THE PRIMAL-DUAL SECOND-ORDER CONE APPROXIMATIONS ALGORITHM FOR SYMMETRIC CONE PROGRAMMING

CHEK BENG CHUA

ABSTRACT. Given any open convex cone K, a logarithmically homogeneous, self-concordant barrier for K and any positive real number $r < 1$, we associate, with each direction $x \in K$, a second-order cone $\hat{K}_r(x)$ containing K. We show that K is the interior of the intersection of the second-order cones $\hat{K}_r(x)$, as x ranges over all directions in K. Using these second-order cones as approximations to cones of symmetric, positive definite matrices, we develop a new polynomial-time primal-dual interior-point algorithm for semidefinite programming. The algorithm is extended to symmetric cone programming via the relation between symmetric cones and Euclidean Jordan algebras.

1. INTRODUCTION

This paper considers the idea of approximating convex conic programming (CCP) problems with simple second-order cone programming (SOCP) problems.

CCP is a generalization of linear programming (LP) in which the nonnegativity constraints are generalized to the conic constraint $x \in \operatorname{cl} K$, where K is a finite-dimensional, open, convex cone and $\operatorname{cl} K$ is its closure. Since every convex programming problem can be homogenized to a CCP problem, the class of CCP problems represents a very wide class of optimization problems.

In SOCP, the cone K is a direct sum of second-order cones. For a simple SOCP problem, K is just a single second-order cone.

The analytical foundation for the study of interior-point methods for CCP was established by Nesterov and Nemirovski [11] in 1994. Instrumental to their work is a special class of barriers which they termed *logarithmically homogeneous, self-concordant barriers*. This class of barriers captures the essential properties of the standard logarithmic barrier $(x_1, \ldots, x_n) \mapsto -\sum_{i=1}^{n} \ln x_i$ responsible for several polynomial-time algorithms for LP.

The Hessian of a logarithmically homogeneous, self-concordant barrier at each $x \in K$ defines an inner product called the *local inner product at x* (see [15]). Under the metric induced by the local inner product at x, the open ball of radius r centered at x is contained entirely in K whenever $r < 1$. The smallest open cone $K_r(x)$ containing this open ball is a second-order cone contained entirely in K. We consider using these second-order cones as local approximates to K. This is an extension of Megiddo's ellipsoidal cone approximation [10] for linear programming. See also Padberg [14].

In the context of semidefinite programming (SDP), where K is a cone of symmetric positive definite matrices, we develop a new primal-dual algorithm that alternates between the primal

2000 *Mathematics Subject Classification.* 90C22 90C25 90C51.

Key words and phrases. second-order cone, semidefinite programming, symmetric cone programming, primal-dual method.

This research was performed as part of the author's Ph.D. study at Cornell University.

problem and its dual (which is again an SDP problem), solving an approximating simple SOCP problem at each step. We call this the *primal-dual second-order cone approximations algorithm*. This algorithm has a provable polynomial complexity bound that matches the best bound known for SDP; See [6, 17] for excellent reviews on SDP.

Semidefinite programming is a subclass of symmetric cone programming, which is the class of CCP problems over symmetric cones (also known as self-dual homogeneous cones or self-scaled cones). Nesterov and Todd [12, 13] were the first to develop efficient primal-dual interior-point algorithms for symmetric cone programming using self-scaled barriers. Symmetric cones are closely related to Euclidean Jordan algebras. Using the relation between symmetric cones and Euclidean Jordan algebras, many interior-point algorithms for semidefinite programming were extended to symmetric cone programming (see [2, 8]). We show that the primal-dual second-order cone approximations algorithm can also be extended to symmetric cone programming.

This paper is organized as follows. We begin Section 2 with a brief introduction to the theory of logarithmically homogeneous, self-concordant barriers. This sets up the discussion on second-order cone approximations in the second part of the section. In Section 3, we look at semidefinite programming and its standard logarithmic barrier. We highlight and prove several properties of the standard logarithmic barrier that are essential to the development and analysis of the primal-dual second-order cone approximations algorithm for semidefinite programming. This is followed by the description and analysis of the algorithm, which shows that the algorithm is polynomial-time. We compare our algorithm with the primal-dual cone affine scaling algorithm in Section 5. Section 4 is devoted to the extension of the algorithm to symmetric cone programming. We briefly discuss the relation between symmetric cones and Euclidean Jordan algebras, and show that this relation provides a simple extension of the primal-dual second-order cone approximations algorithm. Specifically, through Euclidean Jordan algebras, we define standard logarithmic barriers for symmetric cones and show that they possess the same essential properties highlighted in Section 3.

2. Second-Order Cone Approximations

Among all open convex cones, second-order cones form the class of cones that are easiest to deal with. In fact, every simple SOCP problem can be solved analytically (see Section 3.3). In this section, we discuss the concept of approximating an arbitrary open convex cone with a certain class of second-order cones.

Throughout this section, $\mathbb{E}$ denotes the Euclidean n-space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$, and K denotes a regular, open, convex cone in $\mathbb{E}$.

2.1. Logarithmically Homogeneous, Self-Concordant Barriers.

A *convex conic programming problem* is the optimization (i.e., minimization or maximization) of a linear functional over the intersection of a closed convex cone and an affine subspace. Without loss of generality, we may assume that the cone is regular, i.e., it has nonempty interior and does not contain any vector subspace. Mathematically, it can be formulated as

$$\text{(CPP)} \qquad \inf\{\langle \hat{s}, x\rangle : x \in L + \hat{x},\ x \in \operatorname{cl} K\},$$

where $\hat{s}, \hat{x} \in \mathbb{E}$, $L \subset \mathbb{E}$ is a linear subspace and $\operatorname{cl} K$ is the closure of K. The Lagrangian dual problem can be written as

$$\text{(CPD)} \qquad \inf\{\langle \hat{x}, s\rangle : s \in L^{\perp} + \hat{s},\ s \in \operatorname{cl} K^{*}\},$$

where $K^* := \{z \in \mathbb{E} : \langle z, y \rangle > 0,\ \forall 0 \neq y \in \operatorname{cl} K\}$ is the *dual cone* of K and $L^\perp$ is the orthogonal complement of L.

A functional $f : K \to \mathbb{R}$ is a *(nondegenerate and strongly) ϑ-self-concordant barrier* for K if it is a strictly convex, three-times continuously differentiable functional on K that diverges to infinity as its argument approaches a point on the boundary ∂K of K, and satisfies

$$D^3 f(x)[h,h,h] \leq 2(D^2 f(x)[h,h])^{3/2} \tag{1}$$

and

$$Df(x)[h] \leq (\vartheta D^2 f(x)[h,h])^{1/2} \tag{2}$$

for all $x \in K$ and $h \in \mathbb{E}$. Here, $D^k f(x)[h, \ldots, h]$ denotes the k-th directional derivative of f along h, i.e., $\frac{d^k}{dt^k} f(x + th)|_{t=0}$. The parameter ϑ is called the *barrier parameter.*

If f further satisfies

$$f(tx) = f(x) - \vartheta \ln t \tag{3}$$

for all $x \in K$ and $t > 0$, then f is called a *ϑ-logarithmically homogeneous, self-concordant barrier* for K. It can be shown that if f is convex and two-times continuously differentiable, then (3) implies (2). Hence, we can drop (2) in the definition of a logarithmically homogeneous, self-concordant barrier. Let g and H denote the gradient and Hessian of f respectively. It follows straightforwardly from the logarithmic homogeneity of f that

$$g(tx) = \frac{1}{t} g(x), \quad H(tx) = \frac{1}{t^2} H(x),$$
$$H(x)x = -g(x), \quad \text{and} \quad \langle -g(x), x \rangle = \vartheta,$$

for any $x \in K$ and any $t > 0$.

Logarithmically homogeneous, self-concordant barriers were introduced by Nesterov and Nemirovski [11]. All results in this subsection were discussed and proven in [11]. Hence, we do not give proofs nor justifications for these results.

If f is a ϑ-logarithmically homogeneous, self-concordant barrier for K, then it can be shown that $\vartheta \geq 1$, and the *duality map* $x \mapsto -g(x)$ takes K onto its dual cone K^*. Moreover, the functional $f^* : K^* \to \mathbb{R} : s \mapsto -\inf_{x \in K}\{\langle s, x \rangle + f(x)\}$, which is called the *conjugate barrier* of f, is a ϑ-logarithmically homogeneous, self-concordant barrier for K^*. Let g^* and H^* denote the gradient and Hessian of f^* respectively. Once again, it follows straightforwardly from the logarithmic homogeneity of f and its conjugate barrier that

$$f(x) + f^*(-g(x)) = f(-g^*(s)) + f^*(s) = -\vartheta, \tag{4}$$

$$-g^*(-g(x)) = x, \quad -g(-g^*(s)) = s, \quad \text{and} \tag{5}$$

$$H^*(-g(x)) = H(x)^{-1}, \quad H(-g^*(s)) = H^*(s)^{-1}, \tag{6}$$

for any $x \in K$ and any $s \in K^*$.

2.2. **Second-Order Cones and Their Dual Cones.** The *second-order cone in* $\mathbb{R}^n$ is the open, convex cone

$$\left\{x \in \mathbb{R}^n : x_1 > \sqrt{x_2^2 + \cdots + x_n^2}\right\}.$$

Note that the direction $d = [1, 0, \ldots, 0]^T$ can be considered as the center direction of the second-order cone. Indeed, under the usual dot product, the angle between d and any direction along the boundary of the second-order cone in $\mathbb{R}^n$ is half a right angle.

In the Euclidean space $\mathbb{E}$, a *second-order cone* is an open, convex cone of the form

$$\{x \in \mathbb{E} : \langle d, x\rangle' > \|x - \mathrm{Pr}'_d\, x\|'\|d\|'\} = \left\{x \in \mathbb{E} : \langle d, x\rangle' > \frac{1}{\sqrt{2}}\|d\|'\|x\|'\right\}, \tag{7}$$

where $\langle\cdot,\cdot\rangle'$ is an inner product on $\mathbb{E}$, $\|\cdot\|'$ is the norm induced by $\langle\cdot,\cdot\rangle'$, $d \in \mathbb{E}$ is the *center direction*, and Pr'_d is the orthogonal projection under $\langle\cdot,\cdot\rangle'$ onto the subspace spanned by d. Once again, the angle (defined by the inner product $\langle\cdot,\cdot\rangle'$) between d and directions along the boundary of the cone is half a right angle.

For each $r \in (0, 1)$, the cone $\{x \in \mathbb{E} : \langle d, x\rangle' > r\|d\|'\|x\|'\}$ is the second-order cone defined by the inner product $\langle\cdot,\cdot\rangle'' : (x, z) \mapsto \langle M_r x, M_r z\rangle'$ and center direction d, where M_r is the invertible linear operator $x \mapsto x + ((\sqrt{1 - r^2}/r) - 1)\,\mathrm{Pr}'_d\, x$. This can be derived from $M_r d = (\sqrt{1 - r^2}/r)d$, and

$$\begin{aligned}
(\|M_r x\|')^2 &= (\|x - \mathrm{Pr}'_d\, x + \frac{\sqrt{1-r^2}}{r}\,\mathrm{Pr}'_d\, x\|')^2 \\
&= (\|x - \mathrm{Pr}'_d\, x\|')^2 + \frac{1-r^2}{r^2}(\|\mathrm{Pr}'_d\, x\|')^2 \\
&= (\|x\|')^2 - (\|\mathrm{Pr}'_d\, x\|')^2 + \frac{1-r^2}{r^2}(\|\mathrm{Pr}'_d\, x\|')^2 \\
&= (\|x\|')^2 + \frac{1-2r^2}{r^2}\left(\frac{\langle d, x\rangle'}{\|d\|'}\right)^2,
\end{aligned}$$

whence for $x \neq 0$,

$$\begin{aligned}
\langle M_r d, M_r x\rangle' &= \frac{\sqrt{1-r^2}}{r}\langle d, x - \mathrm{Pr}'_d\, x + \frac{\sqrt{1-r^2}}{r}\,\mathrm{Pr}'_d\, x\rangle' \\
&= \frac{1-r^2}{r^2}\langle d, \mathrm{Pr}'_d\, x\rangle' = \frac{1-r^2}{r^2}\langle d, x\rangle'
\end{aligned}$$

and

$$\begin{aligned}
\|M_r d\|'\|M_r x\|' &= \frac{\sqrt{1-r^2}}{r}\|d\|'\sqrt{(\|x\|')^2 + \frac{1-2r^2}{r^2}\left(\frac{\langle d, x\rangle'}{\|d\|'}\right)^2} \\
&= \frac{\sqrt{1-r^2}}{r^2}\|d\|'\|x\|'\sqrt{r^2 + (1-2r^2)\left(\frac{\langle d, x\rangle'}{\|d\|'\|x\|'}\right)^2}
\end{aligned}$$

gives

$$\frac{\langle M_r d, M_r x\rangle'}{\|M_r d\|'\|M_r x\|'} = \frac{\sqrt{1-r^2}\langle d, x\rangle'/\|d\|'\|x\|'}{\sqrt{r^2 + (1-2r^2)\left(\langle d, x\rangle'/\|d\|'\|x\|'\right)^2}},$$

so that

$$\frac{\langle M_r d, M_r x\rangle'}{\|M_r d\|'\|M_r x\|'} > \frac{1}{\sqrt{2}} \iff \frac{\langle d, x\rangle'}{\|d\|'\|x\|'} > r.$$

Thus, our definition of second-order cones includes those whose angle between the center direction and directions along the boundary of the cone is an acute angle different from half a right angle.

It is a well-known fact that when $\langle \cdot, \cdot \rangle' \equiv \langle \cdot, \cdot \rangle$, the second-order cone (7) coincides with its dual cone under $\langle \cdot, \cdot \rangle$. This is stated formally in the following proposition.

Proposition 1. *For any $d \in \mathbb{E}$, the second-order cone*

$$\left\{ x \in \mathbb{E} : \langle d, x \rangle > \frac{1}{\sqrt{2}} \|d\| \|x\| \right\}$$

coincides with its dual cone.

Corollary 2. *For any $d \in \mathbb{E}$, any $r \in (0, 1)$, and any self-adjoint, invertible, linear operator L,*

$$\{x \in \mathbb{E} : \langle d, x \rangle' > r \|d\|' \|x\|'\}^* = \left\{ s \in \mathbb{E} : \langle d, s \rangle > \sqrt{1 - r^2} \|Ld\| \|L^{-1} s\| \right\}$$

where $\langle \cdot, \cdot \rangle' : (x, z) \mapsto \langle Lx, Lz \rangle$.

Proof. For each $r \in (0, 1)$, let M_r denote the invertible linear operator

$$x \mapsto x + \left(\frac{\sqrt{1 - r^2}}{r} - 1 \right) \mathrm{Pr}'_d \, x = x + \left(\frac{\sqrt{1 - r^2}}{r} - 1 \right) L^{-1} \, \mathrm{Pr}_{Ld} \, Lx.$$

As noted in the paragraphs leading to Proposition 1,

$$\begin{aligned} \{x \in \mathbb{E} : \langle d, x \rangle' > r \|d\|' \|x\|'\} &= \left\{ x \in \mathbb{E} : \langle LM_r d, LM_r x \rangle > \frac{1}{\sqrt{2}} \|LM_r d\| \|LM_r x\| \right\} \\ &= M_r^{-1} L^{-1} \left\{ y \in \mathbb{E} : \langle Ld, y \rangle > \frac{1}{\sqrt{2}} \|Ld\| \|y\| \right\} \end{aligned} \tag{8}$$

where we have used $M_r d = (\sqrt{1 - r^2}/r) d$. The adjoint operator M_r^* of M_r is

$$s \mapsto s + \left(\frac{\sqrt{1 - r^2}}{r} - 1 \right) L \, \mathrm{Pr}_{Ld} \, L^{-1} s = L^2 M_r L^{-2} s.$$

We assumed that L is self-adjoint. Hence the dual of the cone in (8) is

$$M_r^* L^* \left\{ y \in \mathbb{E} : \langle Ld, y \rangle > \frac{1}{\sqrt{2}} \|Ld\| \|y\| \right\}^* = L^2 M_r L^{-1} \left\{ w \in \mathbb{E} : \langle Ld, w \rangle > \frac{1}{\sqrt{2}} \|Ld\| \|w\| \right\}$$

by Proposition 1. Note that the inverse operator of M_r is $M_{\sqrt{1-r^2}}$. We then conclude from (8) that this dual cone is precisely

$$L^2 \left\{ w \in \mathbb{E} : \langle d, w \rangle' > \sqrt{1 - r^2} \|d\|' \|w\|' \right\},$$

or equivalently, $\left\{ s \in \mathbb{E} : \langle d, s \rangle > \sqrt{1 - r^2} \|Ld\| \|L^{-1} s\| \right\}$. ∎

2.3. **Second-Order Cone Approximations.** Let f be a ϑ-logarithmically homogeneous, self-concordant barrier for K and let f^* denote its conjugate barrier. Let g (resp. g^*) denote the gradient of f (resp. f^*) and let H (resp. H^*) denote the Hessian of f (resp. f^*).

For each $x \in K$, we define the inner product

$$\langle \cdot, \cdot \rangle_x : \mathbb{E} \times \mathbb{E} \to \mathbb{R} : (u, v) \mapsto D^2 f(x)[u, v] = \langle u, H(x) v \rangle.$$

This is called the *local inner product at x*. The induced norm

$$\|\cdot\|_x : u \mapsto \sqrt{\langle u, u \rangle_x} = \sqrt{\langle u, H(x) u \rangle}.$$

is called the *local norm at x*; see [15].

Similarly for the dual cone K^*, *the local inner product at $s \in K^*$* is

$$\langle \cdot, \cdot \rangle_s^* : \mathbb{E} \times \mathbb{E} \to \mathbb{R} : (u, v) \mapsto D^2 f^*(s)[u, v] = \langle u, H^*(s) v \rangle,$$

and *the local norm at s* is

$$\|\cdot\|_s^* : u \mapsto \sqrt{\langle u, u \rangle_s^*} = \sqrt{\langle u, H^*(s) u \rangle}.$$

For each $x \in K$ and each $r \in (0, \sqrt{\vartheta})$, *the local ball of radius r at x* is the open, convex set

$$\{z \in \mathbb{E} : \|z - x\|_x < r\}.$$

It is denoted by $B_r(x)$. Nesterov and Nemirovski [11, Theorem 2.1.1] showed that $B_1(x) \subset K$ for all $x \in K$.

The smallest open cone containing $B_r(x)$ is $\{z \in \mathbb{E} : \exists t > 0, tz \in B_r(x)\}$. Now $tz \in B_r(x)$ for some $t > 0$ if and only if the quadratic polynomial $t \in \mathbb{R} \mapsto t^2 \|z\|_x^2 - 2t\langle x, z \rangle_x + \|x\|_x^2 - r^2$ has a positive root. Since the polynomial has value $\|x\|_x^2 - r^2 = \vartheta - r^2 > 0$ at $t = 0$, it has a positive root if and only if $\langle x, z \rangle_x^2 > \|z\|_x^2(\|x\|_x^2 - r^2)$ and $\langle x, z \rangle_x > 0$. Thus the smallest open cone can be expressed as

$$\begin{aligned}\left\{tz \in \mathbb{E} : \langle x, z \rangle_x > \sqrt{\|x\|_x^2 - r^2}\|z\|_x\right\} &= \left\{z \in \mathbb{E} : \langle x, z \rangle_x > \sqrt{\vartheta - r^2}\|z\|_x\right\} \\ &= \left\{z \in \mathbb{E} : \langle x, z \rangle_x > \sqrt{1 - r^2/\vartheta}\|x\|_x \|z\|_x\right\},\end{aligned}$$

This cone is called the *second-order cone of radius r along x* and is denoted by $K_r(x)$. It follows immediately from the preceding paragraph that $K_1(x) \subseteq K$. From Corollary 2,

$$\begin{aligned}(K_r(x))^* &= \left\{s \in \mathbb{E} : \langle x, s \rangle > r/\sqrt{\vartheta}\|H(x)^{1/2} x\| \|H(x)^{-1/2} s\|\right\} \\ &= \left\{s \in \mathbb{E} : \langle -g(x), s \rangle_{-g(x)}^* > r\|s\|_{-g(x)}^*\right\},\end{aligned}$$

where $H(X)^{1/2}$ denotes the self-adjoint positive definite square root[1] of $H(X)$.

Similarly for the dual cone K^*, we define, for each $s \in K^*$ and $r > 0$, the *dual second-order cone of radius r along s* as the smallest open cone containing the local ball of radius r at s. It is given by

$$K_r^*(s) = \left\{y \in \mathbb{E} : \langle s, y \rangle_s^* > \sqrt{\vartheta - r^2}\|y\|_s^*\right\}.$$

Since $K_r^*(-g(x)) \subseteq K^*$ for all $x \in K$ and $r \in (0, 1)$, we deduce from elementary duality theory that $(K_r^*(-g(x)))^* \supseteq (K^*)^* = K$. For simplicity of notation, we denote the cone $(K_r^*(-g(x)))^* = K_{\sqrt{\vartheta - r^2}}(x)$ by $\hat{K}_r(x)$.

[1]For any self-adjoint positive definite linear operator H and any integer p and any positive integer q, we denote by $H^{p/q}$ the self-adjoint positive definite linear operator that satisfies $(H^{p/q})^q = H^p$.

Theorem 3. *For any* $r \in (0,1)$,

$$K = \operatorname{int}(X,S) \in K \times K,$$

where $\operatorname{int}(X,S) \in K \times K$ *denotes interior.*

Proof. Since $\hat{K}_r(x) \supseteq K$ for all $x \in K$, we have $K \subseteq \bigcap_{x\in K} \hat{K}_r(x)$. Therefore, it follows that $K \subseteq \operatorname{int}(X,S) \in K \times K$, since K is open.

For the other inclusion, take any $z \notin \operatorname{cl} K$. Since $K = (K^*)^*$, there exists $s \in K^*$ such that $\langle z, s\rangle < 0$. Let $x' = -g^*(s) \in K$. By (5), $s = -g(-g^*(s)) = -g(x') \in K_r^*(-g(x'))$. Together with $\langle z, s\rangle < 0$, we conclude that z lies outside the dual cone $\hat{K}_r(x')$ of $K_r^*(-g(x'))$. Consequently, $\operatorname{cl} K \supseteq \bigcap_{x\in K} \hat{K}_r(x)$. Since K is open, it follows that $K \supseteq \operatorname{int}(X,S) \in K \times K$. □

The theorem suggests that we can approximate K by the "simpler" cone

$$\bigcap_{x\in T} \hat{K}_r(x),$$

where $T \subset K$ is some finite subset. The approximating problem is a second-order cone programming problem. Of course, the problem may not be tractable if $|T|$ is large. (In this case, we may further approximate the second-order cone with a polyhedral cone; see [4].) As an extreme case, we have $|T| = 1$, i.e., we approximate K by a second-order cone. We call the approximations in this extreme case the *second-order cone approximations of* K.

3. The Primal-Dual Second-Order Cone Approximations Algorithm

One of the most well-studied class of CCP problems is *semidefinite programming* (SDP), which is the class of CCP problems over cones of symmetric, positive definite matrices. The study of SDP is largely motivated by its wide applicability; see [3].

In this section, we apply the concept of second-order cone approximations to SDP. We develop a primal-dual interior-point algorithm for SDP, and show that for an underlying cone of n-by-n symmetric, positive definite matrices, the algorithm requires at most $O(\sqrt{n}\ln\frac{1}{\varepsilon})$ iterations to reduce the duality gap by ε. This complexity bound matches the best bound known for SDP.

3.1. **Semidefinite Programming.** Let $\mathbb{S}^n$ denote the space of n-by-n symmetric matrices and $\langle\cdot,\cdot\rangle : \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{R} : (X,Y) \mapsto \operatorname{tr} X^T Y = \sum_{i,j=1}^n X_{ij}Y_{ij}$ denote the *trace inner product.* The norm $\|\cdot\|$ induced by the trace inner product is the Frobenius norm. A *semidefinite programming problem* is the following minimization problem

$$(SDP) \qquad \inf\{\langle \hat{S}, X\rangle : X \in L + \hat{X},\ X \in \operatorname{cl}\mathbb{S}^n_{++}\},$$

where $\hat{X}, \hat{S} \in \mathbb{S}^n$, $L \subset \mathbb{S}^n$ is a vector subspace and $\mathbb{S}^n_{++}$, the cone of symmetric, positive definite n-by-n matrices (also called the *positive definite cone of order* n).

The positive definite cone is self-dual, i.e., $(\mathbb{S}^n_{++})^* = \mathbb{S}^n_{++}$. The Lagrangian dual of the SDP problem is

$$(SDD) \qquad \inf\{\langle \hat{X}, S\rangle : S \in L^\perp + \hat{S},\ S \in \operatorname{cl}\mathbb{S}^n_{++}\}.$$

The *standard logarithmic barrier* for $\mathbb{S}^n_{++}$ is $X \mapsto -\ln\det X$. It is an n-logarithmically homogeneous, self-concordant barrier. Under the trace inner product, its gradient and Hessian are, respectively,

$$g(X) = -X^{-1} \qquad \text{and} \qquad H(X) : \mathbb{S}^n \to \mathbb{S}^n : U \mapsto X^{-1}UX^{-1}.$$

The duality map is $X \mapsto -g(X) = X^{-1}$. The local inner product at $X \in K$ is

$$\langle \cdot, \cdot \rangle_X : (U, V) \mapsto \langle U, H(X)V \rangle = \operatorname{tr} X^{-1}UX^{-1}V,$$

and the local norm at X is

$$\|\cdot\|_X : U \mapsto \sqrt{\operatorname{tr}((X^{-1}U)^2)}.$$

Clearly, the local inner product at the identity matrix I is the trace inner product. Consequently, I is a fixed point of the duality map.

In the next proposition, we state several well-known properties of standard logarithmic barriers for positive definite cones that are essential for the development and analysis of the primal-dual second-order cone approximations algorithm in subsequent subsections. We leave the proof of the proposition to the reader.

Proposition 4. *Let f be the standard logarithmic barrier for a positive definite cone $\mathbb{S}^n_{++}$, and let g and H be its gradient and Hessian respectively. The following are true.*

1. *The conjugate barrier of f differs from f by an additive constant.*
2. *For any $X, Z \in \mathbb{S}^n_{++}$, it holds $H(X)Z \in \mathbb{S}^n_{++}$ and*

$$H(H(X)Z) = H(X)^{-1}H(Z)H(X)^{-1}. \tag{9}$$

3. *For any $X \in \mathbb{S}^n_{++}$, there exists $X_{1/2} \in \mathbb{S}^n_{++}$ such that*

$$H(X_{1/2}) = H(X)^{1/2}. \tag{10}$$

4. *For any $X \in \mathbb{S}^n_{++}$,*

$$H(X)^{-1/2}I = X \tag{11}$$

5. *For any $X \in \mathbb{S}^n_{++}$, there exist positive real constants $\lambda_1, \ldots, \lambda_n$ such that for $p = 1, 2, \ldots,$*

$$\langle H(X)^{-p/2}I, I \rangle = \sum_{i=1}^{n} \lambda_i^p. \tag{12}$$

Proof. For the first statement, direct computation of the conjugate barrier gives $f^* = f - n$. Let $X, Z \in \mathbb{S}^n_{++}$ be arbitrary. Now $H(X)Z = X^{-1}ZX^{-1} \in \mathbb{S}^n_{++}$ and

$$H(H(X)Z)U = (H(X)Z)^{-1}U(H(X)Z)^{-1} = XZ^{-1}XUXZ^{-1}X = H(X)^{-1}H(Z)H(X)^{-1}U$$

for all $U \in \mathbb{S}^n_{++}$ proves the second statement. For the third, we take $X_{1/2} = X^{1/2}$, the symmetric positive definite square root of X. The next statement then follows since $H(X)^{-1/2} = H(X_{1/2})^{-1} = H(X^{-1/2})$. Finally, taking λ_i's to be the eigenvalues of X proves the last statement. □

Proposition 5. *Suppose that*

1. *$\mathbb{E}$ is a Euclidean space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$,*
2. *K is a self-dual, regular, open, convex cone in $\mathbb{E}$,*
3. *f is a ϑ-logarithmically homogeneous, self-concordant barrier for K, and*

4. there exists $E \in K$ such that the local inner product at E coincides with $\langle \cdot, \cdot \rangle$.

Let g and H denote the gradient and Hessian of f, respectively. If f satisfies the properties listed in Proposition 4 with $\mathbb{S}^n_{++}$, n and I replaced by K, ϑ and E respectively, then for any $X, Z \in K$ and $S \in K^* = K$,

$$H(H(X)^{1/2}Z) = H(X)^{-1/2}H(Z)H(X)^{-1/2} \quad \textit{and} \quad \|S\|_{-g(X)} = \|X\|_{-g(S)}.$$

Proof. Let $X, Z \in K$ be arbitrary. It follows from (9) and (10) that

$$\begin{aligned} H(H(X)^{1/2}Z) &= H(H(X_{1/2})Z) \\ &= H(X_{1/2})^{-1}H(Z)H(X_{1/2})^{-1} \\ &= H(X)^{-1/2}H(Z)H(X)^{-1/2}. \end{aligned}$$

Let $S \in K^* = K$ be arbitrary. Let $W = H(S_{1/2})^{-1}X = H(-g(S_{1/2}))X \in K^* = K$ and $Z = H(S_{1/2})W_{1/2} \in K$. Using (9), (10) and (11), we deduce that

$$\begin{aligned} H(Z)X &= H(H(S_{1/2})W_{1/2})X \\ &= H(S_{1/2})^{-1}H(W_{1/2})H(S_{1/2})^{-1}X \\ &= H(S)^{-1/2}H(H(S_{1/2})^{-1}X)^{1/2}H(S_{1/2})^{-1}X \\ &= H(S)^{-1/2}E = S. \end{aligned}$$

Therefore, using (6) and (9), we deduce that

$$\begin{aligned} \|S\|_{-g(X)} &= \langle H(-g(X))S, S\rangle \\ &= \langle H(X)^{-1}H(Z)X, H(Z)X\rangle \\ &= \langle (H(Z)^{-1}H(X)H(Z)^{-1})^{-1}X, X\rangle \\ &= \langle H(H(Z)X)^{-1}X, X\rangle \\ &= \langle H(S)^{-1}X, X\rangle \\ &= \langle H(-g(S))X, X\rangle \\ &= \|X\|_{-g(S)}. \end{aligned}$$

□

Remark 1. The above proposition was first proved for symmetric cones by Nesterov and Todd [13]. The matrix $Z = H(S_{1/2})W_{1/2}$ in the proof is called a *scaling-point* in [13].

3.2. **Description of Algorithm.** In this subsection, we describe a primal-dual interior-point algorithm for SDP.

Assume that both (SDP) and (SDD) have strictly feasible solutions. This implies that the sets of optimal solutions for the primal and dual problems are nonempty and bounded, and the gap between the optimal primal and dual objective values is zero. We may assume further that $\hat{X} \notin L$ and $\hat{S} \notin L^\perp$, for if $\hat{X} \in L$, then the zero matrix is optimal for (SDP), and if $\hat{S} \in L^\perp$, then (SDP) has constant value.

Henceforth in this section, g and H denote the gradient and Hessian of the standard logarithmic barrier for $\mathbb{S}^n_{++}$, respectively.

Fix some $r \in (0, 1)$. Suppose we have a pair of strictly feasible primal-dual solutions $(\tilde{X}, \tilde{S})$; i.e., $\tilde{X}$ and $\tilde{S}$ are positive definite and feasible for (SDP) and (SDD) respectively. As suggested at the end of Section 2.3, we may use the second-order cone $\hat{K}_r(\tilde{X})$ to approximate $K = \mathbb{S}^n_{++}$. Consider the following approximating simple SOCP problem and its dual:

$$(SOCP) \qquad \inf\{\langle \hat{S}, X\rangle : X \in L + \hat{X},\ X \in \operatorname{cl} \hat{K}_r(\tilde{X})\},$$

and

$$(SOCD) \qquad \inf\{\langle \hat{X}, S\rangle : S \in L^{\perp} + \hat{S},\ S \in \operatorname{cl} K_r(-g(\tilde{X}))\},$$

where

$$\begin{aligned} K_r(-g(\tilde{X})) &= \left\{S \in \mathbb{S}^n : \langle -g(\tilde{X}), S\rangle_{-g(\tilde{X})} > \sqrt{n-r^2}\|S\|_{-g(\tilde{X})}\right\} \\ &= \left\{S \in \mathbb{S}^n : \langle \tilde{X}, S\rangle > \sqrt{n-r^2}\|S\|_{-g(\tilde{X})}\right\}, \end{aligned}$$

and

$$\hat{K}_r(\tilde{X}) = (K_r(-g(\tilde{X})))^* = K_{\sqrt{n-r^2}}(\tilde{X}) = \left\{X \in \mathbb{S}^n : \langle \tilde{X}, X\rangle_{\tilde{X}} > r\|X\|_{\tilde{X}}\right\}.$$

Since $\tilde{X} \in K \subseteq \hat{K}_r(\tilde{X})$, $(SOCP)$ is strictly feasible. This implies that $(SOCD)$ has an optimal solution if and only if it is feasible. Clearly, $(SOCD)$ is feasible if and only if there exists $\bar{S} \in L^{\perp} + \hat{S}$ such that the pair $(\tilde{X}, \bar{S})$ is in the set

$$\mathcal{S}_r := \{(X, S) \in K \times K : S \in \operatorname{cl} K_r(-g(X))\}.$$

Therefore if we assume, in addition, that $(\tilde{X}, \tilde{S}) \in \mathcal{S}_r$, then $(SOCD)$ has an optimal solution $\tilde{\tilde{S}}$. Moreover, $\tilde{\tilde{S}} \in \operatorname{cl} K_r(-g(\tilde{X})) \subset K$.

Suppose that we now use the second-order cone $\hat{K}_r(\tilde{\tilde{S}})$ to approximate K. We would then consider the following primal-dual pair of second-order cone approximating problems:

$$(SOCD^*) \qquad \inf\{\langle \hat{X}, S\rangle : S \in L^{\perp} + \hat{S},\ S \in \operatorname{cl} \hat{K}_r(\tilde{\tilde{S}})\},$$

and

$$(SOCP^*) \qquad \inf\{\langle \hat{S}, X\rangle : X \in L + \hat{X},\ X \in \operatorname{cl} K_r(-g(\tilde{\tilde{S}}))\}.$$

Since

$$\begin{aligned} (\tilde{X}, \tilde{\tilde{S}}) &\in \mathcal{S}_r \\ &= \left\{(X, S) \in K \times K : \langle X, S\rangle \geq \sqrt{n-r^2}\|S\|_{-g(X)}\right\} \\ &= \left\{(X, S) \in K \times K : \langle X, S\rangle \geq \sqrt{n-r^2}\|X\|_{-g(S)}\right\} \\ &= \{(X, S) \in K \times K : X \in \operatorname{cl} K_r(-g(S))\}, \end{aligned}$$

the problem $(SOCP^*)$ is feasible, and hence $(SOCP^*)$ has an optimal solution $\tilde{\tilde{X}}$. Clearly, $(\tilde{\tilde{X}}, \tilde{\tilde{S}}) \in \mathcal{S}_r$, so that we can repeat the process with $(\tilde{\tilde{X}}, \tilde{\tilde{S}})$ in place of $(\tilde{X}, \tilde{S})$. We then continue repeating this process until the duality gap between the strictly feasible primal-dual solution pair is below a certain prescribed fraction of the initial duality gap.

This algorithm is called the *primal-dual second-order cone approximations algorithm*. Clearly, all pairs of iterates generated lie in the neighborhood

$$\mathcal{N}_2(r) := \left\{(X, S) \in (K \cap (L + \hat{X})) \times (K \cap (L^{\perp} + \hat{S})) : (X, S) \in \mathcal{S}_r\right\}.$$

Note that $\operatorname{cl} K_r(-g(X))$ is the smallest closed cone containing the local ball of radius r at $-g(X)$. Thus $S \in \operatorname{cl} K_r(-g(X))$ if and only if there exist $\tau \geq 0$ and $W \in \mathbb{S}^n$ such that $S = \tau W$ and $\|W + g(X)\|_{-g(X)} \leq r$. Simplifying the last inequality using $-g(X) = X^{-1}$ and the definition of the local norm yields

$$S \in \operatorname{cl} K_r(-g(X)) \iff \exists \tau \geq 0,\ \|S^{1/2} X S^{1/2} - \tau I\| \leq \tau r,$$

where I is the identity matrix of appropriate size. Consequently

$$\mathcal{N}_2(r) = \left\{ (X, S) \in (L + \hat{X}) \times (L^{\perp} + \hat{S}) : \|S^{1/2} X S^{1/2} - \tau I\| \leq \tau r \text{ for some } \tau \geq 0 \right\}.$$

In the following subsections, we show that the algorithm is polynomial-time.

Primal-Dual Second-Order Cone Approximations Algorithm
Given strictly feasible primal-dual pair $(X^{in}, S^{in}) \in \mathcal{N}_2(r)$ and $\varepsilon > 0$.

(1) Set $k = 0$, $X(0) = X^{in}$ and $S(0) = S^{in}$.
(2) While $\langle X(k), S(k) \rangle > \varepsilon \langle X^{in}, S^{in} \rangle$,
  (a) Solve for the optimal solution $S(k+1)$ of
  $$\min\{\langle \hat{X}, S \rangle : S \in L^{\perp} + \hat{S},\ S \in \operatorname{cl} K_r(-g(X(k)))\}.$$
  (b) Solve for the optimal solution $X(k+1)$ of
  $$\min\{\langle \hat{S}, X \rangle : X \in L + \hat{X},\ X \in \operatorname{cl} K_r(-g(S(k+1)))\}.$$
  (c) Set $k = k + 1$.
(3) Output $(X^{out}, S^{out}) = (X(k), S(k))$.

3.3. **Optimization Over A Second-Order Cone.** In this subsection, we discuss the issue of solving the second-order programming subproblem

$$(SOCP^*) \qquad \min\{\langle \hat{S}, X \rangle : X \in L + \hat{X},\ X \in \operatorname{cl} K_r(-g(\tilde{S}))\}.$$

where $\hat{X} \notin L$, $\hat{S} \notin L^{\perp}$, and $(\tilde{X}, \tilde{S}) \in \mathcal{N}_2(r)$ for some $\tilde{X}$. A similar problem, where $K_r(-g(\tilde{S}))$ is replaced by the second-order cone $\{x \in \mathbb{R}^n : x_1 > \sqrt{x_2^2 + \cdots + x_n^2}\}$, is discussed in [1, pp. 27-28].

Since $(\tilde{X}, \tilde{S}) \in \mathcal{N}_2(r)$, the problem $(SOCP^*)$ has an optimal solution. By replacing $X \in \operatorname{cl} K_r(-g(\tilde{S}))$ with the equivalent pair of constraints

$$\langle \tilde{S}, X \rangle^2 \geq (n - r^2) \|H(\tilde{S})^{-1/2} X\|^2 \quad \text{and} \quad \langle \tilde{S}, X \rangle \geq 0,$$

we deduce the following Fritz John necessary conditions:

$$\mu \hat{S} - \lambda \left(\langle \tilde{S}, X^{opt} \rangle \tilde{S} - (n - r^2) H(\tilde{S})^{-1} X^{opt} \right) - \nu \tilde{S} \in L^{\perp},$$
$$X^{opt} - \hat{X} \in L, \quad \langle \tilde{S}, X^{opt} \rangle^2 - (n - r^2) \|H(\tilde{S})^{-1/2} X^{opt}\|^2 \geq 0, \quad \langle \tilde{S}, X^{opt} \rangle \geq 0,$$
$$\lambda \left(\langle \tilde{S}, X^{opt} \rangle^2 - (n - r^2) \|H(\tilde{S})^{-1/2} X^{opt}\|^2 \right) = 0, \quad \nu \langle \tilde{S}, X^{opt} \rangle = 0,$$
$$\mu, \lambda, \nu \geq 0, \quad (\mu, \lambda, \nu) \neq (0, 0, 0).$$

Under the assumptions $\hat{X} \notin L$, we know that $X^{opt} \neq 0$, and hence $\langle \tilde{S}, X^{opt} \rangle > 0$ as $\tilde{S} \in K$. This further implies that $\nu = 0$ from complementary slackness. Therefore $\lambda > 0$, for otherwise $\mu > 0$, and thus the first condition implies $\hat{S} \in L^{\perp}$, contradicting the assumption

$\hat{S} \notin L^\perp$. By scaling λ to one, and noting that $\mu\tilde{S} - \mu\hat{S} \in L^\perp$, we arrive at the following conditions:

$$\mu\tilde{S} - \Big(\langle \tilde{S}, X^{opt}\rangle \tilde{S} - (n-r^2)H(\tilde{S})^{-1}X^{opt}\Big) \in L^\perp, \tag{13a}$$

$$X^{opt} - \hat{X} \in L, \tag{13b}$$

$$\langle \tilde{S}, X^{opt}\rangle^2 - (n-r^2)\|H(\tilde{S})^{-1/2}X^{opt}\|^2 = 0, \tag{13c}$$

$$\langle \tilde{S}, X^{opt}\rangle > 0, \quad \text{and} \tag{13d}$$

$$\mu \geq 0. \tag{13e}$$

The first two conditions are respectively equivalent to

$$(\mu - \langle \tilde{S}, X^{opt}\rangle)H(\tilde{S})^{-1/2}\tilde{S} + (n-r^2)H(\tilde{S})^{-1/2}X^{opt} \in H(\tilde{S})^{1/2}L^\perp = (H(\tilde{S})^{-1/2}L)^\perp,$$

and

$$H(\tilde{S})^{-1/2}X^{opt} - H(\tilde{S})^{-1/2}\hat{X} \in H(\tilde{S})^{-1/2}L,$$

Therefore,

$$\begin{aligned} H(\tilde{S})^{-1/2}X^{opt} - H(\tilde{S})^{-1/2}\hat{X} &= -\operatorname{Pr}_{H(\tilde{S})^{-1/2}L}(\alpha H(\tilde{S})^{-1/2}\tilde{S} + \hat{X}) \\ &= -\alpha \operatorname{Pr}_{H(\tilde{S})^{-1/2}L} H(\tilde{S})^{-1/2}\tilde{S} - \operatorname{Pr}_{H(\tilde{S})^{-1/2}L}\hat{X}, \end{aligned}$$

where $\alpha = (\mu - \langle \tilde{S}, X^{opt}\rangle)/(n-r^2)$.

Substituting the above expression into (13c) gives a quadratic equation in α. This quadratic equation has two real roots, one of which gives the optimal solution X^{opt}. To determine which root gives the optimal solution, we only need to check the feasibility of the solutions given by the roots and pick the one with the lower objective value.

3.4. **Analysis of Algorithm.** Consider the rate of decrease of the duality gap at each iteration, i.e.,

$$\frac{\langle X(k), S(k)\rangle - \langle X(k+1), S(k+1)\rangle}{\langle X(k), S(k)\rangle}.$$

We claim that the rate of decrease in each iteration is no less than $c_r/\sqrt{n}$, where c_r is a positive constant depending on r.

Theorem 6. *Suppose $n > 2$. Let $(X(k), S(k))$ be the k-th strictly feasible primal-dual pair of iterates of the primal-dual second-order cone approximations algorithm. For all $k \geq 1$,*

$$\frac{\langle X(k), S(k)\rangle - \langle X(k+1), S(k+1)\rangle}{\langle X(k), S(k)\rangle} > \frac{c_r}{\sqrt{n}},$$

where

$$c_r = \frac{(1-r)^2}{16}\min\left\{r^2, \frac{8}{23}\right\} > 0.$$

Proof. Fix a $k \geq 1$. To simplify notation, let $\bar{X} = X(k)$, $\bar{S} = S(k)$, $\bar{\bar{X}} = X(k+1)$ and $\bar{\bar{S}} = S(k+1)$. Clearly,

$$\langle \bar{X}, \bar{S}\rangle \geq \langle \bar{X}, \bar{\bar{S}}\rangle \geq \langle \bar{\bar{X}}, \bar{\bar{S}}\rangle.$$

For each $X \in \mathbb{S}^n_{++}$ and $r \in (0,1)$, let $Q_r(X) : \mathbb{S}^n \to \mathbb{S}^n$ be defined by

$$Q_r(X)U = \langle X, U\rangle X - (n-r^2)H(X)^{-1}U.$$

Consider the optimality conditions (13a), (13b), (13c) and (13e) at $\bar{\bar{S}}$ and $\bar{\bar{X}}$:

$$\mu\hat{X} - Q_r(\bar{X})\bar{\bar{S}} \in L, \qquad \bar{\bar{S}} - \hat{S} \in L^\perp, \qquad \langle \bar{\bar{S}}, Q_r(\bar{X})\bar{\bar{S}}\rangle = 0, \qquad \mu \geq 0,$$

and

$$\nu\hat{S} - Q_r(\bar{\bar{S}})\bar{\bar{X}} \in L^\perp, \qquad \bar{\bar{X}} - \hat{X} \in L, \qquad \langle \bar{\bar{X}}, Q_r(\bar{\bar{S}})\bar{\bar{X}}\rangle = 0, \qquad \nu \geq 0.$$

It follows from these conditions, together with $\bar{X} - \hat{X} \in L$ and $\bar{S} - \hat{S} \in L^\perp$, that

$$\langle Q_r(\bar{X})\bar{\bar{S}} - \mu\bar{X}, \bar{S} - \bar{\bar{S}}\rangle = 0, \tag{14}$$

$$\langle Q_r(\bar{\bar{S}})\bar{\bar{X}} - \nu\bar{\bar{S}}, \bar{X} - \bar{\bar{X}}\rangle = 0, \quad \text{and} \tag{15}$$

$$\langle Q_r(\bar{X})\bar{\bar{S}} - \mu\bar{\bar{X}}, Q_r(\bar{\bar{S}})\bar{\bar{X}} - \nu\bar{\bar{S}}\rangle = 0. \tag{16}$$

These equations imply the following strict inequalities. We prove these inequalities at the end of this subsection.

Claim 7. *It holds*

$$\langle \bar{X}, Q_r(\bar{\bar{S}})\bar{\bar{X}}\rangle > \frac{r^2(1-r)^2}{16n}\langle \bar{X}, \bar{\bar{S}}\rangle\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle, \tag{17}$$

$$\langle \bar{S}, Q_r(\bar{X})\bar{\bar{S}}\rangle > \frac{r^2(1-r)^2}{16n}\langle \bar{X}, \bar{S}\rangle\langle \bar{X}, \bar{\bar{S}}\rangle, \quad \textit{and} \tag{18}$$

$$-\langle Q_r(\bar{X})\bar{\bar{S}}, Q_r(\bar{\bar{S}})\bar{\bar{X}}\rangle < \frac{2\sqrt{2}r^2}{n}\langle \bar{X}, \bar{\bar{S}}\rangle^2\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle. \tag{19}$$

We now consider two cases.

Case 1 : $\mu \leq \frac{\langle \bar{X}, \bar{\bar{S}}\rangle}{\sqrt{n}}$. Using (14) and (18), we deduce

$$\frac{\langle \bar{X}, \bar{\bar{S}}\rangle}{\sqrt{n}}\langle \bar{X}, \bar{S} - \bar{\bar{S}}\rangle \geq \mu\langle \bar{X}, \bar{S} - \bar{\bar{S}}\rangle = \langle \bar{S}, Q_r(\bar{X})\bar{\bar{S}}\rangle > \frac{r^2(1-r)^2}{16n}\langle \bar{X}, \bar{S}\rangle\langle \bar{X}, \bar{\bar{S}}\rangle.$$

Rearranging the terms gives

$$\frac{r^2(1-r)^2}{16\sqrt{n}} < \frac{\langle \bar{X}, \bar{S}\rangle - \langle \bar{X}, \bar{\bar{S}}\rangle}{\langle \bar{X}, \bar{S}\rangle} < \frac{\langle \bar{X}, \bar{S}\rangle - \langle \bar{\bar{X}}, \bar{\bar{S}}\rangle}{\langle \bar{X}, \bar{S}\rangle}.$$

Case 2 : $\mu > \frac{\langle \bar{X}, \bar{\bar{S}}\rangle}{\sqrt{n}}$. In this case

$$\begin{aligned} -\langle Q_r(\bar{X})\bar{\bar{S}}, Q_r(\bar{\bar{S}})\bar{\bar{X}}\rangle\langle \bar{X} - \bar{\bar{X}}, \bar{\bar{S}}\rangle &= \mu\nu\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle\langle \bar{X} - \bar{\bar{X}}, \bar{\bar{S}}\rangle \\ &= \mu\langle \bar{X}, Q_r(\bar{\bar{S}})\bar{\bar{X}}\rangle\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle \\ &> \frac{\langle \bar{X}, \bar{\bar{S}}\rangle}{\sqrt{n}}\frac{r^2(1-r)^2}{16n}\langle \bar{X}, \bar{\bar{S}}\rangle\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle^2, \end{aligned}$$

where the equations follow from (16) and (15), respectively, and the inequality follows from (17). Therefore, by (19), we have

$$\frac{\langle \bar{X}, \bar{\bar{S}}\rangle - \langle \bar{\bar{X}}, \bar{\bar{S}}\rangle}{\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle} > \frac{(1-r)^2}{32\sqrt{2n}}.$$

Hence,

$$\begin{aligned}\frac{\langle \bar{X}, \bar{S}\rangle - \langle \bar{\bar{X}}, \bar{\bar{S}}\rangle}{\langle \bar{X}, \bar{S}\rangle} &\geq 1 - \frac{\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle}{\langle \bar{X}, \bar{\bar{S}}\rangle} \\ &= \frac{1}{\frac{\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle}{\langle \bar{X}, \bar{\bar{S}}\rangle - \langle \bar{\bar{X}}, \bar{\bar{S}}\rangle} + 1} \\ &> \frac{1}{\frac{32\sqrt{2n}}{(1-r)^2} + 1} \\ &= \frac{(1-r)^2}{32\sqrt{2n} + (1-r)^2} \\ &> \frac{(1-r)^2}{46\sqrt{n}}.\end{aligned}$$

Combining both cases, we conclude that

$$\frac{\langle \bar{X}, \bar{S}\rangle - \langle \bar{\bar{X}}, \bar{\bar{S}}\rangle}{\langle \bar{X}, \bar{S}\rangle} > \frac{c_r}{\sqrt{n}},$$

where

$$c_r = \min\left\{\frac{r^2(1-r)^2}{16}, \frac{(1-r)^2}{46}\right\} > 0.$$

□

Once we have a lower bound on the rate of decrease that is inversely proportional to some polynomial in the dimension of the problem, we can invoke the following theorem to conclude that the algorithm is polynomial.

Theorem 8. *Let $\varepsilon \in (0,1)$, $\delta > 0$ and $\omega > 0$ be given. Suppose that a sequence of real numbers $\{\mu_k\}_{k=1}^{\infty}$ satisfies*

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n^{\omega}}\right)\mu_k$$

for all $k = 1, 2, \ldots$. Then there exists an index K with

$$K = O(n^{\omega} \ln \frac{1}{\varepsilon})$$

such that $\mu_K \leq \varepsilon\mu_1$.

Proof. See [18] Theorem 3.2. □

Corollary 9. *Fix $r \in (0,1)$. Given any semidefinite programming problem over symmetric n-by-n matrices and a pair of strictly feasible primal-dual solutions $(X^{in}, S^{in}) \in \mathcal{N}_2(r)$, the primal-dual second-order cone approximations algorithm requires at most $O(\sqrt{n}\ln\frac{1}{\varepsilon})$ iterations to produce a pair of strictly feasible primal-dual solutions (X^{out}, S^{out}) satisfying*

$$\langle X^{out}, S^{out}\rangle \leq \varepsilon\langle X^{in}, S^{in}\rangle.$$

Proof. Apply Theorem 8 with $\delta = c_r$, $\omega = \frac{1}{2}$ and $\mu_k = \langle X(k), S(k)\rangle$, $k = 1, 2, \ldots$. □

We conclude this section with a proof of Claim 7.

Proof of Claim 7. From the necessary optimality conditions (13c) at $\bar{\bar{S}}$ and $\bar{\bar{X}}$, we deduce that

$$\frac{\langle \bar{X}, \bar{\bar{S}} \rangle}{\sqrt{n-r^2}} = \|\bar{\bar{S}}\|_{-g(\bar{X})} = \|\bar{X}\|_{-g(\bar{\bar{S}})}, \quad \text{and} \tag{20}$$

$$\frac{\langle \bar{\bar{X}}, \bar{\bar{S}} \rangle}{\sqrt{n-r^2}} = \|\bar{\bar{X}}\|_{-g(\bar{\bar{S}})} = \|\bar{\bar{S}}\|_{-g(\bar{\bar{X}})}. \tag{21}$$

Since $\langle \bar{\bar{X}}, \bar{\bar{S}} \rangle$, μ and ν are nonnegative, we deduce from (16) that for $U = H(\bar{\bar{S}})^{-1/2}\bar{\bar{X}}$, $V = H(\bar{\bar{S}})^{-1/2}\bar{X}$ and $D = \langle E, V\rangle V - (n-r^2)H(V)^{-1}E$, where $E = I$ is the unique fixed point of the duality map,

$$\begin{aligned}
0 &\leq -\langle Q_r(\bar{X})\bar{\bar{S}}, Q_r(\bar{\bar{S}})\bar{\bar{X}} \rangle \\
&= -\langle \bar{\bar{S}}, \bar{\bar{X}} \rangle \langle Q_r(\bar{X})\bar{\bar{S}}, \bar{\bar{S}} \rangle + (n-r^2)\langle Q_r(\bar{X})\bar{\bar{S}}, H(\bar{\bar{S}})^{-1}\bar{\bar{X}} \rangle \\
&= (n-r^2)\langle Q_r(\bar{X})\bar{\bar{S}}, H(\bar{\bar{S}})^{-1}\bar{\bar{X}} \rangle \\
&= (n-r^2)[\langle \bar{X}, \bar{\bar{S}} \rangle \langle \bar{X}, H(\bar{\bar{S}})^{-1}\bar{\bar{X}} \rangle - (n-r^2)\langle H(\bar{X})^{-1}\bar{\bar{S}}, H(\bar{\bar{S}})^{-1}\bar{\bar{X}} \rangle] \\
&= (n-r^2)\langle \bar{X}, H(\bar{\bar{S}})^{-1/2} I \rangle \langle H(\bar{\bar{S}})^{-1/2}\bar{X}, U \rangle \\
&\quad - (n-r^2)^2 \langle H(\bar{\bar{S}})^{-1/2} H(\bar{X})^{-1} H(\bar{\bar{S}})^{-1/2} I, U \rangle \\
&= (n-r^2)\langle I, V \rangle \langle V, U \rangle - (n-r^2)^2 \langle (H(\bar{\bar{S}})^{1/2} H(\bar{X}) H(\bar{\bar{S}})^{1/2})^{-1} I, U \rangle \\
&= (n-r^2)\langle I, V \rangle \langle V, U \rangle - (n-r^2)^2 \langle H(H(\bar{\bar{S}})^{-1/2}\bar{X})^{-1} I, U \rangle \\
&= (n-r^2)\langle D, U \rangle.
\end{aligned}$$

By (12), there exist positive real constants $\lambda_1, \ldots, \lambda_n$ such that for all positive integers p,

$$\langle H(V)^{-p/2}E, E \rangle = \sum_{i=1}^{n} \lambda_i^p \tag{22}$$

Since $\|V\|^2 = \langle H(\bar{\bar{S}})^{-1}\bar{X}, \bar{X} \rangle = \|\bar{X}\|_{-g(\bar{\bar{S}})}$ and $\langle V, E \rangle = \langle \bar{X}, H(\bar{\bar{S}})^{-1/2}E \rangle = \langle \bar{X}, \bar{\bar{S}} \rangle$, the condition (20) is equivalent to $\langle V, E \rangle = \sqrt{n-r^2}\|V\|$. Therefore,

$$\begin{aligned}
\left(\sum_{i=1}^{n} \lambda_i\right)^2 &= \langle H(V)^{-1/2}E, E \rangle^2 \\
&= \langle V, E \rangle^2 \\
&= (n-r^2)\langle V, V \rangle \\
&= (n-r^2)\langle H(V)^{-1/2}E, H(V)^{-1/2}E \rangle \\
&= (n-r^2)\langle H(V)^{-1}E, E \rangle \\
&= (n-r^2)\sum_{i=1}^{n} \lambda_i^2;
\end{aligned}$$

i.e., the constants $\lambda_1, \ldots, \lambda_n$ satisfy the conditions for Lemmas 12 and 13 in the Appendix. By Lemma 12,

$$\begin{aligned}
\langle D, V\rangle &= \langle E, V\rangle \|V\|^2 - (n - r^2)\langle H(V)^{-1}E, V\rangle \\
&= \frac{1}{n - r^2}\langle E, V\rangle^3 - (n - r^2)\langle H(V)^{-3/2}E, E\rangle \\
&= \frac{1}{n - r^2}\left(\sum_{i=1}^n \lambda_i\right)^3 - (n - r^2)\left(\sum_{i=1}^n \lambda_i^3\right) \\
&\leq -\frac{r^2(1-r)}{n(n-r^2)}\left(\sum_{i=1}^n \lambda_i\right)^3 < 0
\end{aligned}$$

Thus, D is not the origin. Let H^+ be the half-space $\{Z \in \mathbb{S}^n : \langle D, Z\rangle \geq 0\}$. It follows from $(n - r^2)\langle D, U\rangle \geq 0$ that $U \in H^+$. By Lemmas 12 and 13

$$\begin{aligned}
\|D\|^2 &= \langle E, V\rangle^2 \|V\|^2 - 2(n - r^2)\langle E, V\rangle\langle V, H(V)^{-1}E\rangle + (n - r^2)^2\|H(V)^{-1}E\|^2 \\
&= \frac{1}{n - r^2}\langle E, V\rangle^4 - 2(n - r^2)\langle E, V\rangle\langle E, H(V)^{-3/2}E\rangle + (n - r^2)^2\langle E, H(V)^{-2}E\rangle \\
&= \frac{1}{n - r^2}\left(\sum_{i=1}^n \lambda_i\right)^4 - 2(n - r^2)\left(\sum_{i=1}^n \lambda_i\right)\left(\sum_{i=1}^n \lambda_i^3\right) + (n - r^2)^2\left(\sum_{i=1}^n \lambda_i^4\right) \\
&\leq \frac{1}{n - r^2}\left(\sum_{i=1}^n \lambda_i\right)^4 - \frac{2}{n - r^2}\left(1 + \frac{r^2(1-r)}{n}\right)\left(\sum_{i=1}^n \lambda_i\right)^4 \\
&\quad + \frac{1}{n - r^2}\left(1 + \frac{8r^2}{n}\right)\left(\sum_{i=1}^n \lambda_i\right)^4 \\
&= \frac{8r^2}{n(n-r^2)}\left(\sum_{i=1}^n \lambda_i\right)^4 - \frac{2r^2(1-r)}{n(n-r^2)}\left(\sum_{i=1}^n \lambda_i\right)^4 \\
&< \frac{8r^2}{n(n-r^2)}\left(\sum_{i=1}^n \lambda_i\right)^4. \qquad (23)
\end{aligned}$$

Hence, the distance from $\frac{V}{\|V\|}$ to the half-space H^+ is

$$\frac{|\langle D, V\rangle|}{\|D\|\|V\|} = \frac{-\langle D, V\rangle}{\|D\|\|V\|} > \frac{\frac{r^2(1-r)}{n(n-r^2)}(\sum_{i=1}^n \lambda_i)^3}{\left(\sqrt{\frac{8r^2}{n(n-r^2)}}(\sum_{i=1}^n \lambda_i)^2\right)\left(\sqrt{\frac{1}{n-r^2}}(\sum_{i=1}^n \lambda_i)\right)} = \frac{r(1-r)}{\sqrt{8n}}.$$

Since $\frac{U}{\|U\|} \in H^+$, it follows that the distance between $\frac{U}{\|U\|}$ and $\frac{V}{\|V\|}$ is greater than $\frac{r(1-r)}{\sqrt{8n}}$. Therefore,

$$
\begin{aligned}
\frac{r^2(1-r)^2}{8n} &< \|\frac{V}{\|V\|} - \frac{U}{\|U\|}\|^2 \\
&= 2 - 2\frac{\langle V, U\rangle}{\|V\|\|U\|} \\
&= 2 - 2\frac{(n-r^2)\langle V, U\rangle}{\langle E, V\rangle\langle E, U\rangle} \\
&= 2 - 2\frac{(n-r^2)\langle H(\bar{\bar{S}})^{-1}\bar{X}, \bar{\bar{X}}\rangle}{\langle \bar{X}, \bar{\bar{S}}\rangle\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle},
\end{aligned}
$$

which implies that

$$
\frac{r^2(1-r)^2}{16n}\langle \bar{X}, \bar{\bar{S}}\rangle\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle < \langle \bar{X}, \bar{\bar{S}}\rangle\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle - (n-r^2)\langle H(\bar{\bar{S}})^{-1}\bar{X}, \bar{\bar{X}}\rangle = \langle \bar{X}, Q_r(\bar{\bar{S}})\bar{\bar{X}}\rangle. \tag{24}
$$

This proves (17). We can prove (18) in a similar way.

For (19), the left hand side is exactly $(n-r^2)\langle D, U\rangle$, which, by (23), satisfies

$$
\begin{aligned}
(n-r^2)\langle D, U\rangle &= (n-r^2)\Big(\langle D, U - \frac{\langle U, E\rangle}{n}E\rangle + \langle\langle E, V\rangle V - (n-r^2)H(V)^{-1}E\rangle\frac{\langle U, E\rangle}{n}E\Big) \\
&= (n-r^2)\Big(\langle D, U - \frac{\langle U, E\rangle}{n}E\rangle + \frac{\langle U, E\rangle}{n}(\langle E, V\rangle^2 - (n-r^2)\|V\|^2)\Big) \\
&\le (n-r^2)\|D\|\|U - \frac{\langle U, E\rangle}{n}E\| \\
&< (n-r^2)\frac{\sqrt{8r^2}}{\sqrt{n(n-r^2)}}\left(\sum_{i=1}^{n}\lambda_i\right)^2\sqrt{\|U\|^2 - 2\frac{\langle U, E\rangle^2}{n} + \frac{\langle U, E\rangle^2}{n}} \\
&= \frac{\sqrt{8r^2(n-r^2)}}{\sqrt{n}}\langle V, E\rangle^2\sqrt{\frac{\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle^2}{n-r^2} - \frac{\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle^2}{n}} \\
&= \frac{2\sqrt{2}r^2}{n}\langle \bar{X}, \bar{\bar{S}}\rangle^2\langle \bar{\bar{X}}, \bar{\bar{S}}\rangle
\end{aligned}
$$

□

4. Extension To Symmetric Programming

At this point, we note that the development of the primal-dual second-order cone approximations algorithm and its analysis apply to any CCP problem over a self-dual, regular, open, convex cone equipped with a logarithmically homogeneous, self-concordant barrier possessing properties listed in Proposition 4. Although these properties are easily proven for standard logarithmic barriers for positive definite cones, they are not specific to these barriers. Indeed, the standard logarithmic barriers for symmetric cones also possess these properties. This allows for the extension of the primal-dual second-order cone approximations algorithm to symmetric cone programming.

Symmetric cones are self-dual homogeneous cones. A regular, open, convex cone K in an Euclidean space $\mathbb{E}$ with inner product $\langle \cdot, \cdot \rangle$ is *homogeneous* if the group of linear automorphisms of K acts transitively on it, i.e., for every $x, y \in K$, there exists a linear map $A \in L[\mathbb{E}, \mathbb{E}]$ such that $AK = K$ and $Ax = y$. The cone K is *symmetric* if it is self-dual (i.e., $K = K^*$) and homogeneous.

The class of symmetric cones consists of the following five classes of cones, and their direct sums; see [7].

(1) The class of second-order cones.
(2) The class of cones of symmetric, positive definite matrices.
(3) The class of cones of Hermitian, positive definite matrices.
(4) The class of cones of Hermitian, positive definite, quaternion matrices.
(5) An exceptional 27-dimensional cone.

Thus, symmetric cone programming includes LP, SDP and SOCP.

Symmetric cones can also be characterized as cones of squares of Euclidean Jordan algebras. With this characterization, we define the standard logarithmic barriers for symmetric cones and show that these barriers possess the properties listed in Proposition 4.

4.1. **Symmetric Cones and Euclidean Jordan Algebras.** In this subsection, we review concepts in the theory of Euclidean Jordan algebras that are necessary for the purpose of this section. Interested readers are referred to the second and third chapters of [7] for a more comprehensive discussion on the theory of Euclidean Jordan algebras.

Every symmetric cone can be associated with a Euclidean Jordan algebra as follows (see [7]): a cone is symmetric if and only if it is linearly isomorphic to the interior of the cone of squares $K(\mathbb{J})$ of a unique (up to isomorphism) Euclidean Jordan algebra $(\mathbb{J}, \circ)$. The cone of squares of $\mathbb{J}$ can alternatively be described as the set of all elements with nonnegative eigenvalues.

For each $a \in \mathbb{J}$, let $P(a)$ denote its quadratic representation $2L^2(a) - L(a^2)$, where $L(a)$ denotes the linear map $b \in \mathbb{J} \mapsto a \circ b$. The next theorem lists several properties of P.

Theorem 10. *Let $(\mathbb{J}, \circ)$ be an Euclidean Jordan algebra, and let e be its identity element. The following are true.*

1. *For any $a \in K(\mathbb{J})$, the quadratic representation $P(a)$ of a is self-adjoint and positive definite.*
2. *For invertible $a \in \mathbb{J}$,*

$$P(a)K(\mathbb{J}) = K(\mathbb{J}). \tag{25}$$

3. *For any $a \in \mathbb{J}$,*

$$P(a)e = a^2. \tag{26}$$

4. *For any $a \in K(\mathbb{J})$ and any rational t,*

$$P(a^t) = P(a)^t. \tag{27}$$

5. *For any $a, b \in \mathbb{J}$,*

$$P(P(a)b) = P(a)P(b)P(a). \tag{28}$$

6. *For any $a, b \in \mathbb{J}$,*

$$\operatorname{tr} P(a)b = \operatorname{tr} a^2 \circ b. \tag{29}$$

7. *For any $a, b \in \mathbb{J}$,*

$$\det P(a)b = \det a^2 \det b. \tag{30}$$

Proof. Let $a, b \in \mathbb{J}$ be arbitrary.

1. Since $L(a)$ is self-adjoint for all $a \in \mathbb{J}$, so is $P(a)$. From [7], Theorems II.3.1 and III.2.2, we see that the set $\{P(a) : a \in K(\mathbb{J})\}$ is the connected component of the identity map in the set $\{P(a) : a \in \mathbb{J},\ P(a) \text{ invertible}\}$. Thus, $P(a)$ is positive definite for any $a \in K(\mathbb{J})$.
2. See [7], Theorem III.2.2.
3. By definition, $P(a)e = 2L(a)^2 e - L(a^2)e = 2a \circ (a \circ e) - (a \circ a) \circ e = a^2$.
4. From [9], Chapter IV, Theorem 1 and the paragraph following it, we have $P(a)^p = P(a^p)$ for any positive integer p. It then follows that for any positive integer p,
$$P(a) = P((a^{1/p})^p) = P(a^{1/p})^p \implies P(a)^{1/p} = P(a^{1/p}).$$
for any $a \in K(\mathbb{J})$. Consequently, for any $a \in K(\mathbb{J})$ and any positive integers p and q,
$$P(a^{p/q}) = P((a^{1/q})^p) = P(a^{1/q})^p = P(a)^{p/q}.$$
Finally, it follows from [7], Theorem II.3.1 that
$$P(a^{-p/q}) = P((a^{p/q})^{-1}) = P(a^{p/q})^{-1} = P(a)^{-p/q}.$$
5. See [7], Theorem II.3.2.
6. By definition, $\operatorname{tr} P(a)b = 2\langle a, L(a)b\rangle - \operatorname{tr} a^2 \circ b = 2\langle L(a)a, b\rangle - \operatorname{tr} a^2 \circ b = \operatorname{tr} a^2 \circ b$.
7. See [7], Theorem III.4.2.

□

4.2. **Extending the Primal-Dual Second-Order Cone Approximations Algorithm.** Let K be a symmetric cone, and let $\mathbb{J}$ be a Euclidean Jordan algebra such that $K = K(\mathbb{J})$. The *standard logarithmic barrier* for K is the functional

$$f : K(\mathbb{J}) \to \mathbb{R} : a \mapsto -\ln \det(a).$$

Since $\mathbb{J}$ is unique up to isomorphism, the standard logarithmic barrier is well-defined. The standard logarithmic barrier is a r-logarithmically homogeneous, self-concordant barrier for $K(\mathbb{J})$, where $r = \operatorname{rk}(\mathbb{J})$, the rank of $\mathbb{J}$. Under the trace inner product $(a, b) \mapsto \operatorname{tr} a \circ b$, the gradient and Hessian of the standard logarithmic barrier are

$$g(a) = -a^{-1} \qquad \text{and} \qquad H(a) = P(a^{-1}) = P(a)^{-1}$$

(see [7], Theorems III.4.2 and II.3.3). Clearly, the local inner product at the identity element e of $\mathbb{J}$ coincides with the trace inner product, and thus e is the unique fixed point of the duality map $a \mapsto -g(a) = a^{-1}$.

In this subsection, we show that the standard logarithmic barrier of a symmetric cone possesses the properties listed in Proposition 4. This means that the primal-dual second-order cone approximations algorithm and its analysis apply to CCP problems over symmetric cones with their standard logarithmic barriers.

Proposition 11. *Let f be the standard logarithmic barrier for the symmetric cone $K(\mathbb{J})$, and let e be the identity element of $\mathbb{J}$. The following are true.*

1. *The conjugate barrier of f differs from f by an additive constant.*

2. For any $a, b \in K(\mathbb{J})$, $H(a)b \in K(\mathbb{J})$ and $H(H(a)b) = H(a)^{-1}H(b)H(a)^{-1}$.
3. For any $a \in K(\mathbb{J})$, there exists $a_{1/2} \in K(\mathbb{J})$ such that $H(a_{1/2}) = H(a)^{1/2}$.
4. For any $a \in K(\mathbb{J})$, $H(a)^{-1/2}e = a$.
5. For any $a \in K(\mathbb{J})$, there exist $\lambda_1, \ldots, \lambda_n > 0$ such that $\langle H(a)^{-p/2}e, e\rangle = \sum_{i=1}^{r} \lambda_i^p$ for any positive integer p.

Proof. Let $a, b \in K(\mathbb{J})$ be arbitrary.

1. From (4), we see that $f^*(a) = f^*(-g(a^{-1})) = -f(a^{-1}) - r = \ln\det a^{-1} - r = \ln(\det a)^{-1} - r = -\ln\det a - r = f(a) - r$.
2. It follows from (25) that $H(a)b = P(a^{-1})b \in P(a^{-1})K(\mathbb{J}) = K(\mathbb{J})$. Using (28), we deduce that

$$\begin{aligned} H(H(a)b) &= P(P(a^{-1})b)^{-1} = (P(a^{-1})P(b)P(a^{-1}))^{-1} \\ &= P(a^{-1})^{-1}P(b)^{-1}P(a^{-1})^{-1} = H(a)^{-1}H(b)H(a)^{-1}. \end{aligned}$$

3. Let $a_{1/2} = a^{1/2}$. Using (27), we deduce that $H(a_{1/2}) = P(a^{1/2})^{-1} = P(a)^{-1/2} = H(a)^{1/2}$.
4. Using (26) and (27), we deduce that $H(a)^{-1/2}e = P(a)^{1/2}e = P(a^{1/2})e = (a^{1/2})^2 = a$.
5. For $i = 1, \ldots, n$, let $\lambda_i = \lambda_i(a) > 0$, the eigenvalues of a. It follows from (26) and (27) that

$$\begin{aligned} \langle H(a)^{-p/2}e, e\rangle &= \langle P(a)^{p/2}e, e\rangle = \langle P(a^{p/2})e, e\rangle = \langle a^p, e\rangle \\ &= \operatorname{tr} a^p = \sum_{i=1}^{r} \lambda_i(a)^p = \sum_{i=1}^{r} \lambda_i^p \end{aligned}$$

for any positive integer p.

□

5. Comparison with the Primal-Dual Cone Affine Scaling Algorithm

In 1999, Berkelaar, Sturm and Zhang [5] developed the primal-dual cone affine scaling (CAS) algorithm for SDP that combines the primal and dual problems, and uses Nesterov-Todd scaling points to scale the combined primal-dual problem before approximating the resulting problem with a standard second-order cone (i.e., a second-order cone centered along I, the identity matrix). In this section, we compare the primal-dual CAS algorithm with our algorithm.

The first difference is the generality of the concepts behind the algorithms. The primal-dual CAS algorithm is motivated as an extension of primal CAS algorithms to the primal-dual setting. Since the primal-dual CAS algorithm uses the Nesterov-Todd scaling points, which are not known to be well-defined beyond symmetric cone programming, it is reasonable to expect that the primal-dual CAS algorithm can at most be extended to symmetric cone programming. On the other hand, we motivate the concept of second-order cone approximations as a tool for CCP in general, not just for SDP (or symmetric cone programming). Although there are some difficulties in extending our algorithm beyond symmetric cone programming, the generality of our principle concept makes such extensions plausible.

The second difference is in the algorithms themselves. They differ in a fundamental way. Our algorithm strictly alternates between the primal problem and the dual problem. The primal-dual CAS algorithm, like most primal-dual algorithms, instead works in the setting of a combined primal-dual space.

Besides this fundamental difference, the algorithms use different types and sizes of neighborhoods. Our algorithm uses the standard primal-dual neighborhood based on the Frobenius norm, whereas the primal-dual CAS algorithm uses a variant of the standard neighborhood which was shown by Sturm and Zhang [16] to properly contain the standard neighborhood of the same radius. Thus, it seems that our algorithm uses a smaller neighborhood than the primal-dual CAS algorithm. However, a careful analysis shows that while our algorithm allows the neighborhood radius to be arbitrarily close to 1, the primal-dual CAS algorithm produces pairs of iterates that stay in a neighborhood of radius no more than $1/\sqrt{2}$. (This is due to the fact that in the combined primal-dual setting, the underlying cone is the direct product of $\mathbb{S}^n_{++}$ with itself, which implies that the underlying barrier has complexity parameter twice that of the original barrier.)

Despite these differences, the two algorithms are very similar when applied in the same context. Specifically, suppose that we combine the primal and dual problems into the self-dual SDP problem

$$(SDPD) \quad \inf\{\langle \hat{S}, X\rangle + \langle \hat{X}, S\rangle : X \in L + \hat{X},\ S \in L^\perp + \hat{S},\ (X, S) \in (\operatorname{cl}\mathbb{S}^n_{++}) \times (\operatorname{cl}\mathbb{S}^n_{++})\},$$

and apply our algorithm to $(SDPD)$. Since $(SDPD)$ is self-dual, the primal-dual alternation vanishes. Moreover, we are now in the setting of the primal-dual CAS algorithm.

In the primal-dual CAS algorithm, $(SDPD)$ is scaled by the pair of matrices $(W(X,S), W(S,X))$, where $W(X,S)$ denotes the Nesterov-Todd scaling matrix from X to S, before using a standard second-order cone to locally approximate $\mathbb{S}^n_{++}$. Our algorithm, when applied to $(SDPD)$, can be viewed similarly. Instead of scaling with the pair $(W(X,S), W(S,X))$, our algorithms uses the pair (S,X) to scale and then uses a standard second-order cone to locally approximate $\mathbb{S}^n_{++}$. Thus, from this perspective, our algorithm, when applied in the setting of combined primal-dual space, differs from the primal-dual CAS algorithm only in the choice of scaling matrices.

6. Acknowledgements

I thank Professor James Renegar for carefully reading through this paper and providing invaluable suggestions on its presentation. I thank the referees for their useful comments that greatly improved the readability of this paper.

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Program.*, 95:3–51, 2003.
[2] F. Alizadeh and S. H. Schmieta. Potential reduction methods for symmetric cone programming. Technical Report RRR20-99, RUTCOR, Rutgers University, New Brunswick, NJ, USA, September 1999.
[3] A. Ben-Tal and A. S. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS SIAM Ser. Optim. SIAM Publication, Philadelphia, PA, USA, 2001.
[4] A. Ben-Tal and A. S. Nemirovski. On polyhedral approximations of the second-order cone. *Math. Oper. Res.*, 26(2):193–205, 2001.
[5] A. B. Berkelaar, J. F. Sturm, and S. Zhang. Polynomial primal-dual cone affine scaling for semidefinite programming. *Appl. Numer. Math.*, 29(3):317–333, 1999.
[6] S. Boyd and L. Vandenberghe. Semidefinite programming. *SIAM Rev.*, 38:49–95, 1996.

[7] J. Faraut and A. Korányi. *Analysis on Symmetric Cones.* Oxford Press, New York, NY, USA, 1994.
[8] L. Faybusovich. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J. Comput. Appl. Math.*, 86:149–175, 1997.
[9] M. Koecher. *Jordan Algebras and Their Applications.* Lect. Notes. University of Minnesota, Minneapolis, MN, USA, 1962.
[10] N. Megiddo. A variation on Karmarkar's algorithm. Preliminary report, IBM San Jose Research Laboratory, San Jose, CA, USA, 1984.
[11] Yu. E. Nesterov and A. S. Nemirovski. *Interior Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Stud. Appl. Math.* SIAM Publication, Philadelphia, PA, USA, 1994.
[12] Yu. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22:1–46, 1997.
[13] Yu. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8:324–364, 1998.
[14] M. W. Padberg. Solution of a nonlinear programming problem arising in the projective method for linear programming. Manuscript, School of Business and Administration, New York University, New York, NY, USA, 1985.
[15] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization.* MPS SIAM Ser. Optim. SIAM Publication, Philadelphia, PA, USA, 2001.
[16] J. F. Sturm and S. Zhang. An $O(\sqrt{n}L)$ iteration bound primal-dual cone affine scaling algorithm for linear programming. *Math. Program.*, 72:177–194, 1996.
[17] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications.* Springer-Verlag, Berlin-Heidelberg-New York, 2000.
[18] S. J. Wright. *Primal-Dual Interior-Point Methods.* SIAM Publication, Philadelphia, PA, USA, 1997.

Appendix A. Technical Lemmas

Lemma 12. *If $\lambda_1, \ldots, \lambda_m > 0$ $(m > 1)$ satisfies $(\sum_{i=1}^m \lambda_i)^2 = (m - r^2) \sum_{i=1}^m \lambda_i^2$ for some $r \in (0, 1)$, then*

$$(m - r^2)^2 \sum_{i=1}^m \lambda_i^3 \geq \left(1 + \frac{r^2(1 - r)}{m}\right) \left(\sum_{i=1}^m \lambda_i\right)^3.$$

Proof. Let $D := \{\lambda \in \mathbb{R}^m : \sum_{i=1}^m \lambda_i = 1,\ (\sum_{i=1}^m \lambda_i)^2 = (m - r^2) \sum_{i=1}^m \lambda_i^2\}$, which is a closed, bounded subset of the positive orthant $\mathbb{R}^m_{++}$ as it is the intersection of the closed ball of radius $r/(m - r^2)$ centered at $[1/(m - r^2), \ldots, 1/(m - r^2)]^T$ with the affine subspace of vectors whose coordinates sum to one. Due to the positive homogeneity of the inequality, it suffices to show that

$$(m - r^2)^2 \min \left\{ \sum_{i=1}^m \lambda_i^3 : \lambda \in D \right\} \geq 1 + \frac{r^2(1 - r)}{m}.$$

If $m = 2$, then we necessarily have $(\lambda_1, \lambda_2) = ((1 + r/\sqrt{2 - r^2})/2, (1 - r/\sqrt{2 - r^2})/2)$. Consequently, $(2 - r^2)^2(\lambda_1^3 + \lambda_2^3) = 1 + (r^2 - r^4)/2 > 1 + r^2(1 - r)/2$. Henceforth, we assume that $m > 2$.

Consider the real polynomial $p : x \mapsto x^m + \sum_{i=0}^{m-1} a_i x^i$ of degree m with roots $\lambda_1, \ldots, \lambda_m$. Now

$$
\begin{aligned}
a_{m-1} &= -\sum_{i=1}^{m} \lambda_1 = -1 \\
a_{m-2} &= \sum_{1 \le i < j \le m} \lambda_i \lambda_j = \big((\sum_{i=1}^{m} \lambda_i)^2 - \sum_{i=1}^{m} \lambda_i^2 \big)/2 \\
&= \frac{m - r^2 - 1}{2(m - r^2)} \\
a_{m-3} &= -\sum_{1 \le i < j < k \le m} \lambda_i \lambda_j \lambda_k = -\left(2 \sum_{i=1}^{m} \lambda_i^3 - 3 \left(\sum_{i=1}^{m} \lambda_i^2 \right) \left(\sum_{i=1}^{m} \lambda_i \right) + \left(\sum_{i=1}^{m} \lambda \right)^3 \right) /6 \\
&= -\left(2 \sum_{i=1}^{m} \lambda_i^3 - \frac{3}{m - r^2} + 1 \right) /6
\end{aligned}
$$

Since p has m positive real roots, its $(m-3)$-rd derivative

$$
\begin{aligned}
p^{(m-3)}(x) =& [m(m-1)\cdots 4]x^3 - [(m-1)(m-2)\cdots 3]x^2 \\
&+ \frac{(m - r^2 - 1)(m-2)(m-3)\cdots 3}{m - r^2} x \\
&- [(m-3)(m-4)\cdots 2] \frac{2\sum_{i=1}^{m} \lambda_i^3 - \frac{3}{m-r^2} + 1}{6}
\end{aligned}
$$

must have three positive real roots, i.e., the cubic polynomial

$$
\begin{aligned}
\frac{(m - r^2) p^{(m-3)}}{(m-3)(m-4)\cdots 4} : x \mapsto & m(m - r^2)(m-1)(m-2)x^3 - 3(m - r^2)(m-1)(m-2)x^2 \\
&+ 3(m - r^2 - 1)(m-2)x - (m - r^2)(2\sum_{i=1}^{m} \lambda_i^3 - \frac{3}{m - r^2} + 1)
\end{aligned}
$$

has three positive real roots. The two stationary points of the cubic polynomial are $\alpha := \big(1 \pm r/\sqrt{(m-1)(m-r^2)}\big)/m$. Thus, the value of the polynomial must be nonpositive at $\alpha = \big(1 + r/\sqrt{(m-1)(m-r^2)}\big)/m$. Hence,

$$
\begin{aligned}
2(m - r^2) \sum_{i=1}^{m} \lambda_i^3 \ge & m(m - r^2)(m-1)(m-2)\alpha^3 - 3(m - r^2)(m-1)(m-2)\alpha^2 \\
&+ 3(m - r^2 - 1)(m-2)\alpha - (m - r^2)(1 - \frac{3}{m - r^2})
\end{aligned}
\tag{31}
$$

Since α is a root of

$$
\begin{aligned}
x \mapsto & 3m(m - r^2)(m-1)(m-2)x^2 - 6(m - r^2)(m-1)(m-2)x + 3(m - r^2 - 1)(m-2) \\
&= (m-2)\big[3m(m - r^2)(m-1)x^2 - 6(m - r^2)(m-1)x + 3(m - r^2 - 1)\big],
\end{aligned}
$$

we have

$$\alpha^2 = -\frac{(m-r^2-1)}{m(m-r^2)(m-1)} + \frac{2}{m}\alpha \tag{32}$$

$$\begin{aligned}
\implies \alpha^3 &= -\frac{(m-r^2-1)}{m(m-r^2)(m-1)}\alpha + \frac{2}{m}\alpha^2 \\
&= -\frac{(m-r^2-1)}{m(m-r^2)(m-1)}\alpha - \frac{2(m-r^2-1)}{m^2(m-r^2)(m-1)} + \frac{4}{m^2}\alpha
\end{aligned}$$

$$= -\frac{2(m-r^2-1)}{m^2(m-r^2)(m-1)} + \frac{3m^2-3mr^2-3m+4r^2}{m^2(m-r^2)(m-1)}\alpha. \tag{33}$$

From (31),

$$\begin{aligned}
&(m-r^2)^2\sum_{i=1}^m \lambda_i^3 \\
\geq &\frac{(m-r^2)}{2}\begin{pmatrix} m(m-r^2)(m-1)(m-2)\alpha^3 - 3(m-r^2)(m-1)(m-2)\alpha^2 \\ +3(m-r^2-1)(m-2)\alpha - (m-r^2)\left(1-\frac{3}{m-r^2}\right)\end{pmatrix} \\
= &\frac{(m-r^2)}{2}\left(\frac{2(1+r^2)}{m} + 2\left(\frac{2r^2}{m} - r^2\right)\alpha\right)
\end{aligned}$$

where we have used (32) and (33) in the equality. Finally, substituting for α gives

$$\begin{aligned}
(m-r^2)^2\sum_{i=1}^m \lambda_i^3 &\geq (m-r^2)\left(\frac{1}{m} + \frac{2r^2}{m^2} - \frac{r^3(m-2)}{m^2\sqrt{(m-1)(m-r^2)}}\right) \\
&= 1 - \frac{r^2}{m} + \frac{2r^2}{m} - \frac{2r^4}{m^2} - \frac{r^3(m-2)\sqrt{m-r^2}}{m^2\sqrt{m-1}} \\
&> 1 + \frac{r^2(1-r)}{m},
\end{aligned}$$

where the strict inequality follows from

$$2r + \frac{(m-2)\sqrt{m-r^2}}{\sqrt{m-1}} \leq 1 + r^2 + \frac{(m-2)(m-r^2)}{(m-1)} = \frac{m^2-m-(1-r^2)}{m-1} < m.$$

□

Lemma 13. *If $\lambda_1, \dots, \lambda_m > 0$ $(m > 1)$ satisfies $(\sum_{i=1}^m \lambda_i)^2 = (m-r^2)\sum_{i=1}^m \lambda_i^2$ for some $r \in (0,1)$, then*

$$(m-r^2)^3\sum_{i=1}^m \lambda_i^4 \leq \left(1 + \frac{8r^2}{m}\right)\left(\sum_{i=1}^m \lambda_i\right)^4.$$

Proof. Let $D := \{\lambda \in \mathbb{R}^m : \sum_{i=1}^m \lambda_i = 1,\ (\sum_{i=1}^m \lambda_i)^2 = (m-r^2)\sum_{i=1}^m \lambda_i^2\}$, which is a closed, bounded subset of the positive orthant $\mathbb{R}^m_{++}$ (see proof of preceding lemma). Due to the positive homogeneity of the inequality, it suffices to show that

$$(m-r^2)^3 \max\left\{\sum_{i=1}^m \lambda_i^4 : \lambda \in D\right\} \leq \left(1+\frac{8r^2}{m}\right).$$

Since D is compact and $\sum_{i=1}^m \lambda_i^4$ is continuous over D, the above maximum is attained. Let $\tilde{\lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m) \in \mathbb{R}^m_{++}$ be the maximum. Then, $\tilde{\lambda}$ satisfies the following Karush-Kuhn-Tucker optimality conditions:

$$4\lambda_j^3 - \mu - \nu\left(2\sum_{i=1}^m \lambda_i - 2(m-r^2)\lambda_j\right) = 0, \qquad j = 1, \ldots, m$$

$$\sum_{i=1}^m \lambda_i = 1 \qquad \text{and} \qquad \left(\sum_{i=1}^m \lambda_i\right)^2 = (m-r^2)\sum_{i=1}^m \lambda_i^2.$$

If we take the weighted sum of the first m conditions with weights $\tilde{\lambda}_i$, we get

$$4\sum_{i=1}^m \tilde{\lambda}_i^4 - \mu\sum_{i=1}^m \tilde{\lambda}_i - 2\nu\left(\left(\sum_{i=1}^m \tilde{\lambda}_i\right)^2 - (m-r^2)\sum_{i=1}^m \tilde{\lambda}_i^2\right) = 0,$$

which implies that $\mu = 4\sum_{i=1}^m \tilde{\lambda}_i^4$ under the last two conditions. Summing up the first m conditions gives

$$4\sum_{i=1}^m \tilde{\lambda}_i^3 - m\mu - 2\nu r^2 \sum_{i=1}^m \tilde{\lambda}_i = 0,$$

which implies that $2\nu = \frac{1}{r^2}(4\sum_{i=1}^m \tilde{\lambda}_i^3 - m\mu)$. Therefore,

$$\begin{aligned}
2\nu + \mu &= \frac{4\sum_{i=1}^m \tilde{\lambda}_i^3 - m\mu}{r^2} + \mu \\
&= \frac{4\sum_{i=1}^m \tilde{\lambda}_i^3 - (m-r^2)\mu}{r^2} \\
&= \frac{4}{r^2}\left[\sum_{i=1}^m \tilde{\lambda}_i^3 - (m-r^2)\sum_{i=1}^m \tilde{\lambda}_i^4\right].
\end{aligned}$$

Now,

$$\begin{aligned}
\left(\sum_{i=1}^m \tilde{\lambda}_i^3\right)^2 &\leq \left(\sum_{i=1}^m \tilde{\lambda}_i^4\right)\left(\sum_{i=1}^m \tilde{\lambda}_i^2\right) \\
&= (m-r^2)\left(\sum_{i=1}^m \tilde{\lambda}_i^4\right)\left(\sum_{i=1}^m \tilde{\lambda}_i^2\right)^2 \\
&\leq (m-r^2)\left(\sum_{i=1}^m \tilde{\lambda}_i^4\right)\left(\sum_{i=1}^m \tilde{\lambda}_i^3\right)\left(\sum_{i=1}^m \tilde{\lambda}_i\right) \\
\Longrightarrow \qquad \sum_{i=1}^m \tilde{\lambda}_i^3 &\leq (m-r^2)\left(\sum_{i=1}^m \tilde{\lambda}_i^4\right),
\end{aligned}$$

where the first two inequalities follow from the Cauchy-Schwartz inequality. Thus, $2\nu+\mu \leq 0$. Consequently, at least one root of the cubic polynomial $t \mapsto 4t^3 - \mu - \nu(2 - 2(m-r^2)t)$

is nonpositive. It follows that the components of $\tilde{\lambda}$, which are positive roots of this cubic polynomial, can only take the values of the other two roots.

Suppose that p components of $\tilde{\lambda}$ take the value of α and $m-p$ components take the value of β, with $\alpha > \beta > 0$. Since $r > 0$, we deduce from the last Karush-Kuhn-Tucker condition that $0 < p < m$. The condition $\sum_{i=1}^m \tilde{\lambda}_i = 1$ is equivalent to $p\alpha + (m-p)\beta = 1$, and

$$\left(\sum_{i=1}^m \tilde{\lambda}_i\right)^2 = (m-r^2)\left(\sum_{i=1}^m \tilde{\lambda}_i^2\right) \iff 1 = (m-r^2)(p\alpha^2 + (m-p)\beta^2).$$

Together with $p\alpha + (m-p)\beta = 1$, we deduce that $mp(m-r^2)\alpha^2 - 2p(m-r^2)\alpha + p - r^2 = 0$. Solving for α gives two possible values, namely, $\frac{1}{m} \pm \frac{r\sqrt{m-p}}{m\sqrt{p(m-r^2)}}$. If $\alpha < \frac{1}{m}$, then it follows from $p\alpha + (m-p)\beta = 1$ that $\beta > \frac{1}{m} > \alpha$, a contradiction with our assumption. Hence

$$\alpha = \frac{1}{m} + \frac{r\sqrt{m-p}}{m\sqrt{p(m-r^2)}}. \tag{34}$$

From $mp(m-r^2)\alpha^2 - 2p(m-r^2)\alpha + p - r^2 = 0$, we deduce that

$$\alpha^2 = \frac{2}{m}\alpha - \frac{p-r^2}{mp(m-r^2)}, \tag{35}$$

and

$$\begin{aligned}
\alpha^4 &= \frac{4}{m^2}\alpha^2 - \frac{4(p-r^2)}{m^2p(m-r^2)}\alpha + \frac{(p-r^2)^2}{m^2p^2(m-r^2)^2} \\
&= \frac{8}{m^3}\alpha - \frac{4(p-r^2)}{m^3p(m-r^2)} - \frac{4(p-r^2)}{m^2p(m-r^2)}\alpha + \frac{(p-r^2)^2}{m^2p^2(m-r^2)^2} \\
&= \frac{4(mp - 2pr^2 + mr^2)}{m^3p(m-r^2)}\alpha - \frac{(p-r^2)(3mp - 4pr^2 + mr^2)}{m^3p^2(m-r^2)^2} \qquad (36)
\end{aligned}$$

Thus, from $1 = (m-r^2)(p\alpha^2 + (m-p)\beta^2)$, we have

$$\begin{aligned}
&(m-r^2)^3\sum_{i=1}^m \lambda_i^4 \\
=&(m-r^2)^3\left(p\alpha^4 + \frac{(1-(m-r^2)p\alpha^2)^2}{(m-p)(m-r^2)^2}\right) \\
=&\frac{m-r^2}{m-p}(m(m-r^2)^2p\alpha^4 - 2(m-r^2)p\alpha^2 + 1) \\
=&\frac{m-r^2}{m-p}\left(\frac{4(m-r^2)(m-2p)r^2}{m^2}\alpha + \frac{m^2p - mp^2 + 4p^2r^2 - 4pr^4 + mr^4}{m^2p}\right),
\end{aligned}$$

where we have used (35) and (36) for the third equality. Substituting (34) gives

$$(m-r^2)^3\sum_{i=1}^m \lambda_i^4 = \frac{m-r^2}{m}\left(1 + \frac{4r^2}{m} - \frac{8r^4}{m^2} + \frac{r^4}{(m-p)p} + \frac{4(m-2p)r^3\sqrt{(m-r^2)}}{m^2\sqrt{p(m-p)}}\right).$$

Since the last two terms are maximized when $p = 1$, we have

$$\begin{aligned}
(m-r^2)^3 \sum_{i=1}^{m} \lambda_i^4 &\le \frac{m-r^2}{m}\left(1 + \frac{4r^2}{m} - \frac{8r^4}{m^2} + \frac{r^4}{m-1} + \frac{4(m-2)r^3\sqrt{(m-r^2)}}{m^2\sqrt{m-1}}\right) \\
&= 1 + \frac{3r^2}{m} - \frac{12r^4}{m^2} + \frac{8r^6}{m^3} + \frac{r^4(m-r^2)}{m(m-1)} + \frac{4r^3(m-r^2)(m-2)\sqrt{m-r^2}}{m^3\sqrt{m-1}} \\
&< 1 + \frac{3r^2}{m} - \frac{12r^4}{m^2} + \frac{8r^6}{m^3} + \frac{r^4}{m}\left(1 + \frac{2}{m}\right) + \frac{4r^3}{m} \\
&= 1 + \frac{3r^2 + r^4 + 4r^3}{m} - \frac{8r^4}{m^2} + \frac{8r^6}{m^3} \\
&< 1 + \frac{8r^2}{m} - \frac{8r^4}{m^2} + \frac{8r^4}{m^3} \le 1 + \frac{8r^2}{m}.
\end{aligned}$$

□

Division of Mathematical Sciences, Nanyang Technological University, 1 Nanyang Walk, Blk 5 Level 3, Singapore 637616, Singapore

E-mail address: `cbchua@ntu.edu.sg`