix}$$

is rank deficient at $\hat\epsilon$, due to failure of the strict complementarity condition. Hence, at least one of $\operatorname{rank}(X^*(.))$ or $\operatorname{rank}(S^*(.))$ must change at $\hat\epsilon$. $\square$

[1] Recall that a nonlinearity interval may have no extreme point.

Lemma 5.1.12 indicates that at a transition point $\bar{\epsilon}$ which satisfies strict complementarity condition, at least one of $X^*(.)$ or $S^*(.)$ has to be discontinuous, i.e., there exists a sequence $\epsilon_k \to \bar{\epsilon}$ so that $\big(X(\epsilon_k), y(\epsilon_k), S(\epsilon_k)\big)$ has no accumulation point[2] in $\mathrm{ri}\left(\mathcal{P}^*_{\mathrm{SDO}}(\bar{\epsilon}) \times \mathcal{D}^*_{\mathrm{SDO}}(\bar{\epsilon})\right)$. Then the following result is immediate.

Corollary 5.1.2. *At a transition point $\bar{\epsilon}$, at least one of the strict complementarity, primal nondegeneracy, or dual nondegeneracy conditions has to fail.*

Proof. If neither of the conditions fail, then there exists a neighborhood of $\bar{\epsilon}$ on which $(X^*(.), y^*(.), S^*(.))$ is uniquely defined and continuously differentiable. Then the result follows from Lemma 5.1.12. □

As indicated in Lemma 5.1.13, we can prove the continuity of the primal and dual optimal set mappings in a nonlinearity interval under uniqueness condition. In general, however, this statement is not necessarily true. The reason lies in the fact that lim inf of a sequence of faces is not necessarily a face of the feasible set, i.e., it might be a subset of the relative interior of a face.

Lemma 5.1.14. *The primal or dual optimal set mappings might be discontinuous (in a sense of Painlevé-Kuratowski) in a nonlinearity interval.*

Proof. The 3-elliptope is the counterexample. As worked out in Example 5.1.1, both the primal and dual nondegeneracy conditions hold at all $\epsilon \in [0,1] \setminus \{\frac{1}{2}\}$, and at $\epsilon = \frac{1}{2}$ the dual nondegeneracy condition fails. Hence, $\mathcal{P}^*_{\mathrm{SDO}}(.)$ is not continuous at $\epsilon = \frac{1}{2}$ even though $[0,1]$ is a subinterval of a nonlinearity interval. □

Remark 5.1.5. *Lemma 5.1.14 implies that $\mathcal{I}_{\mathrm{inv}}$ does not necessarily coincide with $\mathcal{I}_{\mathrm{non}}$ for a parametric SDO.*

[2] An accumulation point exists, since $\mathcal{P}^*_{\mathrm{SDO}}(.)$ and $\mathcal{D}^*_{\mathrm{SDO}}(.)$ are uniformly bounded near any $\epsilon' \in \mathrm{int}(\mathcal{E})$. This directly follows from the interior point condition, see Lemma 5.1.5.

5.1.3 Sensitivity of the approximation of the optimal partition

In this section, we resort to the perturbation theory of eigenspaces in [160] to measure the sensitivity of the approximation of the optimal partition at a given $\bar{\epsilon} \in \operatorname{int}(\mathcal{E})$. Without loss of generality, we may assume that $\bar{\epsilon} = 0$. Even though multiplicity of the eigenvalues causes discontinuous behavior of the eigenvectors, the range space of these eigenvectors are, in general, less sensitive to the perturbation of matrix entries, see e.g., [160]. Throughout this section, unless stated otherwise, we always assume that μ is positive. For the sake of simplicity, we drop ϵ from the optimal partition and optimal solutions at $\bar{\epsilon} = 0$.

Consider an equivalent form of the perturbed central path equations as follows

$$F_{\mathrm{SD}}\big(X, y, S, \mu, \epsilon\big) := \begin{pmatrix} \mathcal{A}^s \operatorname{svec}(X) - b \\ (\mathcal{A}^s)^T y + \operatorname{svec}(S) - \operatorname{svec}(C) - \epsilon \operatorname{svec}(\bar{C}) \\ \frac{1}{2} \operatorname{svec}(XS + SX - \mu I_n) \end{pmatrix} = 0, \qquad (5.8)$$

$$X, S \succeq 0.$$

It can be shown that $\nabla F_{\mathrm{SD}}\big(X^\mu, y^\mu, S^\mu\big)$ is nonsingular, see e.g., Theorem 3.3 in [32]. As a result, system (5.8) is solvable for all ϵ in a neighborhood of 0. This directly follows from the implicit function theorem and continuity arguments. For every ϵ the unique solution of (5.8) is denoted by $\big(X^\mu(\epsilon), y^\mu(\epsilon), S^\mu(\epsilon)\big)$ and a common eigenvector basis is represented by $Q^\mu(\epsilon)$.

Suppose that for $\bar{\epsilon} = 0$ a central solution $\big(X^\mu, y^\mu, S^\mu\big)$ is given, where $\mu < \tilde{\mu}$ as defined in (3.15). The eigenvectors of X^μ and S^μ can be rearranged so that

$$Q^\mu := \big(Q^\mu_{\mathcal{B}}, Q^\mu_{\mathcal{T}}, Q^\mu_{\mathcal{N}}\big).$$

It is known that $\mathcal{R}\big(Q^\mu_{\mathcal{B}}\big)$, $\mathcal{R}\big(Q^\mu_{\mathcal{T}}\big)$, and $\mathcal{R}\big(Q^\mu_{\mathcal{N}}\big)$ are invariant subspaces of both X^μ

and S^μ, since, e.g., $X^\mu\mathcal{R}\big(Q^\mu_\mathcal{B}\big) \subseteq \mathcal{R}\big(Q^\mu_\mathcal{B}\big)$ and $S^\mu\mathcal{R}\big(Q^\mu_\mathcal{N}\big) \subseteq \mathcal{R}\big(Q^\mu_\mathcal{N}\big)$. We are interested in the variation of $\mathcal{R}\big(Q^\mu_\mathcal{B}\big)$ and $\mathcal{R}\big(Q^\mu_\mathcal{N}\big)$, when ϵ belongs to a sufficiently small neighborhood of 0.

The idea in [160] is to span an invariant subspace of $X^\mu(\epsilon)$, with $\epsilon \in \operatorname{int}(\mathcal{E})$, using the first $n_\mathcal{B}$ columns of $Q^\mu W^\mu_\mathcal{B}$, where

$$W^\mu_\mathcal{B} := \begin{pmatrix} I_{n_\mathcal{B}} & -(V^\mu_\mathcal{B})^T \\ V^\mu_\mathcal{B} & I_{n_\mathcal{T}+n_\mathcal{N}} \end{pmatrix} \begin{pmatrix} \big(I_{n_\mathcal{B}} + (V^\mu_\mathcal{B})^T V^\mu_\mathcal{B}\big)^{-\frac{1}{2}} & 0 \\ 0 & \big(I_{n_\mathcal{T}+n_\mathcal{N}} + V^\mu_\mathcal{B}(V^\mu_\mathcal{B})^T\big)^{-\frac{1}{2}} \end{pmatrix} \tag{5.9}$$

is an $n \times n$ orthogonal matrix and $V^\mu_\mathcal{B} \in \mathbb{R}^{(n_\mathcal{T}+n_\mathcal{N})\times n_\mathcal{B}}$. Hence, the problem is equivalent to choosing $V^\mu_\mathcal{B}$ such that the column space of

$$Y^\mu_\mathcal{B} := \big(Q^\mu_\mathcal{B} + Q^\mu_{\mathcal{T}\cup\mathcal{N}} V^\mu_\mathcal{B}\big)\big(I_{n_\mathcal{B}} + (V^\mu_\mathcal{B})^T V^\mu_\mathcal{B}\big)^{-\frac{1}{2}} \tag{5.10}$$

becomes an invariant subspace of $X^\mu(\epsilon)$. Analogously, an invariant subspace of $S^\mu(\epsilon)$ can be formulated as the column space of

$$Y^\mu_\mathcal{N} := \big(Q^\mu_\mathcal{N} + Q^\mu_{\mathcal{B}\cup\mathcal{T}} V^\mu_\mathcal{N}\big)\big(I_{n_\mathcal{N}} + (V^\mu_\mathcal{N})^T V^\mu_\mathcal{N}\big)^{-\frac{1}{2}},$$

where $V^\mu_\mathcal{N} \in \mathbb{R}^{(n_\mathcal{B}+n_\mathcal{T})\times n_\mathcal{N}}$. Notice that $\mathcal{R}\big(Y^\mu_\mathcal{B}\big)$ and $\mathcal{R}\big(Y^\mu_\mathcal{N}\big)$ are not necessarily approximations, in terms of the discussion in Section 3.1.1, for $\mathcal{B}(\epsilon)$ and $\mathcal{N}(\epsilon)$, respectively, see the discussion after Theorem 5.1.3.

A sufficient condition for the existence of $V^\mu_\mathcal{B}$ and an upper bound on $\|V^\mu_\mathcal{B}\|_2$ are specified in the following theorem adopted from [62], see Theorem 4.11 in [160] for more general results. For the ease of exposition, we have tailored the theorem for central solutions by introducing

$$\Xi^\mu_X := X^\mu(\epsilon) - X^\mu, \qquad \Xi^\mu_S := S^\mu(\epsilon) - S^\mu.$$

Theorem 5.1.2 (Theorem 8.1.10 in [62]). *Let a central solution* $\big(X^\mu, y^\mu, S^\mu\big)$ *with* $\mu < \tilde{\mu}$ *be given. If*

$$\big\|\Xi^\mu_X\big\|_F \leq \frac{\lambda_{[n_\mathcal{B}]}(X^\mu) - \lambda_{[n_\mathcal{B}+1]}(X^\mu)}{5}, \tag{5.11}$$

then there exists $V^\mu_{\mathcal{B}}$ so that

$$\|V^\mu_{\mathcal{B}}\|_F \leq \frac{4\big\|(Q^\mu_{\mathcal{B}})^T \Xi^\mu_X Q^\mu_{\mathcal{T}\cup\mathcal{N}}\big\|_F}{\lambda_{[n_{\mathcal{B}}]}(X^\mu) - \lambda_{[n_{\mathcal{B}}+1]}(X^\mu)}. \tag{5.12}$$

$\square$

Remark 5.1.6. *Using the bounds in Lemma 3.1.1, it is easy to verify that*

$$\lambda_{[n_{\mathcal{B}}]}(X^\mu) - \lambda_{[n_{\mathcal{B}}+1]}(X^\mu) \geq \phi^\mu := \frac{\sigma}{n} - c\sqrt{n}(n\mu)^\gamma, \tag{5.13}$$

where the right hand side in (5.13) is positive when $\mu < \tilde{\mu}$.

The following technical lemma bounds the distance between $\mathcal{R}\big(Q^\mu_{\mathcal{B}}\big)$ and $\mathcal{R}\big(Y^\mu_{\mathcal{B}}\big)$. For brevity, we only state the result for $Y^\mu_{\mathcal{B}}$. The proof can be found in Corollary 8.1.11 in [62].

Lemma 5.1.15. *Let $Y^\mu_{\mathcal{B}}$ be defined as in (5.10). Then we have*

$$\operatorname{dist}\Big(\mathcal{R}\big(Q^\mu_{\mathcal{B}}\big), \mathcal{R}\big(Y^\mu_{\mathcal{B}}\big)\Big) \leq \|V^\mu_{\mathcal{B}}\|_2.$$

5.1.3.1 Upper bound on $\|\Xi^\mu_X\|_F$ and $\|\Xi^\mu_S\|_F$

The application of Theorem 5.1.2 requires an estimate of the effect of the perturbation on the central solutions. Due to the nonsingularity of the Jacobian, an upper bound on $\|\Xi^\mu_X\|_F$ and $\|\Xi^\mu_S\|_F$ can be obtained by applying Kantorovich theorem[3], see Theorem A.4.2, to F_{SD}, as defined in (5.8). To that end, we define

$$\begin{aligned}
\delta^\mu &:= \min\Big\{\lambda_{[n_{\mathcal{B}}]}(X^\mu),\ \lambda_{[n_{\mathcal{N}}]}(S^\mu),\ \lambda_{[n_{\mathcal{B}}+n_{\mathcal{T}}]}(X^\mu)\Big\},\\
\theta^\mu &:= \big\|\nabla F^{-1}_{\mathrm{SD}}\big(X^\mu, y^\mu, S^\mu\big)\big\|_2,\\
\eta^\mu &:= \big\|\nabla F^{-1}_{\mathrm{SD}}\big(X^\mu, y^\mu, S^\mu\big) F_{\mathrm{SD}}\big(X^\mu, y^\mu, S^\mu, \mu, \epsilon\big)\big\|_2.
\end{aligned}$$

[3]The Kantorovich theorem was applied in [122] to (5.8) at $\mu = 0$ under strict complementarity and primal-dual nondegeneracy conditions.

Lemma 5.1.16. *Let* (X^μ, y^μ, S^μ) *be a central solution, where* $\mu < \tilde{\mu}$*. If* ϵ *is chosen in such a way that*

$$|\epsilon| < \min\left\{\frac{\delta^\mu}{2\theta^\mu\|\bar{C}\|_F},\ \frac{1}{2(\theta^\mu)^2\|\bar{C}\|_F}\right\}, \tag{5.14}$$

then $\epsilon \in \operatorname{int}(\mathcal{E})$*, and there exists a central solution* $(X^\mu(\epsilon), y^\mu(\epsilon), S^\mu(\epsilon))$ *so that*

$$\|\Xi_X^\mu\|_F \le \frac{1-\sqrt{1-2|\epsilon|(\theta^\mu)^2\|\bar{C}\|_F}}{\theta^\mu}, \tag{5.15}$$

$$\|\Xi_S^\mu\|_F \le \frac{1-\sqrt{1-2|\epsilon|(\theta^\mu)^2\|\bar{C}\|_F}}{\theta^\mu}. \tag{5.16}$$

Proof. Note that F_{SD} is continuously differentiable, and ∇F_{SD} is Lipschitz continuous with global Lipschitz constant 1, see Lemma 2 in [122]. Furthermore, we have

$$\eta^\mu \le \|\nabla F_{\mathrm{SD}}^{-1}(X^\mu, y^\mu, S^\mu)\|_2 \|F_{\mathrm{SD}}(X^\mu, y^\mu, S^\mu, \mu, \epsilon)\|_2 = |\epsilon|\theta^\mu\|\bar{C}\|_F,$$

where the last equality follows from

$$F_{\mathrm{SD}}(X^\mu, y^\mu, S^\mu, \mu, \epsilon) = \begin{pmatrix} 0 \\ -\epsilon\,\mathrm{svec}(\bar{C}) \\ 0 \end{pmatrix}.$$

Thus, by the condition of Kantorovich theorem, if

$$\eta^\mu\theta^\mu \le |\epsilon|(\theta^\mu)^2\|\bar{C}\|_F \le \frac{1}{2},$$

then there exists a solution $(X^\mu(\epsilon), y^\mu(\epsilon), S^\mu(\epsilon))$ satisfying the equations in (5.8), such that

$$\left\|\left(\mathrm{svec}(X^\mu(\epsilon) - X^\mu);\ y^\mu(\epsilon) - y^\mu;\ \mathrm{svec}(S^\mu(\epsilon) - S^\mu)\right)\right\|_2 \le \frac{1-\sqrt{1-2|\epsilon|(\theta^\mu)^2\|\bar{C}\|_F}}{\theta^\mu}.$$

In particular, this implies that for $i = 1, \ldots, n_{\mathcal{B}} + n_{\mathcal{T}}$

$$\left|\lambda_{[i]}(X^\mu(\epsilon)) - \lambda_{[i]}(X^\mu)\right| \le \|\Xi_X^\mu\|_F \le \frac{1-\sqrt{1-2|\epsilon|(\theta^\mu)^2\|\bar{C}\|_F}}{\theta^\mu}, \tag{5.17}$$

and that for $j = 1, \dots, n_{\mathcal{N}}$

$$\left|\lambda_{[j]}\big(S^{\mu}(\epsilon)\big) - \lambda_{[j]}\big(S^{\mu}\big)\right| \le \left\|\Xi_S^{\mu}\right\|_F \le \frac{1 - \sqrt{1 - 2|\epsilon|(\theta^{\mu})^2\left\|\bar{C}\right\|_F}}{\theta^{\mu}}. \tag{5.18}$$

On the other hand, $X^{\mu}(\epsilon)$ and $S^{\mu}(\epsilon)$ stay positive definite if

$$\begin{aligned}
\left|\lambda_{[i]}\big(X^{\mu}(\epsilon)\big) - \lambda_{[i]}\big(X^{\mu}\big)\right| &< \delta^{\mu}, & i &= 1, \dots, n_{\mathcal{B}} + n_{\mathcal{T}},\\
\left|\lambda_{[j]}\big(S^{\mu}(\epsilon)\big) - \lambda_{[j]}\big(S^{\mu}\big)\right| &< \delta^{\mu}, & j &= 1, \dots, n_{\mathcal{N}},
\end{aligned}$$

which together with (5.17) and (5.18) induces the following bound:

$$\delta^{\mu} > \frac{1 - \big(1 - 2|\epsilon|(\theta^{\mu})^2\left\|\bar{C}\right\|_F\big)}{\theta^{\mu}} \ge \frac{1 - \sqrt{1 - 2|\epsilon|(\theta^{\mu})^2\left\|\bar{C}\right\|_F}}{\theta^{\mu}}, \tag{5.19}$$

where the second inequality in (5.19) follows from $2|\epsilon|(\theta^{\mu})^2\left\|\bar{C}\right\|_F \le 1$. Note that if (5.19) holds, then $\lambda_{[i]}\big(X^{\mu}(\epsilon)\big) > 0$ for $i = n - n_{\mathcal{N}} + 1, \dots, n$ and $\lambda_{[j]}\big(S^{\mu}(\epsilon)\big) > 0$ for $j = n_{\mathcal{N}} + 1, \dots, n$ are immediate from the centrality condition. Consequently, if (5.14) holds, then solution $\big(X^{\mu}(\epsilon), y^{\mu}(\epsilon), S^{\mu}(\epsilon)\big)$ satisfies (5.8), and it is indeed a central solution for the perturbed SDO problem. The proof is complete. □

Condition (5.14) guarantees that $\nabla F_{\mathrm{SD}}\big(X^{\mu}(\epsilon), y^{\mu}(\epsilon), S^{\mu}(\epsilon)\big)$ is nonsingular for every $0 < \mu < \tilde{\mu}$. More specifically, from (5.14) and the Lipschitz continuity of ∇F_{SD} we have

$$\begin{aligned}
&\left\|\nabla F_{\mathrm{SD}}\big(X^{\mu}(\epsilon), y^{\mu}(\epsilon), S^{\mu}(\epsilon)\big) - \nabla F_{\mathrm{SD}}\big(X^{\mu}, y^{\mu}, S^{\mu}\big)\right\|_2\\
&\qquad \le \left\|\big(\mathrm{svec}(X^{\mu}(\epsilon) - X^{\mu});\ y^{\mu}(\epsilon) - y^{\mu};\ \mathrm{svec}(S^{\mu}(\epsilon) - S^{\mu})\big)\right\|_2\\
&\qquad \le \frac{1 - \sqrt{1 - 2|\epsilon|(\theta^{\mu})^2\left\|\bar{C}\right\|_F}}{\theta^{\mu}}\\
&\qquad < \frac{1}{\theta^{\mu}},
\end{aligned}$$

which gives

$$\begin{aligned}&\Big\|\nabla F_{\mathrm{SD}}^{-1}\big(X^\mu,y^\mu,S^\mu\big)\Big(\nabla F_{\mathrm{SD}}\big(X^\mu(\epsilon),y^\mu(\epsilon),S^\mu(\epsilon)\big)-\nabla F_{\mathrm{SD}}\big(X^\mu,y^\mu,S^\mu\big)\Big)\Big\|_2\\ &\qquad \big\|\nabla F_{\mathrm{SD}}^{-1}\big(X^\mu,y^\mu,S^\mu\big)\big\|_2\big\|\nabla F_{\mathrm{SD}}\big(X^\mu(\epsilon),y^\mu(\epsilon),S^\mu(\epsilon)\big)-\nabla F_{\mathrm{SD}}\big(X^\mu,y^\mu,S^\mu\big)\big\|_2<1.\end{aligned}$$

Consequently, it follows from Banach Lemma [39] that

$$\begin{aligned}&\big\|\nabla F_{\mathrm{SD}}^{-1}\big(X^\mu(\epsilon),y^\mu(\epsilon),S^\mu(\epsilon)\big)\big\|_2\\ &\quad\le\frac{\big\|\nabla F_{\mathrm{SD}}^{-1}\big(X^\mu,y^\mu,S^\mu\big)\big\|_2}{1-\Big\|\nabla F_{\mathrm{SD}}^{-1}\big(X^\mu,y^\mu,S^\mu\big)\Big(\nabla F_{\mathrm{SD}}\big(X^\mu(\epsilon),y^\mu(\epsilon),S^\mu(\epsilon)\big)-\nabla F_{\mathrm{SD}}\big(X^\mu,y^\mu,S^\mu\big)\Big)\Big\|_2},\end{aligned}$$

which implies the nonsingularity of $\nabla F_{\mathrm{SD}}\big(X^\mu(\epsilon),y^\mu(\epsilon),S^\mu(\epsilon)\big)$. See also Corollary 1 in [122].

Remark 5.1.7. *Assume that the Jacobian is nonsingular at (X^*,y^*,S^*), i.e., strict complementarity, and both the primal and dual nondegeneracy conditions hold at (X^*,y^*,S^*). Then $\nabla F_{\mathrm{SD}}\big(X^\mu(\epsilon),y^\mu(\epsilon),S^\mu(\epsilon)\big)$ is nonsingular for every $0<\mu<\tilde{\mu}$, see the discussion after Lemma 5.1.16. Furthermore, by the Lipschitz continuity of the Jacobian, Lemma 5.1.7, and the fact that the partition $\mathcal{T}$ does not exist, the right hand side in (5.14) converges to a finite positive value as $\mu\to 0$. Since (X^*,y^*,S^*) is the unique optimal solution, then (5.14) gives a subinterval of a nonlinearity interval at the limit, which contains $\bar{\epsilon}=0$. Consequently, the nonlinearity interval can be estimated using (5.14) and the trajectory of central solutions when $\mu<\tilde{\mu}$. In case that $\theta^\mu\to\infty$, the right hand side of (5.14) converges to 0, providing no information about neither a linearity nor a nonlinearity interval.*

Regardless of the strict complementarity condition, Lemma 5.1.16 gives an estimation of the length of $\mathcal{E}$. The following result is immediate.

Corollary 5.1.3. *The length of $\mathcal{E}$ is bounded below by*

$$\sup_{0<\bar{\mu}<\tilde{\mu}}\min\left\{\frac{\delta^{\bar{\mu}}}{2\theta^{\bar{\mu}}\big\|\bar{C}\big\|_F},\ \frac{1}{2(\theta^{\bar{\mu}})^2\big\|\bar{C}\big\|_F}\right\}.$$

5.1.3.2 Change in $\mathcal{R}\big(Q^\mu_\mathcal{B}\big)$ and $\mathcal{R}\big(Q^\mu_\mathcal{N}\big)$

Using the results in Lemma 5.1.16, we can now derive upper bounds for the sensitivity of $\mathcal{R}\big(Q^\mu_\mathcal{B}\big)$ and $\mathcal{R}\big(Q^\mu_\mathcal{N}\big)$.

Theorem 5.1.3. *Let a central solution $\big(X^\mu, y^\mu, S^\mu\big)$ with $\mu < \tilde{\mu}$ be given. If ϵ is chosen so that*

$$|\epsilon| < \min\left\{\frac{\delta^\mu}{2\theta^\mu\big\|\bar{C}\big\|_F}, \frac{1}{2(\theta^\mu)^2\big\|\bar{C}\big\|_F}, \frac{\phi^\mu}{10\theta^\mu\big\|\bar{C}\big\|_F}\right\} \tag{5.20}$$

holds, then $\epsilon \in \operatorname{int}(\mathcal{E})$, and there exist $V^\mu_\mathcal{B}$ and $V^\mu_\mathcal{N}$ with

$$\|V^\mu_\mathcal{B}\|_F \le \frac{4\Big(1 - \sqrt{1 - 2|\epsilon|(\theta^\mu)^2\big\|\bar{C}\big\|_F}\Big)}{\theta^\mu\phi^\mu},$$
$$\|V^\mu_\mathcal{N}\|_F \le \frac{4\Big(1 - \sqrt{1 - 2|\epsilon|(\theta^\mu)^2\big\|\bar{C}\big\|_F}\Big)}{\theta^\mu\phi^\mu},$$

so that $\mathcal{R}\big(Y^\mu_\mathcal{B}\big)$ and $\mathcal{R}\big(Y^\mu_\mathcal{N}\big)$ are invariant subspaces of $X^\mu(\epsilon)$ and $S^\mu(\epsilon)$, respectively. Furthermore, we have

$$\operatorname{dist}\Big(\mathcal{R}\big(Q^\mu_\mathcal{B}\big), \mathcal{R}\big(Y^\mu_\mathcal{B}\big)\Big) \le \frac{4\Big(1 - \sqrt{1 - 2|\epsilon|(\theta^\mu)^2\big\|\bar{C}\big\|_F}\Big)}{\theta^\mu\phi^\mu},$$
$$\operatorname{dist}\Big(\mathcal{R}\big(Q^\mu_\mathcal{N}\big), \mathcal{R}\big(Y^\mu_\mathcal{N}\big)\Big) \le \frac{4\Big(1 - \sqrt{1 - 2|\epsilon|(\theta^\mu)^2\big\|\bar{C}\big\|_F}\Big)}{\theta^\mu\phi^\mu}. \tag{5.21}$$

Proof. Condition (5.11), after including (5.13) and (5.15), holds if

$$\frac{1 - \sqrt{1 - 2|\epsilon|(\theta^\mu)^2\big\|\bar{C}\big\|_F}}{\theta^\mu} \le \frac{1 - (1 - 2|\epsilon|(\theta^\mu)^2\big\|\bar{C}\big\|_F)}{\theta^\mu} \le \frac{\phi^\mu}{5},$$

which gives the upper bound

$$|\epsilon| \le \frac{\phi^\mu}{10\theta^\mu\big\|\bar{C}\big\|_F}.$$

The upper bounds on $\|V^\mu_\mathcal{B}\|_F$, and on the distance between the subspaces are then immediate from (5.12) and Lemma 5.1.15. The proof for $\|V^\mu_\mathcal{N}\|_F$ and (5.21) is analogous. $\square$

Notice that $\tilde{\mu}$ is actually dependent on ϵ, and it should be denoted by $\tilde{\mu}(\epsilon)$, because the optimal set and thus the condition number σ vary with ϵ. All this hints that $\mathcal{R}(Y^{\mu}_{\mathcal{B}})$ and $\mathcal{R}(Y^{\mu}_{\mathcal{N}})$ cannot be necessarily considered as approximations of $\mathcal{B}(\epsilon)$ and $\mathcal{N}(\epsilon)$, respectively. The reason lies in the fact that the perturbation of the objective vector might give $\mu > \tilde{\mu}(\epsilon)$, which disallows the identification of eigenvectors whose accumulation points form orthonormal bases for the optimal partition, or even ϵ might be a transition point. However, in case that (X^*, y^*, S^*) is unique and strictly complementary, we can provide conditions to ensure that $\mathcal{R}(Y^{\mu}_{\mathcal{B}})$ and $\mathcal{R}(Y^{\mu}_{\mathcal{N}})$ are valid approximations of $\mathcal{B}(\epsilon)$ and $\mathcal{N}(\epsilon)$.

Lemma 5.1.17. *Assume that the Jacobian is nonsingular at (X^*, y^*, S^*). If ϵ satisfies*

$$|\epsilon| < \inf_{0 \le \bar{\mu} < \tilde{\mu}} \min \left\{ \frac{\delta^{\bar{\mu}}}{2\theta^{\bar{\mu}} \|\bar{C}\|_F}, \frac{1}{2(\theta^{\bar{\mu}})^2 \|\bar{C}\|_F}, \frac{\phi^{\bar{\mu}}}{10\theta^{\bar{\mu}} \|\bar{C}\|_F} \right\}, \tag{5.22}$$

then $\mathcal{R}(Y^{\mu}_{\mathcal{B}})$ and $\mathcal{R}(Y^{\mu}_{\mathcal{N}})$, with $\mu < \tilde{\mu}(\epsilon)$, are approximations of $\mathcal{B}(\epsilon)$ and $\mathcal{N}(\epsilon)$.

Proof. Recall from Lemma 5.1.13 and Remark 5.1.7 that the right hand side in (5.22) is positive, and that $\operatorname{rank}\big(X^*(\epsilon)\big) = \operatorname{rank}(X^*)$ and $\operatorname{rank}\big(S^*(\epsilon)\big) = \operatorname{rank}(S^*)$ for all ϵ satisfying (5.22). Furthermore, condition (5.22) guarantees that the central path exists for $(\mathrm{P}^{\epsilon}_{\mathrm{SDO}})$ and $(\mathrm{D}^{\epsilon}_{\mathrm{SDO}})$, and for a sequence $\{\mu_k\} \to 0$ there exists an orthogonal matrix $W^{\mu_k}_{\mathcal{B}}$, defined in (5.9), so that

$$(W^{\mu_k}_{\mathcal{B}})^T (Q^{\mu_k})^T X^{\mu_k}(\epsilon) Q^{\mu_k} W^{\mu_k}_{\mathcal{B}} = D^{\mu_k} := \begin{bmatrix} D^{\mu_k}_{\mathcal{B}} & 0 \\ 0 & D^{\mu_k}_{\mathcal{N}} \end{bmatrix}$$

for sufficiently large k, where $D^{\mu_k}_{\mathcal{B}}$ and $D^{\mu_k}_{\mathcal{N}}$ are positive definite matrices, see Lemma 5.1.16 and Theorem 5.1.3. Since the eigenvalues of $X^{\mu_k}(\epsilon)$ and D^{μ_k} are equal and bounded, then $D^{\mu_k}_{\mathcal{B}}$ has an accumulation point $\tilde{D}_{\mathcal{B}} \succ 0$, and any accumulation point of $D^{\mu_k}_{\mathcal{N}}$ is the zero matrix. Additionally, $W^{\mu_k}_{\mathcal{B}}$, $Y^{\mu_k}_{\mathcal{B}}$, and Q^{μ_k} have accumulation points $\tilde{W}_{\mathcal{B}}$, $\tilde{Y}_{\mathcal{B}}$, and $\tilde{Q}$, since they exist and belong to compact sets. Therefore, we have

$$\mathcal{R}(X^*) = \mathcal{R}\left(\tilde{Q}\tilde{W}_{\mathcal{B}} \begin{bmatrix} \tilde{D}_{\mathcal{B}} & 0 \\ 0 & 0 \end{bmatrix} \tilde{W}_{\mathcal{B}}^T \tilde{Q}^T\right) = \mathcal{R}\left(\tilde{Q}\tilde{W}_{\mathcal{B}} \begin{bmatrix} \tilde{D}_{\mathcal{B}} & 0 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{R}(\tilde{Y}_{\mathcal{B}}\tilde{D}_{\mathcal{B}}) = \mathcal{R}(\tilde{Y}_{\mathcal{B}}),$$

which implies that $\tilde{Y}_{\mathcal{B}}$ forms an orthonormal basis for $\mathcal{B}(\epsilon)$. Analogous results hold for $S^{\mu_k}(\epsilon)$ and $Y_{\mathcal{N}}^{\mu_k}$. Hence, $\mathcal{R}(Y_{\mathcal{B}}^{\mu})$ and $\mathcal{R}(Y_{\mathcal{N}}^{\mu})$ are valid approximations of $\mathcal{B}(\epsilon)$ and $\mathcal{N}(\epsilon)$ when $\mu < \tilde{\mu}(\epsilon)$, see Remark 3.1.7. This completes the proof. $\square$

5.2 Parametric analysis of SOCO

We consider the parametric analysis of a SOCO problem with respect to the perturbation of the objective vector. The primal parametric SOCO problem is phrased as

$$(\mathrm{P}_{\mathrm{SOCO}}^{\epsilon}) \qquad \min\{(c+\epsilon\bar{c})^T x \mid \ Ax = b,\ x \in \mathcal{L}_+^{\bar{n}}\},$$

in which $\bar{c} := (\bar{c}^1; \ldots; \bar{c}^p) \in \mathbb{R}^{\bar{n}}$ is a fixed direction. The dual parametric SOCO problem is given by

$$(\mathrm{D}_{\mathrm{SOCO}}^{\epsilon}) \qquad \max\{b^T y \mid \ A^T y + s = c + \epsilon\bar{c},\ s \in \mathcal{L}_+^{\bar{n}}\}.$$

Associated with $(\mathrm{P}_{\mathrm{SOCO}}^{\epsilon})$ and $(\mathrm{D}_{\mathrm{SOCO}}^{\epsilon})$ we define the primal and dual feasible set mappings as

$$\begin{aligned} \mathcal{P}_{\mathrm{SOCO}}(\epsilon) &:= \{x \mid \ Ax = b,\ x \in \mathcal{L}_+^{\bar{n}}\}, \\ \mathcal{D}_{\mathrm{SOCO}}(\epsilon) &:= \{(y;s) \mid \ A^T y + s = c + \epsilon\bar{c},\ s \in \mathcal{L}_+^{\bar{n}}\}. \end{aligned}$$

Recall that $\mathcal{E}$ denotes the domain of the optimal value function. Analogous to the parametric SDO problem, we rely on the following assumptions. The coefficient matrix A has full row rank, and the interior point condition holds for both $(\mathrm{P}_{\mathrm{SOCO}}^{\epsilon'})$ and

($\mathrm{D}^{\epsilon'}_{\mathrm{SOCO}}$) at all $\epsilon' \in \mathrm{int}(\mathcal{E})$, i.e., there exists a feasible solution $(x^\circ(\epsilon'); y^\circ(\epsilon'); s^\circ(\epsilon'))$ such that $x^{\circ i}(\epsilon'), s^{\circ i}(\epsilon') \in \mathrm{int}(\mathbb{L}^{n_i}_+)$ for all $i = 1, \ldots, p$. Then the primal and dual optimal set mappings are given by

$$\begin{aligned}
\mathcal{P}^*_{\mathrm{SOCO}}(\epsilon) &:= \{x \mid (c + \epsilon\bar{c})^T x = \varphi(\epsilon),\ Ax = b,\ x \in \mathcal{L}^{\bar{n}}_+\},\\
\mathcal{D}^*_{\mathrm{SOCO}}(\epsilon) &:= \{(y, s) \mid b^T y = \varphi(\epsilon),\ A^T y + s = c + \epsilon\bar{c},\ s \in \mathcal{L}^{\bar{n}}_+\}.
\end{aligned}$$

5.2.1 Continuity of optimal solutions for SOCO

The continuity results are analogous to the ones of parametric SDO. For the sake of brevity, we only recall the statements of the lemma and refer the reader to Section 5.1.1. The primal feasible set is invariant with respect to the perturbation, and thus $\mathcal{P}_{\mathrm{SOCO}}(.)$ is continuous at any ϵ. For the dual problem ($\mathrm{D}^{\epsilon}_{\mathrm{SOCO}}$) we define the set-valued mapping $\mathcal{D}^y_{\mathrm{SOCO}} : \mathbb{R} \to \mathbb{R}^m$ as

$$\mathcal{D}^y_{\mathrm{SOCO}}(\epsilon) := \{y \in \mathbb{R}^m \mid c + \epsilon\bar{c} - A^T y \in \mathcal{L}^{\bar{n}}_+\},$$

Since $\mathcal{L}^n_+$ is a closed convex cone, it can be shown that $\mathcal{D}^y_{\mathrm{SOCO}}(.)$ is outer semicontinuous at every $\epsilon \in \mathrm{int}(\mathcal{E})$. Furthermore, we can show that $\mathcal{D}^y_{\mathrm{SOCO}}(.)$ and thus $\mathcal{D}_{\mathrm{SOCO}}(.)$ is inner semicontinuous at every $\epsilon \in \mathrm{int}(\mathcal{E})$.

Lemma 5.2.1. *Let $\bar{y} \in \mathcal{D}^y_{\mathrm{SOCO}}(\epsilon)$ such that $c + \epsilon\bar{c} - A^T\bar{y} \in \mathrm{int}(\mathcal{L}^{\bar{n}}_+)$. Then the set-valued mapping $\mathcal{D}^y_{\mathrm{SOCO}}(.)$ is inner semicontinuous at ϵ.*

Using the interior point condition, we can prove the outer semicontinuity of $\mathcal{P}^*_{\mathrm{SOCO}}(.)$ and $\mathcal{D}^*_{\mathrm{SOCO}}(.)$ at any $\epsilon \in \mathrm{int}(\mathcal{E})$. The proof is immediate from the optimality conditions. An alternative proof is given in Theorem 8 in [82].

Lemma 5.2.2. *The set-valued mapping $\mathcal{P}^*_{\mathrm{SOCO}}(.) \times \mathcal{D}^*_{\mathrm{SOCO}}(.)$ is outer semicontinuous at every $\epsilon \in \mathrm{int}(\mathcal{E})$.*

The optimal set mapping is not necessarily inner semicontinuous, e.g., when either primal or dual nondegeneracy condition fails at ϵ. Nevertheless, a sufficient condition can be given for the continuity of $\mathcal{P}^*_{\text{SOCO}}(.)$ and $\mathcal{D}^*_{\text{SOCO}}(.)$, no matter if the nondegeneracy conditions hold or not.

Lemma 5.2.3. *Assume that the set-valued mapping $\mathcal{P}^*_{\text{SOCO}}(.)$ is single-valued at $\epsilon \in \operatorname{int}(\mathcal{E})$. Then $\mathcal{P}^*_{\text{SOCO}}(.)$ is continuous at ϵ. Analogously, if $\mathcal{D}^*_{\text{SOCO}}(.)$ is single-valued at ϵ, then $\mathcal{D}^*_{\text{SOCO}}(.)$ is continuous at ϵ.*

5.2.2 Sensitivity of the optimal partition for SOCO

In this section, we investigate the sensitivity and the stability of the optimal partition for SOCO while ϵ runs through $\operatorname{int}(\mathcal{E})$. More specifically, we are interested in the characterization of a nonlinearity interval and a transition point.

Let $\pi_{\text{SOCO}}(\epsilon) := (\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{R}(\epsilon), \mathcal{T}(\epsilon))$ denote the optimal partition of $(\mathrm{P}^{\epsilon}_{\text{SOCO}})$ and $(\mathrm{D}^{\epsilon}_{\text{SOCO}})$ at a given ϵ. Analogous to SDO, the case for a parametric SOCO may not be as easy as LO, mostly due to the existence of $\mathcal{R}(\epsilon)$. In fact, the interval $\mathcal{E}$ might contain a subinterval where both the optimal set and the optimal partition change with ϵ while the index sets stay unchanged. This can be actually demonstrated in the following example.

Example 5.2.1. *Consider the following parametric SOCO problem:*

$$
\begin{aligned}
\min \ & -\epsilon x^1_2 - (1-\epsilon) x^1_3 \\
\text{s.t.} \ & \qquad\qquad x^1_1 = 1, \\
& \qquad x^1_3 - x^2_1 = 0, \\
& \qquad x^1_2 - x^2_2 = 1, \\
& \qquad\qquad x^1_1 \geq \sqrt{(x^1_2)^2 + (x^1_3)^2}, \\
& \qquad\qquad x^2_1 \geq |x^2_2|,
\end{aligned}
$$

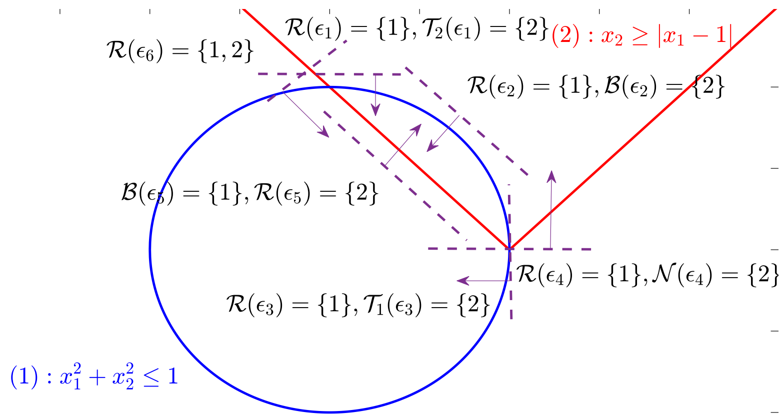

Figure 5.2: Illustration of the optimal partition for a parametric SOCO problem.

where $\mathcal{E} = \mathbb{R}$. *One can check that the interior point condition holds for every* $\epsilon' \in \operatorname{int}(\mathcal{E})$. *On the interval* $(-1, 2)$ *the optimal partition is given by*

$$
\begin{aligned}
&(\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{R}(\epsilon), \mathcal{T}(\epsilon)) = (\emptyset, \emptyset, \{1, 2\}, (\emptyset, \emptyset, \emptyset)), && \epsilon \in (-1, 0),\\
&(\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{R}(\epsilon), \mathcal{T}(\epsilon)) = (\emptyset, \emptyset, \{1\}, (\emptyset, \{2\}, \emptyset)), && \epsilon = 0,\\
&(\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{R}(\epsilon), \mathcal{T}(\epsilon)) = (\{2\}, \emptyset, \{1\}, (\emptyset, \emptyset, \emptyset)), && \epsilon \in (0, 1),\\
&(\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{R}(\epsilon), \mathcal{T}(\epsilon)) = (\emptyset, \emptyset, \{1\}, (\{2\}, \emptyset, \emptyset)), && \epsilon = 1,\\
&(\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{R}(\epsilon), \mathcal{T}(\epsilon)) = (\emptyset, \emptyset, \{1, 2\}, (\emptyset, \emptyset, \emptyset)), && \epsilon \in (1, 2).
\end{aligned}
$$

As demonstrated in Figure 5.2, on the interval $(0, 1)$ *both the primal and the dual optimal set mappings are single-valued, they change continuously, and the strict complementarity holds. Furthermore, the optimal partitions on* $(-1, 0)$ *and* $(1, 2)$ *are constant. Note that the linearity and nonlinearity intervals are separated at* $\epsilon = 0, 1$ *which are transition points. At* $\epsilon = 0$*, both primal and dual nondegeneracy conditions hold, but strict complementarity fails. At* $\epsilon = 1$*, the strict complementarity fails and* $\mathcal{D}^*_{\text{SOCO}}(.)$ *is multiple-valued.*

We should note that even the continuity of primal and dual optimal set mappings does not give a complete characterization of a nonlinearity interval. For instance, in Figure 5.2 both the primal and the dual optimal set mappings are continuous at ϵ_1. However, ϵ_1 is a transition point.

In a linearity interval of SOCO, the index sets of $\pi_{\text{SOCO}}(.)$ remain unchanged. Further, there exists a unique Jordan frame associated with each $i \in \mathcal{R}(.) \cup \mathcal{T}_2(.) \cup \mathcal{T}_3(.)$. Let $\bar{\epsilon}$ belong to a linearity interval $(\alpha_{\text{lin}}, \beta_{\text{lin}})$. Then the extreme points α_{lin} and β_{lin}, analogous to the auxiliary problems in Lemma 5.1.10, can be computed by solving a pair of auxiliary SOCO problems:

$$
\begin{aligned}
\alpha_{\text{lin}}(\beta_{\text{lin}}) := \min_{\epsilon}(\max_{\epsilon}) \quad & \epsilon \\
\text{s.t.} \quad & \sum_{i \in \mathcal{B}(\bar{\epsilon}) \cup \mathcal{T}(\bar{\epsilon})} A_i x^i + \sum_{i \in \mathcal{R}(\bar{\epsilon})} \theta_i A_i x^{*i}(\bar{\epsilon}) = b, \\
& A_i^T y = c^i + \epsilon \bar{c}^i, \quad i \in \mathcal{B}(\bar{\epsilon}), \\
& A_i^T y + \tau_i s^{*i}(\bar{\epsilon}) = c^i + \epsilon \bar{c}^i, \quad i \in \mathcal{R}(\bar{\epsilon}), \\
& A_i^T y + s^i = c^i + \epsilon \bar{c}^i, \quad i \in \mathcal{T}(\bar{\epsilon}), \\
& x^i \in \mathbb{L}_+^{n_i}, \quad i \in \mathcal{B}(\bar{\epsilon}) \cup \mathcal{T}(\bar{\epsilon}), \\
& s^i \in \mathbb{L}_+^{n_i}, \quad i \in \mathcal{N}(\bar{\epsilon}) \cup \mathcal{T}(\bar{\epsilon}), \\
& \theta_i, \tau_i \geq 0, \quad i \in \mathcal{R}(\bar{\epsilon}).
\end{aligned}
$$

A transition point and a nonlinearity interval for SOCO is formally defined as follows.

Definition 5.2.1. *The point $\bar{\epsilon} \in \text{int}(\mathcal{E})$ is called a transition point if for every $\varsigma > 0$ there exists $\epsilon' \in (\bar{\epsilon} - \varsigma, \bar{\epsilon} + \varsigma) \subseteq \text{int}(\mathcal{E})$ such that $\pi_{\text{SOCO}}(\bar{\epsilon}) \neq \pi_{\text{SOCO}}(\epsilon')$, i.e., the index sets in $\pi_{\text{SOCO}}(\bar{\epsilon})$ and $\pi_{\text{SOCO}}(\epsilon')$ are not identical. Furthermore, $\mathcal{I}_{\text{non}}$ is called a nonlinearity interval if both the optimal set and the optimal partition change with ϵ while the index sets of $\pi(\epsilon')$ and $\pi(\epsilon'')$ are identical for any two $\epsilon', \epsilon'' \in \mathcal{I}_{\text{non}}$.*

Equivalently, in a nonlinearity interval the rank of $L(x^{*i}(\epsilon))$ and $L(s^{*i}(\epsilon))$ is stable for all $i = 1, \ldots, p$.

Analogous to parametric SDO, the openness of a nonlinearity interval does not directly follow from the definition. In fact, a nonlinearity interval is a connected component of ϵ which contains $\bar{\epsilon}$ and for which the solution set of

$$
\begin{aligned}
\sum_{i \in \mathcal{B}(\bar{\epsilon}) \cup \mathcal{R}(\bar{\epsilon}) \cup \mathcal{T}_2(\bar{\epsilon})} A_i x^i &= b, && \\
A_i^T y &= c^i + \epsilon \bar{c}^i, && i \in \mathcal{B}(\bar{\epsilon}) \cup \mathcal{T}_1(\bar{\epsilon}) \cup \mathcal{T}_2(\bar{\epsilon}), \\
A_i^T y + s^i &= c^i + \epsilon \bar{c}^i, && i \in \mathcal{R}(\bar{\epsilon}) \cup \mathcal{T}_3(\bar{\epsilon}), \\
x^i \circ s^i &= 0, && i \in \mathcal{R}, \\
x^i &\in \operatorname{int}(\mathbb{L}_+^{n_i}), && i \in \mathcal{B}(\bar{\epsilon}), \\
s^i &\in \operatorname{int}(\mathbb{L}_+^{n_i}), && i \in \mathcal{N}(\bar{\epsilon}), \\
x_1^i &> 0, && i \in \mathcal{R}(\bar{\epsilon}) \cup \mathcal{T}_2(\bar{\epsilon}), \\
s_1^i &> 0, && i \in \mathcal{N}(\bar{\epsilon}) \cup \mathcal{T}_3(\bar{\epsilon})
\end{aligned}
$$

is nonempty. If both $\mathcal{P}^*_{\text{SOCO}}(.)$ and $\mathcal{D}^*_{\text{SOCO}}(.)$ are continuous at every $\epsilon' \in \mathcal{I}_{\text{non}}$, then $\mathcal{I}_{\text{non}}$ is an open interval. The proof is analogous to Corollary 5.1.1.

The continuity arguments can be applied to prove some inclusion properties for the optimal partition. More generally, we can use the conditions in Remark 5.1.2. The following technical lemma is in order.

Lemma 5.2.4. *If $\mathcal{P}^*_{\text{SOCO}}(.)$ is continuous at $\bar{\epsilon}$, then $\mathcal{B}(\epsilon) \subseteq \mathcal{B}(\epsilon')$ for all ϵ' in a neighborhood of $\bar{\epsilon}$. If $\mathcal{D}^*_{\text{SOCO}}(.)$ is continuous at $\bar{\epsilon}$, then $\mathcal{N}(\epsilon) \subseteq \mathcal{N}(\epsilon')$ for all ϵ' in a neighborhood of $\bar{\epsilon}$. Finally, if both $\mathcal{P}^*_{\text{SOCO}}(.)$ and $\mathcal{D}^*_{\text{SOCO}}(.)$ are continuous at $\bar{\epsilon}$, then $\mathcal{R}(\epsilon) \subseteq \mathcal{R}(\epsilon')$ holds as well.*

Proof. Let $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ be a maximally complementary solution. By the inner semicontinuity of the primal and dual optimal set mappings at $\bar{\epsilon}$, for ϵ' sufficiently close to $\bar{\epsilon}$ there exist primal optimal solution $x(\epsilon')$ and dual optimal solution

$(y(\epsilon'); s(\epsilon'))$ close to $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ so that

$$\begin{aligned} x_1^i(\epsilon') - \|x_{2:n_i}^i(\epsilon')\|_2 > 0, \qquad & i \in \mathcal{B}(\bar{\epsilon}), \\ s_1^i(\epsilon') - \|s_{2:n_i}^i(\epsilon')\|_2 > 0, \qquad & i \in \mathcal{N}(\bar{\epsilon}), \end{aligned}$$

which implies that $\mathcal{B}(\bar{\epsilon}) \subseteq \mathcal{B}(\epsilon')$ and $\mathcal{N}(\bar{\epsilon}) \subseteq \mathcal{N}(\epsilon')$. If $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ is continuous at $\bar{\epsilon}$, then we have $x_1^i(\epsilon') > 0$ and $s_1^i(\epsilon') > 0$ for $i \in \mathcal{R}(\bar{\epsilon})$, which by $x^i(\epsilon') \circ s^i(\epsilon') = 0$, implies that $i \in \mathcal{R}(\epsilon')$. This completes the proof. $\square$

The necessity of continuity of both primal and dual optimal set mappings in Lemma 5.2.4 can be illustrated by the following parametric SOCO problem:

$$\begin{aligned} \min \quad & (1-2\epsilon)x_2^1 - x_3^1 \\ \text{s.t.} \quad & x_1^1 = 1, \\ & x_1^2 = 2, \\ & x_2^2 - x_2^1 = 0, \\ & x_3^2 - 2x_3^1 = 0, \\ & x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2}, \\ & x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}. \end{aligned} \tag{5.23}$$

At $\epsilon = 0$ the primal optimal set mapping is single-valued, while the dual optimal set mapping is multiple-valued. However, we can observe that $\mathcal{R}(\epsilon) = \{1\}$ for all $\epsilon \neq 0$ while $\mathcal{R}(0) = \{1, 2\}$. Interestingly, for any sequence $\epsilon_k \to 0$ we have $\liminf_{k\to\infty} \mathcal{D}^*_{\text{SOCO}}(\epsilon_k) \cap \text{ri}(\mathcal{D}^*_{\text{SOCO}}(0)) = \emptyset$, see Figure 5.3. This example also shows that the index sets of $\pi_{\text{SOCO}}(.)$ could be the same around a transition point.

5.2.3 Stability of regularity conditions

By the definition of a nonlinearity interval, the strict complementarity is stable in a nonlinearity interval. However, this may not hold at a transition point.

Lemma 5.2.5. *The strict complementarity condition is not necessarily stable around a transition point.*

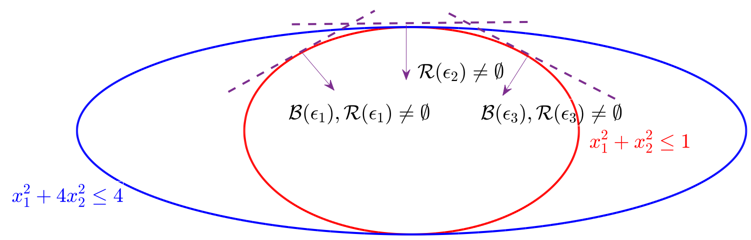

Figure 5.3: Discontinuity of the dual optimal set mapping at a transition point.

Proof. The following counterexample can be given:

$$\begin{array}{ll}
\min & -\epsilon x_2^1 + (1-\epsilon)x_3^1 \\
\text{s.t.} & x_1^1 = 1, \\
& x_4^1 - x_1^2 = -1, \\
& x_2^1 - x_2^2 = 0, \\
& x_3^1 - x_3^2 = 0, \\
& x_3^1 - x_4^1 - x_1^3 = 0, \\
& x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2 + (x_4^1)^2}, \\
& x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}, \\
& x_1^3 \geq 0.
\end{array}$$

For any $\epsilon \neq 0$ the optimal partition is given by $(\emptyset, \emptyset, \{1\}, (\emptyset, \{2\}, \emptyset))$. However, at $\bar{\epsilon} = 0$ we have a transition point and there exists a strictly complementary solution. $\square$

If strict complementarity fails at a given $\bar{\epsilon} \in \text{int}(\mathcal{E})$, then the nondegeneracy conditions do not necessarily hold in a neighborhood of $\bar{\epsilon}$, since the optimal partition might change. Nevertheless, the stability is valid when both primal and dual nondegeneracy conditions simultaneously hold at $\bar{\epsilon}$.

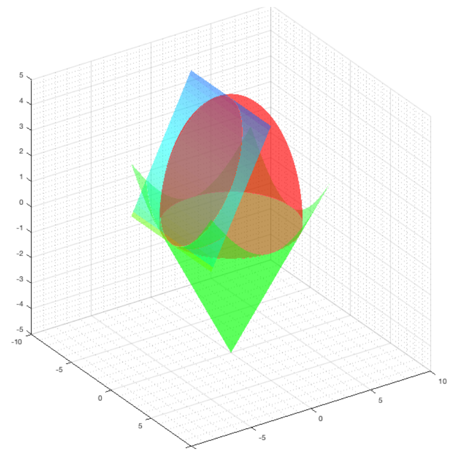

Figure 5.4: Instability of the strict complementarity condition.

Lemma 5.2.6. *Assume that both the primal and the dual nondegeneracy conditions hold for* $(\mathrm{P}_{\mathrm{SOCO}}^{\bar{\epsilon}})$ *and* $(\mathrm{D}_{\mathrm{SOCO}}^{\bar{\epsilon}})$. *Then the nondegeneracy conditions hold in a sufficiently small neighborhood of* $\bar{\epsilon}$.

Proof. By the primal and the dual nondegeneracy conditions at $\bar{\epsilon}$, there exists a unique optimal solution $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$. Since both $\mathcal{P}_{\mathrm{SOCO}}^*(.)$ and $\mathcal{D}_{\mathrm{SOCO}}^*(.)$ are continuous at $\bar{\epsilon}$, for every $\epsilon_k \to \bar{\epsilon}$ and with sufficiently large k there exists a sequence of optimal solutions $(x(\epsilon_k); y(\epsilon_k); s(\epsilon_k))$ converging to $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$. Furthermore, $\bar{P}_i(\epsilon_k)$ can be made arbitrary close to $\bar{P}_i^*(\bar{\epsilon})$ for sufficiently large k. Note that the eigenvalues of $L(x^{*i}(\bar{\epsilon}))$ are given by $x_1^{*i}(\bar{\epsilon}) + \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2$, $x_1^{*i}(\bar{\epsilon}) - \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2$, and $x_1^{*i}(\bar{\epsilon})$ with multiplicity $n_i - 2$. Since $i \in \mathcal{R}(\bar{\epsilon})$, we have one zero eigenvalue and $x_{2:n_i}^{*i}(\bar{\epsilon}) \neq 0$. The eigenvectors associated with $x_1^{*i}(\bar{\epsilon}) + \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2$ and $x_1^{*i}(\bar{\epsilon}) - \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2$ are given by

$$\frac{\sqrt{2}}{2}\left(1; \frac{x_{2:n_i}^{*i}(\bar{\epsilon})}{\|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2}\right), \qquad \frac{\sqrt{2}}{2}\left(1; \frac{-x_{2:n_i}^{*i}(\bar{\epsilon})}{\|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2}\right),$$

respectively. Further, the eigenvectors corresponding to $x_1^{*i}(\bar{\epsilon})$ can be computed by solving the equation system

$$\begin{pmatrix} 0 & \left(x^{*i}_{2:n_i}(\bar\epsilon)\right)^T \\ x^{*i}_{2:n_i}(\bar\epsilon) & 0 \end{pmatrix} \eta = 0,$$

$$\|\eta\|_2 = 1,$$

which can be simplified to

$$\begin{cases} \eta^T_{2:n_i} x^{*i}_{2:n_i}(\bar\epsilon) & = 0 \\ \eta_1 x^{*i}_j(\bar\epsilon) & = 0, \qquad j = 2, \ldots, n_i, \\ \|\eta\|_2 = 1. \end{cases}$$

As $x^{*i}_{2:n_i}(\bar\epsilon) \neq 0$, we have $\eta_1 = 0$. Thus, the eigenvectors belong to the boundary of an $n_i - 2$ dimensional ball, which is the intersection of an $n_i - 1$ dimensional ball and the $n_i - 2$ dimensional subspace. Therefore, a slight change in $x^{*i}(\bar\epsilon)$ also results in a slight change in $\bar P^*_i(\bar\epsilon)$.

Recall from the primal nondegeneracy condition at $\bar\epsilon$ that

$$\left(\left(A_i \bar P^*_i(\bar\epsilon)\right)_{i \in \mathcal{R}(\bar\epsilon) \cup \mathcal{T}_2(\bar\epsilon)},\ A_{\mathcal{B}(\bar\epsilon)} \right) \tag{5.24}$$

has full row rank. By Lemma 5.2.4 we have $\mathcal{B}(\bar\epsilon) \subseteq \mathcal{B}(\epsilon_k)$ and $\mathcal{R}(\bar\epsilon) \subseteq \mathcal{R}(\epsilon_k)$ for sufficiently large k. Furthermore, the continuity of the primal optimal set implies that if $i \in \mathcal{T}_2(\bar\epsilon)$, then we would have $x^i_1(\epsilon_k) > 0$ for sufficiently large k. As a consequence, for sufficiently large k, the matrix

$$\left((A_i \bar P_i(\epsilon_k))_{i \in \mathcal{R}(\bar\epsilon) \cup \mathcal{T}_2(\bar\epsilon) \setminus \{J_k \cup J'_k\}},\ (A_i \bar P_i(\epsilon_k))_{i \in J'_k},\ (A_i \bar P_i(\epsilon_k))_{i \in J_k},\ A_{\mathcal{B}(\bar\epsilon)} \right), \tag{5.25}$$

where

$$J_k := \left\{ i \in \mathcal{T}_2(\bar\epsilon) \mid x^i_1(\epsilon_k) > \|x^i_{2:n_i}(\epsilon_k)\|_2 \right\},$$
$$J'_k := \left\{ i \in \mathcal{T}_2(\bar\epsilon) \mid x^i_1(\epsilon_k) = \|x^i_{2:n_i}(\epsilon_k)\|_2,\ x^i_1(\epsilon_k) > 0 \right\},$$

has full row rank, since (5.25) is obtained from a slight perturbation of (5.24). On the other hand, $x(\epsilon_k)$ is primal nondegenerate if and only if

$$\big((A_i\bar{P}_i(\epsilon_k))_{i\in\mathcal{R}(\bar{\epsilon})\cup\mathcal{T}_2(\bar{\epsilon})\setminus\{J_k\cup J_k'\}},\ (A_i\bar{P}_i(\epsilon_k))_{i\in J_k'},\ A_{\mathcal{B}(\bar{\epsilon})\cup J_k}\big), \tag{5.26}$$

has full row rank. Since (5.25) has full row rank, matrix (5.26) has to be of full row rank too.

The proof for the dual nondegeneracy condition is analogous. The dual nondegeneracy condition at $\bar{\epsilon}$ holds if

$$\big((A_iR_i(s^{*i}(\bar{\epsilon}))_{i\in\mathcal{R}(\bar{\epsilon})\cup\mathcal{T}_3(\bar{\epsilon})},\ A_{\mathcal{B}(\bar{\epsilon})\cup\mathcal{T}_1(\bar{\epsilon})\cup\mathcal{T}_2(\bar{\epsilon})}\big) \tag{5.27}$$

has linearly independent columns. By Lemma 5.2.4, if $i\in\mathcal{T}_1(\bar{\epsilon})$, then i could belong to any partition for ϵ sufficiently close to $\bar{\epsilon}$. However, if $i\in\mathcal{T}_3(\bar{\epsilon})$, then we would get $s_1^i(\epsilon_k)\neq 0$ for sufficiently large k. Then for sufficiently large k the matrix

$$\big((A_iR_is^i(\epsilon_k))_{i\in\mathcal{R}(\bar{\epsilon})\cup\mathcal{T}_3(\bar{\epsilon})\setminus\{K_k\cup K_k'\}},\ (A_iR_is^i(\epsilon_k))_{i\in K_k\cup K_k'},\ A_{\mathcal{B}(\bar{\epsilon})\cup\mathcal{T}_1(\bar{\epsilon})\cup\mathcal{T}_2(\bar{\epsilon})}\big) \tag{5.28}$$

has linearly independent columns, where

$$\begin{aligned}
I_k &:= \big\{i\in\mathcal{T}_1(\bar{\epsilon}) \mid s^i(\epsilon_k)=0\big\},\\
I_k' &:= \big\{i\in\mathcal{T}_1(\bar{\epsilon}) \mid s_1^i(\epsilon_k)=\|s_{2:n_i}^i(\epsilon_k)\|_2,\ s_1^i(\epsilon_k)>0\big\},\\
I_k'' &:= \big\{i\in\mathcal{T}_1(\bar{\epsilon}) \mid s_1^i(\epsilon_k)>\|s_{2:n_i}^i(\epsilon_k)\|_2\big\},\\
J_k'' &:= \big\{i\in\mathcal{T}_2(\bar{\epsilon}) \mid s_1^i(\epsilon_k)=\|s_{2:n_i}^i(\epsilon_k)\|_2,\ s_1^i(\epsilon_k)>0\big\},\\
J_k''' &:= \big\{i\in\mathcal{T}_2(\bar{\epsilon}) \mid s^i(\epsilon_k)=0\big\},\\
K_k &:= \big\{i\in\mathcal{T}_3(\bar{\epsilon}) \mid s_1^i(\epsilon_k)>\|s_{2:n_i}^i(\epsilon_k)\|_2\big\},\\
K_k' &:= \big\{i\in\mathcal{T}_3(\bar{\epsilon}) \mid s_1^i(\epsilon_k)=\|s_{2:n_i}^i(\epsilon_k)\|_2\big\}.
\end{aligned}$$

This obviously holds, because (5.28) and (5.27) are almost identical, and $s(\epsilon_k)$ is sufficiently close to $s^*(\bar{\epsilon})$. Now, $(y(\epsilon_k);s(\epsilon_k))$ is dual nondegenerate if and only if

$$\big((A_iR_is^i(\epsilon_k))_{i\in\mathcal{R}(\bar{\epsilon})\cup\mathcal{T}_3(\bar{\epsilon})\setminus\{K_k\cup K_k'\}},\ (A_iR_is^i(\epsilon_k))_{i\in I_k'\cup J_k''\cup K_k'},\ A_{\mathcal{B}(\bar{\epsilon})\cup\mathcal{T}_1(\bar{\epsilon})\cup\mathcal{T}_2(\bar{\epsilon})\setminus\{I_k'\cup I_k''\cup J_k''\}}\big) \tag{5.29}$$

has linearly independent columns. Note that (5.29) is obtained from (5.28) by removing the columns $A_{I''_k}$ and $(A_i R_i s^i(\epsilon_k))_{i \in K_k}$ and by replacing the columns $A_{I'_k \cup J''_k}$ by $(A_i R_i s^i(\epsilon_k))_{i \in I'_k \cup J''_k}$. Then it is immediate that (5.29) should be full column rank, since otherwise a subset of the columns in (5.28) would be linearly dependent.

Consequently, we have shown that there exists a sequence of optimal solutions

$$(x(\epsilon_k); y(\epsilon_k); s(\epsilon_k)) \to (x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$$

which are primal-dual nondegenerate for sufficiently large k. This completes the proof. □

As a result of Lemma 5.2.6, if both the primal and dual nondegeneracy conditions hold at $\bar{\epsilon}$, then both the primal and dual optimal set mappings are single-valued and the second-order sufficient condition holds at $(x^*(\epsilon'); y^*(\epsilon'); s^*(\epsilon'))$ for all ϵ' in a sufficiently small neighborhood of $\bar{\epsilon}$.

Remark 5.2.1. *Even if both the primal and dual nondegeneracy conditions hold at a given transition point, strict complementarity might still fail at a neighboring interval. For example, it can be verified from the parametric SOCO problem*

$$\begin{array}{rr} \min & -\epsilon x_2^1 + (1-\epsilon) x_3^1 \\ \text{s.t.} & x_1^1 = 1, \\ & x_1^2 - x_4^1 = 1, \\ & x_2^1 - x_2^2 = 0, \\ & x_3^1 - x_3^2 = 0, \\ & x_2^1 - x_1^3 = 0, \\ & x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2 + (x_4^1)^2}, \\ & x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}, \\ & x_1^3 \geq 0, \end{array}$$

that $\bar{\epsilon} = 0$ is a transition point, and both the primal and dual nondegeneracy conditions hold at $\bar{\epsilon} = 0$. However, strict complementarity is absent at every $\epsilon \neq 0$.

Note that the optimal partition is $\mathcal{R}(\epsilon) = \{1\}$, $\mathcal{T}_2(\epsilon) = \{2\}$, and $\mathcal{T}_1(\epsilon) = \{3\}$. In a nonlinearity interval next to the transition point $\epsilon = 0$, the partition $\mathcal{T}_2(\epsilon) = \{2\}$ is stable.

Besides the stability of strict complementarity and primal-dual nondegeneracy conditions, the stability of the elements of $\mathcal{T}(\epsilon)$ can be questionable. We already observed from Lemma 5.2.5 that the subsets $\mathcal{T}_2(\epsilon)$ and $\mathcal{T}_3(\epsilon)$, due to the symmetry between the primal and dual problems, might be nonempty in a nonlinearity interval. However, we conjecture that this is not the case for $\mathcal{T}_1(\epsilon)$.

Conjecture 5.2.1. *For every nonlinearity interval $\mathcal{I}_{\text{non}} \subseteq \operatorname{int}(\mathcal{E})$ we have $\mathcal{T}_1(.) = \emptyset$.*

In other words, if the conjecture is true and if $\mathcal{T}_1(\epsilon) \neq \emptyset$, then ϵ must either belong to a linearity interval or it must be a transition point.

5.2.4 Continuity of solutions in a nonlinearity interval

The behavior of the optimal partition in a nonlinearity interval can be described using continuity arguments under strict complementarity condition. This can be considered as a special case of Lemma 5.1.12.

Lemma 5.2.7. *Assume that $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ is a strictly complementary solution, and both $\mathcal{P}^*_{\text{SOCO}}(.)$ and $\mathcal{D}^*_{\text{SOCO}}(.)$ are continuous at $\bar{\epsilon}$. Then $\bar{\epsilon}$ belongs to the interior of a nonlinearity interval.*

Proof. By the continuity of $\mathcal{P}^*_{\text{SOCO}}(.)$ and $\mathcal{D}^*_{\text{SOCO}}(.)$ and the strict complementarity condition, it holds that $x^*(\epsilon') + s^*(\epsilon') \in \operatorname{int}(\mathcal{L}^n_+)$ for all ϵ' in a sufficiently small neighborhood of $\bar{\epsilon}$. Then by Lemma 5.2.4, the index sets remain unchanged. □

Remark 5.2.2. *Assume that $\mathcal{I}_{\text{non}}$ is an open nonlinearity interval. Since $\pi_{\text{SOCO}}(\epsilon')$ is determined by a maximally complementary optimal solution $(x^*(\epsilon'); y^*(\epsilon'); s^*(\epsilon'))$, for every $\epsilon' \in \mathcal{I}_{\text{non}}$ we have*

$$
\begin{aligned}
x_1^{*i}(\epsilon') - \|x_{2:n_i}^{*i}(\epsilon')\|_2 &> 0, && i \in \mathcal{B}(\epsilon'),\\
s_1^{*i}(\epsilon') - \|s_{2:n_i}^{*i}(\epsilon')\|_2 &> 0, && i \in \mathcal{N}(\epsilon'),\\
x_1^{*i}(\epsilon') &> 0, && i \in \mathcal{R}(\epsilon') \cup \mathcal{T}_2(\epsilon'),\\
s_1^{*i}(\epsilon') &> 0, && i \in \mathcal{R}(\epsilon') \cup \mathcal{T}_3(\epsilon'),\\
x^{*i}(\epsilon') &= 0, && i \in \mathcal{T}_1(\epsilon') \cup \mathcal{T}_3(\epsilon'),\\
s^{*i}(\epsilon') &= 0, && i \in \mathcal{T}_1(\epsilon') \cup \mathcal{T}_2(\epsilon').
\end{aligned}
$$

Hence, in the nonlinearity interval $\mathcal{I}_{\text{non}}$ there is no change in $s^{*i}(\epsilon)$ for $i \in \mathcal{T}_1(\epsilon) \cup \mathcal{T}_2(\epsilon)$ or $x^{*i}(\epsilon)$ for $i \in \mathcal{T}_1(\epsilon) \cup \mathcal{T}_3(\epsilon)$, since otherwise ϵ would be a transition point. Let $\hat{\epsilon}$ be an extreme point of $\mathcal{I}_{\text{non}}$. Recall that near $\hat{\epsilon}$ both the primal and dual optimal set mappings are uniformly bounded. Then for any $\hat{\epsilon} \neq \epsilon_k \to \hat{\epsilon}$ the sequence $(x(\epsilon_k); y(\epsilon_k); s(\epsilon_k))$ has an accumulation point $(\hat{x}; \hat{y}; \hat{s}) \in \mathcal{P}^*_{\text{SOCO}}(\bar{\epsilon}) \times \mathcal{D}^*_{\text{SOCO}}(\bar{\epsilon})$ such that at least one of the following holds:

$$
\begin{aligned}
\hat{x}_1^{i} - \|\hat{x}_{2:n_i}^{i}\|_2 &= 0, && \textit{for some } i \in \mathcal{B}(\hat{\epsilon}),\\
\hat{s}_1^{i} - \|\hat{s}_{2:n_i}^{i}\|_2 &= 0, && \textit{for some } i \in \mathcal{N}(\hat{\epsilon}),\\
\hat{x}^{i} &= 0, && \textit{for some } i \in \mathcal{R}(\hat{\epsilon}) \cup \mathcal{T}_2(\hat{\epsilon}),\\
\hat{s}^{i} &= 0, && \textit{for some } i \in \mathcal{R}(\hat{\epsilon}) \cup \mathcal{T}_3(\hat{\epsilon}).
\end{aligned}
$$

Now, assume that both the primal and the dual optimal set mappings are continuous at $\hat{\epsilon}$. If $\lambda_{\min}(x^{*i}(\epsilon_k))$ has no positive accumulation point for a given $i \in \mathcal{B}(\hat{\epsilon})$, then either $i \in \mathcal{T}_1(\hat{\epsilon})$ or $i \in \mathcal{T}_2(\hat{\epsilon})$ holds. Analogously, if every accumulation point of $\lambda_{\min}(s^{*i}(\epsilon_k))$ is zero for an $i \in \mathcal{N}(\hat{\epsilon})$, then we have $i \in \mathcal{T}_1(\hat{\epsilon})$ or $i \in \mathcal{T}_3(\hat{\epsilon})$.

Using the same arguments for SDO, a nonlinearity interval can be more specifically characterized by relying on the nonsingularity of the Jacobian ∇F_{SO} which is given in (4.45). If at a given $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ the Jacobian is nonsingular, then by Lemma 4.2.9 and the implicit function theorem, see Theorem A.4.3, there exists $\varsigma > 0$ and a unique continuously differentiable mapping $(x^*(.); y^*(.); s^*(.))$ on $(\bar{\epsilon}-\varsigma, \bar{\epsilon}+\varsigma)$ such that $\nabla F_{\text{SO}}((x^*(.); y^*(.); s^*(.)))$ is nonsingular on $(\bar{\epsilon}-\varsigma, \bar{\epsilon}+\varsigma)$. In fact, ς can be made small enough such that the strict complementarity and the nondegeneracy conditions hold. Therefore, by Lemma 5.2.7, $(\bar{\epsilon}-\varsigma, \bar{\epsilon}+\varsigma)$ is a subinterval of a nonlinearity interval $\mathcal{I}_{\text{non}}$.

Now, let $\mathcal{I}_{\text{inv}}$ denote the open interval of maximal length containing $\bar{\epsilon}$ such that $\nabla F_{\text{SO}}((x^*(\epsilon'); y^*(\epsilon'); s^*(\epsilon')))$ is singular for all $\epsilon' \in \mathcal{I}_{\text{inv}}$. If the strict complementarity condition fails at every extreme point of $\mathcal{I}_{\text{inv}}$, then $\mathcal{I}_{\text{inv}}$ coincides with the nonlinearity interval $\mathcal{I}_{\text{non}}$, since at least one of $\operatorname{rank}(L(x^*(.))$ or $\operatorname{rank}(s^*(.))$ changes at every extreme point. The proof is analogous to Lemma 5.1.13. On the other hand, if the strict complementarity condition holds at an extreme point $\hat{\epsilon}$, then at least one of the primal or dual nondegeneracy conditions must fail at $\hat{\epsilon}$. In this case, either $\mathcal{P}^*_{\text{SOCO}}(.)$ or $\mathcal{D}^*_{\text{SOCO}}(.)$ is no longer continuous at $\hat{\epsilon}$. In Lemma 5.1.14, we showed the possibility of this case for a parametric SDO problem. However, we conjecture that $\pi_{\text{SOCO}}(.)$ changes at every extreme point of $\mathcal{I}_{\text{inv}}$.

Conjecture 5.2.2. *Assume that the strict complementarity and the nondegeneracy conditions hold at $\bar{\epsilon}$. Then the open interval $\mathcal{I}_{\text{inv}}$ coincides with the nonlinearity interval $\mathcal{I}_{\text{non}}$ containing $\bar{\epsilon}$.*

If Conjecture 5.2.2 is true, then both the primal and dual optimal set mappings are continuous in a nonlinearity interval, under the conditions of Lemma 4.2.9.

As a result of Lemma 5.2.7, at a transition point $\bar{\epsilon}$ at least one of $\mathcal{P}^*_{\text{SOCO}}(.)$ or $\mathcal{D}^*_{\text{SOCO}}(.)$ has to be discontinuous. Failure of a nondegeneracy condition in the presence of strict complementarity condition signals the discontinuity of primal or dual optimal sets. In general, we have the following result.

Corollary 5.2.1. *At a transition point $\bar{\epsilon}$, at least one of the strict complementarity, the primal nondegeneracy, or the dual nondegeneracy conditions fails.*

Proof. If all the three conditions hold, then ∇F_{SO} is nonsingular at the unique optimal solution. As a consequence, the implicit function theorem is applicable: $(x^*(.); y^*(.); s^*(.))$ is continuously differentiable in a sufficiently small neighborhood of $\bar{\epsilon}$, and thus the result follows from Lemma 5.2.7. □

The converse of the statement in Corollaries 5.1.2 and 5.2.1 is not necessarily true. The following parametric SOCO problem is a counterexample:

$$\begin{aligned}
\min \quad & -\epsilon x_2^1 - (1-\epsilon)x_3^1 \\
\text{s.t.} \quad & x_1^1 = 5, \\
& x_1^2 - x_4^1 = 5, \\
& x_2^2 - x_2^1 = 0, \\
& x_3^2 - x_3^1 = 0, \\
& x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2 + (x_4^1)^2}, \\
& x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2},
\end{aligned} \tag{5.30}$$

where $\mathcal{E} = \mathbb{R}$. For this problem strict complementarity fails at any given ϵ', but both the primal and dual nondegeneracy conditions hold. It can be verified that on $\mathcal{E}$ both the primal and the dual optimal solutions are unique, continuous, and their ranks are stable, i.e., there is no transition point, see Figure 5.5. While this is a SOCO problem, it can be obviously represented as a parametric SDO problem.

Remark 5.2.3. *In contrast to LO and LCQO, where the transition points and non-differentiable points of the optimal value function coincide, see [16], for SDO and SOCO the optimal value function might be infinitely many times differentiable at a transition point. For instance, SOCO problem (5.23) has a strictly complementary solution at $\epsilon = \frac{1}{2}$, and the primal optimal solution is unique. On the intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ both $x^*(.)$ and $s^*(.)$ are unique and continuous, and they have identical*

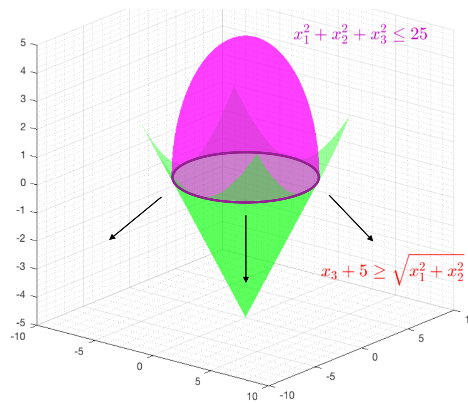

Figure 5.5: Strict complementarity always fails while there is no transition point.

ranks. The optimal value function is given by $\varphi(\epsilon) = -\sqrt{(\epsilon+1)^2+\epsilon^2}$*. At* $\epsilon = \frac{1}{2}$ *there is a transition point, and all the higher order derivatives of the optimal value function exist.*

5.2.5 Computation of a nonlinearity interval

In what follows, we propose an auxiliary problem to compute a subinterval of a nonlinearity interval under strict complementarity condition. To that end, let $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ be a strictly complementary optimal solution. Let us define

$$
\begin{aligned}
\delta_{\mathcal{B}(\bar{\epsilon})} &:= \frac{\sqrt{2}}{2} \min_{i\in\mathcal{B}(\bar{\epsilon})} \{(x_1^{*i}(\bar{\epsilon}) - \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2\},\\
\delta_{\mathcal{N}(\bar{\epsilon})} &:= \frac{\sqrt{2}}{2} \min_{i\in\mathcal{N}(\bar{\epsilon})} \{s_1^{*i}(\bar{\epsilon}) - \|s_{2:n_i}^{*i}(\bar{\epsilon})\|_2\},\\
\delta_{\mathcal{R}(\bar{\epsilon})} &:= \min\left\{ \min_{i\in\mathcal{R}(\bar{\epsilon})} \{x_1^{*i}(\bar{\epsilon})\},\ \min_{i\in\mathcal{R}(\bar{\epsilon})} \{s_1^{*i}(\bar{\epsilon})\}\right\},\\
\delta(\bar{\epsilon}) &:= \min\{\delta_{\mathcal{B}(\bar{\epsilon})}, \delta_{\mathcal{N}(\bar{\epsilon})}, \delta_{\mathcal{R}(\bar{\epsilon})}\}.
\end{aligned}
\tag{5.31}
$$

Then the following lemma is in order.

Lemma 5.2.8. *Let $\bar{\epsilon}$ belong to a nonlinearity interval. Assume that $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ is a strictly complementary optimal solution, and let α_{non} and β_{non} be defined by*

$$\begin{aligned}
\alpha_{\text{non}}(\beta_{\text{non}}) := \min(\max) \quad & \epsilon \\
\text{s.t.} \quad & \sum_{i\in\mathcal{B}(\bar{\epsilon})\cup\mathcal{R}(\bar{\epsilon})} A_i x^i = b, \\
& A_i^T y = c^i + \epsilon \bar{c}^i, \quad i \in \mathcal{B}(\bar{\epsilon}), \\
& A_i^T y + s^i = c^i + \epsilon \bar{c}^i, \quad i \in \mathcal{N}(\bar{\epsilon}) \cup \mathcal{R}(\bar{\epsilon}), \\
& x^i \circ s^i = 0, \qquad i \in \mathcal{R}(\bar{\epsilon}), \\
& \|x - x^*(\bar{\epsilon})\|_2^2 \le \frac{1}{4}\delta^2(\bar{\epsilon}), \\
& \|s - s^*(\bar{\epsilon})\|_2^2 \le \frac{1}{4}\delta^2(\bar{\epsilon}).
\end{aligned} \tag{5.32}$$

If $\alpha_{\text{non}} < \beta_{\text{non}}$, then $[\alpha_{\text{non}}, \beta_{\text{non}}]$ is a subinterval of the nonlinearity interval which contains $\bar{\epsilon}$.

Proof. The last two constraints ensure that every feasible solution of (5.32) is strictly complementary. In fact, we have

$$\begin{aligned}
\|x^i(\epsilon) - x^{*i}(\bar{\epsilon})\|_2 &< \delta(\bar{\epsilon}), & i \in \mathcal{B}(\bar{\epsilon}), \\
\|s^i(\epsilon) - s^{*i}(\bar{\epsilon})\|_2 &< \delta(\bar{\epsilon}), & i \in \mathcal{N}(\bar{\epsilon}).
\end{aligned}$$

Then, using the equality

$$\begin{aligned}
&\Big(\big(x_1^i(\epsilon) - x_1^{*i}(\bar{\epsilon}) - \|x_{2:n_i}^i(\epsilon) - x_{2:n_i}^{*i}(\bar{\epsilon})\|_2\big)^2 + \big(x_1^i(\epsilon) - x_1^{*i}(\bar{\epsilon}) + \|x_{2:n_i}^i(\epsilon) - x_{2:n_i}^{*i}(\bar{\epsilon})\|_2\big)^2\Big)^{\frac{1}{2}} \\
&\qquad = \sqrt{2}\big\|x^i(\epsilon) - x^{*i}(\bar{\epsilon})\big\|_2,
\end{aligned}$$

we can conclude that

$$\Big|x_1^i(\epsilon) - x_1^{*i}(\bar{\epsilon}) - \big\|x_{2:n_i}^i(\epsilon) - x_{2:n_i}^{*i}(\bar{\epsilon})\big\|_2\Big| \le \sqrt{2}\|x^i(\epsilon) - x^{*i}(\bar{\epsilon})\|_2 < \sqrt{2}\delta(\bar{\epsilon}).$$

All this gives

$$\begin{aligned}
\|x_{2:n_i}^i(\epsilon)\|_2 - \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2 - x_1^i(\epsilon) + x_1^{*i}(\bar{\epsilon}) &\le \|x_{2:n_i}^i(\epsilon) - x_{2:n_i}^{*i}(\bar{\epsilon})\|_2 - x_1^i(\epsilon) + x_1^{*i}(\bar{\epsilon}) \\
&< \sqrt{2}\delta(\bar{\epsilon}),
\end{aligned}$$

which implies that

$$0 < x_1^{*i}(\bar{\epsilon}) - \|x_{2:n_i}^{*i}(\bar{\epsilon})\|_2 - \sqrt{2}\delta < x_1^i(\epsilon) - \|x_{2:n_i}^i(\epsilon)\|_2, \qquad i \in \mathcal{B}(\epsilon). \tag{5.33}$$

In a similar manner, we can show that

$$s_1^i(\epsilon) - \|s_{2:n_i}^i(\epsilon)\|_2 > 0, \qquad i \in \mathcal{N}(\epsilon). \tag{5.34}$$

Furthermore, for every $i \in \mathcal{R}(\epsilon)$ it follows from the complementarity condition $x^i(\epsilon) \circ s^i(\epsilon) = 0$ that

$$\begin{aligned} 0 = x(\epsilon)^T s(\epsilon) = x_1(\epsilon)s_1(\epsilon) + (x_{2:n_i}^i(\epsilon))^T s_{2:n_i}^i(\epsilon) &= x_1^i(\epsilon)s_1^i(\epsilon) - \frac{s_1^i(\epsilon)x_{2:n_i}^i(\epsilon)}{x_1^i(\epsilon)} \\ &= \frac{s_1^i(\epsilon)\big((x_1^i(\epsilon))^2 - \|x_{2:n_i}^i(\epsilon)\|_2^2\big)}{x_1^i(\epsilon)}, \end{aligned}$$

which gives $x_1^i(\epsilon) - \|x_{2:n_i}^i(\epsilon)\|_2 = 0$ and analogously, $s_1^i(\epsilon) - \|s_{2:n_i}^i(\epsilon)\|_2 = 0$ for every $i \in \mathcal{R}(\epsilon)$. Hence, it remains to show that $x_1^i(\epsilon) > 0$ and $s_1^i(\epsilon) > 0$ for every $i \in \mathcal{R}(\epsilon)$. To that end, we have

$$|x_1^i(\epsilon) - x_1^{*i}(\bar{\epsilon})| \le \|x^i(\epsilon) - x^{*i}(\bar{\epsilon})\|_2 < \delta(\bar{\epsilon}) \implies x_1^i(\epsilon) > x_1^{*i}(\bar{\epsilon}) - \delta(\bar{\epsilon}) > 0, \tag{5.35}$$

$$|s_1^i(\epsilon) - s_1^{*i}(\bar{\epsilon})| \le \|s^i(\epsilon) - s^{*i}(\bar{\epsilon})\|_2 < \delta(\bar{\epsilon}) \implies s_1^i(\epsilon) > s_1^{*i}(\bar{\epsilon}) - \delta(\bar{\epsilon}) > 0. \tag{5.36}$$

Thus, we have shown that $\pi_{\text{SOCO}}(\bar{\epsilon})$ is weakly identical with $\pi_{\text{SOCO}}(\epsilon')$ for every ϵ' which yields a nonempty solution set and also belong to $[\alpha_{\text{non}}, \beta_{\text{non}}]$. We claim that $[\alpha_{\text{non}}, \beta_{\text{non}}]$ does not contain any transition point. When the index sets of $\pi_{\text{SOCO}}(\epsilon')$ are distinct for all $\epsilon' \in \text{int}(\mathcal{E})$, then the proof is immediate. However, in contrast to the case of LO and LCQO, we may have identical index sets around a transition point. Hence, it remains to show that the latter case does not happen either. Note that the intersection of $\|x^i(\epsilon) - x^{*i}(\bar{\epsilon})\|_2 < \delta(\bar{\epsilon})$ and $\text{bd}(\mathbb{L}_+^{n_i})$ is a connected set for every $i \in \mathcal{R}(\bar{\epsilon})$, since the ball $\|x^i(\epsilon) - x^{*i}(\bar{\epsilon})\|_2 < \delta(\bar{\epsilon})$ does contain the origin of a second-order cone. Analogously, the intersection of $\|s^i(\epsilon) - s^{*i}(\bar{\epsilon})\|_2 < \delta(\bar{\epsilon})$ and

$\mathrm{bd}(\mathbb{L}_+^{n_i})$ is a connected set for every $i \in \mathcal{R}(\bar{\epsilon})$. Consequently, the intersection of the balls $\|x - x^*(\bar{\epsilon})\|_2^2 \le \frac{1}{4}\delta^2(\bar{\epsilon})$ and $\|s - s^*(\bar{\epsilon})\|_2^2 \le \frac{1}{4}\delta^2(\bar{\epsilon})$ with the boundary of primal and dual feasible sets are connected as well. As a result, if $\bar{\epsilon}$ belongs to a nonlinearity interval, then by (5.33) to (5.36), there is no transition point in $[\alpha_{\mathrm{non}}, \beta_{\mathrm{non}}]$. This completes the proof. □

Corollary 5.2.2. *Assume that $\bar{\epsilon}$ is a transition point, and let $(x^*(\bar{\epsilon}); y^*(\bar{\epsilon}); s^*(\bar{\epsilon}))$ be a strictly complementary optimal solution. Then $\alpha_{\mathrm{non}} = \beta_{\mathrm{non}}$.*

Proof. If $\bar{\epsilon}$ is a transition point, then there is no other transition point in $[\alpha_{\mathrm{non}}, \beta_{\mathrm{non}}]$, since the feasible set of (5.32) is connected. □

Remark 5.2.4. *If the nondegeneracy and the strict complementarity conditions hold, then the result of Nayakkankuppam and Overton [122] for SDO can be specialized to estimate a subinterval of the nonlinearity interval. This is in fact the application of Kantorovich theorem to the system of optimality conditions, which gives*

$$|\epsilon - \bar{\epsilon}| < \min\left\{\frac{\delta(\bar{\epsilon})}{2\theta^\mu \left\|\bar{C}\right\|_F}, \ \frac{1}{2(\theta^\mu)^2 \left\|\bar{C}\right\|_F}\right\}.$$

5.2.6 Failure of strict complementarity

Even if strict complementarity fails, there might exist a nonlinearity interval. However, the Jacobian ∇F_{SO} has to be singular in this case, and thus the implicit function theorem is not applicable. Additionally, formulating a problem analogous to (5.32) may not produce the correct bounds when the strict complementarity condition fails. As an example, the following SOCO problem can be given:

$$\begin{aligned}
\min \quad & \epsilon c + (1-\epsilon)c' \\
\text{s.t.} \quad & x_1^1 = 1, \\
x_3^1 - & x_1^2 = a, \\
x_2^1 - & x_2^2 = a+1, \\
& x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2}, \\
& x_1^2 \geq |x_2^2|,
\end{aligned}$$

where

$$\begin{aligned}
c &:= \left(0,\ -2a-1+\sqrt{1-4a-4a^2},\ -2a-1-\sqrt{1-4a-4a^2},\ 0,\ 0\right)^T, \\
c' &:= \left(0,\ -2a-1-\sqrt{1-4a-4a^2},\ -2a-1+\sqrt{1-4a-4a^2},\ 0,\ 0\right)^T.
\end{aligned}$$

For all $0 < a < (\sqrt{2}-1)/2$ and $0 < \epsilon < 1$, the optimal solution is primal-dual nondegenerate, and it satisfies the strict complementarity condition. At $\bar{\epsilon} = 0$ the primal and dual optimal solutions are given by

$$\begin{aligned}
x^a(0) &= \Big(1,\ (2a+1-\sqrt{1-4a-4a^2})/2,\ (2a+1+\sqrt{1-4a-4a^2})/2, \\
&\qquad (1+\sqrt{1-4a-4a^2})/2,\ (-1-\sqrt{1-4a-4a^2})/2\Big)^T, \\
y^a(0) &= (-2,\ 0,\ 0)^T, \\
s^a(0) &= \left(2,\ -2a-1-\sqrt{1-4a-4a^2},\ -2a-1+\sqrt{1-4a-4a^2}\right)^T.
\end{aligned}$$

For all $0 < a < (\sqrt{2}-1)/2$ the optimal partition is given by

$$\begin{aligned}
\pi^a(0) &= (\emptyset, \emptyset, \{1\}, \{2\}), \\
\pi^a(\epsilon) &= (\{2\}, \emptyset, \{1\}, \emptyset), \quad 0 < \epsilon < 1, \\
\pi^a(1) &= (\emptyset, \emptyset, \{1\}, \{2\}),
\end{aligned}$$

i.e., $\pi^a(0) \stackrel{w}{=} \pi^a(1)$ for all $0 < a < (\sqrt{2}-1)/2$. However,

$$\operatorname{dist}\Big(\big(x^a(0); y^a(0); s^a(0)\big),\ \big(x^a(1); y^a(1); s^a(1)\big)\Big) \to 0$$

as $a \to (\sqrt{2}-1)/2$, while $\delta(.)$ stays positive.

Consider ($\mathrm{P}^{\epsilon}_{\mathrm{SOCO}}$) and ($\mathrm{D}^{\epsilon}_{\mathrm{SOCO}}$), and assume that the primal and dual nondegeneracy conditions hold at $\epsilon \in \mathrm{int}(\mathcal{E})$. Recall from Section 4.2.3 that the unique primal optimal solution $x^*(\epsilon)$ can be obtained by solving

$$
\begin{aligned}
(\mathrm{P}^{\epsilon}_{\mathrm{NLO}}) \quad \min \quad & \sum_{i \in \mathcal{B}(\epsilon) \cup \mathcal{R}(\epsilon) \cup \mathcal{T}_2(\epsilon)} (c^i + \epsilon \bar{c}^i)^T \nu^i \\
\text{s.t.} \quad & \sum_{i \in \mathcal{B}(\epsilon) \cup \mathcal{R}(\epsilon) \cup \mathcal{T}_2(\epsilon)} A_i \nu^i = b, \\
& (\nu^i)^T R_i \nu^i = 0, \qquad i \in \mathcal{R}(\epsilon) \cup \mathcal{T}_2(\epsilon), \\
& \nu \in \mathcal{V},
\end{aligned}
$$

where $\nu^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{B}(\epsilon) \cup \mathcal{R}(\epsilon) \cup \mathcal{T}_2(\epsilon)$, and $\mathcal{V}$ is given by

$$
\mathcal{V} = \left\{ \nu \mid \nu_1^i > 0,\ i \in \mathcal{R}(\epsilon) \cup \mathcal{T}_2(\epsilon), \quad \nu^i \in \mathrm{int}(\mathbb{L}_+^{n_i}),\ i \in \mathcal{B}(\epsilon) \right\}.
$$

In a similar manner, the unique optimal solution $(y^*(\epsilon); s^*(\epsilon))$ is the globally optimal solution of

$$
\begin{aligned}
(\mathrm{D}^{\epsilon}_{\mathrm{NLO}}) \quad \min \quad & -b^T w \\
\text{s.t.} \quad & A_i^T w = c^i + \epsilon \bar{c}^i, && i \in \mathcal{B}(\epsilon) \cup \mathcal{T}_1(\epsilon) \cup \mathcal{T}_2(\epsilon), \\
& A_i^T w + z^i = c^i + \epsilon \bar{c}^i, && i \in \mathcal{R}(\epsilon) \cup \mathcal{N}(\epsilon) \cup \mathcal{T}_3(\epsilon), \\
& (z^i)^T R_i z^i = 0, && i \in \mathcal{R}(\epsilon) \cup \mathcal{T}_3(\epsilon), \\
& z \in \mathcal{W},
\end{aligned}
$$

where $w \in \mathbb{R}^m$, $z^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{R}(\epsilon) \cup \mathcal{N}(\epsilon)$, and $\mathcal{W}$ is given by

$$
\mathcal{W} = \left\{ z \mid\ z_1^i > 0,\ i \in \mathcal{R}(\epsilon) \cup \mathcal{T}_3(\epsilon), \quad z^i \in \mathrm{int}(\mathbb{L}_+^{n_i}),\ i \in \mathcal{N}(\epsilon) \right\}.
$$

Since $(y^*(\epsilon); s^*(\epsilon))$ is unique, then ($\mathrm{D}^{\epsilon}_{\mathrm{NLO}}$) has a unique globally optimal solution $(w^*(\epsilon); z^*(\epsilon))$. Analogous to Section 4.2.3, let $u^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{B}(\epsilon) \cup \mathcal{T}(\epsilon) \cup \mathcal{R}(\epsilon) \cup$

$\mathcal{N}(\epsilon)$ and $v \in \mathbb{R}^{|\mathcal{R}(\epsilon)|+|\mathcal{T}_3(\epsilon)|}$ be the Lagrange multipliers associated with the constraints in $(\mathrm{D}^{\epsilon}_{\mathrm{NLO}})$. The first-order optimality conditions for $(\mathrm{D}^{\epsilon}_{\mathrm{NLO}})$ are given by

$$\begin{aligned}
-\sum_{i\in\mathcal{B}(\epsilon)\cup\mathcal{N}(\epsilon)\cup\mathcal{R}(\epsilon)\cup\mathcal{T}_2(\epsilon)} A_i u^i &= b, && \\
-u^i - 2v_i R_i z^i &= 0, && i \in \mathcal{R}(\epsilon)\cup\mathcal{T}_3(\epsilon),\\
-u^i &= 0, && i \in \mathcal{N}(\epsilon),\\
A_i^T w &= c^i + \epsilon\bar{c}^i, && i \in \mathcal{B}(\epsilon)\cup\mathcal{T}_1(\epsilon)\cup\mathcal{T}_2(\epsilon),\\
A_i^T w + z^i &= c^i + \epsilon\bar{c}^i, && i \in \mathcal{R}(\epsilon)\cup\mathcal{N}(\epsilon)\cup\mathcal{T}_3(\epsilon),\\
(z^i)^T R_i z^i &= 0, && i \in \mathcal{R}(\epsilon)\cup\mathcal{T}_3(\epsilon),\\
z &\in \mathcal{W},
\end{aligned} \tag{5.37}$$

Note that u is the concatenation of the column vectors u^i for $i \in \mathcal{B}(\epsilon)\cup\mathcal{T}_2(\epsilon)\cup\mathcal{R}(\epsilon)\cup\mathcal{N}(\epsilon)$. Then for $z^*(\epsilon) \in \mathcal{W}$ there exist Lagrange multipliers $u^*(\epsilon)$ and $v^*(\epsilon)$ so that $\vartheta^*(\epsilon) := (w^*(\epsilon); z^*(\epsilon); u^*(\epsilon); v^*(\epsilon))$ satisfies the first-order optimality conditions (5.37). Such a solution at ϵ is given by

$$\begin{aligned}
w^*(\epsilon) &:= y^*(\epsilon), && \\
z^{*i}(\epsilon) &:= s^{*i}(\epsilon), && i \in \mathcal{R}(\epsilon)\cup\mathcal{N}(\epsilon)\cup\mathcal{T}_3(\epsilon),\\
u^{*i}(\epsilon) &:= -x^{*i}(\epsilon), && i \in \mathcal{B}(\epsilon)\cup\mathcal{T}_2(\epsilon)\cup\mathcal{R}(\epsilon)\cup\mathcal{N}(\epsilon),\\
v_i^*(\epsilon) &:= \frac{1}{2}\frac{x_1^{*i}(\epsilon)}{s_1^{*i}(\epsilon)}, && i \in \mathcal{R}(\epsilon)\cup\mathcal{T}_3(\epsilon).
\end{aligned} \tag{5.38}$$

We redefine the mapping $G : \mathbb{R}^{\bar{n}_c} \times \mathbb{R} \to \mathbb{R}^{\bar{n}_c}$ as follows

$$G(\vartheta,\epsilon) := \begin{pmatrix}
-\sum_{i\in\mathcal{B}(\epsilon)\cup\mathcal{T}_2(\epsilon)\cup\mathcal{R}(\epsilon)\cup\mathcal{N}(\epsilon)} A_i u^i - b & \\
-u^i - 2v_i R_i z^i & i \in \mathcal{R}(\epsilon)\cup\mathcal{T}_3(\epsilon)\\
-u^i & i \in \mathcal{N}(\epsilon)\\
A_i^T w - c^i - \epsilon\bar{c}^i & i \in \mathcal{B}(\epsilon)\cup\mathcal{T}_1(\epsilon)\cup\mathcal{T}_2(\epsilon)\\
A_i^T w + z^i - c^i - \epsilon\bar{c}^i & i \in \mathcal{R}(\epsilon)\cup\mathcal{N}(\epsilon)\cup\mathcal{T}_3(\epsilon)\\
(z^i)^T R_i z^i & i \in \mathcal{R}(\epsilon)\cup\mathcal{T}_3(\epsilon)
\end{pmatrix},$$

in which

$$\bar{n}_c = \sum_{i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}} n_i + \sum_{i \in \mathcal{R} \cup \mathcal{N}} n_i + |\mathcal{R}| + m.$$

Then Lemmas 4.2.4, 4.2.5, and 4.2.6 can be applied to prove the following theorem.

Theorem 5.2.1. *Assume that the primal and dual nondegeneracy conditions hold at $\bar{\epsilon}$. Then the Jacobian of the equality constraints in $(\mathrm{D}^{\bar{\epsilon}}_{\mathrm{NLO}})$ has full row rank, and the second-order sufficient condition holds at $(w^*(\bar{\epsilon}); z^*(\bar{\epsilon}))$. Furthermore, ∇G is nonsingular at $\vartheta^*(\bar{\epsilon})$.* $\square$

By Lemma 5.2.4 there exists $\varsigma > 0$ so that for all $\epsilon' \in (\bar{\epsilon} - \varsigma, \bar{\epsilon} + \varsigma)$ we have

$$\mathcal{B}(\bar{\epsilon}) \subseteq \mathcal{B}(\epsilon'), \qquad \mathcal{N}(\bar{\epsilon}) \subseteq \mathcal{N}(\epsilon'), \qquad \mathcal{R}(\bar{\epsilon}) \subseteq \mathcal{R}(\epsilon').$$

Hence, we only need to investigate the behavior of the second-order cones belonging to $\mathcal{T}(\epsilon')$ to find the optimal partition in the vicinity of the given $\bar{\epsilon}$. The continuity arguments imply that for ϵ' sufficiently close to $\bar{\epsilon}$ we have

$$\begin{aligned} i \in \mathcal{T}_1(\bar{\epsilon}) \quad &\Rightarrow \quad i \in \mathcal{T}(\epsilon'), \\ i \in \mathcal{T}_2(\bar{\epsilon}) \quad &\Rightarrow \quad i \in \mathcal{B}(\epsilon') \cup \mathcal{R}(\epsilon') \cup \mathcal{T}_2(\epsilon'), \\ i \in \mathcal{T}_3(\bar{\epsilon}) \quad &\Rightarrow \quad i \in \mathcal{N}(\epsilon') \cup \mathcal{R}(\epsilon') \cup \mathcal{T}_3(\epsilon'). \end{aligned}$$

Now, we can apply the analytic version of the implicit function theorem to the mapping $G(., \epsilon)$.

Theorem 5.2.2. *Suppose that both the primal and dual nondegeneracy conditions hold at $\bar{\epsilon}$, and $\mathcal{T}_1(\bar{\epsilon}) = \emptyset$. If $\bar{\epsilon}$ belongs to a nonlinearity interval, then all the higher order derivatives of $v_i^*(\bar{\epsilon})$ for all $i \in \mathcal{T}_3(\bar{\epsilon})$ are equal to 0.*

Proof. By Lemma 5.2.6, there exists $\varsigma > o$ so that $(x^*(\epsilon'); y^*(\epsilon'); s^*(\epsilon'))$ is a unique primal-dual optimal solution for every $\epsilon' \in (\bar{\epsilon} - \varsigma, \bar{\epsilon} + \varsigma)$. Furthermore, the mapping (5.38) gives the globally optimal solution of $(\mathrm{D}^{\epsilon'}_{\mathrm{NLO}})$ for every $\epsilon' \in (\bar{\epsilon} - \varsigma, \bar{\epsilon} + \varsigma)$.

If $d^k v_i^*(.)/d\epsilon^k \geq 0$ for all $i \in \mathcal{T}_3(\bar{\epsilon})$ with strict inequality for some i, then from (5.37) we can find an optimal solution $(x(\epsilon'); y(\epsilon'); s(\epsilon'))$ for $(\mathrm{P}^{\epsilon'}_{\mathrm{SOCO}})$ and $(\mathrm{D}^{\epsilon'}_{\mathrm{SOCO}})$, in which $x_1^i(\epsilon'), s_1^i(\epsilon') \neq 0$ for some $i \in \mathcal{T}_3(\bar{\epsilon})$. All this implies that $\pi_{\mathrm{SOCO}}(\epsilon') \neq \pi_{\mathrm{SOCO}}(\bar{\epsilon})$.

If $d^k v_i^*(.)/d\epsilon^k < 0$ for some k and some $i \in \mathcal{T}_3(\bar{\epsilon})$, then we claim that $\pi_{\mathrm{SOCO}}(\epsilon') \neq \pi_{\mathrm{SOCO}}(\bar{\epsilon})$. On the contrary, suppose that $\pi_{\mathrm{SOCO}}(\epsilon') = \pi_{\mathrm{SOCO}}(\bar{\epsilon})$. Since ∇G is nonsingular at $\vartheta^*(\bar{\epsilon})$, by the implicit function theorem there exists $\varsigma' > 0$ and a unique analytic mapping $\chi(.)$ so that $\chi(\epsilon')$ is a KKT solution of $(\mathrm{D}^{\epsilon'}_{\mathrm{NLO}})$ for all $(\epsilon' \in \bar{\epsilon} - \varsigma', \bar{\epsilon} + \varsigma')$ and $\vartheta^*(\bar{\epsilon}) = \chi(\bar{\epsilon})$. Since $\mathcal{P}^*_{\mathrm{SOCO}}(.) \times \mathcal{D}^*_{\mathrm{SOCO}}(.)$ is a continuous set-valued mapping at $\bar{\epsilon}$, the mapping $\vartheta^*(.)$ given by (5.38) is continuous at $\bar{\epsilon}$. Therefore, there exists $\varsigma'' > 0$ so that the analytic mapping $\chi(.)$ and the continuous mapping $\vartheta^*(.)$ coincide on $(\bar{\epsilon} - \varsigma'', \bar{\epsilon} + \varsigma'')$. However, by $x_1^{*i}(\epsilon') = 0$ for all $i \in \mathcal{T}_3(\epsilon')$ and every $\epsilon' \in (\bar{\epsilon} - \varsigma'', \bar{\epsilon} + \varsigma'')$, we have that all the higher order derivatives of $v_i(\bar{\epsilon})$ are 0 for $i \in \mathcal{T}_3(\bar{\epsilon})$, which is a contradiction. $\square$

Theorem 5.2.2 only provides a partial characterization for a transition point of the optimal under primal and dual nondegeneracy conditions. We can get a complete characterization by incorporating the derivative information from the KKT conditions of $(\mathrm{P}^{\bar{\epsilon}}_{\mathrm{NLO}})$. Recall from Lemma 4.2.2 that the second-order sufficient condition holds at the unique solution $x^*(\bar{\epsilon})$ when the dual nondegeneracy condition holds. On the other hand, we can apply the same technique as in Lemma 4.2.4 to show that under the primal nondegeneracy condition, the Jacobian of the equality constraints of $(\mathrm{P}^{\bar{\epsilon}}_{\mathrm{NLO}})$ has full row rank at the unique globally optimal solution. All this implies that the Jacobian of the KKT conditions is nonsingular, and the derivatives of the Lagrange multipliers can be computed. Consequently, we can determine whether $i \in \mathcal{T}_2(\epsilon)$, $i \in \mathcal{R}(\epsilon)$, or $i \in \mathcal{B}(\epsilon)$ for every $i \in \mathcal{T}_2(\bar{\epsilon})$ and every ϵ sufficiently close to $\bar{\epsilon}$.

5.3 Extension to LCO

Our derivations and continuity results for the parametric SDO and SOCO can be naturally extended to LCO using the facial description of the optimal partition presented in Section 1.5. This generalization enables us to study the continuity of optimal set mapping in a nonlinearity interval from the lens of convex geometry.

Let a parametric LCO problem be given by

$$\begin{aligned} &(\mathrm{P}^{\epsilon}_{\mathrm{LCO}}) \quad \min_{x} \big\{ \langle c + \epsilon \bar{c}, x \rangle \mid \mathcal{A}x = b,\ x \in \mathcal{K} \big\}, \\ &(\mathrm{D}^{\epsilon}_{\mathrm{LCO}}) \quad \max_{(y,s)} \big\{ b^T y \mid \mathcal{A}^* y + s = c + \epsilon \bar{c},\ s \in \mathcal{K} \big\}, \end{aligned}$$

where $\bar{c} \in \mathbb{V}$ is a fixed direction, and the rest of parameters are defined in Section 1.2. Recall that $\mathcal{K}$ has to be self-dual, since otherwise the notion of the optimal partition is not well-defined. The primal and dual feasible set mappings are defined as

$$\begin{aligned} \mathcal{P}_{\mathrm{LCO}}(\epsilon) &:= \{ x \mid \mathcal{A}x = b,\ x \in \mathcal{K} \}, \\ \mathcal{D}_{\mathrm{LCO}}(\epsilon) &:= \big\{ (y,s) \mid \mathcal{A}^* y + s = c + \epsilon \bar{c},\ s \in \mathcal{K} \big\}. \end{aligned}$$

We have the same assumptions as in the SDO and SOCO cases. More specifically, we assume that Assumption 1.2.1 and the interior point condition hold for all $\epsilon' \in \mathrm{int}(\mathcal{E})$, i.e., there exists a feasible solution $\big(x^{\circ}(\epsilon'), y^{\circ}(\epsilon'), s^{\circ}(\epsilon')\big)$ such that $x^{\circ}(\epsilon') \in \mathrm{int}(\mathcal{K})$ and $s^{\circ}(\epsilon') \in \mathrm{int}(\mathcal{K})$. Then the primal and dual optimal set mappings are defined as

$$\begin{aligned} \mathcal{P}^*_{\mathrm{LCO}}(\epsilon) &:= \{ x \mid \langle c + \epsilon \bar{c}, x \rangle = \varphi(\epsilon),\ x \in \mathcal{P}_{\mathrm{LCO}}(\epsilon) \}, \\ \mathcal{D}^*_{\mathrm{LCO}}(\epsilon) &:= \big\{ (y,s) \mid b^T y = \varphi(\epsilon),\ (y,s) \in \mathcal{D}_{\mathrm{LCO}}(\epsilon) \big\}. \end{aligned}$$

Therefore, for every $\epsilon \in \mathcal{E}$ there exists a maximally complementary solution, and every $\big(x(\epsilon), y(\epsilon), s(\epsilon)\big) \in \mathcal{P}^*_{\mathrm{LCO}}(\epsilon) \times \mathcal{D}^*_{\mathrm{LCO}}(\epsilon)$ satisfies

$$\begin{aligned} \mathcal{A}x &= b, & x &\in \mathcal{K}, \\ \mathcal{A}^*y + s &= c + \epsilon\bar{c}, & s &\in \mathcal{K}^*, \\ \langle x, s\rangle &= 0. \end{aligned}$$

Let us define $\pi_{\mathrm{LCO}}(\epsilon) := \big(\mathcal{F}_{x^*(\epsilon)}, \mathcal{F}_{s^*(\epsilon)}, \mathcal{G}(\epsilon)\big)$ as the optimal partition of ($\mathrm{P}^{\epsilon}_{\mathrm{LCO}}$) and ($\mathrm{D}^{\epsilon}_{\mathrm{LCO}}$) at a given ϵ. The notions of a transition point and a linearity interval in SDO and SOCO can be extended to LCO in a word-for-word fashion. The two optimal partitions $\pi_{\mathrm{LCO}}(\epsilon')$ and $\pi_{\mathrm{LCO}}(\epsilon'')$ are called identical if $\pi_{\mathrm{LCO}}(\epsilon') = \pi_{\mathrm{LCO}}(\epsilon'')$, i.e.,

$$\mathcal{F}_{x^*(\epsilon')} = \mathcal{F}_{x^*(\epsilon'')}, \qquad \mathcal{F}_{s^*(\epsilon')} = \mathcal{F}_{s^*(\epsilon'')}.$$

For a nonlinearity interval, however, we substitute the rank by the dimensions of minimal faces of $\mathcal{K}$ which contain the primal and dual optimal sets. More specifically, $\pi_{\mathrm{LCO}}(\epsilon')$ and $\pi_{\mathrm{LCO}}(\epsilon'')$ are called weakly identical if

$$\dim\big(\mathcal{F}_{x^*(\epsilon')}\big) = \dim\big(\mathcal{F}_{x^*(\epsilon'')}\big), \qquad \dim\big(\mathcal{F}_{s^*(\epsilon)}\big) = \dim\big(\mathcal{F}_{s^*(\epsilon')}\big),$$

and it is denoted by $\pi_{\mathrm{LCO}}(\epsilon') \stackrel{w}{=} \pi_{\mathrm{LCO}}(\epsilon'')$. Then the definition of a transition point, linearity, and a nonlinearity interval follows from Definitions 5.1.2 and 5.2.1. The stability of strict complementarity in a linearity interval follows from the definition.

Example 5.3.1. *For SDO the minimal faces containing $\mathcal{P}^*(\epsilon)$ and $\mathcal{D}^*(\epsilon)$ are given by*

$$\mathcal{F}_{X^*(\epsilon)} = \big\{Q_{\mathcal{B}(\epsilon)}U_X Q^T_{\mathcal{B}(\epsilon)} \mid U_X \succeq 0\big\}, \qquad \mathcal{F}_{S^*(\epsilon)} = \big\{Q_{\mathcal{N}(\epsilon)}U_S Q^T_{\mathcal{N}(\epsilon)} \mid U_S \succeq 0\big\},$$

where

$$\dim\big(\mathcal{F}_{X^*(\epsilon)}\big) = n_{\mathcal{B}(\epsilon)}(n_{\mathcal{B}(\epsilon)} + 1)/2, \qquad \dim\big(\mathcal{F}_{S^*(\epsilon)}\big) = n_{\mathcal{N}(\epsilon)}(n_{\mathcal{N}(\epsilon)} + 1)/2.$$

Thus, $n_{\mathcal{B}(\epsilon')}$ and $n_{\mathcal{N}(\epsilon')}$ are constant for all ϵ' belonging to a nonlinearity interval.

As demonstrated in Examples 5.1.2 and 5.2.1, linearity intervals exist for LCO, and they can be computed as efficiently as in SDO and SOCO. The LCO counterpart can be found in Section 4 in [184].

Lemma 5.3.1. *Let $\bar{\epsilon}$ belong to a linearity interval $\mathcal{I}_{\mathrm{lin}}$ and $(x^*(\bar{\epsilon}), y^*(\bar{\epsilon}), s^*(\bar{\epsilon}))$ be a maximally complementary optimal solution. Then the extreme points of $\mathcal{I}_{\mathrm{lin}}$, if exist, can be computed by solving*

$$\begin{aligned} \alpha_{lin}(\beta_{\mathrm{lin}}) := \inf(\sup) \qquad & \epsilon \\ \text{s.t.} \qquad & \mathcal{A}x = b, \\ & \mathcal{A}^* y + s = c + \epsilon \bar{c}, \\ & x \in \mathrm{ri}\left(\mathcal{F}_{x^*(\bar{\epsilon})}\right), \\ & s \in \mathrm{ri}\left(\mathcal{F}_{s^*(\bar{\epsilon})}\right). \end{aligned}$$

Remark 5.3.1. *One can extend Lemma 5.1.9 from SDO to LCO. Furthermore, it is easy to show that there exists either a unique primal optimal solution or a primal optimal set associated with a linearity interval. As a consequence, the optimal value function $\varphi(.)$ behaves linearly in a linearity interval for a parametric LCO problem.*

Now, we present the extension of Lemmas 5.1.8 and 5.2.4 to estimate the behavior of the optimal partition $\pi_{\mathrm{LCO}}(\epsilon)$ in a neighborhood of ϵ. The following technical lemma is in order.

Lemma 5.3.2. *Let $\{\mathcal{C}_k\} \subset \mathbb{V}$ be a sequence of closed convex sets. Then for $\liminf_{k\to\infty} \mathcal{C}_k$, if exists, there exists an integer N so that for all $k' \geq N$ we have*

$$\dim\left(\mathcal{C}_{k'}\right) \geq \dim\left(\liminf_{k\to\infty} \mathcal{C}_k\right).$$

Proof. Since $\{\mathcal{C}_k\}$ is a closed convex set for all k, then $\liminf_{k\to\infty} \mathcal{C}_k$ is a closed convex set by Proposition 4.15 in [150]. Let $d := \dim\left(\liminf_{k\to\infty} \mathcal{C}_k\right)$. Then there exist $x^1, \ldots, x^{d+1} \in \liminf_{k\to\infty} \mathcal{C}_k$ so that

$$x^2 - x^1,\ x^3 - x^1,\ \ldots,\ x^{d+1} - x^1$$

are linearly independent. Hence, for $i = 1, \ldots, d+1$ there exists a sequence $\{x_k^i\} \to x^i$ so that $\{x_k^2 - x_k^1,\ x_k^3 - x_k^1,\ \ldots,\ x_k^{d+1} - x_k^1\}$ remain linearly independent for sufficiently large k. Hence, $\mathcal{C}_k$ contains at least $d+1$ affinely independent points, which completes the proof. $\square$

Lemma 5.3.3. *Let $\big(x^*(\bar{\epsilon}), y^*(\bar{\epsilon}), s^*(\bar{\epsilon})\big)$ be a maximally complementary optimal solution. If $\mathcal{P}^*_{\mathrm{LCO}}(.)$ is continuous at $\bar{\epsilon}$, then $\dim\big(\mathcal{F}_{x^*(\epsilon)}\big) \le \dim\big(\mathcal{F}_{x(\epsilon')}\big)$ for all ϵ' sufficiently close to $\bar{\epsilon}$. If $\mathcal{D}^*_{\mathrm{LCO}}(.)$ is continuous at $\bar{\epsilon}$, then $\dim\big(\mathcal{F}_{s^*(\epsilon)}\big) \le \dim\big(\mathcal{F}_{s(\epsilon')}\big)$ for all ϵ' in a small neighborhood of $\bar{\epsilon}$.*

Proof. If $\mathcal{P}^*_{\mathrm{LCO}}(.)$ is continuous at $\bar{\epsilon}$, then $\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{LCO}}(\epsilon_k)$ exists for any sequence $\epsilon_k \to \bar{\epsilon}$. Hence, the result is immediate from Lemma 5.3.2. The proof for $\mathcal{D}^*_{\mathrm{LCO}}(.)$ is analogous. $\square$

We have already showed in Lemma 5.1.14 that the primal or dual optimal set mapping might be discontinuous in a nonlinearity interval for LCO. In other words, we may not arrive at a transition point by simply looking at the discontinuity of the optimal set mapping. However, if there exists a convergent sequence to a strictly complementary optimal solution at $\bar{\epsilon}$, then $\bar{\epsilon}$ has to belong to the interior of a nonlinearity interval.

Lemma 5.3.4. *Let $\big(x^*(\bar{\epsilon}), y^*(\bar{\epsilon}), s^*(\bar{\epsilon})\big)$ be a strictly complementary optimal solution, and assume that both $\mathcal{P}^*_{\mathrm{LCO}}(.)$ and $\mathcal{D}^*_{\mathrm{LCO}}(.)$ are continuous at $\bar{\epsilon}$. Then $\bar{\epsilon}$ belongs to the interior of a linearity or nonlinearity interval.*

Proof. The proof is very similar to Lemma 5.1.12. By the inner semicontinuity of $\mathcal{P}^*_{\mathrm{LCO}}(.)$ and $\mathcal{D}^*_{\mathrm{LCO}}(.)$ at $\bar{\epsilon}$, for ϵ sufficiently close to $\bar{\epsilon}$ there exists a maximally complementary solution $\big(x^*(\epsilon), y^*(\epsilon), s^*(\epsilon)\big)$ so that $x^*(\epsilon) + s^*(\epsilon) \in \operatorname{int}(\mathcal{K})$. Further, we have from Lemma 5.3.3 that

$$\dim(\mathcal{F}_{x^*(\epsilon)}) \geq \dim(\mathcal{F}_{x^*(\bar{\epsilon})}),$$
$$\dim(\mathcal{F}_{s^*(\epsilon)}) \geq \dim(\mathcal{F}_{s^*(\bar{\epsilon})}),$$

and from the strict complementarity condition that

$$\dim(\mathcal{F}_{x^*(\bar{\epsilon})}) + \dim(\mathcal{F}_{s^*(\bar{\epsilon})}) = \dim(\mathbb{V}),$$

which in turn implies

$$\dim\big(\mathcal{F}_{x^*(\epsilon)}\big) = \dim(\mathcal{F}_{x^*(\bar{\epsilon})}),$$
$$\dim\big(\mathcal{F}_{x^*(\epsilon)}\big) = \dim(\mathcal{F}_{s^*(\bar{\epsilon})}).$$

□

The following corollary is immediate from Lemma 5.3.4.

Lemma 5.3.5. *At a transition point $\bar{\epsilon}$, at least one of the strict complementarity, the primal, or dual nondegeneracy conditions fails.*

5.4 Discussion and open questions

In this chapter, we introduced the notion of a nonlinearity interval for parametric SDO and SOCO, and in general parametric LCO. We studied the continuity of optimal set mappings and optimal partition in a nonlinearity interval and provided sufficient conditions to either identify a transition point or compute a subinterval of a nonlinearity interval. We pointed out the 3-elliptope example to show that either primal or dual optimal set mapping might be discontinuous in a nonlinearity interval. Nevertheless, a complete characterization of nonlinearity intervals and transition points of the optimal partition seems to be nontrivial, under general conditions.

The main difficulty arises from the unknown type of discontinuity in a nonlinearity interval. In the 3-elliptope example, the inner semicontinuity fails at $\bar{\epsilon}$, because

$$\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{SDO}}(\epsilon_k) \subset \mathrm{ri}(\mathcal{P}^*_{\mathrm{SDO}}(\bar{\epsilon})) \tag{5.39}$$

for any sequence $\{\epsilon_k\} \to \bar{\epsilon}$. However, it is not known in general whether

$$\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{SDO}}(\epsilon_k) \subset \mathrm{rbd}(\mathcal{P}^*_{\mathrm{SDO}}(\bar{\epsilon})) \quad \text{or} \quad \liminf_{k\to\infty} \mathcal{D}^*_{\mathrm{SDO}}(\epsilon_k) \subset \mathrm{rbd}(\mathcal{D}^*_{\mathrm{SDO}}(\bar{\epsilon})) \tag{5.40}$$

can happen in a nonlinearity interval for a parametric SDO problem. In (5.39), we can obtain partial information about the rank of primal optimal solution near $\bar{\epsilon}$ from the lower semicontinuity of the rank function. However, in general, the discontinuity in (5.40) does not imply information about the rank of a maximally complementary optimal solution at $\bar{\epsilon}$. The same question can be formulated for a parametric SOCO problem and for a parametric LCO problem, in general.

In the light of the above discussion, we should also point out that either (5.39) or (5.40) might result in a transition point. For instance, we would get a transition point at $\bar{\epsilon} = \frac{1}{2}$ in Example 5.1.1, if we add the weakly inactive constraint

$$X_{23} \leq 1$$

to the primal problem. However, we can provide a sufficient condition for the discontinuity (5.40).

Lemma 5.4.1. *Assume that $\mathcal{P}^*_{\mathrm{LCO}}(.)$ is discontinuous of type (5.40) at $\bar{\epsilon}$, and for a sequence $\{\epsilon_k\} \to \bar{\epsilon}$, $\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{LCO}}(\epsilon_k)$, if exists, is attained when k is sufficiently large. Then $\bar{\epsilon}$ is a transition point. The same statement holds for $\mathcal{D}^*_{\mathrm{LCO}}(.)$.*

Proof. Since $\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{LCO}}(\epsilon_k) \subset \mathrm{rbd}(\mathcal{P}^*_{\mathrm{LCO}}(\bar{\epsilon}))$ and $\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{LCO}}(\epsilon_k)$ is attained when k is large enough, then in every neighborhood of $\bar{\epsilon}$ there exists a maximally complementary solution whose dimension of the minimal face is strictly smaller than $\dim\left(\mathcal{F}_{x^*(\bar{\epsilon})}\right)$, which implies that $\bar{\epsilon}$ is a transition point. The proof for $\mathcal{D}^*_{\mathrm{LCO}}(.)$ is analogous. □

The condition in Lemma 5.4.1 is only sufficient for the existence of a transition point, since $\liminf_{k\to\infty} \mathcal{P}^*_{\mathrm{LCO}}(\epsilon_k)$ or $\liminf_{k\to\infty} \mathcal{D}^*_{\mathrm{LCO}}(\epsilon_k)$ may not be attained. For instance, $\bar{\epsilon} = \frac{1}{2}$ is a transition point in the parametric SOCO problem (5.23), while $\liminf_{k\to\infty} \mathcal{D}^*_{\mathrm{LCO}}(\epsilon_k)$ is not attained.

We close this chapter by stating an open question on the continuity of the solutions in a nonlinearity interval:

Open question: Is there a continuous selection through the relative interior of both primal and dual optimal sets in a nonlinearity interval? In other words, can we prove that the discontinuity (5.40) never happens at any ϵ in a nonlinearity interval?

Chapter 6

Conclusions and future research

In this thesis, we studied optimal partition, solution identification, and parametric analysis for three classes of LCO problems, namely SCO, SDO, and SOCO problems. We presented a polynomial time Dikin-type algorithm for SCO. We showed how to approximate the optimal partition of SDO and how to use the approximation in a rounding procedure. Further, we used the optimal partition information to establish the quadratic convergence of Newton's method to the unique optimal solution of SOCO. Finally, we studied the parametric analysis of SDO and SOCO problems and investigated the continuity of the optimal solutions in a nonlinearity interval. Here is a summary of the materials presented in this thesis:

- We generalized the Dikin-type affine scaling method of Jansen et al. [86] to SCO using the notion of Euclidean Jordan algebras. The method starts with an interior feasible solution which is not necessarily centered. In contrast to the primal-dual affine scaling method of Monteiro et al. [117], the method features simultaneously centering and reducing the duality gap. This generalization has an $\mathcal{O}(\xi r L)$ iteration complexity, where ξ and r denotes the measure of proximity and the order of the symmetric cone, respectively. The method was tested against SeDuMi, MOSEK and SDPT3 solvers on a set of 13 SOCO test problems. The numerical experiments showed that the method is viable, robust, and capable of providing accurate solutions even though it is outperformed by the competing methods in terms of the computational time.

- We considered the identification of the optimal partition for SDO where strict complementarity may fail. Using the condition number σ defined in (3.1) and the upper bounds in (3.9), we derived bounds on the magnitude of the eigenvalues of a primal-dual solution on, or in a neighborhood of the central path. We then used the bounds to identify the subsets of the eigenvectors of the interior solutions whose accumulation points form orthonormal bases for the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$. Moreover, we measured the proximity of the approximation of the optimal partition obtained from the bounded sequence of central solutions. For the interior solutions in a neighborhood of the central path, an iteration complexity bound was provided which states that the Dikin-type primal-dual affine scaling algorithm needs at most

$$\left\lceil \xi n \log\left(\mu^{(0)}\left(\min\left\{\frac{1}{n}\left(\frac{\sigma}{\kappa' n^{\frac{3}{2}}\xi}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2\xi},\ \hat{\mu}\right\}\right)^{-1}\right)\right\rceil$$

  iterations to identify the subsets of eigenvectors whose accumulation points are orthonormal bases for $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$. It can be inferred from this complexity bound that even approximation of the optimal partition for SDO is significantly harder than the identification of the optimal partition for LO and LCP.

- We used the approximation of the optimal partition to generate an ϵ-feasible primal-dual solution with zero duality gap for the SDO problem. It is proven that if the duality gap drops below a certain bound, then both the primal and dual solutions satisfy the cone constraints, yielding an approximate maximally complementary solution.

- We revisited the identification of the optimal partition for SOCO and reproduced the bounds for the identification of the partition $\mathcal{T}$ using the error bounds for a linear conic system. Using the optimal partition of a SOCO problem, we established quadratic convergence of Newton's method to the unique optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) without strict complementarity condition. We showed that if the primal and dual nondegeneracy conditions hold, then $\nabla G(\bar{\vartheta})$ is nonsingular. Furthermore, we derived a complexity bound for identifying the quadratic convergence region of Newton's method from a sequence of central solutions. The numerical results confirmed the quadratic convergence of Newton's method to the unique optimal solution of SOCO in the absence of strict complementarity.

- We presented a rounding procedure, analogous to SDO, for an approximate maximally complementary solution. In a special case where $\mathcal{R}, \mathcal{T} = \emptyset$, the rounding procedure gives an exact strictly complementary optimal solution in a strongly polynomial time.

- We studied the parametric analysis and the identification of the optimal partition for SDO problems, when the objective function is perturbed along a fixed direction. We characterized the nonlinearity interval of the optimal partition, where the ranks of primal and dual optimal solutions, belonging to the relative interior of the optimal set, remain constant. Further, we studied the sensitivity of $\mathcal{R}\big(Q^{\mu}_{\mathcal{B}}\big)$ and $\mathcal{R}\big(Q^{\mu}_{\mathcal{N}}\big)$ with respect to ϵ and derived an upper bound on the distance between the invariant subspaces spanned by the approximation of the optimal partition. We showed that if the Jacobian is nonsingular at

$\left(X^*(\bar{\epsilon}), y^*(\bar{\epsilon}), S^*(\bar{\epsilon})\right)$, then $\bar{\epsilon}$ belongs to the interior of a nonlinearity interval.

- We studied the parametric analysis of a SOCO problem, where the objective function is perturbed along a fixed direction. We investigated the continuity of optimal solutions and the behavior of the optimal partition of the problem in a nonlinearity interval. We showed how to compute a subinterval of a nonlinearity interval under strict complementarity condition. Furthermore, under primal and dual nondegeneracy conditions, we showed that a transition point can be identified from the higher-order derivatives of the KKT conditions from the NLO reformulation of the SOCO problem.

- We extended our derivations and continuity results to LCO using the facial description of the optimal partition. We showed that the problem on the continuity of optimal solutions in a nonlinearity interval can be viewed from the lens of convex geometry.

6.1 Future research

The work on sensitivity and stability analysis is in progress. We are applying methods from numerical algebraic geometry in order to exactly compute a nonlinearity interval. Furthermore, we are investigating the continuity and differentiability of both primal and dual optimal set mappings in a nonlinearity interval. Nevertheless, we believe that there is still room for further extension/improvement of the results presented in this thesis by invoking techniques from differential geometry real algebraic geometry:

- Our approach in Section 3.1 only allows for an approximation of the optimal partition from a bounded sequence of interior solutions. It might be possible to derive additional characterization of the optimal partition if we look at the central path as a semi-algebraic set parameterized by μ.

- It is worth investigating the dependence of the condition numbers κ and κ' on the problem data in Section 3.1. The derivation of upper bounds on θ_1, θ_2, κ, and κ' can be another subject of future studies.

- It is worth considering the application of numerical algebraic geometry to establish quadratic convergence to an isolated solution of SOCO without primal or dual nondegeneracy condition.

- To establish quadratic convergence of Newton's method in Section 3.2, we assumed that the optimal partition is known, and that $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$ can be identified from $\mathcal{T}$. The derivation of upper bounds on θ_1, θ_2, κ, and β_1 deserves further research.

- The extension of our sensitivity and stability results to copositive and nonnegative polynomial cones are highly desired. It is still possible to generalize the definition of a nonlinearity interval to these cases, even though the notion of the optimal partition is no longer well-defined.

Appendix A

Appendix

A.1 Jordan algebra

This section gives a brief review of the basic properties of Euclidean Jordan algebras. For the sake of simplicity, we only provide the necessary concepts which will be required in this dissertation. For detailed studies of Euclidean Jordan algebras the reader can consult [43] and [156].

Definition A.1.1. *Let $\mathbb{V}$ be an n-dimensional vector space over the field of real numbers with a bilinear map $(x, s) \to x \circ s$. Then, $(\mathbb{V}, \circ)$ is referred to as the Euclidean Jordan algebra if for all $x, s \in \mathbb{V}$*

1. $x \circ s = s \circ x$,
2. $x \circ (x^2 \circ s) = x^2 \circ (x \circ s)$,
3. $\langle x, x \rangle > 0$ *for all* $x \neq 0$,

where $x^2 = x \circ x$, and $\langle ., . \rangle$ denotes an inner product defined on $\mathbb{V}$. An identity element is defined for a Euclidean Jordan algebra $\mathbb{V}$, if there exists a unique element e, such that $x \circ e = e \circ x = x$ for all $x \in \mathbb{V}$.

Roughly speaking, a Euclidean Jordan algebra is a commutative algebra over the field of real numbers which is not necessarily associative. Nevertheless, Euclidean Jordan algebras are power associative, i.e., $x^{p+q} := x^p \circ x^q$.

For all $x, s \in \mathbb{V}$, the bilinear map $(x, s) \to x \circ s$ is characterized by

$$L(x)s = x \circ s,$$

where $L(x)$ denotes a symmetric matrix. In particular, $L(x)e = x$ and $L(x)x = x^2$. The quadratic representation of x is defined as

$$P_x := 2L^2(x) - L(x^2).$$

Definition A.1.2. *The cone of squares of a Euclidean Jordan algebra $\mathbb{V}$ is defined as*

$$\mathcal{K}(\mathbb{V}) := \{x^2 : x \in \mathbb{V}\},$$

where $x^2 = x \circ x$, and $\mathcal{K}(\mathbb{V})$ is a closed pointed convex cone with nonempty interior.

Example A.1.1. *Let $\mathbb{V}$ be an n-dimensional vector space over the field of real numbers, where the identity element e and $L(x)$ are defined as*

$$e := (1, \mathbf{0}_{n-1})^T, \qquad L(x) := \begin{pmatrix} x_1 & x_{2:n}^T \\ x_{2:n} & x_1 I_{n-1} \end{pmatrix}.$$

The vector space $\mathbb{V}$ endowed with the bilinear map characterized by $L(x)$ is a Euclidean Jordan algebra. Further, $\mathcal{K}(\mathbb{V}) \equiv \mathbb{L}_+^n$, where

$$\mathbb{L}_+^n = \big\{x \in \mathbb{R}^n : \ x_1 \geq \|x_{2:n}\|_2\big\}.$$

In this algebra, the quadratic representation of $x \in \mathcal{K}(\mathbb{V})$ is given by

$$P_x := \begin{pmatrix} \|x\|_2^2 & 2x_1 x_{2:n}^T \\ 2x_1 x_{2:n} & (x_1^2 - \|x_{2:n}\|_2^2) I_{n-1} + 2x_{2:n} x_{2:n}^T \end{pmatrix}.$$

The following lemma specifies some important properties of the quadratic representation. We call x invertible if all the eigenvalues of x are nonzero. We refer the reader to [43] for more details.

Lemma A.1.1 (Proposition II.3.1 in [43], Lemma 8 in [156])**.** *For an invertible $x \in \mathbb{V}$ and integer value k, we have*

1. $P_{x^{-1}} = P_x^{-1}$ *and in general* $P_{x^k} = P_x^k$*,*
2. $P_x x^{-1} = x$*,*
3. $P_x e = x^2$*.*

We now introduce the concept of eigenvalue and spectral decomposition in Euclidean Jordan algebras. Let r be the smallest integer such that the set $\{e, x, x^2, \ldots, x^r\}$ is linearly dependent for $x \in \mathbb{V}$. Then, r is denoted as the degree of x, $\deg(x)$. The rank of $\mathbb{V}$ is defined as the largest value of $\deg(x)$ over $x \in \mathbb{V}$.

A nonzero element $p \in \mathbb{V}$ is called idempotent if $p^2 = p$. Furthermore, an idempotent is primitive if it is not the sum of two other idempotents. In light of these definition, a Jordan frame is defined as a set of primitive idempotents $\{p_1, \ldots, p_r\}$, where $p_i \circ p_j = 0$ for all $i \neq j$ and $p_1 + \ldots + p_r = e$.

Theorem A.1.1 (Theorem III.1.2 in [43])**.** *Let $\mathbb{V}$ be a Euclidean Jordan algebra with rank r. Then each $x \in \mathbb{V}$ can be represented as*

$$x = \lambda_1 p_1 + \ldots + \lambda_r p_r,$$

where $\{p_1, \ldots, p_r\}$ denotes a Jordan frame, and λ_i stands for the eigenvalues of x.

Example A.1.2. *Let $x \in \mathbb{L}^n_+$. It can be easily shown that*

$$x^2 - 2x_1 x + (x_1^2 - \|x_{2:n}\|_2^2)e = 0.$$

This implies that $r = 2$ for this Euclidean Jordan algebra. Hence, the spectral decomposition for an element $x \in \mathbb{L}^n_+$ is given by

$$x = \lambda_1 p_1 + \lambda_2 p_2, \tag{A.1}$$

where

$$\lambda_1 = x_1 - \|x_{2:n}\|_2, \qquad \lambda_2 = x_1 + \|x_{2:n}\|_2,$$
$$p_1 = \frac{1}{2}\begin{pmatrix} 1 \\ -\frac{x_{2:n}}{\|x_{2:n}\|_2} \end{pmatrix}, \qquad p_2 = \frac{1}{2}\begin{pmatrix} 1 \\ \frac{x_{2:n}}{\|x_{2:n}\|_2} \end{pmatrix}.$$

We can now extend the definition of any real valued function $f(.)$ to elements of Euclidean Jordan algebras by

$$f(x) := f(\lambda_1)p_1 + \ldots + f(\lambda_r)p_r.$$

In particular, we have

$$x^{\frac{1}{2}} := \lambda_1^{\frac{1}{2}} p_1 + \ldots + \lambda_r^{\frac{1}{2}} p_r,$$
$$x^{-1} := \lambda_1^{-1} p_1 + \ldots + \lambda_r^{-1} p_r.$$

Note that $x^{-1} \circ x = e$. In light of the definitions given so far, the trace, determinant, and norms of x are formally defined as follows.

Definition A.1.3. *Let $x \in \mathbb{V}$ and $\lambda_1, \ldots, \lambda_r$ be the eigenvalues of x. Then,*

1. $\mathrm{Trace}(x) := \lambda_1 + \lambda_2 + \ldots + \lambda_r$,
2. $\det(x) := \lambda_1 \lambda_2 \ldots \lambda_r$,
3. $\langle x, y \rangle := \mathrm{Trace}(x \circ y)$,
4. $\|x\|_F := \sqrt{\lambda_1^2 + \ldots + \lambda_r^2}$,

5. $\|x\|_2 := \max_i |\lambda_i|$.

We now state an important theorem from [9] which is adopted from [43].

Theorem A.1.2 (Theorem 8.3.6 in [9], [43]). *Let $\mathbb{V}$ be a Euclidean Jordan algebra. Then $\mathbb{V}$ falls into one of the following categories:*

1. *An n-dimensional vector space over the field of real numbers, where*

$$x \circ s := \begin{pmatrix} x^T s \\ x_0 s_{1:n-1} + s_0 x_{1:n-1} \end{pmatrix}.$$

2. *The space of $n \times n$ real symmetric matrices, where*

$$X \circ S := \frac{XS + SX}{2} \tag{A.2}$$

   for symmetric matrices X and S.

3. *The space of $n \times n$ complex Hermitian matrices, where the Jordan product is defined as in (A.2) for Hermitian matrices X and S.*

4. *The space of $n \times n$ Hermitian matrices with quaternion entries, where the Jordan product is defined as in (A.2) for quaternion Hermitian matrices X and S.*

5. *The space of 3×3 Hermitian matrices with octonion entries, known as Albert algebra, where the Jordan product is defined as in (A.2).*

In the rest of this section, we review some technical lemmas (without their proofs) which are necessary for the complexity analysis of the Dikin-type algorithm. From now on, it is assumed that $\mathcal{K}$ is a symmetric cone with a nonempty interior.

Lemma A.1.2 (Lemma 2.13 in [67]). *Assume that* $x, s \in \mathbb{V}$ *and* $\operatorname{Trace}(x \circ s) = 0$. *Then, we have*

$$\frac{1}{4}\|x + s\|_F^2 e \succeq_{\mathcal{K}} x \circ s \succeq_{\mathcal{K}} -\frac{1}{4}\|x + s\|_F^2 e.$$

Lemma A.1.3 (Lemma 2.15 in [67]). *Assume that* $x \circ s \succ_{\mathcal{K}} 0$, *where* $x, s \in \mathbb{V}$. *Then,* $\det(x) \neq 0$.

Lemma A.1.4 (Lemma 2.17 in [67]). *Let* $x \in \mathbb{V}$ *and* $s \succ_{\mathcal{K}} 0$. *Then, we have*

$$\lambda_{\min}(x) \operatorname{Trace}(s) \leq \operatorname{Trace}(x \circ s) \leq \lambda_{\max}(x) \operatorname{Trace}(s).$$

The following lemma points out a nice property of the quadratic representation.

Lemma A.1.5 (Theorem III.2.1 and Proposition III.2.2 in [43]). *Let* $x \in \mathbb{V}$. *Then,* $L(x)$ *is positive definite (semidefinite) if and only if* $x \succ_{\mathcal{K}} 0$ $(x \succeq_{\mathcal{K}})$. *Further,* $P_x \operatorname{int}(\mathcal{K}) = \operatorname{int}(\mathcal{K})$ *if* x *is invertible.*

In fact, Lemma A.1.5 states that for each interior solution $x \succ_{\mathcal{K}} 0$ and $s \succ_{\mathcal{K}} 0$, $P_x s$ is an invertible linear map from $\operatorname{int}(\mathcal{K})$ to $\operatorname{int}(\mathcal{K})$. All this indicates that the NT search directions obtained from (2.15) are well-defined.

Lemma A.1.6 (Proposition 21 in [156]). *Let* $x \succ_{\mathcal{K}} 0$ *and* $s \succ_{\mathcal{K}} 0$, *and* w *be the scaling point of* x *and* s *as defined in* (1.19). *Then,*

1. $\operatorname{Trace}(P_w^{-\frac{1}{2}} x \circ P_w^{\frac{1}{2}} s) = \operatorname{Trace}(x \circ s)$,
2. $P_{x^{\frac{1}{2}}} s \sim P_{s^{\frac{1}{2}}} x$,
3. $P_{\tilde{x}^{\frac{1}{2}}} \tilde{s} \sim P_{x^{\frac{1}{2}}} s$,

where $\tilde{x} := P_w^{-\frac{1}{2}} x$ *and* $\tilde{s} := P_w^{\frac{1}{2}} s$.

Lemma A.1.7 (Lemma 30 in [156]). *Let* $x \succ_{\mathcal{K}} 0$ *and* $s \succ_{\mathcal{K}} 0$. *Then, we have*

1. $\lambda_{\min}(P_x^{\frac{1}{2}} s) \geq \lambda_{\min}(x \circ s)$,
2. $\lambda_{\max}(P_x^{\frac{1}{2}} s) \leq \lambda_{\max}(x \circ s)$.

A.2 Error bounds for mathematical optimization

An error bound has its roots in numerical optimization, where a stopping criterion is needed for termination of iterative methods. However, the application of error bound in the other areas of mathematical optimization has received a great deal of attention. For a set $\mathcal{D}_1 \subseteq \mathbb{V}$, error bound quantifies lower/upper bound on the distance of a given vector to $\mathcal{D}_1$ in terms of a so called residual function. For a system of linear inequalities, the residual function can be simply defined as the amount of violation of the inequalities at a given solution.

Mathematically speaking, for given sets $\mathcal{D}_1$ and $\mathcal{D}_2$ in the inner product space $\mathbb{V}$ the residual function is defined as a real valued function $\mathrm{res} : \mathcal{D}_1 \cup \mathcal{D}_2 \to \mathbb{R}_+$ where $\mathrm{res}(x) = 0$ if and only if $x \in \mathcal{D}_1$. In light of this mathematical definition, the error bound of the set $\mathcal{D}_2$ with respect to $\mathcal{D}_1$ takes the following form

$$\kappa_1 \,\mathrm{res}(x)^{\gamma_1} \leq \mathrm{dist}(x, \mathcal{D}_1) \leq \kappa_2 \,\mathrm{res}(x)^{\gamma_2}, \tag{A.3}$$

where $\gamma_1, \gamma_2 > 0$, κ_1 and κ_2 are positive condition numbers, and $\mathrm{dist}(x, \mathcal{D}_1)$ denotes the distance of a given solution x from $\mathcal{D}_1$ with respect to the Frobenius norm, i.e.,

$$\mathrm{dist}(x, \mathcal{D}_1) = \inf\{\|x - s\|_F \mid s \in \mathcal{D}_1\}.$$

From the mathematical optimization point of view, $\mathcal{D}_1$ could be a system of linear inequalities over a closed convex cone, or a system of nonlinear inequalities. From now on, we only consider the right hand side inequality in (A.3).

The error bound in (A.3) is called Lipschitzian if there exists a condition number $\kappa > 0$ with $\gamma = 1$. If $\gamma \neq 1$, the error bound is called Hölderian. For a given value γ it is not always easy to evaluate the condition number κ. Indeed, this is equivalent to the following optimization problem:

$$\kappa = \sup_{x \in \mathcal{D}_2 \setminus \mathcal{D}_1} \frac{\mathrm{dist}(x, \mathcal{D}_1)}{\mathrm{res}(x)^{\gamma}},$$

which needs to have a finite value. The reader is referred to [130] for more applications of an error bound.

In case that $\mathcal{D}_1$ is defined by a system of linear equalities and inequalities, a Lipschitzian error bound exists due to Hoffman [80], see also [72, 107].

Theorem A.2.1 (Section 2 in [80])**.** *Consider a nonempty convex set $\mathcal{D}_1$ defined by*

$$\mathcal{D}_1 = \{x \in \mathbb{R}^n \mid A_1 x = b_1, \quad A_2 x \leq b_2\},$$

where $A_1 \in \mathbb{R}^{m_1 \times n}$, $A_2 \in \mathbb{R}^{m_2 \times n}$, $b_1 \in \mathbb{R}^{m_1}$, and $b_2 \in \mathbb{R}^{m_2}$. Then there exists a condition number $\kappa > 0$ so that

$$\operatorname{dist}(x, \mathcal{D}_1) \leq \kappa \Big(\|A_1 x - b_1\|_2 + \|[A_2 x - b_2]_+\|_2 \Big), \qquad \forall x \in \mathbb{R}^n, \tag{A.4}$$

where $[\eta]_+ = (\max\{\eta_1, 0\}, \ldots, \max\{\eta_{m_2}, 0\})^T$ for any $\eta \in \mathbb{R}^{m_2}$.

The Lipschitzian error bound (A.4) can be extended for a convex system by imposing the Slater condition [102], and either a constraint qualification [108] or conditions on the boundedness of the solution set [142]. See Section 4 in [130] and the references cited therein for a detailed discussion.

If the Slater condition fails, however, a Lipschitzian error bound may not exist. For instance, consider the convex inequality system $\{x \in \mathbb{R} \mid x^2 \leq 0\}$, whose solution set is a singleton. We can observe that $|x| \leq \kappa x^2$ always fails for any condition number κ. Nevertheless, a Hölderian error bound with $\gamma = \frac{1}{2}$ exists. This is formally stated in the following theorem.

Theorem A.2.2 (Theorem 3.1 in [178])**.** *Consider a nonempty convex quadratic system*

$$\mathcal{D}_1 = \{x \in \mathbb{R}^n \mid g(x) \leq 0\},$$

where $g(x) = \big(g_1(x), \ldots, g_m(x)\big)^T$. *There exists a positive exponent* $\gamma \neq 1$ *and a positive condition number* κ *so that*

$$\operatorname{dist}(x, \mathcal{D}_1) \le \kappa \max\Big\{\|[g(x)]_+\|_2, \|[g(x)]_+\|_2^{\gamma}\Big\}, \qquad \forall x \in \mathbb{R}^n,$$

where $\gamma = 2^{-d(\mathcal{D}_1)}$, $d(\mathcal{D}_1)$ *is called the degree of singularity of the system, and*

$$[g(x)]_+ := \big(\max\{g_1(x), 0\}, \ldots, \max\{g_m(x), 0\}\big)^T.$$

An upper bound on $d(\mathcal{D}_1)$ is given in Corollary 3.1 in [178]. The extension of Theorem A.2.2 to nonconvex quadratic functions yields a local Hölderian error bound.

Theorem A.2.3 (Theorem 2.4 in [104]). *Let* $\mathcal{D}_1$ *be the set of solutions of a (not necessarily convex) quadratic system as*

$$\mathcal{D}_1 = \{x \in \mathbb{R}^n \mid g(x) \le 0\}.$$

If $\mathcal{D}_1$ *is nonempty, then for a given scalar* $\varrho > 0$ *there exist* γ *and a positive condition number* κ *such that*

$$\operatorname{dist}(x, \mathcal{D}_1) \le \kappa\big(\|[g(x)]_+\|_2\big)^{\gamma}, \qquad \forall x \in \mathbb{R}^n,\ \|x\|_2 \le \varrho.$$

In the rest of this section, we present some error bound results for conic and polynomial systems. We employ the error bounds in Chapters 3 and 4 to estimate the distance of a central solution from the optimal set. We refer the reader to [38, 106] for further reading.

A.2.1 Error bound for a linear matrix inequality (LMI) system

An LMI system in conic form is defined as

$$\begin{cases} X \in X_0 + \mathcal{S}, \\ X \succeq 0, \end{cases} \tag{A.5}$$

where X_0 is a symmetric matrix and $\mathcal{S} \subset \mathbb{S}^n$ denotes a linear subspace of symmetric matrices. For system (A.5) we consider a sequence of solutions denoted by X^ζ for $\zeta > 0$ which satisfies

$$\operatorname{dist}(X^\zeta, X_0 + \mathcal{S}) \leq \zeta, \qquad \lambda_{\min}(X^\zeta) \geq -\zeta, \tag{A.6}$$

for all $\zeta > 0$. Further, $\bar{\mathcal{S}}$ is defined as the smallest subspace containing $X_0 + \mathcal{S}$, i.e.,

$$\bar{\mathcal{S}} := \{X \in \mathbb{S}^n \mid X + tX_0 \in \mathcal{S}, \text{ for some } t\}.$$

The following theorem is in order.

Theorem A.2.4 (Theorem 3.3 in [167]). *Let $\{X^\zeta \mid 0 < \zeta \leq 1\}$ be a set of solutions so that $\|X^\zeta\|_F$ is bounded and (A.6) holds for all $0 < \zeta \leq 1$. Then there exist a positive condition number κ independent of ζ and a positive exponent γ such that*

$$\operatorname{dist}\left(X^\zeta, (X_0 + \mathcal{S}) \cap \mathbb{S}^n_+\right) \leq \kappa\zeta^\gamma,$$

where $\gamma = 2^{-d(\bar{\mathcal{S}}, \mathbb{S}^n_+)}$ in which $d(\bar{\mathcal{S}}, \mathbb{S}^n_+)$ denotes the degree of singularity of the linear subspace $\bar{\mathcal{S}}$.

In simple words, the degree of singularity [167] is defined as the minimum number of facial reduction steps to get the minimal face of the positive semidefinite cone which contains the optimal set. See [132] for a simple derivation of the facial reduction algorithm.

Theorem A.2.5 (Theorem 3.6 in [167]). *For a linear subspace $\bar{\mathcal{S}} \subset \mathbb{S}^n$, we have*

$$d(\bar{\mathcal{S}}, \mathbb{S}^n_+) \leq \min\left\{n - 1, \dim(\bar{\mathcal{S}}), \dim(\bar{\mathcal{S}}^\perp)\right\}.$$

Example A.2.1. *We can show that the upper bound given in Theorem A.2.5 is* *indeed tight. To do so, consider the following LMI system*

$$\begin{cases} X_{11} = 0, \\ X_{kk} = x_{1(k+1)}, \qquad k = 2, \dots, n-1, \\ X \succeq 0, \end{cases}$$

where the set of feasible solutions is given by

$$X = \begin{pmatrix} \mathbf{0}_{(n-1)\times(n-1)} & 0 \\ 0 & X_{nn} \end{pmatrix}, \qquad X_{nn} \geq 0.$$

Using the facial reduction procedure in [167], we can see that the number of facial reduction steps is $n-1$ for all $n \geq 2$. Due to the lengthy discussion, we omit the details here and refer the interested reader to Section 3 in [167] for a simple demonstration of the facial reduction algorithm. Additional examples of the facial reduction for SDO problems can be found in [25].

A.2.2 Error bound for a linear conic system

The Hölderian error bound given in Section A.2.1 can be extended for a linear conic system. Let $\mathcal{S}$ be a linear subspace of $\mathbb{R}^{\bar{n}}$, $x_0 \in \mathbb{R}^{\bar{n}}$, and $\mathcal{K} \subset \mathbb{R}^{\bar{n}}$ be a Cartesian product of p second-order and q positive semidefinite cones as

$$\mathcal{K} := \mathbb{L}_+^{n_1} \times \dots \times \mathbb{L}_+^{n_p} \times \mathbb{S}_+^{n_{p+1}} \times \dots \times \mathbb{S}_+^{n_{p+q}},$$

where $\bar{n} = \sum_{i=1}^{p} n_i + \sum_{i=p+1}^{p+q} \frac{n_i(n_i+1)}{2}$. Then a linear conic system is defined as

$$\begin{cases} x \in x_0 + \mathcal{S}, \\ x \in \mathcal{K}. \end{cases} \tag{A.7}$$

At a given solution x the amount of constraint violation is given by

$$\operatorname{dist}\big(x,\ x_0 + \mathcal{S}\big) + [-\lambda_{\min}(x)]_+.$$

If $x^i \in \mathbb{L}_+^{n_i}$, then $\lambda_{\min}(x^i) = x_1^i - \|x_{2:n_i}^i\|_2$, see Section 1.1 for the definition of an eigenvalue in case of SOCO. The following Hölderian error bound is well-known from Theorem 7.4.2 in [103].

Theorem A.2.6 (Theorem 7.4.2 in [103])**.** *Let x^ζ with $0 < \zeta \leq 1$ be a bounded set of solutions so that for all $0 < \zeta \leq 1$ we have*

$$\operatorname{dist}\left(x^\zeta,\ \bar{\mathcal{S}}\right) \leq \zeta, \quad \lambda_{\min}\left(x^\zeta\right) \geq -\zeta, \tag{A.8}$$

where $\bar{\mathcal{S}}$ is the minimal linear subspace which contains $x_0 + \mathcal{S}$, i.e.,

$$\bar{\mathcal{S}} := \{x \mid x + tx_0 \in \mathcal{S},\ \textit{for some } t \in \mathbb{R}\}.$$

and

$$\operatorname{dist}\left(x^\zeta,\ \bar{\mathcal{S}}\right) = \min\left\{\|\psi - x^\zeta\|_2 \mid \psi \in \bar{\mathcal{S}}\right\}.$$

Then there exist a positive condition number κ independent of ζ and a positive exponent γ so that

$$\operatorname{dist}\left(x^\zeta,\ (x_0 + \mathcal{S}) \cap \mathcal{K}\right) \leq \kappa\zeta^\gamma,$$

where $\gamma = 2^{-d(\bar{\mathcal{S}},\mathcal{K})}$ and $d(\bar{\mathcal{S}}, \mathcal{K})$ denotes the degree of singularity of the subspace $\bar{\mathcal{S}}$.

The definition of the degree of singularity is analogous to Section A.2.1. The degree of singularity of a linear conic system is zero if it satisfies the interior point condition. An upper bound on the degree of singularity is given in Theorem A.2.7.

Theorem A.2.7. *For the linear conic system (A.7) we have*

$$d(\bar{\mathcal{S}}, \mathcal{K}) \leq \min\left\{p + \sum_{i=p+1}^{p+q} (n_i - 1),\ \dim(\bar{\mathcal{S}}),\ \dim(\bar{\mathcal{S}}^\perp)\right\}. \tag{A.9}$$

A.2.3 Error bound for a polynomial system

The Hölderian error bound in Theorem A.2.2 can be extended for a solution set defined by a system of polynomial mappings as follows

$$\mathcal{D}_1 := \big\{x \in \mathbb{R}^n \mid g_1(x) \le 0, \ldots, g_{m_1}(x) \le 0,\ h_1(x) = 0, \ldots, h_{m_2}(x) = 0\big\},$$

in which g_j for $j = 1, \ldots, m_1$ and h_k for $k = 1, \ldots, m_2$ are polynomials with real coefficients.

Theorem A.2.8 (Theorem 2.2 in [102])**.** *Assume that $\mathcal{D}_1 \neq \emptyset$. There exist exponents $\gamma > 0$ and $\gamma' \ge 0$, and a condition number $\kappa > 0$ such that*

$$\operatorname{dist}(x, \mathcal{D}_1) \le \kappa\big(1 + \|x\|_2\big)^{\gamma'}\big(\|[g(x)]_+\|_2 + \|h(x)\|_2\big)^{\gamma}, \qquad \forall x \in \mathbb{R}^n,$$

where

$$\begin{aligned} [g(x)]_+ &:= \big(\max\{g_1(x), 0\}, \ldots, \max\{g_{m_1}(x), 0\}\big)^T, \\ h(x) &:= (h_1(x), \ldots, h_{m_2}(x))^T. \end{aligned}$$

A.3 A lower bound on σ

In this section, we derive a lower bound on the condition number σ defined in (3.1). To do so, we resort to a technical lemma in [137].

An integral polynomial map $f : \mathbb{R}^s \to \mathbb{R}^t$ is defined as a map consisting of polynomial functions f^i of degree d_i with integer coefficients. We consider a solution set $V(f)$ defined as

$$V(f) := \{x \mid f^i(x)\ \Delta_i\ 0, \ \ \forall\, i\},$$

where Δ_i stands for one of the relations $\{>, =, \geq\}$. Depending on the polynomial map f, the solution set $V(f)$ could be connected or disconnected. For this polynomial map L_f denotes the binary length of the largest absolute value of the coefficients of the polynomials, where the binary length of an integer n is defined as

$$\ell(n) := 1 + \left\lceil \log_2(|n| + 1) \right\rceil, \tag{A.10}$$

in which $\log_2(.)$ stands for the logarithm to the base 2.

The next lemma shows that there exists a sphere $B(0, r)$ which circumscribes some solutions from every connected component of $V(f)$.

Lemma A.3.1 (Lemma 3.1 in [137])**.** *Suppose that the polynomials in the polynomial map f have maximum degree d, i.e., $d := \max_i\{d_i\}$ with $d \geq 2$. Then every connected component of $V(f)$ intersects the sphere $\{x \mid \|x\|_2 \leq r\}$, where $\log_2(r) = L_f(td)^s$.*

Lemma A.3.2. *Let the SDO problems* ($\mathrm{P_{SDO}}$) *and* ($\mathrm{D_{SDO}}$) *be given by integer data, L denote the binary length of the largest absolute value of the entries in b, C, and A^i for $i = 1, \ldots, m$. Then, for the condition number σ we have*

$$\sigma \geq \min\left\{\frac{1}{r_{\mathcal{P}^*_{\mathrm{SDO}}} \sum_{i=1}^m \|A^i\|_F}, \frac{1}{r_{\mathcal{D}^*_{\mathrm{SDO}}}}\right\}, \tag{A.11}$$

where

$$\begin{aligned}
\log_2(r_{\mathcal{P}^*_{\mathrm{SDO}}}) &= (L+2)\Big(\max\{n,3\}(6n^2+2n+m)\Big)^{5n^2+2m},\\
\log_2(r_{\mathcal{D}^*_{\mathrm{SDO}}}) &= (L+2)\Big(\max\{n,3\}(7n^2+2n+2m)\Big)^{6n^2+m}.
\end{aligned}$$

Proof. Recall from (3.2) and (3.3) that

$$\sigma_{\mathcal{B}} \geq \lambda_{\min}(Q_{\mathcal{B}}^T \tilde{X} Q_{\mathcal{B}}), \quad \sigma_{\mathcal{N}} \geq \lambda_{\min}(Q_{\mathcal{N}}^T \tilde{X} Q_{\mathcal{N}}), \qquad \forall (\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^*_{\mathrm{SDO}} \times \mathcal{D}^*_{\mathrm{SDO}},$$

which motivates us to find a solution in the relative interior of the optimal set. We apply the definition of the analytic center of the optimal set to find a solution in the relative interior of the optimal set, and we then derive a lower bound on its minimum eigenvalue. It should be noted that Ramana [137] used this definition to compute a lower bound on the volume of a sphere inscribed in the feasible set of a so called strict semidefinite feasibility problem.

Throughout the proof, we can assume that $n_{\mathcal{B}}, n_{\mathcal{N}} > 0$. By Theorem 1.5.1, any primal-dual optimal pair is a solution to the following LMI system

$$\begin{cases} A^i \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T & = b_i, \quad i = 1, \ldots, m, \\ C - \sum_{i=1}^{m} y_i A^i & = Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T, \\ U_X, U_S \succeq 0, \end{cases} \tag{A.12}$$

where $U_X \in \mathbb{S}_+^{n_{\mathcal{B}}}$ and $U_S \in \mathbb{S}_+^{n_{\mathcal{N}}}$ as defined in Theorem 1.5.1, and $Q_{\mathcal{B}}$ and $Q_{\mathcal{N}}$ are assumed to be known. Therefore, since $n_{\mathcal{B}}, n_{\mathcal{N}} > 0$, we obtain the set of maximally complementary optimal solutions if we add the constraints $U_X, U_S \succ 0$ to (A.12), i.e.,

$$\begin{cases} A^i \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T & = b_i, \quad i = 1, \ldots, m, \\ C - \sum_{i=1}^{m} y_i A^i & = Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T, \\ U_X, U_S \succ 0. \end{cases} \tag{A.13}$$

Then for a given orthonormal basis $Q_{\mathcal{B}}$, the analytic center of the primal optimal set can be computed by solving

$$\begin{aligned} \max \quad & \log(\det(U_{X^a})) \\ \text{s.t.} \quad & A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T = b_i, \quad i = 1, \ldots, m, \\ & U_{X^a} \succ 0. \end{aligned} \tag{A.14}$$

Problem (A.14) is convex with a strictly concave objective function over the cone of positive definite matrices, which by $n_{\mathcal{B}} > 0$ induces the existence of a unique optimal solution for (A.14). Further, there exists a vector of Lagrange multipliers $u \in \mathbb{R}^m$ so that the following system of optimality conditions has a solution:

$$\begin{cases} U_{X^a}^{-1} - \sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}} & = 0, \\ A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T & = b_i, \quad i = 1, \dots, m, \\ U_{X^a} \succ 0. \end{cases} \tag{A.15}$$

For any solution (U_{X^a}, u) of (A.15), which is unique in terms of U_{X^a} but not necessarily in terms of u, $X^a := Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T$ is the analytic center of the primal optimal set. To derive a lower bound on the minimum eigenvalue of X^a, we have from (A.15) that

$$\begin{aligned} \lambda_{\min}(U_{X^a}) = \frac{1}{\lambda_{\max}\Big(\sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\Big)} &\geq \frac{1}{\big\|\sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\big\|_F} \\ &\geq \frac{1}{\sum_{i=1}^m |u_i| \|Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\|_F} \\ &\geq \frac{1}{\sum_{i=1}^m |u_i| \|A^i\|_F}, \end{aligned} \tag{A.16}$$

where we have used the triangle inequality and the fact that $\|Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\|_F \leq \|A^i\|_F$. Note that the bound (A.16) depends on an upper bound on $|u_i|$ which itself relies on $Q_{\mathcal{B}}$. In reality, however, $Q_{\mathcal{B}}$ is not known a priori, since it is determined by solutions in the relative interior of the optimal set. Hence, the idea is to characterize all possible orthonormal bases for $\mathcal{B}$, i.e., to characterize the properties of $\Gamma_{\mathcal{B}}$, in the optimality conditions (A.15) to describe the analytic center of the optimal set. Then a direct application of Lemma A.3.1 to the embedded set yields an upper bound on $|u_i|$.

Assume that $Q_{\mathcal{B}}$ is an unknown orthonormal basis in (A.14), i.e., $Q_{\mathcal{B}}$ is still an orthonormal basis for $\mathcal{B}$ but acts as an unknown in (A.14), which leads to a nonconvex optimization problem in $Q_{\mathcal{B}}$ and U_{X^a}. Then, problem (A.14) can equivalently be written, see e.g., Theorem 2.1 in [57], as

$$\max_{Q_{\mathcal{B}}\in\Gamma_{\mathcal{B}}} \max_{U_{X^a}\succ 0} \Big\{ \log(\det(U_{X^a})) \mid A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T = b_i, \quad i = 1,\ldots,m \Big\}. \tag{A.17}$$

Any optimal solution $(Q_{\mathcal{B}}, U_{X^a})$ of (A.14) is also optimal for (A.17) and vice versa. This is due to the fact that the optimal solution of the inner maximization problem in (A.17) is attained. By Lemma 1.5.3, Theorem 1.5.1 and (A.13), the set $\Gamma_{\mathcal{B}}$ is compact, and it is equivalent to the set of all $Q_{\mathcal{B}}$ with orthonormal columns by which (A.13) is feasible. Since the unique optimal solution of the inner maximization problem in (A.17) is attained, and its set of Lagrange multipliers is nonempty, then (A.15) with $\Gamma_{\mathcal{B}}$ describes the analytic center of the primal optimal set, see Section 4.2 in [57] for a similar argument in the context of the generalized Benders decomposition.

Now, we apply Lemma A.3.1 to the above embedded set. Let

$$\vartheta_p := (U_{X^a}, u, Z_X, U_S, y, Q_{\mathcal{B}}, Q_{\mathcal{N}}),$$

where $Z_X \in \mathbb{R}^{n_{\mathcal{B}}\times n_{\mathcal{B}}}$. We then define the integral polynomial map

$$f_p : \mathbb{R}^{n_{\mathcal{B}}\times n_{\mathcal{B}}} \times \mathbb{R}^m \times \mathbb{R}^{n_{\mathcal{B}}\times n_{\mathcal{B}}} \times \mathbb{R}^{n_{\mathcal{N}}\times n_{\mathcal{N}}} \times \mathbb{R}^m \times \mathbb{R}^{n\times n_{\mathcal{B}}} \times \mathbb{R}^{n\times n_{\mathcal{N}}} \to \mathbb{R}^{t_p}$$

as defined below

$$f_p(\vartheta_p) := \begin{pmatrix} \operatorname{vec}\Big(Z_X - \sum_{i=1}^{m} u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\Big) \\ \operatorname{vec}\big(U_{X^a} Z_X - I_{n_{\mathcal{B}}}\big) \\ A^1 \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T - b_1 \\ \vdots \\ A^m \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T - b_m \\ \operatorname{vec}\Big(C - \sum_{i=1}^{m} y_i A^i - Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T\Big) \\ \operatorname{vec}\big(Q_{\mathcal{B}}^T Q_{\mathcal{B}} - I_{n_{\mathcal{B}}}\big) \\ \operatorname{vec}\big(Q_{\mathcal{N}}^T Q_{\mathcal{N}} - I_{n_{\mathcal{N}}}\big) \\ \operatorname{vec}\big(Q_{\mathcal{B}}^T Q_{\mathcal{N}}\big) \end{pmatrix}, \tag{A.18}$$

where $t_p = 3n_{\mathcal{B}}^2 + n_{\mathcal{N}}^2 + n_{\mathcal{B}} n_{\mathcal{N}} + n^2 + m$. Note that the symmetry of Z_X and U_S follows from the symmetry of A^i and C, and the symmetry of U_{X^a} follows from the symmetry of Z_X. Moreover, we define the solution set Ω_p to enforce the positive definiteness of U_{X^a} and U_S as follows

$$\Omega_p := \Big\{\vartheta_p \mid \det(U_{X^a}[i]) > 0, \quad \det(U_S[j]) > 0, \quad i = 1, \ldots, n_{\mathcal{B}},\ j = 1, \ldots, n_{\mathcal{N}}\Big\}, \tag{A.19}$$

in which $U_{X^a}[i]$ denotes the i^{th} leading principal submatrix of U_{X^a}. Indeed, the strict inequalities in (A.19) are necessary and sufficient for the positive definiteness of U_{X^a} and U_S. By the interior point condition, the solution set $V(f_p) \cap \Omega_p$, where $V(f_p) = \big\{\vartheta_p \mid f_p(\vartheta_p) = 0\big\}$, is nonempty but not necessarily a singleton. Then, from every solution $\vartheta_p \in V(f_p) \cap \Omega_p$, we can extract a solution $(U_{X^a}, u, Q_{\mathcal{B}})$ which is the analytic center of the primal optimal set, since it satisfies the constraints in (A.15).

The solution set Ω_p is characterized by $n_{\mathcal{B}} + n_{\mathcal{N}}$ integer polynomials of the maximum degree $\max\{n_{\mathcal{B}}, n_{\mathcal{N}}\}$. Since the symmetry of the matrices U_{X^a}, Z_X, and U_S is not presumed for f_p and Ω_p, the coefficients of the polynomial functions are bounded above by twice the largest absolute value of the entries in b, C, and A^i for $i = 1,\dots,m$. For instance, the coefficients of $\det(U_{X^a}[i])$ are just 1, but $u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}$ has some polynomial terms with coefficients twice the off-diagonal entries of A^i. Hence, the binary length of the largest absolute value of the coefficients in (A.18) and (A.19) is bounded above by $L + \ell(2) - 1 = L + 2$, see Section 3.1 in [137].

Consequently, by applying Lemma A.3.1 to the set $V(f_p) \cap \Omega_p$, we can conclude that there exists a solution $\vartheta_p \in V(f_p) \cap \Omega_p$ so that $\|\vartheta_p\|_2 \leq r_{\mathcal{P}^*_{\mathrm{SDO}}}$, where

$$
\begin{aligned}
\log_2(r_{\mathcal{P}^*_{\mathrm{SDO}}}) &= (L+2)(\bar{t}_p \bar{d}_p)^{\bar{s}_p},\\
\bar{d}_p &:= \max\{n_{\mathcal{B}}, n_{\mathcal{N}}, 3\} \leq \max\{n, 3\},\\
\bar{t}_p &:= t_p + n_{\mathcal{B}} + n_{\mathcal{N}} = 3n_{\mathcal{B}}^2 + n_{\mathcal{N}}^2 + n_{\mathcal{B}} n_{\mathcal{N}} + n_{\mathcal{B}} + n_{\mathcal{N}} + n^2 + m \leq 6n^2 + 2n + m,\\
\bar{s}_p &:= 2n_{\mathcal{B}}^2 + n_{\mathcal{N}}^2 + n(n_{\mathcal{B}} + n_{\mathcal{N}}) + 2m \leq 5n^2 + 2m,
\end{aligned}
$$

in which $\bar{s}_p$ denotes the total number of variables in the polynomial map f_p, and $\bar{d}_p$ is the maximum degree of the polynomials in f_p and the polynomials defining Ω_p. As a result, there exists u so that $|u_i| \leq \|u\|_2 \leq r_{\mathcal{P}^*_{\mathrm{SDO}}}$. Then, using the inequality (A.16), we get

$$
\sigma_{\mathcal{B}} \geq \lambda_{\min}(U_{X^a}) \geq \frac{1}{\sum_{i=1}^m |u_i| \|A^i\|_F} \geq \frac{1}{r_{\mathcal{P}^*_{\mathrm{SDO}}} \sum_{i=1}^m \|A^i\|_F}.
$$

This completes the first part of the proof. In a similar fashion, we can use the same reasoning as in the primal side to derive a lower bound on $\sigma_{\mathcal{N}}$. Notice that for a given orthonormal basis $Q_{\mathcal{N}}$, the analytic center of the dual optimal set can be obtained by solving

$$
\begin{aligned}
\max \quad & \log(\det(U_{S^a}))\\
\text{s.t.} \quad & \sum_{i=1}^m y_i^a A^i + Q_{\mathcal{N}} U_{S^a} Q_{\mathcal{N}}^T = C, \qquad\qquad (A.20)\\
& U_{S^a} \succ 0,
\end{aligned}
$$

which is a convex optimization problem with strictly concave objective function. The optimality conditions for (A.20) are given by

$$\begin{cases} U_{S^a}^{-1} - Q_{\mathcal{N}}^T W Q_{\mathcal{N}} & = 0, \\ A^i \bullet W & = 0, \quad i = 1, \ldots, m, \\ \sum_{i=1}^m y_i^a A^i + Q_{\mathcal{N}} U_{S^a} Q_{\mathcal{N}}^T & = C, \\ U_{S^a} \succ 0, \end{cases} \tag{A.21}$$

where W is an $n \times n$ symmetric matrix. Note that the symmetry of A^i induces the symmetry of U_{S^a} but not necessarily the symmetry[1] of W. Then the optimality conditions (A.21) imply

$$\lambda_{\min}(U_{S^a}) = \frac{1}{\lambda_{\max}(Q_{\mathcal{N}}^T W Q_{\mathcal{N}})} \geq \frac{1}{\|Q_{\mathcal{N}}^T W Q_{\mathcal{N}}\|_F} \geq \frac{1}{\|W\|_F}. \tag{A.22}$$

Let $\vartheta_d := \big(U_{S^a}, y^a, U_X, Z_S, W, Q_{\mathcal{B}}, Q_{\mathcal{N}}\big)$, where $Z_S \in \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}}$, and consider the solution set

$$V(f_d) := \Big\{\vartheta_d \mid f_d(\vartheta_d) = 0\Big\},$$

where the integral polynomial map

$$f_d : \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}} \times \mathbb{R}^m \times \mathbb{R}^{n_{\mathcal{B}} \times n_{\mathcal{B}}} \times \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n_{\mathcal{B}}} \times \mathbb{R}^{n \times n_{\mathcal{N}}} \rightarrow \mathbb{R}^{t_d}$$

[1]Note that there is no need to add a symmetrization constraint. One can easily check that $\frac{W+W^T}{2}$ is a symmetric feasible solution for (A.21).

is defined as

$$
f_d(\vartheta_d) := \begin{pmatrix} \mathrm{vec}\big(Z_S - Q_{\mathcal{N}}^T W Q_{\mathcal{N}}\big) \\ \mathrm{vec}\big(U_{S^a} Z_S - I_{n_{\mathcal{N}}}\big) \\ A^1 \bullet W \\ \vdots \\ A^m \bullet W \\ A^1 \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T - b_1 \\ \vdots \\ A^m \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T - b_m \\ \mathrm{vec}\Big(C - \sum_{i=1}^m y_i^a A^i - Q_{\mathcal{N}} U_{S^a} Q_{\mathcal{N}}^T\Big) \\ \mathrm{vec}\big(W - W^T\big) \\ \mathrm{vec}\big(Q_{\mathcal{B}}^T Q_{\mathcal{B}} - I_{n_{\mathcal{B}}}\big) \\ \mathrm{vec}\big(Q_{\mathcal{N}}^T Q_{\mathcal{N}} - I_{n_{\mathcal{N}}}\big) \\ \mathrm{vec}\big(Q_{\mathcal{B}}^T Q_{\mathcal{N}}\big) \end{pmatrix}, \tag{A.23}
$$

in which $t_d = n_{\mathcal{B}}^2 + 3n_{\mathcal{N}}^2 + n_{\mathcal{B}} n_{\mathcal{N}} + 2n^2 + 2m$. By the interior point condition, the set of solutions of $V(f_d) \cap \Omega_d$ is nonempty, where Ω_d is defined as

$$
\Omega_d := \Big\{\vartheta_d \mid \det(U_X[i]) > 0, \ \ \det(U_{S^a}[j]) > 0, \ \ i = 1, \ldots, n_{\mathcal{B}}, \ j = 1, \ldots, n_{\mathcal{N}}\Big\}.
$$

Then, analogous to the primal case, from a solution $\vartheta_d \in V(f_d) \cap \Omega_d$ we can get a solution $(U_{S^a}, y^a, W, Q_{\mathcal{N}})$ with symmetric W, which is the analytic center of the dual optimal set. Therefore, Lemma A.3.1 implies the existence of $\vartheta_d \in V(f_d) \cap \Omega_d$

so that $\|\vartheta_d\|_2 \leq r_{\mathcal{D}^*_{\mathrm{SDO}}}$, where

$$
\begin{aligned}
\log_2(r_{\mathcal{D}^*_{\mathrm{SDO}}}) &= (L+2)(\bar{t}_d \bar{d}_d)^{\bar{s}_d},\\
\bar{d}_d &:= \max\{n_{\mathcal{B}}, n_{\mathcal{N}}, 3\} \leq \max\{n, 3\},\\
\bar{t}_d &:= t_d + n_{\mathcal{B}} + n_{\mathcal{N}} = n_{\mathcal{B}}^2 + 3n_{\mathcal{N}}^2 + n_{\mathcal{B}} n_{\mathcal{N}} + n_{\mathcal{B}} + n_{\mathcal{N}} + 2n^2 + 2m\\
&\leq 7n^2 + 2n + 2m,\\
\bar{s}_d &:= n_{\mathcal{B}}^2 + 2n_{\mathcal{N}}^2 + n^2 + n(n_{\mathcal{B}} + n_{\mathcal{N}}) + m \leq 6n^2 + m,
\end{aligned}
$$

in which $\bar{s}_d$ and $\bar{d}_d$ are defined analogously as in the primal side. As a result, a lower bound on $\sigma_{\mathcal{N}}$ is given by using $\|W\|_F \leq r_{\mathcal{D}^*_{\mathrm{SDO}}}$ and (A.22). This completes the proof. $\square$

Remark A.3.1. *For the special case $n_{\mathcal{B}} = 0$ we get $\sigma = \sigma_{\mathcal{N}}$ by (3.1), and thus the lower bound (A.11) is still valid. Indeed, any dual feasible solution is also dual optimal for this special case. Thus, to derive a lower bound on $\sigma_{\mathcal{N}}$ we only need to compute the analytic center of the dual feasible set $\mathcal{D}_{\mathrm{SDO}}$, i.e.,*

$$
\begin{aligned}
\max \quad & \log\big(\det(S^a)\big)\\
\text{s.t.} \quad & \sum_{i=1}^{m} y_i^a A^i + S^a = C, \\
& S^a \succ 0.
\end{aligned}
\tag{A.24}
$$

It it easy to verify that the application of Lemma A.3.1 to the system of optimality conditions of (A.24) gives an integral polynomial map with strictly fewer number of polynomials and variables than (A.23), which yields a smaller $r_{\mathcal{D}^*_{\mathrm{SDO}}}$.

A.4 Theorems of the Newton method

Quadratic convergence of Newton's method:

Theorem A.4.1 (Theorem 5.2.1 in [39]). *Consider a continuously differentiable mapping $G : \mathbb{R}^n \to \mathbb{R}^n$ on an open convex set $\mathcal{D} \subseteq \mathbb{R}^n$. Let $x^* \in \mathbb{R}^n$ be a root of $G(x) = 0$ so that $B_r(x^*) \subseteq \mathcal{D}$ for some $r > 0$. If ∇G is Lipschitz continuous with constant τ on $B_r(x^*)$ and $\|\nabla G(x^*)^{-1}\|_2 \leq \theta$ for some $\theta > 0$, then for a given $x^{(0)} \in B_{r_n}(x^*)$, where*

$$r_n := \min\left\{r, \frac{1}{2\theta\tau}\right\}, \tag{A.25}$$

the Newton iterates $x^{(k)}$ are well-defined and converge to x^* so that

$$\|x^{(k+1)} - x^*\|_2 \leq \theta\tau\|x^{(k)} - x^*\|_2^2, \qquad k \geq 0.$$

Kantorovich theorem:

Theorem A.4.2 (Theorem 5.3.1 in [39]). *Given a solution $x_0 \in \mathbb{R}^n$, let $G : \mathbb{R}^n \to \mathbb{R}^n$ be a mapping which is continuously differentiable on $B_r(x_0)$. Assume that $\nabla G(x_0)$ is nonsingular and Lipschitz continuous with Lipschitz constant τ on $B_r(x_0)$. Furthermore, let*

$$\left\|\nabla G^{-1}(x_0)\right\|_2 \leq \theta, \qquad \left\|\nabla G^{-1}(x_0)G(x_0)\right\|_2 \leq \eta.$$

for some $\theta, \eta > 0$. If $\tau\theta\eta \leq \frac{1}{2}$ and $(1 - \sqrt{1 - 2\tau\theta\eta})/(\theta\tau) \leq r$, then there exists a solution x^* to $G(x) = 0$ such that

$$\|x^* - x_0\|_2 \leq \frac{1 - \sqrt{1 - 2\tau\theta\eta}}{\theta\tau}.$$

Implicit function theorem:

Theorem A.4.3 (Theorem 2.4.1 in [50]). *Assume that $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a k-times continuously differentiable mapping. Furthermore, assume that $(\bar{x}, \bar{\omega}) \in \mathbb{R}^n \times \mathbb{R}^m$, $F(\bar{x}, \bar{\omega}) = 0$, and the Jacobian of F with respect to x is nonsingular at $(\bar{x}, \bar{\omega})$. Then there exist $\varsigma > 0$ and a unique k-times continuously differentiable mapping $\Phi : B_\varsigma(\bar{\omega}) \to \mathbb{R}^n$ so that $\Phi(\bar{\omega}) = \bar{x}$ and $F(\Phi(\omega), \omega) = 0$ for every $\omega \in B_\varsigma(\bar{\omega})$.*

Bibliography

[1] I. Adler and F. Alizadeh, *Primal-dual interior point algorithms for convex quadratically constrained and semidefinite optimization problems*, Tech. Rep. RRR-111-95, RUTCOR, Rutgers University, USA, 1995.

[2] I. Adler and R. D. C. Monteiro, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Mathematical Programming, 50 (1991), pp. 29–51.

[3] ——, *A geometric view of parametric linear programming*, Algorithmica, 8 (1992), pp. 161–176.

[4] F. Alizadeh, *Combinatorial Optimization with Interior Point Methods and Semidefinite Matrices*, PhD thesis, University of Minnesota, 1991.

[5] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming, 95 (2003), pp. 3–51.

[6] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, *Complementarity and nondegeneracy in semidefinite programming*, Mathematical Programming, 77 (1997), pp. 111–128.

[7] F. Alizadeh, J.-P. A. 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), pp. 746–768.

[8] F. Alizadeh and S. Schmieta, *Optimization with semidefinite, quadratic and linear constraints*, Tech. Rep. RRR 23-97, RUTCOR, Rutgers University, USA, 1997.

[9] F. Alizadeh and S. Schmieta, *Symmetric cones, potential reduction methods and word-by-word extensions*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., vol. 27 of International Series in Operations Research & Management Science, Springer, 2000, pp. 195–233.

[10] B. Bank, J. Guddat, and D. Klatte, *Non-linear Parametric Optimization*, Birkhäuser Verlag, 1983.

[11] G. Barker and D. Carlson, *Cones of diagonally dominant matrices*, Pacific Journal of Mathematics, 57 (1975), pp. 15–32.

[12] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons, 2006.

[13] A. Ben-Israel, A. Ben-Tal, and S. Zlobec, *Optimality in convex programming: A feasible directions approach*, in Optimality and Stability in Mathematical Programming, M. Guignard, ed., Springer, Berlin, Heidelberg, 1982, pp. 16–38.

[14] A. Ben-Tal and A. Nemirovskii, *Lectures on Modern Convex Optimization*, SIAM, 2001.

[15] C. Berge, *Topological Spaces: Including a Treatment of Multi-valued Functions, Vector Spaces, and Convexity*, Dover Publications, 1997.

[16] A. Berkelaar, B. Jansen, K. Roos, and T. Terlaky, *Sensitivity analysis in (degenerate) quadratic programming*, Tech. Rep. 96-26, Delft University of Technology, Netherlands, 1996.

[17] A. B. Berkelaar, K. Roos, and T. Terlaky, *The optimal set and optimal partition approach to linear and quadratic programming*, in Advances in Sensitivity Analysis and Parametric Programming, T. Gal and H. J. Greenberg, eds., vol. 6 of International Series in Operations Research & Management Science, Springer, 1997, pp. 159–202.

[18] J. H. Bigelow and N. Z. Shapiro, *Implicit function theorems for mathematical programming and for systems of inequalities*, Mathematical Programming, 6 (1974), pp. 141–156.

[19] G. Blekherman, P. A. Parrilo, and R. R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*, SIAM, 2013.

[20] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Sensitivity analysis of optimization problems under second order regular constraints*, Mathematics of Operations Research, 23 (1998), pp. 806–831.

[21] J. F. Bonnans and H. Ramírez, *Perturbation analysis of second-order cone programming problems*, Mathematical Programming, 104 (2005), pp. 205–227.

[22] J. F. Bonnans and A. Shapiro, *Optimization problems with perturbations: A guided tour*, SIAM Review, 40 (1998), pp. 228–264.

[23] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, 2000.

[24] Z. X. Chan and D. Sun, *Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming*, SIAM Journal on Optimization, 19 (2008), pp. 370–396.

[25] Y.-L. Cheung, *Preprocessing and Reduction for Semidefinite Programming via Facial Reduction: Theory and Practice*, PhD thesis, Department of Combinatorics & Optimization, University of Waterloo, 2013.

[26] Y.-L. Cheung, S. Schurr, and H. Wolkowicz, *Preprocessing and regularization for degenerate semidefinite programs*, in Computational and Analytical Mathematics, In Honor of Jonathan Borwein's 60th Birthday, D. Bailey, H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, eds., vol. 50 of Springer Proceedings in Mathematics & Statistics, Springer, 2013, pp. 613–634.

[27] Y.-L. Cheung and H. Wolkowicz, *Sensitivity analysis of semidefinite programs without strong duality*, tech. rep., 2014. http://www.optimization-online.org/DB_HTML/2014/06/4416.html.

[28] X. Chi and S. Liu, *A non-interior continuation method for second-order cone programming*, Optimization, 58 (2009), pp. 965–979.

[29] ——, *A one-step smoothing Newton method for second-order cone programming*, Journal of Computational and Applied Mathematics, 223 (2009), pp. 114 – 123.

[30] J. M. Danskin, *The theory of max-min, with applications*, SIAM Journal on Applied Mathematics, 14 (1966), pp. 641–664.

[31] G. B. Dantzig, J. Folkman, and N. Shapiro, *On the continuity of the*

minimum set of a continuous function, Journal of Mathematical Analysis and Applications, 17 (1967), pp. 519 – 548.

[32] E. de Klerk, *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, vol. 65 of the Series Applied Optimization, Springer, 2002.

[33] E. de Klerk, J. Peng, C. Roos, and T. Terlaky, *A scaled Gauss–Newton primal-dual search direction for semidefinite optimization*, SIAM Journal on Optimization, 11 (2001), pp. 870–888.

[34] E. de Klerk, C. Roos, and T. Terlaky, *Initialization in semidefinite programming via a self-dual skew-symmetric embedding*, Operations Research Letters, 20 (1997), pp. 213–221.

[35] E. de Klerk, C. Roos, and T. Terlaky, *Polynomial primal-dual affine scaling algorithms in semidefinite programming*, Journal of Combinatorial Optimization, 2 (1998), pp. 51–69.

[36] S. Dempe, *Directional differentiability of optimal solutions under Slater's condition*, Mathematical Programming, 59 (1993), pp. 49–69.

[37] D. den Hertog, *Interior Point Approach to Linear, Quadratic and Convex Programming*, vol. 277 of Series Mathematics and its Applications, Springer, 1994.

[38] S. Deng and H. Hu, *Computable error bounds for semidefinite programming*, Journal of Global Optimization, 14 (1999), pp. 105–115.

[39] J. E. Dennis Jr and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, 1983.

[40] A. Dontchev and R. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM Journal on Optimization, 6 (1996), pp. 1087–1105.

[41] J. P. Evans and F. J. Gould, *Stability in nonlinear programming*, Operations Research, 18 (1970), pp. 107–118.

[42] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. 1, Springer, 2003.

[43] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, 1994.

[44] L. Faybusovich, *Euclidean Jordan algebras and interior point algorithms*, Positivity, 1 (1997), pp. 331–357.

[45] L. Faybusovich, *Linear systems in Jordan algebras and primal-dual interior-point algorithms*, Journal of Computational and Applied Mathematics, 86 (1997), pp. 149 – 175.

[46] L. Faybusovich, *A Jordan algebraic approach to potential-reduction algorithms*, Mathematische Zeitschrift, 239 (2002), pp. 117–129.

[47] A. Fiacco and G. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, vol. 65 of the series Applied Optimization, 1990.

[48] A. V. Fiacco, *Convergence properties of local solutions of sequences of mathematical programming problems in general spaces*, Journal of Optimization Theory and Applications, 13 (1974), pp. 1–12.

[49] A. V. Fiacco, *Sensitivity analysis for nonlinear programming using penalty methods*, Mathematical Programming, 10 (1976), pp. 287–311.

[50] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, vol. 165, Academic Press Inc., 1983.

[51] K. Frisch, *The logarithmic potential method for solving linear programming problems*, Tech. Rep. TR91-13, Department of Industrial Engineering and Management Sciences, Northwestern University, USA, 1991.

[52] J. Gauvin, *The generalized gradient of a marginal function in mathematical programming*, Mathematics of Operations Research, 4 (1979), pp. 458–463.

[53] J. Gauvin and F. Dubeau, *Differential properties of the marginal function in mathematical programming*, in Optimality and Stability in Mathematical Programming, M. Guignard, ed., Springer, Berlin, Heidelberg, 1982, pp. 101–119.

[54] ——, *Some examples and counterexamples for the stability analysis of nonlinear programming problems*, in Sensitivity, Stability and Parametric Analysis, A. V. Fiacco, ed., Springer, Berlin, Heidelberg, 1984, pp. 69–78.

[55] J. Gauvin and R. Janin, *Directional behaviour of optimal solutions in nonlinear mathematical programming*, Mathematics of Operations Research, 13 (1988), pp. 629–649.

[56] J. Gauvin and J. Tolle, *Differential stability in nonlinear programming*, SIAM Journal on Control and Optimization, 15 (1977), pp. 294–311.

[57] A. M. Geoffrion, *Generalized Benders decomposition*, Journal of Optimization Theory and Applications, 10 (1972), pp. 237–260.

[58] D. Goldfarb and K. Scheinberg, *Interior point trajectories in semidefinite programming*, SIAM Journal on Optimization, 8 (1998), pp. 871–886.

[59] D. Goldfarb and K. Scheinberg, *On parametric semidefinite programming*, Applied Numerical Mathematics, 29 (1999), pp. 361 – 377.

[60] A. Goldman and A. Tucker, *Theory of linear programming*, in Linear Equalities and Related Systems, H. Kuhn and A. Tucker, eds., Princeton University Press, 1956, pp. 53–97.

[61] E. Gol'shtein, *Theory of Convex Programming*, vol. 36 of Translations of Mathematical Monographs, American Mathematical Society, 1972.

[62] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 2013.

[63] C. C. Gonzaga, *Path-following methods for linear programming*, SIAM Review, 34 (1992), pp. 167–224.

[64] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, V. Blondel, S. Boyd, and H. Kimura, eds., Lecture Notes in Control and Information Sciences, Springer, 2008, pp. 95–110. http://stanford.edu/~boyd/graph_dcp.html.

[65] ——, *CVX: Matlab software for disciplined convex programming, version 2.1*. http://cvxr.com/cvx, Mar. 2014.

[66] H. J. Greenberg, *The use of the optimal partition in a linear programming solution for postoptimal analysis*, Operations Research Letters, 15 (1994), pp. 179 – 185.

[67] G. Gu, M. Zangiabadi, and C. Roos, *Full nesterov-todd step infeasible interior-point method for symmetric optimization*, European Journal of Operational Research, 214 (2011), pp. 473 – 484.

[68] O. Güler, *Limiting behavior of weighted central paths in linear programming*, Mathematical Programming, 65 (1994), pp. 347–363.

[69] O. Güler, *Barrier functions in interior point methods*, Mathematics of Operations Research, 21 (1996), pp. 860–885.

[70] O. Güler, *Foundations of Optimization*, Springer, 2010.

[71] O. Güler, D. den Hertog, C. Roos, T. Terlaky, and T. Tsuchiya, *Degeneracy in interior point methods for linear programming: a survey*, Annals of Operations Research, 46 (1993), pp. 107–138.

[72] O. Güler, A. Hoffman, and U. Rothblum, *Approximations to solutions to systems of linear inequalities*, SIAM Journal on Matrix Analysis and Applications, 16 (1995), pp. 688–696.

[73] J.-P. Haeberly, *Remarks on nondegeneracy in mixed semidefinite-quadratic programming*, 1998.

[74] M. Halická, *Analytical properties of the central path at boundary point in linear programming*, Mathematical Programming, 84 (1999), pp. 335–355.

[75] ——, *Analyticity of the central path at the boundary point in semidefinite programming*, European Journal of Operational Research, 143 (2002), pp. 311–324.

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

[77] C. Helmberg, *Semidefinite programming*, European Journal of Operational Research, 137 (2002), pp. 461 – 482.

[78] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM Journal on Optimization, 6 (1996), pp. 342–361.

[79] J.-B. Hiriart-Urruty and H. Y. Le, *A variational approach of the rank function*, TOP, 21 (2013), pp. 207–240.

[80] A. Hoffman, *On approximate solutions of systems of linear inequalities*, Journal of Research of the National Bureau of Standards, 49 (1952), pp. 263–265.

[81] W. Hogan, *Directional derivatives for extremal-value functions with applications to the completely convex case*, Operations Research, 21 (1973), pp. 188–209.

[82] W. W. Hogan, *Point-to-set maps in mathematical programming*, SIAM Review, 15 (1973), pp. 591–603.

[83] W. W. Hogan, *The continuity of the perturbation function of a convex program*, Operations Research, 21 (1973), pp. 351–352.

[84] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 2012.

[85] T. Illés, J. Peng, C. Roos, and T. Terlaky, *A strongly polynomial rounding procedure yielding a maximally complementary solution for $P_*(\kappa)$ linear complementarity problems*, SIAM Journal on Optimization, 11 (2000), pp. 320–340.

[86] B. Jansen, C. Roos, and T. Terlaky, *A polynomial primal-dual Dikin-type algorithm for linear programming*, Mathematics of Operations Research, 21 (1996), pp. 341–353.

[87] B. Jansen, K. Roos, and T. Terlaky, *An interior point method approach to postoptimal and parametric analysis in linear programming*, Tech. Rep. 92-21, Delft University of Technology, Netherlands, 1993.

[88] K. Jittorntrum, *Solution point differentiability without strict complementarity in nonlinear programming*, in Sensitivity, Stability and Parametric Analysis, A. V. Fiacco, ed., Springer, Berlin, Heidelberg, 1984, pp. 127–138.

[89] ——, *Sequential Algorithms in Nonlinear Programming*, PhD thesis, Australian National University, Canberra, 1993.

[90] C. Kanzow, I. Ferenczi, and M. Fukushima, *On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity*, SIAM Journal on Optimization, 20 (2009), pp. 297–320.

[91] C. Kanzow and C. Nagel, *Quadratic convergence of a nonsmooth Newton-type method for semidefinite programs without strict complementarity*, SIAM Journal on Optimization, 15 (2005), pp. 654–672.

[92] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.

[93] D. Klatte and B. Kummer, *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, vol. 60, Springer, 2002.

[94] M. Kojima, *Strongly stable stationary solutions in nonlinear programs*, in Analysis and Computation of Fixed Points, S. M. Robinson, ed., Academic Press, 1980, pp. 93 – 138.

[95] 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), pp. 86–125.

[96] L. Kong, *Quadratic convergence of a smoothing Newton method for symmetric cone programming without strict complementarity*, Positivity, 16 (2012), pp. 297–319.

[97] L. Kong, L. Tunçel, and N. Xiu, *Equivalent conditions for Jacobian nonsingularity in linear symmetric cone programming*, Journal of Optimization Theory and Applications, 148 (2011), pp. 364–389.

[98] S. Kruk, M. Muramatsu, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, *The Gauss-Newton direction in semidefinite programming*, Optimization Methods and Software, 15 (2001), pp. 1–28.

[99] E.-Y. Lee, *Extension of Rotfel'd theorem*, Linear Algebra and its Applications, 435 (2011), pp. 735 – 741.

[100] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Applications of second-order cone programming*, Linear Algebra and its Applications, 284 (1998), pp. 193 – 228.

[101] L.-Z. Lu and C. Pearce, *Some new bounds for singular values and eigenvalues of matrix products*, Annals of Operations Research, 98 (2000), pp. 141–148.

[102] X.-D. Luo and Z.-Q. Luo, *Extension of Hoffman's error bound to polynomial systems*, SIAM Journal on Optimization, 4 (1994), pp. 383–392.

[103] Z.-Q. Luo and J. F. Sturm, *Error bounds for mixed semidefinite and second order cone programming*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., vol. 27 of International Series in Operations Research & Management Science, Springer, 2000, pp. 163–190.

[104] ——, *Error bounds for quadratic systems*, in High performance optimization, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., vol. 33 of the series Applied Optimization, Springer, 2000, pp. 383–404.

[105] Z.-Q. Luo, J. F. Sturm, and S. Zhang, *Duality results for conic convex programming*, Tech. Rep. 9719/A, Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands, 1997.

[106] 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), pp. 59–81.

[107] O. Mangasarian and T. Shiau, *Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems*, SIAM Journal on Control and Optimization, 25 (1987), pp. 583–595.

[108] O. L. Mangasarian, *A condition number for differentiable convex inequalities*, Mathematics of Operations Research, 10 (1985), pp. 175–179.

[109] S. Mehrotra and Y. Ye, *Finding an interior point in the optimal face of linear programs*, Mathematical Programming, 62 (1993), pp. 497–515.

[110] H. Mittelmann, *An independent benchmarking of SDP and SOCP solvers*, Mathematical Programming, 95 (2003), pp. 407–430.

[111] S. Mizuno, *A superlinearly convergent infeasible-interior-point algorithm for geometrical LCPs without a strictly complementary condition*, Mathematics of Operations Research, 21 (1996), pp. 382–400.

[112] A. Mohammad-Nezhad and T. Terlaky, *On the identification of optimal partition for semidefinite optimization*, Tech. Rep. 17T-008, Dept. of ISE, Lehigh U., 2017. http://www.optimization-online.org/DB_HTML/2016/09/5626.html.

[113] A. Mohammad-Nezhad and T. Terlaky, *A polynomial primal-dual affine scaling algorithm for symmetric conic optimization*, Computational Optimization and Applications, 66 (2017), pp. 577–600.

[114] A. Mohammad-Nezhad and T. Terlaky, *Quadratic convergence to the optimal solution of second-order conic optimization without strict complementarity*, Tech. Rep. 17T-014, Department of Industrial and Systems Engineering, Lehigh University, 2017. http://www.optimization-online.org/DB_HTML/2017/10/6300.html.

[115] ——, *A rounding procedure for semidefinite optimization*, Tech. Rep. 17T-009, Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, USA, 2017. http://www.optimization-online.org/DB_HTML/2017/07/6114.html.

[116] ——, *Parametric analysis of semidefinite optimization*, 2018. arXiv:1808.00587 https://arxiv.org/abs/1808.00587.

[117] R. Monteiro, I. Adler, and M. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*, Mathematics of Operations Research, 15 (1990), pp. 191–214.

[118] R. Monteiro and M. J. Todd, *Path-following methods*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., vol. 27 of International Series in Operations Research & Management Science, Springer, 2000, pp. 267–306.

[119] R. D. Monteiro and T. Tsuchiya, *Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions*, Mathematical Programming, 88 (2000), pp. 61–83.

[120] R. D. C. Monteiro, *Primal–dual path-following algorithms for semidefinite programming*, SIAM Journal on Optimization, 7 (1997), pp. 663–678.

[121] R. D. C. Monteiro, *First- and second-order methods for semidefinite programming*, Mathematical Programming, 97 (2003), pp. 209–244.

[122] M. V. Nayakkankuppam and M. L. Overton, *Conditioning of semidefinite programs*, Mathematical Programming, 85 (1999), pp. 525–540.

[123] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, vol. 13, Society for Industrial and Applied Mathematics, 1994.

[124] Y. Nesterov and M. J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research, 22 (1997), pp. 1–42.

[125] Y. E. Nesterov and M. J. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM Journal on Optimization, 8 (1998), pp. 324–364.

[126] J. Nie, K. Ranestad, and B. Sturmfels, *The algebraic degree of semidefinite programming*, Mathematical Programming, 122 (2010), pp. 379–405.

[127] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, second ed., 2006.

[128] M. Nunez and R. Freund, *Condition-measure bounds on the behavior of the central trajectory of a semidefinite program*, SIAM Journal on Optimization, 11 (2001), pp. 818–836.

[129] J. Ortega and W. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, vol. 30 of Classics in Applied Mathematics, SIAM, 1970.

[130] J.-S. Pang, *Error bounds in mathematical programming*, Mathematical Programming, 79 (1997), pp. 299–332.

[131] G. Pataki, *The geometry of semidefinite programming*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., vol. 27 of International Series in Operations Research & Management Science, Springer, 2000, pp. 29–65.

[132] ——, *Strong duality in conic linear programming: Facial reduction and extended duals*, in Computational and Analytical Mathematics, D. Bailey, H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, eds., vol. 50 of Springer Proceedings in Mathematics & Statistics, Springer, 2013, pp. 613–634.

[133] J. Peña and V. Roshchina, *A complementarity partition theorem for multifold conic systems*, Mathematical Programming, 142 (2013), pp. 579–589.

[134] J. Peng, C. Roos, and T. Terlaky, *Self-Regularity: A New Paradigm for Primal-Dual Interior Point Algorithms*, Princeton Series in Applied Mathematics, 2002.

[135] F. A. Potra, *A superlinearly convergent predictor-corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{n}L)$-iteration complexity*, Mathematical Programming, 100 (2004), pp. 317–337.

[136] F. A. Potra and S. J. Wright, *Interior-point methods*, Journal of Computational and Applied Mathematics, 124 (2000), pp. 281–302.

[137] M. V. Ramana, *An algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems*, PhD thesis, Johns Hopkins University, 1993.

[138] M. V. Ramana, *An exact duality theory for semidefinite programming and its complexity implications*, Mathematical Programming, 77 (1997), pp. 129–162.

[139] B. Rangarajan, *Polynomial convergence of infeasible-interior point methods over symmetric cones*, SIAM Journal on Optimization, 16 (2006), pp. 1211–1229.

[140] J. Renegar, *A Mathematical View of Interior-Point Methods in Convex Optimization*, SIAM, 2001.

[141] S. Robinson, *Stability theory for systems of inequalities, part II: Differentiable nonlinear systems*, SIAM Journal on Numerical Analysis, 13 (1976), pp. 497–513.

[142] S. M. Robinson, *An application of error bounds for convex programming in a linear space*, SIAM Journal on Control, 13 (1975), pp. 271–273.

[143] S. M. Robinson, *Generalized equations and their solutions, part I: Basic theory*, in Point-to-Set Maps and Mathematical Programming, P. Huard, ed., Springer, Berlin, Heidelberg, 1979, pp. 128–141.

[144] S. M. Robinson, *Strongly regular generalized equations*, Mathematics of Operations Research, 5 (1980), pp. 43–62.

[145] S. M. Robinson, *Generalized equations and their solutions, part II: Applications to nonlinear programming*, in Optimality and Stability in Mathematical Programming, M. Guignard, ed., Springer, Berlin, Heidelberg, 1982, pp. 200–221.

[146] S. M. Robinson and R. H. Day, *A sufficient condition for continuity of optimal sets in mathematical programming*, Journal of Mathematical Analysis and Applications, 45 (1974), pp. 506 – 511.

[147] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[148] ——, *Conjugate Duality and Optimization*, vol. Regional Series in Applied Mathematics, SIAM, 1974.

[149] R. T. Rockafellar and A. Dontchev, *Implicit Functions and Solution Mappings*, Springer, 2014.

[150] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317, Springer, 2009.

[151] C. Roos, T. Terlaky, and J.-P. Vial, *Interior Point Methods for Linear Optimization*, Springer, 2005.

[152] K. Scheinberg, *Parametric linear semidefinite programming*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe,

eds., vol. 27 of International Series in Operations Research & Management Science, Springer, 2000, pp. 93–110.

[153] S. Schmieta, *Application of Jordan algebras to the design and analysis of interior-point algorithms for linear, quadratically constrained quadratic, and semidefinite programming*, PhD thesis, RUTCOR, Rutgers University, USA, 1999.

[154] S. Schmieta and F. Alizadeh, *Extension of commutative class of primal-dual interior-point algorithms to symmetric cones*, Tech. Rep. RRR 13-99, RUTCOR, Rutgers University, USA, 1999.

[155] ——, *Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones*, Mathematics of Operations Research, 26 (2001), pp. 543–564.

[156] ——, *Extension of polynomial time interior-point algorithms to symmetric cones*, Mathematical Programming, 96 (2003), pp. 409–438.

[157] Y. Sekiguchi and H. Waki, *Perturbation analysis of singular semidefinite programs and its applications to control problems*, 2016. arXiv:1607.05568.

[158] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Mathematical Programming, 77 (1997), pp. 301–320.

[159] M. H. Stern and D. M. Topkis, *Rates of stability in nonlinear programming*, Operations Research, 24 (1976), pp. 462–476.

[160] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review, 15 (1973), pp. 727–764.

[161] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.

[162] J. Stoer and M. Wechs, *Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity*, Mathematical Programming, 83 (1998), pp. 407–423.

[163] ——, *On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems*, Numerische Mathematik, 81 (1999), pp. 631–645.

[164] J. Stoer, M. Wechs, and S. Mizuno, *High order infeasible-interior-point methods for solving sufficient linear complementarity problems*, Mathematics of Operations Research, 23 (1998), pp. 832–862.

[165] J. F. Sturm, *Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists*, Mathematics of Operations Research, 24 (1999), pp. 72–94.

[166] J. F. Sturm, *Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones*, Optimization Methods and Software, 11 (1999), pp. 625–653. Available at http://sedumi.ie.lehigh.edu/.

[167] J. F. Sturm, *Error bounds for linear matrix inequalities*, SIAM Journal on Optimization, 10 (2000), pp. 1228–1248.

[168] J. F. Sturm, *Similarity and other spectral relations for symmetric cones*, Linear Algebra and Its Applications, 312 (2000), pp. 135–154.

[169] J. F. Sturm, *Implementation of interior point methods for mixed semidefinite and second order cone optimization problems*, Optimization Methods and Software, 17 (2002), pp. 1105–1154.

[170] J. F. Sturm and S. Zhang, *On sensitivity of central solutions in semidefinite programming*, Mathematical Programming, 90 (2001), pp. 205–227.

[171] T. Terlaky and Z. Wang, *On the identification of the optimal partition of second order cone optimization problems*, SIAM Journal on Optimization, 24 (2014), pp. 385–414.

[172] M. J. Todd, *Semidefinite optimization*, Acta Numerica, 10 (2001), pp. 515–560.

[173] T. Tsuchiya, *A polynomial primal-dual path-following algorithm for second-order cone programming*, Tech. Rep. 649, The Institute of Statistical Mathematics, Tokyo, Japan, 1997.

[174] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, 11 (1999), pp. 141–182.

[175] L. Tunçel and H. Wolkowicz, *Strong duality and minimal representations for cone optimization*, Computational Optimization and Applications, 53 (2012), pp. 619–648.

[176] R. H. Tütüncü, K. C. Toh, and M. J. Todd, *Solving semidefinite-quadratic-linear programs using SDPT3*, Mathematical Programming, 95 (2003), pp. 189–217.

[177] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Review, 38 (1996), pp. 49–95.

[178] T. Wang and J.-S. Pang, *Global error bounds for convex quadratic inequality systems*, Optimization, 31 (1994), pp. 1–12.

[179] H. Wei and H. Wolkowicz, *Generating and measuring instances of hard semidefinite programs*, Mathematical Programming, 125 (2010), pp. 31–45.

[180] Y. Ye, *On the finite convergence of interior-point algorithms for linear programming*, Mathematical Programming, 57 (1992), pp. 325–335.

[181] Y. Ye, *Toward probabilistic analysis of interior-point algorithms for linear programming*, Mathematics of Operations Research, 19 (1994), pp. 38–52.

[182] Y. Ye and K. Anstreicher, *On quadratic and $O(\sqrt{n}L)$ convergence of a predictor–corrector algorithm for LCP*, Mathematical Programming, 62 (1993), pp. 537–551.

[183] Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang, *A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming*, Mathematical Programming, 59 (1993), pp. 151–162.

[184] E. Yildirim, *Unifying optimal partition approach to sensitivity analysis in conic optimization*, Journal of Optimization Theory and Applications, 122 (2004), pp. 405–423.

[185] E. A. Yildirim and M. Todd, *Sensitivity analysis in linear programming and semidefinite programming using interior-point methods*, Mathematical Programming, 90 (2001), pp. 229–261.

[186] Y. Zhang, *On extending primal-dual interior-point algorithms from linear programming to semidefinite programming*, SIAM Journal on Optimization, 8 (1998), pp. 356–386.

[187] S. Zlobec, R. Gardner, and A. Ben-Israel, *Regions of stability for arbitrarily perturbed convex programs*, in Mathematical Programming with Data Perturbations I, A. Fiacco, ed., M. Dekker, 1982, pp. 69–89.

Vita

Ali Mohammad-Nezhad is a doctoral candidate in the department of Industrial and Systems Engineering at Lehigh University. He obtained his Bachelor and Master degrees in Industrial Engineering in 2008 and 2011, respectively, from Golpayegan College of Engineering and Sharif University of Technology, Iran. His research interest revolves around conic optimization, polynomial optimization, and real algebraic geometry. He is a Rossin Doctoral fellow since 2016, and he was awarded Honorable Mention in the Van Hoesen Family Best Publication Award competition in 2016. He spent the Summer of 2016 as an Operations Research intern and the Summer 2017 as an Operations Research fellow at SAS Institute. After graduation, Ali will join the Nonconvex Optimization group in the school of Industrial Engineering at Purdue University as a postdoctoral research assistant.

Index