A rounding procedure for semidefinite optimization

Ali Mohammad-Nezhad and Tamás Terlaky

Department of Industrial and Systems Engineering, Lehigh University, USA

ISE Technical Report 17T-009

A rounding procedure for semidefinite optimization

Ali Mohammad-Nezhad[1] and Tamás Terlaky[1]

[1]Department of Industrial and Systems Engineering, Lehigh University, USA

July 02, 2017

Abstract

In this paper, we review the concept of the optimal partition and its identification for semidefinite optimization. In contrast to linear optimization and linear complementarity problem, it is impossible to identify the optimal partition of semidefinite optimization exactly. Instead, the sets of eigenvectors converging to an orthonormal bases for the optimal partition can be identified from an interior solution, which is either a central solution, or a solution in a neighborhood of the central path. Using these sets of eigenvectors, we propose a rounding procedure to generate an approximate maximally complementary solution of semidefinite optimization. The procedure generates a rounded primal-dual solution from an interior solution by solving two least squares problems. We show that if the complementarity gap drops below a certain bound, then the rounded primal-dual solution satisfies the cone constraints.

1 Introduction

The optimal partition for linear optimization (LO) and linear complementarity problems (LCPs) is quite well-known in the literature of interior point methods (IPMs). From the complementarity condition for LO, which reduces to $x_j s_j = 0$ for $j = 1, \dots, n$, it can be seen that for every optimal solution (x, y, s), for all $j = 1, \dots, n$, either x_j or s_j has to be zero. Further, there always exists a strictly complementary optimal solution, i.e., an optimal solution with $x_j + s_j > 0$ for every $j = 1, \dots, n$. Then, for LO, the optimal partition is defined as the two disjoint sets B and N:

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

where $B \cup N = \{1, \dots, n\}$. We can use the optimal partition to characterize the primal and dual optimal sets, $\mathcal{P}^*_{lo}$ and $\mathcal{D}^*_{lo}$, respectively, as

$$\begin{aligned} \mathcal{P}^*_{lo} &:= \{ x \mid Ax = b, \ x_j \geq 0, \ \forall j \in B, \ \ x_j = 0, \ \ \forall j \in N \}, \\ \mathcal{D}^*_{lo} &:= \{ (y, s) \mid A^T y + s = c, \ s_j \geq 0, \ \forall j \in N, \ \ s_j = 0, \ \ \forall j \in B \}. \end{aligned}$$

When dealing with LCPs, there might exist a third set T defined as

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

Then we have both $x_j = 0$ and $s_j = 0$ for all $j \in T$ for every optimal solution of the LCP. Goldfarb and Scheinberg [4] extended the concept of the optimal partition to semidefinite optimization (SDO), and Yildirim [17] derived a facial description of the optimal partition for linear conic optimization with self-dual cones. Recently, Mohammad-Nezhad and Terlaky [8] have investigated the identification of the optimal partition for SDO. The optimal partition for second-order conic optimization (SOCO) and its identification was studied by Bonnans and Ramírez [2] and Terlaky and Wang [14], respectively. Peña and Roshchina [11] established the complementarity partition for a multifold homogeneous conic system. The optimal partition information can be used in a so called rounding procedure to generate either a maximally or strictly complementary optimal solution, see e.g., Ye [16] and Roos, Terlaky, and Vial [12] for LO, and Illés, Peng, Roos, and Terlaky [7] and the references in there for LCP.

In this paper, we propose a rounding procedure for the identification of an approximate maximally complementary solution, which uses the sets of eigenvectors converging to an orthonormal bases for the optimal partition. The rounding procedure generates a primal-dual solution with zero complementarity gap by solving two least squares problems. We show that if the complementarity gap is less than a certain bound, then the primal-dual solution satisfies the cone constraints. The rest of this paper is built as follows. In Section 2, we review the concept of the optimal partition and its identification for SDO. In Section 3, we present a rounding procedure to generate an approximate maximally complementary solution and provide feasibility bounds for the primal and dual solutions. Section 4 presents the same approach but for solutions in a neighborhood of the central path. Our concluding remarks are stated in Section 5.

2 Preliminaries

An SDO problem in standard form is written as

$$(P) \qquad p^* := \min\left\{\langle C, X\rangle \mid \langle A^i, X\rangle = b_i, \ \ i = 1, \dots, m, \ X \succeq 0\right\},$$

where $C, X, A_i \in \mathbb{S}^n$ for $i = 1, \dots, m$, $b \in \mathbb{R}^m$, $\mathbb{S}^n$ denotes the linear space of $n \times n$ symmetric matrices, and the inner product is defined as $\langle C, X\rangle := \operatorname{tr}(CX)$. The dual SDO problem is given by

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

Let $\mathcal{P}$ and $\mathcal{D}$ denote the primal and dual feasible sets, and $\mathcal{P}^*$ and $\mathcal{D}^*$ be the primal and dual optimal sets, respectively. It is assumed that the matrices A^i for $i = 1, \dots, m$ are linearly independent, and the interior point condition holds, i.e., there exists $(X, y, S) \in \mathcal{P} \times \mathcal{D}$ with $X, S \succ 0$. The interior point condition ensures that the primal and dual problems are solvable, and at optimality there is no duality gap. Therefore, for every primal-dual optimal solution, the complementarity condition $XS = 0$ holds. The primal-dual optimal solution (X^*, y^*, S^*) is maximally complementary if $\operatorname{rank}(X^* + S^*)$ is maximal over the optimal set, or equivalently, if $(X^*, y^*, S^*) \in \operatorname{ri}(\mathcal{P}^* \times \mathcal{D}^*)$, where $\operatorname{ri}(.)$ denotes the relative interior of a set. A maximally complementary optimal solution (X^*, y^*, S^*) is strictly complementary if $X^* + S^* \succ 0$. By the definition of a maximally complementary optimal solution, all $X^* \in \operatorname{ri}(\mathcal{P}^*)$ have the same range space, see e.g., Lemma 2.3 in [3] or Lemma 3.1 in [4]. The same is true for S^* with $(y^*, S^*) \in \operatorname{ri}(\mathcal{D}^*)$.

Let $(X^*, y^*, S^*) \in \operatorname{ri}(\mathcal{P}^* \times \mathcal{D}^*)$ be a maximally complementary optimal solution and $\mathcal{R}(.)$ denote the range space. Since X^* and S^* commute by the complementarity condition, they are simultaneously diagonalizable, i.e.,

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

where Q^* is the common orthogonal matrix, and $\Lambda(X^*)$ and $\Lambda(S^*)$ are diagonal matrices of 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 indicates that $\mathcal{R}(X^*)$ and $\mathcal{R}(S^*)$ are equivalent to the span of the eigenvectors associated with the positive eigenvalues of X^* and S^*, respectively. The subspaces $\mathcal{R}(X^*)$ and $\mathcal{R}(S^*)$ are orthogonal by the complementarity condition. If there exists a strictly complementary optimal solution, then the subspaces $\mathcal{R}(X^*)$ and $\mathcal{R}(S^*)$ span $\mathbb{R}^n$. Otherwise, the orthogonal complement of $\mathcal{R}(X^*) + \mathcal{R}(S^*)$, which we call $\mathcal{T}$, is nonzero. Let $\mathcal{B} := \mathcal{R}(X^*)$ and $\mathcal{N} := \mathcal{R}(S^*)$. The partition $(\mathcal{B}, \mathcal{N}, \mathcal{T})$ of $\mathbb{R}^n$ is called the optimal partition of an SDO problem. Since $(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^* \times \mathcal{D}^*)$, the optimal partition is invariant with respect to the choice of (X^*, y^*, S^*).

Let $Q := [Q_\mathcal{B}, Q_\mathcal{N}, Q_\mathcal{T}]$ be an orthonormal bases for the optimal partition. For instance, the eigenvectors associated with the positive eigenvalues of X^* and S^* can be chosen as an orthonormal bases for $\mathcal{B}$ and $\mathcal{N}$, respectively. The following theorem characterizes the primal and dual optimal sets of an SDO problem. For the ease of exposition, we define $n_\mathcal{B} := \dim(\mathcal{B})$, $n_\mathcal{N} := \dim(\mathcal{N})$, and $n_\mathcal{T} := \dim(\mathcal{T})$.

Theorem 1 (Theorem 2.7 in [3])**.** *The primal and dual optimal solutions of SDO can be represented by*

$$X = Q_\mathcal{B} U_X Q_\mathcal{B}^T, \qquad S = Q_\mathcal{N} U_S Q_\mathcal{N}^T,$$

where $U_X \in \mathbb{S}_+^{n_\mathcal{B}}$ and $U_S \in \mathbb{S}_+^{n_\mathcal{N}}$, in which $\mathbb{S}_+^{n_\mathcal{B}}$ (and similarly $\mathbb{S}_+^{n_\mathcal{N}}$) denotes the cone of $n_\mathcal{B} \times n_\mathcal{B}$ positive semidefinite matrices. If $n_\mathcal{B} > 0$ and $X^* \in \mathrm{ri}(\mathcal{P}^*)$, then $U_{X^*} \succ 0$. Analogously, if $n_\mathcal{N} > 0$ and $(y^*, S^*) \in \mathrm{ri}(\mathcal{D}^*)$, then $U_{S^*} \succ 0$.

As a consequence of Theorem 1 we have

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

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

Now, we investigate the identification of the optimal partition for SDO along the central path, see e.g., [1] and [9]. For an SDO problem the central path is defined as the set of solutions to

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

where $XS = \mu I_n$ is called the centrality condition, and I_n is the identity matrix of size n. For a given $\mu > 0$, a unique solution $(X(\mu), y(\mu), S(\mu))$ exists, so called a central solution, under the linear independence of A^i for $i = 1, \ldots, m$ and the interior point condition. Halická et al. [5] proved that the central path converges to a maximally complementary optimal solution.

We define some condition numbers as

$$\begin{aligned} \sigma_\mathcal{B} &:= \begin{cases} \max_{X\in\mathcal{P}^*} \lambda_{\min}(Q_\mathcal{B}^T X Q_\mathcal{B}), & \mathcal{B} \neq \{0\}, \\ \infty, & \mathcal{B} = \{0\}, \end{cases} \\ \sigma_\mathcal{N} &:= \begin{cases} \max_{(y,S)\in\mathcal{D}^*} \lambda_{\min}(Q_\mathcal{N}^T S Q_\mathcal{N}), & \mathcal{N} \neq \{0\}, \\ \infty, & \mathcal{N} = \{0\}, \end{cases} \\ \sigma &:= \min\{\sigma_\mathcal{B}, \sigma_\mathcal{N}\}. \end{aligned}$$

It is proved in Lemma 3 in [8] that the condition number σ is positive. In Appendix A in [8], a positive lower bound is provided for σ.

Consider an orthogonal transformation of $(X(\mu), y(\mu), S(\mu))$ with respect to Q denoted by

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

where $\hat{X}(\mu) := Q^T X(\mu) Q$ and $\hat{S}(\mu) := Q^T S(\mu) Q$. As $\mu \to 0$ we have

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

To derive upper bounds for the vanishing blocks of $\hat{X}(\mu)$ and $\hat{S}(\mu)$, we resort to Theorem 3.3 in [13] which provides a Hölderian error bound, see e.g., [15], for the linear matrix inequality (LMI) system obtained from the optimal set. The Hölderian bound depends on the degree of singularity of the LMI system. In simple words, the degree of singularity [13] is defined as the number of facial reduction steps to recover the minimal face of the positive semidefinite cone which contains the optimal set. See [10] for a simple derivation of the facial reduction algorithm.

Let $\gamma = 2^{-d_s}$, in which d_s denotes the degree of singularity of the minimal subspace containing the optimal set, and c is a positive condition number. In the following theorem, we use the error bound result and the properties of the central solutions to provide bounds for the eigenvalues of the central solutions. We can assume w.l.o.g. that both $n_{\mathcal{B}}$ and $n_{\mathcal{N}}$ are positive. In our notation, $\lambda_{[i]}(X)$ denotes the i^{th} largest eigenvalue of X so that

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

Theorem 2 (Theorem 7 in [8])**.** *For a central solution $(X(\mu), y(\mu), S(\mu))$ with $n\mu \leq 1$ we have*

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

$$\lambda_{[i]}(S(\mu)) \geq \frac{\sigma}{n}, \qquad \lambda_{[n-i+1]}(X(\mu)) \leq \frac{n\mu}{\sigma}, \qquad i = 1, \ldots, n_{\mathcal{N}}. \tag{2}$$

Furthermore, we have

$$\begin{aligned} \lambda_{[n-i+1]}(X(\mu)) &\leq c\sqrt{n}(n\mu)^{\gamma}, \quad \lambda_{[i]}(S(\mu)) \geq \frac{\mu}{c\sqrt{n}(n\mu)^{\gamma}}, \; i = 1, \ldots, n_{\mathcal{N}} + n_{\mathcal{T}}, \\ \lambda_{[n-i+1]}(S(\mu)) &\leq c\sqrt{n}(n\mu)^{\gamma}, \quad \lambda_{[i]}(X(\mu)) \geq \frac{\mu}{c\sqrt{n}(n\mu)^{\gamma}}, \; i = 1, \ldots, n_{\mathcal{B}} + n_{\mathcal{T}}. \end{aligned} \tag{3}$$

If $n_{\mathcal{T}} > 0$, then we have $c \geq 1$, and

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

If μ is small enough so that

$$\mu < \min\left\{\frac{1}{n}\left(\frac{\sigma}{cn^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}}, \; \frac{\sigma^2}{n^2}, \; \frac{1}{n}\right\}, \tag{4}$$

then we can identify the sets of eigenvectors which converge to an orthonormal bases for $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$.

We can observe from Theorem 2 that the eigenvectors of a central solution $(X(\mu), y(\mu), S(\mu))$ fall into three categories:

- $q_i(\mu)$ for which $\lambda_i(X(\mu))$ converges to a positive value and $\lambda_i(S(\mu))$ converges to 0;
- $q_i(\mu)$ for which $\lambda_i(S(\mu))$ converges to a positive value and $\lambda_i(X(\mu))$ converges to 0;
- $q_i(\mu)$ for which both $\lambda_i(X(\mu))$ and $\lambda_i(S(\mu))$ converge to 0,

where $\lambda_i(X(\mu))$ and $\lambda_i(S(\mu))$ are associated with the eigenvector $q_i(\mu)$. As $\mu \to 0$, the above sets of eigenvectors converge to an orthonormal bases for $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$ (not necessarily to the Q chosen in Section 2), see Section 3.3 in [3].

The sets of eigenvectors converging to an orthonormal bases for the optimal partition can be identified for a solution in a neighborhood of the central path. However, note that since X and S are not necessarily simultaneously diagonalizable, the sets of eigenvectors derived from X and S are not identical in contrast to the case for a central solution, see Section 4. It is immediate from the similarity of XS and $X^{\frac{1}{2}}SX^{\frac{1}{2}}$ that XS has real positive eigenvalues. Thus, the proximity of (X, y, S) to the central path can be measured (i.e., Section 6.4 in [3]) by

$$\kappa(XS) := \frac{\lambda_{\max}(XS)}{\lambda_{\min}(XS)}, \quad (X, y, S) \in \mathrm{ri}(\mathcal{P} \times \mathcal{D}),$$

where $\kappa(XS) = 1$ iff (X, y, S) is on the central path. Then a neighborhood of the central path is given by

$$\mathcal{N}_\kappa(\tau) = \Big\{(X, y, S) \in \mathrm{ri}(\mathcal{P} \times \mathcal{D}) \mid \kappa(XS) \leq \tau\Big\},$$

in which $\tau > 1$. Let $\tau_2 := \tau\tau_1$ for some $\tau_1 > 0$ so that $\tau_2 \geq 1 \geq \tau_1 > 0$. Then for $(X, y, S) \in \mathcal{N}_\kappa(\tau)$ we have

$$\tau_1\lambda_{\min}(XS) \leq \lambda_{[i]}(XS) \leq \tau_2\lambda_{\min}(XS), \quad i = 1, \ldots, n.$$

Using the error bound result and the application of Weyl theorem[1], see e.g., Theorem 4.3.7 in [6], the following theorem generalizes the bounds given in Theorem 2 to an approximate solution $(X, y, S) \in \mathcal{N}_\kappa(\tau)$.

Theorem 3 (Theorem 11 in [8])**.** *Let $(X, y, S) \in \mathcal{N}_\kappa(\tau)$ with $n\mu \leq 1$, where $\mu := \frac{\langle X,S\rangle}{n}$. Then we have*

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

Further, we have

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

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

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

If μ is sufficiently small so that

$$\mu < \min\left\{\frac{1}{n}\left(\frac{\sigma}{cn^{\frac{3}{2}}\tau}\right)^{\frac{1}{\gamma}}, \frac{\sigma^2}{n^2\tau}, \ \frac{1}{n}\right\},$$

then we can identify the sets of eigenvectors which converge to an orthonormal bases for $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$.

3 A rounding procedure for central solutions

Since we do not have the exact orthonormal bases for $\mathcal{B}$, $\mathcal{N}$ and $\mathcal{T}$, an exact solution of an SDO cannot be obtained from a given central solution even if it is very close to the optimal set. Nevertheless, from the current central solution $(X(\mu), y(\mu), S(\mu))$ for a sufficiently small μ, we can make a projection onto the boundary of the positive semidefinite cone to generate a solution with zero complementarity gap but with approximate primal-dual feasibility. Such solutions are called approximate maximally complementary solutions.

[1]to derive an upper bound for $\lambda_{\min}(XS)$.

Suppose that a central solution $(X(\mu), y(\mu), S(\mu))$ is given, where μ satisfies (4). The eigenvectors of $X(\mu)$ and $S(\mu)$ can be rearranged so that

$$Q(\mu) := [Q_{\mathcal{B}}(\mu), Q_{\mathcal{T}}(\mu), Q_{\mathcal{N}}(\mu)],$$

e.g., $Q_{\mathcal{B}}(\mu)$ denotes the set of eigenvectors of $Q(\mu)$ converging to an orthonormal basis for $\mathcal{B}$. Let (X^*, y^*, S^*) be a maximally complementary optimal solution, $\hat{X}^* := Q(\mu)^T X^* Q(\mu)$, and $\hat{S}^* := Q(\mu)^T S^* Q(\mu)$, i.e.,

$$\hat{X}^* := \begin{bmatrix} \hat{X}^*_{\mathcal{B}} & \hat{X}^*_{\mathcal{BT}} & \hat{X}^*_{\mathcal{BN}} \\ \hat{X}^*_{\mathcal{TB}} & \hat{X}^*_{\mathcal{T}} & \hat{X}^*_{\mathcal{TN}} \\ \hat{X}^*_{\mathcal{NB}} & \hat{X}^*_{\mathcal{NT}} & \hat{X}^*_{\mathcal{N}} \end{bmatrix}, \quad \hat{S}^* := \begin{bmatrix} \hat{S}^*_{\mathcal{B}} & \hat{S}^*_{\mathcal{BT}} & \hat{S}^*_{\mathcal{BN}} \\ \hat{S}^*_{\mathcal{TB}} & \hat{S}^*_{\mathcal{T}} & \hat{S}^*_{\mathcal{TN}} \\ \hat{S}^*_{\mathcal{NB}} & \hat{S}^*_{\mathcal{NT}} & \hat{S}^*_{\mathcal{N}} \end{bmatrix}.$$

Further, let $\bar{A}^i := Q(\mu)^T A^i Q(\mu)$, and $\Lambda(X(\mu)) := Q(\mu)^T X(\mu) Q(\mu)$ be a diagonal matrix, i.e.,

$$\bar{A}^i := \begin{bmatrix} \bar{A}^i_{\mathcal{B}} & \bar{A}^i_{\mathcal{BT}} & \bar{A}^i_{\mathcal{BN}} \\ \bar{A}^i_{\mathcal{TB}} & \bar{A}^i_{\mathcal{T}} & \bar{A}^i_{\mathcal{TN}} \\ \bar{A}^i_{\mathcal{NB}} & \bar{A}^i_{\mathcal{NT}} & \bar{A}^i_{\mathcal{N}} \end{bmatrix}, \quad \Lambda(X(\mu)) := \begin{bmatrix} \Lambda_{\mathcal{B}}(X(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(X(\mu)) & 0 \\ 0 & 0 & \Lambda_{\mathcal{N}}(X(\mu)) \end{bmatrix}.$$

Then from the primal feasibility constraints we have

$$\langle \bar{A}^i, \hat{X}^* \rangle = b_i, \quad i = 1, \ldots, m, \tag{5}$$

$$\langle \bar{A}^i, \Lambda(X(\mu)) \rangle = b_i, \quad i = 1, \ldots, m, \tag{6}$$

By subtracting (6) from (5) for each i for $\Delta X_{\mathcal{B}}(\mu) = \hat{X}^*_{\mathcal{B}} - \Lambda_{\mathcal{B}}(X(\mu))$ we get

$$\langle \bar{A}^i_{\mathcal{B}}, \Delta X_{\mathcal{B}}(\mu) \rangle = \langle \bar{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu)) \rangle + \langle \bar{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu)) \rangle + \xi_i, \tag{7}$$

where the residual term ξ_i is

$$\xi_i = -\langle \bar{A}^i_{\mathcal{N}}, \hat{X}^*_{\mathcal{N}} \rangle - \langle \bar{A}^i_{\mathcal{T}}, \hat{X}^*_{\mathcal{T}} \rangle - 2\Big(\langle \bar{A}^i_{\mathcal{BT}}, \hat{X}^*_{\mathcal{BT}} \rangle + \langle \bar{A}^i_{\mathcal{BN}}, \hat{X}^*_{\mathcal{BN}} \rangle + \langle \bar{A}^i_{\mathcal{TN}}, \hat{X}^*_{\mathcal{TN}} \rangle \Big).$$

Analogously, let $\bar{C} = Q(\mu)^T C Q(\mu)$, and

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

Then for the dual constraints we get

$$\sum_{i=1}^{m} y_i^* \bar{A}^i + \hat{S}^* = \bar{C}, \tag{8}$$

$$\sum_{i=1}^{m} y_i(\mu) \bar{A}^i + \Lambda(S(\mu)) = \bar{C}. \tag{9}$$

By subtracting (9) from (8) for $\Delta y_i(\mu) = y_i^* - y_i(\mu)$ and $\Delta S_{\mathcal{N}}(\mu) = \hat{S}^*_{\mathcal{N}} - \Lambda_{\mathcal{N}}(S(\mu))$ we get

$$\sum_{i=1}^{m} \Delta y_i(\mu) \bar{A}^i + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S_{\mathcal{N}}(\mu) \end{bmatrix} = \begin{bmatrix} \Lambda_{\mathcal{B}}(S(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S(\mu)) & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \hat{S}^*_{\mathcal{B}} & \hat{S}^*_{\mathcal{BT}} & \hat{S}^*_{\mathcal{BN}} \\ \hat{S}^*_{\mathcal{TB}} & \hat{S}^*_{\mathcal{T}} & \hat{S}^*_{\mathcal{TN}} \\ \hat{S}^*_{\mathcal{NB}} & \hat{S}^*_{\mathcal{NT}} & 0 \end{bmatrix}. \tag{10}$$

We aim to find a primal-dual solution with zero complementarity gap which is approximately primal-dual feasible. Both the right hand side in (7) and the right hand side matrix in (10) depend on the chosen maximally complementary optimal solution. Therefore, the system of equations in (7) and (10) may not be solvable if we drop the unknown terms. Instead, we solve two least squares problems to obtain search directions for the primal-dual solutions.

3.1 Primal least squares problem

For the primal problem we solve

$$\begin{array}{ll} \min & \|e\|^2 + \|\Delta X\|^2 \\ \text{s.t.} & \langle \bar{A}^i_{\mathcal{B}}, \Delta X\rangle - e_i = \langle \bar{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\rangle + \langle \bar{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\rangle, \quad i = 1,\dots,m, \end{array} \tag{11}$$

where $\|.\|$ denotes the l_2-norm. We may assume that $\bar{A}^i_{\mathcal{B}} \neq 0$ for some i. Otherwise, the optimal solution of (11) would give $\Delta X^* = 0$, and thus the effect of the vanishing terms is absorbed in the error of primal infeasibility.

The optimal solution $(e^*, \Delta X^*)$ to the auxiliary problem (11) yields

$$\tilde{X}_{\mathcal{B}} := \Lambda_{\mathcal{B}}(X(\mu)) + \Delta X^*$$

so that

$$\langle \bar{A}^i_{\mathcal{B}}, \tilde{X}_{\mathcal{B}}\rangle = b_i + e^*_i, \qquad i = 1,\dots,m.$$

Thus, $\tilde{X}_{\mathcal{B}}$ has ϵ_p infeasibility for the primal constraints, where

$$\epsilon_p := \|e^*\|. \tag{12}$$

Let

$$\mathcal{A}^v := \begin{bmatrix} \operatorname{vec}(\bar{A}^1)^T \\ \operatorname{vec}(\bar{A}^2)^T \\ \vdots \\ \operatorname{vec}(\bar{A}^m)^T \end{bmatrix}, \quad \mathcal{A}^s_{\mathcal{B}} := \begin{bmatrix} \operatorname{svec}(\bar{A}^1_{\mathcal{B}})^T \\ \operatorname{svec}(\bar{A}^2_{\mathcal{B}})^T \\ \vdots \\ \operatorname{svec}(\bar{A}^m_{\mathcal{B}})^T \end{bmatrix}, \quad r(n) := \frac{n(n+1)}{2},$$

in which the operator vec : $\mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ is the concatenation of the columns of a matrix, and svec : $\mathbb{S}^n \to \mathbb{R}^{r(n)}$ is the concatenation of the upper triangular part of a matrix, where the off-diagonal entries are multiplied by $\sqrt{2}$. Note that $\mathcal{A}^s_{\mathcal{B}}$ might be rank deficient. Then the auxiliary problem (11) reduces to

$$\min \quad \|\mathcal{A}^s_{\mathcal{B}}\Delta x - \eta\|^2 + \|\Delta x\|^2, \tag{13}$$

where $\Delta x = \operatorname{svec}(\Delta X)$, and $\eta_i = \langle \bar{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\rangle + \langle \bar{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\rangle$ denotes the vanishing term for $i = 1,\dots,m$, which should be zero for all optimal solutions. Lemma 1 establishes upper bounds for $\|e^*\|$ and $\|\Delta X^*\|$. The bounds depend on the parameter π_p defined as

$$\pi_p := \prod_{k=1}^{r(n_{\mathcal{B}})} \left\| \left((\mathcal{A}^s_{\mathcal{B}})^T \mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})} \right)_{.k} \right\|,$$

where $\left((\mathcal{A}^s_{\mathcal{B}})^T \mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)_{.k}$ denotes the k^{th} column of $(\mathcal{A}^s_{\mathcal{B}})^T \mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$. Using the upper bounds in Lemma 1 , we show in Theorem 4 that $\tilde{X}_{\mathcal{B}} \succ 0$ for sufficiently small μ.

Lemma 1. *Let* $(e^*, \Delta X^*)$ *be the unique optimal solution to* (11). *Then we have*

$$\begin{aligned} \|\Delta X^*\| &\leq 2\pi_p \sqrt{r(n_{\mathcal{B}})} \|\mathcal{A}^v\|^2 \max\left\{ \frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma \right\}, \\ \epsilon_p = \|e^*\| &\leq 2\|\mathcal{A}^v\| \left(\pi_p \sqrt{r(n_{\mathcal{B}})} \|\mathcal{A}^v\|^2 + 1 \right) \max\left\{ \frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma \right\}. \end{aligned}$$

Proof. The optimality conditions for (13) are given by

$$\big((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)\Delta x = (\mathcal{A}^s_{\mathcal{B}})^T\eta, \tag{14}$$

where $(\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})} \succ 0$. All this means that the system of equations (14) has a unique solution. The solution Δx^* of (14) can be computed using Cramer's rule [6]:

$$\Delta x^*_j = \frac{\det\Big(\big((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)^{(j)}\Big)}{\det\Big((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\Big)}, \quad j = 1,\dots,r(n_{\mathcal{B}}),$$

in which the matrix $\big((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)^{(j)}$ in the nominator is obtained by substituting the j^{th} column of $(\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$ by $(\mathcal{A}^s_{\mathcal{B}})^T\eta$. Noting that[2] $\det((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}) \geq 1$, we can deduce from Hadamard's inequality [6] that

$$|\Delta x^*_j| \leq \Big|\det\Big(\big((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\big)^{(j)}\Big)\Big| \leq \big\|(\mathcal{A}^s_{\mathcal{B}})^T\eta\big\| \prod_{k=1,\ k\neq j}^{r(n_{\mathcal{B}})} \big\|((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})})_{.k}\big\|,$$

for $j = 1,\dots,r(n_{\mathcal{B}})$. Since the diagonal entries of $(\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$ are greater than or equal to 1, the norm of each column is at least 1, and thus a uniform bound for all $j = 1,\dots,r(n_{\mathcal{B}})$ can be derived as

$$|\Delta x^*_j| \leq \big\|(\mathcal{A}^s_{\mathcal{B}})^T\eta\big\| \prod_{k=1,\ k\neq j}^{r(n_{\mathcal{B}})} \big\|((\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})})_{.k}\big\| \leq \pi_p\|(\mathcal{A}^s_{\mathcal{B}})^T\eta\|. \tag{15}$$

Noting that $\|\bar{A}^i_{\mathcal{N}}\| \leq \|\bar{A}^i\| = \|A^i\|$ and $\|\bar{A}^i_{\mathcal{T}}\| \leq \|\bar{A}^i\| = \|A^i\|$, we can conclude from (2) and (3) that

$$\begin{aligned} |\langle \bar{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\rangle| &\leq \frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}\|A^i\|, && i = 1,\dots,m,\\ |\langle \bar{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\rangle| &\leq c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\|A^i\|, && i = 1,\dots,m, \end{aligned}$$

which yields the upper bound

$$|\eta_i| \leq 2\|A^i\|\max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}, \quad i = 1,\dots,m. \tag{16}$$

Consequently, from the bounds in (15) and (16) it follows that

$$|\Delta x^*_j| \leq \pi_p\|(\mathcal{A}^s_{\mathcal{B}})^T\eta\| \leq 2\pi_p\|\mathcal{A}^v\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}, \quad j = 1,\dots,r(n_{\mathcal{B}}),$$

where we have used $\|\mathcal{A}^v\|^2 = \sum_{i=1}^m\|A^i\|^2$, and the inequality $\|\mathcal{A}^s_{\mathcal{B}}\| \leq \|\mathcal{A}^v\|$. As a result, we get

$$\begin{aligned} \|e^*\| &= \|\mathcal{A}^s_{\mathcal{B}}\Delta x^* - \eta\| \leq \|\mathcal{A}^s_{\mathcal{B}}\|\|\Delta x^*\| + \|\eta\|\\ &\leq 2\|\mathcal{A}^v\|\left(\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^v\|^2 + 1\right)\max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}. \end{aligned}$$

This completes the proof. □

[2]This is true regardless of data type, since the eigenvalues of $(\mathcal{A}^s_{\mathcal{B}})^T\mathcal{A}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$ are at least 1.

3.2 Dual least squares problem

Let E denote a residual matrix as

$$E := \begin{bmatrix} E_{\mathcal{B}} & E_{\mathcal{B}\mathcal{T}} & E_{\mathcal{B}\mathcal{N}} \\ E_{\mathcal{T}\mathcal{B}} & E_{\mathcal{T}} & E_{\mathcal{T}\mathcal{N}} \\ E_{\mathcal{N}\mathcal{B}} & E_{\mathcal{N}\mathcal{T}} & 0 \end{bmatrix}, \tag{17}$$

which is defined in accordance with the unknown right hand side matrix in (10). Then the auxiliary problem for an approximate dual solution is formulated as

$$\begin{aligned} \min \quad & \|E\|^2 + \|\Delta y\|^2 + \|\Delta S\|^2 \\ \text{s.t.} \quad & \sum_{i=1}^{m} \Delta y_i \bar{A}^i + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S \end{bmatrix} - E = \begin{bmatrix} \Lambda_{\mathcal{B}}(S(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S(\mu)) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{18}$$

The optimal solution $(E^*, \Delta y^*, \Delta S^*)$ gives $\tilde{y}_i := y_i(\mu) + \Delta y_i^*$ for $i = 1, \dots, m$ and $\tilde{S}_{\mathcal{N}} := \Lambda_{\mathcal{N}}(S(\mu)) + \Delta S^*$ with ϵ_d infeasibility for the dual constraints, where

$$\epsilon_d := \|E^*\|. \tag{19}$$

For the sake of clarity in what follows, the auxiliary problem (18) is represented in vector form. To do so, we consider the vec(.) operator for each block of $\bar{A}_i$, i.e., $\mathcal{A}^v_{\mathcal{B}}$, $\mathcal{A}^v_{\mathcal{N}}$, $\mathcal{A}^v_{\mathcal{T}}$, $\mathcal{A}^v_{\mathcal{B}\mathcal{T}}$, $\mathcal{A}^v_{\mathcal{B}\mathcal{N}}$, and $\mathcal{A}^v_{\mathcal{T}\mathcal{N}}$. Therefore, the auxiliary problem (18) can be simplified to the following least squares problem

$$\begin{aligned} \min \quad & \|(\mathcal{A}^v_{\mathcal{B}})^T \Delta y - \zeta_{\mathcal{B}}\|^2 + \|(\mathcal{A}^v_{\mathcal{T}})^T \Delta y - \zeta_{\mathcal{T}}\|^2 + \|(\mathcal{A}^v_{\mathcal{N}})^T \Delta y\|^2 + \|\Delta y\|^2 \\ & + 2\|(\mathcal{A}^v_{\mathcal{B}\mathcal{T}})^T \Delta y\|^2 + 2\|(\mathcal{A}^v_{\mathcal{B}\mathcal{N}})^T \Delta y\|^2 + 2\|(\mathcal{A}^v_{\mathcal{T}\mathcal{N}})^T \Delta y\|^2, \end{aligned} \tag{20}$$

where $\zeta_{\mathcal{B}} = \text{vec}(\Lambda_{\mathcal{B}}(S(\mu)))$ and $\zeta_{\mathcal{T}} = \text{vec}(\Lambda_{\mathcal{T}}(S(\mu)))$. Lemma 2 establishes upper bounds for ϵ_d and $\|\Delta S^*\|$. For the upper bounds we define the parameter π_d as

$$\pi_d := \prod_{k=1}^{m} \|\mathcal{H}_{.k}\|,$$

where

$$\begin{aligned} \mathcal{H} := {} & \mathcal{A}^v_{\mathcal{B}}(\mathcal{A}^v_{\mathcal{B}})^T + \mathcal{A}^v_{\mathcal{T}}(\mathcal{A}^v_{\mathcal{T}})^T + \mathcal{A}^v_{\mathcal{N}}(\mathcal{A}^v_{\mathcal{N}})^T + 2\mathcal{A}^v_{\mathcal{B}\mathcal{T}}(\mathcal{A}^v_{\mathcal{B}\mathcal{T}})^T + 2\mathcal{A}^v_{\mathcal{B}\mathcal{N}}(\mathcal{A}^v_{\mathcal{B}\mathcal{N}})^T \\ & + 2\mathcal{A}^v_{\mathcal{T}\mathcal{N}}(\mathcal{A}^v_{\mathcal{T}\mathcal{N}})^T + I_m, \end{aligned}$$

and $\mathcal{H}_{.k}$ denotes the k^{th} column of $\mathcal{H}$. Observe that $\mathcal{H}$ is a positive definite matrix. Further, Theorem 4 proves that $\tilde{S}_{\mathcal{N}} \succ 0$ for sufficiently small μ.

Lemma 2. *Problem* (18) *has a unique optimal solution* $(E^*, \Delta y^*, \Delta S^*)$*, which satisfies*

$$\begin{aligned} \|\Delta S^*\| &\le 2\pi_d\sqrt{m}\|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}, \\ \epsilon_d = \|E^*\| &\le \sqrt{2}(4\pi_d\sqrt{m}\|\mathcal{A}^v\|^2 + 1) \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}. \end{aligned}$$

Proof. For the sake of brevity we define

$$\varphi := \mathcal{A}^v_{\mathcal{B}}\zeta_{\mathcal{B}} + \mathcal{A}^v_{\mathcal{T}}\zeta_{\mathcal{T}}.$$

The optimality conditions for (20) can be written as

$$\mathcal{H}\Delta y = \varphi,$$

where $\mathcal{H} \succ 0$, and thus (20) has a unique solution. This solution can be computed by using Cramer's rule as follows

$$\Delta y_i^* = \frac{\det(\mathcal{H}^{(i)})}{\det(\mathcal{H})}, \quad i = 1, \ldots, m,$$

where the matrix $\mathcal{H}^{(i)}$ in the nominator is obtained by substituting the i^{th} column of $\mathcal{H}$ by φ. Note that $\lambda_{\min}(\mathcal{H}) \geq 1$, which implies $\det(\mathcal{H}) \geq 1$. Therefore, we get

$$|\Delta y_i^*| \leq |\det(\mathcal{H}^{(i)})| \leq \|\varphi\| \prod_{k=1,\ k\neq i}^{m} \|\mathcal{H}_{.k}\| \leq \pi_d \|\varphi\|, \quad i = 1, \ldots, m,$$

where the second inequality follows from Hadamard's inequality. Note that $\prod_{k=1,k\neq i}^{m} \|\mathcal{H}_{.k}\| \leq \pi_d$, since the diagonal entries of $\mathcal{H}$ are at least 1. Furthermore, we have from (1) and (3) that

$$\|\varphi\| \leq \|\mathcal{A}_{\mathcal{B}}^v \zeta_{\mathcal{B}}\| + \|\mathcal{A}_{\mathcal{T}}^v \zeta_{\mathcal{T}}\| \leq 2\|\mathcal{A}^v\| \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\},$$

which leads to

$$|\Delta y_i^*| \leq 2\pi_d \|\mathcal{A}^v\| \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}, \quad i = 1, \ldots, m. \tag{21}$$

Consequently, from (21) we have

$$\|\Delta S^*\| = \|(\mathcal{A}_{\mathcal{N}}^v)^T \Delta y^*\| \leq 2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}.$$

Note that $\|QEQ^T\| = \|E\|$. Then using (21) we can also derive bounds for the components of the residual matrix as follows

$$\begin{aligned}
\|E_{\mathcal{B}}^*\| &= \|(\mathcal{A}_{\mathcal{B}}^v)^T \Delta y^* - \zeta_{\mathcal{B}}\| \leq 2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\} \\
&\quad + \frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma} \leq (2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 + 1) \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}, \\
\|E_{\mathcal{T}}^*\| &= \|(\mathcal{A}_{\mathcal{T}}^v)^T \Delta y^* - \zeta_{\mathcal{T}}\| \leq 2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\} \\
&\quad + c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma} \leq (2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 + 1) \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}, \\
\|E_{\mathcal{BT}}^*\|, \|E_{\mathcal{BN}}^*\|, \|E_{\mathcal{TN}}^*\| &\leq 2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}.
\end{aligned}$$

Then we get

$$\begin{aligned}
\|E^*\|^2 &\leq \left(2\Big(2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 + 1\Big)^2 + 6\Big(2\pi_d \sqrt{m} \|\mathcal{A}^v\|^2\Big)^2\right) \left(\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}\right)^2 \\
&\leq 2\Big(4\pi_d \sqrt{m} \|\mathcal{A}^v\|^2 + 1\Big)^2 \left(\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}\right)^2,
\end{aligned}$$

from which the bound for ϵ_d follows. □

3.3 Cone feasibility

As specified by Lemmas 1 and 2, $\left(Q_{\mathcal{B}}(\mu)\tilde{X}_{\mathcal{B}}Q_{\mathcal{B}}(\mu)^T, \tilde{y}, Q_{\mathcal{N}}(\mu)\tilde{S}_{\mathcal{N}}Q_{\mathcal{N}}(\mu)^T\right)$ yields a complementary solution for the primal and dual SDO problems. This primal-dual pair has $\epsilon := \max\{\epsilon_p, \epsilon_d\}$ infeasibility w.r.t. the linear constraints. Theorem 4 shows that for a sufficiently small μ, the rounding procedure yields a primal-dual solution with $\tilde{X}_{\mathcal{B}}, \tilde{S}_{\mathcal{N}} \succ 0$.

Theorem 4. *Let $\vartheta_1 := 2n^2\|\mathcal{A}^v\|^2$, $\vartheta_2 := 2cn\sqrt{nn_{\mathcal{T}}}\|\mathcal{A}^v\|^2$, and let*

$$\tilde{\mu} := \min\left\{\frac{\sigma^2}{\vartheta_1 \max\{\pi_p\sqrt{r(n_{\mathcal{B}})n_{\mathcal{N}}},\ \pi_d\sqrt{mn_{\mathcal{B}}}\}}, \frac{1}{n}\left(\frac{\sigma}{\vartheta_2 \max\{\pi_p\sqrt{r(n_{\mathcal{B}})},\ \pi_d\sqrt{m}\}}\right)^{\frac{1}{\gamma}}\right\}.$$

If $\mu \leq \tilde{\mu}$, then we have $\tilde{X}_{\mathcal{B}}, \tilde{S}_{\mathcal{N}} \succ 0$.

Proof. We only need to show that for $\mu \leq \tilde{\mu}$ the rounding procedure results in $\tilde{X}_{\mathcal{B}}, \tilde{S}_{\mathcal{N}} \succ 0$. Noting that

$$|\lambda_{\min}(\Delta X^*)| \leq \|\Delta X^*\|, \quad |\lambda_{\min}(\Delta S^*)| \leq \|\Delta S^*\|,$$

we can conclude from (1) and (2), and Lemmas 1 and 2 that

$$\begin{aligned}
\lambda_{\min}(\tilde{X}_{\mathcal{B}}) &\geq \lambda_{\min}(\Lambda_{\mathcal{B}}(X(\mu))) + \lambda_{\min}(\Delta X^*) \\
&\geq \frac{\sigma}{n} - 2\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}, \\
\lambda_{\min}(\tilde{S}_{\mathcal{N}}) &\geq \lambda_{\min}(\Lambda_{\mathcal{N}}(S(\mu)) + \lambda_{\min}(\Delta S^*) \\
&\geq \frac{\sigma}{n} - 2\pi_d\sqrt{m}\|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\}.
\end{aligned}$$

Consequently, $\tilde{X}_{\mathcal{B}}, \tilde{S}_{\mathcal{N}} \succ 0$ holds if

$$\begin{aligned}
2\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\} &< \frac{\sigma}{n}, \\
2\pi_d\sqrt{m}\|\mathcal{A}^v\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^{\gamma}\right\} &< \frac{\sigma}{n}.
\end{aligned}$$

These inequalities hold if $\mu \leq \tilde{\mu}$. The proof is complete. □

Remark 1. *Computing an ϵ-feasible maximally complementary solution requires $\mathcal{O}(\max\{n_{\mathcal{B}}^6, m^3\})$ arithmetic operations. In fact, solving* (11) *and* (18) *is equivalent to solving two linear systems of equations, using the Gauss elimination method, with $r(n_{\mathcal{B}})$ and m variables, respectively.*

3.4 Rounding procedure

Using the bounds given in Lemmas 1 and 2, and Theorem 4, we can outline a simple procedure which yields an approximate maximally complementary solution.

Algorithm 1 Rounding procedure for SDO

Parameters
- Desired tolerance for primal infeasibility according to the bound for $\|e^*\|$.
- Desired tolerance for dual infeasibility according to the bound for $\|E^*\|$.

Input
- A central solution $(X(\mu), y(\mu), S(\mu))$, where $\mu \leq \tilde{\mu}$.

Do
- Solve the primal least squares problem (11) to get $\tilde{X}_{\mathcal{B}}$.
- Solve the dual least squares problem (18) to get $(\tilde{y}, \tilde{S}_{\mathcal{N}})$.

return Approximate maximally complementary solution $\left(Q_{\mathcal{B}}(\mu)\tilde{X}_{\mathcal{B}}Q_{\mathcal{B}}(\mu)^T, \tilde{y}, Q_{\mathcal{N}}(\mu)\tilde{S}_{\mathcal{N}}Q_{\mathcal{N}}(\mu)^T\right)$.

4 A rounding procedure for approximate solutions

The rounding procedure obtains an approximate maximally complementary solution from a central solution sufficiently close to the optimal set. Let $(X, y, S) \in \mathcal{N}_\kappa(\tau)$, where $X = M\Lambda(X)M^T$ and $S = P\Lambda(S)P^T$, and M and P are orthogonal matrices. To extend the rounding procedure to solutions in the neighborhood of the central path, we need to choose the eigenvectors either from M or P, because X and S do not commute[3]. To do so, we can solve the primal least squares problem (11), where $X(\mu)$ and $Q(\mu)$ are replaced by X and M, respectively, in the definition of $\bar{A}^i$, $\bar{C}$, and the right hand side in (11). We then solve the following least squares problem to compute a dual solution:

$$\begin{aligned} \min \quad & \|E\|^2 + \|\Delta S\|^2 + \|\Delta y\|^2 \\ \text{s.t.} \quad & \sum_{i=1}^m \Delta y_i \bar{A}^i + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S \end{bmatrix} - E = \begin{bmatrix} M_\mathcal{B}^T S M_\mathcal{B} & M_\mathcal{B}^T S M_\mathcal{T} & M_\mathcal{B}^T S M_\mathcal{N} \\ M_\mathcal{T}^T S M_\mathcal{B} & M_\mathcal{T}^T S M_\mathcal{T} & M_\mathcal{T}^T S M_\mathcal{N} \\ M_\mathcal{N}^T S M_\mathcal{B} & M_\mathcal{N}^T S M_\mathcal{T} & 0 \end{bmatrix}, \end{aligned} \tag{22}$$

where E is defined as in (17). Let $\left(M_\mathcal{B}\tilde{X}_\mathcal{B}M_\mathcal{B}^T, \tilde{y}, M_\mathcal{N}\tilde{S}_\mathcal{N}M_\mathcal{N}^T\right)$ be the updated primal-dual solution after applying the search directions from (11) and (22), where

$$\tilde{X}_\mathcal{B} = \Lambda_\mathcal{B}(X) + \Delta X^*, \qquad \tilde{S}_\mathcal{N} = M_\mathcal{N}^T S M_\mathcal{N} + \Delta S^*, \qquad \tilde{y}_i = y_i(\mu) + \Delta y_i^*, \qquad \forall\, i = 1, \ldots, m.$$

We can show, in a similar manner as in Section 3, that $\left(M_\mathcal{B}\tilde{X}_\mathcal{B}M_\mathcal{B}^T, \tilde{y}, M_\mathcal{N}\tilde{S}_\mathcal{N}M_\mathcal{N}^T\right)$ becomes an approximate complementary solution if the complementarity gap $\langle X, S\rangle$ is sufficiently small.

Alternatively, we may fix the basis at P and solve (18) to compute a dual solution, where $(y(\mu), S(\mu))$ and $Q(\mu)$ are replaced by (y, S) and P, respectively, in the definition of $\bar{A}^i$ and the right hand side in (18). Afterwards, we solve

$$\begin{aligned} \min \quad & \|e\|^2 + \|\Delta X\|^2 \\ \text{s.t.} \quad & \langle \bar{A}_\mathcal{B}^i, \Delta X\rangle - e_i = \bar{b}_i, \quad i = 1, \ldots, m, \end{aligned} \tag{23}$$

where

$$\bar{b}_i = \langle \bar{A}_\mathcal{T}^i, P_\mathcal{T}^T X P_\mathcal{T}\rangle + \langle \bar{A}_\mathcal{N}^i, P_\mathcal{N}^T X P_\mathcal{N}\rangle + 2\langle \bar{A}_{\mathcal{B}\mathcal{T}}^i, P_\mathcal{B}^T X P_\mathcal{T}\rangle + 2\langle \bar{A}_{\mathcal{B}\mathcal{N}}^i, P_\mathcal{B}^T X P_\mathcal{N}\rangle + 2\langle \bar{A}_{\mathcal{T}\mathcal{N}}^i, P_\mathcal{T}^T X P_\mathcal{N}\rangle.$$

Let $\left(P_\mathcal{B}\tilde{X}_\mathcal{B}P_\mathcal{B}^T, \tilde{y}, P_\mathcal{N}\tilde{S}_\mathcal{N}P_\mathcal{N}^T\right)$ be the new primal-dual solution after applying the search directions from (18) and (23), where

$$\tilde{X}_\mathcal{B} = P_\mathcal{B}^T X P_\mathcal{B} + \Delta X^*, \qquad \tilde{S}_\mathcal{N} = \Lambda_\mathcal{N}(S) + \Delta S^*, \qquad \tilde{y}_i = y_i(\mu) + \Delta y_i^*, \qquad \forall\, i = 1, \ldots, m.$$

Then for sufficiently small complementarity gap $\langle X, S\rangle$, we can show that $\left(P_\mathcal{B}\tilde{X}_\mathcal{B}P_\mathcal{B}^T, \tilde{y}, P_\mathcal{N}\tilde{S}_\mathcal{N}P_\mathcal{N}^T\right)$ is an approximate maximally complementary solution. The approach to derive the feasibility bounds is analogous to Section 3.

5 Concluding remarks

In this paper, we proposed a rounding procedure to generate an approximate maximally complementary solution of SDO. The sets of eigenvectors converging to an orthonormal bases for the optimal partition are identified from an interior solution, which is either a central solution, or a solution in a neighborhood of the central path. Using these sets of eigenvectors, the procedure generates an ϵ-feasible primal-dual solution with zero complementarity gap. It is proved that if the complementarity gap drops below a certain bound, then the primal-dual solution satisfies the cone constraints, yielding an approximate maximally complementary solution.

[3]For a given central solution $(X(\mu), y(\mu), S(\mu))$ we get the same sets of eigenvectors from $X(\mu)$ and $S(\mu)$.

Acknowledgements

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

References

[1] F. Alizadeh. *Combinatorial Optimization with Interior Point Methods and Semidefinite Matrices*. PhD thesis, University of Minnesota, 1991.

[2] J. F. Bonnans and H. Ramírez C. Perturbation analysis of second-order cone programming problems. *Mathematical Programming*, 104(2):205–227, 2005.

[3] E. de Klerk. *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, volume 65 of Series Applied Optimization. Springer, 2006.

[4] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. *SIAM Journal on Optimization*, 8(4):871–886, 1998.

[5] M. Halická, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM Journal on Optimization*, 12(4):1090–1099, 2002.

[6] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2012.

[7] T. Illés, J. Peng, C. Roos, and T. Terlaky. A strongly polynomial rounding procedure yielding a maximally complementary solution for $P_*(\kappa)$ linear complementarity problems. *SIAM Journal on Optimization*, 11(2):320–340, 2000.

[8] A. Mohammad-Nezhad and T. Terlaky. On the identification of optimal partition for semidefinite optimization. Technical Report 17T-008, Department of Industrial and Systems Engineering, Lehigh University, 2017.

[9] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13. Society for Industrial and Applied Mathematics, 1994.

[10] G. Pataki. Strong duality in conic linear programming: Facial reduction and extended duals. In D. Bailey, H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, editors, *Computational and Analytical Mathematics*, volume 50 of Springer Proceedings in Mathematics & Statistics, pages 613–634. Springer, 2013.

[11] J. Peña and V. Roshchina. A complementarity partition theorem for multifold conic systems. *Mathematical Programming*, 142(1):579–589, 2013.

[12] C. Roos, T. Terlaky, and J.-P. Vial. *Interior Point Methods for Linear Optimization*. Springer, 2005.

[13] J. F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, 2000.

[14] T. Terlaky and Z. Wang. On the identification of the optimal partition of second order cone optimization problems. *SIAM Journal on Optimization*, 24(1):385–414, 2014.

[15] T. Wang and J.-S. Pang. Global error bounds for convex quadratic inequality systems. *Optimization*, 31(1):1–12, 1994.

[16] Y. Ye. On the finite convergence of interior-point algorithms for linear programming. *Mathematical Programming*, 57(1):325–335, 1992.

[17] E. Yildirim. Unifying optimal partition approach to sensitivity analysis in conic optimization. *Journal of Optimization Theory and Applications*, 122(2):405–423, 2004.