



ISE



Industrial and
Systems Engineering

On the identification of the optimal partition for semidefinite optimization

ALI MOHAMMAD-NEZHAD AND TAMÁS TERLAKY

Department of Industrial and Systems Engineering, Lehigh University, USA

ISE Technical Report 17T-008



LEHIGH
UNIVERSITY.

On the identification of the optimal partition for semidefinite optimization

Ali Mohammad-Nezhad and Tamás Terlaky

^aDepartment of Industrial and Systems Engineering, Lehigh University, Harold S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA 18015-1582

ARTICLE HISTORY

Compiled March 10, 2018

ABSTRACT

The concept of the optimal partition was originally introduced for linear optimization and linear complementarity problems and subsequently extended to semidefinite optimization. For linear optimization and sufficient linear complementarity problems, the optimal partition and a maximally complementary optimal solution can be identified in strongly polynomial time. In this paper, we consider the identification of the optimal partition of semidefinite optimization, for which we provide an approximation from a bounded sequence of solutions on, or in a neighborhood of the central path. 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.

KEYWORDS

Semidefinite optimization; Optimal partition; Maximally complementary optimal solution; Degree of singularity

1. Introduction

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

$$C \bullet X := \text{trace}(CX),$$

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

$$(P) \quad z_p^* := \min \{ C \bullet X \mid A^i \bullet X = b_i, \quad i = 1, \dots, m, \quad X \succeq 0 \},$$

where A^i for $i = 1, \dots, m$ are $n \times n$ symmetric matrices, $b \in \mathbb{R}^m$, and $X \in \mathbb{S}_+^n$, where \mathbb{S}_+^n denotes the cone of $n \times n$ positive semidefinite matrices. The dual SDO problem is given by

$$(D) \quad z_{d^*} := \max \left\{ b^T y \mid \sum_{i=1}^m y_i A^i + S = C, \ S \succeq 0, \ y \in \mathbb{R}^m \right\}.$$

Let \mathcal{P} and \mathcal{D} denote the primal and dual feasible sets, respectively, as follows

$$\begin{aligned} \mathcal{P} &:= \{X \mid A^i \bullet X = b_i, \ i = 1, \dots, m, \ X \succeq 0\}, \\ \mathcal{D} &:= \{(y, S) \mid \sum_{i=1}^m y_i A^i + S = C, \ S \succeq 0\}. \end{aligned}$$

In light of this notation, the primal and dual optimal sets are defined as

$$\begin{aligned} \mathcal{P}^* &:= \{X \mid X \in \mathcal{P}, \ C \bullet X = z_{p^*}\}, \\ \mathcal{D}^* &:= \{(y, S) \mid (y, S) \in \mathcal{D}, \ b^T y = z_{d^*}\}. \end{aligned}$$

Throughout this paper, the following assumptions are made:

Assumption 1. A^i for $i = 1, \dots, m$ are linearly independent.

Assumption 2. The interior point condition holds, i.e., there exists $(X^\circ, y^\circ, S^\circ) \in \mathcal{P} \times \mathcal{D}$ with $X^\circ, S^\circ \succ 0$.

Assumption 1 guarantees that y is uniquely determined for a given dual solution S . Assumption 2 ensures that the primal and dual optimal sets are nonempty, bounded, and that strong duality holds. 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 [de Klerk et al. \(1997\)](#) 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. Second-order conic optimization (SOCO) problems can be embedded in SDO formulation. See e.g., [Vandenberghe and Boyd \(1996\)](#) for a detailed description of the problems. Analogous to LO, using interior point methods (IPMs), SDO problems can be solved in polynomial time, though they require significantly more computational effort per iteration. The extension of IPMs from LO to SDO was pioneered by [Nesterov and Nemirovskii \(1994\)](#), and [Alizadeh \(1991\)](#). The main idea of primal-dual path following IPMs is to follow the central path, which is defined as the set of solutions of

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

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 $(X(\mu), y(\mu), S(\mu))$ to this system exists, and it is uniquely defined under the interior point condition and the linear independence of A^i for $i = 1, \dots, m$. For $0 \leq \mu \leq \bar{\mu}$, where $\bar{\mu} > 0$, the set of solutions of (1) is bounded, and the trajectory of the central solutions has accumulation points in the relative interior of the optimal set (Goldfarb and Scheinberg [1998]; de Klerk et al. [1997]; Luo et al. [1998]). A proof was given by Halická et al. [2002] for the fact that the central path converges to a maximally complementary optimal solution.

Definition 1.1 (Adopted from Definition 2.7 in de Klerk [2006]). Let $(X^*, y^*, S^*) \in \mathcal{P}^* \times \mathcal{D}^*$. Then (X^*, y^*, S^*) is a maximally complementary optimal pair if $\text{rank}(X^* + S^*)$ is maximal over the optimal set.

Definition 1.2 (Adopted from Definition 2.7 in de Klerk [2006]). A maximally complementary pair (X^*, y^*, S^*) is strictly complementary if $X^* + S^* \succ 0$.

The analyticity and limiting behavior of the central path for SDO have been extensively studied in the literature. Luo et al. [1998] established the superlinear convergence of an IPM for SDO under the strict complementarity assumption 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. [1998] and de Klerk et al. [1997] under the strict complementarity condition. Goldfarb and Scheinberg [1998] showed, under the strict complementarity and primal-dual nondegeneracy conditions, that the first order derivatives of the central path converge as $\mu \rightarrow 0$. However, the first order derivatives may be unbounded if strict complementarity fails to hold. Using the strict complementarity assumption only, Halická [2002] showed the extension of the analyticity of the central path to $\mu = 0$.

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 (Güler et al. [1993]). It would be helpful, like in LO and linear complementarity problem (LCP) (Roos et al. [2005]; Illés et al. [2000]), 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.

The notion of the optimal partition was originally introduced for LO and LCPs. Ye [1992] 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. [2005] presented a rounding procedure which uses the optimal partition information to identify a strictly complementary optimal solution. Under the same conditions, Illés et al. [2000] considered the identification of the optimal partition for sufficient LCPs and proposed a strongly polynomial rounding procedure to a maximally complementary optimal solution. The concept of the optimal partition was extended to SDO by Goldfarb and Scheinberg [1998] and to general linear conic optimization by Yildirim [2004]. Bonnans and Ramírez [2005] established another algebraic definition of the optimal partition for SOCO. Peña and Roshchina [2013] extended the idea of the complementarity partition for a linear system to a homogeneous convex conic system formed by regular closed convex cones. Recently, Terlaky and Wang [2014] have studied the identification of the optimal partition for SOCO. The optimal partition provides unique information about the optimal set of an SDO problem, regardless

of nondegeneracy and strict complementarity conditions.

In this paper, we consider the identification of the optimal partition for SDO. 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 aim to 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.

The rest of this paper is organized as follows. In Section 2, we review the concepts of the optimal partition and complementarity. Our main results are presented in Sections 3 and 4. In Section 3, we analyze the magnitude of the eigenvalues of the solutions on the central path based on a condition number and error bound result for linear matrix inequalities (LMIs). The latter bound enables us to determine the subsets of the eigenvectors of the central solutions whose accumulation points form orthonormal bases for the subspaces of the optimal partition. Furthermore, we measure the accuracy of the approximation of the optimal partition. In Section 4, we extend the identification results to solutions in a neighborhood of the central path and provide an iteration complexity bound for the identification of the above sets of eigenvectors. Finally, our conclusions are presented in Section 5.

Throughout this paper, \mathbb{S}^n denotes the space of symmetric matrices of size n . An arbitrary optimal solution is denoted by $(\tilde{X}, \tilde{y}, \tilde{S})$, and any maximally complementary optimal solution is indicated by $*$ superscript. Furthermore, the limit point of the central path and the analytic center of the optimal set are denoted by (X^{**}, y^{**}, S^{**}) and (X^a, y^a, S^a) , respectively. The subscript $[i]$ in our notation means the i^{th} largest component of a vector. For instance, $\lambda_{[i]}(X)$ denotes the i^{th} largest eigenvalue of X so that

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

In particular, $\lambda_{\min}(X) := \lambda_{[n]}(X)$ and $\lambda_{\max}(X) := \lambda_{[1]}(X)$ stand for the minimum and maximum eigenvalues of X , respectively. The kernel and range of a linear operator are denoted by $\text{Ker}(\cdot)$ and $\mathcal{R}(\cdot)$, respectively, and $\text{ri}(\cdot)$ stands for the relative interior of a set. Note that if A is a matrix, then $\mathcal{R}(A)$ denotes the column space of A . For a linear subspace \mathcal{L} , \mathcal{L}^\perp denotes the orthogonal complement of \mathcal{L} . By an orthogonal matrix we mean a square matrix whose columns are orthonormal, i.e., the columns are orthogonal, and they have unit length. Finally, $\|\cdot\|$ denotes the Frobenius norm and $\|\cdot\|_2$ serves as the l_2 norm and the induced 2-norm for the vectors and matrices, respectively.

2. The optimal partition for SDO

Consider the optimality conditions for (P) and (D) . Since the interior point condition holds, for optimality the KKT conditions (Nocedal and Wright [2006]) are necessary and sufficient for (P) and (D) , which are written as

$$\begin{aligned} A^i \bullet X &= b_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^m A^i y_i + S &= C, \\ XS &= 0, \quad X, S \succeq 0, \end{aligned}$$

where $XS = 0$ is referred to as the complementarity condition. A solution (X, y, S) which satisfies $XS = 0$ is called complementary.

Note that strict complementarity may fail in SDO, i.e., an SDO problem might have no strictly complementary optimal solution. See Goldfarb and Scheinberg (1998); de Klerk (2006) for further details. A maximally complementary optimal pair can be equivalently defined as a primal-dual optimal solution in the relative interior of the optimal set. As a result, all $X^* \in \text{ri}(\mathcal{P}^*)$ have the same range space. Analogously, all S^* have identical range spaces, where $(y^*, S^*) \in \text{ri}(\mathcal{D}^*)$, see e.g., Lemma 2.3 in de Klerk (2006) or Lemma 3.1 in Goldfarb and Scheinberg (1998).

Let $\mathcal{B} := \mathcal{R}(X^*)$ and $\mathcal{N} := \mathcal{R}(S^*)$, where (X^*, y^*, S^*) is a maximally complementary optimal solution. We define $n_{\mathcal{B}} := \dim(\mathcal{B})$ and $n_{\mathcal{N}} := \dim(\mathcal{N})$. Then, it follows from the above equivalence, that $\mathcal{R}(\tilde{X}) \subseteq \mathcal{B}$ and $\mathcal{R}(\tilde{S}) \subseteq \mathcal{N}$ for all $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*$. By the complementarity condition, the subspaces \mathcal{B} and \mathcal{N} are orthogonal, and this implies that $n_{\mathcal{B}} + n_{\mathcal{N}} \leq n$. In case of strict complementarity, the subspaces \mathcal{B} and \mathcal{N} span \mathbb{R}^n . Otherwise, there exists a subspace \mathcal{T} , which is the orthogonal complement to $\mathcal{B} + \mathcal{N}$, i.e., \mathbb{R}^n is partitioned into three mutually orthogonal subspaces \mathcal{B} , \mathcal{N} , and \mathcal{T} . In a similar manner, we define $n_{\mathcal{T}} := \dim(\mathcal{T})$.

Definition 2.1. The partition $(\mathcal{B}, \mathcal{T}, \mathcal{N})$ of \mathbb{R}^n is called the optimal partition of an SDO problem.

Consider a maximally complementary optimal solution (X^*, y^*, S^*) . By the complementarity condition, X^* and S^* commute, and thus they have a common eigenvector basis Q^* , i.e., we can represent X^* and S^* as

$$X^* = Q^* \Lambda(X^*) (Q^*)^T, \quad S^* = Q^* \Lambda(S^*) (Q^*)^T,$$

where $\Lambda(X^*)$ and $\Lambda(S^*)$ are diagonal matrices containing the eigenvalues of X^* and S^* , respectively. Then we have

$$\mathcal{R}(X^*) = \mathcal{R}(Q^* \Lambda(X^*)), \quad \mathcal{R}(S^*) = \mathcal{R}(Q^* \Lambda(S^*)),$$

which implies that the range spaces are spanned by the eigenvectors associated with the positive eigenvalues. In particular, 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} .

Remark 1. If the interior point condition fails for either (P) or (D) , but a primal-dual optimal solution exists, and the duality gap is 0, then the optimal partition of (P) and (D) can be recovered from the optimal partition of the problem in self-dual embedding format, see [de Klerk et al. \(1998\)](#).

Let $Q := [Q_{\mathcal{B}}, Q_{\mathcal{T}}, Q_{\mathcal{N}}]$ be an orthonormal basis partitioned according to \mathcal{B} , \mathcal{T} , and \mathcal{N} . Now, the following theorem is in order.

Theorem 2.2 (Theorem 2.7 in [de Klerk \(2006\)](#)). *For every primal-dual optimal solution $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ we can represent \tilde{X} and \tilde{S} as*

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

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 $X^* \in \text{ri}(\mathcal{P}^*)$, then there exists $U_{X^*} \succ 0$. Similarly, if $n_{\mathcal{N}} > 0$ and $(y^*, S^*) \in \text{ri}(\mathcal{D}^*)$, then there exists $U_{S^*} \succ 0$.

Notice the necessity of the condition $n_{\mathcal{B}} > 0$ or $n_{\mathcal{N}} > 0$ in Theorem [2.2](#). For instance, if $n_{\mathcal{B}} = 0$, then we have $\mathcal{P}^* = \text{ri}(\mathcal{P}^*) = \{0\}$, which implies $U_{X^*} = 0$.

Remark 2. 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 (P) , 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 $(X^*, y^*, S^*) \in \text{ri}(\mathcal{P}^* \times \mathcal{D}^*)$ with respect to Q reveals the optimal partition as

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

where $U_{X^*} \succ 0$ and $U_{S^*} \succ 0$ if $n_{\mathcal{B}}, n_{\mathcal{N}} > 0$. As a result of Theorem [2.2](#) we have

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

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 2.3. *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 [de Klerk \(2006\)](#). The result follows by noting that the set of orthogonal matrices

is compact. The compactness of $\Gamma_{\mathcal{N}}$ follows analogously. \square

The optimal partition plays a central role in the parametric and sensitivity analysis of SDO problems, see e.g., [Goldfarb and Scheinberg \(1999\)](#) and [Yildirim \(2004\)](#).

3. On the identification of the optimal partition along the central path

In this section, we aim to provide a characterization of the optimal partition using the eigenvectors of a central solution, when μ is sufficiently close to 0. In Subsection [3.1](#), we define a condition number and employ an error bound result for LMIs to derive an upper bound on the distance of a central solution to the optimal set. In Subsection [3.2](#), we proceed with the approximation of the optimal partition using the condition number and the error bound result specified in Subsection [3.1](#). In Subsection [3.3](#), we measure the accuracy of the approximation of the optimal partition.

3.1. Condition number and error bound

Recall that the central path for (P) and (D) is defined by [\(1\)](#), and assume that $Q_{\mathcal{B}}$ and $Q_{\mathcal{N}}$ are known. To derive bounds on the magnitudes of the eigenvalues of $X(\mu)$ and $S(\mu)$ as $\mu \rightarrow 0$, we define a condition number σ as

$$\sigma := \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}, \quad (2)$$

where

$$\sigma_{\mathcal{B}} := \begin{cases} \max_{\tilde{X} \in \mathcal{P}^*} \lambda_{\min}(Q_{\mathcal{B}}^T \tilde{X} Q_{\mathcal{B}}) \\ \quad = \max_{\bar{Q}_{\mathcal{B}} \in \Gamma_{\mathcal{B}}} \max_{\tilde{X} \in \mathcal{P}^*} \lambda_{\min}(\bar{Q}_{\mathcal{B}}^T \tilde{X} \bar{Q}_{\mathcal{B}}), & \text{if } n_{\mathcal{B}} > 0, \\ \infty, & \text{if } n_{\mathcal{B}} = 0, \end{cases} \quad (3)$$

$$\sigma_{\mathcal{N}} := \begin{cases} \max_{(\tilde{y}, \tilde{S}) \in \mathcal{D}^*} \lambda_{\min}(Q_{\mathcal{N}}^T \tilde{S} Q_{\mathcal{N}}) \\ \quad = \max_{\bar{Q}_{\mathcal{N}} \in \Gamma_{\mathcal{N}}} \max_{(\tilde{y}, \tilde{S}) \in \mathcal{D}^*} \lambda_{\min}(\bar{Q}_{\mathcal{N}}^T \tilde{S} \bar{Q}_{\mathcal{N}}), & \text{if } n_{\mathcal{N}} > 0, \\ \infty, & \text{if } n_{\mathcal{N}} = 0. \end{cases} \quad (4)$$

The condition number σ is indeed a generalization of the analogous condition number from LO, as introduced by [Ye \(1994\)](#).

Lemma 3.1. *The condition number σ is positive.*

Proof. By the interior point condition, $\mathcal{P}^* \times \mathcal{D}^*$ is nonempty and compact, see e.g., Lemma 3.2 in [Goldfarb and Scheinberg \(1998\)](#). Thus, σ is well-defined by Remark [2](#). Assume that $n_{\mathcal{B}} > 0$. Then there exists $\tilde{X} \in \mathcal{P}^*$ so that $\lambda_{\min}(Q_{\mathcal{B}}^T \tilde{X} Q_{\mathcal{B}}) > 0$. By the compactness of \mathcal{P}^* and the continuity of the eigenvalues, there exists $\bar{X} \in \mathcal{P}^*$ so that

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

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. In Appendix [A](#), we provide a positive lower bound on the condition number σ as

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

where

$$\begin{aligned} \log_2(r_{\mathcal{P}^*}) &= (L+2) \left(\max\{n, 3\} (6n^2 + 2n + m) \right)^{5n^2+2m}, \\ \log_2(r_{\mathcal{D}^*}) &= (L+2) \left(\max\{n, 3\} (7n^2 + 2n + 2m) \right)^{6n^2+m}, \end{aligned}$$

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.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 3.2. Consider Khachiyan's example which is adopted from [Ramana \(1997\)](#):

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

This problem can be represented in dual form (D) if we define

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

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

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

From the linear matrix inequalities 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

$i = 1, \dots, \bar{m}$. The optimal solution is unique, and it is given by $y_{i+1}^* = 4^{2^i - 2^{\bar{m}}}$ for $i = 0, \dots, \bar{m}$.

Since $y_1^* y_{i+1}^* = 4(y_i^*)^2$ and $G_i(\cdot)$ is a 2×2 matrix, we get

$$\lambda_{\max}(G_i(y^*)) = \text{trace}(G_i(y^*)) = y_1^* + y_{i+1}^* = 4^{1-2^{\bar{m}}} + 4^{2^i - 2^{\bar{m}}}, \quad i = 1, \dots, \bar{m}.$$

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

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

□

In what follows, we resort to a Hölderian error bound result for an LMI system from Theorem 3.3 in [Sturm \(2000\)](#), see also Lemma [B.1](#) in Appendix [B](#). In Lemma [3.3](#), we employ the error bound result to specify an upper bound on the distance of a central solution to the optimal set. In Subsection [3.2](#), we use this upper bound along with the condition number σ to derive bounds on the magnitude of the eigenvalues of the central solutions.

Before stating Lemma [3.3](#), we need to introduce two Hoffman condition numbers associated with the primal and dual optimal sets. To that end, let $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ be a primal-dual optimal solution. Then the primal and dual optimal sets can be equivalently written as

$$\begin{cases} \mathcal{A}X = b, \\ X \bullet \tilde{S} = 0, \\ X \succeq 0, \end{cases} \quad \begin{cases} \mathcal{A}^*y + S = C, \\ S \bullet \tilde{X} = 0, \\ S \succeq 0, \end{cases} \quad (5)$$

where $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$ is a linear map given by

$$\mathcal{A}X = \begin{bmatrix} A^1 \bullet X \\ \vdots \\ A^m \bullet X \end{bmatrix},$$

and \mathcal{A}^* is the adjoint operator of \mathcal{A} , see also Section 4 in [Sturm \(2000\)](#). Then the minimal subspaces containing the primal and dual optimal sets are given by

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

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

$$(X(\mu) - \tilde{X}) \bullet (S(\mu) - \tilde{S}) = 0,$$

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

$$X(\mu) \bullet \tilde{S} + \tilde{X} \bullet S(\mu) = n\mu.$$

All this implies that $0 \leq X(\mu) \bullet \tilde{S} \leq n\mu$ and $0 \leq \tilde{X} \bullet S(\mu) \leq n\mu$. Then the application of the Hoffman error bound (Hoffman 1952) gives

$$\begin{aligned} & \text{dist} \left(X(\mu), \{X \in \tilde{X} + \text{Ker}(\mathcal{A}) \mid X \bullet \tilde{S} = 0\} \right) \\ &= \text{dist} \left(X(\mu), \{X \mid \mathcal{A}X = b, X \bullet \tilde{S} = 0\} \right) \\ &\leq \theta_1 \left(\|\mathcal{A}X(\mu) - b\|_2 + X(\mu) \bullet \tilde{S} \right) = \theta_1 X(\mu) \bullet \tilde{S} \leq \theta_1 n\mu, \end{aligned} \quad (6)$$

where $\theta_1 > 0$ denotes the Hoffman condition number, which depends on \mathcal{A} and \tilde{S} , only. Analogously, we can derive

$$\begin{aligned} & \text{dist} \left(S(\mu), \{S \in \tilde{S} + \mathcal{R}(\mathcal{A}^*) \mid S \bullet \tilde{X} = 0\} \right) \\ &= \text{dist} \left(S(\mu), \{S \mid \exists y \in \mathbb{R}^m, \mathcal{A}^*y + S = C, S \bullet \tilde{X} = 0\} \right) \\ &\leq \theta_2 (\|\mathcal{A}^*y(\mu) + S(\mu) - C\| + S(\mu) \bullet \tilde{X}) = \theta_2 S(\mu) \bullet \tilde{X} \leq \theta_2 n\mu, \end{aligned} \quad (7)$$

where $\theta_2 > 0$ is the Hoffman condition number, which is dependent on \mathcal{A} and \tilde{X} .

Now, we present the following lemma, as planned.

Lemma 3.3. *Let $(X(\mu), y(\mu), S(\mu))$ be a central solution with*

$$\mu \leq \hat{\mu} := \frac{1}{n} \min \left\{ \theta_1^{-1}, \theta_2^{-1} \right\}. \quad (8)$$

Then there exist $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^ \times \mathcal{D}^*$, a positive condition number c independent of μ , and an exponent $\gamma > 0$ so that*

$$\|X(\mu) - X_\mu\| \leq c(n\mu)^\gamma, \quad \|S(\mu) - S_\mu\| \leq c(n\mu)^\gamma, \quad (9)$$

where γ depends on the degree of singularity¹ of $\bar{\mathcal{A}}_{\mathcal{P}^}$ and $\bar{\mathcal{A}}_{\mathcal{D}^*}$.*

Proof. The bounds in (9) can be established easily by applying the error bound result, as stated in Lemma B.1 to the LMIs in (5). As defined by (I), the set of central solutions $(X(\mu), y(\mu), S(\mu))$ for $0 < \mu \leq \hat{\mu}$ is bounded, see Lemma 3.2 in de Klerk (2006). Therefore, from Lemma B.1 and the compactness of the optimal set it follows the existence of $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$, positive condition numbers c_1 and c_2 both independent of μ , and positive exponents γ_1 and γ_2 so that

$$\|X(\mu) - X_\mu\| \leq c_1(n\mu)^{\gamma_1}, \quad \|S(\mu) - S_\mu\| \leq c_2(n\mu)^{\gamma_2},$$

where $\gamma_1 = 2^{-d(\bar{\mathcal{A}}_{\mathcal{P}^*}, \mathbb{S}_+^n)}$ and $\gamma_2 = 2^{-d(\bar{\mathcal{A}}_{\mathcal{D}^*}, \mathbb{S}_+^n)}$, in which $d(\bar{\mathcal{A}}_{\mathcal{P}^*}, \mathbb{S}_+^n)$ and $d(\bar{\mathcal{A}}_{\mathcal{D}^*}, \mathbb{S}_+^n)$ denote the degree of singularity of the subspaces $\mathcal{A}_{\mathcal{P}^*}$ and $\mathcal{A}_{\mathcal{D}^*}$, respectively, as intro-

¹In this context, the degree of singularity (Sturm 2000) 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.

duced in Appendix [B](#). Setting $\gamma := \min\{\gamma_1, \gamma_2\}$ and $c := \max\{c_1, c_2\}$ we get the result as desired. \square

Remark 4. From Lemma [B.2](#) in Appendix [B](#) we can get a nontrivial upper bound $n - 1$ on the degree of singularity. Therefore, we have $\gamma \geq \frac{1}{2^{n-1}}$ for $n \geq 2$. However, we are not aware of any method to compute an upper bound on the condition number c .

Remark 5. For the special case $\mathcal{T} = \{0\}$, the degree of singularity is at most 1 ([Sturm 2000](#)). For instance, this special case happens when we embed an LO problem in SDO. Then Lemma [3.3](#) 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 $\mathcal{O}(n\mu)$.

3.2. Approximation of the optimal partition

Consider the orthogonal transformation of $X(\mu)$ with respect to Q denoted by

$$\hat{X}(\mu) := \begin{bmatrix} \hat{X}_{\mathcal{B}}(\mu) & \hat{X}_{\mathcal{B}\mathcal{T}}(\mu) & \hat{X}_{\mathcal{B}\mathcal{N}}(\mu) \\ \hat{X}_{\mathcal{T}\mathcal{B}}(\mu) & \hat{X}_{\mathcal{T}}(\mu) & \hat{X}_{\mathcal{T}\mathcal{N}}(\mu) \\ \hat{X}_{\mathcal{N}\mathcal{B}}(\mu) & \hat{X}_{\mathcal{N}\mathcal{T}}(\mu) & \hat{X}_{\mathcal{N}}(\mu) \end{bmatrix}, \quad (10)$$

where $\hat{X}(\mu) := Q^T X(\mu) Q$. The orthogonal transformation of $S(\mu)$ is defined analogously. Since the central path converges to a maximally complementary optimal solution, from the orthogonal transformation in [\(10\)](#) we have

$$\lim_{\mu \rightarrow 0} \hat{X}_{\mathcal{B}}(\mu) = U_{X^{**}}, \quad \text{and} \quad \lim_{\mu \rightarrow 0} \hat{S}_{\mathcal{N}}(\mu) = U_{S^{**}},$$

and

$$\lim_{\mu \rightarrow 0} Q_{\mathcal{T} \cup \mathcal{N}}^T X(\mu) Q_{\mathcal{T} \cup \mathcal{N}} = 0, \quad \lim_{\mu \rightarrow 0} Q_{\mathcal{B} \cup \mathcal{T}}^T S(\mu) Q_{\mathcal{B} \cup \mathcal{T}} = 0,$$

where $\hat{X}_{\mathcal{B}}(\mu) = Q_{\mathcal{B}}^T X(\mu) Q_{\mathcal{B}}$ and $\hat{S}_{\mathcal{N}}(\mu) = Q_{\mathcal{N}}^T S(\mu) Q_{\mathcal{N}}$. The following lemma establishes upper bounds on the vanishing blocks of $\hat{X}(\mu)$ and $\hat{S}(\mu)$.

Lemma 3.4. *Let $(X(\mu), y(\mu), S(\mu))$ be a central solution with $\mu \leq \hat{\mu}$. Then we have*

$$\begin{aligned} \text{trace}(\hat{X}_{\mathcal{N}}(\mu)) &\leq \frac{n\mu}{\sigma}, & \text{trace}(\hat{S}_{\mathcal{B}}(\mu)) &\leq \frac{n\mu}{\sigma}, \\ \|Q_{\mathcal{T} \cup \mathcal{N}}^T X(\mu) Q_{\mathcal{T} \cup \mathcal{N}}\| &\leq c(n\mu)^\gamma, & \|Q_{\mathcal{B} \cup \mathcal{T}}^T S(\mu) Q_{\mathcal{B} \cup \mathcal{T}}\| &\leq c(n\mu)^\gamma. \end{aligned}$$

Proof. By the compactness of \mathcal{P}^* , and the continuity of the eigenvalues, there exists $\bar{X} \in \mathcal{P}^*$ so that $\sigma_{\mathcal{B}} = \lambda_{\min}(Q_{\mathcal{B}}^T \bar{X} Q_{\mathcal{B}})$ as defined in [\(3\)](#). Analogously, it follows from [\(4\)](#) that $\sigma_{\mathcal{N}} = \lambda_{\min}(Q_{\mathcal{N}}^T \bar{S} Q_{\mathcal{N}})$ for some $(\bar{y}, \bar{S}) \in \mathcal{D}^*$. Since $\sigma = \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}$, then there exists $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

$$\lambda_{\min}(U_{\bar{X}}) \geq \sigma, \quad \lambda_{\min}(U_{\bar{S}}) \geq \sigma, \quad (11)$$

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

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

which by (1) and optimality of \bar{X} and \bar{S} gives

$$X(\mu) \bullet \bar{S} + \bar{X} \bullet S(\mu) = n\mu.$$

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

$$X(\mu) \bullet \bar{S} + \bar{X} \bullet S(\mu) = \hat{X}_{\mathcal{N}}(\mu) \bullet U_{\bar{S}} + U_{\bar{X}} \bullet \hat{S}_{\mathcal{B}}(\mu) = n\mu,$$

where $\hat{S}_{\mathcal{B}}(\mu) = Q_{\mathcal{B}}^T S(\mu) Q_{\mathcal{B}}$ and $\hat{X}_{\mathcal{N}}(\mu) = Q_{\mathcal{N}}^T X(\mu) Q_{\mathcal{N}}$. Therefore, the positive definiteness of $\hat{X}_{\mathcal{N}}(\mu)$ gives rise to $\hat{X}_{\mathcal{N}}(\mu) \bullet U_{\bar{S}} \leq n\mu$. Furthermore, from the inequality $\lambda_{\min}(U_{\bar{S}}) \text{trace}(\hat{X}_{\mathcal{N}}(\mu)) \leq \hat{X}_{\mathcal{N}}(\mu) \bullet U_{\bar{S}}$, it immediately follows that

$$\lambda_{\min}(U_{\bar{S}}) \text{trace}(\hat{X}_{\mathcal{N}}(\mu)) \leq n\mu,$$

which by the lower bounds (11) gives

$$\text{trace}(\hat{X}_{\mathcal{N}}(\mu)) \leq \frac{n\mu}{\sigma}.$$

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

$$\text{trace}(\hat{S}_{\mathcal{B}}(\mu)) \leq \frac{n\mu}{\sigma}.$$

From Lemma 3.3, there exists $(X_{\mu}, y_{\mu}, S_{\mu}) \in \mathcal{P}^* \times \mathcal{D}^*$ so that (9) holds. Recall from Theorem 2.2 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

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

and

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

which completes the proof. \square

Let $X(\mu) = Q(\mu) \Lambda(X(\mu)) Q^T(\mu)$ and $S(\mu) = Q(\mu) \Lambda(S(\mu)) Q^T(\mu)$ be the eigenvalue decompositions of $X(\mu)$ and $S(\mu)$, where $Q(\mu)$ denotes a common eigenvector basis. We show in Theorems 3.6 and 3.7 that it is possible to identify the subsets of columns of $Q(\mu)$ whose accumulation points are orthonormal bases for the subspaces \mathcal{B} , \mathcal{N} , and \mathcal{T} , when μ is sufficiently small.

Lemma 3.5 (Theorem 4.5 in [Stewart and Sun \(1990\)](#)). Let $X \in \mathbb{S}^n$ and $Y \in \mathbb{R}^{n \times k}$. Then we have

$$\begin{aligned} \lambda_{[n-k+1]}(X) + \dots + \lambda_{[n]}(X) &= \min_Y \text{trace}(Y^T X Y), \\ \text{s.t. } Y^T Y &= I_k. \end{aligned}$$

Theorem 3.6. For a central solution $(X(\mu), y(\mu), S(\mu))$ with $\mu \leq \hat{\mu}$, where $\hat{\mu}$ is given by [\(8\)](#), it holds that:

(1) For $i = 1, \dots, n_{\mathcal{B}}$ we have

$$\lambda_{[n-i+1]}(S(\mu)) \leq \frac{n\mu}{\sigma}, \quad \lambda_{[i]}(X(\mu)) \geq \frac{\sigma}{n}. \quad (12)$$

(2) For $i = 1, \dots, n_{\mathcal{N}}$ we have

$$\lambda_{[i]}(S(\mu)) \geq \frac{\sigma}{n}, \quad \lambda_{[n-i+1]}(X(\mu)) \leq \frac{n\mu}{\sigma}. \quad (13)$$

Furthermore, we have

$$\begin{aligned} \lambda_{[n-i+1]}(X(\mu)) &\leq c\sqrt{n}(n\mu)^\gamma, & \lambda_{[i]}(S(\mu)) &\geq \frac{\mu}{c\sqrt{n}(n\mu)^\gamma}, & i = 1, \dots, n_{\mathcal{N}} + n_{\mathcal{T}}, \\ \lambda_{[n-i+1]}(S(\mu)) &\leq c\sqrt{n}(n\mu)^\gamma, & \lambda_{[i]}(X(\mu)) &\geq \frac{\mu}{c\sqrt{n}(n\mu)^\gamma}, & i = 1, \dots, n_{\mathcal{B}} + n_{\mathcal{T}}. \end{aligned} \quad (14)$$

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

$$c \geq \frac{(\min\{\theta_1^{-1}, \theta_2^{-1}\})^{\frac{1}{2}-\gamma}}{n}, \quad \frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}.$$

Proof. Recall that $\hat{S}_{\mathcal{B}}(\mu) = Q_{\mathcal{B}}^T S(\mu) Q_{\mathcal{B}}$ and $\hat{X}_{\mathcal{N}}(\mu) = Q_{\mathcal{N}}^T X(\mu) Q_{\mathcal{N}}$ as defined in [\(10\)](#). Then it follows from Lemma [3.5](#) that

$$\begin{aligned} \lambda_{[n-n_{\mathcal{B}}+1]}(S(\mu)) + \dots + \lambda_{[n]}(S(\mu)) &\leq \text{trace}(\hat{S}_{\mathcal{B}}(\mu)) \leq \frac{n\mu}{\sigma}, \\ \lambda_{[n-n_{\mathcal{N}}+1]}(X(\mu)) + \dots + \lambda_{[n]}(X(\mu)) &\leq \text{trace}(\hat{X}_{\mathcal{N}}(\mu)) \leq \frac{n\mu}{\sigma}. \end{aligned}$$

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

$$\begin{aligned} \lambda_{[n-i+1]}(S(\mu)) &\leq \frac{n\mu}{\sigma}, & i = 1, \dots, n_{\mathcal{B}}, \\ \lambda_{[n-i+1]}(X(\mu)) &\leq \frac{n\mu}{\sigma}, & i = 1, \dots, n_{\mathcal{N}}. \end{aligned}$$

Further, from the centrality condition $\Lambda(X(\mu))\Lambda(S(\mu)) = \mu I_n$, we can observe that the i^{th} largest eigenvalue of $X(\mu)$ and the i^{th} smallest eigenvalue of $S(\mu)$ have the

same eigenvector, which implies $\lambda_{[i]}(X(\mu))\lambda_{[n-i+1]}(S(\mu)) = \mu$. Hence, we can derive

$$\begin{aligned}\lambda_{[i]}(X(\mu)) &\geq \frac{\sigma}{n}, & i = 1, \dots, n_{\mathcal{B}}, \\ \lambda_{[i]}(S(\mu)) &\geq \frac{\sigma}{n}, & i = 1, \dots, n_{\mathcal{N}}.\end{aligned}$$

It follows from Lemmas [3.4](#) and [3.5](#) and $\text{trace}(X) \leq \sqrt{n}\|X\|$ that

$$\begin{aligned}\frac{1}{\sqrt{n}} \left(\lambda_{[n-n_{\mathcal{N}}-n_{\mathcal{T}}+1]}(X(\mu)) + \dots + \lambda_{[n]}(X(\mu)) \right) &\leq \|Q_{\mathcal{T} \cup \mathcal{N}}^T X(\mu) Q_{\mathcal{T} \cup \mathcal{N}}\| \leq c(n\mu)^\gamma, \\ \frac{1}{\sqrt{n}} \left(\lambda_{[n-n_{\mathcal{B}}-n_{\mathcal{T}}+1]}(S(\mu)) + \dots + \lambda_{[n]}(S(\mu)) \right) &\leq \|Q_{\mathcal{B} \cup \mathcal{T}}^T S(\mu) Q_{\mathcal{B} \cup \mathcal{T}}\| \leq c(n\mu)^\gamma,\end{aligned}$$

which by the centrality condition gives [\(14\)](#).

By [\(14\)](#), if $n_{\mathcal{T}} > 0$, there exist $n_{\mathcal{T}}$ eigenvalues of $X(\mu)$ and $n_{\mathcal{T}}$ eigenvalues of $S(\mu)$ which stay within the interval $\left[\frac{\mu}{c\sqrt{n}(n\mu)^\gamma}, c\sqrt{n}(n\mu)^\gamma \right]$, and thus both converge to 0 as $\mu \rightarrow 0$. Then it holds that

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

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

$$c \geq \frac{(\min\{\theta_1^{-1}, \theta_2^{-1}\})^{\frac{1}{2}-\gamma}}{n}, \quad \gamma \leq \frac{1}{2}.$$

The lower bound $\gamma \geq \frac{1}{2^{n-1}}$ follows from Lemma [B.2](#) in Appendix [B](#). This completes the proof. \square

Since the central path is an analytic curve, the eigenvalues of $X(\mu)$ and $S(\mu)$ 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 [3.6](#) that the eigenvalues of a central solution $(X(\mu), y(\mu), S(\mu))$ 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 $Q(\mu)$. Let $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$ denote the subsets of columns of $Q(\mu)$ corresponding to the above subsets of eigenvalues, respectively. Since the central path converges to a maximally complementary optimal solution, the accumulation points of $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$, when $\mu \rightarrow 0$, form orthonormal bases for the subspaces \mathcal{B} , \mathcal{T} , and \mathcal{N} , respectively. Section 3.3 in [de Klerk \(2006\)](#) elucidates the details. The following theorem specifies an upper bound on μ which allows for the identification of $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$.

Theorem 3.7. *If μ satisfies*

$$\mu < \tilde{\mu} := \min \left\{ \frac{1}{n} \left(\frac{\sigma}{cn^{\frac{3}{2}}} \right)^{\frac{1}{\gamma}}, \frac{\sigma^2}{n^2}, \hat{\mu} \right\}, \quad (15)$$

then we can identify $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$ from $Q(\mu)$.

Proof. From inequalities (12) and (13), we can deduce that the $n_{\mathcal{B}}$ largest eigenvalues of $X(\mu)$ stay positive while the $n_{\mathcal{B}}$ smallest eigenvalues of $S(\mu)$ will converge to 0. Similarly, the $n_{\mathcal{N}}$ largest eigenvalues of $S(\mu)$ will remain positive while the last $n_{\mathcal{N}}$ eigenvalues of $X(\mu)$ converge to 0 as $\mu \rightarrow 0$. Inequalities (14) also hint that, if $n_{\mathcal{T}} > 0$, then there should exist a set of $n_{\mathcal{T}}$ eigenvalues of $X(\mu)$ and $S(\mu)$ which stay within the interval $\left[\frac{\mu}{c\sqrt{n}(n\mu)^{\gamma}}, c\sqrt{n}(n\mu)^{\gamma} \right]$. Recall that the i^{th} largest eigenvalue of $X(\mu)$ and the i^{th} smallest eigenvalue of $S(\mu)$ have the same eigenvector. Thus, we can identify $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$ if

$$\frac{n\mu}{\sigma} < \frac{\mu}{c\sqrt{n}(n\mu)^{\gamma}}, \quad c\sqrt{n}(n\mu)^{\gamma} < \frac{\sigma}{n}, \quad \frac{n\mu}{\sigma} < \frac{\sigma}{n}, \quad (16)$$

which, by $\frac{\mu}{c\sqrt{n}(n\mu)^{\gamma}} \leq c\sqrt{n}(n\mu)^{\gamma}$, is equivalent to

$$\mu < \frac{1}{n} \left(\frac{\sigma}{cn^{\frac{3}{2}}} \right)^{\frac{1}{\gamma}}. \quad (17)$$

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 (9). This completes the proof. \square

Remark 6. In general, we do not know in advance if the strict complementarity condition holds for a given instance of SDO. Note that (16) and (17) imply that if $n_{\mathcal{T}} > 0$, then we have

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

If $n_{\mathcal{T}} = 0$, then we can make improvement on the bound (15). In fact, the bounds in (14) may provide no further information compared to (12) and (13) for small values of μ . Hence, in order to identify $Q_{\mathcal{B}}(\mu)$ and $Q_{\mathcal{N}}(\mu)$ it is enough to have

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

which reduces the bound (15) to $\mu < \frac{\sigma^2}{n^2}$. This bound matches the one for LO, see Section 3.3.3 in Roos et al. (2005).

3.3. Proximity to the optimal partition

We can provide more information about the optimal partition of the problem by measuring the proximity of $\mathcal{R}(Q_{\mathcal{B}}(\mu))$ and $\mathcal{R}(Q_{\mathcal{N}}(\mu))$ to the subspaces \mathcal{B} and \mathcal{N} , respectively, for $\mu < \tilde{\mu}$. To that end, we use the approach in [Cheung et al. \(2013\)](#) which measures the distance between a primal optimal solution $X \in \mathcal{P}^*$ and its projection onto $Q_{\mathcal{B} \cup \mathcal{T}}(\mu) \mathbb{S}_+^{n_{\mathcal{B}}+n_{\mathcal{T}}} Q_{\mathcal{B} \cup \mathcal{T}}^T(\mu)$, which is a face of the positive semidefinite cone, see Proposition 2.2.14 in [Cheung \(2013\)](#) for its proof. In fact, \mathcal{P}^* is contained in the minimal face $Q_{\mathcal{B}} \mathbb{S}_+^{n_{\mathcal{B}}} Q_{\mathcal{B}}^T$ which itself is a face of $Q_{\mathcal{B} \cup \mathcal{T}} \mathbb{S}_+^{n_{\mathcal{B}}+n_{\mathcal{T}}} Q_{\mathcal{B} \cup \mathcal{T}}^T$. Analogously, we measure the distance between \tilde{S} , where $(\tilde{y}, \tilde{S}) \in \mathcal{D}^*$, and its projection onto $Q_{\mathcal{T} \cup \mathcal{N}}(\mu) \mathbb{S}_+^{n_{\mathcal{T}}+n_{\mathcal{N}}} Q_{\mathcal{T} \cup \mathcal{N}}^T(\mu)$.

The following technical lemma is in order.

Lemma 3.8. *Let $(X(\mu), y(\mu), S(\mu))$ be given so that $\mu \leq \hat{\mu}$. Then we have*

$$\sup_{\tilde{X} \in \mathcal{P}^* \setminus \{0\}} \frac{S(\mu) \bullet \tilde{X}}{\|\tilde{X}\|} \leq c(n\mu)^\gamma,$$

$$\sup_{(\tilde{y}, \tilde{S}) \in \mathcal{D}^*, \tilde{S} \neq 0} \frac{X(\mu) \bullet \tilde{S}}{\|\tilde{S}\|} \leq c(n\mu)^\gamma.$$

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

$$\frac{S(\mu) \bullet \tilde{X}}{\|\tilde{X}\|} = \frac{(S(\mu) - \tilde{S} + \tilde{S}) \bullet \tilde{X}}{\|\tilde{X}\|} = \frac{(S(\mu) - \tilde{S}) \bullet \tilde{X}}{\|\tilde{X}\|} \leq \|S(\mu) - \tilde{S}\|.$$

Therefore, we get

$$\sup_{\tilde{X} \in \mathcal{P}^* \setminus \{0\}} \frac{S(\mu) \bullet \tilde{X}}{\|\tilde{X}\|} \leq \min_{(\tilde{y}, \tilde{S}) \in \mathcal{D}^*} \|S(\mu) - \tilde{S}\| \leq c(n\mu)^\gamma,$$

where the last inequality follows from Lemma [3.3](#). The proof for the second part follows analogously. \square

Theorem 3.9. *Let $(X(\mu), y(\mu), S(\mu))$ be given so that $\mu < \tilde{\mu}$. Then for all $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ we have*

$$\|\tilde{X} - \tilde{X}_{\mathcal{F}_{\mathcal{B}\mathcal{T}}}\| \leq \sqrt{2} \|\tilde{X}\| \sqrt{\frac{cn(n\mu)^\gamma}{\sigma}}, \quad (18)$$

$$\|\tilde{S} - \tilde{S}_{\mathcal{F}_{\mathcal{T}\mathcal{N}}}\| \leq \sqrt{2} \|\tilde{S}\| \sqrt{\frac{cn(n\mu)^\gamma}{\sigma}}, \quad (19)$$

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

$$\mathcal{F}_{\mathcal{B}\mathcal{T}} := Q_{\mathcal{B} \cup \mathcal{T}}(\mu) \mathbb{S}_+^{n_{\mathcal{B}}+n_{\mathcal{T}}} Q_{\mathcal{B} \cup \mathcal{T}}^T(\mu),$$

$$\mathcal{F}_{\mathcal{T}\mathcal{N}} := Q_{\mathcal{T} \cup \mathcal{N}}(\mu) \mathbb{S}_+^{n_{\mathcal{T}}+n_{\mathcal{N}}} Q_{\mathcal{T} \cup \mathcal{N}}^T(\mu).$$

Proof. If $\tilde{X} = 0$ or $\tilde{S} = 0$, then $\tilde{X}_{\mathcal{F}_{\mathcal{B}\mathcal{T}}} = 0$ or $\tilde{S}_{\mathcal{F}_{\mathcal{T}\mathcal{N}}} = 0$, and thus the inequalities (18) and (19) trivially hold. Note that the projection of \tilde{X} onto the face $\mathcal{F}_{\mathcal{B}\mathcal{T}}$ is the optimal solution to

$$\begin{aligned}\tilde{X}_{\mathcal{F}_{\mathcal{B}\mathcal{T}}} &:= \operatorname{argmin}_{U \succeq 0} \left\| \tilde{X} - Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) U Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \right\| \\ &= \operatorname{argmin}_{U \succeq 0} \left\| \begin{bmatrix} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \tilde{X} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) - U & Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \tilde{X} Q_{\mathcal{N}}(\mu) \\ Q_{\mathcal{N}}^T(\mu) \tilde{X} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) & Q_{\mathcal{N}}^T(\mu) \tilde{X} Q_{\mathcal{N}}(\mu) \end{bmatrix} \right\|,\end{aligned}$$

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

$$\begin{aligned}\|\tilde{X} - \tilde{X}_{\mathcal{F}_{\mathcal{B}\mathcal{T}}}\| &= \|\tilde{X} - Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) U^* Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu)\| \\ &= \|\tilde{X} - Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \tilde{X} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu)\| \\ &= \sqrt{\|\tilde{X}\|^2 - \|Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \tilde{X} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu)\|^2} \\ &\leq \|\tilde{X}\| \sqrt{1 - \frac{\|Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \tilde{X} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu)\|^2}{\|\tilde{X}\|^2}}.\end{aligned}$$

Thus, it only remains to derive a lower bound on

$$\frac{\|Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \tilde{X} Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu)\|}{\|\tilde{X}\|}. \quad (20)$$

Let us define

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

Then from Lemma 3.8 we get

$$\begin{aligned}Q_{\mathcal{N}}(\mu) \Lambda_{\mathcal{N}}(S(\mu)) Q_{\mathcal{N}}^T(\mu) \bullet \tilde{X} &\leq Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu) \Lambda_{\mathcal{B}\mathcal{U}\mathcal{T}}(S(\mu)) Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) \bullet \tilde{X} \\ &\quad + Q_{\mathcal{N}}(\mu) \Lambda_{\mathcal{N}}(S(\mu)) Q_{\mathcal{N}}^T(\mu) \bullet \tilde{X} \\ &= S(\mu) \bullet \tilde{X} \\ &\leq c(n\mu)^\gamma \|\tilde{X}\|.\end{aligned}$$

All this implies that

$$\begin{aligned}\min \quad & \|Q_{\mathcal{B}\mathcal{U}\mathcal{T}}^T(\mu) X Q_{\mathcal{B}\mathcal{U}\mathcal{T}}(\mu)\| \\ \text{s.t.} \quad & Q_{\mathcal{N}}(\mu) \Lambda_{\mathcal{N}}(S(\mu)) Q_{\mathcal{N}}^T(\mu) \bullet X \leq c(n\mu)^\gamma, \\ & \|X\| = 1, \\ & X \succeq 0,\end{aligned} \quad (21)$$

gives a lower bound on (20). Let $\check{X} := Q^T(\mu)XQ(\mu)$, where

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

Then auxiliary problem (21) is equivalent to

$$\begin{aligned} \min \quad & \|\check{X}_{\mathcal{B} \cup \mathcal{T}}\| \\ \text{s.t.} \quad & \Lambda_{\mathcal{N}}(S(\mu)) \bullet \check{X}_{\mathcal{N}} \leq c(n\mu)^\gamma, \\ & \|\check{X}_{\mathcal{B} \cup \mathcal{T}}\|^2 + \|\check{X}_{\mathcal{N}}\|^2 + 2\|\check{X}_{(\mathcal{B} \cup \mathcal{T})\mathcal{N}}\|^2 = 1, \\ & \check{X} \succeq 0. \end{aligned} \tag{22}$$

Since $\check{X} \succeq 0$, we can use the inequality² $\|\check{X}_{(\mathcal{B} \cup \mathcal{T})\mathcal{N}}\|^2 \leq \|\check{X}_{\mathcal{B} \cup \mathcal{T}}\| \|\check{X}_{\mathcal{N}}\|$ to derive a relaxation of (22) as

$$\begin{aligned} \min \quad & \|\check{X}_{\mathcal{B} \cup \mathcal{T}}\| \\ \text{s.t.} \quad & \Lambda_{\mathcal{N}}(S(\mu)) \bullet \check{X}_{\mathcal{N}} \leq c(n\mu)^\gamma, \\ & \|\check{X}_{\mathcal{B} \cup \mathcal{T}}\| + \|\check{X}_{\mathcal{N}}\| \geq 1, \\ & \check{X}_{\mathcal{B} \cup \mathcal{T}} \succeq 0, \\ & \check{X}_{\mathcal{N}} \succeq 0. \end{aligned} \tag{23}$$

Finally, from the constraints in (23) we get

$$\|\check{X}_{\mathcal{B} \cup \mathcal{T}}\| \geq 1 - \|\check{X}_{\mathcal{N}}\| \geq 1 - \frac{c(n\mu)^\gamma}{\lambda_{[n_{\mathcal{N}}]}(S(\mu))} \geq 1 - \frac{cn(n\mu)^\gamma}{\sigma} \tag{24}$$

$$> 1 - \frac{1}{\sqrt{n}} > 0, \tag{25}$$

in which (24) follows from (13) as well as

$$\lambda_{\min}(\Lambda_{\mathcal{N}}(S(\mu)))\|\check{X}_{\mathcal{N}}\| \leq \Lambda_{\mathcal{N}}(S(\mu)) \bullet \check{X}_{\mathcal{N}} \leq c(n\mu)^\gamma,$$

and (25) results from $\mu < \tilde{\mu}$. In a similar way as in Cheung et al. (2013), it can be shown that $1 - \frac{c(n\mu)^\gamma}{\lambda_{[n_{\mathcal{N}}]}(S(\mu))}$ is indeed the optimal value of (21). Consequently, we can

²The validity of this inequality can be verified by squaring both sides of $\left\| \begin{bmatrix} A & X \\ X^T & B \end{bmatrix} \right\| \leq \|A\| + \|B\|$, which is valid for all positive semidefinite $\begin{bmatrix} A & X \\ X^T & B \end{bmatrix}$. See Theorem 2.1 and Remark 2.3 in Lee (2011) for more general results.

conclude that

$$\begin{aligned}
\|\tilde{X} - \tilde{X}_{\mathcal{B}\mathcal{T}}\| &\leq \|\tilde{X}\| \sqrt{1 - \frac{\|Q_{\mathcal{B}\mathcal{T}}^T(\mu)\tilde{X}Q_{\mathcal{B}\mathcal{T}}(\mu)\|^2}{\|\tilde{X}\|^2}} \\
&\leq \|\tilde{X}\| \sqrt{2\left(\frac{cn(n\mu)^\gamma}{\sigma}\right) - \left(\frac{cn(n\mu)^\gamma}{\sigma}\right)^2} \\
&\leq \sqrt{2}\|\tilde{X}\| \sqrt{\frac{cn(n\mu)^\gamma}{\sigma}}.
\end{aligned}$$

Analogously, we can prove that

$$\|\tilde{S} - \tilde{S}_{\mathcal{F}_{\mathcal{T}\mathcal{N}}}\| \leq \|\tilde{S}\| \sqrt{1 - \frac{\|Q_{\mathcal{T}\mathcal{N}}^T(\mu)\tilde{S}Q_{\mathcal{T}\mathcal{N}}(\mu)\|^2}{\|\tilde{S}\|^2}} \leq \sqrt{2}\|\tilde{S}\| \sqrt{\frac{cn(n\mu)^\gamma}{\sigma}}.$$

□

Under the assumption of primal-dual uniqueness, we can provide an upper bound on the distance between the subspaces \mathcal{B} and $\mathcal{R}(Q_{\mathcal{B}}(\mu))$, which are of the same dimension if $\mu < \tilde{\mu}$. The distance between two subspaces Ψ_1 and Ψ_2 of \mathbb{R}^n with the same dimension is defined as

$$\text{dist}(\Psi_1, \Psi_2) := \|P_{\Psi_1} - P_{\Psi_2}\|_2,$$

where P_{Ψ_1} and P_{Ψ_2} denote the orthogonal projections onto Ψ_1 and Ψ_2 , respectively, see Section 2.5.3 in [Golub and Van Loan \(2013\)](#).

Theorem 3.10. *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 [2](#). Then there exist $\rho > 0$ and $v > 0$ such that*

$$\begin{aligned}
\text{dist}(\mathcal{B}, \mathcal{R}(Q_{\mathcal{B}}(\mu))) &\leq \min\{2\rho(\sqrt{n}\mu)^v, 1\}, \\
\text{dist}(\mathcal{T}, \mathcal{R}(Q_{\mathcal{T}}(\mu))) &\leq \min\{2\rho(\sqrt{n}\mu)^v, 1\}, \\
\text{dist}(\mathcal{N}, \mathcal{R}(Q_{\mathcal{N}}(\mu))) &\leq \min\{2\rho(\sqrt{n}\mu)^v, 1\}.
\end{aligned}$$

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 [Golub and Van Loan \(2013\)](#). Then we get

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

which implies

$$\text{dist}(\mathcal{B}, \mathcal{R}(Q_{\mathcal{B}}(\mu))) \leq 2 \min_{\bar{Q}_{\mathcal{B}} \in \Gamma_{\mathcal{B}}} \|Q_{\mathcal{B}}(\mu) - \bar{Q}_{\mathcal{B}}\|_2. \quad (26)$$

The set of primal-dual optimal solutions can be represented as a system of polynomial equations and inequalities (Alizadeh et al. 1998). Let \mathcal{C} be the set of all solutions $(\text{vec}(\tilde{Q}); \text{diag}(\Lambda(\tilde{X})); \tilde{y}; \text{diag}(\Lambda(\tilde{S})))$ of this system so that $\tilde{Q}\Lambda(\tilde{X})\tilde{Q}^T \in \mathcal{P}^*$ and $(\tilde{y}, \tilde{Q}\Lambda(\tilde{S})\tilde{Q}^T) \in \mathcal{D}^*$, where $\text{vec}(\cdot)$ is the concatenation of the columns of a matrix, and $\text{diag}(\cdot)$ denotes the vector of diagonal entries of a square matrix. By the centrality condition, a central solution $(X(\mu), y(\mu), S(\mu))$ violates the constraints $\Lambda_{ii}(\tilde{X})\Lambda_{ii}(\tilde{S}) = 0$ by μ for $i = 1, \dots, n$. Since the set of central solutions $(X(\mu), y(\mu), S(\mu))$ with $\mu < \tilde{\mu}$ is contained in a compact set, it follows from (26) and Lemma C.1 that there exist $v > 0$ and $\rho > 0$ such that

$$\begin{aligned} \text{dist}(\mathcal{B}, \mathcal{R}(Q_{\mathcal{B}}(\mu))) &\leq 2 \inf_{(Q^*\Lambda(X^*)(Q^*)^T, y^*, Q^*\Lambda(S^*)(Q^*)^T) \in \text{ri}(\mathcal{P}^* \times \mathcal{D}^*)} \left\| \begin{pmatrix} \text{vec}(Q(\mu) - Q^*); \\ \text{diag}(\Lambda(X(\mu)) - \Lambda(X^*)); y(\mu) - y^*; \text{diag}(\Lambda(S(\mu)) - \Lambda(S^*)) \end{pmatrix} \right\|_2 \\ &= 2 \text{dist} \left(\begin{pmatrix} \text{vec}(Q(\mu)); \text{diag}(\Lambda(X(\mu))); y(\mu); \text{diag}(\Lambda(S(\mu))) \end{pmatrix}, \mathcal{C} \right) \\ &\leq 2\rho(\sqrt{n}\mu)^v, \end{aligned} \quad (27)$$

where the equality (27) follows from the uniqueness of the optimal solution. The proofs for the subspaces \mathcal{T} and \mathcal{N} are analogous. \square

Remark 7. Since \mathcal{C} is a compact set, there exists a solution $(\text{vec}(\tilde{Q}_{\mu}); \text{diag}(\Lambda(\tilde{X}_{\mu})); \tilde{y}_{\mu}; \text{diag}(\Lambda(\tilde{S}_{\mu})))$ of \mathcal{C} whose distance from $(\text{vec}(Q(\mu)); \text{diag}(\Lambda(X(\mu))); y(\mu); \text{diag}(\Lambda(S(\mu))))$ is minimal. The assumption of uniqueness in Theorem 3.10 can be released if there exists a sequence of common eigenvector bases of maximally complementary solutions converging to \tilde{Q}_{μ} . More precisely, assume that $Q_k \rightarrow \tilde{Q}_{\mu}$ for a convergent sequence $(X_k, y_k, S_k) \rightarrow (\tilde{X}_{\mu}, \tilde{y}_{\mu}, \tilde{S}_{\mu})$ such that $(X_k, y_k, S_k) \in \text{ri}(\mathcal{P}^* \times \mathcal{D}^*)$ 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 (27) holds, and thus the upper bounds in Theorem 3.10 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 (Goldfarb and Scheinberg 1999):

$$\begin{aligned} A^1 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}, \\ b^T &= [1, 0, 0]. \end{aligned}$$

The primal optimal set can be described as

$$X_\delta^* = \begin{bmatrix} 1 & 2(\delta-1) & 2(\delta-1) \\ 2(\delta-1) & 4(1-\delta) & 4(1-\delta) \\ 2(\delta-1) & 4(1-\delta) & 4(1-\delta) \end{bmatrix}, \quad 0 \leq \delta \leq 1,$$

and the unique dual optimal solution is

$$S^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, \quad y^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

One can verify that the eigenvalues of X_δ^* for $0 \leq \delta \leq 1$ 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_δ^*, y^*, S^*) is strictly complementary, and that the multiplicity of the positive eigenvalues of X_δ^* and S^* are 1. Hence, for all $0 < \delta < 1$, the eigenvalue decompositions of X_δ^* 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}_μ and \tilde{S}_μ have the unique common eigenvector basis

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

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

4. On the identification of the optimal partition 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). 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 $(X^\circ, y^\circ, S^\circ) \in \text{ri}(\mathcal{P} \times \mathcal{D})$ given by a primal-dual path-following IPM, where $X^\circ = M\Lambda(X^\circ)M^T$ and $S^\circ = P\Lambda(S^\circ)P^T$ are eigenvalue decompositions of X° and S° , respectively, and M and P are orthogonal matrices. In contrast to the result of Theorem 3.7, 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 [Nesterov and Todd \(1997\)](#) for the definition of D . Then the eigenvalue decomposition of S° yields

$$X^\circ = DP\Lambda(S^\circ)P^T D. \quad (28)$$

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 [\(28\)](#) 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

$$M = D^{-1} P \Sigma^{-\frac{1}{2}} N,$$

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 can be measured (see e.g., Section 6.4 in [de Klerk \(2006\)](#)) by

$$\kappa(X^\circ S^\circ) := \frac{\lambda_{\max}(X^\circ S^\circ)}{\lambda_{\min}(X^\circ S^\circ)}, \quad (X^\circ, y^\circ, S^\circ) \in \text{ri}(\mathcal{P} \times \mathcal{D}). \quad (29)$$

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 [\(29\)](#) that $\kappa(X^\circ S^\circ) \geq 1$, 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

$$\mathcal{N}_\kappa(\xi) := \left\{ (X^\circ, y^\circ, S^\circ) \in \text{ri}(\mathcal{P} \times \mathcal{D}) \mid \kappa(X^\circ S^\circ) \leq \xi \right\}, \quad (30)$$

where $\xi > 1$. Then for $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$ we have

$$\lambda_{\min}(X^\circ S^\circ) \leq \lambda_{[i]}(X^\circ S^\circ) \leq \xi \lambda_{\min}(X^\circ S^\circ), \quad i = 1, \dots, n. \quad (31)$$

Here, we use the application of Weyl theorem³ in [Lu and Pearce \(2000\)](#) to provide an upper bound on $\lambda_{\min}(X^\circ S^\circ)$.

Lemma 4.1 (Corollary 2.3 in [Lu and Pearce \(2000\)](#)). *Let X and S be two $n \times n$ symmetric positive semidefinite matrices. Then for $j \leq \min\{\text{rank}(X), \text{rank}(S)\}$ we have*

$$\min_{1 \leq i \leq j} \{\lambda_{[i]}(X) \lambda_{[j-i+1]}(S)\} \geq \lambda_{[j]}(XS) \geq \max_{j \leq i \leq n} \{\lambda_{[i]}(X) \lambda_{[n+j-i]}(S)\}. \quad (32)$$

³See Theorem 4.3.7 in [Horn and Johnson \(2012\)](#).

Lemma 4.2. Let $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$. Then we have

$$\lambda_{[i]}(X^\circ)\lambda_{[n-i+1]}(S^\circ) \geq \lambda_{\min}(X^\circ S^\circ), \quad i = 1, \dots, n. \quad (33)$$

Proof. The proof is straightforward from the first inequality in (32) and the positive definiteness of X° and S° . In fact, for the special case $k = n$ there holds that

$$\min \left\{ \lambda_{[1]}(X^\circ)\lambda_{[n]}(S^\circ), \lambda_{[2]}(X^\circ)\lambda_{[n-1]}(S^\circ), \dots, \lambda_{[n]}(X^\circ)\lambda_{[1]}(S^\circ) \right\} \geq \lambda_{\min}(X^\circ S^\circ),$$

which completes the proof. \square

The following theorem generalizes the bounds derived in Theorem 3.6 to an approximate solution $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$. Analogous to the case of central solutions, we let $M := (M_{\mathcal{B}}, M_{\mathcal{T}}, M_{\mathcal{N}})$ and $P := (P_{\mathcal{B}}, P_{\mathcal{T}}, P_{\mathcal{N}})$ be the subsets of columns of M and P , respectively, associated with the eigenvalues of X° and S° whose accumulation points are positive and zero.

Theorem 4.3. Let $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$ and $\mu := \frac{X^\circ \bullet S^\circ}{n}$. Then there exist a positive condition number c' independent of μ and a positive exponent γ so that

(1) For $i = 1, \dots, n_{\mathcal{B}}$ we have

$$\lambda_{[n-i+1]}(S^\circ) \leq \frac{n\mu}{\sigma}, \quad \lambda_{[i]}(X^\circ) \geq \frac{\sigma}{n\xi}.$$

(2) For $i = 1, \dots, n_{\mathcal{N}}$ we have

$$\lambda_{[i]}(S^\circ) \geq \frac{\sigma}{n\xi}, \quad \lambda_{[n-i+1]}(X^\circ) \leq \frac{n\mu}{\sigma}.$$

Furthermore, we have

$$\begin{aligned} \lambda_{[n-i+1]}(X^\circ) &\leq c' \sqrt{n}(n\mu)^\gamma, & \lambda_{[i]}(S^\circ) &\geq \frac{\mu}{c' \sqrt{n}\xi(n\mu)^\gamma}, & i = 1, \dots, n_{\mathcal{N}} + n_{\mathcal{T}}, \\ \lambda_{[n-i+1]}(S^\circ) &\leq c' \sqrt{n}(n\mu)^\gamma, & \lambda_{[i]}(X^\circ) &\geq \frac{\mu}{c' \sqrt{n}\xi(n\mu)^\gamma}, & i = 1, \dots, n_{\mathcal{B}} + n_{\mathcal{T}}. \end{aligned}$$

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

$$\frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}.$$

If μ satisfies

$$\mu < \min \left\{ \frac{1}{n} \left(\frac{\sigma}{c' n^{\frac{3}{2}} \xi} \right)^{\frac{1}{\gamma}}, \frac{\sigma^2}{n^2 \xi}, \hat{\mu} \right\}, \quad (34)$$

then we can identify $M_{\mathcal{B}}$, $M_{\mathcal{T}}$, and $M_{\mathcal{N}}$ from X° , and $P_{\mathcal{B}}$, $P_{\mathcal{T}}$, and $P_{\mathcal{N}}$ from S° .

Proof. The proof technique can be traced back to Theorem 3.6 fairly easily. Let $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ which satisfies (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

$$X^\circ \bullet \bar{S} + \bar{X} \bullet S^\circ = \hat{X}_{\mathcal{N}}^\circ \bullet U_{\bar{S}} + U_{\bar{X}} \bullet \hat{S}_{\mathcal{B}}^\circ = X^\circ \bullet S^\circ,$$

where $\hat{S}_{\mathcal{B}}^\circ = Q_{\mathcal{B}}^T S^\circ Q_{\mathcal{B}}$ and $\hat{X}_{\mathcal{N}}^\circ = Q_{\mathcal{N}}^T X^\circ Q_{\mathcal{N}}$. Using the inequality $\lambda_{\min}(U_{\bar{S}}) \text{trace}(\hat{X}_{\mathcal{N}}^\circ) \leq \hat{X}_{\mathcal{N}}^\circ \bullet U_{\bar{S}}$ and the positive definiteness of X° and S° we have

$$\begin{aligned} \lambda_{\min}(U_{\bar{X}}) \text{trace}(\hat{S}_{\mathcal{B}}^\circ) &\leq X^\circ \bullet S^\circ \Rightarrow \text{trace}(\hat{S}_{\mathcal{B}}^\circ) \leq \frac{n\mu}{\sigma}, \\ \lambda_{\min}(U_{\bar{S}}) \text{trace}(\hat{X}_{\mathcal{N}}^\circ) &\leq X^\circ \bullet S^\circ \Rightarrow \text{trace}(\hat{X}_{\mathcal{N}}^\circ) \leq \frac{n\mu}{\sigma}, \end{aligned}$$

where the latter inequalities follow from (11). Now, Lemma 3.5 can be applied to get

$$\begin{aligned} \lambda_{[n-n_{\mathcal{B}}+1]}(S^\circ) + \dots + \lambda_{[n]}(S^\circ) &\leq \text{trace}(\hat{S}_{\mathcal{B}}^\circ) \leq \frac{n\mu}{\sigma}, \\ \lambda_{[n-n_{\mathcal{N}}+1]}(X^\circ) + \dots + \lambda_{[n]}(X^\circ) &\leq \text{trace}(\hat{X}_{\mathcal{N}}^\circ) \leq \frac{n\mu}{\sigma}, \end{aligned}$$

which by $X^\circ, S^\circ \succ 0$ implies

$$\begin{aligned} \lambda_{[n-i+1]}(S^\circ) &\leq \frac{n\mu}{\sigma}, & i = 1, \dots, n_{\mathcal{B}}, \\ \lambda_{[n-i+1]}(X^\circ) &\leq \frac{n\mu}{\sigma}, & i = 1, \dots, n_{\mathcal{N}}. \end{aligned} \tag{35}$$

Recall from (31) that

$$n\mu = X^\circ \bullet S^\circ \leq n\xi \lambda_{\min}(X^\circ S^\circ),$$

which yields

$$\frac{\lambda_{\min}(X^\circ S^\circ)}{\mu} \geq \frac{1}{\xi}. \tag{36}$$

Then (33) and (36) can be applied to (35) to derive lower bounds on the eigenvalues of X° and S° :

$$\begin{aligned} \lambda_{[i]}(X^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(S^\circ)} \geq \frac{\sigma \lambda_{\min}(X^\circ S^\circ)}{n\mu} \geq \frac{\sigma}{n\xi}, & i = 1, \dots, n_{\mathcal{B}}, \\ \lambda_{[i]}(S^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(X^\circ)} \geq \frac{\sigma \lambda_{\min}(X^\circ S^\circ)}{n\mu} \geq \frac{\sigma}{n\xi}, & i = 1, \dots, n_{\mathcal{N}}. \end{aligned}$$

For the subspace \mathcal{T} we should note that

$$\{(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi) \mid X^\circ \bullet S^\circ \leq \min\{\theta_1^{-1}, \theta_2^{-1}\}\},$$

where θ_1 and θ_2 are defined as in (6) and (7), is a bounded set by the interior point condition and the linear independence of A^i for $i = 1, \dots, m$, see e.g., Lemma 3.1 in de Klerk (2006). Furthermore, the amount of constraint violation with respect to

the LMI system (5) for $(X^\circ, y^\circ, S^\circ)$ is equal to $n\mu$. Hence, the result of Lemma B.1 is still valid, i.e., for $0 < \mu \leq \hat{\mu}$ there exist $(X_{\mu^\circ}, y_{\mu^\circ}, S_{\mu^\circ}) \in \mathcal{P}^* \times \mathcal{D}^*$, a positive condition number c' independent of μ , and a positive exponent γ so that

$$\|X^\circ - X_{\mu^\circ}\| \leq c'(n\mu)^\gamma, \quad \|S^\circ - S_{\mu^\circ}\| \leq c'(n\mu)^\gamma, \quad (37)$$

where c' and γ are defined as in Lemma 3.3. Analogous to the proof of Theorem 3.6, we can observe, using the orthogonal transformation Q , that

$$\begin{aligned} \|Q_{\mathcal{T} \cup \mathcal{N}}^T X^\circ Q_{\mathcal{T} \cup \mathcal{N}}\| &= \left\| \begin{bmatrix} \hat{X}_{\mathcal{T}}^\circ & \hat{X}_{\mathcal{T}\mathcal{N}}^\circ \\ \hat{X}_{\mathcal{N}\mathcal{T}}^\circ & \hat{X}_{\mathcal{N}}^\circ \end{bmatrix} \right\| \leq \|X^\circ - X_{\mu^\circ}\| \leq c'(n\mu)^\gamma, \\ \|Q_{\mathcal{B} \cup \mathcal{T}}^T S^\circ Q_{\mathcal{B} \cup \mathcal{T}}\| &= \left\| \begin{bmatrix} \hat{S}_{\mathcal{B}}^\circ & \hat{S}_{\mathcal{B}\mathcal{T}}^\circ \\ \hat{S}_{\mathcal{T}\mathcal{B}}^\circ & \hat{S}_{\mathcal{T}}^\circ \end{bmatrix} \right\| \leq \|S^\circ - S_{\mu^\circ}\| \leq c'(n\mu)^\gamma. \end{aligned} \quad (38)$$

Then it follows from Lemma 3.5 and (38) that

$$\begin{aligned} \lambda_{[n-n_{\mathcal{N}}-n_{\mathcal{T}}+1]}(X^\circ) + \dots + \lambda_{[n]}(X^\circ) &\leq c'\sqrt{n}(n\mu)^\gamma, \\ \lambda_{[n-n_{\mathcal{B}}-n_{\mathcal{T}}+1]}(S^\circ) + \dots + \lambda_{[n]}(S^\circ) &\leq c'\sqrt{n}(n\mu)^\gamma, \end{aligned}$$

and consequently,

$$\begin{aligned} \lambda_{[n-i+1]}(X^\circ) &\leq c'\sqrt{n}(n\mu)^\gamma, & i = 1, \dots, n_{\mathcal{N}} + n_{\mathcal{T}}, \\ \lambda_{[n-i+1]}(S^\circ) &\leq c'\sqrt{n}(n\mu)^\gamma, & i = 1, \dots, n_{\mathcal{B}} + n_{\mathcal{T}}. \end{aligned}$$

Using the bounds in (33) and (36) we can derive

$$\begin{aligned} \lambda_{[i]}(X^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(S^\circ)} \geq \frac{\lambda_{\min}(X^\circ S^\circ)}{c'\sqrt{n}(n\mu)^\gamma} \geq \frac{\mu}{c'\sqrt{n}\xi(n\mu)^\gamma}, & i = 1, \dots, n_{\mathcal{B}} + n_{\mathcal{T}}, \\ \lambda_{[i]}(S^\circ) &\geq \frac{\lambda_{\min}(X^\circ S^\circ)}{\lambda_{[n-i+1]}(X^\circ)} \geq \frac{\lambda_{\min}(X^\circ S^\circ)}{c'\sqrt{n}(n\mu)^\gamma} \geq \frac{\mu}{c'\sqrt{n}\xi(n\mu)^\gamma}, & i = 1, \dots, n_{\mathcal{N}} + n_{\mathcal{T}}. \end{aligned}$$

In the sequel, using the same argument as in Theorem 3.7, we can identify the subsets of columns of M and P whose accumulation points form orthonormal bases for \mathcal{B} , \mathcal{T} and \mathcal{N} if

$$\frac{n\mu}{\sigma} < \frac{\mu}{c'\sqrt{n}\xi(n\mu)^\gamma}, \quad c'\sqrt{n}(n\mu)^\gamma < \frac{\sigma}{n\xi}. \quad (39)$$

Considering the case $\mathcal{T} = \{0\}$, we can represent (39) as

$$\mu < \min \left\{ \frac{1}{n} \left(\frac{\sigma}{c'n^{\frac{3}{2}}\xi} \right)^{\frac{1}{\gamma}}, \frac{\sigma^2}{n^2\xi} \right\}.$$

Including the condition $\mu \leq \hat{\mu}$ gives the result as desired. Further, if $n_{\mathcal{T}} > 0$, from

$$\frac{\mu}{c'\sqrt{n}\xi(n\mu)^\gamma} \leq c'\sqrt{n}(n\mu)^\gamma \text{ we get}$$

$$(c')^2 n^2 \xi \geq (n\mu)^{1-2\gamma}, \quad \forall 0 < \mu \leq \hat{\mu},$$

which implies

$$\frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}.$$

This completes the proof. \square

Corollary 4.4. *Let $(X^{(0)}, y^{(0)}, S^{(0)}) \in \mathcal{N}_\kappa(\xi)$ be an initial solution, $\mu^{(0)} := \frac{X^{(0)} \bullet S^{(0)}}{n}$, and $\log(\cdot)$ denote the natural logarithm. Then the Dikin-type primal-dual affine scaling method with steplength $\alpha = \frac{1}{\xi\sqrt{n}}$ and the neighborhood (30) (see Section 6.6 in de Klerk (2006)) needs at most*

$$\left\lceil \xi n \log \left(\mu^{(0)} \left(\min \left\{ \frac{1}{n} \left(\frac{\sigma}{c' n^{\frac{3}{2}} \xi} \right)^{\frac{1}{\gamma}}, \frac{\sigma^2}{n^2 \xi}, \hat{\mu} \right\} \right)^{-1} \right) \right\rceil$$

iterations to get an $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$ which allows to identify $(M_{\mathcal{B}}, M_{\mathcal{T}}, M_{\mathcal{N}})$ and $(P_{\mathcal{B}}, P_{\mathcal{T}}, P_{\mathcal{N}})$.

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 de Klerk (2006). 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 (34) multiplied by n . \square

Remark 8. In (37), we employed the same exponent γ as in (9) but a different condition number c' . In fact, the primal and dual systems in (5) are used for both Theorems 3.6 and 4.3. However, it is not known whether c and c' are identical or of the same order.

5. Concluding remarks

In this paper, we considered the identification of the optimal partition for SDO where strict complementarity may fail. Using the condition number σ defined in (2) and the upper bounds in (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

$$\left\lceil \xi n \log \left(\mu^{(0)} \left(\min \left\{ \frac{1}{n} \left(\frac{\sigma}{c' n^{\frac{3}{2}} \xi} \right)^{\frac{1}{\gamma}}, \frac{\sigma^2}{n^2 \xi}, \hat{\mu} \right\} \right)^{-1} \right) \right\rceil$$

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 provided a positive lower bound for the condition number σ . Even though the lower bound 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. However, 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.

Our approach 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 . Moreover, it is worth investigating the dependence of the condition numbers c and c' on the problem data. The derivation of upper bounds on θ_1 , θ_2 , c , and c' can be another subject of future studies.

Funding

This work is supported by the Air force Office of Scientific Research (AFOSR) Grant # FA9550-15-1-0222.

References

- Alizadeh F. 1991. Combinatorial optimization with interior point methods and semidefinite matrices [dissertation]. University of Minnesota.
- Alizadeh F, Haeberly JPA, Overton ML. 1998. Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM Journal on Optimization*. 8(3):746–768.
- Bonnans JF, Ramírez H. 2005. Perturbation analysis of second-order cone programming problems. *Mathematical Programming*. 104(2):205–227.
- Cheung YL. 2013. Preprocessing and reduction for semidefinite programming via facial reduction: Theory and practice [dissertation]. Department of Combinatorics & Optimization, University of Waterloo.
- Cheung YL, Schurr S, Wolkowicz H. 2013. Preprocessing and regularization for degenerate semidefinite programs. In: Bailey D, Bauschke H, Borwein P, Garvan F, Théra M, Vanderwerff J, Wolkowicz H, editors. *Computational and analytical mathematics, in honor of Jonathan Borwein’s 60th birthday*. vol. 50 of Springer Proceedings in Mathematics & Statistics. Springer; p. 613–634.
- de Klerk E. 2006. Aspects of semidefinite programming: Interior point algorithms and selected applications. vol. 65. Springer Science & Business Media.
- de Klerk E, Roos C, Terlaky T. 1997. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Operations Research Letters*. 20(5):213–221.
- de Klerk E, Roos C, Terlaky T. 1998. Infeasible-start semidefinite programming algorithms via self dual embeddings. *Fields Institute Communications*. 18:215–236.

- Geoffrion AM. 1972. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*. 10(4):237–260.
- Goldfarb D, Scheinberg K. 1998. Interior point trajectories in semidefinite programming. *SIAM Journal on Optimization*. 8(4):871–886.
- Goldfarb D, Scheinberg K. 1999. On parametric semidefinite programming. *Applied Numerical Mathematics*. 29(3):361 – 377.
- Golub GH, Van Loan CF. 2013. *Matrix computations*. The Johns Hopkins University Press.
- Güler O, den Hertog D, Roos C, Terlaky T, Tsuchiya T. 1993. Degeneracy in interior point methods for linear programming: a survey. *Annals of Operations Research*. 46(1):107–138.
- Halická M. 2002. Analyticity of the central path at the boundary point in semidefinite programming. *European Journal of Operational Research*. 143(2):311–324.
- Halická M, de Klerk E, Roos C. 2002. On the convergence of the central path in semidefinite optimization. *SIAM Journal on Optimization*. 12(4):1090–1099.
- Hoffman A. 1952. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*. 49(4):263–265.
- Horn RA, Johnson CR. 2012. *Matrix analysis*. Cambridge University Press.
- Illés T, Peng J, Roos C, Terlaky T. 2000. A strongly polynomial rounding procedure yielding a maximally complementary solution for $P_*(\kappa)$ linear complementarity problems. *SIAM Journal on Optimization*. 11(2):320–340.
- Lee EY. 2011. Extension of Rotfel’d theorem. *Linear Algebra and its Applications*. 435(4):735 – 741.
- Lu LZ, Pearce C. 2000. Some new bounds for singular values and eigenvalues of matrix products. *Annals of Operations Research*. 98(1):141–148.
- Luo XD, Luo ZQ. 1994. Extension of Hoffman’s error bound to polynomial systems. *SIAM Journal on Optimization*. 4(2):383–392.
- Luo ZQ, Sturm JF, Zhang S. 1998. Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming. *SIAM Journal on Optimization*. 8(1):59–81.
- Nesterov Y, Nemirovskii A. 1994. Interior-point polynomial algorithms in convex programming. vol. 13. Society for Industrial and Applied Mathematics.
- Nesterov Y, Todd MJ. 1997. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*. 22(1):1–42.
- Nocedal J, Wright S. 2006. *Numerical optimization*. 2nd ed. Springer Science & Business Media.
- Peña J, Roshchina V. 2013. A complementarity partition theorem for multifold conic systems. *Mathematical Programming*. 142(1):579–589.
- Ramana MV. 1993. An algorithmic analysis of multiquadratic and semidefinite programming problems [dissertation]. Johns Hopkins University.
- Ramana MV. 1997. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*. 77(1):129–162.
- Roos C, Terlaky T, Vial JP. 2005. *Interior point methods for linear optimization*. Springer Science & Business Media.
- Stewart GW, Sun Jg. 1990. *Matrix perturbation theory*. Academic Press.
- Sturm JF. 2000. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*. 10(4):1228–1248.
- Terlaky T, Wang Z. 2014. On the identification of the optimal partition of second order cone optimization problems. *SIAM Journal on Optimization*. 24(1):385–414.
- Vandenbergh L, Boyd S. 1996. Semidefinite programming. *SIAM Review*. 38(1):49–95.
- Ye Y. 1992. On the finite convergence of interior-point algorithms for linear programming. *Mathematical Programming*. 57(1):325–335.
- Ye Y. 1994. Toward probabilistic analysis of interior-point algorithms for linear programming. *Mathematics of Operations Research*. 19(1):38–52.
- Yildirim E. 2004. Unifying optimal partition approach to sensitivity analysis in conic optimization. *Journal of Optimization Theory and Applications*. 122(2):405–423.

Appendix A. A lower bound on σ

In this section, we derive a lower bound on the condition number σ defined in (2). To do so, we resort to a technical lemma in Ramana (1993).

An integral polynomial map $f : \mathbb{R}^s \rightarrow \mathbb{R}^t$ 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

$$V(f) := \{x \mid f^i(x) \Delta_i 0, \quad \forall i\},$$

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

$$l(n) := 1 + \lceil \log_2(|n| + 1) \rceil,$$

in which $\log_2(\cdot)$ 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.1 (Lemma 3.1 in Ramana (1993)). *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.2. *Let the SDO problems (P) and (D) 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 $i = 1, \dots, m$. Then, for the condition number σ we have*

$$\sigma \geq \min \left\{ \frac{1}{r_{\mathcal{P}^*} \sum_{i=1}^m \|A^i\|}, \frac{1}{r_{\mathcal{D}^*}} \right\}, \quad (\text{A1})$$

where

$$\begin{aligned} \log_2(r_{\mathcal{P}^*}) &= (L + 2) \left(\max\{n, 3\} (6n^2 + 2n + m) \right)^{5n^2 + 2m}, \\ \log_2(r_{\mathcal{D}^*}) &= (L + 2) \left(\max\{n, 3\} (7n^2 + 2n + 2m) \right)^{6n^2 + m}. \end{aligned}$$

Proof. Recall from (3) and (4) that

$$\sigma_{\mathcal{B}} \geq \lambda_{\min}(Q_{\mathcal{B}}^T \tilde{X} Q_{\mathcal{B}}), \quad \sigma_{\mathcal{N}} \geq \lambda_{\min}(Q_{\mathcal{N}}^T \tilde{X} Q_{\mathcal{N}}), \quad \forall (\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*,$$

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 (1993) 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 2.2, any primal-

dual optimal pair is a solution to the following LMI system

$$\begin{cases} A^i \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T &= b_i, \quad i = 1, \dots, m, \\ C - \sum_{i=1}^m y_i A^i &= Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T, \\ U_X, U_S &\succeq 0, \end{cases} \quad (\text{A2})$$

where $U_X \in \mathbb{S}_+^{n_{\mathcal{B}}}$ and $U_S \in \mathbb{S}_+^{n_{\mathcal{N}}}$ are as defined in Theorem 2.2, 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 (A2), i.e.,

$$\begin{cases} A^i \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T &= b_i, \quad i = 1, \dots, m, \\ C - \sum_{i=1}^m y_i A^i &= Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T, \\ U_X, U_S &\succ 0. \end{cases} \quad (\text{A3})$$

Then for a given orthonormal basis $Q_{\mathcal{B}}$, the analytic center of the primal optimal set can be computed by solving

$$\begin{aligned} \max \quad & \log(\det(U_{X^a})) \\ \text{s.t.} \quad & A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T = b_i, \quad i = 1, \dots, m, \\ & U_{X^a} \succ 0. \end{aligned} \quad (\text{A4})$$

Problem (A4) 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 (A4). Further, there exists a vector of Lagrange multipliers $u \in \mathbb{R}^m$ so that the following system of optimality conditions has a solution:

$$\begin{cases} U_{X^a}^{-1} - \sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}} &= 0, \\ A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T &= b_i, \quad i = 1, \dots, m, \\ U_{X^a} &\succ 0. \end{cases} \quad (\text{A5})$$

For any solution (U_{X^a}, u) of (A5), 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 (A5) that

$$\begin{aligned} \lambda_{\min}(U_{X^a}) &= \frac{1}{\lambda_{\max}\left(\sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\right)} \geq \frac{1}{\left\|\sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\right\|} \\ &\geq \frac{1}{\sum_{i=1}^m |u_i| \|Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\|} \\ &\geq \frac{1}{\sum_{i=1}^m |u_i| \|A^i\|}, \end{aligned} \quad (\text{A6})$$

where we have used the triangle inequality and the fact that $\|Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\| \leq \|A^i\|$. Note that the bound (A6) depends on an upper bound on $|u_i|$ which itself relies on $Q_{\mathcal{B}}$. In general, 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 (A5) to describe the analytic center of the optimal set. Then a direct application of Lemma A.1 to the embedded set yields an upper bound on $|u_i|$.

Assume that $Q_{\mathcal{B}}$ is an unknown orthonormal basis in (A4), i.e., $Q_{\mathcal{B}}$ is still an orthonormal basis for \mathcal{B} but acts as an unknown in (A4), which leads to a nonconvex optimization problem in $Q_{\mathcal{B}}$ and U_{X^a} . Then, problem (A4) can equivalently be written, see e.g., Theorem 2.1 in Geoffrion (1972), as

$$\max_{Q_{\mathcal{B}} \in \Gamma_{\mathcal{B}}} \max_{U_{X^a} \succ 0} \left\{ \log(\det(U_{X^a})) : A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T = b_i, \quad i = 1, \dots, m \right\}. \quad (\text{A7})$$

Any optimal solution $(Q_{\mathcal{B}}, U_{X^a})$ of (A4) is also optimal for (A7) and vice versa. This is due to the fact that the optimal solution of the inner maximization problem in (A7) is attained. By Lemma 2.3, Theorem 2.2 and (A3), the set $\Gamma_{\mathcal{B}}$ is compact, and it is equivalent to the set of all $Q_{\mathcal{B}}$ with orthonormal columns by which (A3) is feasible. Since the unique optimal solution of the inner maximization problem in (A7) is attained, and its set of Lagrange multipliers is nonempty, then (A5) with $\Gamma_{\mathcal{B}}$ describes the analytic center of the primal optimal set, see Section 4.2 in Geoffrion (1972) for a similar argument in the context of the generalized Benders decomposition.

Now, we apply Lemma A.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

$$f_p : \mathbb{R}^{n_{\mathcal{B}} \times n_{\mathcal{B}}} \times \mathbb{R}^m \times \mathbb{R}^{n_{\mathcal{B}} \times n_{\mathcal{B}}} \times \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}} \times \mathbb{R}^m \times \mathbb{R}^{n \times n_{\mathcal{B}}} \times \mathbb{R}^{n \times n_{\mathcal{N}}} \rightarrow \mathbb{R}^{t_p}$$

as defined below

$$f_p(\vartheta_p) := \begin{bmatrix} \text{vec}\left(Z_X - \sum_{i=1}^m u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}\right) \\ \text{vec}(U_{X^a} Z_X - I_{n_{\mathcal{B}}}) \\ A^1 \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T - b_1 \\ \vdots \\ A^m \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T - b_m \\ \text{vec}\left(C - \sum_{i=1}^m y_i A^i - Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T\right) \\ \text{vec}(Q_{\mathcal{B}}^T Q_{\mathcal{B}} - I_{n_{\mathcal{B}}}) \\ \text{vec}(Q_{\mathcal{N}}^T Q_{\mathcal{N}} - I_{n_{\mathcal{N}}}) \\ \text{vec}(Q_{\mathcal{B}}^T Q_{\mathcal{N}}) \end{bmatrix}, \quad (\text{A8})$$

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

$$\Omega_p := \left\{ \vartheta_p \mid \det(U_{X^a}[i]) > 0, \det(U_S[j]) > 0, \quad i = 1, \dots, n_B, j = 1, \dots, n_N \right\}, \quad (\text{A9})$$

in which $U_{X^a}[i]$ denotes the i^{th} leading principal submatrix of U_{X^a} . Indeed, the strict inequalities in (A9) 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 $V(f_p) = \{\vartheta_p \mid f_p(\vartheta_p) = 0\}$, 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_B) which is the analytic center of the primal optimal set, since it satisfies the constraints in (A5).

The solution set Ω_p is characterized by $n_B + n_N$ integer polynomials of degree at most $\max\{n_B, n_N\}$. 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 $u_i Q_B^T A^i Q_B$ 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 (A8) and (A9) is bounded above by $L + l(2) - 1 = L + 2$, see Section 3.1 in Ramana (1993).

Consequently, by applying Lemma A.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 $\|\vartheta_p\|_2 \leq r_{\mathcal{P}^*}$, where

$$\begin{aligned} \log_2(r_{\mathcal{P}^*}) &= (L + 2)(\bar{t}_p \bar{d}_p)^{\bar{s}_p}, \\ \bar{d}_p &:= \max\{n_B, n_N, 3\} \leq \max\{n, 3\}, \\ \bar{t}_p &:= t_p + n_B + n_N = 3n_B^2 + n_N^2 + n_B n_N + n_B + n_N + n^2 + m \leq 6n^2 + 2n + m, \\ \bar{s}_p &:= 2n_B^2 + n_N^2 + n(n_B + n_N) + 2m \leq 5n^2 + 2m, \end{aligned}$$

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 $|u_i| \leq \|u\|_2 \leq r_{\mathcal{P}^*}$. Then, using inequality (A6), we get

$$\sigma_B \geq \lambda_{\min}(U_{X^a}) \geq \frac{1}{\sum_{i=1}^m |u_i| \|A^i\|} \geq \frac{1}{r_{\mathcal{P}^*} \sum_{i=1}^m \|A^i\|}.$$

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 σ_N . Notice that for a given orthonormal basis Q_N , the analytic center of the dual optimal set can be obtained by solving

$$\begin{aligned} \max \quad & \log(\det(U_{S^a})) \\ \text{s.t.} \quad & \sum_{i=1}^m y_i^a A^i + Q_N U_{S^a} Q_N^T = C, \\ & U_{S^a} \succ 0, \end{aligned} \quad (\text{A10})$$

which is a convex optimization problem with strictly concave objective function. The

optimality conditions for (A10) are given by

$$\begin{cases} U_{S^a}^{-1} - Q_{\mathcal{N}}^T W Q_{\mathcal{N}} &= 0, \\ A^i \bullet W &= 0, \quad i = 1, \dots, m, \\ \sum_{i=1}^m y_i^a A^i + Q_{\mathcal{N}} U_{S^a} Q_{\mathcal{N}}^T &= C, \\ U_{S^a} \succ 0, \end{cases} \quad (\text{A11})$$

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 symmetry⁴ of W . Then the optimality conditions (A11) imply

$$\lambda_{\min}(U_{S^a}) = \frac{1}{\lambda_{\max}(Q_{\mathcal{N}}^T W Q_{\mathcal{N}})} \geq \frac{1}{\|Q_{\mathcal{N}}^T W Q_{\mathcal{N}}\|} \geq \frac{1}{\|W\|}. \quad (\text{A12})$$

Let $\vartheta_d := (U_{S^a}, y^a, U_X, Z_S, W, Q_{\mathcal{B}}, Q_{\mathcal{N}})$, where $Z_S \in \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}}$, and consider the solution set

$$V(f_d) := \left\{ \vartheta_d \mid f_d(\vartheta_d) = 0 \right\},$$

where the integral polynomial map

$$f_d : \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}} \times \mathbb{R}^m \times \mathbb{R}^{n_{\mathcal{B}} \times n_{\mathcal{B}}} \times \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n_{\mathcal{B}}} \times \mathbb{R}^{n \times n_{\mathcal{N}}} \rightarrow \mathbb{R}^{t_d}$$

is defined as

$$f_d(\vartheta_d) := \begin{bmatrix} \text{vec}(Z_S - Q_{\mathcal{N}}^T W Q_{\mathcal{N}}) \\ \text{vec}(U_{S^a} Z_S - I_{n_{\mathcal{N}}}) \\ A^1 \bullet W \\ \vdots \\ A^m \bullet W \\ A^1 \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T - b_1 \\ \vdots \\ A^m \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T - b_m \\ \text{vec}\left(C - \sum_{i=1}^m y_i^a A^i - Q_{\mathcal{N}} U_{S^a} Q_{\mathcal{N}}^T\right) \\ \text{vec}(W - W^T) \\ \text{vec}(Q_{\mathcal{B}}^T Q_{\mathcal{B}} - I_{n_{\mathcal{B}}}) \\ \text{vec}(Q_{\mathcal{N}}^T Q_{\mathcal{N}} - I_{n_{\mathcal{N}}}) \\ \text{vec}(Q_{\mathcal{B}}^T Q_{\mathcal{N}}) \end{bmatrix}, \quad (\text{A13})$$

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

$$\Omega_d := \left\{ \vartheta_d \mid \det(U_X[i]) > 0, \det(U_{S^a}[j]) > 0, \quad i = 1, \dots, n_{\mathcal{B}}, \quad j = 1, \dots, n_{\mathcal{N}} \right\}.$$

⁴Note 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 (A11).

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.1](#) implies the existence of $\vartheta_d \in V(f_d) \cap \Omega_d$ so that $\|\vartheta_d\|_2 \leq r_{\mathcal{D}^*}$, where

$$\begin{aligned} \log_2(r_{\mathcal{D}^*}) &= (L+2)(\bar{t}_d \bar{d}_d)^{\bar{s}_d}, \\ \bar{d}_d &:= \max\{n_{\mathcal{B}}, n_{\mathcal{N}}, 3\} \leq \max\{n, 3\}, \\ \bar{t}_d &:= t_d + n_{\mathcal{B}} + n_{\mathcal{N}} = n_{\mathcal{B}}^2 + 3n_{\mathcal{N}}^2 + n_{\mathcal{B}}n_{\mathcal{N}} + n_{\mathcal{B}} + n_{\mathcal{N}} + 2n^2 + 2m \leq 7n^2 + 2n + 2m, \\ \bar{s}_d &:= n_{\mathcal{B}}^2 + 2n_{\mathcal{N}}^2 + n^2 + n(n_{\mathcal{B}} + n_{\mathcal{N}}) + m \leq 6n^2 + m, \end{aligned}$$

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 $\|W\| \leq r_{\mathcal{D}^*}$ and [\(A12\)](#). This completes the proof. \square

Remark 9. For the special case $n_{\mathcal{B}} = 0$ we get $\sigma = \sigma_{\mathcal{N}}$ by [\(2\)](#), and thus the lower bound [\(A1\)](#) 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} , i.e.,

$$\begin{aligned} \max \quad & \log(\det(S^a)) \\ \text{s.t.} \quad & \sum_{i=1}^m y_i^a A^i + S^a = C, \\ & S^a \succ 0. \end{aligned} \tag{A14}$$

It is easy to verify that the application of Lemma [A.1](#) to the system of optimality conditions of [\(A14\)](#) gives an integral polynomial map with strictly fewer number of polynomials and variables than [\(A13\)](#), which yields a smaller $r_{\mathcal{D}^*}$.

Example A.3. From [\(A1\)](#) we get a doubly exponentially small lower bound on σ . Consider the SDO problem in Example [3.2](#) for which we have $\sigma \leq 20 \times 4^{-2\bar{m}}$. Given $n_{\mathcal{B}} \leq 2\bar{m}+1$, $n_{\mathcal{N}} \leq 2\bar{m}+1$, $\|A^1\| = \sqrt{\bar{m}} + 8$, $\|A^{i+1}\| = 3$ for $i = 1, \dots, \bar{m}-1$, $\|A^{\bar{m}+1}\| = \sqrt{2}$, and $L = l(2) = 1 + \lceil \log_2(3) \rceil = 3$, we can compute the lower bound [\(A1\)](#). To do so, we have

$$\begin{aligned} \bar{t}_p &\leq 6(2\bar{m}+1)^2 + 2(2\bar{m}+1) + \bar{m} + 1, & \bar{t}_d &\leq 7(2\bar{m}+1)^2 + 2(2\bar{m}+1) + 2\bar{m} + 2, \\ \bar{s}_p &\leq 5(2\bar{m}+1)^2 + 2\bar{m} + 2, & \bar{s}_d &\leq 6(2\bar{m}+1)^2 + \bar{m} + 1, \\ \bar{d}_p &= \bar{d}_d \leq 2\bar{m} + 1, \\ \sum_{i=1}^m \|A^i\| &= \sqrt{\bar{m}} + 8 + 3(\bar{m}-1) + \sqrt{2}. \end{aligned}$$

Therefore, we get

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

Consequently,

$$\sigma \geq \min \left\{ \left(\sqrt{\bar{m}} + 8 + 3(\bar{m} - 1) + \sqrt{2} \right) 2^{-5 \times (48\bar{m}^3 + 82\bar{m}^2 + 47\bar{m} + 9)^{20\bar{m}^2 + 22\bar{m} + 7}}, \right. \\ \left. 2^{-5 \times (56\bar{m}^3 + 96\bar{m}^2 + 56\bar{m} + 11)^{24\bar{m}^2 + 25\bar{m} + 7}} \right\}.$$

Appendix B. Error bound for an LMI system

An LMI system is defined as

$$\begin{cases} X \in D_0 + \mathcal{L}, \\ X \succeq 0, \end{cases} \quad (\text{B1})$$

where D_0 is a symmetric matrix and $\mathcal{L} \subset \mathbb{S}^n$ denotes a linear subspace of symmetric matrices. For system (B1), we consider a sequence of solutions denoted by $X(\epsilon)$ for $\epsilon > 0$ which satisfies

$$\text{dist}(X(\epsilon), D_0 + \mathcal{L}) \leq \epsilon, \quad \lambda_{\min}(X(\epsilon)) \geq -\epsilon, \quad (\text{B2})$$

for all $\epsilon > 0$, where $\text{dist}(\cdot, \cdot)$ denotes the distance function with respect to a norm. Further, $\bar{\mathcal{L}}$ is defined as the smallest subspace containing $D_0 + \mathcal{L}$, i.e.,

$$\bar{\mathcal{L}} := \{X \in \mathbb{S}^n \mid X + \beta D_0 \in \mathcal{L} \text{ for some } \beta\}.$$

The following lemma is in order.

Lemma B.1 (Adopted from Theorem 3.3 in Sturm (2000)). *Let $\{X(\epsilon) \mid 0 < \epsilon \leq 1\}$ be a set of solutions so that $\|X(\epsilon)\|$ is bounded and (B2) holds for all $0 < \epsilon \leq 1$. Then there exist a positive condition number c independent of ϵ and a positive exponent γ such that*

$$\text{dist}\left(X(\epsilon), (D_0 + \mathcal{L}) \cap \mathbb{S}_+^n\right) \leq c\epsilon^\gamma,$$

where $\gamma = 2^{-d(\bar{\mathcal{L}}, \mathbb{S}_+^n)}$ in which $d(\bar{\mathcal{L}}, \mathbb{S}_+^n)$ denotes the degree of singularity of the linear subspace $\bar{\mathcal{L}}$.

Lemma B.2 (Adopted from Theorem 3.6 in Sturm (2000)). *For a linear subspace $\bar{\mathcal{L}} \subset \mathbb{S}^n$, we have*

$$d(\bar{\mathcal{L}}, \mathbb{S}_+^n) \leq \min\{n - 1, \dim(\bar{\mathcal{L}}), \dim(\bar{\mathcal{L}}^\perp)\}.$$

Example B.3. We can show that the upper bound given in Lemma B.2 is indeed tight. To do so, consider the following LMI system

$$\begin{cases} X_{11} = 0, \\ X_{kk} = x_{1(k+1)}, \quad k = 2, \dots, n-1, \\ X \succeq 0, \end{cases}$$

where the set of feasible solutions is given by

$$X = \begin{bmatrix} \mathbf{0}_{(n-1) \times (n-1)} & 0 \\ 0 & X_{nn} \end{bmatrix}, \quad X_{nn} \geq 0.$$

Using the facial reduction procedure in [Sturm \(2000\)](#), 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 [Sturm \(2000\)](#) for a simple demonstration of the facial reduction algorithm. Additional examples of the facial reduction for SDO problems can be found in [Cheung \(2013\)](#).

Appendix C. Error bound for a polynomial system

In this section, we present the extension of the Hoffman error bound to a polynomial system from [Luo and Luo \(1994\)](#), where the solution set \mathcal{C} is defined as

$$\mathcal{C} := \{x \in \mathbb{R}^n \mid g_1(x) \leq 0, \dots, g_{m_1}(x) \leq 0, h_1(x) = 0, \dots, h_{m_2}(x) = 0\},$$

in which g_j for $j = 1, \dots, m_1$ and h_k for $k = 1, \dots, m_2$ are polynomials with real coefficients.

Lemma C.1 (Adopted from Theorem 2.2 in [Luo and Luo \(1994\)](#)). *Assume that $\mathcal{C} \neq \emptyset$. Then there exist a positive condition number ρ and exponents $v > 0$ and $v' \geq 0$ so that*

$$\text{dist}(x, \mathcal{C}) \leq \rho(1 + \|x\|_2)^{v'} (\|[g(x)]_+\|_2 + \|h(x)\|_2)^v, \quad \forall x \in \mathbb{R}^n,$$

where

$$\begin{aligned} [g(x)]_+ &:= (\max\{g_1(x), 0\}, \dots, \max\{g_{m_1}(x), 0\})^T, \\ h(x) &:= (h_1(x), \dots, h_{m_2}(x))^T. \end{aligned}$$