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[a,*], Tamás Terlaky[a,1]

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

Abstract

Recently, Mohammad-Nezhad and Terlaky studied the identification of the optimal partition for semidefinite optimization. An approximation of the optimal partition was obtained from a bounded sequence of solutions on, or in a neighborhood of the central path. Here, we use the approximation of the optimal partition in a rounding procedure to generate an approximate maximally complementary solution. The procedure generates a rounded primal-dual solution from an interior solution, sufficiently close to the optimal set, by solving two least square problems.

Keywords: Semidefinite optimization, Optimal partition, Interior point methods, Maximally complementary solution
2010 MSC: 90C22, 90C25, 90C51

1. Introduction

The notion of optimal partition for linear optimization (LO) and linear complementarity problems (LCPs) is well-known in the literature of interior point methods (IPMs). For LO, the Goldman-Tucker theorem [5] proves the existence of strictly complementary solutions, and hence the partition of the index set of the variables to two disjoint complementary sets. For LCPs [9], only maximally complementary solutions exist. Hence, the optimal partition consists of three disjoint sets.

Goldfarb and Scheinberg [4] extended the concept of the optimal partition to semidefinite optimization (SDO), and Yildirim [17] presented a facial description of the optimal partition for linear conic optimization with self-dual cones. The optimal partition for second-order conic optimization was studied by Bonnans and Ramírez [1]. Peña and Roshchina [11] established the complementarity partition for a multifold homogeneous conic system.

The optimal partition provides a characterization of the optimal set, and it is uniquely defined for SDO problems with strong duality. The optimal partition information can be used in a so called rounding procedure to generate either a maximally or strictly complementary solution, see e.g., Ye [16] and Roos, Terlaky, and Vial [12] for LO, and Illés, Peng, Roos, and Terlaky [9] and the references there in for LCP. Recently, Terlaky and Wang [15] and Mohammad-Nezhad and Terlaky [10] have investigated the identification of the optimal partition for SOCO and SDO, respectively. The rationale behind the identification of the optimal partition is closely interconnected with the limiting behavior of the central path and the existence of a maximally complementary solution. In [10], the approximation of the optimal partition was obtained from the eigenvectors of interior solutions, on or in a neighborhood of the central path, whose accumulation points form orthonormal bases for the subspaces of the optimal partition.

In this paper, we revisit the identification of the optimal partition given in [10] for SDO. Our contribution is a rounding procedure for an approximate maximally complementary solution of SDO. We use the approximation of the optimal partition from [10] to generate an approximate maximally complementary solution through solving a pair of auxiliary problems. To the best of our knowledge, the identification of the optimal partition and rounding procedures for SDO have not been studied before in the literature of IPMs. The rest of this paper is built as follows. In Section 2, we review the concepts of the complementarity and optimal partition 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. Concluding remarks and topics for future studies are presented in Section 4.

Throughout this paper, $\mathbb{S}^n$ is the space of $n \times n$ symmetric matrices, $\mathbb{S}^n_+$ is the cone of $n \times n$ positive semidefinite matrices, and $\operatorname{ri}(.)$ denotes the relative interior of a set. For a symmetric matrix $\lambda_{[i]}(.)$ denotes the i^{th} largest eigenvalue. Furthermore, $\mathcal{R}(.)$ stands for the column space and $\operatorname{Null}(.)$ denotes the null space of a matrix. The Frobenius norm is indicated by $\|.\|$, and the l_2 norm of a vector is denoted by $\|.\|_2$. We use the notation $(., ., \ldots, .)$ to indicate the side by side arrangement of the column vectors and matrices.

*Corresponding author
Email addresses: alm413@lehigh.edu (Ali Mohammad-Nezhad), terlaky@lehigh.edu (Tamás Terlaky)

2. Preliminaries

A pair of primal and dual SDO problems in standard form can be phrased as

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

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

where $C, X, A^i \in \mathbb{S}^n$ for $i = 1,\dots,m$, $b \in \mathbb{R}^m$, and the inner product is defined as $\langle C, X\rangle := \operatorname{tr}(CX)$. 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. The following assumptions are made throughout this paper:

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

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

The IPC is standard in the literature of IPMs for linear and conic optimization [12]. It ensures that the primal and dual problems are solvable, and strong duality holds, i.e., for every primal-dual optimal solution we have $XS = 0$. The primal-dual optimal pair (X^*, y^*, S^*) is called maximally complementary (MC) if $\operatorname{rank}(X^* + S^*)$ is maximal over the optimal set, or equivalently, if $(X^*, y^*, S^*) \in \operatorname{ri}(\mathcal{P}^*\times\mathcal{D}^*)$, see e.g., Lemma 2.3 in [2]. An MC solution (X^*, y^*, S^*) is called strictly complementary if $X^* + S^* \succ 0$. All $X^* \in \operatorname{ri}(\mathcal{P}^*)$ have the same range space, see e.g., Lemma 2.3 in [2] or Lemma 3.1 in [4]. The same is true for S^* with $(y^*, S^*) \in \operatorname{ri}(\mathcal{D}^*)$. Recall [2] that an SDO problem may have no strictly complementary solution.

2.1. Optimal partition of SDO

Let $(X^*, y^*, S^*) \in \operatorname{ri}(\mathcal{P}^*\times\mathcal{D}^*)$ be an MC solution. Since X^* and S^* commute by the complementarity condition, they are simultaneously diagonalizable, i.e., there exists an orthogonal matrix Q^* so that $X^* = Q^*\Lambda(X^*)(Q^*)^T$ and $S^* = Q^*\Lambda(S^*)(Q^*)^T$, where $\Lambda(X^*)$ and $\Lambda(S^*)$ are diagonal matrices of the eigenvalues. Then $\mathcal{R}(X^*) = \mathcal{R}(Q^*\Lambda(X^*))$ and $\mathcal{R}(S^*) = \mathcal{R}(Q^*\Lambda(S^*))$ indicate that the subspaces $\mathcal{R}(X^*)$ and $\mathcal{R}(S^*)$ are orthogonal, and they are spanned by the eigenvectors associated with the positive eigenvalues of X^* and S^*, respectively. Let us define $\mathcal{B} := \mathcal{R}(X^*)$, $\mathcal{N} := \mathcal{R}(S^*)$, and $\mathcal{T} := \left(\mathcal{R}(X^*) + \mathcal{R}(S^*)\right)^\perp$, where $\perp$ denotes the orthogonal complement of a subspace. Then $(\mathcal{B}, \mathcal{T}, \mathcal{N})$, which partitions $\mathbb{R}^n$, is called the optimal partition of an SDO problem. It is immediate that $\mathcal{T} = \{0\}$ when a strictly complementary solution exists. Since $(X^*, y^*, S^*) \in \operatorname{ri}(\mathcal{P}^*\times\mathcal{D}^*)$, the optimal partition is invariant w.r.t. the choice of (X^*, y^*, S^*).

We consider $Q := (Q_\mathcal{B}, Q_\mathcal{T}, Q_\mathcal{N})$ as an orthonormal basis partitioned according to the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$. For instance, the eigenvectors associated with the positive eigenvalues of X^* and S^* can be chosen as orthonormal bases for $\mathcal{B}$ and $\mathcal{N}$, respectively. Theorem 1 characterizes $\mathcal{P}^*$ and $\mathcal{D}^*$. For brevity, we define $n_\mathcal{B} := \dim(\mathcal{B})$, $n_\mathcal{T} := \dim(\mathcal{T})$, and $n_\mathcal{N} := \dim(\mathcal{N})$.

Theorem 1 (Theorem 2.7 in [2])**.** *The primal and dual optimal solutions of SDO can be represented by $(Q_\mathcal{B} U_{\check{X}} Q_\mathcal{B}^T, \check{y}, Q_\mathcal{N} U_{\check{S}} Q_\mathcal{N}^T)$, where $U_{\check{X}} \in \mathbb{S}_+^{n_\mathcal{B}}$ and $U_{\check{S}} \in \mathbb{S}_+^{n_\mathcal{N}}$. If $n_\mathcal{B} > 0$ and $X^* \in \operatorname{ri}(\mathcal{P}^*)$, then $U_{X^*} \succ 0$. Analogously, if $n_\mathcal{N} > 0$ and $(y^*, S^*) \in \operatorname{ri}(\mathcal{D}^*)$, then $U_{S^*} \succ 0$.* □

From Theorem 1, for any $(\check{X}, \check{y}, \check{S}) \in \mathcal{P}^*\times\mathcal{D}^*$ we have $Q_{\mathcal{T}\cup\mathcal{N}}^T \check{X} Q_{\mathcal{T}\cup\mathcal{N}} = 0$ and $Q_{\mathcal{B}\cup\mathcal{T}}^T \check{S} Q_{\mathcal{B}\cup\mathcal{T}} = 0$, 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 1. At least one of $n_\mathcal{B}$ or $n_\mathcal{N}$ has to be positive by the IPC. The case with $n_\mathcal{B} = 0$ or $n_\mathcal{N} = 0$ turns out to be fairly trivial, and it is easy to show that this trivial case can actually happen. Hence, we can assume w.l.o.g. that both $n_\mathcal{B}$ and $n_\mathcal{N}$ are positive.

2.2. Approximation of the optimal partition

The central path is defined as the set of solutions to

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

where $XS = \mu I_n$ is the centrality condition, and I_n stands for the identity matrix of size n. Assuming linear independence of A^i for $i = 1,\dots,m$ and the IPC, for any given $\mu > 0$ the solution $(X(\mu), y(\mu), S(\mu))$, so called central solution, is unique. It readily follows from the centrality condition that $X(\mu)$ and $S(\mu)$ commute, and thus they have a common eigenvector basis. Furthermore, it is well-known that the central path converges to an MC solution [6].

Remark 2. The concept of the optimal partition is well-defined only when strong duality holds, and without the IPC the central path does not exist. It is known that the IPC can be made w.l.o.g., by using the self-dual embedding model, see e.g., [3]. Note that in this case, the embedding model is always well-posed (in terms of the IPC) even if the original problem is not.

Recall the condition number σ from [10]:

$$\sigma_\mathcal{B} := \begin{cases} \max_{\check{X}\in\mathcal{P}^*} \lambda_{\min}(Q_\mathcal{B}^T \check{X} Q_\mathcal{B}), & \mathcal{B} \neq \{0\},\\ \infty, & \mathcal{B} = \{0\}, \end{cases} \tag{1}$$

$$\sigma_\mathcal{N} := \begin{cases} \max_{(\check{y},\check{S})\in\mathcal{D}^*} \lambda_{\min}(Q_\mathcal{N}^T \check{S} Q_\mathcal{N}), & \mathcal{N} \neq \{0\},\\ \infty, & \mathcal{N} = \{0\}, \end{cases} \tag{2}$$

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

Condition number σ quantifies the magnitude of the eigenvalues on the central path. By the IPC and the compactness of the optimal set, there exists an MC solution, and the optimal values in (1) or (2) are attained. Hence, σ is well-defined and positive. See Appendix 1 in [10] for a positive lower bound on σ. Moreover, Section 5 in [10] adds more comments on the magnitude of σ and complexity of computing an exact solution of SDO.

Let $\hat{X}(\mu) := Q^T X(\mu) Q$ and $\hat{S}(\mu) := Q^T S(\mu) Q$ be an orthogonal transformation of $(X(\mu), y(\mu), S(\mu))$ w.r.t. Q

(see Section 2.1), where

$$\hat{X}(\mu) := \begin{bmatrix} \hat{X}_{\mathcal{B}}(\mu) & \hat{X}_{\mathcal{BT}}(\mu) & \hat{X}_{\mathcal{BN}}(\mu) \\ \hat{X}_{\mathcal{TB}}(\mu) & \hat{X}_{\mathcal{T}}(\mu) & \hat{X}_{\mathcal{TN}}(\mu) \\ \hat{X}_{\mathcal{NB}}(\mu) & \hat{X}_{\mathcal{NT}}(\mu) & \hat{X}_{\mathcal{N}}(\mu) \end{bmatrix}, \tag{3}$$

and $\hat{S}(\mu)$ is defined analogously. In [10], the derivation of bounds for $Q^T_{\mathcal{T}\cup\mathcal{N}} X(\mu) Q_{\mathcal{T}\cup\mathcal{N}}$ and $Q^T_{\mathcal{B}\cup\mathcal{T}} S(\mu) Q_{\mathcal{B}\cup\mathcal{T}}$ relies on an error bound result, see Theorem 3.3 in [14], for the following linear matrix inequality (LMI) systems

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

where $(\check{X}, \check{y}, \check{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ is a primal-dual optimal pair,

$$\mathcal{A}^s := \big(\operatorname{svec}(A^1), \dots, \operatorname{svec}(A^m)\big)^T,$$

and svec(.) is an isometry between $\mathbb{S}^n$ and $\mathbb{R}^{n(n+1)/2}$ stacking the upper triangular part of a symmetric matrix, in which the off-diagonal entries are multiplied by $\sqrt{2}$, see e.g., [2]. Recall from the feasibility conditions that

$$\langle X(\mu) - \check{X}, S(\mu) - \check{S} \rangle = 0, \tag{4}$$

which, by the optimality of $\check{X}$ and $\check{S}$, implies

$$\operatorname{svec}(\check{S})^T \operatorname{svec}\big(X(\mu)\big) \le n\mu, \quad \operatorname{svec}(\check{X})^T \operatorname{svec}\big(S(\mu)\big) \le n\mu.$$

Therefore, by the orthogonal projection of $\operatorname{svec}\big(X(\mu)\big)$ and $\operatorname{svec}\big(S(\mu)\big)$ onto the affine subspaces

$$\bar{\mathcal{L}}_{\mathcal{P}} := \Big\{ x \mid x \in \operatorname{svec}(\check{X}) + \operatorname{Null}(\mathcal{A}^s),\ \operatorname{svec}(\check{S})^T x = 0 \Big\},$$
$$\bar{\mathcal{L}}_{\mathcal{D}} := \Big\{ s \mid s \in \operatorname{svec}(\check{S}) + \mathcal{R}\big((\mathcal{A}^s)^T\big),\ \operatorname{svec}(\check{X})^T s = 0 \Big\},$$

we get

$$\begin{aligned} \operatorname{dist}\Big(\operatorname{svec}\big(X(\mu)\big), \bar{\mathcal{L}}_{\mathcal{P}}\Big) &\le \theta_1 n\mu, \\ \operatorname{dist}\Big(\operatorname{svec}\big(S(\mu)\big), \bar{\mathcal{L}}_{\mathcal{D}}\Big) &\le \theta_2 n\mu, \end{aligned} \tag{5}$$

where dist(.,.) is induced by the Frobenius norm, and θ_1 and θ_2 depend on $\mathcal{A}^s$ and $\check{S}$, and $\mathcal{A}^s$ and $\check{X}$, respectively. Interestingly, θ_1 and θ_2 can be interpreted as Hoffman [7] condition numbers. Now, we derive an upper bound on the distance between a central solution and the optimal set.

Lemma 2. *Let $(X(\mu), y(\mu), S(\mu))$ with*

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

be given. Then there exists $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

$$\big\|X(\mu) - X_\mu\big\| \le c(n\mu)^\gamma, \quad \big\|S(\mu) - S_\mu\big\| \le c(n\mu)^\gamma, \tag{6}$$

where $c > 0$ is independent of μ, and $\gamma \ge 2^{1-n}$.

Proof. It is known that the set of central solutions $(X(\mu), y(\mu), S(\mu))$ for $0 < \mu \le \hat{\mu}$ is bounded, see Lemma 3.2 in [2]. If $\mu \le \hat{\mu}$ holds, then from (5) we get

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

Hence, Theorem 3.3 and Lemma 3.6 in [14] are applicable, i.e., by the compactness of the optimal set there exists $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

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

where c_1 and c_2 are positive numbers both independent of μ, and $\gamma_1, \gamma_2 \ge 2^{1-n}$. We get the result by setting $\gamma := \min\{\gamma_1, \gamma_2\}$ and $c := \max\{c_1, c_2\}$. □

Remark 3. The exponents γ_1 and γ_2 depend on the degree of singularity [14] of the minimal subspaces which contain the primal and dual optimal sets, respectively. We are not aware of any upper bound on c. We refer the reader to Section 4 in [14] for further details.

Now, consider the eigenvalue decompositions $X(\mu) = Q(\mu)\Lambda(X(\mu))Q^T(\mu)$ and $S(\mu) = Q(\mu)\Lambda(S(\mu))Q^T(\mu)$, where $Q(\mu)$ is a common eigenvector basis. In the sequel, we prove that as $\mu \to 0$, the eigenvalues of $X(\mu)$ and $S(\mu)$ fall into three categories: 1) $\lambda_{[i]}(X(\mu))$ converges to a positive value and $\lambda_{[n-i+1]}(S(\mu))$ converges to 0; 2) $\lambda_{[i]}(S(\mu))$ converges to a positive value and $\lambda_{[n-i+1]}(X(\mu))$ converges to 0; 3) both $\lambda_{[i]}(X(\mu))$ and $\lambda_{[n-i+1]}(S(\mu))$ converge to 0, where the i^{th} largest eigenvalue of $X(\mu)$ and the i^{th} smallest eigenvalue of $S(\mu)$ have an identical eigenvector. Associated with the above categories of eigenvalues are $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$ as the respective subsets of columns of $Q(\mu)$. It is well-known that for a given sequence $\{\mu_t\}$, the accumulation points of $Q_{\mathcal{B}}(\mu_t)$, $Q_{\mathcal{T}}(\mu_t)$, and $Q_{\mathcal{N}}(\mu_t)$ form orthonormal bases for the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$, respectively, see e.g., Section 3.3 in [2]. Consequently, a sufficiently small μ allows for the identification of $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$. Theorem 3 presents lower and upper bounds for the eigenvalues of $X(\mu)$ and $S(\mu)$ using the condition number σ and the error bounds in (6).

Theorem 3. *Let a central solution $(X(\mu), y(\mu), S(\mu))$ with $\mu \le \hat{\mu}$ be given. Then we have*

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

Furthermore, if μ is small enough so that

$$\mu < \tilde{\mu} := \min\left\{ \frac{1}{n}\left(\frac{\sigma}{cn^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2},\ \hat{\mu} \right\}, \tag{7}$$

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

Proof. We first derive bounds for the vanishing blocks of $\hat{X}(\mu)$ and $\hat{S}(\mu)$ defined in (3). It follows from the compactness of the optimal set and the continuity of the eigenvalues that a primal-dual optimal solution $(\bar{X}, \bar{y}, \bar{S}) \in \mathcal{P}^* \times \mathcal{D}^*$ exists such that $\sigma_{\mathcal{B}} = \lambda_{\min}(Q^T_{\mathcal{B}} \bar{X} Q_{\mathcal{B}})$ and $\sigma_{\mathcal{N}} = \lambda_{\min}(Q^T_{\mathcal{N}} \bar{S} Q_{\mathcal{N}})$. Therefore, we have

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

where $U_{\bar{X}} = Q^T_{\mathcal{B}} \bar{X} Q_{\mathcal{B}}$ and $U_{\bar{S}} = Q^T_{\mathcal{N}} \bar{S} Q_{\mathcal{N}}$. By (4) and the optimality of $\bar{X}$ and $\bar{S}$ we have

$$\begin{aligned} &\langle X(\mu), \bar{S} \rangle + \langle \bar{X}, S(\mu) \rangle \\ &= \langle Q^T_{\mathcal{N}} X(\mu) Q_{\mathcal{N}}, U_{\bar{S}} \rangle + \langle U_{\bar{X}}, Q^T_{\mathcal{B}} S(\mu) Q_{\mathcal{B}} \rangle = n\mu. \end{aligned}$$

Since $X(\mu) \succ 0$, it is immediate that

$$\lambda_{\min}(U_{\bar{S}}) \operatorname{tr}\left(\hat{X}_{\mathcal{N}}(\mu)\right) \le \langle \hat{X}_{\mathcal{N}}(\mu), U_{\bar{S}}\rangle \le n\mu,$$

which together with (8) implies $\operatorname{tr}\left(\hat{X}_{\mathcal{N}}(\mu)\right) \le n\mu/\sigma$. Further, by Lemma 2 and Theorem 1 there exists $(X_\mu, y_\mu, S_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$ so that (6) holds for $\mu \le \hat{\mu}$ and that $X_\mu = Q_{\mathcal{B}} U_{X_\mu} Q_{\mathcal{B}}^T$ for unique $U_{X_\mu} \succeq 0$. All this gives

$$\frac{1}{\sqrt{n}} \operatorname{tr}\left(Q_{\mathcal{T}\cup\mathcal{N}}^T X(\mu) Q_{\mathcal{T}\cup\mathcal{N}}\right) \le \left\|Q_{\mathcal{T}\cup\mathcal{N}}^T X(\mu) Q_{\mathcal{T}\cup\mathcal{N}}\right\| \le \left\|X(\mu) - X_\mu\right\| \le c(n\mu)^\gamma.$$

In a similar fashion, we can derive upper bounds on $\operatorname{tr}\left(\hat{S}_{\mathcal{B}}(\mu)\right)$ and $\operatorname{tr}\left(Q_{\mathcal{B}\cup\mathcal{T}}^T S(\mu) Q_{\mathcal{B}\cup\mathcal{T}}\right)$.

In the sequel, taking into account the centrality condition

$$\lambda_{[i]}(X(\mu))\lambda_{[n-i+1]}(S(\mu)) = \mu,$$

and the application of Theorem 4.5 in [13](i.e., for $X \in \mathbb{S}^n$ and $Y \in \mathbb{R}^{n\times k}$ we have $\lambda_{[n-k+1]}(X) + \ldots + \lambda_{[n]}(X) = \min\{\operatorname{tr}(Y^TXY) \mid Y^TY = I_n\}$) to $\operatorname{tr}\left(\hat{S}_{\mathcal{B}}(\mu)\right)$, $\operatorname{tr}\left(\hat{X}_{\mathcal{N}}(\mu)\right)$, $\operatorname{tr}\left(Q_{\mathcal{T}\cup\mathcal{N}}^T X(\mu) Q_{\mathcal{T}\cup\mathcal{N}}\right)$, and $\operatorname{tr}\left(Q_{\mathcal{B}\cup\mathcal{T}}^T S(\mu) Q_{\mathcal{B}\cup\mathcal{T}}\right)$ yield the lower and upper bounds. For instance, we get

$$\lambda_{[n-n_{\mathcal{B}}+1]}(S(\mu)) + \ldots + \lambda_{[n]}(S(\mu)) \le \operatorname{tr}\left(\hat{S}_{\mathcal{B}}(\mu)\right) \le \frac{n\mu}{\sigma},$$

which gives upper and lower bounds on the $n_{\mathcal{B}}$ smallest and the $n_{\mathcal{B}}$ largest eigenvalues of $S(\mu)$ and $X(\mu)$, respectively:

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

Finally, it can be deducted from the bounds that as $\mu \to 0$, the $n_{\mathcal{B}}$ and $n_{\mathcal{N}}$ largest eigenvalues of $X(\mu)$ and $S(\mu)$ remain positive while the $n_{\mathcal{B}}$ and $n_{\mathcal{N}}$ smallest eigenvalues of $S(\mu)$ and $X(\mu)$ converge to 0, respectively. Additionally, if $n_{\mathcal{T}} > 0$, then there exist a set of $n_{\mathcal{T}}$ eigenvalues of $X(\mu)$ and $S(\mu)$ both belonging to the interval $I := \left[\mu/c\sqrt{n}(n\mu)^\gamma,\ c\sqrt{n}(n\mu)^\gamma\right]$. Thus, we can identify $Q_{\mathcal{B}}(\mu)$, $Q_{\mathcal{T}}(\mu)$, and $Q_{\mathcal{N}}(\mu)$ when $\mu \le \hat{\mu}$ and the intervals I, $(0, n\mu/\sigma]$, and $[\sigma/n, \infty)$ are disjoint, i.e., when

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

which gives the upper bound (7). □

Remark 4. Theorem 3 is used in Section 3 to derive bounds for the feasibility of an approximate MC solution. Throughout the following sections, $\mathcal{R}\big(Q_{\mathcal{B}}(\mu)\big)$, $\mathcal{R}\big(Q_{\mathcal{T}}(\mu)\big)$, and $\mathcal{R}\big(Q_{\mathcal{N}}(\mu)\big)$ with $\mu < \tilde{\mu}$ are referred to as approximations of the subspaces $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$, respectively.

3. A rounding procedure for central solutions

We present a rounding procedure which can be viewed as a generalization of the approach in [12]. Even though only approximations of $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$ are available, from a central solution with sufficiently small μ we can make a projection onto the boundary of the positive semidefinite cone to generate a complementary solution with approximate primal-dual feasibility. Such solutions are called approximate MC solutions.

Suppose that a central solution $(X(\mu), y(\mu), S(\mu))$ and $Q(\mu) := (Q_{\mathcal{B}}(\mu), Q_{\mathcal{T}}(\mu), Q_{\mathcal{N}}(\mu))$ are given, where $\mu < \tilde{\mu}$ as defined in Theorem 3. Let (X^*, y^*, S^*) be an MC solution and define $\hat{X}^* := Q^T(\mu)X^*Q(\mu)$, $\hat{S}^* := Q^T(\mu)S^*Q(\mu)$, and $\hat{A}^i := Q^T(\mu)A^iQ(\mu)$. Then from the primal feasibility constraints we have

$$\langle \hat{A}^i, \hat{X}^*\rangle = b_i, \qquad i = 1, \ldots, m, \tag{9}$$
$$\big\langle \hat{A}^i, \Lambda(X(\mu))\big\rangle = b_i, \qquad i = 1, \ldots, m. \tag{10}$$

By subtracting (10) from (9) for each i, for $\Delta X_{\mathcal{B}}(\mu) = \hat{X}^*_{\mathcal{B}} - \Lambda_{\mathcal{B}}(X(\mu))$ we get

$$\begin{aligned}\langle \hat{A}^i_{\mathcal{B}}, \Delta X_{\mathcal{B}}(\mu)\rangle &= \big\langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\big\rangle \\ &\quad + \big\langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\big\rangle + \xi_i,\end{aligned} \tag{11}$$

where the residual term ξ_i is given by

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

Analogously, let $\hat{C} = Q^T(\mu)CQ(\mu)$. Then we get

$$\sum_{i=1}^m y_i^* \hat{A}^i + \hat{S}^* = \hat{C}, \tag{12}$$
$$\sum_{i=1}^m y_i(\mu)\hat{A}^i + \Lambda(S(\mu)) = \hat{C}. \tag{13}$$

By subtracting (13) from (12), 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

$$\begin{aligned}&\sum_{i=1}^m \Delta y_i(\mu)\hat{A}^i + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S_{\mathcal{N}}(\mu)\end{bmatrix} \\ &= \begin{bmatrix} \Lambda_{\mathcal{B}}(S(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S(\mu)) & 0 \\ 0 & 0 & 0\end{bmatrix} - \begin{bmatrix} \hat{S}^*_{\mathcal{B}} & \hat{S}^*_{\mathcal{BT}} & \hat{S}^*_{\mathcal{BN}} \\ \hat{S}^*_{\mathcal{TB}} & \hat{S}^*_{\mathcal{T}} & \hat{S}^*_{\mathcal{TN}} \\ \hat{S}^*_{\mathcal{NB}} & \hat{S}^*_{\mathcal{NT}} & 0\end{bmatrix}.\end{aligned} \tag{14}$$

Both the right hand sides in (11) and (14) depend on the chosen MC solution. Thus, the systems (11) and (14) may not be solvable if we drop the unknown terms. Instead, we can solve two least square problems to obtain search directions towards primal and dual solutions.

Remark 5. For an LO problem embedded in SDO, $X(\mu)$ and $S(\mu)$ are diagonal matrices. When the optimal partition is known, the coordinates of variables can be rearranged to get $Q(\mu) = I_n$. All this implies that the unknown terms in (11) and (14) are just zero for LO.

3.1. Primal least square problem

For the primal problem we solve

$$\begin{aligned}\min \quad & \|\Delta X\|^2 + \|e\|_2^2 \\ \text{s.t.} \quad & \langle \hat{A}^i_{\mathcal{B}}, \Delta X\rangle - e_i = \big\langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\big\rangle \\ & \qquad + \big\langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\big\rangle, \quad i = 1, \ldots, m.\end{aligned} \tag{15}$$

We may assume that $\hat{A}^i_{\mathcal{B}} \neq 0$ for some i. Otherwise, the optimal solution of (15) would give $\Delta X^* = 0$, and thus the effect of the vanishing terms is absorbed in primal infeasibility. For LO we have $\hat{A}^i_{\mathcal{B}} \neq 0$ for some i, since otherwise the primal optimal solution would be trivial.

The optimal solution $(\Delta X^*, e^*)$ to the auxiliary problem (15) yields $\tilde{X}_{\mathcal{B}} := \Lambda_{\mathcal{B}}(X(\mu)) + \Delta X^*$ so that

$$\langle \hat{A}^i_{\mathcal{B}}, \tilde{X}_{\mathcal{B}}\rangle = b_i + e_i^*, \qquad i = 1, \ldots, m.$$

Thus, $\tilde{X}_{\mathcal{B}}$ is ϵ_p infeasible for the primal constraints with

$$\epsilon_p := \|e^*\|_2.$$

Let $r(n) := n(n+1)/2$ and define

$$\hat{\mathcal{A}}^s_{\mathcal{B}} := \left(\text{svec}(\hat{A}^1_{\mathcal{B}}), \ldots, \text{svec}(\hat{A}^m_{\mathcal{B}})\right)^T,$$

$$\mathcal{A} := \left(\text{vec}(A^1), \ldots, \text{vec}(A^m)\right)^T,$$

in which vec : $\mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ is the concatenation of the columns of a matrix. Then problem (15) reduces to

$$\min \quad \|\Delta x\|_2^2 + \|\hat{\mathcal{A}}^s_{\mathcal{B}} \Delta x - \eta\|_2^2, \tag{16}$$

where $\Delta x = \text{svec}(\Delta X)$, and $\eta_i = \langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\rangle + \langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\rangle$ denotes the vanishing term, which should be zero for all optimal solutions. Lemma 4 establishes upper bounds on $\|\Delta X^*\|$ and $\|e^*\|_2$. The bounds depend on the constant

$$\pi_p := \prod_{k=1}^{r(n_{\mathcal{B}})} \left\| \left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)_{.k} \right\|_2,$$

where $\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)_{.k}$ denotes the k^{th} column of $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$. Using the upper bounds in Lemma 4, Theorem 6 proves that $\tilde{X}_{\mathcal{B}} \succ 0$ for sufficiently small μ.

Lemma 4. *Let* $(\Delta X^*, e^*)$ *be the unique optimal solution of* (15)*. Then we have*

$$\|\Delta X^*\| \le 2\pi_p \sqrt{r(n_{\mathcal{B}})} \|\mathcal{A}\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{n n_{\mathcal{T}}}(n\mu)^\gamma\right\},$$

$$\epsilon_p \le 2\|\mathcal{A}\| \left(\pi_p \sqrt{r(n_{\mathcal{B}})} \|\mathcal{A}\|^2 + 1\right) \times \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{n n_{\mathcal{T}}}(n\mu)^\gamma\right\}.$$

Proof. The optimality conditions for (16) are given by

$$\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)\Delta x = (\hat{\mathcal{A}}^s_{\mathcal{B}})^T \eta, \tag{17}$$

where $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})} \succ 0$. The unique solution of (17) can be computed using Cramer's rule [8]:

$$\Delta x^*_j = \frac{\det\left(\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)^{(j)}\right)}{\det\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)}, \quad j = 1, \ldots, r(n_{\mathcal{B}}),$$

in which the matrix $\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)^{(j)}$ in the nominator is obtained by substituting the j^{th} column of $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}$ by $(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \eta$. Noting that $\det\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right) \ge 1$, we can deduce from Hadamard's inequality [8] that for $j = 1, \ldots, r(n_{\mathcal{B}})$

$$\begin{aligned} |\Delta x^*_j| &\le \left|\det\left(\left((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})}\right)^{(j)}\right)\right| \\ &\le \left\|(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \eta\right\|_2 \prod_{\substack{k=1\\k\ne j}}^{r(n_{\mathcal{B}})} \left\|((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})})_{.k}\right\|_2 \end{aligned}$$

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

$$\begin{aligned} |\Delta x^*_j| &\le \left\|(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \eta\right\|_2 \prod_{\substack{k=1\\k\ne j}}^{r(n_{\mathcal{B}})} \left\|((\hat{\mathcal{A}}^s_{\mathcal{B}})^T \hat{\mathcal{A}}^s_{\mathcal{B}} + I_{r(n_{\mathcal{B}})})_{.k}\right\|_2 \\ &\le \pi_p \|(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \eta\|_2. \end{aligned} \tag{18}$$

Then we can conclude from Theorem 3 that

$$\left|\langle \hat{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu))\rangle\right| \le \frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma} \|A^i\|,$$

$$\left|\langle \hat{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu))\rangle\right| \le c\sqrt{n n_{\mathcal{T}}}(n\mu)^\gamma \|A^i\|$$

for $i = 1, \ldots, m$, which yields the upper bound

$$|\eta_i| \le 2\|A^i\| \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{n n_{\mathcal{T}}}(n\mu)^\gamma\right\}. \tag{19}$$

Consequently, from (18) and (19) it follows that

$$\begin{aligned} |\Delta x^*_j| &\le \pi_p \|(\hat{\mathcal{A}}^s_{\mathcal{B}})^T \eta\|_2 \\ &\le 2\pi_p \|\mathcal{A}\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{n n_{\mathcal{T}}}(n\mu)^\gamma\right\} \end{aligned}$$

for $j = 1, \ldots, r(n_{\mathcal{B}})$, where we have used the inequality $\|\hat{\mathcal{A}}^s_{\mathcal{B}}\| \le \|\mathcal{A}\|$. As a result, we get

$$\begin{aligned} \|e^*\|_2 &= \|\hat{\mathcal{A}}^s_{\mathcal{B}} \Delta x^* - \eta\|_2 \le \|\hat{\mathcal{A}}^s_{\mathcal{B}}\| \|\Delta x^*\|_2 + \|\eta\|_2 \\ &\le 2\|\mathcal{A}\| \left(\pi_p \sqrt{r(n_{\mathcal{B}})} \|\mathcal{A}\|^2 + 1\right) \\ &\quad \times \max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{n n_{\mathcal{T}}}(n\mu)^\gamma\right\}. \end{aligned}$$

This completes the proof. □

3.2. Dual least square problem

Let E denote a slack matrix as

$$E := \begin{bmatrix} E_{\mathcal{B}} & E_{\mathcal{BT}} & E_{\mathcal{BN}} \\ E_{\mathcal{TB}} & E_{\mathcal{T}} & E_{\mathcal{TN}} \\ E_{\mathcal{NB}} & E_{\mathcal{NT}} & 0 \end{bmatrix}, \tag{20}$$

which is defined in accordance with the unknown right hand side matrix in (14). Then the auxiliary problem for an approximate dual solution is formulated as

$$\begin{aligned} \min \quad & \|\Delta y\|_2^2 + \|\Delta S\|^2 + \|E\|^2 \\ \text{s.t.} \quad & \sum_{i=1}^m \Delta y_i \hat{A}^i + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S \end{bmatrix} - E \\ & = \begin{bmatrix} \Lambda_{\mathcal{B}}(S(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S(\mu)) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{21}$$

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

$$\epsilon_d := \|E^*\|.$$

For the sake of clarity, in what follows, auxiliary problem (21) is represented in vector form. To do so, analogous to the definition of $\mathcal{A}$, we apply the vec(.) operator to each block of $\hat{A}_i$ to form $\hat{\mathcal{A}}_{\mathcal{B}}$, $\hat{\mathcal{A}}_{\mathcal{N}}$, $\hat{\mathcal{A}}_{\mathcal{T}}$, $\hat{\mathcal{A}}_{\mathcal{BT}}$, $\hat{\mathcal{A}}_{\mathcal{BN}}$, and $\hat{\mathcal{A}}_{\mathcal{TN}}$. Thus, problem (21) reduces to the least square problem

$$\begin{aligned} \min \quad & \|\Delta y\|_2^2 + \left\|\hat{\mathcal{A}}^T_{\mathcal{N}} \Delta y\right\|_2^2 + \left\|\hat{\mathcal{A}}^T_{\mathcal{B}} \Delta y - \zeta_{\mathcal{B}}\right\|_2^2 \\ & + \left\|\hat{\mathcal{A}}^T_{\mathcal{T}} \Delta y - \zeta_{\mathcal{T}}\right\|_2^2 + 2\left\|\hat{\mathcal{A}}^T_{\mathcal{BT}} \Delta y\right\|_2^2 \\ & + 2\left\|\hat{\mathcal{A}}^T_{\mathcal{BN}} \Delta y\right\|_2^2 + 2\left\|\hat{\mathcal{A}}^T_{\mathcal{TN}} \Delta y\right\|_2^2, \end{aligned} \tag{22}$$

where $\zeta_{\mathcal{B}} = \text{vec}\left(\Lambda_{\mathcal{B}}(S(\mu))\right)$ and $\zeta_{\mathcal{T}} = \text{vec}\left(\Lambda_{\mathcal{T}}(S(\mu))\right)$. Lemma 5 establishes upper bounds on ϵ_d and $\|\Delta S^*\|$. For the upper bounds we define the positive definite matrix

$$\begin{aligned} \mathcal{H} := \; & \hat{\mathcal{A}}_{\mathcal{B}} \hat{\mathcal{A}}^T_{\mathcal{B}} + \hat{\mathcal{A}}_{\mathcal{T}} \hat{\mathcal{A}}^T_{\mathcal{T}} + \hat{\mathcal{A}}_{\mathcal{N}} \hat{\mathcal{A}}^T_{\mathcal{N}} + 2\hat{\mathcal{A}}_{\mathcal{BT}} \hat{\mathcal{A}}^T_{\mathcal{BT}} \\ & + 2\hat{\mathcal{A}}_{\mathcal{BN}} \hat{\mathcal{A}}^T_{\mathcal{BN}} + 2\hat{\mathcal{A}}_{\mathcal{TN}} \hat{\mathcal{A}}^T_{\mathcal{TN}} + I_m \end{aligned}$$

and constant

$$\pi_d := \prod_{k=1}^{m} \|\mathcal{H}_{.k}\|_2,$$

where $\mathcal{H}_{.k}$ denotes the k^{th} column of $\mathcal{H}$. Theorem 6 proves that for sufficiently small μ we have $\tilde{S}_{\mathcal{N}} \succ 0$.

Lemma 5. *Problem (21) has a unique optimal solution $(\Delta y^*, \Delta S^*, E^*)$, which satisfies*

$$\|\Delta S^*\| \le 2\pi_d\sqrt{m}\|\mathcal{A}\|^2 \max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\},$$

$$\epsilon_d \le \sqrt{2}(4\pi_d\sqrt{m}\|\mathcal{A}\|^2 + 1)\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}.$$

Proof. For the sake of simplicity we define

$$\varphi := \hat{\mathcal{A}}_{\mathcal{B}}\zeta_{\mathcal{B}} + \hat{\mathcal{A}}_{\mathcal{T}}\zeta_{\mathcal{T}}.$$

The optimality conditions for (22) can be written as $\mathcal{H}\Delta y = \varphi$, where $\mathcal{H} \succ 0$. The unique solution of this system can be computed by using Cramer's rule as follows

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

where matrix $\mathcal{H}^{(i)}$ in the nominator is obtained by substituting the i^{th} column of $\mathcal{H}$ by φ. Therefore, for $i = 1, \dots, m$, using $\det(\mathcal{H}) \ge 1$ we get

$$|\Delta y_i^*| \le |\det(\mathcal{H}^{(i)})| \le \|\varphi\|_2 \prod_{\substack{k=1\\k\neq i}}^{m} \|\mathcal{H}_{.k}\|_2 \le \pi_d\|\varphi\|_2,$$

where the second inequality follows from Hadamard's inequality. Note that $\prod_{k=1,k\neq i}^m \|\mathcal{H}_{.k}\|_2 \le \pi_d$, since the diagonal entries of $\mathcal{H}$ are at least 1. Furthermore, from Theorem 3 we get

$$\begin{aligned}\|\varphi\|_2 &\le \|\hat{\mathcal{A}}_{\mathcal{B}}\zeta_{\mathcal{B}}\|_2 + \|\hat{\mathcal{A}}_{\mathcal{T}}\zeta_{\mathcal{T}}\|_2\\ &\le 2\|\mathcal{A}\|\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\},\end{aligned}$$

which leads to

$$|\Delta y_i^*| \le 2\pi_d\|\mathcal{A}\|\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}. \tag{23}$$

Consequently, from (23) we have

$$\begin{aligned}\|\Delta S^*\| &= \|(\hat{\mathcal{A}}_{\mathcal{N}})^T\Delta y^*\|_2\\ &\le 2\pi_d\sqrt{m}\|\mathcal{A}\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}.\end{aligned}$$

Then we derive bounds on the components of E^* as follows

$$\begin{aligned}\|E^*_{\mathcal{B}}\| &= \|\hat{\mathcal{A}}_{\mathcal{B}}^T\Delta y^* - \zeta_{\mathcal{B}}\|_2\\ &\le 2\pi_d\sqrt{m}\|\mathcal{A}\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\} + \frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}\\ &\le (2\pi_d\sqrt{m}\|\mathcal{A}\|^2 + 1)\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\},\end{aligned}$$

$$\begin{aligned}\|E^*_{\mathcal{T}}\| &= \|\hat{\mathcal{A}}_{\mathcal{T}}^T\Delta y^* - \zeta_{\mathcal{T}}\|_2\\ &\le 2\pi_d\sqrt{m}\|\mathcal{A}\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}\\ &\quad + c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\\ &\le (2\pi_d\sqrt{m}\|\mathcal{A}\|^2 + 1)\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\},\end{aligned}$$

$$\begin{aligned}&\|E^*_{\mathcal{BT}}\|, \|E^*_{\mathcal{BN}}\|, \|E^*_{\mathcal{TN}}\|\\ &\le 2\pi_d\sqrt{m}\|\mathcal{A}\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma},\ c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\}.\end{aligned}$$

Finally, we get

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

which gives the upper bound on ϵ_d. □

3.3. Cone feasibility

Solving (15) and (21) yields a complementary solution $\left(Q_{\mathcal{B}}(\mu)\tilde{X}_{\mathcal{B}}Q_{\mathcal{B}}^T(\mu), \tilde{y}, Q_{\mathcal{N}}(\mu)\tilde{S}_{\mathcal{N}}Q_{\mathcal{N}}^T(\mu)\right)$ with $\epsilon := \max\{\epsilon_p, \epsilon_d\}$ infeasibility w.r.t. the linear constraints. Theorem 6 shows that for sufficiently small μ, the rounding procedure yields a primal-dual solution with $\tilde{X}_{\mathcal{B}}, \tilde{S}_{\mathcal{N}} \succ 0$.

Theorem 6. *Let $\vartheta_1 := 2n^2\|\mathcal{A}\|^2$, $\vartheta_2 := 2cn^{\frac{3}{2}}\sqrt{n_{\mathcal{T}}}\|\mathcal{A}\|^2$, and*

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

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

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

$$|\lambda_{\min}(\Delta X^*)| \le \|\Delta X^*\|, \qquad |\lambda_{\min}(\Delta S^*)| \le \|\Delta S^*\|,$$

we conclude from Theorem 3, and Lemmas 4 and 5 that

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

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

$$2\pi_p\sqrt{r(n_{\mathcal{B}})}\|\mathcal{A}\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{N}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\} < \frac{\sigma}{n},$$

$$2\pi_d\sqrt{m}\|\mathcal{A}\|^2\max\left\{\frac{n\sqrt{n_{\mathcal{B}}}\mu}{\sigma}, c\sqrt{nn_{\mathcal{T}}}(n\mu)^\gamma\right\} < \frac{\sigma}{n}.$$

□

3.4. Rounding procedure

Now, we can outline a simple procedure which yields an approximate MC solution.

Algorithm 1 Rounding procedure for SDO

Input $(X(\mu), y(\mu), S(\mu))$, where $\mu < \mu^r$ is fixed.
Do Solve least square problem (15) to get $\tilde{X}_{\mathcal{B}}$.
Solve least square problem (21) to get $(\tilde{y}, \tilde{S}_{\mathcal{N}})$.
Return $(Q_{\mathcal{B}}(\mu)\tilde{X}_{\mathcal{B}}Q_{\mathcal{B}}^T(\mu), \tilde{y}, Q_{\mathcal{N}}(\mu)\tilde{S}_{\mathcal{N}}Q_{\mathcal{N}}^T(\mu))$.

Even though Algorithm 1 produces an approximate MC solution and needs the eigenvalue decomposition of a central solution, it relies on solving two linear systems of equations with better conditioned coefficient matrices than the Jacobian of the Newton system.

Remark 6. For a fixed μ the orthogonal matrix $Q(\mu)$, due to multiplicity of the eigenvalues, may not be unique. Then the solution given by Algorithm 1 varies with the choice of $Q(\mu)$. Obviously, this cannot happen for LO.

Remark 7. As indicated in Remark 5, Algorithm 1 can be inferred as an extension of the method in [12]. Computing an ϵ-feasible MC solution requires $\mathcal{O}(\max\{n_{\mathcal{B}}^6, m^3\})$ arithmetic operations. In fact, this is equivalent to solving two linear systems of equations, using the Gauss elimination method, with $r(n_{\mathcal{B}})$ and m variables, respectively.

4. Concluding remarks and conclusion

4.1. A rounding procedure for approximate solutions

We recall from [10] that the identification results can be extended for approximate solutions, which is best suited for primal-dual IPMs, since they generate a sequence of interior solutions in a neighborhood of the central path. This motivates the extension of the rounding procedure for solutions in a neighborhood of the central path.

Let $(X^\circ, y^\circ, S^\circ)$ be an interior solution, and consider the eigenvalue decompositions $X^\circ = M\Lambda(X^\circ)M^T$ and $S^\circ = P\Lambda(S^\circ)P^T$, where M and P are orthogonal matrices. Further, let $M := (M_{\mathcal{B}}, M_{\mathcal{T}}, M_{\mathcal{N}})$ and $P := (P_{\mathcal{B}}, P_{\mathcal{T}}, P_{\mathcal{N}})$, where the subsets of columns of M and P correspond to the optimal partition. When $\langle X^\circ, S^\circ\rangle$ is sufficiently small, we can identify $M_{\mathcal{B}}$, $M_{\mathcal{T}}$, and $M_{\mathcal{N}}$ from X°, and $P_{\mathcal{B}}$, $P_{\mathcal{T}}$, and $P_{\mathcal{N}}$ from S°. Then we need to choose the eigenvectors either from M or P, because X° and S° do not necessarily commute. The rest of the procedure follows, e.g., by replacing $X(\mu)$ and $Q(\mu)$ by X° and M in the definition of $\hat{A}^i$, $\hat{C}$, and the right hand sides in (15) and (21).

4.2. Conclusion and future studies

In this paper, we used the approximation of the optimal partition obtained from a sequence of bounded interior solutions, on or in a neighborhood of the central path, to generate an ϵ-feasible primal-dual solution with zero complementarity gap. It is proven that if the complementarity gap drops below a certain bound, then both the primal and dual solutions satisfy the cone constraints, yielding an approximate MC solution.

It is worth investigating the sensitivity of the invariant subspaces spanned by the columns of $Q_{\mathcal{B}}(\mu)$ and $Q_{\mathcal{N}}(\mu)$ w.r.t. the perturbation of right hand side and objective vectors. Parametric analysis of SDO and SOCO problems will be studied in a forthcoming paper.

Acknowledgements

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

References

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

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

[3] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Oper. Res. Lett.*, 20(5):213–221, 1997.

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

[5] A. Goldman and A. Tucker. Theory of linear programming. In H. Kuhn and A. Tucker, editors, *Linear Equalities and Related Systems*, pages 53–97. Princeton University Press, 1956.

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

[7] A. Hoffman. On approximate solutions of systems of linear inequalities. *J. Res. Natl. Bur. Stand.*, 49(4):263–265, 1952.

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

[9] 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 J. Optim.*, 11(2):320–340, 2000.

[10] A. Mohammad-Nezhad and T. Terlaky. On the identification of optimal partition for semidefinite optimization. Technical Report 17T-008, Dept. of ISE, Lehigh U., 2017. `http://www.optimization-online.org/DB_HTML/2016/09/5626.html`.

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

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

[13] G. W. Stewart and J.-g. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

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

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

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

[17] E. Yildirim. Unifying optimal partition approach to sensitivity analysis in conic optimization. *J. Optim. Theory Appl.*, 122(2):405–423, 2004.