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

[a]Department of Industrial and Systems Engineering, Lehigh University, 200 W. Packer Avenue, Bethlehem, PA 18015-1582

ARTICLE HISTORY
Compiled October 23, 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, from a central solution sufficiently close to the optimal set, 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

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

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

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

CONTACT Ali Mohammad-Nezhad. Email: alm413@lehigh.edu; Tel.: +1(610)758-2903.

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

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

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

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

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

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

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[1]. 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,

[1] By strong duality we mean that both the primal and dual problems have optimal solutions with equal objective values, see e.g., Theorem 5.81 in Bonnans and Shapiro (2000).

which is defined as the set of solutions of

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

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^μ, y^μ, S^μ) to this system exists, and it is uniquely defined under the interior point condition and the linear independence of A^i for $i = 1, \ldots, m$, see Theorem 3.1 in de Klerk (2006). For $0 \leq \mu \leq \bar{\mu}$, where $\bar{\mu} > 0$, the set of solutions of (1) is bounded, see Lemma 3.2 in de Klerk (2006), 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 (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 $\operatorname{rank}(X^* + S^*)$ is maximal over the optimal set.

Definition 1.2 (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 $\mu \to 0$. However, the first order derivatives may be unbounded if strict complementarity fails to hold. Using the strict complementarity condition only, Halická (2002) showed the extension of the analyticity of the central path to $\mu = 0$.

1.1. Motivation and related works

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, from a central solution sufficiently close to the optimal set. 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.

1.2. Contributions

In this paper, we consider the identification of the optimal partition for an SDO problem. As an extension from LO, the optimal partition of SDO is a 3-tuple of mutually orthogonal subspaces $\mathcal{B}, \mathcal{T}, \mathcal{N} \subseteq \mathbb{R}^n$ such that $\mathcal{B} + \mathcal{T} + \mathcal{N} = \mathbb{R}^n$. Our goal is to approximate the optimal partition of an SDO problem using the limiting behavior of the central path and a bounded sequence of interior solutions in a neighborhood of the central path. We show how the complexity of approximating the optimal partition depends on condition numbers of the problem. We quantify the magnitude of the eigenvalues of an interior solution (a central solution, or a solution in a neighborhood of the central path) 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. In Theorems 3.8, 3.9, and 4.3 we use bounds on the eigenvalues to identify the subsets of the eigenvectors of the interior solutions whose accumulation points form orthonormal bases for the subspaces of the optimal partition. This is referred to as an *approximation of the optimal partition*. We show that even approximation of the optimal partition is notably more expensive than the identification of the optimal partition for LO. Finally, in Theorems 3.11 and 3.13 we evaluate the proximity of the approximation of the optimal partition to the true optimal partition.

1.3. Organization of the paper

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 and directions for future research are presented in Section 5.

1.4. Notation

Throughout this paper, $\mathbb{S}^n$ denotes the space of symmetric matrices of size n. We adopt the notation $(.;.;\ldots;.)$ to indicate the concatenation of column vectors. 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

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

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. Associated with a symmetric matrix $\operatorname{svec}: \mathbb{S}^n \to \mathbb{R}^{n(n+1)/2}$ is a linear bijection which stacks the upper triangular part into a vector by multiplying the off-diagonal entries of the symmetric matrix by $\sqrt{2}$, i.e.,

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

and $\operatorname{smat}: \mathbb{R}^{n(n+1)/2} \to \mathbb{S}^n$ denotes the inverse of $\operatorname{svec}(.)$. Note that $\operatorname{svec}(.)$ is a linear isometry between $\mathbb{S}^n$ and $\mathbb{R}^{n(n+1)/2}$. The kernel and range of a matrix are denoted by $\operatorname{Ker}(.)$ and $\mathcal{R}(.)$, respectively, and $\operatorname{ri}(.)$ stands for the relative interior of a convex set. For a linear subspace $\mathcal{L}$, $\mathcal{L}^\perp$ denotes the orthogonal complement of $\mathcal{L}$. Finally, the Frobenius norm of a matrix is denoted by $\|.\|$, and $\|.\|_2$ 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

$$\begin{aligned} A^i \bullet X &= b_i, \quad i = 1, \ldots, m, \\ \sum_{i=1}^m A^i y_i + S &= C, \\ XS &= 0, \quad X, S \succeq 0, \end{aligned} \tag{2}$$

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 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 $X^* \in \operatorname{ri}(\mathcal{P}^*)$ have the same range space. Analogously, all S^* have identical range spaces, where $(y^*, S^*) \in \operatorname{ri}(\mathcal{D}^*)$, 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

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

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

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

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*

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

where $U_{\tilde{X}} \in \mathbb{S}_+^{n_{\mathcal{B}}}$ and $U_{\tilde{S}} \in \mathbb{S}_+^{n_{\mathcal{N}}}$. If $n_{\mathcal{B}} > 0$ and $X^* \in \mathrm{ri}(\mathcal{P}^*)$, then there exists $U_{X^*} \succ 0$. Similarly, if $n_{\mathcal{N}} > 0$ and $(y^*, S^*) \in \mathrm{ri}(\mathcal{D}^*)$, 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 $\mathcal{P}^* = \mathrm{ri}(\mathcal{P}^*) = \{0\}$, 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 $(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^* \times \mathcal{D}^*)$ with respect to Q reveals the optimal partition as

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

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

$$\begin{aligned} Q_{\mathcal{T}\cup\mathcal{N}}^T \tilde{X} Q_{\mathcal{T}\cup\mathcal{N}} &= 0, \qquad \forall\ \tilde{X} \in \mathcal{P}^*, \\ Q_{\mathcal{B}\cup\mathcal{T}}^T \tilde{S} Q_{\mathcal{B}\cup\mathcal{T}} &= 0, \qquad \forall\ (\tilde{y}, \tilde{S}) \in \mathcal{D}^*, \end{aligned} \tag{3}$$

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

Remark 3. From now on, unless stated otherwise, we use a fixed orthonormal basis Q both to represent the optimal solutions and to make an orthogonal transformation with.

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 provide a characterization of the optimal partition using the eigenvectors of a central solution, when μ is sufficiently close to 0. 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. Consider the orthogonal transformation of X^μ with respect to Q denoted by

$$\hat{X}^\mu := \begin{pmatrix} \hat{X}^\mu_\mathcal{B} & \hat{X}^\mu_{\mathcal{B}\mathcal{T}} & \hat{X}^\mu_{\mathcal{B}\mathcal{N}} \\ \hat{X}^\mu_{\mathcal{T}\mathcal{B}} & \hat{X}^\mu_\mathcal{T} & \hat{X}^\mu_{\mathcal{T}\mathcal{N}} \\ \hat{X}^\mu_{\mathcal{N}\mathcal{B}} & \hat{X}^\mu_{\mathcal{N}\mathcal{T}} & \hat{X}^\mu_\mathcal{N} \end{pmatrix}, \tag{4}$$

where $\hat{X}^\mu := Q^T X^\mu Q$. The orthogonal transformation of S^μ is defined analogously. Since the central path converges to a maximally complementary optimal solution, from the orthogonal transformation in (4) we have

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

and

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

where $\hat{X}^\mu_{\mathcal{B}} = Q^T_{\mathcal{B}} X^\mu Q_{\mathcal{B}}$ and $\hat{S}^\mu_{\mathcal{N}} = Q^T_{\mathcal{N}} S^\mu Q_{\mathcal{N}}$, see (3). We define a condition number and employ an error bound result for LMIs to derive bounds on the magnitude of vanishing blocks of $\hat{X}^\mu$ and $\hat{S}^\mu$ as $\mu \to 0$.

3.1. Bounds on $\hat{S}^\mu_{\mathcal{B}}$ and $\hat{X}^\mu_{\mathcal{N}}$

To derive bounds on $\hat{S}^\mu_{\mathcal{B}}$ and $\hat{X}^\mu_{\mathcal{N}}$, we define a condition number σ as

$$\sigma := \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}, \tag{5}$$

where

$$\sigma_{\mathcal{B}} := \begin{cases} \max\limits_{\tilde{X}\in\mathcal{P}^*} \lambda_{\min}(Q^T_{\mathcal{B}} \tilde{X} Q_{\mathcal{B}}) \\ \qquad = \max\limits_{\bar{Q}_{\mathcal{B}}\in\Gamma_{\mathcal{B}}} \max\limits_{\tilde{X}\in\mathcal{P}^*} \lambda_{\min}(\bar{Q}^T_{\mathcal{B}} \tilde{X} \bar{Q}_{\mathcal{B}}), & \text{if } n_{\mathcal{B}} > 0, \\ \infty, & \text{if } n_{\mathcal{B}} = 0, \end{cases} \tag{6}$$

$$\sigma_{\mathcal{N}} := \begin{cases} \max\limits_{(\tilde{y},\tilde{S})\in\mathcal{D}^*} \lambda_{\min}(Q^T_{\mathcal{N}} \tilde{S} Q_{\mathcal{N}}) \\ \qquad = \max\limits_{\bar{Q}_{\mathcal{N}}\in\Gamma_{\mathcal{N}}} \max\limits_{(\tilde{y},\tilde{S})\in\mathcal{D}^*} \lambda_{\min}(\bar{Q}^T_{\mathcal{N}} \tilde{S} \bar{Q}_{\mathcal{N}}), & \text{if } n_{\mathcal{N}} > 0, \\ \infty, & \text{if } n_{\mathcal{N}} = 0. \end{cases} \tag{7}$$

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. Thus, σ is well-defined by Remark 2. Assume that $n_{\mathcal{B}} > 0$. Then there exists $\check{X} \in \mathcal{P}^*$ so that $\lambda_{\min}(Q^T_{\mathcal{B}} \check{X} Q_{\mathcal{B}}) > 0$. By the compactness of $\mathcal{P}^*$ and the continuity of the eigenvalues, there exists $\bar{X} \in \mathcal{P}^*$ so that

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

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

Remark 4. 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

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

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

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

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

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

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

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

From the 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

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

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

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

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

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

□

Now, the following lemma is in order.

Lemma 3.3. *For a given central solution* (X^μ, y^μ, S^μ) *we have*

$$\operatorname{trace}(\hat{X}^\mu_{\mathcal{N}}) \leq \frac{n\mu}{\sigma}, \qquad \operatorname{trace}(\hat{S}^\mu_{\mathcal{B}}) \leq \frac{n\mu}{\sigma}.$$

Proof. By the compactness of $\mathcal{P}^*$ and the continuity of the eigenvalues, there exists $\bar{X} \in \mathcal{P}^*$ so that $\sigma_{\mathcal{B}} = \lambda_{\min}(Q^T_{\mathcal{B}} \bar{X} Q_{\mathcal{B}})$ as defined in (6). Analogously, it follows from (7) that $\sigma_{\mathcal{N}} = \lambda_{\min}(Q^T_{\mathcal{N}} \bar{S} Q_{\mathcal{N}})$ for some $(\bar{y}, \bar{S}) \in \mathcal{D}^*$. Since $\sigma = \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}$, there exists $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

$$\lambda_{\min}(U_{\bar{X}}) \geq \sigma, \qquad \lambda_{\min}(U_{\bar{S}}) \geq \sigma, \tag{8}$$

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

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

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

$$X^\mu \bullet \bar{S} + \bar{X} \bullet S^\mu = n\mu. \tag{9}$$

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

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

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

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

which by the lower bounds (8) gives

$$\operatorname{trace}(\hat{X}^\mu_{\mathcal{N}}) \le \frac{n\mu}{\sigma}.$$

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

$$\operatorname{trace}(\hat{S}^\mu_{\mathcal{B}}) \le \frac{n\mu}{\sigma}.$$

The proof is complete. □

3.2. Bounds on $Q^T_{\mathcal{T}\cup\mathcal{N}} X^\mu Q_{\mathcal{T}\cup\mathcal{N}}$ and $Q^T_{\mathcal{B}\cup\mathcal{T}} S^\mu Q_{\mathcal{B}\cup\mathcal{T}}$

In order to estimate the magnitude of $Q^T_{\mathcal{T}\cup\mathcal{N}} X^\mu Q_{\mathcal{T}\cup\mathcal{N}}$ and $Q^T_{\mathcal{B}\cup\mathcal{T}} S^\mu Q_{\mathcal{B}\cup\mathcal{T}}$ we derive an upper bound on the distance of a central solution to the optimal set. Toward this end, we resort to a Hölderian error bound result for an LMI system from Theorem 3.3 in Sturm (2000). An LMI system is defined as

$$\begin{cases} X \in D_0 + \mathcal{L}, \\ X \succeq 0, \end{cases} \tag{10}$$

where D_0 is a symmetric matrix and $\mathcal{L} \subset \mathbb{S}^n$ denotes a linear subspace of symmetric matrices. Further, $\bar{\mathcal{L}}$ is defined as the smallest subspace containing $D_0 + \mathcal{L}$, i.e.,

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

Lemma 3.4 yields an upper bound on the distance of an approximate solution to the solution set obtained from (10).

Lemma 3.4 (Theorem 3.3 in Sturm (2000)). *Let $\{X^\epsilon \mid 0 < \epsilon \le 1\}$ be a set of solutions so that $\|X^\epsilon\|$ is bounded and*

$$\operatorname{dist}\big(X^\epsilon, D_0 + \mathcal{L}\big) \le \epsilon, \qquad \lambda_{\min}(X^\epsilon) \ge -\epsilon \tag{11}$$

holds for all $0 < \epsilon \le 1$, where

$$\operatorname{dist}\big(X^\epsilon, D_0 + \mathcal{L}\big) := \min\big\{\|X - X^\epsilon\| \mid X \in D_0 + \mathcal{L}\big\}.$$

Then there exist a positive condition number c independent of ϵ and a positive exponent γ such that

$$\operatorname{dist}\big(X^\epsilon, (D_0 + \mathcal{L}) \cap \mathbb{S}^n_+\big) \le c\epsilon^\gamma,$$

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

[2] In 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 $D_0 + \mathcal{L}$.

In Lemma 3.5, we employ the error bound result to specify an upper bound on the distance of a central solution to the optimal set. 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 are given by

$$\begin{cases} \mathbf{A}\,\mathrm{svec}(X) = b, \\ \mathrm{svec}(\tilde{S})^T\,\mathrm{svec}(X) = 0, \\ X \succeq 0, \end{cases} \qquad \begin{cases} \mathbf{A}^T y + \mathrm{svec}(S) = \mathrm{svec}(C), \\ \mathrm{svec}(\tilde{X})^T\,\mathrm{svec}(S) = 0, \\ S \succeq 0, \end{cases} \tag{12}$$

where

$$\mathbf{A} := \big(\mathrm{svec}(A^1), \dots, \mathrm{svec}(A^m)\big)^T.$$

By invoking the null and range spaces of $\mathbf{A}$, it is easy to see that the set of all optimal solutions (X, S) is obtained as the set of solutions of the following LMIs

$$\begin{cases} X \in \tilde{X} + \mathrm{smat}\big(\mathrm{Ker}(\mathbf{A})\big) \cap (\mathbb{R}\tilde{S})^\perp, \\ X \succeq 0, \end{cases} \qquad \begin{cases} S \in \tilde{S} + \mathrm{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp, \\ S \succeq 0, \end{cases} \tag{13}$$

where $\mathrm{smat}\big(\mathrm{Ker}(\mathbf{A})\big) \cap (\mathbb{R}\tilde{S})^\perp$ and $\mathrm{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp$ are linear subspaces of $\mathbb{S}^n$, see also Section 4 in Sturm (2000). The minimal subspaces containing the affine subspaces $\tilde{X} + \mathrm{smat}\big(\mathrm{Ker}(\mathbf{A})\big) \cap (\mathbb{R}S)^\perp$ and $\tilde{S} + \mathrm{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp$ are given by

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

in which $\mathbb{R}\tilde{X}$ and $\mathbb{R}\tilde{S}$ denote the set of all multiples of $\tilde{X}$ and $\tilde{S}$, respectively.

By using the Hoffman error bound (Hoffman 1952) we can derive upper bounds on the distance of X^μ and S^μ to the affine subspaces in (13). Recall from (9) that

$$X^\mu \bullet \tilde{S} + \tilde{X} \bullet S^\mu = n\mu,$$

which implies $0 \le X^\mu \bullet \tilde{S} \le n\mu$ and $0 \le \tilde{X} \bullet S^\mu \le n\mu$. Then the application of the Hoffman error bound gives

$$\begin{aligned} &\mathrm{dist}\big(X^\mu, \tilde{X} + \mathrm{smat}\big(\mathrm{Ker}(\mathbf{A})\big) \cap (\mathbb{R}\tilde{S})^\perp\big) \\ &\quad = \mathrm{dist}\big(\mathrm{svec}(X^\mu),\ \big\{x \mid \mathbf{A}x = b,\ \mathrm{svec}(\tilde{S})^T x = 0\big\}\big) \\ &\quad \le \theta_1\big(\|\mathbf{A}\,\mathrm{svec}(X^\mu) - b\|_2 + \mathrm{svec}(X^\mu)^T\,\mathrm{svec}(\tilde{S})\big) \\ &\quad = \theta_1 X^\mu \bullet \tilde{S} \le \theta_1 n\mu, \end{aligned} \tag{14}$$

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

$$\begin{aligned}
&\operatorname{dist}\big(S^\mu, \tilde{S} + \operatorname{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp\big) \\
&\quad = \operatorname{dist}\big(\operatorname{svec}(S^\mu), \big\{s \mid \exists\, y \in \mathbb{R}^m, \mathbf{A}^T y + s = \operatorname{svec}(C), \operatorname{svec}(\tilde{X})^T s = 0\big\}\big) \\
&\quad \le \theta_2\big(\|\mathbf{A}^T y^\mu + \operatorname{svec}(S^\mu) - \operatorname{svec}(C)\|_2 + \operatorname{svec}(S^\mu)^T \operatorname{svec}(\tilde{X})\big) \\
&\quad = \theta_2 S^\mu \bullet \tilde{X} \le \theta_2 n\mu,
\end{aligned} \tag{15}$$

where $\theta_2 > 0$ is the Hoffman condition number, which is dependent on $\mathbf{A}$ and $\tilde{X}$. Now, we present the following lemma, as planned.

Lemma 3.5. *Let $\big(X^\mu, y^\mu, S^\mu\big)$ be a central solution with*

$$\mu \le \hat{\mu} := \frac{1}{n} \min\Big\{\theta_1^{-1}, \theta_2^{-1}\Big\}. \tag{16}$$

Then there exist a positive condition number c independent of μ and an exponent $\gamma > 0$ so that

$$\begin{aligned}
\operatorname{dist}\big(X^\mu, \mathcal{P}^*\big) &\le c(n\mu)^\gamma, \\
\operatorname{dist}\big(S^\mu, \{S \mid \exists\, y \in \mathbb{R}^m \ \text{ s.t. } (y, S) \in \mathcal{D}^*\}\big) &\le c(n\mu)^\gamma,
\end{aligned} \tag{17}$$

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

Proof. The bounds in (17) can be established easily by applying the error bound result, as stated in Lemma 3.4, to the LMIs in (12). As defined by (1), the set of central solutions (X^μ, y^μ, S^μ) for $0 \le \mu \le \hat{\mu}$ is bounded. Additionally, from (14), (15), and (16) we get

$$\begin{aligned}
\operatorname{dist}\big(X^\mu, \tilde{X} + \operatorname{smat}\big(\operatorname{Ker}(\mathbf{A})\big) \cap (\mathbb{R}\tilde{S})^\perp\big) &\le 1, \\
\operatorname{dist}\big(S^\mu, \tilde{S} + \operatorname{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp\big) &\le 1,
\end{aligned}$$

which satisfy the conditions in (11). Therefore, we can conclude from Lemma 3.4 that there exist positive condition numbers c_1 and c_2, both independent of μ, and positive exponents γ_1 and γ_2 so that

$$\begin{aligned}
\operatorname{dist}\big(X^\mu, \big(\tilde{X} + \operatorname{smat}\big(\operatorname{Ker}(\mathbf{A})\big) \cap (\mathbb{R}\tilde{S})^\perp\big) \cap \mathbb{S}^n_+\big) &\le c_1(n\mu)^{\gamma_1}, \\
\operatorname{dist}\big(S^\mu, \big(\tilde{S} + \operatorname{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp\big) \cap \mathbb{S}^n_+\big) &\le c_2(n\mu)^{\gamma_2},
\end{aligned}$$

where $\gamma_1 = 2^{-d(\bar{\mathcal{L}}_{\mathcal{P}^*}, \mathbb{S}^n_+)}$ and $\gamma_2 = 2^{-d(\bar{\mathcal{L}}_{\mathcal{D}^*}, \mathbb{S}^n_+)}$, in which $d(\bar{\mathcal{L}}_{\mathcal{P}^*}, \mathbb{S}^n_+)$ and $d(\bar{\mathcal{L}}_{\mathcal{D}^*}, \mathbb{S}^n_+)$ denote the degree of singularity of the subspaces $\bar{\mathcal{L}}_{\mathcal{P}^*}$ and $\bar{\mathcal{L}}_{\mathcal{D}^*}$, respectively. Setting $\gamma := \min\{\gamma_1, \gamma_2\}$ and $c := \max\{c_1, c_2\}$ we get the result of the lemma. □

Remark 5. For the special case $\mathcal{T} = \{0\}$ there exists a stronger bound on the distance of a central solution to the optimal set. In this case, the central path converges to the analytic center of the optimal set. Hence, a direct application of Theorem 3.5 in (Luo et al. 1998) to the central solutions gives

$$\|X^\mu - X^a\| = \mathcal{O}(\mu), \quad \text{and} \quad \|S^\mu - S^a\| = \mathcal{O}(\mu)$$

for any SDO problem with strict complementarity condition, while Lemma 3.5 gives an upper bound $\mathcal{O}(\sqrt{n\mu})$. This is due to the fact that the degree of singularity of $\bar{\mathcal{L}}_{\mathcal{P}^*}$ and $\bar{\mathcal{L}}_{\mathcal{D}^*}$ is at most 1 when the strict complementarity condition holds, see Sturm (2000).

Remark 6. From Theorem 3.6 in Sturm (2000) we can get a nontrivial upper bound $n-1$ on the degree of singularity[3]. Therefore, we have $\gamma \geq 2^{1-n}$ for $n \geq 2$. However, we are not aware of any method to compute an upper bound on the condition number c.

Now, we use the upper bounds in (17) to derive bounds on the vanishing blocks $Q^T_{\mathcal{T}\cup\mathcal{N}} X^\mu Q_{\mathcal{T}\cup\mathcal{N}}$ and $Q^T_{\mathcal{B}\cup\mathcal{T}} S^\mu Q_{\mathcal{B}\cup\mathcal{T}}$.

Lemma 3.6. *Let $\left(X^\mu, y^\mu, S^\mu\right)$ be a central solution with $\mu \leq \hat{\mu}$, where $\hat{\mu}$ is defined in (16). Then we have*

$$\left\|Q^T_{\mathcal{T}\cup\mathcal{N}} X^\mu Q_{\mathcal{T}\cup\mathcal{N}}\right\| \leq c(n\mu)^\gamma, \quad \text{and} \quad \left\|Q^T_{\mathcal{B}\cup\mathcal{T}} S^\mu Q_{\mathcal{B}\cup\mathcal{T}}\right\| \leq c(n\mu)^\gamma.$$

Proof. From Lemma 3.5 and compactness of the optimal set it follows the existence of $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$ such that (17) holds. Recall from Theorem 2.2 that X_μ can be represented as $Q_\mathcal{B} U_{X_\mu} Q^T_\mathcal{B}$ where $U_{X_\mu} \succeq 0$. Thus, we have

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

and

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

which completes the proof. □

3.3. Approximation of the optimal partition

Let $X^\mu = Q^\mu \Lambda(X^\mu)(Q^\mu)^T$ and $S^\mu = Q^\mu \Lambda(S^\mu)(Q^\mu)^T$ be the eigenvalue decompositions of X^μ and S^μ, where Q^μ denotes a common eigenvector basis. We show in Theorems 3.8 and 3.9 that it is possible to identify the subsets of columns of Q^μ whose accumulation points are orthonormal bases for the subspaces $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$, when μ is sufficiently small. To do so, we need Lemma 3.7 to derive bounds on the eigenvalues of central solutions, which will be presented in Theorem 3.8.

Lemma 3.7 (Theorem 4.5 in Stewart and Sun (1990))**.** *Let $X \in \mathbb{S}^n$ and $Y \in \mathbb{R}^{n\times k}$.*

[3] An example was provided by Sturm (2000) which needs $n-1$ facial reduction steps, see Example 2 in Sturm (2000).

Then we have

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

$\square$

Theorem 3.8. *For a central solution (X^μ, y^μ, S^μ) with $\mu \le \hat{\mu}$, where $\hat{\mu}$ is given by (16), it holds that:*

$$\lambda_{[n-i+1]}(S^\mu) \le \frac{n\mu}{\sigma}, \qquad \lambda_{[i]}(X^\mu) \ge \frac{\sigma}{n}, \qquad i=1,\ldots,n_{\mathcal{B}}, \tag{18}$$

$$\lambda_{[n-i+1]}(X^\mu) \le \frac{n\mu}{\sigma}, \qquad \lambda_{[i]}(S^\mu) \ge \frac{\sigma}{n}, \qquad i=1,\ldots,n_{\mathcal{N}}, \tag{19}$$

$$\frac{\mu}{c\sqrt{n}(n\mu)^\gamma} \le \lambda_{[i]}(X^\mu), \lambda_{[n-i+1]}(S^\mu) \le c\sqrt{n}(n\mu)^\gamma, \quad i=n_{\mathcal{B}}+1,\ldots,n_{\mathcal{B}}+n_{\mathcal{T}}. \tag{20}$$

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

$$c \ge \frac{\left(\min\{\theta_1^{-1}, \theta_2^{-1}\}\right)^{\frac{1}{2}-\gamma}}{n}, \qquad \frac{1}{2^{n-1}} \le \gamma \le \frac{1}{2}.$$

Proof. Recall that $\hat{S}^\mu_{\mathcal{B}} = Q^T_{\mathcal{B}} S^\mu Q_{\mathcal{B}}$ and $\hat{X}^\mu_{\mathcal{N}} = Q^T_{\mathcal{N}} X^\mu Q_{\mathcal{N}}$ as defined in (4). Then it follows from Lemmas 3.3 and 3.7 that

$$\begin{aligned}\lambda_{[n-n_{\mathcal{B}}+1]}(S^\mu)+\ldots+\lambda_{[n]}(S^\mu) &\le \operatorname{trace}(\hat{S}^\mu_{\mathcal{B}}) \le \frac{n\mu}{\sigma},\\ \lambda_{[n-n_{\mathcal{N}}+1]}(X^\mu)+\ldots+\lambda_{[n]}(X^\mu) &\le \operatorname{trace}(\hat{X}^\mu_{\mathcal{N}}) \le \frac{n\mu}{\sigma},\end{aligned}$$

which, by $\lambda_{\min}(X^\mu), \lambda_{\min}(S^\mu) > 0$, give

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

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

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

In a similar manner, one can conclude from Lemmas 3.6 and 3.7 and $\operatorname{trace}(X) \le$

$\sqrt{n}\|X\|$ that

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

which, by the centrality condition, give

$$\lambda_{[n-i+1]}(X^\mu)\le c\sqrt{n}(n\mu)^\gamma,\qquad \lambda_{[i]}(S^\mu)\ge\frac{\mu}{c\sqrt{n}(n\mu)^\gamma},\qquad i=1,\ldots,n_{\mathcal{N}}+n_{\mathcal{T}},$$
$$\lambda_{[n-i+1]}(S^\mu)\le c\sqrt{n}(n\mu)^\gamma,\qquad \lambda_{[i]}(X^\mu)\ge\frac{\mu}{c\sqrt{n}(n\mu)^\gamma},\qquad i=1,\ldots,n_{\mathcal{B}}+n_{\mathcal{T}}.$$

This completes the first part of the proof.

By (20), if $n_{\mathcal{T}}>0$, there exist $n_{\mathcal{T}}$ eigenvalues of X^μ and $n_{\mathcal{T}}$ eigenvalues of S^μ which stay within the interval $\big[\mu/(c\sqrt{n}(n\mu)^\gamma),\ c\sqrt{n}(n\mu)^\gamma\big]$, and thus both converge to 0 as $\mu\to 0$. Then it holds that

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

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

$$c\ge\frac{\big(\min\{\theta_1^{-1},\theta_2^{-1}\}\big)^{\frac{1}{2}-\gamma}}{n},\qquad\gamma\le\frac{1}{2}.$$

The proof is complete. $\square$

Since the central path is an analytic curve, the eigenvalues of X^μ and S^μ are continuous functions of μ, and the eigenvalues of central solutions converge to the eigenvalues of the limit point of the central path. Hence, one can observe from Theorem 3.8 that as $\mu\to 0$, the eigenvalues of a central solution (X^μ,y^μ,S^μ) naturally separate into three subsets:

(1) $\lambda_{[i]}(X^\mu)$ converges to a positive value and $\lambda_{[n-i+1]}(S^\mu)$ converges to 0,
(2) both $\lambda_{[i]}(X^\mu)$ and $\lambda_{[n-i+1]}(S^\mu)$ converge to 0,
(3) $\lambda_{[i]}(S^\mu)$ converges to a positive value and $\lambda_{[n-i+1]}(X^\mu)$ converges to 0.

Recall that X^μ and S^μ have a common eigenvector basis Q^μ. If μ is sufficiently small, then Q^μ can be represented as

$$Q^\mu=(Q^\mu_{\mathcal{B}},Q^\mu_{\mathcal{T}},Q^\mu_{\mathcal{N}}),$$

where $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$ denote the subsets of columns of Q^μ corresponding to the above three subsets of eigenvalues, respectively. The following theorem shows how small μ should be in order to identify $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$ from Q^μ.

Theorem 3.9. *Let (X^μ, y^μ, S^μ) and a common orthogonal matrix Q^μ be given. If μ satisfies*

$$\mu < \tilde{\mu} := \min\left\{\frac{1}{n}\left(\frac{\sigma}{cn^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2},\ \hat{\mu}\right\}, \tag{21}$$

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

Proof. From inequalities (18) and (19), we can deduce that the $n_{\mathcal{B}}$ largest eigenvalues of X^μ stay positive while the $n_{\mathcal{B}}$ smallest eigenvalues of S^μ will converge to 0. Similarly, the $n_{\mathcal{N}}$ largest eigenvalues of S^μ will remain positive while the last $n_{\mathcal{N}}$ eigenvalues of X^μ converge to 0 as $\mu \to 0$. Inequalities (20) also hint that, if $n_{\mathcal{T}} > 0$, there should exist a set of $n_{\mathcal{T}}$ eigenvalues of X^μ and S^μ which stay within the interval $\left[\mu/(c\sqrt{n}(n\mu)^\gamma),\ c\sqrt{n}(n\mu)^\gamma\right]$. Recall that the i^{th} largest eigenvalue of X^μ and the i^{th} smallest eigenvalue of S^μ have the same eigenvector. Thus, if the intervals $\left[\mu/(c\sqrt{n}(n\mu)^\gamma),\ c\sqrt{n}(n\mu)^\gamma\right]$, $(0, n\mu/\sigma]$, and $[\sigma/n, \infty)$ are disjoint, i.e.,

$$\frac{n\mu}{\sigma} < \frac{\mu}{c\sqrt{n}(n\mu)^\gamma}, \qquad c\sqrt{n}(n\mu)^\gamma < \frac{\sigma}{n}, \qquad \frac{n\mu}{\sigma} < \frac{\sigma}{n}, \tag{22}$$

and that $\mu \leq \hat{\mu}$ holds, then we can identify $Q^\mu_{\mathcal{B}}$, $Q^\mu_{\mathcal{T}}$, and $Q^\mu_{\mathcal{N}}$. Consequently, when $\mu < \tilde{\mu}$, we can identify the vanishing eigenvalues of X^μ and S^μ by comparing the magnitude of the eigenvalues to the lower and upper bounds given in (18) to (20). This completes the proof. □

Since the central path converges to a maximally complementary optimal solution, for a given sequence $\{\mu_k\}$ the accumulation points of $Q^{\mu_k}_{\mathcal{B}}$, $Q^{\mu_k}_{\mathcal{T}}$, and $Q^{\mu_k}_{\mathcal{N}}$ form orthonormal bases for the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$, respectively, see Section 3.3 in de Klerk (2006). For every $\mu < \tilde{\mu}$ we refer to $(\mathcal{R}(Q^\mu_{\mathcal{B}}), \mathcal{R}(Q^\mu_{\mathcal{T}}), \mathcal{R}(Q^\mu_{\mathcal{N}}))$ as an approximation of the optimal partition $(\mathcal{B}, \mathcal{T}, \mathcal{N})$.

Remark 7. In general, we do not know in advance if the strict complementarity condition holds for a given instance of SDO. If $n_{\mathcal{T}} > 0$, then (22) implies that

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

In case that $n_{\mathcal{T}} = 0$, we can improve the bound (21). In fact, the bounds in (20) may provide no further information compared to (18) and (19) for small values of μ. Hence, in order to identify $Q^\mu_{\mathcal{B}}$ and $Q^\mu_{\mathcal{N}}$ it is enough to have

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

which reduces the bound (21) to $\mu < \sigma^2/n^2$. This bound matches the one for LO, see Section 3.3.3 in Roos et al. (2005).

3.4. Proximity to the true optimal partition

We can provide more information about the optimal partition of the problem by measuring the proximity of $\mathcal{R}(Q^\mu_{\mathcal{B}})$ and $\mathcal{R}(Q^\mu_{\mathcal{N}})$ to the subspaces $\mathcal{B}$ and $\mathcal{N}$, respectively, for $\mu < \tilde{\mu}$. To that end, we use the approach in Cheung et al. (2013) which measures the distance between a primal optimal solution $X \in \mathcal{P}^*$ and its projection onto $Q^{\mu}_{\mathcal{B}\cup\mathcal{T}}\mathbb{S}^{n_{\mathcal{B}}+n_{\mathcal{T}}}_{+}(Q^{\mu}_{\mathcal{B}\cup\mathcal{T}})^T$, which is a face of the positive semidefinite cone[4]. In fact, $\mathcal{P}^*$ is contained in the minimal face $Q_{\mathcal{B}}\mathbb{S}^{n_{\mathcal{B}}}_{+}Q^T_{\mathcal{B}}$ which itself is a face of $Q_{\mathcal{B}\cup\mathcal{T}}\mathbb{S}^{n_{\mathcal{B}}+n_{\mathcal{T}}}_{+}Q^T_{\mathcal{B}\cup\mathcal{T}}$. Analogously, we measure the distance between $\tilde{S}$, where $(\tilde{y}, \tilde{S}) \in \mathcal{D}^*$, and its projection onto $Q^{\mu}_{\mathcal{T}\cup\mathcal{N}}\mathbb{S}^{n_{\mathcal{T}}+n_{\mathcal{N}}}_{+}(Q^{\mu}_{\mathcal{T}\cup\mathcal{N}})^T$.

The following technical lemma is in order.

Lemma 3.10. *Let $(X^{\mu}, y^{\mu}, S^{\mu})$ with $\mu \le \hat{\mu}$ be given, and assume that $\mathcal{P}^*, \mathcal{D}^* \neq \{0\}$. Then we have*

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

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

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

Therefore, we get

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

where the last inequality follows from Lemma 3.5. The proof for the second part follows analogously. □

Theorem 3.11. *Let $(X^{\mu}, y^{\mu}, S^{\mu})$ be given so that $\mu < \tilde{\mu}$, where $\tilde{\mu}$ is defined in (21). Then for all $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ we have*

$$\|\tilde{X} - \tilde{X}_{\mathcal{F}_{\mathcal{B}\mathcal{T}}}\| \le \sqrt{2}\|\tilde{X}\|\sqrt{\frac{cn(n\mu)^{\gamma}}{\sigma}}, \tag{23}$$
$$\|\tilde{S} - \tilde{S}_{\mathcal{F}_{\mathcal{T}\mathcal{N}}}\| \le \sqrt{2}\|\tilde{S}\|\sqrt{\frac{cn(n\mu)^{\gamma}}{\sigma}}, \tag{24}$$

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

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

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 (23) and (24) trivially hold. Note that the projection of $\tilde{X}$ onto the face $\mathcal{F}_{\mathcal{B}\mathcal{T}}$ is the optimal

[4] See Proposition 2.2.14 in Cheung (2013) for its proof.

solution to

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

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

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

Thus, it only remains to derive a lower bound on

$$\frac{\left\|(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \tilde{X} Q^\mu_{\mathcal{B}\cup\mathcal{T}}\right\|}{\|\tilde{X}\|}. \tag{25}$$

Let us define

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

Then we have

$$\begin{aligned}Q^\mu_{\mathcal{N}} \Lambda_{\mathcal{N}}(S^\mu)(Q^\mu_{\mathcal{N}})^T \bullet \tilde{X} &\le Q^\mu_{\mathcal{B}\cup\mathcal{T}} \Lambda_{\mathcal{B}\cup\mathcal{T}}(S^\mu)(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T \bullet \tilde{X} + Q^\mu_{\mathcal{N}} \Lambda_{\mathcal{N}}(S^\mu)(Q^\mu_{\mathcal{N}})^T \bullet \tilde{X} \\ &= S^\mu \bullet \tilde{X} \le c(n\mu)^\gamma \|\tilde{X}\|,\end{aligned} \tag{26}$$

where the last inequality follows from Lemma 3.10. Since (26) holds for any $\tilde{X} \in \mathcal{P}^*$, a lower bound on (25) is given by

$$\begin{aligned}\min \quad & \left\|(Q^\mu_{\mathcal{B}\cup\mathcal{T}})^T X Q^\mu_{\mathcal{B}\cup\mathcal{T}}\right\| \\ \text{s.t.} \quad & Q^\mu_{\mathcal{N}} \Lambda_{\mathcal{N}}(S^\mu)(Q^\mu_{\mathcal{N}})^T \bullet X \le c(n\mu)^\gamma, \\ & \|X\| = 1, \\ & X \succeq 0.\end{aligned} \tag{27}$$

Let $\check{X} := (Q^\mu)^T X Q^\mu$, where

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

Then auxiliary problem (27) is equivalent to

$$\begin{aligned} \min \quad & \|\check{X}_{\mathcal{B}\cup\mathcal{T}}\| \\ \text{s.t.} \quad & \Lambda_{\mathcal{N}}(S^\mu) \bullet \check{X}_{\mathcal{N}} \le c(n\mu)^\gamma, \\ & \|\check{X}_{\mathcal{B}\cup\mathcal{T}}\|^2 + \|\check{X}_{\mathcal{N}}\|^2 + 2\|\check{X}_{(\mathcal{B}\cup\mathcal{T})\mathcal{N}}\|^2 = 1, \\ & \check{X} \succeq 0. \end{aligned} \tag{28}$$

Since $\check{X} \succeq 0$, we can use the inequality[5] $\left\|\check{X}_{(\mathcal{B}\cup\mathcal{T})\mathcal{N}}\right\|^2 \le \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\|\left\|\check{X}_{\mathcal{N}}\right\|$ to derive a relaxation of (28) as

$$\begin{aligned} \min \quad & \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\| \\ \text{s.t.} \quad & \Lambda_{\mathcal{N}}(S^\mu) \bullet \check{X}_{\mathcal{N}} \le c(n\mu)^\gamma, \\ & \left\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\right\| + \left\|\check{X}_{\mathcal{N}}\right\| \ge 1, \\ & \check{X}_{\mathcal{B}\cup\mathcal{T}} \succeq 0, \\ & \check{X}_{\mathcal{N}} \succeq 0. \end{aligned} \tag{29}$$

Finally, from the constraints in (29) we get

$$\|\check{X}_{\mathcal{B}\cup\mathcal{T}}\| \ge 1 - \|\check{X}_{\mathcal{N}}\| \ge 1 - \frac{c(n\mu)^\gamma}{\lambda_{[n_{\mathcal{N}}]}(S^\mu)} \ge 1 - \frac{cn(n\mu)^\gamma}{\sigma} \tag{30}$$

$$> 1 - \frac{1}{\sqrt{n}} > 0, \tag{31}$$

in which (30) follows from (19) as well as

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

and (31) results from $\mu < \tilde{\mu}$. In a similar way as in Cheung et al. (2013), it can be shown that $1 - c(n\mu)^\gamma/\lambda_{[n_{\mathcal{N}}]}(S^\mu)$ is indeed the optimal value of (27). Consequently, we can conclude that

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

Analogously, we can prove that

$$\left\|\tilde{S} - \tilde{S}_{\mathcal{F}_{\mathcal{TN}}}\right\| \le \|\tilde{S}\|\sqrt{1 - \frac{\|(Q^\mu_{\mathcal{T}\cup\mathcal{N}})^T\tilde{S}Q^\mu_{\mathcal{T}\cup\mathcal{N}}\|^2}{\|\tilde{S}\|^2}} \le \sqrt{2}\|\tilde{S}\|\sqrt{\frac{cn(n\mu)^\gamma}{\sigma}},$$

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

which completes the proof. $\square$

Under the assumption of primal-dual uniqueness, we provide an upper bound on the distance between the subspaces $(\mathcal{B}, \mathcal{T}, \mathcal{N})$ and $(\mathcal{R}(Q^\mu_{\mathcal{B}}), \mathcal{R}(Q^\mu_{\mathcal{T}}), \mathcal{R}(Q^\mu_{\mathcal{N}}))$, which are of the same dimension if $\mu < \tilde{\mu}$. The distance between two subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$ of $\mathbb{R}^n$ with the same dimension, as defined e.g., in Section 2.5.3 in Golub and Van Loan (2013), is

$$\operatorname{dist}(\mathcal{L}_1, \mathcal{L}_2) := \| \operatorname{Proj}_{\mathcal{L}_1} - \operatorname{Proj}_{\mathcal{L}_2} \|_2,$$

where $\operatorname{Proj}_{\mathcal{L}_1}$ and $\operatorname{Proj}_{\mathcal{L}_2}$ denote the orthogonal projections onto $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively. By definition, $0 \leq \operatorname{dist}(\mathcal{L}_1, \mathcal{L}_2) \leq 1$.

We use an error bound result for analytic systems from Luo and Pang (1994) to derive upper bounds on $\operatorname{dist}(\mathcal{B}, \mathcal{R}(Q^\mu_{\mathcal{B}}))$, $\operatorname{dist}(\mathcal{T}, \mathcal{R}(Q^\mu_{\mathcal{T}}))$, and $\operatorname{dist}(\mathcal{N}, \mathcal{R}(Q^\mu_{\mathcal{N}}))$, see also Luo and Luo (1994).

Lemma 3.12 (Theorem 2.2 in Luo and Pang (1994))**.** *Let a solution set $\mathcal{C}$ be defined as*

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

in which g_j for $j = 1, \ldots, m_1$ and h_k for $k = 1, \ldots, m_2$ are analytic functions on an open set $\mathcal{X} \subseteq \mathbb{R}^n$. If $\mathcal{C} \neq \emptyset$, then for each compact set $\bar{\mathcal{X}} \subset \mathcal{X}$ there exist a positive condition number ρ and an exponent $\upsilon > 0$ such that

$$\operatorname{dist}(x, \mathcal{C}) \leq \rho \big(\| [g(x)]_+ \|_2 + \| h(x) \|_2 \big)^{\upsilon}, \qquad \forall x \in \bar{\mathcal{X}},$$

where

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

$\square$

Using eigenvalue decompositions of X and S, the optimality conditions (2) can be cast into an analytic system, since every function is polynomial, see e.g., Alizadeh et al. (1998). Let

$$\mathcal{C}^* := \Big\{ \big(\operatorname{vec}(\tilde{Q}); \operatorname{diag}(\Lambda(\tilde{X})); \tilde{y}; \operatorname{diag}(\Lambda(\tilde{S})) \big) \mid \tilde{Q}\Lambda(\tilde{X})\tilde{Q}^T \in \mathcal{P}^*,\ (\tilde{y}, \tilde{Q}\Lambda(\tilde{S})\tilde{Q}^T) \in \mathcal{D}^* \Big\},$$

where $\operatorname{vec}(.)$ is the concatenation of the columns of a matrix, and $\operatorname{diag}(.)$ denotes the vector of diagonal entries of a square matrix. Then the following theorem is in order.

Theorem 3.13. *Assume that the primal-dual optimal solution is unique, and let a central solution (X^μ, y^μ, S^μ) with $\mu < \tilde{\mu}$ be given, where $\tilde{\mu}$ is defined in (21). Then* *there exist $\rho > 0$ and $\upsilon > 0$ such that*

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

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

$$\begin{aligned}
\operatorname{dist}\big(\mathcal{B}, \mathcal{R}(Q^\mu_{\mathcal{B}})\big) &= \big\|Q^\mu_{\mathcal{B}}(Q^\mu_{\mathcal{B}})^T - Q_{\mathcal{B}}Q^T_{\mathcal{B}}\big\|_2\\
&= \big\|Q^\mu_{\mathcal{B}}(Q^\mu_{\mathcal{B}})^T - Q_{\mathcal{B}}Q^T_{\mathcal{B}} - Q^\mu_{\mathcal{B}}Q^T_{\mathcal{B}} + Q^\mu_{\mathcal{B}}Q^T_{\mathcal{B}}\big\|_2\\
&= \big\|Q^\mu_{\mathcal{B}}\big((Q^\mu_{\mathcal{B}})^T - Q^T_{\mathcal{B}}\big) + \big(Q^\mu_{\mathcal{B}} - Q_{\mathcal{B}}\big)Q^T_{\mathcal{B}}\big\|_2\\
&\le \big\|Q^\mu_{\mathcal{B}}\big\|_2\big\|Q^\mu_{\mathcal{B}} - Q_{\mathcal{B}}\big\|_2 + \big\|Q^\mu_{\mathcal{B}} - Q_{\mathcal{B}}\big\|_2\big\|Q_{\mathcal{B}}\big\|_2\\
&\le 2\big\|Q^\mu_{\mathcal{B}} - Q_{\mathcal{B}}\big\|_2,\\
&\le 2\big\|Q^\mu - Q^*\big\|_2,\\
&\le 2\big\|\big(\operatorname{vec}(Q^\mu - Q^*); \operatorname{diag}(\Lambda(X^\mu) - \Lambda(X^*)); y^\mu - y^*; \operatorname{diag}(\Lambda(S^\mu) - \Lambda(S^*))\big)\big\|_2
\end{aligned}$$

for every $\big(Q^*\Lambda(X^*)(Q^*)^T, y^*, Q^*\Lambda(S^*)(Q^*)^T\big) \in \operatorname{ri}(\mathcal{P}^* \times \mathcal{D}^*)$. Consequently,

$$\operatorname{dist}\big(\mathcal{B}, \mathcal{R}(Q^\mu_{\mathcal{B}})\big) \le 2\operatorname{dist}\Big(\big(\operatorname{vec}(Q^\mu); \operatorname{diag}(\Lambda(X^\mu)); y^\mu; \operatorname{diag}(\Lambda(S^\mu))\big),\ \mathcal{C}^*\Big), \tag{32}$$

where the inequality (32) follows from the uniqueness of the optimal solution, since in this case $\operatorname{ri}(\mathcal{P}^* \times \mathcal{D}^*) = \mathcal{P}^* \times \mathcal{D}^*$.

By the centrality condition, a central solution (X^μ, y^μ, S^μ) violates the constraints $\Lambda_{ii}(\tilde{X})\Lambda_{ii}(\tilde{S}) = 0$ by μ for $i = 1, \dots, n$. Due to the fact that the set of central solutions (X^μ, y^μ, S^μ) with $\mu < \tilde{\mu}$ is contained in a compact set, it follows from Lemma 3.12 that there exist $\rho, \upsilon > 0$ such that

$$\begin{aligned}
\operatorname{dist}\Big(\big(\operatorname{vec}(Q^\mu); \operatorname{diag}(\Lambda(X^\mu)); y^\mu; \operatorname{diag}(\Lambda(S^\mu))\big),\ \mathcal{C}^*\Big) &\le \rho\Big(\big\|\operatorname{diag}\big(\Lambda(X^\mu)\Lambda(S^\mu)\big)\big\|_2\Big)^\upsilon\\
&\le \rho(\sqrt{n}\mu)^\upsilon,
\end{aligned}$$

which completes the proof for $\mathcal{B}$. The proof for the subspaces $\mathcal{T}$ and $\mathcal{N}$ are analogous. $\square$

Remark 8. The condition number ρ and the exponent υ in Theorem 3.13 depend on the problem data and the size of the compact set which contains the set of central solutions for $\mu < \tilde{\mu}$. Unlike the exponent γ in Lemma 3.5, there is no known estimate for υ.

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

$$\big(\operatorname{vec}(\tilde{Q}_\mu); \operatorname{diag}(\Lambda(\tilde{X}_\mu)); \tilde{y}_\mu; \operatorname{diag}(\Lambda(\tilde{S}_\mu))\big) \in \mathcal{C}^*,$$

whose distance from $\big(\operatorname{vec}(Q^\mu); \operatorname{diag}(\Lambda(X^\mu)); y^\mu; \operatorname{diag}(\Lambda(S^\mu))\big)$ is minimal. The assumption of uniqueness in Theorem 3.13 can be released if there exists a sequence of

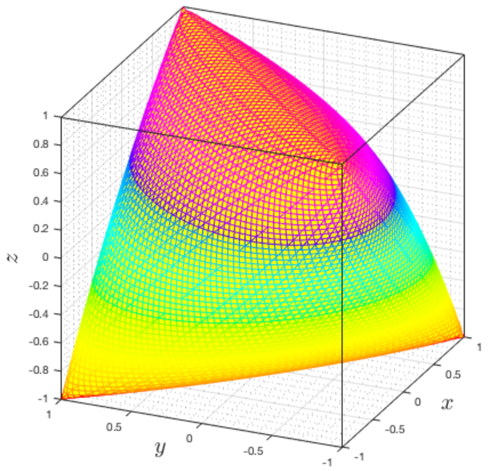

Figure 1.: The illustration of a 3-elliptope.

common eigenvector bases of maximally complementary optimal solutions converging to $\tilde{Q}_\mu$. 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 minimization of a linear objective over a 3-elliptope as illustrated by Figure 1:

$$\min\left\{ -2z \mid \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0, \quad (x,y,z) \in \mathbb{R}^3 \right\}. \tag{33}$$

The primal optimal set is given by

$$\tilde{X}(\delta) = \delta \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} + (1-\delta) \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \delta \in [0,1],$$

and the unique dual optimal solution is given by

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

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

$$\begin{aligned} \lambda_{[1]}(\tilde{X}(\delta)) &= \frac{1}{2}\sqrt{32\delta^2 - 32\delta + 9} + \frac{3}{2}, \\ \lambda_{[2]}(\tilde{X}(\delta)) &= -\frac{1}{2}\sqrt{32\delta^2 - 32\delta + 9} + \frac{3}{2}, \\ \lambda_{[3]}(\tilde{X}(\delta)) &= 0. \end{aligned}$$

Observe that for all $0 < \delta < 1$, $(\tilde{X}(\delta), \tilde{y}, \tilde{S})$ is strictly complementary, and that the multiplicity of the positive eigenvalues of $\tilde{X}(\delta)$ and $\tilde{S}$ are 1. Hence, for all $0 < \delta < 1$, the eigenvalue decompositions of $\tilde{X}(\delta)$ and $\tilde{S}$ are unique up to the sign of columns of the orthogonal matrices. At $\delta = 0$ the optimal solution is not strictly complementary, but $\tilde{X}(0)$ and $\tilde{S}$ have the unique common eigenvector basis

$$\tilde{Q}(0) = \begin{pmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix}.$$

Therefore, for a given $\delta_k \to 0$ there exists a sequence of strictly complementary optimal solutions $(\tilde{X}(\delta_k), \tilde{y}, \tilde{S})$ and a sequence $\tilde{Q}(\delta_k)$ such that $\tilde{Q}(\delta_k) \to \tilde{Q}(0)$.

Remark 9. We should note that unlike the SDO problem (33), there may be no sequence of common eigenvector bases with such a desired property. For example, if we change (33) to

$$\min\Bigg\{ -2z \mid \begin{pmatrix} 1 & x & y & 0 \\ x & 1 & z & 0 \\ y & z & 1 & 0 \\ 0 & 0 & 0 & 2-x-y \end{pmatrix} \succeq 0, \quad (x,y,z) \in \mathbb{R}^3 \Bigg\},$$

then the primal optimal set and the unique dual optimal solution are given as

$$\tilde{X}(\delta) = \delta \begin{pmatrix} 1 & -1 & -1 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} + (1-\delta) \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \delta \in [0,1],$$

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

For all $\delta \in (0,1) \setminus \{\frac{1}{2}\}$ the optimal solution $(\tilde{X}(\delta), \tilde{y}, \tilde{S})$ is strictly complementary and the positive eigenvalues of $\tilde{X}(\delta)$ and $\tilde{S}$ are distinct. However, $\tilde{X}(0)$ and $\tilde{S}$ have more than one common eigenvector basis.

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 extend the results of Theorems 3.8 and 3.9 by considering solutions in a neighborhood[6] of the central path, given by

$$\mathcal{N}_\kappa(\xi) := \Big\{ (X^\circ, y^\circ, S^\circ) \in \mathrm{ri}(\mathcal{P} \times \mathcal{D}) \mid \kappa(X^\circ S^\circ) \leq \xi \Big\}, \tag{34}$$

[6] See e.g., Section 6.4 in de Klerk (2006).

where

$$\kappa(X^\circ S^\circ) := \frac{\lambda_{\max}(X^\circ S^\circ)}{\lambda_{\min}(X^\circ S^\circ)}, \qquad (X^\circ, y^\circ, S^\circ) \in \mathrm{ri}(\mathcal{P} \times \mathcal{D}), \tag{35}$$

and $\xi > 1$. 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 (35) that $\kappa(X^\circ S^\circ) \geq 1$, and equality holds only when $(X^\circ, y^\circ, S^\circ)$ is on the central path. Then, for $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$ we have

$$\lambda_{\min}(X^\circ S^\circ) \leq \lambda_{[i]}(X^\circ S^\circ) \leq \xi \lambda_{\min}(X^\circ S^\circ), \qquad i = 1, \ldots, n. \tag{36}$$

We use the application of Weyl theorem[7] in Lu and Pearce (2000) to provide an upper bound on $\lambda_{\min}(X^\circ S^\circ)$, as presented in Lemma 4.2.

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 $j \leq \min\{\mathrm{rank}(X),\ \mathrm{rank}(S)\}$ we have*

$$\min_{1 \leq i \leq j} \{\lambda_{[i]}(X)\lambda_{[j-i+1]}(S)\} \geq \lambda_{[j]}(XS) \geq \max_{j \leq i \leq n} \{\lambda_{[i]}(X)\lambda_{[n+j-i]}(S)\}. \tag{37}$$

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

$$\lambda_{[i]}(X^\circ)\lambda_{[n-i+1]}(S^\circ) \geq \lambda_{\min}(X^\circ S^\circ), \qquad i = 1, \ldots, n. \tag{38}$$

Proof. The proof is straightforward from the first inequality in (37) and the positive definiteness of X° and S°. In fact, for the special case $k = n$ there holds that

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

which completes the proof. □

Consider a solution $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$, and let $X^\circ = M\Lambda(X^\circ)M^T$ and $S^\circ = P\Lambda(S^\circ)P^T$ be eigenvalue decompositions of X° and S°, respectively, where M and P are orthogonal matrices. Analogous to the case of central solutions, we let $M := (M_\mathcal{B}, M_\mathcal{T}, M_\mathcal{N})$ and $P := (P_\mathcal{B}, P_\mathcal{T}, P_\mathcal{N})$, respectively, partitioned according to the eigenvalues of X° and S° whose accumulation points are positive and zero. Since X° and S° do not necessarily commute, an accumulation point of M is not necessarily identical with an accumulation point of P.

The following theorem generalizes the bounds derived in Theorems 3.8 and 3.9 to an approximate solution $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$.

Theorem 4.3. *Let $(X^\circ, y^\circ, S^\circ) \in \mathcal{N}_\kappa(\xi)$ with $\mu \leq \hat{\mu}$, where $\mu := X^\circ \bullet S^\circ / n$ and $\hat{\mu}$ is defined in (16). Then there exist a positive condition number c' independent of μ and*

[7] See Theorem 4.3.7 in Horn and Johnson (2012).

a positive exponent γ *so that*

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

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

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

Furthermore, if μ *satisfies*

$$\mu < \min\left\{\frac{1}{n}\left(\frac{\sigma}{c' n^{\frac{3}{2}}\xi}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2\xi},\ \hat{\mu}\right\}, \tag{39}$$

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 Theorems 3.8 and 3.9 fairly easily. Let $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ which satisfies (8) 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

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

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

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

where the latter inequalities follow from (8). Now, Lemma 3.7 can be applied to get

$$\begin{aligned}
\lambda_{[n-n_{\mathcal{B}}+1]}(S^\circ) + \ldots + \lambda_{[n]}(S^\circ) \le \operatorname{trace}(\hat{S}^\circ_{\mathcal{B}}) \le \frac{n\mu}{\sigma},\\
\lambda_{[n-n_{\mathcal{N}}+1]}(X^\circ) + \ldots + \lambda_{[n]}(X^\circ) \le \operatorname{trace}(\hat{X}^\circ_{\mathcal{N}}) \le \frac{n\mu}{\sigma},
\end{aligned}$$

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

$$\begin{aligned}
\lambda_{[n-i+1]}(S^\circ) \le \frac{n\mu}{\sigma}, \quad i = 1,\ldots,n_{\mathcal{B}},\\
\lambda_{[n-i+1]}(X^\circ) \le \frac{n\mu}{\sigma}, \quad i = 1,\ldots,n_{\mathcal{N}}.
\end{aligned} \tag{40}$$

Recall from (36) that

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

which yields

$$\frac{\lambda_{\min}(X^\circ S^\circ)}{\mu} \geq \frac{1}{\xi}. \tag{41}$$

Then, using (38), (40), and (41), we can derive lower bounds on the eigenvalues of X° and S°:

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

We employ an analogue of Lemma 3.5 to derive bounds on $Q_{\mathcal{T}\cup\mathcal{N}}^T X^\circ Q_{\mathcal{T}\cup\mathcal{N}}$ and $Q_{\mathcal{B}\cup\mathcal{T}}^T S^\circ Q_{\mathcal{B}\cup\mathcal{T}}$. Note that the amount of constraint violation with respect to the LMI system (12) for $(X^\circ, y^\circ, S^\circ)$ is equal to $n\mu$, where $\mu = X^\circ \bullet S^\circ/n$. Then it is easy to verify that

$$\begin{aligned}
\operatorname{dist}\big(X^\circ, \tilde{X} + \operatorname{smat}\big(\operatorname{Ker}(\mathbf{A})\big) \cap (\mathbb{R}\tilde{S})^\perp\big) &\leq \theta_1 n\mu, \\
\operatorname{dist}\big(S^\circ, \tilde{S} + \operatorname{smat}\big(\mathcal{R}(\mathbf{A}^T)\big) \cap (\mathbb{R}\tilde{X})^\perp\big) &\leq \theta_2 n\mu,
\end{aligned}$$

where θ_1 and θ_2 are the same[8] Hoffman condition numbers defined in (14) and (15). Moreover,

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

is a bounded set by the interior point condition and the linear independence of A^i for $i = 1, \ldots, m$, see e.g., Lemma 3.1 in de Klerk (2006). Hence, for $0 < \mu \leq \hat{\mu}$ the result of Lemma 3.4 is still valid, i.e., 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

$$\big\|X^\circ - X_{\mu^\circ}\big\| \leq c'(n\mu)^\gamma, \qquad \big\|S^\circ - S_{\mu^\circ}\big\| \leq c'(n\mu)^\gamma, \tag{42}$$

where c' and γ are defined as in Lemma 3.5. Analogous to the proof of Theorem 3.8,

[8] Recall that a Hoffman condition number is only dependent on the left hand side of a linear system, see Güler (2010).

we can observe, using the orthogonal transformation Q, that

$$\begin{aligned}
\frac{1}{\sqrt{n}}\,\mathrm{trace}(Q_{\mathcal{T}\cup\mathcal{N}}^T X^\circ Q_{\mathcal{T}\cup\mathcal{N}}) &\le \left\|Q_{\mathcal{T}\cup\mathcal{N}}^T X^\circ Q_{\mathcal{T}\cup\mathcal{N}}\right\| \\
&= \left\| \begin{pmatrix} \hat{X}_{\mathcal{T}}^\circ & \hat{X}_{\mathcal{T}\mathcal{N}}^\circ \\ \hat{X}_{\mathcal{N}\mathcal{T}}^\circ & \hat{X}_{\mathcal{N}}^\circ \end{pmatrix} \right\| \le \left\|X^\circ - X_{\mu^\circ}\right\| \le c'(n\mu)^\gamma, \\
\frac{1}{\sqrt{n}}\,\mathrm{trace}(Q_{\mathcal{B}\cup\mathcal{T}}^T S^\circ Q_{\mathcal{B}\cup\mathcal{T}}) &\le \left\|Q_{\mathcal{B}\cup\mathcal{T}}^T S^\circ Q_{\mathcal{B}\cup\mathcal{T}}\right\| \\
&= \left\| \begin{pmatrix} \hat{S}_{\mathcal{B}}^\circ & \hat{S}_{\mathcal{B}\mathcal{T}}^\circ \\ \hat{S}_{\mathcal{T}\mathcal{B}}^\circ & \hat{S}_{\mathcal{T}}^\circ \end{pmatrix} \right\| \le \left\|S^\circ - S_{\mu^\circ}\right\| \le c'(n\mu)^\gamma.
\end{aligned} \tag{43}$$

Then it follows from Lemma 3.7 and (43) that

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

which, by the bounds in (38) and (41), yield

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

This completes the first part of the proof.

Finally, using the same argument as in Theorem 3.9, 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

$$\frac{n\mu}{\sigma} < \frac{\mu}{c'\sqrt{n}\xi(n\mu)^\gamma}, \qquad c'\sqrt{n}(n\mu)^\gamma < \frac{\sigma}{n\xi}, \quad \text{and} \quad \frac{n\mu}{\sigma} < \frac{\sigma}{n\xi}, \tag{44}$$

which give (39). If $n_{\mathcal{T}} > 0$, then from (44) it is immediate that

$$(c')^2 n^2 \xi \ge (n\mu)^{1-2\gamma}, \qquad \forall\ 0 < \mu \le \hat{\mu},$$

which implies

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

This completes the proof. $\square$

Corollary 4.4. *Let $(X^{(0)}, y^{(0)}, S^{(0)}) \in \mathcal{N}_\kappa(\xi)$ be an initial solution, $\mu^{(0)} := X^{(0)} \bullet S^{(0)}/n$, and $\log(.)$ denote the natural logarithm. Then the Dikin-type primal-dual affine scaling method with steplength $\alpha = 1/(\xi\sqrt{n})$ and the neighborhood (34) (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 $\alpha = 1/(\xi\sqrt{n})$, see Theorem 6.1 in 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 (39) multiplied by n. $\square$

Analogues of Theorems 3.11 and 3.13 can be presented for interior solutions in a neighborhood of the central path. To that end, we only need to replace $(Q^\mu_\mathcal{B}, Q^\mu_\mathcal{T}, Q^\mu_\mathcal{N})$ by $(M_\mathcal{B}, M_\mathcal{T}, M_\mathcal{N})$ or $(P_\mathcal{B}, P_\mathcal{T}, P_\mathcal{N})$ to define the projection of $\tilde{X}$ and $\tilde{S}$ for any $(\tilde{X}, \tilde{y}, \tilde{S}) \in \mathcal{P}^* \times \mathcal{D}^*$. Analogous to Theorem 3.13, we can derive an upper bound on the proximity of $\big(\mathcal{R}(M_\mathcal{B}), \mathcal{R}(M_\mathcal{T}), \mathcal{R}(M_\mathcal{N})\big)$, or $\big(\mathcal{R}(P_\mathcal{B}), \mathcal{R}(P_\mathcal{T}), \mathcal{R}(P_\mathcal{N})\big)$, to $(\mathcal{B}, \mathcal{T}, \mathcal{N})$. For the sake of brevity, we do not present the details here.

Remark 10. In (42), we employed the same exponent γ as in (17) but a different condition number c'. In fact, the primal and dual systems in (12) are used for both Theorems 3.8 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 (5) and the upper bounds in (17), 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 on 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 on, or in a neighborhood of the central path. 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' is 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, Minneapolis, MN, USA.

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.

Bonnans JF, Shapiro A. 2000. Perturbation Analysis of Optimization Problems. Springer, New York, NY, USA.

Cheung YL. 2013. Preprocessing and reduction for semidefinite programming via facial reduction: Theory and practice [dissertation]. Department of Combinatorics & Optimization, University of Waterloo, Waterloo, ON, Canada.

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, New York, NY, USA; p. 613–634.

de Klerk E. 2006. Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. vol. 65. Springer, New York, NY, USA.

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. In: Pardalos P, Wolkowicz H, editors. Topics in Semidefinite and Interior Point Methods. vol. 18. Fields Institute Communications, AMS, Providence, RI; p. 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, Baltimore, MD, USA.

Güler O. 2010. Foundations of Optimization. Springer, New York, NY, USA.

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, Cambridge, UK.

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, Pang JS. 1994. Error bounds for analytic systems and their applications. Mathematical Programming. 67(1):1–28.

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, Philadelphia, PA, USA.

Nocedal J, Wright S. 2006. Numerical Optimization. 2nd ed. Springer, New York, NY, USA.

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, Baltimore, MD, USA.

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, New York, NY, USA.

Stewart GW, Sun Jg. 1990. Matrix Perturbation Theory. Academic Press, San Diego, CA, USA.

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.

Vandenberghe 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 (5). To do so, we resort to a technical lemma in Ramana (1993).

An integral polynomial map $f : \mathbb{R}^s \to \mathbb{R}^t$ is defined as a map consisting of polynomial functions f^i of degree d_i with integer coefficients. We consider a solution set $\mathcal{V}(f)$ defined as

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

where Δ_i stands for one of the relations $\{>, =, \geq\}$. Depending on the polynomial map f, the solution set $\mathcal{V}(f)$ 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 $\log_2(.)$ 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 $\mathcal{V}(f)$.

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 $\mathcal{V}(f)$ 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 and 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*

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

where

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

Proof. Recall from (6) and (7) that

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

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

$$\begin{cases} A^i \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T & = b_i, \qquad i = 1, \ldots, m, \\ C - \displaystyle\sum_{i=1}^{m} y_i A^i & = Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T, \\ U_X, U_S \succeq 0, \end{cases} \tag{A2}$$

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.,

$$\begin{cases} A^i \bullet Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T & = b_i, \qquad i = 1, \ldots, m, \\ C - \displaystyle\sum_{i=1}^{m} y_i A^i & = Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T, \\ U_X, U_S \succ 0. \end{cases} \tag{A3}$$

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

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

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:

$$\begin{cases} U_{X^a}^{-1} - \sum_{i=1}^{m} u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}} & = 0, \\ A^i \bullet Q_{\mathcal{B}} U_{X^a} Q_{\mathcal{B}}^T & = b_i, \qquad i = 1, \ldots, m, \\ U_{X^a} \succ 0. \end{cases} \tag{A5}$$

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

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

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

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

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

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

as defined below

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

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 $\mathcal{U}_p$ to enforce the positive definiteness of

U_{X^a} and U_S as follows

$$\mathcal{U}_p := \Big\{\vartheta_p \mid \det(U_{X^a}[i]) > 0, \quad \det(U_S[j]) > 0, \quad i = 1, \dots, n_{\mathcal{B}}, \ j = 1, \dots, n_{\mathcal{N}}\Big\}, \quad \text{(A9)}$$

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 $\mathcal{V}(f_p) \cap \mathcal{U}_p$, where $\mathcal{V}(f_p) = \big\{\vartheta_p \mid f_p(\vartheta_p) = 0\big\}$, is nonempty but not necessarily a singleton. Then, from every solution $\vartheta_p \in \mathcal{V}(f_p) \cap \mathcal{U}_p$, we can extract a solution $(U_{X^a}, u, Q_{\mathcal{B}})$ which is the analytic center of the primal optimal set, since it satisfies the constraints in (A5).

The solution set $\mathcal{U}_p$ is characterized by $n_{\mathcal{B}} + n_{\mathcal{N}}$ integer polynomials of degree at most $\max\{n_{\mathcal{B}}, n_{\mathcal{N}}\}$. Since the symmetry of the matrices U_{X^a}, Z_X, and U_S is not presumed for f_p and $\mathcal{U}_p$, the coefficients of the polynomial functions are bounded above by twice the largest absolute value of the entries in b, C, and A^i for $i = 1, \dots, m$. For instance, the coefficients of $\det(U_{X^a}[i])$ are just 1, but $u_i Q_{\mathcal{B}}^T A^i Q_{\mathcal{B}}$ has some polynomial terms with coefficients twice the off-diagonal entries of A^i. Hence, the binary length of the largest absolute value of the coefficients in (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 $\mathcal{V}(f_p) \cap \mathcal{U}_p$, we can conclude that there exists a solution $\vartheta_p \in \mathcal{V}(f_p) \cap \mathcal{U}_p$ so that $\|\vartheta_p\|_2 \le r_{\mathcal{P}^*}$, where

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

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 $\mathcal{U}_p$. As a result, there exists u so that $|u_i| \le \|u\|_2 \le r_{\mathcal{P}^*}$. Then, using inequality (A6), we get

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

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 $\sigma_{\mathcal{N}}$. Notice that for a given orthonormal basis $Q_{\mathcal{N}}$, the analytic center of the dual optimal set can be obtained by solving

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

which is a convex optimization problem with strictly concave objective function. The optimality conditions for (A10) are given by

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

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 symmetry[9] of W. Then the optimality conditions (A11) imply

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

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

$$\mathcal{V}(f_d) := \Big\{\vartheta_d \mid f_d(\vartheta_d) = 0\Big\},$$

where the integral polynomial map

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

is defined as

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

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 $\mathcal{V}(f_d) \cap \mathcal{U}_d$ is nonempty, where $\mathcal{U}_d$ is defined as

$$\mathcal{U}_d := \Big\{\vartheta_d \mid \det(U_X[i]) > 0,\ \det(U_{S^a}[j]) > 0, \quad i = 1, \ldots, n_{\mathcal{B}},\ j = 1, \ldots, n_{\mathcal{N}}\Big\}.$$

[9]Note that there is no need to add a symmetrization constraint. One can easily check that $\frac{W+W^T}{2}$ is a symmetric feasible solution for (A11).

Then, analogous to the primal case, from a solution $\vartheta_d \in \mathcal{V}(f_d) \cap \mathcal{U}_d$ 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 $\vartheta_d \in \mathcal{V}(f_d) \cap \mathcal{U}_d$ so that $\|\vartheta_d\|_2 \le r_{\mathcal{D}^*}$, where

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

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

Remark 11. For the special case $n_{\mathcal{B}} = 0$ we get $\sigma = \sigma_{\mathcal{N}}$ by (5), 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.,

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

It it 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 $\sigma \le 20 \times 4^{-2^{\bar{m}}}$. Given $n_{\mathcal{B}} \le 2\bar{m}+1$, $n_{\mathcal{N}} \le 2\bar{m}+1$, $\|A^1\| = \sqrt{\bar{m}+8}$, $\|A^{i+1}\| = 3$ for $i = 1, \ldots, \bar{m}-1$, $\|A^{\bar{m}+1}\| = \sqrt{2}$, and $L = l(2) = 1 + \lceil \log_2(3) \rceil = 3$, we can compute the lower bound (A1). To do so, we have

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

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,

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