A primal-dual interior-point method based on various selections of displacement step for second-order cone programming

Baha Alzalg[†*]

[†]*Department of Mathematics, The University of Jordan, Amman 11942, Jordan*
Submission date: June 1, 2017

Abstract In this paper, a primal-dual interior-point method equipped with various selections of the displacement step are derived for solving second-order cone programming problems. We first establish the existence and uniqueness of the optimal solution of the corresponding perturbed problem and then demonstrate its convergence to the optimal solution of the original problem. Next, we present four different selections to calculate the displacement step. We also establish the convergence of the proposed algorithm and present its complexity result. Finally, the four selections of calculating the displacement step are compared in numerical examples to show the efficiency of the proposed algorithm.

Keywords Second-order cone programming, interior-point methods, primal-dual methods, central trajectory methods, Jordan algebras
AMS Classification 90C30, 90C46, 90C51, 17A15

1 Introduction

Interior-point methods [1–14] are considered one of the effective methods developed for solving large and important classes of conic programs, such as linear programming, second-order cone programming (SOCP) and semidefinite programming. In particular, primal-dual interior-point methods [14–16] are among the most effective known methods for solving conic programming problems.

We can separate two types of interior-point methods: projective methods [17–20] and (feasible or infeasible) central trajectory methods [21–23]. Touil et al. [1] proposed feasible central trajectory methods for semidefinite programming. The purpose of this paper is to exploit the algebraic structure of the second-order cones to derive a feasible primal-dual central trajectory algorithm for SOCP by utilizing the work of Touil et al. [1] for semidefinite programming.

The general scheme of the central trajectory methods is as follows. We associate a perturbed problem to the SOCP problem. Then, we draw a path of the centers defined by the perturbed KKT optimality conditions. After that, Newton's method is applied to treat the corresponding perturbed equations in order to obtain a descent search direction. Throughout our development, we propose

*Email: b.alzalg@ju.edu,jo
The work of the author was supported in part by the Deanship of Scientific Research at the University of Jordan.

four different selections of calculating the displacement step and establish the complexity analysis of the proposed algorithm.

We introduce some notations used to write the primal and dual forms of SOCP and the perturbed problem. We use "," for adjoining scalars, vectors and matrices in a row, and use ";" for adjoining them in a column. So, for instance, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors, we have

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = (\boldsymbol{u}^\mathsf{T}, \boldsymbol{v}^\mathsf{T})^\mathsf{T} = (\boldsymbol{u}; \boldsymbol{v}).$$

For each vector $\boldsymbol{v} \in \mathbb{R}^n$ whose first entry is indexed with 0, we write $\widetilde{\boldsymbol{v}}$ for the subvector consisting of entries 1 through $n-1$; therefore $\boldsymbol{v} = (v_0; \widetilde{\boldsymbol{v}}) \in \mathbb{R} \times \mathbb{R}^{n-1}$. We denote by $\mathcal{E}^n$ the n^{th}-dimensional real vector space $\mathbb{R} \times \mathbb{R}^{n-1}$ whose elements $\boldsymbol{v}$ are indexed from 0. The n^{th}-dimensional *second-order cone* is defined as

$$\mathcal{E}_+^n := \{\boldsymbol{v} \in \mathcal{E}^n : v_0 \geq \|\widetilde{\boldsymbol{v}}\|\},$$

where $\|\cdot\|$ denotes the Euclidean norm. The set $\text{int}\, \mathcal{E}_+^n := \{\boldsymbol{v} \in \mathcal{E}^n : v_0 > \|\widetilde{\boldsymbol{v}}\|\}$ is called the *interior* of $\mathcal{E}_+^n$. The *determinant* of $\boldsymbol{v} \in \mathcal{E}^n$ is defined as $\det(\boldsymbol{v}) := v_0^2 - \|\bar{v}\|^2$. Note that $\boldsymbol{v} \in \mathcal{E}_+^n$ ($\boldsymbol{v} \in \text{int}\, \mathcal{E}_+^n$) if and only if $\det(\boldsymbol{v}) \geq 0$ ($\det(\boldsymbol{v}) > 0$). We define the *logarithmic barrier function* $\ell(\cdot)$ on $\mathcal{E}_+^n$ as

$$\ell(\boldsymbol{v}) := \begin{cases} -\ln\det(\boldsymbol{v}), & \text{if } \boldsymbol{v} \in \text{int}\, \mathcal{E}_+^n; \\ +\infty, & \text{otherwise.} \end{cases}$$

Let $r \geq 1$ be an integer. For each $i = 1, 2, \ldots, r$, let n and n_i be positive integers such that $n = \sum_{i=1}^r n_i$. Let $\mathcal{E}_{r+}^n$ denote the Cartesian product of r second-order cones $\mathcal{E}_+^{n_1}, \mathcal{E}_+^{n_2}, \ldots, \mathcal{E}_+^{n_r}$. That is, $\mathcal{E}_{r+}^n := \mathcal{E}_+^{n_1} \times \mathcal{E}_+^{n_2} \times \cdots \times \mathcal{E}_+^{n_r}$. Similarly, we have $\text{int}\, \mathcal{E}_{r+}^n := \text{int}\, \mathcal{E}_+^{n_1} \times \text{int}\, \mathcal{E}_+^{n_2} \times \cdots \times \text{int}\, \mathcal{E}_+^{n_r}$ and $\mathcal{E}_r^n := \mathcal{E}^{n_1} \times \mathcal{E}^{n_2} \times \cdots \times \mathcal{E}^{n_r}$. We define the *logarithmic barrier function* $\ell(\cdot)$ on $\mathcal{E}_{r+}^n$ as

$$\ell(\boldsymbol{v}) := \begin{cases} -\displaystyle\sum_{i=1}^{r} \ln\det(\boldsymbol{v}_i), & \text{if } \boldsymbol{v} := (\boldsymbol{v}_1; \boldsymbol{v}_2; \ldots; \boldsymbol{v}_r) \in \text{int}\, \mathcal{E}_+^n; \\ +\infty, & \text{otherwise.} \end{cases}$$

Now, we are ready to write the primal and dual forms of SOCP and the perturbed problem. Let $\boldsymbol{y}, \boldsymbol{b} \in \mathbb{R}^m$. Let also $\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{z} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$ be such that they are all conformally partitioned as

$$\boldsymbol{x} := (\boldsymbol{x}_1; \boldsymbol{x}_2; \ldots; \boldsymbol{x}_r),\ \boldsymbol{s} := (\boldsymbol{s}_1; \boldsymbol{s}_2; \ldots; \boldsymbol{s}_r),\ \boldsymbol{c} := (\boldsymbol{c}_1; \boldsymbol{c}_2; \ldots; \boldsymbol{c}_r), \text{ and } A := (A_1, A_2, \ldots, A_r),$$

where $\boldsymbol{x}_i, \boldsymbol{s}_i, \boldsymbol{c}_i \in \mathcal{E}^{n_i}$ and $A_i \in \mathbb{R}^{m_i \times n_i}$ for $i = 1, 2, \ldots, r$. The *SOCP problem and its dual* in multi-block structures are defined as

$$(P_i) \quad \begin{array}{ll} \min & \boldsymbol{c}_1^\mathsf{T}\boldsymbol{x}_1 + \cdots + \boldsymbol{c}_r^\mathsf{T}\boldsymbol{x}_r \\ \text{s.t.} & A_1\boldsymbol{x}_1 + \cdots + A_r\boldsymbol{x}_r = \boldsymbol{b}, \\ & \boldsymbol{x}_i \in \mathcal{E}_+^{n_i},\ i = 1, \ldots, r; \end{array} \qquad (D_i) \quad \begin{array}{ll} \max & \boldsymbol{b}^\mathsf{T}\boldsymbol{y} \\ \text{s.t.} & A_i{}^\mathsf{T}\boldsymbol{y} + \boldsymbol{s}_i = \boldsymbol{c}_i,\ i = 1, \ldots, r, \\ & \boldsymbol{s}_i \in \mathcal{E}_+^{n_i},\ \ i = 1, \ldots, r. \end{array}$$

We can rewrite the pair (P_i, D_i) compactly as

$$(P) \quad \begin{array}{ll} \min & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \succeq \boldsymbol{0}; \end{array} \qquad (D) \quad \begin{array}{ll} \max & \boldsymbol{b}^\mathsf{T}\boldsymbol{y} \\ \text{s.t.} & A^\mathsf{T}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \\ & \boldsymbol{s} \succeq \boldsymbol{0}, \end{array}$$

where $\boldsymbol{x}, \boldsymbol{s} \succeq \boldsymbol{0}$ means that $\boldsymbol{x}, \boldsymbol{s} \in \mathcal{E}_{r+}^n$. We can similarly write $\boldsymbol{x}, \boldsymbol{s} \succ \boldsymbol{0}$ to mean that $\boldsymbol{x}, \boldsymbol{s} \in \text{int}\, \mathcal{E}_{r+}^n$.

Now, we define the following feasibility sets:

$$
\begin{array}{rcl}
\mathcal{F}_P & := & \left\{ \boldsymbol{x} \in \mathcal{E}_r^n : A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \succeq \boldsymbol{0} \right\}, \\
\mathcal{F}_P^\circ & := & \left\{ \boldsymbol{x} \in \mathcal{E}_r^n : A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \succ \boldsymbol{0} \right\}, \\
\mathcal{F}_D & := & \left\{ (\boldsymbol{y}; \boldsymbol{s}) \in \mathbb{R}^m \times \mathcal{E}_r^n : A^\mathsf{T}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c},\ \boldsymbol{s} \succeq \boldsymbol{0} \right\}, \\
\mathcal{F}_D^\circ & := & \left\{ (\boldsymbol{y}; \boldsymbol{s}) \in \mathbb{R}^m \times \mathcal{E}_r^n : A^\mathsf{T}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c},\ \boldsymbol{s} \succ \boldsymbol{0} \right\}.
\end{array}
$$

Next, we make two assumptions about the primal-dual pair (P, D).

Assumption 1.1. The m rows of the matrix A are linearly independent.

Assumption 1.2. The sets $\mathcal{F}_P^\circ$ and $\mathcal{F}_D^\circ$ are nonempty.

Let $\mu > 0$ be a barrier parameter. Solving problem (P) is equivalent to solving the perturbed problem

$$
\left(P_\mu\right) \qquad \begin{array}{ll} \min & f_\mu(\boldsymbol{x}) := \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \mu\, \ell(\boldsymbol{x}) + 2r\mu \ln \mu \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b}, \end{array}
$$

or to solving the perturbed dual problem

$$
\left(D_\mu\right) \qquad \begin{array}{ll} \max & g_\mu(\boldsymbol{y}, \boldsymbol{s}) := \boldsymbol{b}^\mathsf{T}\boldsymbol{y} - \mu\, \ell(\boldsymbol{s}) - 2r\mu \ln \mu \\ \text{s.t.} & A^\mathsf{T}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}. \end{array}
$$

This paper is organized as follows. In Section 2, we briefly review the basics of the Jordan algebra associated with the second-order cones and provide some preliminary results. In Section 3, we first consider the existence and uniqueness of the optimal solution of Problem (P_μ), then we prove the convergence of Problem (P_μ) to Problem (P). In other words, we show that the optimal solution of (P_μ) approaches the optimal solution of (P) as μ goes to 0. After that, we apply Newton's method to Problem (P_μ). In Section 4, we propose four different selections of calculating the appropriate displacement step. Section 5 is devoted to present a feasible primal-dual central trajectory algorithm followed by its complexity analysis. In Section 6, we compare the four selections of computing the displacement step and show the efficiency of the proposed algorithm in some numerical examples. Section 7 contains some concluding remarks.

2 Background and preliminary results

In this section, we provide the necessary background for the subsequent sections.

2.1 The algebraic structure of the second-order cone

In this subsection, we review the basics of the Jordan algebra associated with the second-order cone. We refer the reader to [24, Section 4] and [25] for more details of this algebra.

The *spectral decomposition* of $\boldsymbol{x}$ in $\mathcal{E}^n$ is a decomposition of $\boldsymbol{x}$ into *eigenvectors* (denoted by $\boldsymbol{c}_1(\boldsymbol{x})$ and $\boldsymbol{c}_2(\boldsymbol{x})$) together with *eigenvalues* (denoted by $\lambda_1(\boldsymbol{x})$ and $\lambda_2(\boldsymbol{x})$) as follows

$$
\boldsymbol{x} = \underbrace{(x_0 + \|\tilde{\boldsymbol{x}}\|)}_{\lambda_1(\boldsymbol{x})} \underbrace{\left(\frac{1}{2}\right)\left(1; \frac{\tilde{\boldsymbol{x}}}{\|\tilde{\boldsymbol{x}}\|}\right)}_{\boldsymbol{c}_1(\boldsymbol{x})} + \underbrace{(x_0 - \|\tilde{\boldsymbol{x}}\|)}_{\lambda_2(\boldsymbol{x})} \underbrace{\left(\frac{1}{2}\right)\left(1; -\frac{\tilde{\boldsymbol{x}}}{\|\tilde{\boldsymbol{x}}\|}\right)}_{\boldsymbol{c}_2(\boldsymbol{x})}.
$$

The *square* of $\boldsymbol{x}$ is defined as

$$\boldsymbol{x}^2 := \lambda_1^2(\boldsymbol{x})\boldsymbol{c}_1(\boldsymbol{x}) + \lambda_2^2(\boldsymbol{x})\boldsymbol{c}_2(\boldsymbol{x}) = \begin{bmatrix} x_0^2 + \|\tilde{\boldsymbol{x}}\|^2 \\ 2x_0\tilde{\boldsymbol{x}} \end{bmatrix}.$$

Note that $\boldsymbol{x}^2$ can be written as $\boldsymbol{x}^2 = \text{Arw}(\boldsymbol{x})\boldsymbol{x}$, where $\text{Arw}(\boldsymbol{x})$ is the *arrow-shaped matrix* in $\mathbb{R}^{n\times n}$ associated with each vector $\boldsymbol{x} \in \mathcal{E}^n$ and is defined as

$$\text{Arw}(\boldsymbol{x}) := \begin{bmatrix} x_0 & \tilde{\boldsymbol{x}}^\mathsf{T} \\ \tilde{\boldsymbol{x}} & x_0 I \end{bmatrix}.$$

Note that $\boldsymbol{x} \in \mathcal{E}_+^n$ ($\boldsymbol{x} \in \text{int } \mathcal{E}_+^n$) if and only if $\text{Arw}(\boldsymbol{x})$ is positive semidefinite ($\text{Arw}(\boldsymbol{x})$ is positive definite).

We define the *bilinear map* $\circ : \mathcal{E}^n \times \mathcal{E}^n \longrightarrow \mathcal{E}^n$ as

$$\boldsymbol{x} \circ \boldsymbol{y} := \text{Arw}(\boldsymbol{x})\boldsymbol{y} = \begin{bmatrix} \boldsymbol{x}^\mathsf{T}\boldsymbol{y} \\ x_0\tilde{\boldsymbol{y}} + y_0\tilde{\boldsymbol{x}} \end{bmatrix}$$

for every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^n$. It is known that the space $\mathcal{E}^n$ under the product "$\circ$" forms a Euclidean Jordan algebra (see [25] for definition) with *identity vector* $\boldsymbol{e}_n := (1; \boldsymbol{0}) \in \mathcal{E}^n$ under the inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\mathsf{T}\boldsymbol{y}$.

We define the *trace* and *determinant* of $\boldsymbol{x}$ as

$$\text{trace}(\boldsymbol{x}) := \lambda_1(\boldsymbol{x}) + \lambda_2(\boldsymbol{x}) = 2x_0 \quad \text{and} \quad \det(\boldsymbol{x}) := \lambda_1(\boldsymbol{x})\lambda_2(\boldsymbol{x}) = x_0^2 - \|\tilde{\boldsymbol{x}}\|^2.$$

Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a continuous function. We can define the function f on $\mathcal{E}^n$ as

$$f(\boldsymbol{x}) := f(\lambda_1(\boldsymbol{x}))\boldsymbol{c}_1(\boldsymbol{x}) + f(\lambda_2(\boldsymbol{x}))\boldsymbol{c}_2(\boldsymbol{x}).$$

In particular, $\boldsymbol{x}^n$ for an integer n, which is defined recursively as $\boldsymbol{x}^n := \boldsymbol{x} \circ \boldsymbol{x}^{n-1}$, can be redefined as

$$\boldsymbol{x}^n := \lambda_1^n(\boldsymbol{x})\boldsymbol{c}_1(\boldsymbol{x}) + \lambda_2^n(\boldsymbol{x})\boldsymbol{c}_2(\boldsymbol{x}).$$

For instance, the *inverse* of $\boldsymbol{x} \in \mathcal{E}^n$ is defined as

$$\boldsymbol{x}^{-1} := \frac{1}{\lambda_1(\boldsymbol{x})}\boldsymbol{c}_1(\boldsymbol{x}) + \frac{1}{\lambda_2(\boldsymbol{x})}\boldsymbol{c}_2(\boldsymbol{x}) = \frac{1}{\det(\boldsymbol{x})}R\boldsymbol{x}, \quad \text{where} \quad R := \begin{bmatrix} 1 & \boldsymbol{0}^\mathsf{T} \\ \boldsymbol{0} & -I_{n-1} \end{bmatrix}.$$

Observe that $\boldsymbol{x}^{-1} \circ \boldsymbol{x} = \boldsymbol{e}_n$. The vector $\boldsymbol{x} \in \mathcal{E}^n$ is called *invertible* if $\boldsymbol{x}^{-1}$ is defined (i.e., $\det(\boldsymbol{x}) \neq 0$) and *noninvertible* otherwise. The vector $\boldsymbol{x} \in \mathcal{E}^n$ is called *positive semidefinite* (*positive definite*) if $\boldsymbol{x} \in \mathcal{E}_+^n$ ($\boldsymbol{x} \in \text{int } \mathcal{E}_+^n$). Note that every positive definite vector is invertible and its inverse is also positive definite.

The *Frobenius norm* of $\boldsymbol{x} \in \mathcal{E}^n$ is defined as

$$\|\boldsymbol{x}\|_F := \sqrt{\lambda_1^2(\boldsymbol{x}) + \lambda_2^2(\boldsymbol{x})} = 2\|\boldsymbol{x}\|.$$

For any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^n$, we have

$$\|\boldsymbol{x} \circ \boldsymbol{y}\|_F \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|_F \leq \|\boldsymbol{x}\|_F \, \|\boldsymbol{y}\|_F \quad \text{and} \quad \|\boldsymbol{x} + \boldsymbol{y}\|_F^2 = \|\boldsymbol{x}\|_F^2 + \|\boldsymbol{y}\|_F^2 + 8\boldsymbol{x}^\mathsf{T}\boldsymbol{y}. \tag{1}$$

The *quadratic representation* of $\boldsymbol{x} \in \mathcal{E}^n$, $Q_{\boldsymbol{x}} : \mathcal{E}^n \longrightarrow \mathcal{E}^n$, is defined as

$$Q_{\boldsymbol{x}} := 2\text{Arw}^2(\boldsymbol{x}) - \text{Arw}(\boldsymbol{x}^2) = \begin{bmatrix} \|\boldsymbol{x}\|^2 & 2x_0\tilde{\boldsymbol{x}}^\mathsf{T} \\ 2x_0\tilde{\boldsymbol{x}} & \det(\boldsymbol{x})I + 2\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^\mathsf{T} \end{bmatrix}.$$

The map $Q.$ will play an important role for developing our analysis. It is known [24, Theorem 9] that $Q_{\boldsymbol{p}}\left(\mathcal{E}^n_{r+}\right) = \mathcal{E}^n_{r+}$ for any nonsingular vector $\boldsymbol{p} \in \mathcal{E}^n_{+}$. Note that $Q_{\boldsymbol{p}^{-1}} = Q_{\boldsymbol{p}}^{-1}$ if $\boldsymbol{p}$ is nonsingular.

Note that $\boldsymbol{c}_1(\boldsymbol{x})$ and $\boldsymbol{c}_2(\boldsymbol{x})$ satisfy the properties

$$\boldsymbol{c}_1(\boldsymbol{x}) \circ \boldsymbol{c}_2(\boldsymbol{x}) = \boldsymbol{0},\ \boldsymbol{c}_1^2(\boldsymbol{x}) = \boldsymbol{c}_1(\boldsymbol{x}),\ \boldsymbol{c}_2^2(\boldsymbol{x}) = \boldsymbol{c}_2(\boldsymbol{x}),\ \text{ and }\ \boldsymbol{c}_1(\boldsymbol{x}) + \boldsymbol{c}_2(\boldsymbol{x}) = \boldsymbol{e}_n. \tag{2}$$

Any pair of vectors $\{\boldsymbol{c}_1, \boldsymbol{c}_2\}$ that satisfies the properties in (2) is called a *Jordan frame*. We say that vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ *operator commute* iff they share a Jordan frame, i.e., $\boldsymbol{x} = \lambda_1(\boldsymbol{x})\boldsymbol{c}_1 + \lambda_2(\boldsymbol{x})\boldsymbol{c}_2$ and $\boldsymbol{y} = \lambda_1(\boldsymbol{y})\boldsymbol{c}_1 + \lambda_1(\boldsymbol{y})\boldsymbol{c}_2$ for a Jordan frame $\{\boldsymbol{c}_1, \boldsymbol{c}_2\}$. Therefore, $\boldsymbol{c}_1(\boldsymbol{x}) = \boldsymbol{c}_1(\boldsymbol{y})$ and $\boldsymbol{c}_2(\boldsymbol{x}) = \boldsymbol{c}_2(\boldsymbol{y})$ when $\boldsymbol{x}$ and $\boldsymbol{y}$ operator commute.

All the above notions are also used in the block sense as follows. Let $\boldsymbol{x} = (\boldsymbol{x}_1; \boldsymbol{x}_2; \ldots; \boldsymbol{x}_r)$ and $\boldsymbol{y} = (\boldsymbol{y}_1; \boldsymbol{y}_2; \ldots; \boldsymbol{y}_r)$ with and $\boldsymbol{x}_i, \boldsymbol{y}_i \in \mathcal{E}^{n_i}$ for $i = 1, 2, \ldots, r$. Then

- $\boldsymbol{x} \circ \boldsymbol{y} := (\boldsymbol{x}_1 \circ \boldsymbol{y}_1; \boldsymbol{x}_2 \circ \boldsymbol{y}_2; \ldots; \boldsymbol{x}_r \circ \boldsymbol{y}_r)$.
- $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{y} := {\boldsymbol{x}_1}^{\mathsf{T}}\boldsymbol{y}_1 + {\boldsymbol{x}_2}^{\mathsf{T}}\boldsymbol{y}_2 + \cdots + {\boldsymbol{x}_r}^{\mathsf{T}}\boldsymbol{y}_r$.
- $\operatorname{Arw}(\boldsymbol{x}) := \operatorname{Arw}(\boldsymbol{x}_1) \oplus \operatorname{Arw}(\boldsymbol{x}_2) \oplus \cdots \oplus \operatorname{Arw}(\boldsymbol{x}_r)$.
- $Q_{\boldsymbol{x}} := Q_{\boldsymbol{x}_1} \oplus Q_{\boldsymbol{x}_2} \oplus \cdots \oplus Q_{\boldsymbol{x}_r}$.
- $\boldsymbol{e} := (\boldsymbol{e}_{n_1}; \boldsymbol{e}_{n_2}; \ldots; \boldsymbol{e}_{n_r})$ is the identity vector of $\mathcal{E}^n_r$.
- $f(\boldsymbol{x}) := (f(\boldsymbol{x}_1); f(\boldsymbol{x}_2); \ldots; f(\boldsymbol{x}_r))$. In particular, $\boldsymbol{x}^{-1} := \left({\boldsymbol{x}_1}^{-1}; {\boldsymbol{x}_2}^{-1}; \ldots; {\boldsymbol{x}_r}^{-1}\right)$.
- $\|\boldsymbol{x}\|_F^2 := \|\boldsymbol{x}_1\|_F^2 + \|\boldsymbol{x}_2\|_F^2 + \cdots + \|\boldsymbol{x}_r\|_F^2$.
- $\boldsymbol{x}$ and $\boldsymbol{y}$ operator commute iff $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ operator commute for each $i = 1, 2$.

Note that $\boldsymbol{x} \in \mathcal{E}^n_{r+}$ has $2r$ eigenvalues (including multiplicities). These eigenvalues are comprised of the union of the eigenvalues of each $\boldsymbol{x}_i$ for $i = 1, 2, \ldots, r$.

2.2 Preliminary statistical inequalities

In this part, we review some statistical inequalities needed for the derivation of some of the results in the paper. Let $x_1, x_2, \ldots, x_n \in \mathbb{R}$ be a sample of size n, then its *mean* $\bar{x}$ and its *standard deviation* $\sigma_{\boldsymbol{x}}$ are respectively defined as

$$\bar{x} := \frac{1}{n}\sum_{k=1}^{n} x_k \quad \text{and} \quad \sigma_{\boldsymbol{x}}^2 := \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \bar{x}^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2.$$

The first statement in the following proposition is due to [31] and the second statement is due to [3, Theorem 5].

Proposition 2.1. *Assume that $\boldsymbol{x} \in \mathbb{R}^n$, then we have*

$$\bar{x} - \sigma_{\boldsymbol{x}}\sqrt{n-1} \leq \min_{1\leq k\leq n} x_k \leq \bar{x} - \frac{\sigma_{\boldsymbol{x}}}{\sqrt{n-1}}, \quad \textit{and} \quad \bar{x} + \frac{\sigma_{\boldsymbol{x}}}{\sqrt{n-1}} \leq \max_{1\leq k\leq n} x_k \leq \bar{x} + \sigma_{\boldsymbol{x}}\sqrt{n-1}.$$

In particular, if $x_k > 0$ for all $k = 1, 2, \ldots, n$, then we also have

$$n\ln\left(\bar{x} - \sigma_{\boldsymbol{x}}\sqrt{n-1}\right) \leq A \leq \sum_{k=1}^{n}\ln(x_k) \leq B \leq n\ln(\bar{x}), \quad \textit{where}$$

$$A = (n-1)\ln\left(\bar{x} + \frac{\sigma_{\boldsymbol{x}}}{\sqrt{n-1}}\right) + \ln\left(\bar{x} - \sigma_{\boldsymbol{x}}\sqrt{n-1}\right), \quad \textit{and} \quad B = (n-1)\ln\left(\bar{x} - \frac{\sigma_{\boldsymbol{x}}}{\sqrt{n-1}}\right) + \ln\left(\bar{x} + \sigma_{\boldsymbol{x}}\sqrt{n-1}\right).$$

3 Global convergence and Newton's method

In this section, we first prove the existence and uniqueness of the optimal solution of Problem (P_μ) and its convergence to Problem (P) as μ goes to zero. Then, we apply Newton's method to Problem (P_μ).

3.1 Convergence of the perturbed problem (P_μ) to (P)

In order to elaborate the convergence view, we start with the following introductory definition.

Definition 3.1. Let f be a real-valued function defined on a metric space X and $\alpha \geq 0$. Then

(a) The set $C_\alpha(f) := \{x \in X : f(x) \leq \alpha\}$ is called the *α-level set* of f.

(b) The function f is called *inf-compact* if the level sets $C_\alpha(f)$ are compact for all $\alpha > 0$.

(c) The *recession function* of f is the function $(f)_\infty : X \longrightarrow \mathbb{R}$ defined by

$$(f)_\infty(\Delta x) := \lim_{t\to+\infty} \frac{f(x + t\Delta x) - f(x)}{t}.$$

(d) The *recession cone* of f is the 0-level set of the recession function of f, denoted by $C_0\left((f)_\infty\right)$.

Note that the function f_μ is a strictly convex function for any $\mu > 0$. So, if an optimal solution of Problem (P_μ) exists then it should be unique. In order to show that Problem (P_μ) has a solution, it suffices to show that the function f_μ is inf-compact, which in turn takes us to verify that the following cone of recession

$$C_0\left(\left(f_\mu\right)_\infty\right) := \left\{\Delta x \in \mathcal{E}^n_{r+} : \left(f_\mu\right)_\infty(\Delta x) \leq 0\right\}$$

is nothing but the singleton zero. This will be seen in Lemma 3.2 which essentially depends on the following lemma.

Lemma 3.1. *Let $x \in \mathcal{E}^n_r$ and $t \in \mathbb{R}$ be such that $x \succ \mathbf{0}$ and $x + t\Delta x \succ \mathbf{0}$, then $\left(f_\mu\right)_\infty(\Delta x) = c^{\mathsf{T}}\Delta x$.*

Proof. The result is trivial if x belongs to the boundary of $\mathcal{E}^n_{r+}$. Let $x \in \operatorname{int} \mathcal{E}^n_{r+}$, then

$$f_\mu(x) = c^{\mathsf{T}} x - \mu \ln\det(x) + 2r\mu \ln \mu.$$

Let also $\omega(x) := x^{-1/2} \circ \left(\Delta x \circ x^{-1/2}\right)$, it follows that

$$\begin{aligned}
\left(f_\mu\right)_\infty(\Delta x) &= \lim_{t\to+\infty} \frac{f_\mu(x + t\Delta x) - f_\mu(x)}{t} \\
&= \lim_{t\to+\infty} \frac{c^{\mathsf{T}}(x + t\Delta x) - \mu \ln\det(x + t\Delta x) - c^{\mathsf{T}} x + \mu \ln\det(x)}{t} \\
&= c^{\mathsf{T}}\Delta x - \mu \lim_{t\to+\infty} \frac{\ln\det(x + t\Delta x) - \ln\det(x)}{t} \\
&= c^{\mathsf{T}}\Delta x - \mu \lim_{t\to+\infty} \frac{\ln\det\left(x^{1/2} \circ \left((e + t\omega(x)) \circ x^{1/2}\right)\right) - \ln\det(x)}{t} \\
&= c^{\mathsf{T}}\Delta x - \mu \lim_{t\to+\infty} \frac{\ln\left(\det(x)\det\left(e + t\omega(x)\right)\right) - \ln\det(x)}{t} \\
&= c^{\mathsf{T}}\Delta x - \mu \lim_{t\to+\infty} \frac{\ln\det\left(e + t\omega(x)\right)}{t} \\
&= c^{\mathsf{T}}\Delta x - \mu \lim_{t\to+\infty} \sum_{i=1}^{2r} \frac{\ln\left(1 + t\lambda_i(\omega)\right)}{t} \\
&= c^{\mathsf{T}}\Delta x.
\end{aligned}$$

This completes the proof. $\square$

Lemma 3.2. $C_0\left(\left(f_\mu\right)_\infty\right) = \left\{0\right\}$.

Proof. Suppose in the contrary that $C_0((f_\mu)_\infty) \neq \{0\}$, then there exists $\Delta\boldsymbol{x} \in \mathcal{E}^n_+ - \{\mathbf{0}\}$ such that $\Delta\boldsymbol{x} \in C_0((f_\mu)_\infty)$. Using Lemma 3.1, this implies that $\boldsymbol{c}^{\mathsf{T}}\Delta\boldsymbol{x} \leq 0$. Based on Assumption 1.2, there exists $(\boldsymbol{y};\hat{\boldsymbol{s}}) \in \mathcal{F}_D^\circ$ such that

$$0 < \hat{\boldsymbol{s}}^{\mathsf{T}}\Delta\boldsymbol{x} = \left(\boldsymbol{c} - A^{\mathsf{T}}\boldsymbol{y}\right)^{\mathsf{T}}\Delta\boldsymbol{x} = \boldsymbol{c}^{\mathsf{T}}\Delta\boldsymbol{x} - \boldsymbol{y}^{\mathsf{T}}A\Delta\boldsymbol{x} = \boldsymbol{c}^{\mathsf{T}}\Delta\boldsymbol{x},$$

where the last equality holds because $\Delta\boldsymbol{x}$ is a descent direction, that is $A\Delta\boldsymbol{x} = \mathbf{0}$. Thus, we have arrived to a contradiction which comes from the simultaneous positivity and non-positivity of the sign of $\boldsymbol{c}^{\mathsf{T}}\Delta\boldsymbol{x}$. $\square$

The following lemma proves the convergence of Problem (P_μ) to Problem (P) when μ approaches zero.

Lemma 3.3. *Let $\bar{\boldsymbol{x}}_\mu$ be an optimal primal solution of Problem (P_μ), then $\bar{\boldsymbol{x}} = \lim_{\mu\to 0}\bar{\boldsymbol{x}}_\mu$ is an optimal solution of Problem (P).*

Proof. Let $f_\mu(\boldsymbol{x}) := f(\boldsymbol{x},\mu)$ and $f(\boldsymbol{x}) := f(\boldsymbol{x},0)$. Since the function $f_\mu(\boldsymbol{x})$ is differentiable and convex, there exists an optimal solution $\bar{\boldsymbol{x}}_\mu$ of Problem (P_μ) such that

$$\nabla_{\boldsymbol{x}} f_\mu(\bar{\boldsymbol{x}}_\mu) = \nabla_{\boldsymbol{x}} f(\bar{\boldsymbol{x}}_\mu,\mu) = 0.$$

Then, for all $\boldsymbol{x} \in \mathcal{F}_P^\circ$, we have that

$$\begin{array}{rcl} f(\boldsymbol{x}) & \geq & f(\bar{\boldsymbol{x}}_\mu,\mu) + (\boldsymbol{x} - \bar{\boldsymbol{x}}_\mu)^{\mathsf{T}}\nabla_{\boldsymbol{x}} f(\bar{\boldsymbol{x}}_\mu,\mu) + (0-\mu)\dfrac{\partial}{\partial\mu} f(\bar{\boldsymbol{x}}_\mu,\mu) \\ & \geq & f(\bar{\boldsymbol{x}}_\mu,\mu) + \mu\ln\det\bar{\boldsymbol{x}}_\mu - 2r\mu\ln\mu - 2r\mu \\ & \geq & \boldsymbol{c}^{\mathsf{T}}\bar{\boldsymbol{x}}_\mu - \mu\ln\det\bar{\boldsymbol{x}}_\mu + 2r\mu\ln\mu + \mu\ln\det\bar{\boldsymbol{x}}_\mu - 2r\mu\ln\mu - 2r\mu \\ & \geq & \boldsymbol{c}^{\mathsf{T}}\bar{\boldsymbol{x}}_\mu - 2r\mu. \end{array}$$

Since $\boldsymbol{x}$ was arbitrary in $\mathcal{F}_P^\circ$, this implies that $\min_{\boldsymbol{x}\in\mathcal{F}_P^\circ} f(\boldsymbol{x}) \geq \boldsymbol{c}^{\mathsf{T}}\bar{\boldsymbol{x}}_\mu - 2r\mu \geq \boldsymbol{c}^{\mathsf{T}}\bar{\boldsymbol{x}}_\mu = f(\bar{\boldsymbol{x}}_\mu)$. On the other side, we have $f(\bar{\boldsymbol{x}}_\mu) \geq \min_{\boldsymbol{x}\in\mathcal{F}_P^\circ} f(\boldsymbol{x})$. As μ goes to 0, it immediately follows that $f(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathcal{F}_P^\circ} f(\boldsymbol{x})$. Thus, $\bar{\boldsymbol{x}}$ is an optimal solution of Problem (P). The proof is complete. $\square$

3.2 Newton's method and commutative directions

Since the objective function of Problem (P_μ) is strictly convex, the KKT conditions are necessary and sufficient to characterize an optimal solution of Problem (P_μ). Consequently, the points $\bar{\boldsymbol{x}}_\mu$ and $(\bar{\boldsymbol{y}}_\mu,\bar{\boldsymbol{s}}_\mu)$ are optimal solutions of (P_μ) and (D_μ) respectively if and only if they satisfy the perturbed nonlinear system

$$\begin{array}{rclcrcl} A\boldsymbol{x} & = & \boldsymbol{b}, & & \boldsymbol{x} & \succ & \mathbf{0}, \\ A^{\mathsf{T}}\boldsymbol{y} + \boldsymbol{s} & = & \boldsymbol{c}, & & \boldsymbol{s} & \succ & \mathbf{0}, \\ \boldsymbol{x}\circ\boldsymbol{s} & = & \mu\boldsymbol{e}, & & \mu & > & 0. \end{array} \tag{3}$$

We call the set of all solutions of system (3), denoted by $(\boldsymbol{x}_\mu,\boldsymbol{y}_\mu,\boldsymbol{s}_\mu)$ with $\mu > 0$, the *central path* or *central trajectory*. We say that a point $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{s})$ is near to the central trajectory if it belongs to the set $\mathcal{N}_\theta(\mu)$, which is defined as

$$\mathcal{N}_\theta(\mu) := \left\{(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s}) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ : d_F(\boldsymbol{x},\boldsymbol{s}) \leq \theta\mu\right\}, \ \text{ where } \ d_F(\boldsymbol{x},\boldsymbol{s}) := \|Q_{\boldsymbol{x}^{1/2}}\boldsymbol{s} - \mu\boldsymbol{e}\|_F \ \text{ and } \ \theta \in (0,1).$$

Now, we can apply Newton's method to system (3) and obtain the following linear system

$$\begin{array}{rcl} A\Delta \boldsymbol{x} &=& \mathbf{0}, \\ A^{\mathsf{T}}\Delta \boldsymbol{y} + \Delta \boldsymbol{s} &=& \mathbf{0}, \\ \boldsymbol{x} \circ \Delta \boldsymbol{s} + \Delta \boldsymbol{x} \circ \boldsymbol{s} &=& \sigma\mu \boldsymbol{e} - \boldsymbol{x} \circ \boldsymbol{s}. \end{array} \tag{4}$$

where $(\Delta \boldsymbol{x}; \Delta \boldsymbol{y}; \Delta \boldsymbol{s}) \in \mathcal{E}_r^n \times \mathbb{R}^m \times \mathcal{E}_r^n$ is the search direction, $\mu = \frac{1}{2r}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{s}$ is the normalized duality gap corresponding to $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ and $\sigma \in (0, 1)$ is the centering parameter.

Note that the strict second-order cone inequalities $\boldsymbol{x}, \boldsymbol{s} \succ \mathbf{0}$ imply that $d_F(\boldsymbol{x}, \boldsymbol{s}) \leq \|\boldsymbol{x} \circ \boldsymbol{s} - \mu \boldsymbol{e}\|_F$ with equality holds when $\boldsymbol{x}$ and $\boldsymbol{s}$ operator commute [33, Lemma 30]. In fact, it is known that many interesting properties become apparent for the analysis of interior-point methods when $\boldsymbol{x}$ and $\boldsymbol{s}$ operator commute. The *commutative class* is denoted by $\mathcal{C}(\boldsymbol{x}, \boldsymbol{s})$ and is defined as

$$\mathcal{C}(\boldsymbol{x}, \boldsymbol{s}) := \left\{ \boldsymbol{p} \in \mathcal{E}_r^n : \boldsymbol{p} \text{ is nonsingular, } Q_{\boldsymbol{p}}\boldsymbol{x} \text{ and } Q_{\boldsymbol{p}^{-1}}\boldsymbol{s} \text{ operator commute} \right\}. \tag{5}$$

From [33, Lemma 28], the equality $\boldsymbol{x} \circ \boldsymbol{s} = \mu \boldsymbol{e}$ holds if and only if the equality $(Q_{\boldsymbol{p}}\boldsymbol{x}) \circ (Q_{\boldsymbol{p}^{-1}}\boldsymbol{s}) = \mu \boldsymbol{e}$ holds, for any nonsingular vector $\boldsymbol{p} \in \mathcal{E}_r^n$. Therefore, for any given invertible vector $\boldsymbol{p} \in \mathcal{E}^n$, the system (3) is equivalent to the system

$$\begin{array}{rclcrcl} A\boldsymbol{x} &=& \boldsymbol{b}, & & \boldsymbol{x} &\succ& \mathbf{0}, \\ A^{\mathsf{T}}\boldsymbol{y} + \boldsymbol{s} &=& \boldsymbol{c}, & & \boldsymbol{s} &\succ& \mathbf{0}, \\ (Q_{\boldsymbol{p}}\boldsymbol{x}) \circ (Q_{\boldsymbol{p}^{-1}}\boldsymbol{s}) &=& \mu \boldsymbol{e}, & & \mu &>& 0. \end{array} \tag{6}$$

Let $\boldsymbol{v} \in \mathcal{E}_r^n$. With respect to a nonsingular vector $\boldsymbol{p} \in \mathcal{E}_r^n$, we define the scaling vectors $\overline{\boldsymbol{v}}$ and $\underline{\boldsymbol{v}}$ and the scaling matrix $\underline{A}$ as

$$\overline{\boldsymbol{v}} := Q_{\boldsymbol{p}}\boldsymbol{v}, \;\; \underline{\boldsymbol{v}} := Q_{\boldsymbol{p}^{-1}}\boldsymbol{v}, \;\; \text{and} \;\; \underline{A} := Q_{\boldsymbol{p}}A.$$

Using this change of variables and the fact that $Q_{\boldsymbol{p}}\left(\mathcal{E}_{r+}^n\right) = \mathcal{E}_{r+}^n$, we conclude that the system (4) is equivalent to the following Newton system

$$\begin{array}{rcl} \underline{A}\overline{\Delta \boldsymbol{x}} &=& \boldsymbol{b} - \underline{A}\overline{\boldsymbol{x}}, \\ \underline{A}^{\mathsf{T}}\Delta \boldsymbol{y} + \underline{\Delta \boldsymbol{s}} &=& \underline{\boldsymbol{c}} - \underline{\boldsymbol{s}} - \underline{A}^{\mathsf{T}}\boldsymbol{y}, \\ \overline{\boldsymbol{x}} \circ \underline{\Delta \boldsymbol{s}} + \overline{\Delta \boldsymbol{x}} \circ \underline{\boldsymbol{s}} &=& \sigma\mu \boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}. \end{array} \tag{7}$$

Here, the normalized duality gap is $\mu = \frac{1}{2r}\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} = \frac{1}{2r}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{s}$. In fact,

$$\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} = \left(Q_{\boldsymbol{p}}\boldsymbol{x}\right)^{\mathsf{T}} Q_{\boldsymbol{p}^{-1}}\boldsymbol{s} = \boldsymbol{x}^{\mathsf{T}}Q_{\boldsymbol{p}}Q_{\boldsymbol{p}^{-1}}\boldsymbol{s} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{s}. \tag{8}$$

Solving the scaled Newton system (7) yields the search direction $\left(\overline{\Delta \boldsymbol{x}}; \Delta \boldsymbol{y}; \underline{\Delta \boldsymbol{s}}\right)$. Then, we apply the inverse scaling to $\left(\overline{\Delta \boldsymbol{x}}; \underline{\Delta \boldsymbol{s}}\right)$ to obtain the Newton direction $(\Delta \boldsymbol{x}; \Delta \boldsymbol{s})$. Note that the search direction $\left(\overline{\Delta \boldsymbol{x}}; \Delta \boldsymbol{y}; \underline{\Delta \boldsymbol{s}}\right)$ belongs to the so-called the *MZ family of directions* (due to Monteiro [34] and Zhang [35]). In fact, such a way of scaling originally proposed for semidefinite programming by Monteiro [34] and Zhang [35], and after that it was generalized for general symmetric cone programming by Schmieta and Alizadeh [33].

Clearly, the set $\mathcal{C}(\boldsymbol{x}, \boldsymbol{s})$ defined in (5) is a subclass of the MZ family of search directions. Our focus in this paper is in vectors $\boldsymbol{p} \in \mathcal{C}(\boldsymbol{x}, \boldsymbol{s})$. We discuss the following three choices of $\boldsymbol{p}$ (see [33, Section 3]): We may choose $\boldsymbol{p} = \boldsymbol{x}^{1/2}$ to obtain $\overline{\boldsymbol{x}} = \boldsymbol{e}$, and we may choose $\boldsymbol{p} = \boldsymbol{s}^{-1/2}$ to obtain $\underline{\boldsymbol{s}} = \boldsymbol{e}$. These two choices of directions are called the *HRVW/KSH/M directions* (due to Helmberg et al. [36], Monteiro [34] and Kojima et al. [37]). The third choice of $\boldsymbol{p}$ is given by $\boldsymbol{p} = \left(Q_{\boldsymbol{x}^{1/2}}\left(Q_{\boldsymbol{x}^{1/2}}\boldsymbol{s}\right)^{-1/2}\right)^{-1/2}$, which yields $Q_{\boldsymbol{p}}^2\boldsymbol{x} = \boldsymbol{s}$, and therefore $\underline{\boldsymbol{s}} = Q_{\boldsymbol{p}^{-1}}\boldsymbol{s} = Q_{\boldsymbol{p}}\boldsymbol{x} = \overline{\boldsymbol{x}}$. This choice of directions is called the *NT direction* (due to Nesterov and Todd [38]).

4 Selections of the displacement step

We start by motivating the need for introducing the so-called displacement step. Note that the positive definiteness of the vectors $\boldsymbol{x}^+ = \boldsymbol{x} + \Delta\boldsymbol{x}$ and $\boldsymbol{s}^+ = \boldsymbol{s} + \Delta\boldsymbol{s}$ is not always achieved. In order to circumvent this difficulty, we introduce a parameter $\alpha > 0$, for which we call the *displacement step*, then we redefine $\boldsymbol{x}^+, \boldsymbol{y}^+$ and $\boldsymbol{s}^+$ as

$$\boldsymbol{x}^+ := \boldsymbol{x} + \alpha\Delta\boldsymbol{x},\ \boldsymbol{y}^+ := \boldsymbol{y} + \alpha\Delta\boldsymbol{y},\ \text{and } \boldsymbol{s}^+ := \boldsymbol{s} + \alpha\Delta\boldsymbol{s}.$$

Calculating the displacement step using classical line search methods is undesirable and even generally impossible [1,3]. In this section, we propose four different selections of computing the appropriate displacement step. Each one of the following four lemmas gives a selection to calculate the displacement step α.

Lemma 4.1 (The first selection lemma)**.** *Let* $(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s}) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ$. *If* $\alpha = \rho\min\{\alpha_x, \alpha_s\}$ *with* $0 < \rho < 1$, *then* $\boldsymbol{x}^+, \boldsymbol{s}^+ \succ \boldsymbol{0}$, *where for* $\boldsymbol{v} \in \{\boldsymbol{x}, \boldsymbol{s}\}$ *we have*

$$\alpha_v = \begin{cases} \dfrac{-1}{\bar{\lambda}_v - \delta_v\sqrt{2r-1}} - \varepsilon, & \textit{if } \left(\dfrac{-1}{\bar{\lambda}_v - \delta_v\sqrt{2r-1}} > 0 \textit{ and } \min\limits_{1\le i\le 2r}\lambda_i\left(\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)\right) < 0\right); \\ \varepsilon, & \textit{if } \left(\dfrac{-1}{\bar{\lambda}_v - \delta_v\sqrt{2r-1}} < 0 \textit{ and } \min\limits_{1\le i\le 2r}\lambda_i\left(\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)\right) < 0\right); \\ 1, & \textit{if } \min\limits_{1\le i\le 2r}\lambda_i\left(\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)\right) > 0, \end{cases} \tag{9}$$

where

$$\bar{\lambda}_v = \frac{1}{2r}\sum_{i=1}^{2r}\lambda_i\left(\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)\right),\quad \delta_v^2 = \frac{1}{n}\sum_{i=1}^{2r}\lambda_i^2\left(\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)\right) - \bar{\lambda}_v^2,$$

and ϵ *is a small positive real. Here,* $\lambda_i\left(\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)\right), i = 1,2,\ldots,2r$, *are the eigenvalues of the vector* $\boldsymbol{v}^{-1/2}\circ\left(\Delta\boldsymbol{v}\circ\boldsymbol{v}^{-1/2}\right)$.

Proof. Since $\boldsymbol{x} \succ \boldsymbol{0}$, the vectors $\boldsymbol{x}^{\pm 1/2}$ are well-defined and we have $\boldsymbol{x}^{\pm 1/2} \succ \boldsymbol{0}$. Note that

$$\boldsymbol{x}^+ = \boldsymbol{x} + \alpha\,\Delta\boldsymbol{x} = \boldsymbol{x}^{1/2}\circ\boldsymbol{x}^{1/2} + \alpha\,\Delta\boldsymbol{x} = \boldsymbol{x}^{1/2}\circ\left(\left(\boldsymbol{e} + \alpha\,\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right)\circ\boldsymbol{x}^{1/2}\right).$$

Therefore, as $\boldsymbol{x}^{-1/2} \succ \boldsymbol{0}$, we have that

$$\begin{aligned} \boldsymbol{x}^+ \succ \boldsymbol{0} &\iff \boldsymbol{x}^{-1/2}\circ\boldsymbol{x}^+ \succ \boldsymbol{0} \\ &\iff \left(\boldsymbol{x}^{-1/2}\circ\boldsymbol{x}^+\right)\circ\boldsymbol{x}^{-1/2} \succ \boldsymbol{0} \\ &\iff \boldsymbol{e} + \alpha\,\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right) \succ \boldsymbol{0} \\ &\iff 1 + \alpha\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right) > 0,\ \text{for } i = 1,2,\ldots,2r. \end{aligned}$$

Hence,

$$1 + \alpha\min_{1\le i\le 2r}\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right) > 0 \implies \alpha\min_{1\le i\le 2r}\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right) > -1.$$

If $\min_{1\le i\le 2r}\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right) < 0$, then $\alpha < -1/\min_{1\le i\le 2r}\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right)$. Using Proposition 2.1, we conclude that

$$\alpha < \frac{-1}{\min\limits_{1\le i\le 2r}\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right)} \quad\text{and}\quad \frac{-1}{\bar{\lambda}_x - \delta_x\sqrt{2r-1}} \le \frac{-1}{\min\limits_{1\le i\le 2r}\lambda_i\left(\boldsymbol{x}^{-1/2}\circ\left(\Delta\boldsymbol{x}\circ\boldsymbol{x}^{-1/2}\right)\right)}.$$

This gives the expression of α_x.

Applying the same procedure, we obtain α_s. Finally, we choose $\alpha = \rho \min\{\alpha_x, \alpha_s\}$ with $0 < \rho < 1$. The proof is complete. □

Lemma 4.2 (The second selection lemma). *Let $(x; y; s) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ$. If $\alpha = \rho \min\{\alpha_x, \alpha_s\}$ with $0 < \rho < 1$, then $x^+, s^+ \succ \mathbf{0}$, where for $v \in \{x, s\}$ we have*

$$\alpha_v = \begin{cases} \dfrac{-1}{\bar{\lambda}_v - \delta_v \sqrt{2r-1}} - \varepsilon, & \text{if } \left(\dfrac{-1}{\bar{\lambda}_v - \delta_v \sqrt{2r-1}} > 0 \text{ and } \min\limits_{1 \le i \le 2r} \lambda_i\left(v^{-1} \circ \Delta v\right) < 0\right); \\ \varepsilon, & \text{if } \left(\dfrac{-1}{\bar{\lambda}_v - \delta_v \sqrt{2r-1}} < 0 \text{ and } \min\limits_{1 \le i \le 2r} \lambda_i\left(v^{-1} \circ \Delta v\right) < 0\right); \\ 1, & \text{if } \min\limits_{1 \le i \le 2r} \lambda_i\left(v^{-1} \circ \Delta v\right) > 0, \end{cases} \tag{10}$$

where

$$\bar{\lambda}_v = \frac{1}{2r} \sum_{i=1}^{2r} \lambda_i\left(v^{-1} \circ \Delta v\right), \quad \delta_v^2 = \frac{1}{n} \sum_{i=1}^{2r} \lambda_i^2\left(v^{-1} \circ \Delta v\right) - \bar{\lambda}_v^2,$$

and ϵ is a small positive real. Here, $\lambda_i\left(v^{-1} \circ \Delta v\right), i = 1, 2, \ldots, 2r$, are the eigenvalues of the vector $v^{-1} \circ \Delta v$.

Proof. Since $x \succ \mathbf{0}$, the vector x^{-1} is invertible and positive definite (i.e., $x^{-1} \succ \mathbf{0}$). Note that

$$x^+ = x + \alpha\, \Delta x = x \circ \left(e + \alpha\, x^{-1} \circ \Delta x\right).$$

Therefore, as $x^{-1} \succ \mathbf{0}$, we have that

$$\begin{aligned} x^+ \succ \mathbf{0} &\iff x^{-1} \circ x^+ \succ \mathbf{0} \\ &\iff e + \alpha\, x^{-1} \circ \Delta x \succ \mathbf{0} \\ &\iff 1 + \alpha \lambda_i\left(x^{-1} \circ \Delta x\right) > 0, \text{ for } i = 1, 2, \ldots, 2r. \end{aligned}$$

Hence,

$$1 + \alpha \min_{1 \le i \le 2r} \lambda_i\left(x^{-1} \circ \Delta x\right) > 0 \implies \alpha \min_{1 \le i \le 2r} \lambda_i\left(x^{-1} \circ \Delta x\right) > -1.$$

If $\min_{1 \le i \le 2r} \lambda_i\left(x^{-1} \circ \Delta x\right) < 0$, then $\alpha < -1/\min_{1 \le i \le 2r} \lambda_i\left(x^{-1} \circ \Delta x\right)$. Using Proposition 2.1, we conclude that

$$\alpha < \frac{-1}{\min\limits_{1 \le i \le 2r} \lambda_i\left(x^{-1} \circ \Delta x\right)} \quad \text{and} \quad \frac{-1}{\bar{\lambda}_x - \delta_x \sqrt{2r-1}} \le \frac{-1}{\min\limits_{1 \le i \le 2r} \lambda_i\left(x^{-1} \circ \Delta x\right)}.$$

This gives the expression of α_x.

Applying the same procedure, we obtain α_s. Finally, we choose $\alpha = \rho \min\{\alpha_x, \alpha_s\}$ with $0 < \rho < 1$. The proof is complete. □

Lemma 4.3 (The third selection lemma). *Let $(x; y; s) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ$. If $\alpha = \rho \min\{\alpha_x, \alpha_s\}$ with $0 < \rho < 1$, then $x^+, s^+ \succ \mathbf{0}$, where for $v \in \{x, s\}$ we have*

$$\alpha_v = \begin{cases} -\dfrac{\bar{\lambda}_v - \delta_v \sqrt{2r-1}}{\bar{\lambda}_{\Delta v} - \delta_{\Delta v} \sqrt{2r-1}} - \varepsilon, & \text{if } \left(-\dfrac{\bar{\lambda}_v - \delta_v \sqrt{2r-1}}{\bar{\lambda}_{\Delta v} - \delta_{\Delta v} \sqrt{2r-1}} > 0 \text{ and } \min\limits_{1 \le i \le 2r} \lambda_i(\Delta v) < 0\right); \\ \varepsilon, & \text{if } \left(-\dfrac{\bar{\lambda}_v - \delta_v \sqrt{2r-1}}{\bar{\lambda}_{\Delta v} - \delta_{\Delta v} \sqrt{2r-1}} \text{ and } \min\limits_{1 \le i \le 2r} \lambda_i(\Delta v) < 0\right); \\ 1, & \text{if } \min\limits_{1 \le i \le 2r} \lambda_i(\Delta v) > 0, \end{cases} \tag{11}$$

where

$$\bar{\lambda}_v = \frac{1}{2r}\sum_{i=1}^{2r}\lambda_i(v), \quad \bar{\lambda}_{\Delta v} = \frac{1}{2r}\sum_{i=1}^{2r}\lambda_i(\Delta v), \quad \delta_v^2 = \frac{1}{2r}\sum_{i=1}^{2r}\lambda_i^2(v) - \bar{\lambda}_v^2, \quad \delta_{\Delta v}^2 = \frac{1}{2r}\sum_{i=1}^{2r}\lambda_i^2(\Delta v) - \bar{\lambda}_{\Delta v}^2,$$

and ϵ is a small positive real. Here, $\lambda_i(\boldsymbol{v}), i = 1, 2, \ldots, 2r$, are the eigenvalues of the vector $\boldsymbol{v}$, and $\lambda_i(\Delta\boldsymbol{v}), i = 1, 2, \ldots, 2r$, are the eigenvalues of the vector $\Delta\boldsymbol{v}$,

Proof. It is known that $\boldsymbol{x}^+ = \boldsymbol{x} + \alpha\Delta\boldsymbol{x} \succ \boldsymbol{0}$ if and only if $\min_{1\le i\le 2r}\lambda_i(\boldsymbol{x}^+) > 0$. Here, $\lambda_i(\boldsymbol{x}^+), i = 1, 2, \ldots, 2r$, are the eigenvalues of the vector $\boldsymbol{x}^+$. It is also known (see for example [39]) that

$$\min_{1\le i\le 2r}\lambda_i(\boldsymbol{x}^+) \ge \min_{1\le i\le 2r}\lambda_i(\boldsymbol{x}) + \min_{1\le i\le 2r}\lambda_i(\Delta\boldsymbol{x}).$$

Then, it is enough to find α such that

$$\min_{1\le i\le 2r}\lambda_i(\boldsymbol{x}^+) \ge \min_{1\le i\le 2r}\lambda_i(\boldsymbol{x}) + \alpha \min_{1\le i\le 2r}\lambda_i(\Delta\boldsymbol{x}) > 0.$$

Hence

$$\min_{1\le i\le 2r}\lambda_i(\Delta\boldsymbol{x}) < 0 \implies \alpha < -\frac{\min_{1\le i\le 2r}\lambda_i(\boldsymbol{x})}{\min_{1\le i\le 2r}\lambda_i(\Delta\boldsymbol{x})}.$$

Using Proposition 2.1, we find

$$\alpha < -\frac{\min_{1\le i\le 2r}\lambda_i(\boldsymbol{x})}{\min_{1\le i\le 2r}\lambda_i(\Delta\boldsymbol{x})} \quad \text{and} \quad -\frac{\bar{\lambda}_x - \delta_x\sqrt{2r-1}}{\bar{\lambda}_{\Delta x} - \delta_{\Delta x}\sqrt{2r-1}} \le -\frac{\min_{1\le i\le 2r}\lambda_i(\boldsymbol{x})}{\min_{1\le i\le 2r}\lambda_i(\Delta\boldsymbol{x})}.$$

This gives the expression of α_x.

Applying the same procedure, we obtain α_s. Finally, we choose $\alpha = \rho\min\{\alpha_x, \alpha_s\}$ with $0 < \rho < 1$. The proof is complete. □

Lemma 4.4 (The fourth selection lemma)**.** *Let $(\boldsymbol{x}; \boldsymbol{y}; \boldsymbol{s}) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ$. If $\alpha = \rho\min\{\alpha_x, \alpha_s\}$ with $0 < \rho < 1$, then $\boldsymbol{x}^+, \boldsymbol{s}^+ \succ \boldsymbol{0}$, where for $\boldsymbol{v} \in \{\boldsymbol{x}, \boldsymbol{s}\}$ we have*

$$\alpha_v = \begin{cases} \min\limits_{i\in I_v} \dfrac{\left\|\widetilde{\boldsymbol{v}_i}\right\|_1 - (\boldsymbol{v}_i)_0}{(\Delta\boldsymbol{v}_i)_0 - \left\|\widetilde{\Delta\boldsymbol{v}_i}\right\|_1}, & \text{if} \quad I_v \neq \emptyset; \\ +\infty, & \text{if} \quad I_v = \emptyset, \end{cases} \tag{12}$$

where

$$I_v := \left\{ i \in \{1, 2, \ldots, r\} : (\Delta\boldsymbol{v}_i)_0 - \left\|\widetilde{\Delta\boldsymbol{v}_i}\right\|_1 < 0 \right\}.$$

Proof. Recall that a vector $\boldsymbol{w} \in \mathcal{E}_r^n$ is positive definite (i.e., $\boldsymbol{w} \succ \boldsymbol{0}$) if and only if $(\boldsymbol{w}_i)_0 > \|\widetilde{\boldsymbol{w}_i}\|$ (i.e., $\boldsymbol{w}_i \succ \boldsymbol{0}$) for $i = 1, 2, \ldots, r$. Let $\|\widetilde{\boldsymbol{w}_i}\|_1 := \sum_{j=1}^{n_i-1}|(\widetilde{\boldsymbol{w}_i})_j|$ denote the *Taxicab norm* of the subvector $\widetilde{\boldsymbol{w}_i}$, then it is known that $\|\widetilde{\boldsymbol{w}_i}\| \le \|\widetilde{\boldsymbol{w}_i}\|_1$. It follows that $\boldsymbol{w} \succ \boldsymbol{0}$ if $(\boldsymbol{w}_i)_0 > \|\widetilde{\boldsymbol{w}_i}\|_1$ for $i = 1, 2, \ldots, r$.

Then, $\boldsymbol{x}^+ = \boldsymbol{x} + \alpha\Delta\boldsymbol{x} \succ \boldsymbol{0}$ if

$$\left(\boldsymbol{x}_i^+\right)_0 > \left\|\widetilde{\boldsymbol{x}_i^+}\right\|_1, \quad \forall \;\; i = 1, 2, \ldots, r.$$

Because, for $i = 1, 2, \ldots, r$, we have

$$\left\|\widetilde{\boldsymbol{x}_i^+}\right\|_1 = \left\|\widetilde{\boldsymbol{x}_i} + \alpha\widetilde{\Delta\boldsymbol{x}_i}\right\|_1 \le \left\|\widetilde{\boldsymbol{x}_i}\right\|_1 + \alpha\left\|\widetilde{\Delta\boldsymbol{x}_i}\right\|_1,$$

it is enough to find α such that

$$(\boldsymbol{x}_i)_0 + \alpha\,(\Delta \boldsymbol{x}_i)_0 > \left\|\widetilde{\boldsymbol{x}_i}\right\|_1 + \alpha\left\|\widetilde{\Delta \boldsymbol{x}_i}\right\|_1, \quad \forall\ i = 1, 2, \ldots, r,$$

or equivalently

$$\alpha\left((\Delta \boldsymbol{x}_i)_0 - \left\|\widetilde{\Delta \boldsymbol{x}_i}\right\|_1\right) > \left\|\widetilde{\boldsymbol{x}_i}\right\|_1 - (\boldsymbol{x}_i)_0, \quad \forall\ i = 1, 2, \ldots, r.$$

This gives the expression of $\alpha_{\boldsymbol{x}}$.

Applying the same procedure for the vector $\boldsymbol{s}^+ = \boldsymbol{s} + \alpha\Delta\boldsymbol{s}$, we obtain $\alpha_{\boldsymbol{s}}$. Finally, we choose $\alpha = \rho\min\{\alpha_{\boldsymbol{x}}, \alpha_{\boldsymbol{s}}\}$ with $0 < \rho < 1$. The proof is complete. □

5 The central trajectory algorithm and its convergence

In this section, we present the central trajectory algorithm for solving the SOCP problem and present its convergence result. The algorithm is formally stated in Algorithm 5.1.

Algorithm 5.1. The central trajectory algorithm for SOCP problem.

Begin algorithm

Initialize $k = 0, \boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{s}^{(0)}, \mu^{(0)}, \epsilon, \sigma, \theta$

Ensure: $\left(\boldsymbol{x}^{(0)}; \boldsymbol{y}^{(0)}; \boldsymbol{s}^{(0)}\right) \in \mathcal{N}_\theta(\mu)$, $\epsilon > 0$, $\sigma, \theta \in (0, 1)$

While $\boldsymbol{x}^{(k)\top}\boldsymbol{s}^{(k)} \geq \epsilon$ **do**

Choose $\boldsymbol{p}^{(k)} \in C\left(\boldsymbol{x}^{(k)}, \boldsymbol{s}^{(k)}\right)$

Compute $\left(\overline{\boldsymbol{x}^{(k)}}; \boldsymbol{y}^{(k)}; \underline{\boldsymbol{s}^{(k)}}\right)$ by applying scaling to $\left(\boldsymbol{x}^{(k)}; \boldsymbol{y}^{(k)}; \boldsymbol{s}^{(k)}\right)$

Let $\mu^{(k)} := \frac{1}{2r}\overline{\boldsymbol{x}^{(k)}}^{\top}\underline{\boldsymbol{s}^{(k)}}$, $\boldsymbol{h}^{(k)} := \sigma\mu^{(k)}\boldsymbol{e} - \overline{\boldsymbol{x}^{(k)}} \circ \underline{\boldsymbol{s}^{(k)}}$ and $\Psi^{(k)} := \frac{1}{\mu}\underline{A}\ \overline{\boldsymbol{x}^{(k)2}}\,\underline{A}^{\top}$

Compute $\left(\overline{\Delta\boldsymbol{x}^{(k)}}; \Delta\boldsymbol{y}^{(k)}; \underline{\Delta\boldsymbol{s}^{(k)}}\right)$ by solving the scaled Newton system (7) to get

$$\left(\overline{\Delta\boldsymbol{x}^{(k)}};\ \Delta\boldsymbol{y}^{(k)};\ \underline{\Delta\boldsymbol{s}^{(k)}}\right) := \left(\left(\boldsymbol{h}^{(k)} - \overline{\boldsymbol{x}^{(k)}} \circ \underline{\Delta\boldsymbol{s}^{(k)}}\right) \circ \underline{\boldsymbol{s}^{(k)}}^{-1};\ -\Psi^{(k)^{-1}}\underline{A}\left(\underline{\boldsymbol{s}^{(k)}}^{-1} \circ \boldsymbol{h}^{(k)}\right);\ -\underline{A}^{\top}\Delta\boldsymbol{y}^{(k)}\right)$$

Compute $\left(\Delta\boldsymbol{x}^{(k)}; \Delta\boldsymbol{y}^{(k)}; \Delta\boldsymbol{s}^{(k)}\right)$ by applying inverse scaling to $\left(\overline{\Delta\boldsymbol{x}^{(k)}}; \Delta\boldsymbol{y}^{(k)}; \underline{\Delta\boldsymbol{s}^{(k)}}\right)$

Calculate the displacement step $\alpha^{(k)}$ by one of the selections in (9), (10), (11) or (12)

Set the new iterate according to

$$\left(\boldsymbol{x}^{(k+1)};\ \boldsymbol{y}^{(k+1)};\ \boldsymbol{s}^{(k+1)}\right) := \left(\boldsymbol{x}^{(k)} + \alpha^{(k)}\Delta\boldsymbol{x}^{(k)};\ \boldsymbol{y}^{(k)} + \alpha^{(k)}\Delta\boldsymbol{y}^{(k)};\ \boldsymbol{s}^{(k)} + \alpha^{(k)}\Delta\boldsymbol{s}^{(k)}\right)$$

Set $k = k + 1$

End while

End algorithm

Algorithm 5.1 selects a sequence of step-size $\{\alpha^{(k)}\}$ and centrality parameters $\{\sigma^{(k)}\}$ according to the following rule: for all $k \geq 0$, we take $\alpha^{(k)} = 1 - \delta/\sqrt{2r}$, where $\delta \in \left[0, \sqrt{2r}\right)$ as it will be seen in the next section. In the rest of this section, we present the convergence result of Algorithm 5.1. The proof of this convergence result depends essentially on the following lemma.

Lemma 5.1. *Let* $(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s}) \in int\ \mathcal{E}^n_{r+} \times \mathbb{R}^m \times int\ \mathcal{E}^n_{r+}$, $\left(\overline{\boldsymbol{x}};\boldsymbol{y};\underline{\boldsymbol{s}}\right)$ *be obtained by applying scaling to* $(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s})$, *and* $\left(\overline{\Delta\boldsymbol{x}};\Delta\boldsymbol{y};\underline{\Delta\boldsymbol{s}}\right)$ *be a solution of the system (7). Then we have*

(a) $\overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{\Delta\boldsymbol{s}} = 0.$

(b) $\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\Delta\boldsymbol{s}} + \overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} = \frac{1}{2}\text{trace}(\boldsymbol{h})$, *where* $\boldsymbol{h} = \sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}$ *such that* $\sigma \in (0,1)$ *and* $\mu = \frac{1}{2r}\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}}$.

(c) $\overline{\boldsymbol{x}}^{+\mathsf{T}}\underline{\boldsymbol{s}}^{+} = \left(1 - \alpha\left(1 - \frac{\sigma}{2}\right)\right)\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}}$, $\forall\alpha \in \mathbb{R}$, *where* $\overline{\boldsymbol{x}}^{+} = \overline{\boldsymbol{x}} + \alpha\overline{\Delta\boldsymbol{x}}$ *and* $\underline{\boldsymbol{s}}^{+} = \underline{\boldsymbol{s}} + \alpha\underline{\Delta\boldsymbol{s}}$.

(d) $\boldsymbol{x}^{+\mathsf{T}}\boldsymbol{s}^{+} = \left(1 - \alpha\left(1 - \frac{\sigma}{2}\right)\right)\boldsymbol{x}^{\mathsf{T}}\boldsymbol{s}$, $\forall\alpha \in \mathbb{R}$, *where* $\boldsymbol{x}^{+} = \boldsymbol{x} + \alpha\Delta\boldsymbol{x}$ *and* $\boldsymbol{s}^{+} = \boldsymbol{s} + \alpha\Delta\boldsymbol{s}$.

Proof. By the first two equations of the system (7), we get

$$\overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{\Delta\boldsymbol{s}} = -\overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{A}^{\mathsf{T}}\Delta\boldsymbol{y} = -\left(\underline{A}\,\overline{\Delta\boldsymbol{x}}\right)^{\mathsf{T}}\Delta\boldsymbol{y} = 0.$$

This proves item (a). We prove item (b) by noting that

$$\begin{array}{rcl}
\text{trace}(\boldsymbol{h}) & = & \text{trace}\left(\sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right) \\
& = & \text{trace}\left(\overline{\boldsymbol{x}} \circ \underline{\Delta\boldsymbol{s}} + \overline{\Delta\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right) \\
& = & \text{trace}\left(\overline{\boldsymbol{x}} \circ \underline{\Delta\boldsymbol{s}}\right) + \text{trace}\left(\overline{\Delta\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right) \\
& = & 2\left(\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\Delta\boldsymbol{s}} + \overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}}\right),
\end{array}$$

where we used the last equation of the system (7) to obtain the first equality. To prove item (c), note that

$$\begin{array}{rcl}
\overline{\boldsymbol{x}}^{+\mathsf{T}}\underline{\boldsymbol{s}}^{+} & = & \left(\overline{\boldsymbol{x}} + \alpha\overline{\Delta\boldsymbol{x}}\right)^{\mathsf{T}}\left(\underline{\boldsymbol{s}} + \alpha\underline{\Delta\boldsymbol{s}}\right) \\
& = & \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} + \alpha\left(\overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} + \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\Delta\boldsymbol{s}}\right) + \alpha^2\overline{\Delta\boldsymbol{x}}^{\mathsf{T}}\underline{\Delta\boldsymbol{s}} \\
& = & \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} + \frac{1}{2}\alpha\ \text{trace}\left(\sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right) \\
& = & \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} + \frac{1}{2}\alpha\sigma\mu\ \text{trace}\left(\boldsymbol{e}\right) - \frac{1}{2}\alpha\ \text{trace}\left(\overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right) \\
& = & \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} + \alpha\sigma\mu r - \alpha\overline{\boldsymbol{x}}^{\mathsf{T}}\overline{\boldsymbol{s}} \\
& = & \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} + \frac{1}{2}\alpha\sigma\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} - \alpha\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} \\
& = & \left(1 - \alpha\left(1 - \frac{\sigma}{2}\right)\right)\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}},
\end{array}$$

where the third equality follows from items (a) and (b).

Finally, item (d) follows from item (c) and the fact that $\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{s}$ (see (8)), and similarly that $\overline{\boldsymbol{x}}^{+\mathsf{T}}\underline{\boldsymbol{s}}^{+} = \boldsymbol{x}^{+\mathsf{T}}\boldsymbol{s}^{+}$. The proof is complete. □

Now, we are ready to present and prove the convergence result of Algorithm 5.1.

Theorem 5.1. *Let* $(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s})$ *and* $(\boldsymbol{x}^{+};\boldsymbol{y}^{+};\boldsymbol{s}^{+})$ *be strictly feasible solutions of the pair of problems* (P_μ, D_μ) *with* $(\boldsymbol{x}^{+};\boldsymbol{y}^{+};\boldsymbol{s}^{+}) = (\boldsymbol{x} + \alpha\Delta\boldsymbol{x};\ \boldsymbol{y} + \alpha\Delta\boldsymbol{y};\ \boldsymbol{s} + \alpha\Delta\boldsymbol{s})$, *where* α *is a displacement step and* $(\Delta\boldsymbol{x};\Delta\boldsymbol{y};\Delta\boldsymbol{s})$ *is the Newton direction. Then we have*

(a) $\boldsymbol{x}^{+\mathsf{T}}\boldsymbol{s}^{+} < \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}}$.

(b) $f_\mu(\boldsymbol{x}^{+}) < f_\mu(\boldsymbol{x})$.

Proof. Note that

$$\boldsymbol{x}^{+\mathsf{T}}\boldsymbol{s}^{+} = \left(1-\alpha\left(1-\frac{\sigma}{2}\right)\right)\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}} < \overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}},$$

where the equality follows from item (d) of Lemma 5.1 and the strict inequality follows from $\left(1-\alpha\left(1-\frac{\sigma}{2}\right)\right) < 1$ (as $\alpha > 0$ and $\sigma \in (0,1)$). This proves item (a).

To prove item (b), note that

$$f_\mu(\boldsymbol{x}^+) \simeq f_\mu(\boldsymbol{x}) + \nabla_{\boldsymbol{x}} f_\mu(\boldsymbol{x})^{\mathsf{T}}(\boldsymbol{x}^+ - \boldsymbol{x}),$$

and hence

$$f_\mu(\boldsymbol{x}^+) - f_\mu(\boldsymbol{x}) \simeq \alpha \nabla_{\boldsymbol{x}} f_\mu(\boldsymbol{x})^{\mathsf{T}} \Delta\boldsymbol{x}.$$

Since

$$\nabla_{\boldsymbol{x}} f_\mu(\boldsymbol{x}) = -\nabla^2_{\boldsymbol{xx}} f_\mu(\boldsymbol{x}) \Delta\boldsymbol{x},$$

we have

$$f_\mu(\boldsymbol{x}^+) - f_\mu(\boldsymbol{x}) \simeq -\alpha\, \Delta\boldsymbol{x}^{\mathsf{T}} \nabla^2_{\boldsymbol{xx}} f_\mu(\boldsymbol{x})^{\mathsf{T}} \Delta\boldsymbol{x} < 0,$$

where the strict inequality follows from the positive definiteness of the Hessian matrix $\nabla^2_{\boldsymbol{xx}} f_\mu(\boldsymbol{x})$ (as f_μ is strictly convex). Thus, $f_\mu(\boldsymbol{x}^+) < f_\mu(\boldsymbol{x})$. The proof is complete. □

6 Complexity analysis

In this section, we analyze the complexity of the proposed central trajectory algorithm for SOCP. More specifically, we prove that the iteration-complexity of Algorithm 5.1 is bounded by

$$O\left(\sqrt{2r}\ln\left[\epsilon^{-1}\boldsymbol{x}^{(0)\mathsf{T}}\boldsymbol{s}^{(0)}\right]\right).$$

Our proof depends essentially on the following two lemmas.

Lemma 6.1. *Let $(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s}) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ$, $\left(\overline{\boldsymbol{x}};\boldsymbol{y};\underline{\boldsymbol{s}}\right)$ be obtained by applying scaling to $(\boldsymbol{x};\boldsymbol{y};\boldsymbol{s})$ with $\boldsymbol{h} = \sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}$, and $\left(\overline{\Delta\boldsymbol{x}};\Delta\boldsymbol{y};\underline{\Delta\boldsymbol{s}}\right)$ be a solution of the system (7). For any $\alpha \in \mathbb{R}$, we set*

$$\begin{array}{l}(\boldsymbol{x}(\alpha);\boldsymbol{y}(\alpha);\boldsymbol{s}(\alpha)) := (\overline{\boldsymbol{x}};\boldsymbol{y};\underline{\boldsymbol{s}}) + \alpha(\overline{\Delta\boldsymbol{x}};\Delta\boldsymbol{y};\underline{\Delta\boldsymbol{s}}),\\ \mu(\alpha) := \frac{1}{2r}\boldsymbol{x}(\alpha)^{\mathsf{T}}\boldsymbol{s}(\alpha),\\ \boldsymbol{q}(\alpha) := \boldsymbol{x}(\alpha)\circ\boldsymbol{s}(\alpha) - \mu(\alpha)\boldsymbol{e}.\end{array}$$

Then

$$\boldsymbol{q}(\alpha) = (1-\alpha)(\overline{\boldsymbol{x}}\circ\underline{\boldsymbol{s}} - \mu\boldsymbol{e}) + \alpha^2\,\overline{\Delta\boldsymbol{x}}\circ\underline{\Delta\boldsymbol{s}}. \tag{13}$$

Proof. Given $\alpha \in \mathbb{R}$, using item (c) of Lemma 5.1, we have

$$\boldsymbol{x}(\alpha)^{\mathsf{T}}\boldsymbol{s}(\alpha) = (1-\alpha+\alpha\sigma)\,\overline{\boldsymbol{x}}^{\mathsf{T}}\underline{\boldsymbol{s}}, \text{ and hence } \mu(\alpha) = (1-\alpha+\alpha\sigma)\mu.$$

Thus, we get

$$\begin{array}{rcl}\boldsymbol{q}(\alpha) & = & \boldsymbol{x}(\alpha)\circ\boldsymbol{s}(\alpha) - \mu(\alpha)\boldsymbol{e}\\ & = & \left(\overline{\boldsymbol{x}} + \alpha\overline{\Delta\boldsymbol{x}}\right)\circ\left(\underline{\boldsymbol{s}} + \alpha\underline{\Delta\boldsymbol{s}}\right) - (1-\alpha+\alpha\sigma)\mu\boldsymbol{e}\\ & = & (1-\alpha)(\overline{\boldsymbol{x}}\circ\underline{\boldsymbol{s}} - \mu\boldsymbol{e}) + \alpha\underbrace{(\overline{\boldsymbol{x}}\circ\underline{\boldsymbol{s}} - \sigma\mu\boldsymbol{e})}_{-\boldsymbol{h}} + \alpha\underbrace{\left(\overline{\boldsymbol{x}}\circ\underline{\Delta\boldsymbol{s}} + \overline{\Delta\boldsymbol{x}}\circ\underline{\boldsymbol{s}}\right)}_{\boldsymbol{h}} + \alpha^2\,\overline{\Delta\boldsymbol{x}}\circ\underline{\Delta\boldsymbol{s}}\\ & = & (1-\alpha)(\overline{\boldsymbol{x}}\circ\underline{\boldsymbol{s}} - \mu\boldsymbol{e}) + \alpha^2\,\overline{\Delta\boldsymbol{x}}\circ\underline{\Delta\boldsymbol{s}}.\end{array}$$

This completes the proof. □

Lemma 6.2. *Let* $(x;y;s) \in \mathcal{F}_P^\circ \times \mathcal{F}_D^\circ$*,* $\left(\overline{x};y;\underline{s}\right)$ *be obtained by applying scaling to* $(x;y;s)$ *such that* $\left\|\overline{x}\circ\underline{s}-\mu e\right\| \leq \theta\mu$*, for some* $\theta\in[0,1)$ *and* $\mu>0$*. Let also* $\left(\overline{\Delta x};\Delta y;\underline{\Delta s}\right)$ *be a solution of the system (7),* $h=\sigma\mu e-\overline{x}\circ\underline{s}$*,* $\delta_x := \mu\left\|\overline{\Delta x}\circ\overline{x}^{-1}\right\|_F$*,* $\delta_s := \left\|\overline{x}\circ\underline{\Delta s}\right\|_F$*. Then, we have*

$$\delta_x\delta_s \leq \frac{1}{2}\left(\delta_x^2+\delta_s^2\right) \leq \frac{\|h\|_F^2}{2(1-\theta)^2}. \tag{14}$$

Proof. By the last equation of system (7) and from the operator commutativity, we have

$$h = \overline{x}\circ\underline{\Delta s} + \overline{\Delta x}\circ\underline{s} = \overline{x}\circ\underline{\Delta s} + \mu\overline{\Delta x}\circ\overline{x}^{-1} + \left(\overline{\Delta x}\circ\overline{x}^{-1}\right)\circ\left(\overline{x}\circ\underline{s}-\mu e\right).$$

It immediately follows that

$$\begin{array}{rcl} \|h\|_F & \geq & \left\|\overline{x}\circ\underline{\Delta s}+\mu\,\overline{\Delta x}\circ\overline{x}^{-1}\right\|_F - \left\|\overline{\Delta x}\circ\overline{x}^{-1}\right\|_F\left\|\overline{x}\circ\underline{s}-\mu e\right\| \\ & \geq & \left\|\overline{x}\circ\underline{\Delta s}+\mu\,\overline{\Delta x}\circ\overline{x}^{-1}\right\|_F - \theta\delta_x \\ & \geq & \sqrt{\left\|\overline{x}\circ\underline{\Delta s}\right\|_F^2+\left\|\mu\,\overline{\Delta x}\circ\overline{x}^{-1}\right\|_F^2} - \theta\delta_x \\ & = & \sqrt{\delta_x^2+\delta_s^2}-\theta\delta_x \\ & \geq & (1-\theta)\sqrt{\delta_x^2+\delta_s^2}, \end{array} \tag{15}$$

where the second inequality follows from the assumption that $\left\|\overline{x}\circ\underline{s}-\mu e\right\| \leq \theta\mu$, and the third inequality follows from (1) the fact that

$$\left(\overline{x}\circ\underline{\Delta s}\right)^{\mathsf{T}}\left(\overline{\Delta x}\circ\overline{x}^{-1}\right) = \frac{1}{2}\text{trace}\left(\left(\overline{x}\circ\underline{\Delta s}\right)\circ\left(\overline{\Delta x}\circ\overline{x}^{-1}\right)\right) = \frac{1}{2}\text{trace}\left(\underline{\Delta s}\circ\overline{\Delta x}\right) = \overline{\Delta x}^{\mathsf{T}}\underline{\Delta s},$$

which is essentially zero due to item (a) of Lemma 5.1.

The right-hand side inequality in (14) follows by noting that $(\delta_x-\delta_s)^2\geq 0$, and the left-hand side inequality in (14) follows from the last inequality in (15). The proof is complete. □

The following theorem analyzes the behavior of one iteration of Algorithm 5.1.

Theorem 6.1. *Let* $\theta\in(0,1)$ *and* $\delta\in\left[0,\sqrt{2r}\right)$ *be given such that*

$$\frac{\theta^2+\delta^2}{2(1-\theta)^2\left(1-\frac{\delta}{\sqrt{2r}}\right)} \leq \theta \leq \frac{1}{2}. \tag{16}$$

Suppose that $\left(\overline{x};y;\underline{s}\right)\in\mathcal{N}_\theta(\mu)$ *and let* $\left(\overline{\Delta x};\Delta y;\underline{\Delta s}\right)$ *denote the solution of system (7) with* $h=\sigma\mu e-\overline{x}\circ\underline{s}$ *and* $\sigma = 1-\frac{\delta}{\sqrt{2r}}$*. Then, we have*

(a) $\overline{x}^{+\mathsf{T}}\underline{s}^+ = \left(1-\frac{\delta}{\sqrt{2r}}\right)\overline{x}^{\mathsf{T}}\underline{s}$.

(b) $\left(\overline{x}^+;y^+;\underline{s}^+\right) = \left(\overline{x};y;\underline{s}\right)+\left(\overline{\Delta x};\Delta y;\underline{\Delta s}\right)\in\mathcal{N}_\theta(\mu)$.

(c) $(x^+;y^+;s^+) = (x;y;s)+(\Delta x;\Delta y;\Delta s)\in\mathcal{N}_\theta(\mu)$.

Proof. Item (a) follows directly from item (c) of Lemma 5.1 with $\alpha = 1$ and $\sigma = 1 - \frac{\delta}{\sqrt{2r}}$. We now prove item (b). Define

$$\mu^+ := \frac{1}{2r}\, \overline{\boldsymbol{x}}^{+\mathsf{T}} \underline{\boldsymbol{s}}^+ = \left(1 - \frac{\delta}{2r}\right)\mu \tag{17}$$

and let $\left(\overline{\boldsymbol{x}}; \boldsymbol{y}; \underline{\boldsymbol{s}}\right) \in \mathcal{N}_\theta(\mu)$, we then have

$$\left\|\sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right\|_F^2 \le \left\|(\sigma - 1)\mu\boldsymbol{e}\right\|_F^2 + \left\|\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right\|_F^2 \le \left((\sigma - 1)^2 2r + \theta^2\right)\mu^2 = \left(\delta^2 + \theta^2\right)\mu^2. \tag{18}$$

Since $\|\overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}} - \mu\boldsymbol{e}\| \le \theta\mu$ and $\boldsymbol{h} = \sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}$, using Lemma 6.2 it follows that

$$\left\|\overline{\Delta\boldsymbol{x}} \circ \overline{\boldsymbol{x}}^{-1}\right\|_F \left\|\overline{\boldsymbol{x}} \circ \underline{\Delta\boldsymbol{s}}\right\|_F \le \frac{\left\|\sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right\|_F^2}{2(1-\theta)^2\mu}. \tag{19}$$

Defining $\boldsymbol{q}^+ := \boldsymbol{q}(1) = \overline{\boldsymbol{x}}^+ \circ \underline{\boldsymbol{s}}^+ - \mu^+\boldsymbol{e}$ and using (13) with $\alpha = 1$, (19), (18), (16) and (17), we get

$$\|\boldsymbol{q}^+\|_F = \left\|\overline{\Delta\boldsymbol{x}} \circ \underline{\Delta\boldsymbol{s}}\right\|_F \le \left\|\overline{\Delta\boldsymbol{x}} \circ \overline{\boldsymbol{x}}^{-1}\right\|_F \left\|\overline{\boldsymbol{x}} \circ \underline{\Delta\boldsymbol{s}}\right\|_F \le \frac{\left\|\sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right\|_F^2}{2(1-\theta)^2\mu} \le \frac{(\delta^2 + \theta^2)\mu}{2(1-\theta)^2} \le \theta\left(1 - \frac{\delta}{\sqrt{2r}}\right)\mu = \theta\mu^+.$$

Consequently,

$$\left\|\overline{\boldsymbol{x}}^+ \circ \underline{\boldsymbol{s}}^+ - \mu^+\boldsymbol{e}\right\|_F \le \theta\mu^+. \tag{20}$$

By using the right-hand side inequality in (14), and using (18) and (16), we have

$$\left\|\overline{\Delta\boldsymbol{x}} \circ \overline{\boldsymbol{x}}^{-1}\right\|_F \le \frac{\left\|\sigma\mu\boldsymbol{e} - \overline{\boldsymbol{x}} \circ \underline{\boldsymbol{s}}\right\|_F}{(1-\theta)\mu} \le \frac{\sqrt{\delta^2 + \theta^2}}{(1-\theta)} \le \sqrt{2\theta\left(1 - \frac{\delta}{\sqrt{2r}}\right)} < 1,$$

where the strict inequality follows from $\theta \le \frac{1}{2}$ and $0 < 1 - \frac{\delta}{\sqrt{2r}} < 1$.

One can easily see that $\left\|\overline{\Delta\boldsymbol{x}} \circ \overline{\boldsymbol{x}}^{-1}\right\|_F < 1$ implies that $\boldsymbol{e} + \overline{\Delta\boldsymbol{x}} \circ \overline{\boldsymbol{x}}^{-1} \succ \boldsymbol{0}$, and therefore

$$\overline{\boldsymbol{x}}^+ = \overline{\Delta\boldsymbol{x}} + \overline{\boldsymbol{x}} = \left(\boldsymbol{e} + \overline{\Delta\boldsymbol{x}} \circ \overline{\boldsymbol{x}}^{-1}\right) \circ \overline{\boldsymbol{x}} \succ \boldsymbol{0}.$$

Note that, from (20), we have $\lambda_{\min}\left(\overline{\boldsymbol{x}}^+ \circ \underline{\boldsymbol{s}}^+\right) \ge (1-\theta)\mu^+ > 0$, and therefore $\overline{\boldsymbol{x}}^+ \circ \underline{\boldsymbol{s}}^+ \succ \boldsymbol{0}$. Since $\overline{\boldsymbol{x}}^+ \succ \boldsymbol{0}$ and $\overline{\boldsymbol{x}}^+$ and $\underline{\boldsymbol{s}}^+$ operator commute, we conclude that $\underline{\boldsymbol{s}}^+ \succ \boldsymbol{0}$. Using the first equation of system (7), we get

$$\underline{A}\overline{\boldsymbol{x}}^+ = \underline{A}\left(\overline{\boldsymbol{x}} + \overline{\Delta\boldsymbol{x}}\right) = \underline{A}\overline{\boldsymbol{x}} + \underline{A}\overline{\Delta\boldsymbol{x}} = \boldsymbol{b}, \quad \text{and hence} \quad \overline{\boldsymbol{x}}^+ \in \mathcal{F}_P^\circ.$$

By using the second equation of system (7), we get

$$\underline{A}^\mathsf{T}\boldsymbol{y}^+ + \overline{\boldsymbol{s}}^+ = \underline{A}^\mathsf{T}(\boldsymbol{y} + \Delta\boldsymbol{y}) + \left(\overline{\boldsymbol{s}} + \overline{\Delta\boldsymbol{s}}\right) = \underline{A}^\mathsf{T}\boldsymbol{y} + \overline{\boldsymbol{s}} + \underline{A}^\mathsf{T}\Delta\boldsymbol{y} + \overline{\Delta\boldsymbol{s}} = \boldsymbol{c}, \quad \text{and hence} \quad \left(\boldsymbol{y}^+; \overline{\boldsymbol{s}}^+\right) \in \mathcal{F}_D^\circ.$$

Thus, in view of (20), we deduce that $\left(\overline{\boldsymbol{x}}^+; \boldsymbol{y}^+; \overline{\boldsymbol{s}}^+\right) \in \mathcal{N}_\theta^\circ(\mu)$. Item (b) is therefore established. Item (c) follows from item (b) and [33, Proposition 29]. The proof is now complete. □

Corollary 6.1. *Let θ and δ as given in Theorem 6.1 and $\left(\boldsymbol{x}^0; \boldsymbol{y}^0; \boldsymbol{s}^{(0)}\right) \in \mathcal{N}_\theta(\mu)$. Then Algorithm 5.1 generates a sequence of points $\left\{\left(\boldsymbol{x}^k; \boldsymbol{y}^k; \boldsymbol{s}^{(k)}\right)\right\} \subset \mathcal{N}_\theta(\mu)$ such that*

$$\boldsymbol{x}^{(k)^\mathsf{T}} \boldsymbol{s}^{(k)} = \left(1 - \frac{\delta}{\sqrt{2r}}\right)^k \boldsymbol{x}^{(0)^\mathsf{T}} \boldsymbol{s}^{(0)}, \quad \forall k \geq 0.$$

Moreover, given a tolerance $\epsilon > 0$, Algorithm 5.1 computes an iterate $\left\{\left(\boldsymbol{x}^k; \boldsymbol{y}^k; \boldsymbol{s}^{(k)}\right)\right\}$ satisfying $\boldsymbol{x}^{(k)^\mathsf{T}} \boldsymbol{s}^{(k)} \leq \epsilon$ in at most $K = O\left(\sqrt{2r} \ln\left[\epsilon^{-1} \boldsymbol{x}^{(0)^\mathsf{T}} \boldsymbol{s}^{(0)}\right]\right)$ iterations.

Proof. Looking recursively at item (a) of Theorem 6.1, for each k we have that

$$\boldsymbol{x}^{(k)^\mathsf{T}} \boldsymbol{s}^{(k)} = \left(1 - \frac{\delta}{\sqrt{2r}}\right)^k \boldsymbol{x}^{(0)^\mathsf{T}} \boldsymbol{s}^{(0)} \leq \epsilon.$$

By taking natural algorithm of both sides, we get

$$k \ln\left(1 - \frac{\delta}{\sqrt{2r}}\right) \leq \ln\left(\frac{\epsilon}{\boldsymbol{x}^{(0)^\mathsf{T}} \boldsymbol{s}^{(0)}}\right),$$

which holds only if

$$k\left(-\frac{\delta}{\sqrt{2r}}\right) \leq \ln\left(\frac{\epsilon}{\boldsymbol{x}^{(0)^\mathsf{T}} \boldsymbol{s}^{(0)}}\right), \quad \text{or equivalently,} \quad k \geq K \geq \left[\delta^{-1} \sqrt{2r} \, \ln\left(\frac{\boldsymbol{x}^{(0)^\mathsf{T}} \boldsymbol{s}^{(0)}}{\epsilon}\right)\right].$$

The result is established. □

7 Numerical experiments

In this section, we present some numerical experiments to demonstrate the efficiency of our algorithm. Our numerical experiments are carried out on a PC with Intel(R) Dual CPU at 2.20 GHz and 2 GB of physical memory. The PC runs MATLAB Version: 7.4.0.287 (R2007a) on Windows XP Enterprise 32-bit operating system.

The proposed primal-dual central trajectory algorithm was tested on randomly generated problems. We generate a random matrix $A \in \mathbb{R}^{n \times m}$ with full row rank and random vectors $\boldsymbol{x}, \boldsymbol{s} \in \mathcal{E}_+^n$ and $\boldsymbol{y} \in \mathbb{R}^m$. We let $\boldsymbol{b} = A\boldsymbol{x}$ and $\boldsymbol{c} = A^\mathsf{T}\boldsymbol{y} + \boldsymbol{s}$. Let $\boldsymbol{x}^{(0)} = \boldsymbol{e}_n, \boldsymbol{y}^{(0)} = \boldsymbol{0} \in \mathbb{R}^m$ and $\boldsymbol{s}^{(0)} = \boldsymbol{e}_n$ be initial points. Since the set of strictly feasible solutions of (P) and (D) are non-empty, the generated test problems (P) and (D) have optimal solutions and their optimal values are equal. In all of our experiments, we used the NT direction for choosing the scaling vector $\boldsymbol{p}^{(k)} \in C\left(\boldsymbol{x}^{(k)}, \boldsymbol{s}^{(k)}\right)$ because this direction has the moral excellence of being primal-dual symmetric. The parameters used in Algorithm 5.1 were as follows: $\epsilon = 1.3500e - 06, \sigma = 0.1$ and $\rho = 0.99$.

In our experiments, we generated test problems with size $n(= 2m)$ from 100 to 1200 and $r = 1$. The random test problems of each n are generated 5 times, and hence we have 60 random problems in total. As a well-known interior-point method for solving SOCP, the SDPT3 solver [40] was used in our experiments for comparison purposes. We present the results of our numerical results in Table 7.1, in which "(n, m)" denotes the size of problems, "1$^{\text{st}}$ Sel." denotes the first selection for calculating the displacement step, "2$^{\text{nd}}$ Sel." denotes the second selection for calculating the displacement step, "3$^{\text{rd}}$ Sel." denotes the third selection for calculating the displacement step,

"4^{th} Sel." denotes the fourth selection for calculating the displacement step, "Iter" denotes the number of iterations taken to obtain the optimal solution, and "CPU(s)" denotes the time (in seconds) required to obtain the optimal solution. Note that the values of "Iter" and "CPU(s)" are the average of 5 runs for each n.

Problem size	1^{st} Sel. Alg. 5.1		2^{nd} Sel. Alg. 5.1		3^{rd} Sel. Alg. 5.1		4^{th} Sel. Alg. 5.1		SDPT3	
(m,n)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)
$(50,100)$	5.6	0.0931	5.6	0.0993	8.3	0.1539	5.2	0.0881	6.5	0.1132
$(100,200)$	5.8	0.3107	5.7	0.2981	8.4	0.6913	5.4	0.2840	6.6	0.4347
$(150,300)$	6.1	1.0134	6.2	0.9408	9.4	1.3423	5.8	0.7761	6.6	1.1759
$(200,400)$	6.3	2.2150	6.4	2.0935	8.9	2.8031	6.2	1.8420	7.2	2.1036
$(250,500)$	6.4	2.6019	6.5	2.7142	9.5	3.2401	6.1	2.2245	7.5	2.9518
$(300,600)$	6.4	3.8210	6.4	3.9101	9.8	6.1309	6.3	3.2091	7.9	4.7671
$(350,700)$	6.5	6.0134	6.6	6.2013	9.4	7.8431	6.4	5.5219	7.7	6.9383
$(400,800)$	6.8	8.1394	6.7	8.3405	9.7	9.8410	6.4	7.3490	8.4	9.2521
$(450,900)$	6.8	10.951	6.8	11.302	10.0	14.028	6.7	9.1918	8.5	12.3401
$(500,1000)$	6.9	15.213	6.9	16.029	10.2	18.721	6.7	13.879	7.9	17.310
$(550,1100)$	7.0	16.090	7.2	16.881	11.1	22.411	6.9	15.109	8.3	19.894
$(600,1200)$	7.3	20.132	7.4	21.094	11.4	27.913	7.1	18.938	8.7	23.431

Table 7.1. Comparison of Algorithm 5.1 (equipped with four selections) and SDPT3 on SOCPs.

In view of Table 7.1, we can see that Algorithm 5.1 with the first, second or fourth selections are able to give an optimal solution with a favorable running time and requires less number of iterations than the SDPT3 solver; this is probably due to the fact that the SDPT3 solver uses infeasible initial points. We can also see that the fourth selection for calculating the displacement step has remarkable superiority to other three selections in terms of number of iterations and running time. It is interesting to note that our conclusion exactly matches the same conclusion obtained from the numerical experiments in [1, Section 5] for solving semidefinite programming problems.

8 Conclusions

In this paper, we have presented a primal-dual central trajectory interior-point method for solving the second-order cone programming problem. We have demonstrated the existence and uniqueness of the optimal solution of the corresponding perturbed problem and have proven its convergence to the optimal solution of the original problem when the barrier parameter goes to zero. Then we have applied Newton's method to find a new iterative point by computing a good descent direction. The inconvenience lies in the high computational cost motivated us to avoid using several methods, such as the line search methods, to calculate the displacement step. Alternatively, in this paper, we have proposed a new approach based on four new selections to calculate the displacement step.

After stating the central trajectory algorithm, we have analyzed the convergence of the proposed algorithm and have proven that the complexity for short-step is bounded by $O(\sqrt{2r}\ln[\epsilon^{-1}\boldsymbol{x}^{(0)^{\mathsf{T}}}\boldsymbol{s}^{(0)}])$ iterations. Our numerical results have demonstrated the efficiency of our approach and have shown the convergence of the four proposed selections to the optimal solution of the problem. By looking at the computed time and number of iterations, the numerical results have shown that the fourth selection is the best selection that can be chosen to reach the optimal solution. Finally, we point out the analysis and results of this paper can be extended word-by-word to general symmetric cone programming using the machinery of Euclidean Jordan algebras.

References

[1] I. Touil, D. Benterki, A. Yassine, A feasible primal-dual interior point method for linear semidefinite programming, J. Comput. Appl. Math. 312 (2017) 216–230.

[2] D. Benterki, J.P. Crouzeix, B. Merikhi, A numerical feasible interior point method for linear semidefinite programs, RAIRO Oper. Res. 41 (2007) 49–59.

[3] J.P. Crouzeix, B. Merikhi, A logarithm barrier method for semidefinite programming, RAIRO Oper. Res. 42 (2008) 123–139.

[4] L. Mennichea, D. Benterki, A logarithmic barrier approach for linear programming, J. Comput. Appl. Math. 312 (2017) 267–275.

[5] D. Benterki, J.-P. Crouzeix, B. Merikhi, A numerical implementation of an interior point method for semi-definite programming. Pesquisa Operacional 23 (2003) 49–59.

[6] Z. Kebbiche, A. Keraghel, A. Yassine, Extension of a projective interior point method for linearly constrained convex programming, Appl. Math. Comput. 193 (2007) 553–559.

[7] D. Hertog, Interior Point Approach to Linear Quadratic and Convex Programming, Kluwer Academic Publishers, Dordrecht, 1994.

[8] Y.E. Nesterov, A. Nemiroveskii, Interior-Point Polynomial Algorithms in Convex Programming, SIAM, 1994.

[9] C. Roos, T. Terlaky, J.P. Vial, Theory and Algorithms for Linear Optimization: An Interior Point Approach, W. John & Sons, 1997.

[10] Y. Ye, Interior Point Algorithms: Theory and Analysis, in: Wiley-Interscience Series in Discrete Mathematics Optimization, John wiley & Sons, New York, 1997.

[11] R. Naseri, A. Valinejad, An extended variant of Karmarkars interior point algorithm, Appl. Math. Comput. 184 (2007) 737–742.

[12] M. Achache, A weighted path-following method for the linear complementarity problem, Studia Univ. Babes-Bolyai Ser. Inform. 48 (2004) 61–73.

[13] Z. Darvay, A weighted-path-following method for linear optimization, Studia Univ. Babes-Bolyai Ser. Inform. 47 (2002) 3–12.

[14] F. Alizadeh, J.-P. Haberly, M.-L. Overton, Primal-dual interior-point methods for semidefinite programming, convergence rates, stability and numerical results. SIAM J. Optim. 8 (1998) 746–768.

[15] S. Wright, Primal-Dual Interior Point Methods, SIAM, Philadelphia, 1997.

[16] I. Lustig, Computational experience with a primal-dual interior point method for linear programming, Linear Alg. Appl. 152 (1991) 191–222.

[17] D. Benterki, A. Leulmi, An improving procedure of the interior projective method for linear programming, Appl. Math. Comput. 199 (2008) 811–819.

[18] M. Dodani, A. Babu, Karmarkar's projective method for linear programming: A computational appraisal, Comput. Indust. Eng. 16 (1989) 189–206.

[19] M. Todd, Y. Wang, On combined phase 1-phase 2 projective methods for linear programming, Algorithmica. 9 (1993) 64–83.

[20] P. Gahinet, A. Nemirovski, The projective method for solving linear matrix inequalities, Math. Prog. 77 (1997) 163–190.

[21] S. Kettab, D. Benterki, A relaxed logarithmic barrier method for semidefinite programming, RAIRO Oper. Res. 42 (2015) 555–568.

[22] R. Behling, C. Gonzaga, G. Haeser, Primal-dual relationship between LevenbergMarquardt and central trajectories for linearly constrained convex optimization, J. Optim. Theory Appl. 162 (2014) 705–717.

[23] N. Gould, D. Orban, D. Robinson, Trajectory-following methods for large-scale degenerate convex quadratic programming, Math. Prog. Comput. 5 (2013) 113–142.

[24] F. Alizadeh, D. Goldfarb, Second-order cone programming. Math. Program. Ser. B 95 (2003) 3–51.

[25] J. Faraut, A. Korányi, Analysis on Symmetric Cones, Oxford University Press, Oxford, UK, 1994.

[26] Y.-J. Hans, D. Mittelmann, Interior point methods for second-order cone programming and OR applications. Comput. Optim. Appl. 28 (2004) 255–285.

[27] B. Kheirfam, A corrector-predictor path-following method for second-order cone optimization. Int. J. Computer Math. 93 (2016) 2064–2078.

[28] Z. Hao, Z. Wan, X. Chi, J. Chen, A power penalty method for second-order cone nonlinear complementarity problems. J. Comput. Appl. Math. 290 (2015) 136–149.

[29] A. Auslender, An exact penalty method for nonconvex problems covering, in particular, nonlinear programming, semidefinite programming, and second-order cone programming. SIAM J. Optim. 25 (2015) 1732–1759.

[30] A. Keraghel, Etude adaptative et comparative des principales variantes de l'algorithme de Karmarkar (Thse de Doctorat), Universit Joseph Fourier, Grenoble, France, 1989.

[31] H. Wolkowicz, G.-P.-H. Styan, Bounds for eigenvalues using traces. Linear Algebra Appl. 29 (1980) 471–506.

[32] J. Tang, G. He, L. Dong, L. Fang, A new one-step smoothing newton method for second-order cone programming. Applications of Mathematics, 57 (2012) 311–331.

[33] S.H. Schmieta, F. Alizadeh, Extension of primal-dual interior point methods to symmetric cones. Math. Program. Ser. A 96 (2003) 409–438.

[34] R.D. Monteiro, Primal-dual path-following algorithms for semidefinite programming. SIAM J. Optim. 7 (1997) 663–678.

[35] Y. Zhang, On extending primal-dual interior-point algorithms from linear programing to semidefinite programming. SIAM J. Optim. 8 (1998) 356–386.

[36] C. Helmberg, F. Rendl, R.J. Vanderbei, H. Wolkowicz, An interior-point methods for stochastic semidefinite programming. SIAM J. Optim. 6 (1996) 342–361.

[37] M. Kojima, S. Shindoh, S. Hara, Interior-point methods for the monotone linear complementarity problem in symmetric matrices. SIAM J. Optim. 7 (9) (1997) 86–125.

[38] Y.E. Nesterov, M.J. Todd, Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. 8 (2) (1998) 324–364.

[39] H. Lütkepohl, Handbook of Matrices, Humboldt-Universität zu Berlin, Germany, 1996.

[40] K.C. Toh, M.J. Todd, R.H Tutuncu, SDPT3 version 4.0–a MATLAB software for semidefinite-quadratic-linear programming, (2009). Available online at: http://www.math.nus.edu.sg/~mattohkc/sdpt3.html