Primal-dual path-following algorithms for circular programming

Baha Alzalg*

Department of Mathematics, The University of Jordan, Amman 11942, Jordan
July, 2015

Abstract

Circular programming problems are a new class of convex optimization problems in which we minimize linear function over the intersection of an affine linear manifold with the Cartesian product of circular cones. It has very recently been discovered that, unlike what has previously been believed, circular programming is a special case of symmetric programming, where it lies between second-order cone programming and semidefinite programming. Monteiro [SIAM J. Optim. 7 (1997) 663–678] introduced primal-dual path-following algorithms for solving semidefinite programming problems. Alizadeh and Goldfarb [Math. Program. Ser. A 95 (2003) 3–51] introduced primal-dual path-following algorithms for solving second-order cone programming problems. In this paper, we utilize their work and use the machinery of Euclidean Jordan algebras to derive primal-dual path-following interior point algorithms for circular programming problems. We prove polynomial convergence of the proposed algorithms by showing that the circular logarithmic barrier is a strongly self-concordant barrier.

Keywords: Circular cones; Second-order cone programming; Interior point methods; Euclidean Jordan algebra; Self-concordance

1 Introduction

We introduce and study the primal-dual pair of *circular programming* (CP) problems:

$$\begin{array}{llcll} \min & \langle \boldsymbol{c}, \boldsymbol{x} \rangle_\theta & \qquad & \max & \boldsymbol{b}^\mathsf{T} \boldsymbol{y} \\ \text{s.t.} & (A\boldsymbol{x})_\theta = \boldsymbol{b}, & & \text{s.t.} & A^\mathsf{T}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \\ & \boldsymbol{x} \in \mathcal{Q}^n_\theta, & & & \boldsymbol{s} \in \mathcal{Q}^n_\theta, \boldsymbol{y} \in \mathbb{R}^m, \end{array} \tag{1}$$

where $\theta \in (0, \pi/2)$ is a given angle, the inner product $\langle \boldsymbol{c}, \boldsymbol{x} \rangle_\theta$ denotes the *circular inner product* between $\boldsymbol{c}$ and $\boldsymbol{x}$, which we define as $\langle \boldsymbol{c}, \boldsymbol{x} \rangle_\theta := \boldsymbol{c}^\mathsf{T} I^2_{\theta,n} \boldsymbol{x}$, the matrix-vector product $(A\boldsymbol{x})_\theta$ denotes the *circular matrix-vector product* between A and $\boldsymbol{x}$, which we define as $(A\boldsymbol{x})_\theta := A I^2_{\theta,n} \boldsymbol{x}$, the matrix $I_{\theta,n}$ is called the *circular identity matrix*, which we define as

$$I_{\theta,n} := \begin{bmatrix} 1 & \boldsymbol{0}^\mathsf{T} \\ \boldsymbol{0} & \cot\theta \, I_{n-1} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

*Corresponding author: `b.alzalg@ju.edu.jo`

to act as a generalization of the (standard) identity matrix $I_n = I_{\frac{\pi}{4},n}$, the cone $\mathcal{Q}_\theta^n$ denotes the *circular cone* [1, 2] of dimension n, which is defined as

$$\mathcal{Q}_\theta^n := \left\{ \left[\begin{array}{c} x_0 \\ \bar{\boldsymbol{x}} \end{array} \right] \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \cot\theta \, \|\bar{\boldsymbol{x}}\| \right\}, \tag{2}$$

and, finally, the norm $\|\cdot\|$ denotes the standard Euclidean norm.

An important special case is that in which $\theta = \frac{\pi}{4}$. In this case, the circular cone $\mathcal{Q}_\theta^n$ reduces to the well-known second-order cone $\mathcal{Q}^n$, which is defined as

$$\mathcal{Q}^n := \left\{ \left[\begin{array}{c} x_0 \\ \bar{\boldsymbol{x}} \end{array} \right] \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \|\bar{\boldsymbol{x}}\| \right\} = \mathcal{Q}_{\frac{\pi}{4}}^n,$$

the circular identity matrix $I_{\theta,n}$ reduces to the identity matrix I_n, the circular inner product between $\boldsymbol{c}$ and $\boldsymbol{x}$ (i.e., $\langle \boldsymbol{c}, \boldsymbol{x} \rangle_\theta = \boldsymbol{c}^\mathsf{T} I_{\theta,n}^2 \boldsymbol{x}$) reduces to the standard inner product between $\boldsymbol{c}$ and $\boldsymbol{x}$ (i.e., $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}$), the circular matrix-vector product between A and $\boldsymbol{x}$ (i.e., $A I_{\theta,n}^2 \boldsymbol{x}$) reduces to the (standard) matrix-vector product between A and $\boldsymbol{x}$ (i.e., $A\boldsymbol{x}$), and, therefore, the CP problems (1) reduces to second-order cone programming problems [3]. So, CP includes second-order cone programming as a special case. We will also see that the CP is a special case of semidefinite programming [4].

It is known that circular cone is a closed, convex, pointed, solid cone. It was also popularly known (see for example [1, 2]) that the dual of the circular cone (2), denoted by $\mathcal{Q}_\theta^{n\star}$, is the circular cone

$$\mathcal{Q}_{\frac{\pi}{2}-\theta}^n = \left\{ \left[\begin{array}{c} x_0 \\ \bar{\boldsymbol{x}} \end{array} \right] \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \cot\left(\frac{\pi}{2} - \theta\right) \|\bar{\boldsymbol{x}}\| \right\} = \left\{ \left[\begin{array}{c} x_0 \\ \bar{\boldsymbol{x}} \end{array} \right] \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \tan\theta \, \|\bar{\boldsymbol{x}}\| \right\}.$$

In [10], we have shown that, unlike popularly perceived, this fact (i.e., the fact that $\mathcal{Q}_\theta^{n\star} = \mathcal{Q}_{\frac{\pi}{2}-\theta}^n$) is no longer true. Moreover, we have shown that, under the circular inner product, the circular cone $\mathcal{Q}_\theta^n$ is self-dual (i.e., $\mathcal{Q}_\theta^{n\star} = \mathcal{Q}_\theta^n$; see Lemma 2) and homogeneous, hence it becomes a symmetric cone. In fact, there is a one-to-one correspondence between Euclidean Jordan algebras and symmetric cones. As a result, the circular cone is indeed the cone of squares of some Euclidean Jordan algebra. This motivates us to establish a Euclidean Jordan algebra associated with the circular cone. In [10], we have set up the Jordan algebra associated with the circular cone by establishing a new spectral decomposition (different and much more efficient than what has been established in [1]) associated with this cone and establishing a circular Jordan multiplication associated with its algebra. We have also demonstrated that this algebra forms a Euclidean Jordan algebra with the circular inner product.

Being a symmetric cone, circular programming becomes one of the special cases of symmetric programming [5]. In fact, linear programming, second-order cone programming, CP and semidefinite programming are considered the most important subclasses of symmetric programming. The applications of the circular cone lie in various real-world engineering problems, for example the optimal grasping manipulation problems for multi-fingered robots [6, 7, 8]. In addition, CP is applied in the perturbation analysis of second-order cone programming problems [9].

Monteiro [11] introduced primal-dual path-following algorithms for solving semidefinite programming (see also Zhang [12]). Alizadeh and Goldfarb [3] introduced primal-dual path-following algorithms for solving second-order cone programs. These primal-dual interior point algorithms and their analysis have been extended by Schmieta and Alizadeh [5] for solving optimization problems over all symmetric cones. The purpose of this paper is to utilize the work of Monteiro [11], Alizadeh and Goldfarb [3] and Schmieta and Alizadeh [5] to derive primal-dual path-following algorithms for the primal-dual pair of CP problems (1). We prove the polynomial convergence results of the proposed algorithms by showing that the logarithmic barrier in the circular case is a strongly self-concordant barrier [13].

This paper is organized as follows. We end this section by introducing some notations that will be used in the sequel, and then proving the self-duality of the circular cone. The Euclidean Jordan algebra of the circular cones mentioned above is the substance of Section 2. In Section 3, we introduce the logarithmic barrier in the circular case for our problem formulation, and then state the self-concordance property of this barrier. Based on this property, we write the optimality conditions for the CP problems (1) and describe the commutative class of directions for the central path in Section 4, then we present short-, semi-long-, and long-step variants of the path-following algorithm for CP in Section 5. Section 6 is devoted to prove the self-concordance property stated in Section 3. The last section contains some concluding remarks.

1.1 Notations

We use "," for adjoining vectors and matrices in a row, and use ";" for adjoining them in a column. So, for example, if $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$ are vectors, we have

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix} = (\boldsymbol{x}^{\mathsf{T}}, \boldsymbol{y}^{\mathsf{T}}, \boldsymbol{z}^{\mathsf{T}})^{\mathsf{T}} = (\boldsymbol{x}; \boldsymbol{y}; \boldsymbol{z}).$$

We use $\mathbb{R}$ to denote the field of real numbers. For each vector $\boldsymbol{x} \in \mathbb{R}^n$ whose first entry is indexed with 0, we write $\overline{\boldsymbol{x}}$ for the subvector consisting of entries 1, through $n-1$ (therefore $\boldsymbol{x} = (x_0; \overline{\boldsymbol{x}}) \in \mathbb{R} \times \mathbb{R}^{n-1}$). We let $\mathcal{E}^n$ denote the n dimensional real vector space $\mathbb{R} \times \mathbb{R}^{n-1}$ whose elements $\boldsymbol{x}$ are indexed with 0. For a vector $\boldsymbol{x} \in \mathcal{E}^n$ and a matrix $A \in \mathbb{R}^{m \times n}$, we let

$$\begin{array}{ll} \boldsymbol{x}_\theta := (x_0; \cot\theta\, \overline{\boldsymbol{x}}) = I_{\theta,n} \boldsymbol{x}, & A_\theta := A I_{\theta,n}, \\ \boldsymbol{x}_{\theta^2} := (x_0; \cot^2\theta\, \overline{\boldsymbol{x}}) = I_{\theta,n}^2 \boldsymbol{x}, & A_{\theta^2} := A I_{\theta,n}^2. \end{array} \tag{3}$$

Hence $\overline{\boldsymbol{x}_\theta} = \cot\theta\, \overline{\boldsymbol{x}}$, and $\|\overline{\boldsymbol{x}_\theta}\| = \cot\theta \|\overline{\boldsymbol{x}}\|$. Therefore, we can now rewrite the circular cone $\mathcal{Q}_\theta^n$ as follows

$$\mathcal{Q}_\theta^n = \{\boldsymbol{x} \in \mathcal{E}^n : x_0 \geq \|\overline{\boldsymbol{x}_\theta}\|\}.$$

As a result,

$$\boldsymbol{x} \in \mathcal{Q}_\theta^n \quad \text{if and only if} \quad \boldsymbol{x}_\theta \in \mathcal{Q}^n. \tag{4}$$

We can also rewrite the circular inner product between two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^n$ as follows

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_\theta := \boldsymbol{x}^{\mathsf{T}} I_{\theta,n}^2 \boldsymbol{y} = \boldsymbol{x}_\theta^{\mathsf{T}} \boldsymbol{y}_\theta = \boldsymbol{x}_{\theta^2}^{\mathsf{T}} \boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}} \boldsymbol{y}_{\theta^2} = x_0 y_0 + \cot^2\theta\, \overline{\boldsymbol{x}}^{\mathsf{T}} \overline{\boldsymbol{y}}, \tag{5}$$

and rewrite the circular matrix-vector product between a matrix $A \in \mathbb{R}^{n \times n}$ and a vector $\boldsymbol{x} \in \mathcal{E}^n$ as follows

$$(A\boldsymbol{x})_\theta := A I_{\theta,n}^2 \boldsymbol{x} = A_\theta \boldsymbol{x}_\theta = A_{\theta^2} \boldsymbol{x}. \tag{6}$$

In general, we define the *circular matrix product* between two matrices $A, B \in \mathbb{R}^{n \times n}$ as follows

$$(AB)_\theta := A I_{\theta,n}^2 B = A_{\theta^2} B.$$

1.2 Self-duality of the circular cone

Lemma 1 illustrates why the circular cone $\mathcal{L}_t^n$, with the standard inner product, is not self-dual, and illustrates the reason why it is popularly known that the dual of $\mathcal{Q}_\theta^{n\star}$ is $\mathcal{Q}_{\frac{\pi}{2}-\theta}^n$.

Lemma 1. Under the standard inner product, the circular cone $\mathcal{Q}_\theta^{n\star}$ is not self-dual. Moreover, $\mathcal{Q}_\theta^{n\star} = \mathcal{Q}_{\frac{\pi}{2}-\theta}^n$.

Proof. Under the standard inner product, the dual of $\mathcal{Q}_\theta^n$ is defined as

$$\mathcal{Q}_\theta^{n\star} := \{\boldsymbol{x} \in \mathcal{E}^n : \boldsymbol{x}^\mathsf{T}\boldsymbol{y} \geq 0,\ \forall \boldsymbol{y} \in \mathcal{Q}_\theta^n\}.$$

We first show that $\mathcal{Q}_{\frac{\pi}{2}-\theta}^n \subseteq \mathcal{Q}_\theta^{n\star}$. Let $\boldsymbol{x} = (x_0; \overline{\boldsymbol{x}}) \in \mathcal{Q}_{\frac{\pi}{2}-\theta}^n$, we show that $\boldsymbol{x} \in \mathcal{Q}_\theta^{n\star}$ by verifying that $\boldsymbol{x}^\mathsf{T}\boldsymbol{y} \geq 0$ for any $\boldsymbol{y} \in \mathcal{Q}_\theta^n$. So let $\boldsymbol{y} = (y_0; \overline{\boldsymbol{y}}) \in \mathcal{Q}_\theta^n$. Then

$$\begin{array}{rcl}
\boldsymbol{x}^\mathsf{T}\boldsymbol{y} & = & x_0\, y_0 + \overline{\boldsymbol{x}}^\mathsf{T}\overline{\boldsymbol{y}} \\
& \geq & \left(\cot\left(\frac{\pi}{2} - \theta\right) \|\overline{\boldsymbol{x}}\|\right)\ (\cot\theta\, \|\overline{\boldsymbol{y}}\|) + \overline{\boldsymbol{x}}^\mathsf{T}\overline{\boldsymbol{y}} \\
& = & (\tan\theta\, \|\overline{\boldsymbol{x}}\|)\, (\cot\theta\, \|\overline{\boldsymbol{y}}\|) + \overline{\boldsymbol{x}}^\mathsf{T}\overline{\boldsymbol{y}} \\
& = & \|\overline{\boldsymbol{x}}\| \|\overline{\boldsymbol{y}}\| + \overline{\boldsymbol{x}}^\mathsf{T}\overline{\boldsymbol{y}} \\
& \geq & |\overline{\boldsymbol{x}}^\mathsf{T}\overline{\boldsymbol{y}}| + \overline{\boldsymbol{x}}^\mathsf{T}\overline{\boldsymbol{y}} \geq 0,
\end{array}$$

where the first inequality follows from the fact that $\boldsymbol{x} \in \mathcal{Q}_{\frac{\pi}{2}-\theta}^n$ and $\boldsymbol{y} \in \mathcal{Q}_\theta^n$, and the last follows from Hölder's inequality. Thus, $\mathcal{Q}_{\frac{\pi}{2}-\theta}^n \subseteq \mathcal{Q}_\theta^{n\star}$.

Now, we prove that $\mathcal{Q}_\theta^{n\star} \subseteq \mathcal{Q}_{\frac{\pi}{2}-\theta}^n$. Let $\boldsymbol{y} = (y_0; \overline{\boldsymbol{y}}) \in \mathcal{Q}_\theta^{n\star}$, and consider

$$\boldsymbol{x} := (\tan\theta \|\overline{\boldsymbol{y}}\|; -\overline{\boldsymbol{y}}) = \left(\cot\left(\frac{\pi}{2} - \theta\right) \|\overline{\boldsymbol{y}}\|; -\overline{\boldsymbol{y}}\right) \in \mathcal{Q}_{\frac{\pi}{2}-\theta}^n.$$

Then by using Hölder's inequality, where the equality is attained, we obtain

$$0 \leq \boldsymbol{x}^\mathsf{T}\boldsymbol{y} = \tan\theta\ y_0 \|\overline{\boldsymbol{y}}\| - \overline{\boldsymbol{y}}^\mathsf{T}\overline{\boldsymbol{y}} = \tan\theta\ y_0 \|\overline{\boldsymbol{y}}\| - \|\overline{\boldsymbol{y}}\|^2 = \|\overline{\boldsymbol{y}}\|(\tan\theta\ y_0 - \|\overline{\boldsymbol{y}}\|).$$

This implies that $\tan\theta\, y_0 \geq \|\overline{\boldsymbol{y}}\|$, or equivalently $y_0 \geq \cot\theta\, \|\overline{\boldsymbol{y}}\|$. Hence, $\boldsymbol{y} \in \mathcal{Q}_\theta^n$. This demonstrates that $\mathcal{Q}_{\frac{\pi}{2}-\theta}^n = \mathcal{Q}_\theta^{n\star}$ and completes the proof. □

In fact, the inner product that should be considered with the circular cones is the circular inner product defined in (5). Under the circular inner product, the dual of $\mathcal{Q}_\theta^n$ is defined as

$$\mathcal{Q}_\theta^{n\star} := \{\boldsymbol{x} \in \mathcal{E}^n : \boldsymbol{x}_\theta^\mathsf{T}\boldsymbol{y}_\theta \geq 0,\ \forall \boldsymbol{y} \in \mathcal{Q}_\theta^n\}.$$

The proof of the following lemma can be found in [10, Section 3] and is similar in its composition to the proof of Lemma 1. We prefer to present a shorter and more direct proof because it takes the advantage of using our notations.

Lemma 2. Under the circular inner product, the circular cone $\mathcal{Q}_\theta^n$ is self-dual. That is, $\mathcal{Q}_\theta^{n\star} = \mathcal{Q}_\theta^n$.

Proof. The proof follows from the following equivalences:

$$\begin{array}{rcll}
\boldsymbol{x} \in \mathcal{Q}_\theta^n & \Longleftrightarrow & \boldsymbol{x}_\theta \in \mathcal{Q}^n & \text{(by (4))} \\
& \Longleftrightarrow & \boldsymbol{x}_\theta \in \mathcal{Q}^{n\star} & \text{(as } \mathcal{Q}^n = \mathcal{Q}^{n\star}) \\
& \Longleftrightarrow & \boldsymbol{x}_\theta^\mathsf{T}\boldsymbol{y}_\theta \geq 0,\ \forall \boldsymbol{y}_\theta \in \mathcal{Q}^n & \text{(by definition of } \mathcal{Q}^{n\star}) \\
& \Longleftrightarrow & \boldsymbol{x}_\theta^\mathsf{T}\boldsymbol{y}_\theta \geq 0,\ \forall \boldsymbol{y} \in \mathcal{Q}_\theta^n & \text{(by (4))} \\
& \Longleftrightarrow & \boldsymbol{x}_\theta \in \mathcal{Q}_\theta^{n\star} & \text{(by definition of } \mathcal{Q}_\theta^{n\star}).
\end{array}$$

□

2 The Euclidean Jordan algebra of the circular cone

Let $\boldsymbol{x} \in \mathcal{E}^n$. The *circular spectral decomposition* of $\boldsymbol{x}$ with respect to the angle $\theta \in (0, \pi/2)$ is obtained as follows [10]

$$\begin{aligned}
\boldsymbol{x} &= (x_0 + \cot\theta \|\overline{\boldsymbol{x}}\|)\left(\tfrac{1}{2}\right)\begin{bmatrix} 1 \\ \dfrac{\tan\theta\,\overline{\boldsymbol{x}}}{\|\overline{\boldsymbol{x}}\|} \end{bmatrix} + (x_0 - \cot\theta \|\overline{\boldsymbol{x}}\|)\left(\tfrac{1}{2}\right)\begin{bmatrix} 1 \\ -\dfrac{\tan\theta\,\overline{\boldsymbol{x}}}{\|\overline{\boldsymbol{x}}\|} \end{bmatrix} \\
&= \underbrace{(x_0 + \|\overline{\boldsymbol{x}_\theta}\|)}_{\lambda_{\theta,1}(\boldsymbol{x})} \underbrace{\left(\frac{1}{2}\right)\begin{bmatrix} 1 \\ \dfrac{\overline{\boldsymbol{x}}}{\|\overline{\boldsymbol{x}_\theta}\|} \end{bmatrix}}_{\boldsymbol{c}_{\theta,1}(\boldsymbol{x})} + \underbrace{(x_0 - \|\overline{\boldsymbol{x}_\theta}\|)}_{\lambda_{\theta,2}(\boldsymbol{x})} \underbrace{\left(\frac{1}{2}\right)\begin{bmatrix} 1 \\ -\dfrac{\overline{\boldsymbol{x}}}{\|\overline{\boldsymbol{x}_\theta}\|} \end{bmatrix}}_{\boldsymbol{c}_{\theta,2}(\boldsymbol{x})},
\end{aligned} \tag{7}$$

where $\boldsymbol{x}_\theta$ is defined in (3). The spectral decomposition (7) is associated with the circular cone, and it is viewed as a generalization of the spectral decomposition in [3, Section 4] which is associated with the second-order cone. Under the circular spectral decomposition (7), we have that

$$\operatorname{trace}(\boldsymbol{x}) := \lambda_{\theta,1}(\boldsymbol{x}) + \lambda_{\theta,2}(\boldsymbol{x}) = 2x_0, \quad \det_\theta(\boldsymbol{x}) := \lambda_{\theta,1}(\boldsymbol{x})\lambda_{\theta,2}(\boldsymbol{x}) = x_0^2 - \|\overline{\boldsymbol{x}_\theta}\|^2,$$

and that $\boldsymbol{c}_{\theta,1} + \boldsymbol{c}_{\theta,2} = \boldsymbol{e} := (1; \mathbf{0})$ which is the *identity* element of $\mathcal{E}^n$. It is quite easy to see that $\operatorname{trace}(\boldsymbol{e}) = 2, \det(\boldsymbol{e}) = 1, \lambda_{\theta,1}(\boldsymbol{c}_{\theta,1}) = \lambda_{\theta,1}(\boldsymbol{c}_{\theta,2}) = 1, \lambda_{\theta,2}(\boldsymbol{c}_{\theta,1}) = \lambda_{\theta,2}(\boldsymbol{c}_{\theta,2}) = 0, \boldsymbol{c}_{\theta,1} = R\boldsymbol{c}_{\theta,2}$, and $\boldsymbol{c}_{\theta,2} = R\boldsymbol{c}_{\theta,1}$, where R is the reflection matrix

$$R := \begin{bmatrix} 1 & \mathbf{0}^\mathsf{T} \\ \mathbf{0} & -I_{n-1} \end{bmatrix}. \tag{8}$$

For any real valued continuous function f_θ, we define the image of $\boldsymbol{x}$ under f_θ with respect to θ as

$$f_\theta(\boldsymbol{x}) := f_\theta(\lambda_{\theta,1}(\boldsymbol{x}))\boldsymbol{c}_{\theta,1}(\boldsymbol{x}) + f_\theta(\lambda_{\theta,2}(\boldsymbol{x}))\boldsymbol{c}_{\theta,2}(\boldsymbol{x}).$$

For instance, for $p \in \mathbb{R}$, $\boldsymbol{x}^p := \lambda_{\theta,1}^p(\boldsymbol{x})\boldsymbol{c}_{\theta,1}(\boldsymbol{x}) + \lambda_{\theta,2}^p(\boldsymbol{x})\boldsymbol{c}_{\theta,2}(\boldsymbol{x})$. In particular, with a little calculation, we can obtain

$$\boldsymbol{x}^{-1} := \frac{1}{\lambda_{\theta,1}(\boldsymbol{x})}\boldsymbol{c}_{\theta,1}(\boldsymbol{x}) + \frac{1}{\lambda_{\theta,2}(\boldsymbol{x})}\boldsymbol{c}_{\theta,2}(\boldsymbol{x}) = \frac{1}{\det_\theta(\boldsymbol{x})}\begin{bmatrix} x_0 \\ -\overline{\boldsymbol{x}} \end{bmatrix} = \frac{R}{\det_\theta(\boldsymbol{x})}\,\boldsymbol{x},$$

which is called the *inverse* of $\boldsymbol{x}$ (provided that $\boldsymbol{x}$ is invertible, i.e., $\det_\theta(\boldsymbol{x}) \neq 0$).

The *Frobenius norm* with respect to θ of $\boldsymbol{x}$ is defined as $\|\boldsymbol{x}\|_{\theta,F} := \sqrt{\lambda_{\theta,1}^2(\boldsymbol{x}) + \lambda_{\theta,2}^2(\boldsymbol{x})}$, and the *2-norm* of $\boldsymbol{x}$ with respect to θ is defined as $\|\boldsymbol{x}\|_{\theta,2} := \max\{|\lambda_{\theta,1}(\boldsymbol{x})|, |\lambda_{\theta,2}(\boldsymbol{x})|\}$. It is clear that $\|\boldsymbol{x}\|_{\theta,2} \leq \|\boldsymbol{x}\|_{\theta,F}$.

We call $\boldsymbol{x} \in \mathcal{E}^n$ positive semidefinite if $\boldsymbol{x} \in \mathcal{Q}_\theta^n$ (i.e., $\lambda_{1,2}(\boldsymbol{x}) \geq 0$), and positive definite if $\boldsymbol{x} \in \operatorname{Int}(\mathcal{Q}_\theta^n)$ (i.e., $\lambda_{1,2}(\boldsymbol{x}) > 0$). We write $\boldsymbol{x} \succeq_\theta \mathbf{0}$ to mean that $\boldsymbol{x}$ is positive semidefinite, and $\boldsymbol{x} \succ_\theta \mathbf{0}$ to mean that $\boldsymbol{x}$ is positive definite. We also write $\boldsymbol{x} \succeq_\theta \boldsymbol{y}$ or $\boldsymbol{y} \preceq_\theta \boldsymbol{x}$ to mean that $\boldsymbol{x} - \boldsymbol{y} \succeq_\theta \mathbf{0}$, and $\boldsymbol{x} \succ_\theta \boldsymbol{y}$ or $\boldsymbol{y} \prec_\theta \boldsymbol{x}$ to mean that $\boldsymbol{x} - \boldsymbol{y} \succ_\theta \mathbf{0}$. For real symmetric matrices of order n, X and Y, we write $X \succeq 0$ ($X \succ 0$) to mean that X is positive semidefinite (X is positive definite), and $X \succeq Y$ ($X \succ Y$) or $Y \preceq X$ ($Y \prec X$) to mean that $X - Y \succeq 0$ ($X - Y \succ 0$).

The *arrow-shaped matrix* $\operatorname{Arw}_\theta(\boldsymbol{x})$ associated with $\boldsymbol{x} \in \mathcal{E}^n$ with respect to θ is defined as [10]

$$\operatorname{Arw}_\theta(\boldsymbol{x}) := \begin{bmatrix} x_0 & \cot^2\theta\,\overline{\boldsymbol{x}}^\mathsf{T} \\ \overline{\boldsymbol{x}} & x_0 I_{n-1} \end{bmatrix} = \begin{bmatrix} x_0 & \overline{\boldsymbol{x}_{\theta^2}}^\mathsf{T} \\ \overline{\boldsymbol{x}} & x_0 I_{n-1} \end{bmatrix}. \tag{9}$$

Note that $\mathrm{Arw}_\theta(\boldsymbol{x})\boldsymbol{e} = \boldsymbol{x}, \mathrm{Arw}_\theta(\boldsymbol{x})\boldsymbol{x} = \boldsymbol{x}^2$ and $\mathrm{Arw}(\boldsymbol{e}) = I_n$. Note also that $\boldsymbol{x} \succeq_\theta \boldsymbol{0}$ ($\boldsymbol{x} \succ_\theta \boldsymbol{0}$) if and only if the matrix $\mathrm{Arw}_\theta(\boldsymbol{x}) \succeq \boldsymbol{0}$ ($\mathrm{Arw}_\theta(\boldsymbol{x}) \succ \boldsymbol{0}$). This explains why circular programming is a special case of semidefinite programming.

The *Jordan multiplication* between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ with respect to θ is defined as [10]

$$(\boldsymbol{x} \circ \boldsymbol{y})_\theta := \begin{bmatrix} \boldsymbol{x}_\theta^\top \boldsymbol{y}_\theta \\ x_0\overline{\boldsymbol{y}} + y_0\overline{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_{\theta^2}^\top \boldsymbol{y} \\ x_0\overline{\boldsymbol{y}} + y_0\overline{\boldsymbol{x}} \end{bmatrix} = \mathrm{Arw}_\theta(\boldsymbol{x})\boldsymbol{y} = \mathrm{Arw}_\theta(\boldsymbol{x})\mathrm{Arw}_\theta(\boldsymbol{y})\boldsymbol{e}. \tag{10}$$

It is clear that the definitions of the circular arrow-shaped matrix in (9) and the circular Jordan multiplication in (10) generalize the corresponding ones in [3, Section 4] associated with the second-order cone.

Observe that $\boldsymbol{c}_{\theta,1}^2 = (\boldsymbol{c}_{\theta,1} \circ \boldsymbol{c}_{\theta,1})_\theta = \boldsymbol{c}_{\theta,1}, \boldsymbol{c}_{\theta,2}^2 = (\boldsymbol{c}_{\theta,2} \circ \boldsymbol{c}_{\theta,2})_\theta = \boldsymbol{c}_{\theta,2}, (\boldsymbol{c}_{\theta,1} \circ \boldsymbol{c}_{\theta,2})_\theta = \boldsymbol{0}$. Therefore, $\{\boldsymbol{c}_{\theta,1}, \boldsymbol{c}_{\theta,2}\}$ is a Jordan frame. The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of $\mathcal{E}^n$ are *simultaneously decomposed* if they share a Jordan frame, i.e., $\boldsymbol{x} = \lambda_{\theta,1}(\boldsymbol{x})\boldsymbol{c}_{\theta,1}(\boldsymbol{x}) + \lambda_{\theta,2}(\boldsymbol{x})\boldsymbol{c}_{\theta,2}(\boldsymbol{x})$ and $\boldsymbol{y} = \omega_{\theta,1}(\boldsymbol{x})\boldsymbol{c}_{\theta,1}(\boldsymbol{x}) + \omega_{\theta,2}(\boldsymbol{x})\boldsymbol{c}_{\theta,2}(\boldsymbol{x})$ for a Jordan frame $\{\boldsymbol{c}_{\theta,1}, \boldsymbol{c}_{\theta,2}\}$. We say $\boldsymbol{x}$ and $\boldsymbol{y}$ *operator commute* with respect to θ if for all $\boldsymbol{z} \in \mathcal{E}^n$, we have that $(\boldsymbol{x} \circ (\boldsymbol{y} \circ \boldsymbol{z})_\theta)_\theta = (\boldsymbol{y} \circ (\boldsymbol{x} \circ \boldsymbol{z})_\theta)_\theta$. Two vectors in $\mathcal{E}^n$ operator commute if and only if they are simultaneously decomposed (see [5, Theorem 27]).

One can easily see that, for any $\alpha, \beta \in \mathbb{R}$, $(\boldsymbol{x} \circ (\alpha\boldsymbol{y} + \beta\boldsymbol{z}))_\theta = \alpha(\boldsymbol{x} \circ \boldsymbol{y})_\theta + \beta(\boldsymbol{x} \circ \boldsymbol{z})_\theta$ and $((\alpha\boldsymbol{y} + \beta\boldsymbol{z}) \circ \boldsymbol{x})_\theta = \alpha(\boldsymbol{y} \circ \boldsymbol{x})_\theta + \beta(\boldsymbol{z} \circ \boldsymbol{x})_\theta$. Note that $(\boldsymbol{x} \circ \boldsymbol{e})_\theta = \boldsymbol{x}, (\boldsymbol{x} \circ \boldsymbol{x}^{-1})_\theta = \boldsymbol{e}, \boldsymbol{x}^p = (\boldsymbol{x}^{p-1} \circ \boldsymbol{x})_\theta$ for any nonnegative integer $p \geq 1$, and $(\boldsymbol{x}^p \circ \boldsymbol{x}^q)_\theta = \boldsymbol{x}^{p+q}$ for any nonnegative integer $p, q \geq 1$. Therefore, the algebra $(\mathcal{E}^n, \theta, \circ)$ is power associative (it is not associative though). In fact, one can also see that

$$(\boldsymbol{x} \circ \boldsymbol{y})_\theta = (\boldsymbol{y} \circ \boldsymbol{x})_\theta \ \text{(commutativity)} \quad \text{and} \quad (\boldsymbol{x} \circ (\boldsymbol{x}^2 \circ \boldsymbol{y})_\theta)_\theta = (\boldsymbol{x}^2 \circ (\boldsymbol{x} \circ \boldsymbol{y})_\theta)_\theta \ \text{(Jordan's axiom)}.$$

This shows that the algebra $(\mathcal{E}^n, \theta, \circ)$ is a Jordan algebra with the circular Jordan multiplication $(\cdot \circ \cdot)_\theta$ defined in (10). Moreover, we can also show that the Jordan algebra $(\mathcal{E}^n, \theta, \circ)$ is a Euclidean Jordan algebra under the circular inner product $\langle \cdot, \cdot \rangle_\theta$ defined in (5). We have the following theorem [10].

Theorem 1. *The cone of squares of the Euclidean Jordan algebra $(\mathcal{E}^n, \theta, \circ)$ is the circular cone $\mathcal{Q}_\theta^n$.*

The *quadratic operator* $Q_{\theta,\boldsymbol{x},\boldsymbol{z}} : \mathcal{E}^n \times \mathcal{E}^n \longrightarrow \mathcal{E}^n$ associated with the pair $(\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{E}^n \times \mathcal{E}^n$ with respect to θ is given by

$$\begin{array}{rcl} Q_{\theta,\boldsymbol{x},\boldsymbol{z}} & := & \mathrm{Arw}_\theta(\boldsymbol{x})\mathrm{Arw}(\boldsymbol{z}) + \mathrm{Arw}(\boldsymbol{z})\mathrm{Arw}_\theta(\boldsymbol{x}) - \mathrm{Arw}(\boldsymbol{x} \circ \boldsymbol{z}) \\ & = & \begin{bmatrix} \boldsymbol{x}_\theta^\top \boldsymbol{z}_\theta & \cot^2\theta(x_0\bar{\boldsymbol{z}}^\top + z_0\bar{\boldsymbol{x}}^\top) \\ x_0\bar{\boldsymbol{z}} + z_0\bar{\boldsymbol{x}} & x_0 z_0 I_{n-1} + (\bar{\boldsymbol{x}}_\theta\bar{\boldsymbol{z}}_\theta^\top + \bar{\boldsymbol{z}}_\theta\bar{\boldsymbol{x}}_\theta^\top - \bar{\boldsymbol{x}}_\theta^\top\bar{\boldsymbol{z}}_\theta I_{n-1}) \end{bmatrix}. \end{array}$$

The *quadratic representation* $Q_{\theta,\boldsymbol{x}} : \mathcal{E}^n \longrightarrow \mathcal{E}^n$ associated with $\boldsymbol{x} \in \mathcal{E}^n$ with respect to θ is given by

$$Q_{\theta,\boldsymbol{x}} := 2\mathrm{Arw}^2(\boldsymbol{x}) - \mathrm{Arw}(\boldsymbol{x}^2) = \begin{bmatrix} \|\boldsymbol{x}_\theta\|^2 & 2\cot^2\theta\, x_0\bar{\boldsymbol{x}}^\top \\ 2x_0\bar{\boldsymbol{x}} & \det_\theta(\boldsymbol{x}) I_{n-1} + 2\bar{\boldsymbol{x}}_\theta\bar{\boldsymbol{x}}_\theta^\top \end{bmatrix} = 2\boldsymbol{x}\boldsymbol{x}^\top I_{\theta,n} - \det_\theta(\boldsymbol{x})R,$$

where R is the reflection matrix defined in (8). Note that $Q_{\theta,\boldsymbol{x}}\boldsymbol{e} = \boldsymbol{x}^2, Q_{\theta,\boldsymbol{x}}\boldsymbol{x}^{-1} = \boldsymbol{x}$, and $Q_{\theta,\boldsymbol{e}} = I_n$.

We finally present some handy tools needed for our computations.

Lemma 3. *[10, Theorem 7] Let $\boldsymbol{x}, \boldsymbol{u} \in \mathcal{E}^n$, and $\boldsymbol{y} = \boldsymbol{y}(\boldsymbol{x})$ be a function of $\boldsymbol{x}$ in $\mathcal{J}$. Then*

1. *The gradient* $\nabla_{\boldsymbol{x}} \ln\det_\theta \boldsymbol{x} = 2\, I^2_{\theta,n}\boldsymbol{x}^{-1} = 2\,(\boldsymbol{x}^{-1})_{\theta^2}$*, provided that* $\det_\theta(\boldsymbol{x})$ *is positive (so* $\boldsymbol{x}$ *is invertible). More generally,* $\nabla_{\boldsymbol{x}} \ln\det_\theta \boldsymbol{y} = 2(\nabla_{\boldsymbol{x}}\boldsymbol{y})^{\mathsf{T}}_{\theta^2}\boldsymbol{y}^{-1} = 2(\nabla_{\boldsymbol{x}}\boldsymbol{y})^{\mathsf{T}}(\boldsymbol{y}^{-1})_{\theta^2}$*, provided that* $\det_\theta \boldsymbol{y}$ *is positive.*

2. *The Hessian* $\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \ln\det_\theta \boldsymbol{x} = -2\, I^2_{\theta,n} Q_{\theta,\boldsymbol{x}^{-1}}$*. Hence the gradient* $\nabla_{\boldsymbol{x}}\boldsymbol{x}^{-1} = -Q_{\theta,\boldsymbol{x}^{-1}}$*, provided that* $\boldsymbol{x}$ *is invertible. More generally,* $\nabla_{\boldsymbol{x}}\boldsymbol{y}^{-1} = -Q_{\theta,\boldsymbol{y}^{-1}}\nabla_{\boldsymbol{x}}\boldsymbol{y}$ *provided that* $\boldsymbol{y}$ *is invertible.*

3 Problem formulation and self-concordance properties

In CP problems, we minimize a linear function over the intersection of an affine linear manifold with the Cartesian product of circular cones:

$$\mathcal{Q}^n_{(\theta_1,\theta_2,\ldots,\theta_r)} := \mathcal{Q}^{n_1}_{\theta_1} \times \mathcal{Q}^{n_2}_{\theta_2} \times \cdots \times \mathcal{Q}^{n_r}_{\theta_r},$$

where $n = n_1 + n_2 + \cdots + n_r$. In this paper, without loss of generality, we assume that $r = 1$, hence

$$\mathcal{Q}^{n_1,\ldots,n_r}_{\theta_1,\ldots,\theta_r} = \mathcal{Q}^{n_1}_{\theta_1} = \mathcal{Q}^n_\theta.$$

We use the Euclidean Jordan algebraic characterization of circular cones to define the primal-dual pair of the CP problems. Let $\theta \in (0, \pi/2)$ be a given angle, by using the notations introduced in Subsection 1.1, we define the CP problem and its dual as

$$\begin{array}{llcll} \min & \boldsymbol{c}^{\mathsf{T}}_{\theta^2}\boldsymbol{x} & \qquad & \max & \boldsymbol{b}^{\mathsf{T}}\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{a}^{(i)\mathsf{T}}_{\theta^2}\boldsymbol{x} = b_i, i = 1, 2, \ldots, m, & & \text{s.t.} & \displaystyle\sum_{i=1}^m y_i \boldsymbol{a}_i + \boldsymbol{s} = \boldsymbol{c}, \\ & \boldsymbol{x} \succeq_\theta \boldsymbol{0}, & & & \boldsymbol{s} \succeq_\theta \boldsymbol{0}, \boldsymbol{y} \in \mathbb{R}^m, \end{array} \tag{11}$$

where $\boldsymbol{c}, \boldsymbol{a}^{(i)} \in \mathcal{E}^n$ for $i = 1, 2, \ldots, m$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{x}$ is the primal variable, and $\boldsymbol{y}$ and $\boldsymbol{s}$ are the dual variables. The pair (11) can be compactly rewritten as

$$\begin{array}{llcll} \min & \boldsymbol{c}^{\mathsf{T}}_{\theta^2}\boldsymbol{x} & \qquad & \max & \boldsymbol{b}^{\mathsf{T}}\boldsymbol{y} \\ \text{s.t.} & A_{\theta^2}\boldsymbol{x} = \boldsymbol{b}, & & \text{s.t.} & A^{\mathsf{T}}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \\ & \boldsymbol{x} \succeq_\theta \boldsymbol{0}, & & & \boldsymbol{s} \succeq_\theta \boldsymbol{0}, \boldsymbol{y} \in \mathbb{R}^m, \end{array} \tag{12}$$

where $A := (\boldsymbol{a}^{(1)}; \boldsymbol{a}^{(2)}; \ldots; \boldsymbol{a}^{(m)})$ is a matrix that maps $\mathcal{E}^n$ into $\mathbb{R}^m$, and A^{T} is its transpose. We call $\boldsymbol{x} \in \mathcal{E}^n$ primal feasible if $A_{\theta^2}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \succeq_\theta \boldsymbol{0}$. Similarly, $(\boldsymbol{s}, \boldsymbol{y}) \in \mathcal{E}^n \times \mathbb{R}^m$ is called dual feasible if $A^{\mathsf{T}}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}$ and $\boldsymbol{s} \succeq_\theta \boldsymbol{0}$.

Now, we consider the self-concordance properties for CP. First, we define the logarithmic barrier for the circular case and compute the partial derivatives of this barrier. We then state and prove our self-concordance result.

We define the following feasibility sets:

$$\begin{array}{lcl} \mathcal{F}^0_\theta(P) & := & \left\{\boldsymbol{x} \in \mathcal{E}^n : A_{\theta^2}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \succ_\theta \boldsymbol{0}\right\}; \\ \mathcal{F}^0_\theta(D) & := & \left\{(\boldsymbol{s}, \boldsymbol{y}) \in \mathcal{E}^n \times \mathbb{R}^m : A^{\mathsf{T}}\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \boldsymbol{s} \succ_\theta \boldsymbol{0}\right\}; \\ \mathcal{F}_\theta & := & \mathcal{F}^0_\theta(P) \cap \mathcal{F}^0_\theta(D). \end{array}$$

Now we make two assumptions.

Assumption 1. The matrix A has a full row rank.

Assumption 2. $\mathcal{F}_\theta$ is not empty.

Assumption 1 is for convenience. Assumption 2 requires that both primal and dual CP problems (12) contain strictly feasible solutions (i.e., positive definite feasible points), which guarantees strong duality for CPs and therefore implies that the CP problems (12) have unique solutions.

We define the logarithmic barrier [13] on the interior of the feasible set of the dual CP problem (12):

$$\begin{array}{ll} \max & \boldsymbol{b}^\mathsf{T}\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{s}(\boldsymbol{y}) := A^\mathsf{T}\boldsymbol{y} - \boldsymbol{c} \succeq_\theta \mathbf{0}. \end{array}$$

Let

$$\mathcal{H}_\theta := \left\{ \boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{s}(\boldsymbol{y}) \succeq_\theta \mathbf{0} \right\}, \text{ hence } \operatorname{Int}(\mathcal{H}_\theta) = \left\{ \boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{s}(\boldsymbol{y}) \succ_\theta \mathbf{0} \right\}.$$

Under Assumption 2, the set $\operatorname{Int}(\mathcal{H}_\theta)$ is nonempty. The *circular logarithmic barrier* for $\operatorname{Int}(\mathcal{H}_\theta)$ is the function $f_\theta : \operatorname{Int}(\mathcal{H}_\theta) \to \mathbb{R}$ defined as

$$f_\theta(\boldsymbol{y}) := -\ln \det_\theta(\boldsymbol{s}(\boldsymbol{y})), \quad \forall\, \boldsymbol{y} \in \operatorname{Int}(\mathcal{H}_\theta).$$

To state the result on the self-concordance of the logarithmic barrier $f_\theta(\cdot)$ for CP, we need the following definition.

Definition 1. [13, Definition 2.1.1] Let E be a finite-dimensional real vector space, G be an open nonempty convex subset of E, and let g be a $\mathcal{C}^3$, convex mapping from G to $\mathbb{R}$. Then g is called *α-self-concordant* on G with the parameter $\alpha > 0$ if for every $\boldsymbol{y} \in G$ and $\boldsymbol{h} \in E$, the following inequality holds:

$$|\nabla^3_{\boldsymbol{yyy}} g(\boldsymbol{y})[\boldsymbol{h}, \boldsymbol{h}, \boldsymbol{h}]| \;\; \leq \;\; 2\alpha^{-1/2}(\nabla^2_{\boldsymbol{yy}} g(\boldsymbol{y})[\boldsymbol{h}, \boldsymbol{h}])^{3/2}.$$

An α-self-concordant function g on G is called *strongly α-self-concordant* if g tends to infinity for any sequence approaching a boundary point of G.

We note that in the above definition the set G is assumed to be open. However, relative openness would be sufficient to apply the definition. See also [13, Item A, Page 57].

We now present the following important result in the self-concordance of the circular logarithmic barrier.

Theorem 2. *The logarithmic barrier function $f_\theta(\cdot)$ is a strongly 1-self-concordant barrier for $\mathcal{H}_\theta$.*

Theorem 2 implies the existence of polynomial-time interior-point algorithms for circular programming [13]. This is the substance of Sections 4 and 5. We prove Theorem 2 in Section 6. Proving Theorem 2 establishes the polynomial convergence results of the proposed algorithms.

4 Newton's method and commutative directions

In the primal dual pair of CP problems (12), the matrix A is defined to map $\mathcal{E}^n$ into $\mathbb{R}^m$, and its transpose, A^T, is defined to map $\mathbb{R}^m$ into $\mathcal{E}^n$ such that $\boldsymbol{x}^\mathsf{T}(A^\mathsf{T}_{\theta^2}\boldsymbol{y}) = (A_{\theta^2}\boldsymbol{x})^\mathsf{T}\boldsymbol{y}$. Indeed, we can prove weak and strong duality properties for the pair (12) as justification for referring to them as a primal dual pair. The following lemma generalizes [3, Lemma 15].

Lemma 4. (Complementarity conditions) *Suppose that $\boldsymbol{x}, \boldsymbol{s} \succeq_\theta \mathbf{0}$. Then*

$$\boldsymbol{x}_\theta^\mathsf{T}\boldsymbol{s}_\theta = \boldsymbol{x}^\mathsf{T}\boldsymbol{s}_{\theta^2} = 0 \;\; \textit{if and only if} \;\; (\boldsymbol{x} \circ \boldsymbol{s})_\theta = \mathbf{0}.$$

Proof. Since

$$(\boldsymbol{x} \circ \boldsymbol{s})_\theta = \begin{bmatrix} \boldsymbol{x}_\theta^\mathsf{T} \boldsymbol{s}_\theta \\ x_0 \overline{\boldsymbol{s}} + s_0 \overline{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} x_0 s_0 + \cot^2\theta\, \overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}} \\ x_0 \overline{\boldsymbol{s}} + s_0 \overline{\boldsymbol{x}} \end{bmatrix},$$

it is enough to prove that $\boldsymbol{x}_\theta^\mathsf{T} \boldsymbol{s}_\theta = x_0 s_0 + \cot^2\theta\, \overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}} = 0$ implies that $x_0 \overline{\boldsymbol{s}} + s_0 \overline{\boldsymbol{x}} = \boldsymbol{0}$. If $x_0 = 0$ or $s_0 = 0$, then $\boldsymbol{x} = \boldsymbol{0}$ or $\boldsymbol{s} = \boldsymbol{0}$ and hence result is trivial. Therefore, we need only to consider the case where x_0 and s_0 are strictly greater than zero. By Cauchy-Schwartz inequality and the assumption that $\boldsymbol{x} \succeq_\theta \boldsymbol{0}$ and $\boldsymbol{s} \succeq_\theta \boldsymbol{0}$, we have

$$\cot^2\theta\, \overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}} \geq -\cot^2\theta\, |\overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}}| \geq -\cot^2\theta\, \|\overline{\boldsymbol{x}}\| \|\overline{\boldsymbol{s}}\| \geq -x_0 s_0. \tag{13}$$

Consequently, $x_0 s_0 + \cot^2\theta\, \overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}} \geq 0$. Now, $x_0 s_0 + \cot^2\theta\, \overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}} = 0$ if and only if $\cot^2\theta\, \overline{\boldsymbol{x}}^\mathsf{T} \overline{\boldsymbol{s}} = -x_0 s_0$, therefore if and only if the inequalities in (13) are satisfied as equalities. But if this is true then either $\boldsymbol{x} = \boldsymbol{0}$ or $\boldsymbol{s} = \boldsymbol{0}$, in which case $x_0 \overline{\boldsymbol{s}} + s_0 \overline{\boldsymbol{x}} = \boldsymbol{0}$, or $\boldsymbol{x} \neq \boldsymbol{0}$ and $\boldsymbol{s} \neq \boldsymbol{0}$, $\overline{\boldsymbol{x}} = -\alpha \overline{\boldsymbol{s}}$, where $\alpha > 0$, and $x_0 = \|\overline{\boldsymbol{x}}\| = \alpha \|\overline{\boldsymbol{s}}\| = \alpha s_0$, i.e., $\overline{\boldsymbol{x}} + \frac{x_0}{s_0} \overline{\boldsymbol{s}} = \boldsymbol{0}$. The proof is complete. □

As a result, the complementary slackness condition for the CP problems (12) is given by the equation $(\boldsymbol{x} \circ \boldsymbol{s})_\theta = \boldsymbol{0}$. Thus, the corresponding linear system is:

$$\begin{aligned} & A_{\theta^2} \boldsymbol{x} = \boldsymbol{b}, \\ & A^\mathsf{T} \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \\ & (\boldsymbol{x} \circ \boldsymbol{s})_\theta = \boldsymbol{0}, \\ & \boldsymbol{x}, \boldsymbol{s} \succeq_\theta \boldsymbol{0}. \end{aligned}$$

The logarithmic barrier problem associated with the primal problem in (12) is the problem:

$$\begin{array}{ll} \min & \boldsymbol{c}_{\theta^2}^\mathsf{T} \boldsymbol{x} - \mu_\theta \ln \det_\theta(\boldsymbol{x}) \\ \text{s.t.} & A_{\theta^2} \boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \succ_\theta \boldsymbol{0}, \end{array} \tag{14}$$

where the barrier parameter $\mu_\theta := \frac{1}{2} \boldsymbol{x}_\theta^\mathsf{T} \boldsymbol{s}_\theta > 0$ is the normalized duality gap. The Lagrangian dual of (14) is the problem:

$$\begin{array}{ll} \max & \boldsymbol{b}^\mathsf{T} \boldsymbol{y} + \mu_\theta \ln \det_\theta(\boldsymbol{s}) \\ \text{s.t.} & A^\mathsf{T} \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \\ & \boldsymbol{s} \succ_\theta \boldsymbol{0}, \end{array} \tag{15}$$

which is the logarithmic barrier problem associated with the dual problem in (12).

As the CP problems (14) and (15) are, respectively, concave and convex, $\boldsymbol{x}$ and $(\boldsymbol{y}, \boldsymbol{s})$ are optimal solutions to (14) and (15), respectively, if and only if they satisfy the following optimality conditions:

$$\begin{aligned} & A_{\theta^2} \boldsymbol{x} = \boldsymbol{b}, \\ & A^\mathsf{T} \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \\ & (\boldsymbol{x} \circ \boldsymbol{s})_\theta = \sigma \mu_\theta \boldsymbol{e}, \\ & \boldsymbol{x}, \boldsymbol{s} \succ_\theta \boldsymbol{0}. \end{aligned} \tag{16}$$

Now, we apply Newton's method to this system to get the following linear system:

$$\begin{aligned} & A_{\theta^2} \Delta \boldsymbol{x} = \boldsymbol{b} - A_{\theta^2} \boldsymbol{x}, \\ & A^\mathsf{T} \Delta \boldsymbol{y} + \Delta \boldsymbol{s} = \boldsymbol{c} - \boldsymbol{s} - A^\mathsf{T} \boldsymbol{y}, \\ & (\Delta \boldsymbol{x} \circ \boldsymbol{s})_\theta + (\boldsymbol{x} \circ \Delta \boldsymbol{s})_\theta = \sigma \mu_\theta \boldsymbol{e} - (\boldsymbol{x} \circ \boldsymbol{s})_\theta, \end{aligned} \tag{17}$$

where $(\Delta \boldsymbol{x}, \Delta \boldsymbol{s}, \Delta \boldsymbol{y}) \in \mathcal{E}^n \times \mathcal{E}^n \times \mathbb{R}^m$ and $\sigma \in [0, 1]$ is a centering parameter.

We study short-, semi-long-, and long-step algorithms associated with the following centrality measures defined for $(\boldsymbol{x}, \boldsymbol{s}) \in \text{Int}(\mathcal{Q}_\theta^n) \times \text{Int}(\mathcal{Q}_\theta^n)$ with respect to θ:

$$\begin{array}{lcl} d_{\theta,F}(\boldsymbol{x},\boldsymbol{s}) & := & \|Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s} - \mu_\theta \boldsymbol{e}\|_F = \sqrt{\Big(\lambda_{\theta,1}(Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}) - \mu_\theta\Big)^2 + \Big(\lambda_{\theta,2}(Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}) - \mu_\theta\Big)^2}, \\ d_{\theta,2}(\boldsymbol{x},\boldsymbol{s}) & := & \|Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s} - \mu_\theta \boldsymbol{e}\|_2 = \max\Big\{\Big|\lambda_{\theta,1}(Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}) - \mu_\theta\Big|,\ \Big|\lambda_{\theta,2}(Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}) - \mu_\theta\Big|\Big\}, \\ d_{\theta,-\infty}(\boldsymbol{x},\boldsymbol{s}) & := & \mu_\theta - \min\Big\{\lambda_{\theta,1}(Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}),\ \lambda_{\theta,2}(Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s})\Big\}. \end{array}$$

Let $\gamma \in (0,1)$ be a given constant. With respect to θ, we define the following neighborhoods of the central path for CP:

$$\begin{array}{lcl} \mathcal{N}_{\theta,F}(\gamma) & := & \left\{(\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}) \in \mathcal{F}_\theta^0(P) \times \mathcal{F}_\theta^0(D) : d_{\theta,F}(\boldsymbol{x},\boldsymbol{s}) \leq \gamma\mu_\theta\right\}, \\ \mathcal{N}_{\theta,2}(\gamma) & := & \left\{(\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}) \in \mathcal{F}_\theta^0(P) \times \mathcal{F}_\theta^0(D) : d_{\theta,2}(\boldsymbol{x},\boldsymbol{s}) \leq \gamma\mu_\theta\right\}, \\ \mathcal{N}_{\theta,-\infty}(\gamma) & := & \left\{(\boldsymbol{x},\boldsymbol{s},\boldsymbol{y}) \in \mathcal{F}_\theta^0(P) \times \mathcal{F}_\theta^0(D) : d_{\theta,-\infty}(\boldsymbol{x},\boldsymbol{s}) \leq \gamma\mu_\theta\right\}. \end{array} \tag{18}$$

Note that, by item i of [5, Proposition 21], $Q_{\theta,\boldsymbol{s}^{1/2}}\, \boldsymbol{x}$ and $Q_{\theta,\boldsymbol{s}^{1/2}}\, \boldsymbol{x}$ operator commute, and thus the centrality measures $d_{\theta,\cdot}(\boldsymbol{x},\boldsymbol{s})$ and their corresponding neighborhoods $\mathcal{N}_{\theta,\cdot}(\gamma)$ are symmetric with respect to $\boldsymbol{x}$ and $\boldsymbol{s}$. Note also that, by item ii of [5, Proposition 21], $Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}$ and $Q_{\theta,\widetilde{\boldsymbol{x}}^{1/2}}\, \widetilde{\boldsymbol{s}}$ have the same eigenvalues, and that all neighborhoods can be defined in terms of the eigenvalues of $Q_{\theta,\boldsymbol{x}^{1/2}}\, \boldsymbol{s}$, therefore the three neighborhoods defined in (18) are scaling invariant, i.e. $(\boldsymbol{x},\boldsymbol{s})$ is in the neighborhood if and only if $(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{s}})$ is. Furthermore, it is easy to see that

$$d_{\theta,F}(\boldsymbol{x},\boldsymbol{s}) \geq d_{\theta,2}(\boldsymbol{x},\boldsymbol{s}) \geq d_{\theta,-\infty}(\boldsymbol{x},\boldsymbol{s}), \text{ which implies that } \mathcal{N}_{\theta,F}(\gamma) \subseteq \mathcal{N}_{\theta,2}(\gamma) \subseteq \mathcal{N}_{\theta,-\infty}(\gamma) \subseteq \mathcal{Q}_\theta^n \times \mathcal{Q}_\theta^n.$$

One can see that the vectors $\boldsymbol{x}$ and $\boldsymbol{s}$ may not operator commute, hence the statement that "$(\boldsymbol{x} \circ \boldsymbol{s})_\theta = \sigma\mu_\theta \boldsymbol{e}$ implies $\boldsymbol{x} = \sigma\mu_\theta \boldsymbol{s}^{-1}$" may not true. Such a statement holds if $\boldsymbol{x}$ and $\boldsymbol{s}$ operator commute (see [14, Chapter II]). So, we need to scale the optimality conditions (17) so that the scaled vectors operator commute (or equivalently, the scaled vectors are simultaneously decomposed). We use an effective way of scaling proposed originally by Monteiro [11] and Zhang [12] for semidefinite programming, and then generalized by Schmieta and Alizadeh [5] for general symmetric programming.

From now on, with respect to θ and $\boldsymbol{p} \succ_\theta \boldsymbol{0}$, we define

$$\widetilde{\boldsymbol{x}} := Q_{\theta,\boldsymbol{p}}\, \boldsymbol{x},\ \widetilde{\boldsymbol{s}} := Q_{\theta,\boldsymbol{p}^{-1}}\, \boldsymbol{s},\ \widetilde{\boldsymbol{c}} := Q_{\theta,\boldsymbol{p}^{-1}}\, \boldsymbol{c},\ \widetilde{\boldsymbol{c}_{\theta^2}} := Q_{\theta,\boldsymbol{p}^{-1}}\, \boldsymbol{c}_{\theta^2},\ \widetilde{A} := A\, Q_{\theta,\boldsymbol{p}^{-1}}, \text{ and } \widetilde{A_{\theta^2}} := A_{\theta^2}\, Q_{\theta,\boldsymbol{p}^{-1}}.$$

Note that, using item 2 of [5, Lemma 8], we have $Q_{\theta,\boldsymbol{p}} Q_{\theta,\boldsymbol{p}^{-1}} = Q_{\theta,\boldsymbol{p}}(Q_{\theta,\boldsymbol{p}})^{-1} = I_n$ for $\boldsymbol{p} \succ_\theta \boldsymbol{0}$. With this change of variables, Problem (14) becomes

$$\begin{array}{ll} \min & \widetilde{\boldsymbol{c}_{\theta^2}}^{\mathsf{T}} \widetilde{\boldsymbol{x}} - \mu_\theta \ln \det_\theta(\widetilde{\boldsymbol{x}}) \\ \text{s.t.} & \widetilde{A_{\theta^2}}\, \widetilde{\boldsymbol{x}} = \boldsymbol{b}, \\ & \widetilde{\boldsymbol{x}} \succ_\theta \boldsymbol{0}, \end{array} \tag{19}$$

and Problem (15) becomes

$$\begin{array}{ll} \max & \boldsymbol{b}^{\mathsf{T}} \boldsymbol{y} + \mu_\theta \ln \det_\theta(\widetilde{\boldsymbol{s}}) \\ \text{s.t.} & \widetilde{A}^{\mathsf{T}} \boldsymbol{y} + \widetilde{\boldsymbol{s}} = \widetilde{\boldsymbol{c}}, \\ & \widetilde{\boldsymbol{s}} \succ_\theta \boldsymbol{0}, \end{array} \tag{20}$$

Note that Problems (14) and (19) have the same maximizer, but their optimal objective values are equal up to a constant. Similarly, Problems (15) and (20) have the same minimizer but their optimal objective values differ by a constant. We have the following lemma and proposition.

Lemma 5 (Lemma 28, [5]). *Let $\boldsymbol{p} \in \mathcal{E}^n$ be invertible. Then $(\boldsymbol{x} \circ \boldsymbol{s})_\theta = \sigma\mu_\theta \boldsymbol{e}$ if and only if $(\widetilde{\boldsymbol{x}} \circ \widetilde{\boldsymbol{s}})_\theta = \sigma\mu_\theta \boldsymbol{e}$.*

Proposition 1. *The point $(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y})$ satisfies the optimality conditions (16) if and only if the point $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{s}}, \boldsymbol{y})$ satisfies the relaxed optimality conditions*

$$\begin{array}{l} \widetilde{A_{\theta^2}}\widetilde{\boldsymbol{x}} = \boldsymbol{b}, \\ \widetilde{A}^{\mathsf{T}}\boldsymbol{y} + \widetilde{\boldsymbol{s}} = \widetilde{\boldsymbol{c}}, \\ (\widetilde{\boldsymbol{x}} \circ \widetilde{\boldsymbol{s}})_\theta = \sigma\mu_\theta \boldsymbol{e}, \\ \widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{s}} \succ_\theta \boldsymbol{0}. \end{array}$$

Proof. The proof follows from Lemma 5, the fact that $Q_{\theta,\boldsymbol{p}}(\mathcal{Q}^n_\theta) = \mathcal{Q}^n_\theta$, and likewise, as an operator, $Q_{\theta,\boldsymbol{p}}(\mathrm{Int}(\mathcal{Q}^n_\theta)) = \mathrm{Int}(\mathcal{Q}^n_\theta)$, because $\mathcal{Q}^n_\theta$ is a symmetric cone (see also [10, Theorem 6]). □

As a result of Proposition 1, we conclude that the set directions $(\Delta\boldsymbol{x}, \Delta\boldsymbol{s}, \Delta\boldsymbol{y})$ satisfies the optimality conditions (17) if and only if the set directions $(\widetilde{\Delta\boldsymbol{x}}, \widetilde{\Delta\boldsymbol{s}}, \Delta\boldsymbol{y})$ satisfies the following relaxed optimality conditions:

$$\begin{array}{l} \widetilde{A_{\theta^2}}\, \widetilde{\Delta\boldsymbol{x}} = \boldsymbol{b} - \widetilde{A_{\theta^2}}\, \widetilde{\boldsymbol{x}}, \\ \widetilde{A}^{\mathsf{T}}\Delta\boldsymbol{y} + \widetilde{\Delta\boldsymbol{s}} = \widetilde{\boldsymbol{c}} - \widetilde{\boldsymbol{s}} - \widetilde{A}^{\mathsf{T}}\boldsymbol{y}, \\ (\widetilde{\Delta\boldsymbol{x}} \circ \widetilde{\boldsymbol{s}})_\theta + (\widetilde{\boldsymbol{x}} \circ \widetilde{\Delta\boldsymbol{s}})_\theta = \sigma\mu_\theta\boldsymbol{e} - (\widetilde{\boldsymbol{x}} \circ \widetilde{\boldsymbol{s}})_\theta. \end{array} \tag{21}$$

To compute the Newton directions $(\widetilde{\Delta\boldsymbol{x}}, \widetilde{\Delta\boldsymbol{s}}, \Delta\boldsymbol{y})$, we write the system of equations (21) in the following block matrix form:

$$\begin{bmatrix} \widetilde{A_{\theta^2}} & \boldsymbol{0}^{\mathsf{T}} & \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{0} & \widetilde{A}^{\mathsf{T}} & I_n \\ \mathrm{Arw}_\theta(\widetilde{\boldsymbol{s}}) & \boldsymbol{0}^{\mathsf{T}} & \mathrm{Arw}_\theta(\widetilde{\boldsymbol{x}}) \end{bmatrix} \begin{bmatrix} \widetilde{\Delta\boldsymbol{x}} \\ \Delta\boldsymbol{y} \\ \widetilde{\Delta\boldsymbol{s}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_{\theta,p} \\ \boldsymbol{r}_d \\ \boldsymbol{r}_{\theta,c} \end{bmatrix}, \tag{22}$$

where

$$\boldsymbol{r}_{\theta,p} := \boldsymbol{b} - \widetilde{A_{\theta^2}}\, \widetilde{\boldsymbol{x}}, \;\; \boldsymbol{r}_d := \widetilde{\boldsymbol{c}} - \widetilde{\boldsymbol{s}} - \widetilde{A}^{\mathsf{T}}\boldsymbol{y}, \;\; \text{and} \;\; \boldsymbol{r}_{\theta,c} := \sigma\mu_\theta\boldsymbol{e} - (\widetilde{\boldsymbol{x}} \circ \widetilde{\boldsymbol{s}})_\theta.$$

Solving (22) by applying block Gaussian elimination, we obtain

$$\begin{array}{rcl} \Delta\boldsymbol{y} & = & (\widetilde{A_{\theta^2}}\mathrm{Arw}_\theta^{-1}(\widetilde{\boldsymbol{s}})\mathrm{Arw}_\theta(\widetilde{\boldsymbol{x}})\widetilde{A}^{\mathsf{T}})^{-1}(\boldsymbol{r}_{\theta,p} + \widetilde{A_{\theta^2}}\mathrm{Arw}_\theta^{-1}(\widetilde{\boldsymbol{s}})(\mathrm{Arw}_\theta(\widetilde{\boldsymbol{x}})\boldsymbol{r}_d - \boldsymbol{r}_{\theta,c})), \\ \widetilde{\Delta\boldsymbol{s}} & = & \boldsymbol{r}_d - \widetilde{A}^{\mathsf{T}}\Delta\boldsymbol{y}, \\ \widetilde{\Delta\boldsymbol{x}} & = & \mathrm{Arw}_\theta^{-1}(\widetilde{\boldsymbol{s}})\left(\boldsymbol{r}_{\theta,c} - \mathrm{Arw}_\theta(\widetilde{\boldsymbol{x}})\widetilde{\Delta\boldsymbol{s}}\right). \end{array} \tag{23}$$

As we can see, each choice of $\boldsymbol{p}$ leads to a different search direction. As we mentioned earlier, we are interested in the class of $\boldsymbol{p}$ for which the scaled elements are simultaneously decomposed. Hence, it is suffices to choose $\boldsymbol{p}$ so that $\widetilde{\boldsymbol{x}}$ and $\widetilde{\boldsymbol{s}}$ operator commute. That is, we restrict our attention to the following set of scalings:

$$\mathcal{C}_\theta(\boldsymbol{x}, \boldsymbol{s}) := \{\boldsymbol{p} \succ_\theta \boldsymbol{0} : \; \widetilde{\boldsymbol{x}} \text{ and } \widetilde{\boldsymbol{s}} \text{ operator commute}\}.$$

We introduce the following definition [3].

Definition 2. The set of directions $(\Delta\boldsymbol{x}, \Delta\boldsymbol{s}, \Delta\boldsymbol{y})$ arising from those $\boldsymbol{p} \in \mathcal{C}(\boldsymbol{x}, \boldsymbol{s})$ is called the *commutative class of directions*, and a direction in this class is called a *commutative direction*.

The following three choices of $\boldsymbol{p}$ are the most common in practice [5]:

- **Choice I (The HRVW/KSH/M direction):** We may choose $\boldsymbol{p} = \boldsymbol{s}^{1/2}$ and obtain $\widetilde{\boldsymbol{s}} = \boldsymbol{e}$. This choice is analogue of the XS direction in semidefinite programming, and is known as the HRVW/KSH/M direction (it was introduced by Helmberg, Rendl, Vanderbei, and Wolkowicz [15], and Kojima, Shindoh, and Hara [16] independently, and then rediscovered by Monteiro [11]).

- **Choice II (The dual HRVW/KSH/M direction):** We may choose $\boldsymbol{p} = \boldsymbol{x}^{-1/2}$ and obtain $\widetilde{\boldsymbol{x}} = \boldsymbol{e}$. This choice of directions arises by switching the roles of X and S; it is analogue of the SX direction in semidefinite programming, and is known as the dual HRVW/KSH/M direction.

- **Choice III (The NT direction):** We choose $\boldsymbol{p}$ in such a way that $\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{s}}$. In this case we choose
$$\boldsymbol{p} = \left(Q_{\theta,\boldsymbol{x}^{1/2}}\left(Q_{\theta,\boldsymbol{x}^{1/2}}\ \boldsymbol{s}\right)^{-1/2}\right)^{-1/2} = \left(Q_{\theta,\boldsymbol{s}^{-1/2}}\left(Q_{\theta,\boldsymbol{s}^{1/2}}\ \boldsymbol{x}\right)^{1/2}\right)^{-1/2}.$$
This choice of directions was introduced by Nesterov and Todd [17, 18] and is known as the NT direction.

After we compute the Newton directions $(\widetilde{\Delta \boldsymbol{x}}, \widetilde{\Delta \boldsymbol{s}}, \Delta \boldsymbol{y})$ using (23), we can compute the Newton directions $(\Delta \boldsymbol{x}, \Delta \boldsymbol{s}, \Delta \boldsymbol{y})$ by applying the inverse scaling to $(\widetilde{\Delta \boldsymbol{x}}, \widetilde{\Delta \boldsymbol{s}}, \Delta \boldsymbol{y})$ (see Algorithm 1). The corresponding $(\Delta \boldsymbol{x}, \Delta \boldsymbol{s}, \Delta \boldsymbol{y})$ is different from the one obtained by solving system (17) directly. In fact, the former depends on $\boldsymbol{p}$, while the latter is yielded as a special case when $\boldsymbol{p} = \boldsymbol{e}$. It is clear that $\boldsymbol{p} = \boldsymbol{e}$ may not be in $\mathcal{C}(\boldsymbol{x}, \boldsymbol{s})$.

5 The path-following algorithms for solving CPs

In this section we introduce short-, semi-long-, and long-step path-following algorithms for solving the CP problems (12). This class is stated formally in Algorithm 1.

Algorithm 1: The path-following algorithm for solving CP (12)

Require: $\epsilon \in (0,1)$, $\sigma \in (0,1)$, $\gamma \in (0,1)$, $\left(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{s}^{(0)}\right) \in \mathcal{N}_{\theta,\cdot}(\gamma)$

set $\mu_\theta^{(0)} = \frac{1}{2}\, \boldsymbol{x}_\theta^{(0)\mathsf{T}} \boldsymbol{s}_\theta^{(0)}$ and $k = 0$

while $\mu_\theta^{(k)} \leq \epsilon \mu_\theta^{(0)}$ **do**

choose a scaling vector $\boldsymbol{p}_k \in \mathcal{C}_\theta\left(\boldsymbol{x}^{(k)}, \boldsymbol{s}^{(k)}\right)$

let $\left(\widetilde{\boldsymbol{x}}^{(k)}, \widetilde{\boldsymbol{s}}^{(k)}, \boldsymbol{y}^{(k)}\right) = \left(Q_{\theta,\boldsymbol{p}_k} \boldsymbol{x}^{(k)}, Q_{\theta,\boldsymbol{p}_k^{-1}} \boldsymbol{s}^{(k)}, \boldsymbol{y}^{(k)}\right)$

compute $\left(\widetilde{\Delta \boldsymbol{x}}^{(k)}, \widetilde{\Delta \boldsymbol{s}}^{(k)}, \Delta \boldsymbol{y}^{(k)}\right)$ using (23)

let $\left(\Delta \boldsymbol{x}^{(k)}, \Delta \boldsymbol{s}^{(k)}, \Delta \boldsymbol{y}^{(k)}\right) = \left(Q_{\theta,\boldsymbol{p}_k^{-1}} \widetilde{\Delta \boldsymbol{x}}^{(k)}, Q_{\theta,\boldsymbol{p}_k} \widetilde{\Delta \boldsymbol{s}}^{(k)}, \Delta \boldsymbol{y}^{(kj)}\right)$

choose the largest step length $\mu_\theta^{(k)}$ such that

$\left(\boldsymbol{x}^{(k+1)}, \boldsymbol{y}^{(k+1)}, \boldsymbol{s}^{(k+1)}\right) = \left(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{s}^{(k)}\right) + \mu_\theta^{(k)}\left(\Delta \boldsymbol{x}^{(k)}, \Delta \boldsymbol{s}^{(k)}, \Delta \boldsymbol{y}^{(k)}\right) \in \mathcal{N}_{\theta,\cdot}(\gamma)$

set $\mu_\theta^{(k+1)} = \frac{1}{2}\, \boldsymbol{x}_\theta^{(k+1)\mathsf{T}} \boldsymbol{s}_\theta^{(k+1)}$ and $k = k + 1$

end while

In Algorithm 1, ϵ is the desired accuracy of the solution. The general variant of the algorithm is determined by the choice σ and the neighborhood as follows:

- The short-step algorithm is obtained by choosing $\sigma = 1 - \delta/\sqrt{r}$, with $\delta \in (0, 1)$, and $\mathcal{N}_{\theta,F}(\gamma)$ as the neighborhood.
- The semi-long-step algorithm is obtained by choosing $\sigma \in (0, 1)$ and $\mathcal{N}_{\theta,2}(\gamma)$ as the neighborhood.
- The long-step algorithm is obtained by choosing $\sigma \in (0, 1)$ and $\mathcal{N}_{\theta,-\infty}(\gamma)$ as the neighborhood.

The following theorem gives polynomial convergence results for Algorithm 1.

Theorem 3. *Consider Algorithm 1. If the NT direction is used at every iteration, then the short-step algorithm will terminate in $\mathcal{O}(\sqrt{2} \log \epsilon^{-1})$ iterations, and the semi-long and long-step algorithm will terminate in $\mathcal{O}(2 \log \epsilon^{-1})$ iterations. If the HRVW/KSH/M direction or the dual HRVW/KSH/M direction is used at every iteration, then the short-step algorithm will terminate in $\mathcal{O}(\sqrt{2} \log \epsilon^{-1})$ iterations, the semi-long-step algorithm will terminate in $\mathcal{O}(2 \log \epsilon^{-1})$ iterations, and long-step algorithm will terminate in $\mathcal{O}(2\sqrt{2} \log \epsilon^{-1})$ iterations.*

Theorem 3 is a consequence of Theorem 2 (which will be proved in the remaining part of this paper) and [5, Theorem 37] where the underlying symmetric cone is the circular cone.

6 Proving self-concordance for the circular logarithmic barrier

In this section, we prove Theorem 2. To do so, we first obtain a representation for the gradient, the Hessian and the third directional derivative of the logarithmic barrier in the circular case.

Recall that the circular logarithmic barrier is defined as $f_\theta(\boldsymbol{y}) = -\ln \det_\theta(\boldsymbol{s}(\boldsymbol{y}))$ for all $\boldsymbol{y}$ in $\text{Int}(\mathcal{H}_\theta)$, where $s(\boldsymbol{y}) = A^\mathsf{T}\boldsymbol{y} - \boldsymbol{c}$ and $\mathcal{H}_\theta = \{\boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{s}(\boldsymbol{y}) \succeq_\theta \boldsymbol{0}\}$. Throughout the following proof, we let $\boldsymbol{y} \in \mathbb{R}^m$ be such that $\boldsymbol{s}(\boldsymbol{y}) \succ_\theta \boldsymbol{0}$.

Proof of Theorem 2. Note that

$$\nabla_{\boldsymbol{y}} \boldsymbol{s}(\boldsymbol{y}) = \nabla_{\boldsymbol{y}}(A^\mathsf{T}\boldsymbol{y} - \boldsymbol{c}) = A^\mathsf{T}.$$

Then, by using item 1 of Lemma 3 and applying chain rule, we have

$$\nabla_{\boldsymbol{y}} f_\theta(\boldsymbol{y}) = -\nabla_{\boldsymbol{y}} \ln \det_\theta(\boldsymbol{s}(\boldsymbol{y})) = -(\nabla_{\boldsymbol{y}} \boldsymbol{s})^\mathsf{T}_{\theta^2} \boldsymbol{s}^{-1} = -A_{\theta^2} \boldsymbol{s}^{-1},$$

and consequently, using item 2 of Lemma 3 and applying chain rule, we also have

$$\nabla^2_{\boldsymbol{yy}} f_\theta(\boldsymbol{y}) = -A_{\theta^2} \nabla_{\boldsymbol{y}} \boldsymbol{s}^{-1} = A_{\theta^2} Q_{\theta,\boldsymbol{s}^{-1}} \nabla_{\boldsymbol{y}} \boldsymbol{s} = A_{\theta^2} Q_{\theta,\boldsymbol{s}^{-1}} A^\mathsf{T}.$$

Let

$$H = H(\boldsymbol{y}) := \nabla^2_{\boldsymbol{yy}} f_\theta(\boldsymbol{y}) = A_{\theta^2} Q_{\theta,\boldsymbol{s}^{-1}} A^\mathsf{T}. \tag{24}$$

Note that the Hessian matrix H is positive definite under Assumption 1 and the assumption that $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{y}) \succ_\theta \boldsymbol{0}$.

Let $\boldsymbol{h} \in \mathbb{R}^m, \boldsymbol{h} \neq \boldsymbol{0}, \bar{\boldsymbol{a}}_i := Q_{\theta,\boldsymbol{s}^{-1/2}} \boldsymbol{a}_i, \bar{A} := A Q_{\theta,\boldsymbol{s}^{-1/2}}$, and let $\boldsymbol{p} := \bar{A}\boldsymbol{h}$ be a vector in $\mathcal{E}^n$ with the eigenvalues $\lambda_{\theta,1}(\boldsymbol{p})$ and $\lambda_{\theta,2}(\boldsymbol{p})$. Using (24), we have

$$\boldsymbol{h}^\mathsf{T} H \boldsymbol{h} = (\bar{A}\boldsymbol{h})^\mathsf{T}_{\theta^2} \bar{A}\boldsymbol{h} = \|\bar{A}\boldsymbol{h}\|^2_{\theta,F} = \|\boldsymbol{p}\|^2_{\theta,F} = \lambda^2_{\theta,1}(\boldsymbol{p}) + \lambda^2_{\theta,2}(\boldsymbol{p}). \tag{25}$$

Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be two vectors in $\mathcal{E}^n$ such that $\boldsymbol{u}$ is invertible. We have

$$\begin{array}{rcl} \nabla_{\boldsymbol{u}} Q_{\theta,\boldsymbol{u}^{-1}}[\boldsymbol{v}] & = & \nabla_{\boldsymbol{u}} \left(2\mathrm{Arw}(\boldsymbol{u}^{-1})^2 - \mathrm{Arw}(\boldsymbol{u}^{-2})\right)[\boldsymbol{v}] \\ & = & -2\mathrm{Arw}\left(Q_{\theta,\boldsymbol{u}^{-1}}\boldsymbol{v}\right)\mathrm{Arw}(\boldsymbol{u}^{-1}) - 2\mathrm{Arw}(\boldsymbol{u}^{-1})\mathrm{Arw}\left(Q_{\theta,\boldsymbol{u}^{-1}}\boldsymbol{v}\right) + 2\mathrm{Arw}\left(Q_{\theta,\boldsymbol{u}^{-1}}\boldsymbol{v} \circ \boldsymbol{u}^{-1}\right) \\ & = & -2Q_{\theta,(Q_{\theta,\boldsymbol{u}^{-1}}\boldsymbol{v}),\boldsymbol{u}^{-1}}. \end{array}$$

Note that $\boldsymbol{s}^{-1/2}$ is the unique positive definite vector having $(\boldsymbol{s}^{-1/2})^2 = \boldsymbol{s}^{-1}$. Plugging in $\boldsymbol{s}(\boldsymbol{y})$ for $\boldsymbol{u}$ in the positive equation yields

$$\nabla_{\boldsymbol{s}} Q_{\theta,\boldsymbol{s}(\boldsymbol{y})^{-1}}[\boldsymbol{u}] = -2Q_{\theta,(Q_{\theta,\boldsymbol{s}^{-1}}A\boldsymbol{u}),\boldsymbol{s}^{-1}} = -2Q_{\theta,\boldsymbol{s}^{-1/2}}\mathrm{Arw}(\bar{A}\boldsymbol{u})Q_{\theta,\boldsymbol{s}^{-1/2}}. \tag{26}$$

Using (24) and (26), we have

$$\nabla_{\boldsymbol{y}} H[\boldsymbol{u}] = -2\bar{A}_{\theta^2}\mathrm{Arw}(\bar{A}\boldsymbol{u})\bar{A}^{\mathsf{T}}.$$

Then

$$\nabla_{\boldsymbol{y}}(\boldsymbol{h}^{\mathsf{T}} H \boldsymbol{h})[\boldsymbol{u}] = -2\boldsymbol{p}_{\theta^2}^{\mathsf{T}}(\mathrm{Arw}(\bar{A}\boldsymbol{u})\boldsymbol{p}).$$

Therefore,

$$\nabla^3_{\boldsymbol{yyy}} f_\theta[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}] = \nabla_{\boldsymbol{y}}(\boldsymbol{h}^{\mathsf{T}} H \boldsymbol{h})[\boldsymbol{h}] = -2\,\boldsymbol{p}_{\theta^2}^{\mathsf{T}}(\mathrm{Arw}(\boldsymbol{p})\boldsymbol{p}). \tag{27}$$

It is immediate from the fact that $\|\mathrm{Arw}(\boldsymbol{p})\|_{\theta,2} = \|\boldsymbol{p}\|_{\theta,2}$, the fact that $\|\boldsymbol{p}\|_{\theta,2} \leq \|\boldsymbol{p}\|_{\theta,F}$, and (27) that

$$\begin{array}{rcl} |\nabla^3_{\boldsymbol{yyy}} f_\theta[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}]| & = & 2\,|\boldsymbol{p}_{\theta^2}^{\mathsf{T}}(\mathrm{Arw}(\boldsymbol{p})\boldsymbol{p})| \\ & \leq & 2\,\|\mathrm{Arw}(\boldsymbol{p})\|_{\theta,2}\;\boldsymbol{p}_{\theta^2}^{\mathsf{T}}\boldsymbol{p} \\ & = & 2\,\|\boldsymbol{p}\|_{\theta,2}\;\boldsymbol{p}_{\theta^2}^{\mathsf{T}}\boldsymbol{p} \\ & = & 2\,\|\boldsymbol{p}\|_{\theta,2}\left(\lambda^2_{\theta,1}(\boldsymbol{p}) + \lambda^2_{\theta,2}(\boldsymbol{p})\right) \\ & = & 2\,\|\boldsymbol{p}\|_{\theta,2}\,\boldsymbol{h}^{\mathsf{T}} H(\boldsymbol{x})\boldsymbol{h} \\ & = & 2\,\|\boldsymbol{p}\|_{\theta,2}\,|\nabla^2_{\boldsymbol{yy}} f_\theta[\boldsymbol{h},\boldsymbol{h}]| \\ & \leq & 2\,\|\boldsymbol{p}\|_{\theta,F}\,|\nabla^2_{\boldsymbol{yy}} f_\theta[\boldsymbol{h},\boldsymbol{h}]| \\ & = & 2\,(|\nabla^2_{\boldsymbol{yy}} f_\theta[\boldsymbol{h},\boldsymbol{h}])^{3/2}, \end{array} \tag{28}$$

where we used (25) to obtain the third and last equalities and (24) to obtain the fourth equality.

It follows from (28), and

$$\nabla_{\boldsymbol{x}} f_\theta(\boldsymbol{x})^{\mathsf{T}}(\nabla^2_{\boldsymbol{xx}} f_\theta(\boldsymbol{x}))^{-1}\nabla_{\boldsymbol{x}} f_\theta(\boldsymbol{x}) = \boldsymbol{e}_{\theta^2}^{\mathsf{T}}\;\bar{A}^{\mathsf{T}}(\bar{A}_{\theta^2}\bar{A}^{\mathsf{T}})^{-1}\bar{A}\boldsymbol{e} \leq \boldsymbol{e}^{\mathsf{T}}\boldsymbol{e} = 1,$$

that f_θ is a 1-self-concordant barrier for $\mathcal{H}_\theta$. In addition, it is clear that f tends to infinity for any sequence approaching a boundary point of $\mathcal{H}_\theta$. Thus, the logarithmic barrier f_θ is a strongly 1-self-concordant barrier for $\mathcal{H}_\theta$. The result is established. □

7 Concluding remarks

Contrary to the beliefs in the optimization community, it has seen that circular programming is a symmetric programming with a special structure. There is a particular Euclidean Jordan algebra that underlies the analysis of interior point algorithms for the circular programming. In this paper, we have used the machinery of this particular Euclidean Jordan algebra to derive polynomial-time path-following algorithms for circular programming problems. We have also proved the polynomial convergence results of the proposed algorithms by showing that the circular logarithmic barrier is a strongly self-concordant barrier.

References

[1] J.C. Zhou, J.S. Chen, Properties of circular cone and spectral factorization associated with circular cone, Nonlinear convex anal. 214 (2013) 807–816.

[2] J. Zhou, J. Chen, H. Hung, Circular cone convexity and some inequalities associated with circular cones, J. Ineq. Appl. 571 (2013) 1–17.

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

[4] M.J. Todd, Semidefinite optimization, Acta Numer. 10 (2001) 515–560.

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

[6] C.H. Ko, J.S. Chen, Optimal grasping manipulation for multifingered robots using semismooth Newton method, Math. Probl. Eng. (2013) 1–9.

[7] S. Boyd, B. Wegbreit, Fast computation of optimal contact forces, IEEE Trans. Robot. 23 (2007) 1117–1132.

[8] B. Leon, A. Morales, J. Sancho-Bru, Robot Grasping Foundations, In From Robot to Human Grasping Simulation, Springer International Publishing, 2014.

[9] J.F. Bonnans , H.R. Cabrera, Perturbation analysis of second-order cone programming problems, Math. Program. Ser. B 104 (2005) 205–227.

[10] B. Alzalg, The circular cone: A new symmetric cone. Submitted for publication.

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

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

[13] Yu.E. Nesterov, A.S. Nemirovskii, Interior Point Polynomial Algorithms in Convex Programming, SIAM Publications, Philadelphia, PA, 1994.

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

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

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

[17] Yu.E. Nesterov, M.J. Todd, Self-scaled barriers and interior-point methods for convex programming, Math. Oper. Res. 22 (1997) 1–42.

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