A T-algebraic approach to primal-dual interior-point algorithms

CHEK BENG CHUA

Abstract. Three primal-dual interior-point algorithms for homogeneous cone programming are presented. They are a short-step algorithm, a large-update algorithm, and a predictor-corrector algorithm. These algorithms are described and analyzed based on a characterization of homogeneous cone via T-algebra. The analysis show that the algorithms have polynomial iteration complexity.

Key words and phrases. Homogeneous cone programming; T-algebra; Primal-dual interior-point algorithm.

1. Introduction

Primal-dual interior-point algorithms—first designed for linear programming (see, e.g., [19]), and subsequently extended to semidefinite programming (see, e.g., [18, Part II]) and symmetric cone programming (see, e.g., [13])—are the most widely used interior-point algorithms. At the same time, they are able to achieve the best iteration complexity bound known to date. These strongly motivates further development of primal-dual interior-point algorithms for wider classes of optimization problems.

Various studies to extend primal-dual algorithms beyond symmetric cone programming involves the v-space approach; i.e., the use of scaling points. Tunçel [15] showed that primal-dual symmetric algorithms that use the v-space approach can only be designed for symmetric cone programming. By dropping primal-dual symmetry, Tunçel [16] designed primal-dual algorithms based on the v-space approach. However, polynomial iteration complexity bounds were established only for symmetric cone programming.

This paper focuses on primal-dual interior-point algorithms for homogeneous cone programming.

Homogeneous cone programming, which will be formally defined in the next section, is a class of convex programming problems. It includes, as a proper sub-class, symmetric cone programming. While there is a finite number of non-isomorphic symmetric cones of each dimension, this number is uncountable for homogeneous cones when the dimension is at least eleven [17]. Thus the collection of homogeneous cones is significantly larger than the subclass of symmetric cones.

On the other hand, the author [4] proved that every homogeneous cone can be represented as the intersection of some cone of symmetric positive definite matrices with a suitable linear subspace. Thus all homogeneous cone programming problems can be formulated as semidefinite programming problems, which are special cases of symmetric cone programming problems. However it is argued in the same paper that there are advantages in designing algorithms specifically for homogeneous cone programming as the ranks of these cones—which determines the iteration complexity of interior-point algorithms [9]—are always no more than the ranks of the representing positive definite cones.

The primal-dual algorithms in this paper rely on a characterization of homogeneous cones via T-algebras.

Algebraic structures of convex cones had been successfully exploited in the design and analysis of interior-point algorithms for convex conic programming. This is most notable for symmetric cone programming [1, 3, 8, 14]) where Jordan algebraic structures of symmetric cones were used. The primal-dual algorithms and their analyses in this paper can be applied to symmetric cone programming as a special case of homogeneous cones programming. These these algorithms differ from those in [1] in the choice of search direction. The search direction in this paper is a generalization of the S-Ch-MT search direction for semidefinite programming described in [11], which was motivated via theoretical complexity and computational trials. On the other hand, as a search direction for homogeneous cone programming, this generalization is a natural consequence of the T-algebraic structure of homogeneous cones.

This paper is organized as follows. Section 2 begins with a formal definition of homogeneous cone programming, which is followed by a brief introduction to T-algebras. The quadratic representations in Euclidean Jordan algebras are generalized to T-algebras. Section 2 then concludes with a discussion on dual homogeneous cones. Section 3 defines the primal-dual central path of a pair of primal-dual homogeneous cone programming problems that is based on optimal barriers for the problems, and establishes various technical results needed for the design and analysis of path-following algorithms. Section 4 presents a general framework for primal-dual path-following algorithms, and describes three specializations—a short-step algorithm, a large-update algorithm, and a predictor-corrector algorithm—with proofs of their respective iteration complexity.

2. Homogeneous Cone Programming

The main object of study in this paper is a class of convex programming problems known as *homogeneous cone programming* or *convex conic programming over homogeneous cones*. We begin with a description of the programming problems in this class.

Let $\mathbb{E}$ denote a Euclidean space with inner product $\langle \cdot, \cdot \rangle$.

Definition 1 (Homogeneous cone)**.** A *homogeneous cone* is a pointed, open, convex cone $\mathcal{K} \in \mathbb{E}$ whose automorphism group $\mathrm{Aut}(\mathcal{K})$ acts transitively on it; i.e., for every pair $(\mathbf{x}, \mathbf{y}) \in \mathcal{K} \oplus \mathcal{K}$, there exists a linear map $\boldsymbol{\mathcal{A}} : \mathbb{E} \to \mathbb{E}$ satisfying $\boldsymbol{\mathcal{A}}(\mathcal{K}) = \mathcal{K}$ and $\boldsymbol{\mathcal{A}}(\mathbf{x}) = \mathbf{y}$.

Henceforth, $\mathcal{K}$ denotes a homogeneous cone in $\mathbb{E}$, $\widehat{\mathbf{x}}, \widehat{\mathbf{s}}$ denote vectors in $\mathbb{E}$ and $\mathcal{L}$ denotes a linear subspace of $\mathbb{E}$.

We consider the following homogeneous cone programming problem

$$\begin{array}{ll} \inf\limits_{\mathbf{x}} & \langle \widehat{\mathbf{s}}, \mathbf{x} \rangle \\ \text{subject to} & \mathbf{x} \in \mathcal{L} + \{\widehat{\mathbf{x}}\}, \quad \mathbf{x} \in cl(\mathcal{K}) \end{array} \tag{HCP}$$

and its Lagrangian dual problem

$$\begin{array}{ll} \inf\limits_{\mathbf{s}} & \langle \widehat{\mathbf{x}}, \mathbf{s} \rangle \\ \text{subject to} & \mathbf{s} \in \mathcal{L}^{\perp} + \{\widehat{\mathbf{s}}\}, \quad \mathbf{s} \in cl(\mathcal{K}^{\sharp}) \end{array} \tag{HCD}$$

where $\mathcal{L}^\perp$ denotes the orthogonal complement of $\mathcal{L}$ in $\mathbb{E}$ and $\mathcal{K}^\sharp$ denotes the (open) dual cone $\{\mathbf{s} \in \mathbb{E} : \langle \mathbf{x}, \mathbf{s} \rangle > 0 \ \forall \mathbf{x} \in cl(\mathcal{K}) \setminus \{\mathbf{0}\}\}$. Since $\mathcal{K}$ is pointed, the dual cone $\mathcal{K}^\sharp$ is nonempty.

2.1. **T-algebras.** We begin with some basics of T-algebras; see [17] and [2, Chapter 2] for more details.

Definition 2 (Matrix algebra)**.** A *matrix algebra* $\mathfrak{A}$ is a bi-graded algebra $\bigoplus_{i,j=1}^r \mathfrak{A}_{ij}$ over the reals satisfying

$$\mathfrak{A}_{ij}\mathfrak{A}_{kl} \subseteq \begin{cases} \mathfrak{A}_{il} & \text{if } j = k, \\ \{\mathbf{0}\} & \text{if } j \neq k. \end{cases}$$

The positive integer r is called the *rank* of the matrix algebra. For each $\mathbf{a} \in \mathfrak{A}$, we denote by $\mathbf{a}_{ij}$ its component in $\mathfrak{A}_{ij}$.

Definition 3 (Involution)**.** An *involution* $(\cdot)^*$ of a matrix algebra $\mathfrak{A}$ of rank r is a linear automorphism on $\mathfrak{A}$ that is

(1) involutory (i.e., $(\mathbf{a}^*)^* = \mathbf{a}$ for all $\mathbf{a} \in \mathfrak{A}$),
(2) anti-homomorphic (i.e., $(\mathbf{ab})^* = \mathbf{b}^*\mathbf{a}^*$ for all $\mathbf{a}, \mathbf{b} \in \mathfrak{A}$),

and further satisfies

$$\mathfrak{A}_{ij}^* = \mathfrak{A}_{ji} \ (1 \leq i, j \leq r).$$

Definition 4 (T-algebra)**.** A *T-algebra of rank r* is a matrix algebra $\mathfrak{A}$ of rank r with involution $(\cdot)^*$ satisfying the following seven axioms:

I. For each $i \in \{1, \dots, r\}$, the subalgebra $\mathfrak{A}_{ii}$ is isomorphic to the reals.
II. For each $\mathbf{a} \in \mathfrak{A}$ and each $i, j \in \{1, \dots, r\}$,

$$\mathbf{a}_{ji}\mathbf{e}_i = \mathbf{a}_{ji} \text{ and } \quad \mathbf{e}_i\mathbf{a}_{ij} = \mathbf{a}_{ij},$$

where $\mathbf{e}_i$ denotes the unit of $\mathfrak{A}_{ii}$.
III. For each $\mathbf{a}, \mathbf{b} \in \mathfrak{A}$ and each $i, j \in \{1, \dots, r\}$,

$$\mathbf{a}_{ij}\mathbf{b}_{ji} = \mathbf{b}_{ji}\mathbf{a}_{ij}.$$

IV. For each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{A}$ and each $i, j, k \in \{1, \dots, r\}$,

$$\mathbf{a}_{ij}(\mathbf{b}_{jk}\mathbf{c}_{ki}) = (\mathbf{a}_{ij}\mathbf{b}_{jk})\mathbf{c}_{ki}.$$

V. For each $\mathbf{a} \in \mathfrak{A}$ and each $i \in \{1, \dots, r\}$,

$$\mathbf{a}_{ij}^*\mathbf{a}_{ij} \geq 0$$

with equality when and only when $\mathbf{a}_{ij} = \mathbf{0}$.
VI. For each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{A}$ and each $i, j, k, l \in \{1, \dots, r\}$ with $i \leq j \leq k \leq l$,

$$\mathbf{a}_{ij}(\mathbf{b}_{jk}\mathbf{c}_{kl}) = (\mathbf{a}_{ij}\mathbf{b}_{jk})\mathbf{c}_{kl}.$$

VII. For each $\mathbf{a}, \mathbf{b} \in \mathfrak{A}$ and each $i, j, k, l \in \{1, \dots, r\}$ with $i \leq j \leq k$ and $l \leq k$,

$$\mathbf{a}_{ij}(\mathbf{b}_{jk}\mathbf{b}_{lk}^*) = (\mathbf{a}_{ij}\mathbf{b}_{jk})\mathbf{b}_{lk}^*.$$

Definition 5 (Inessential change of grading)**.** A *change in the grading* of a T-algebra $\mathfrak{A}$ is the replacement of each subspace $\mathfrak{A}_{ij}$ by $\mathfrak{A}_{\pi(i),\pi(j)}$ where π is a permutation of $\{1, \dots, r\}$. An *inessential change in the grading* of a T-algebra $\mathfrak{A}$ is one that preserves the subspace of upper triangular (or equivalently, lower triangular) elements. In another words, the permutation π in an inessential change satisfies

$$(i < j) \wedge (\pi(i) > \pi(j)) \implies \mathfrak{A}_{ij} = \{\mathbf{0}\}.$$

Definition 6 (Isomorphic T-algebras)**.** A T-algebra $\mathfrak{A}$ of rank r is said to be *isomorphic* to another T-algebra of the same rank if, after an inessential change of grading of $\mathfrak{A}$, there is an isomorphism of the bi-graded algebras with involution.

Definition 7 (Associated cone)**.** The *cone associated with a T-algebra* $\mathfrak{A}$ of rank r is

$$\{\mathbf{t}\mathbf{t}^* : \mathbf{t} \in \mathfrak{A},\ \mathbf{t}_{ij} = \mathbf{0}\ \forall 1 \le j < i \le r,\ \mathbf{t}_{ii} > 0\ \forall 1 \le i \le r\}.$$

It is denoted by $\mathcal{K}(\mathfrak{A})$.[1]

Example 1. *Every $(n+2)$-dimensional T-algebra of rank 2 is isomorphic to the T-algebra $\mathfrak{A} = \mathfrak{A}_{11} \oplus \mathfrak{A}_{12} \oplus \mathfrak{A}_{21} \oplus \mathfrak{A}_{22} = \mathbb{R} \oplus \mathbb{R}^n \oplus \mathbb{R}^n \oplus \mathbb{R}$ defined by*

$$(\mathbf{ab})_{ii} = \mathbf{a}_{ij}^T \mathbf{b}_{ji} \quad (i, j \in \{1, 2\},\ i \ne j)$$

By Corollary IV.1.5 of [7]*, the subset of elements of $\mathfrak{A}$ satisfying $\mathbf{a}^* = \mathbf{a}$, when equipped with the* Jordan product

$$(\mathbf{a}, \mathbf{b}) \mapsto \frac{\mathbf{ab} + \mathbf{ba}}{2},$$

is a Euclidean Jordan algebra of rank 2*.*

Example 2. *Let $\mathfrak{H}$ be a Euclidean Hurwitz algebra (i.e., algebra of real numbers, complex numbers, quaternions or octonions). We shall use $\Re(x)$ and $\bar{x}$ to denote, respectively, the real part and the conjugate of $x \in \mathfrak{H}$.*

Let $\mathfrak{A}_{ij} = \mathfrak{H}$ for each $i, j \in \{1, \ldots, r\}$ with $i \ne j$, $\mathfrak{A}_{ii} = \mathbb{R}$ for $i \in \{1, \ldots, r\}$, and define the matrix algebra $\mathfrak{A} = \bigoplus_{i,j=1}^r \mathfrak{A}_{ij}$ by

$$(\mathbf{ab})_{ij} = \begin{cases} \sum_{k=1}^r \Re(\mathbf{a}_{ik}\mathbf{b}_{kj}) & \text{if } i = j, \\ \sum_{k=1}^r \mathbf{a}_{ik}\mathbf{b}_{kj} & \text{if } i \ne j, \end{cases}$$

so that it satisfies Axioms I and II in the definition of a T-algebra. It is straightforward to check that the unary operator $(\cdot)^* : \mathfrak{A} \to \mathfrak{A}$ defined by

$$\mathbf{a}_{ij}^* = \bar{\mathbf{a}}_{ij}$$

is an involution. Moreover, Proposition V.1.2 of [7] *shows that Axioms III–V are also satisfied.*

For $r = 3$, Axiom VI holds since at least one of $\mathbf{a}_{ij}$, $\mathbf{b}_{jk}$ and $\mathbf{c}_{kl}$ is a real number. Finally, Axiom VII holds since Euclidean Hurwitz algebras are alternative (i.e., $x(xy) = (xx)y$ and $(yx)x = y(xx)$, or equivalently, the sub-algebra generated by any two elements is associative), and both x and $\bar{x}$ are in the sub-algebra generated by $x - \Re(x)$ for each $x \in \mathfrak{H}$. Hence $\mathfrak{A}$ is a T-algebra. By Theorem V.3.7 of [7]*, the subset of elements of $\mathfrak{A}$ satisfying $\mathbf{a}^* = \mathbf{a}$, when equipped with the Jordan product, is a Euclidean Jordan algebra of rank* 3*.*

For $r > 3$, suppose that $\mathfrak{H}$ is associative. Then $\mathfrak{A}$ is clearly a T-algebra. As before, by Theorem V.3.7 of [7]*, the subset of elements of $\mathfrak{A}$ satisfying $\mathbf{a}^* = \mathbf{a}$, when equipped with the Jordan product, is a Euclidean Jordan algebra of rank r.*

This example, together with the preceding example, covers all simple Euclidean Jordan algebras (see [7, Chapter V]*).*

[1] When we write $\mathbf{t}\mathbf{t}^* \in \mathcal{K}(\mathfrak{A})$, we always mean that $\mathbf{t}$ satisfies the conditions in the definition of $\mathcal{K}(\mathfrak{A})$. We shall see later that such $\mathbf{t}$ is uniquely determined by the product $\mathbf{t}\mathbf{t}^*$.

Theorem 1 (Characterization of homogeneous cones)**.** *A cone is homogeneous if and only if it is linearly isomorphic to the cone associated with some T-algebra. Moreover the T-algebra is uniquely determined, up to isomorphism, by the homogeneous cone.*

Proof. See Proposition 1 and Theorem 4 of [17]. □

Henceforth,

(1) $\mathfrak{A}$ shall denote a T-algebra of rank r with involution $(\cdot)^*$ such that $\mathcal{K}(\mathfrak{A}) = \mathcal{K}$;
(2) $\mathbf{e}_i$ shall denote the unit of the subalgebra $\mathfrak{A}_{ii}$ and ρ_i shall denote the linear function on $\mathfrak{A}$ satisfying
$$\mathbf{a}_{ii} = \rho_i(\mathbf{a})\mathbf{e}_i \ (\mathbf{a} \in \mathfrak{A})$$
for each $i \in \{1, \ldots, r\}$;
(3) $\mathbf{e}$ shall denote the element in $\mathfrak{A}$ satisfying
$$\mathbf{e}_{ii} = \mathbf{e}_i \ (1 \le i \le r) \text{ and } \mathbf{e}_{ij} = \mathbf{e}_{ji} = \mathbf{0} \ (1 \le i < j \le r);$$
(4) $\operatorname{tr}(\cdot)$ shall denote the linear function
$$\mathbf{a} \in \mathfrak{A} \mapsto \sum_{i=1}^{r} \rho_i(\mathbf{a});$$
(5) $\langle \cdot, \cdot \rangle$ shall denote the bilinear function
$$(\mathbf{a}, \mathbf{b}) \in \mathfrak{A} \oplus \mathfrak{A} \mapsto \operatorname{tr}(\mathbf{a}^*\mathbf{b});$$
(6) $\mathfrak{T}$ shall denote the subalgebra
$$\{\mathbf{a} \in \mathfrak{A} : \mathbf{a}_{ij} = \mathbf{0} \ (1 \le j < i \le r)\}$$
of upper triangular elements (so that $\mathcal{K} = \mathcal{K}(\mathfrak{A}) = \{\mathbf{t}\mathbf{t}^* : \mathbf{t} \in \mathfrak{T}, \ \rho_i(\mathbf{t}) > 0 \ \forall i\}$); and
(7) $\mathfrak{H}$ shall denote the subspace
$$\{\mathbf{a} \in \mathfrak{A} : \mathbf{a}_{ij} = \mathbf{a}_{ji}^* \ (1 \le j < i \le r)\}$$
of "hermitian" elements (so that the linear span of $\mathcal{K}$ is $\mathfrak{H}$).

Remark 1. For each $i \in \{1, \ldots, r\}$, since $(\cdot)^*$ is an involutory linear automorphism on $\mathfrak{A}_{ii}$, which is isomorphic to the reals, it must be the identity map. Thus the linear function $\operatorname{tr}(\cdot)$ is invariant under involution. Subsequently $\langle \cdot, \cdot \rangle$ is symmetric.

Remark 2. Axiom II is equivalent to $\mathbf{e}$ being the unit of the T-algebra $\mathfrak{A}$.

Remark 3. Axiom III is equivalent to $\operatorname{tr}(\mathbf{ab}) = \operatorname{tr}(\mathbf{ba})$ for all $\mathbf{a}, \mathbf{b} \in \mathfrak{A}$. This is further equivalent to $\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{a}^*, \mathbf{b}^* \rangle$.

Remark 4. Axiom V is equivalent to $\operatorname{tr}(\mathbf{aa}^*) \ge 0$ for all $\mathbf{a} \in \mathfrak{A}$ with equality when and only when $\mathbf{a} = \mathbf{0}$. This is further equivalent to $\langle \cdot, \cdot \rangle$ being positive definite. Thus $\langle \cdot, \cdot \rangle$ is an inner product of $\mathfrak{A}$. We shall denote by $\|\cdot\|$ the induced Euclidean norm $\mathbf{a} \mapsto \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$. We shall view $\mathcal{K}$ as a cone in the Euclidean space $(\mathfrak{H}, \langle \cdot, \cdot \rangle)$.

Remark 5. Axiom IV is equivalent to $\operatorname{tr}(\mathbf{a}(\mathbf{bc})) = \operatorname{tr}((\mathbf{ab})\mathbf{c})$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{A}$. This is further equivalent to $\langle \mathbf{ab}, \mathbf{c} \rangle = \langle \mathbf{b}, \mathbf{a}^*\mathbf{c} \rangle$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{A}$. Together with Axiom V, we have that the adjoint of the linear map $\mathbf{b} \in \mathfrak{A} \mapsto \mathbf{ab}$ under the inner product $\langle \cdot, \cdot \rangle$ is the linear map $\mathbf{b} \in \mathfrak{A} \mapsto \mathbf{a}^*\mathbf{b}$ for each $\mathbf{a} \in \mathfrak{A}$. Together with Remark 3, we deduce
$$\langle \mathbf{ab}, \mathbf{c} \rangle = \langle \mathbf{b}^*\mathbf{a}^*, \mathbf{c}^* \rangle = \langle \mathbf{a}^*, \mathbf{bc}^* \rangle = \langle \mathbf{a}, \mathbf{cb}^* \rangle \ \ (\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{A});$$

i.e., the adjoint of the linear map $\mathbf{b} \in \mathfrak{A} \mapsto \mathbf{ba}$ under the inner product $\langle \cdot, \cdot \rangle$ is the linear map $\mathbf{b} \in \mathfrak{A} \mapsto \mathbf{ba}^*$ for each $\mathbf{a} \in \mathfrak{A}$.

Remark 6. Axiom VI is equivalent to $\mathbf{t}(\mathbf{uw}) = (\mathbf{tu})\mathbf{w}$ for all $\mathbf{t}, \mathbf{u}, \mathbf{w} \in \mathfrak{T}$. Taking involution gives the equivalent statement: $\mathbf{t}(\mathbf{uw}) = (\mathbf{tu})\mathbf{w}$ for all $\mathbf{t}, \mathbf{u}, \mathbf{w} \in \mathfrak{T}^*$.

Remark 7. Axiom VII is equivalent to $\mathbf{t}(\mathbf{uu}^*) = (\mathbf{tu})\mathbf{u}^*$ for all $\mathbf{t}, \mathbf{u} \in \mathfrak{T}$. Taking involution gives the equivalent statement: $(\mathbf{u}^*\mathbf{u})\mathbf{t} = \mathbf{u}^*(\mathbf{ut})$ for all $\mathbf{t}, \mathbf{u} \in \mathfrak{T}^*$. Polarization gives the equivalent statements:

(1) $\mathbf{t}(\mathbf{uw}^*) + \mathbf{t}(\mathbf{wu}^*) = (\mathbf{tu})\mathbf{w}^* + (\mathbf{tw})\mathbf{u}^*$ for all $\mathbf{t}, \mathbf{u}, \mathbf{w} \in \mathfrak{T}$, and
(2) $(\mathbf{u}^*\mathbf{w})\mathbf{t} + (\mathbf{w}^*\mathbf{u})\mathbf{t} = \mathbf{u}^*(\mathbf{wt}) + \mathbf{w}^*(\mathbf{ut})$ for all $\mathbf{t}, \mathbf{u}, \mathbf{w} \in \mathfrak{T}^*$.

Theorem 2 (Sub-multiplicativity of Euclidean norm; c.f. Proposition 2.10 of [2].)**.** *The norm $\|\cdot\|$ is sub-multiplicative.*

Proof. We first show that for all $\mathbf{a}, \mathbf{b} \in \mathfrak{A}$,

$$\|\mathbf{a}_{ij}\mathbf{b}_{jk}\| \leq \|\mathbf{a}_{ij}\| \, \|\mathbf{b}_{jk}\| \;\; \forall 1 \leq i, j, k \leq r. \tag{2.1}$$

Since the norm is invariant under involution, it suffices to prove the above inequality for $i \leq k$. Thus we consider three cases

Case 1: $1 \leq i \leq j \leq k \leq r$. In this case, equality holds as

$$\begin{aligned} \|\mathbf{a}_{ij}\mathbf{b}_{jk}\|^2 &= \langle \mathbf{a}_{ij}\mathbf{b}_{jk}, \mathbf{a}_{ij}\mathbf{b}_{jk} \rangle \\ &= \left\langle \mathbf{a}_{ij}, (\mathbf{a}_{ij}\mathbf{b}_{jk})\mathbf{b}_{jk}^* \right\rangle && \text{(Remark 5)} \\ &= \left\langle \mathbf{a}_{ij}, \mathbf{a}_{ij}(\mathbf{b}_{jk}\mathbf{b}_{jk}^*) \right\rangle && \text{(Axiom VII)} \\ &= \left\langle \mathbf{a}_{ij}^*\mathbf{a}_{ij}, \mathbf{b}_{jk}\mathbf{b}_{jk}^* \right\rangle && \text{(Axiom IV)} \\ &= \rho_j(\mathbf{a}_{ij}^*\mathbf{a}_{ij})\rho_j(\mathbf{b}_{jk}\mathbf{b}_{jk}^*) \\ &= \|\mathbf{a}_{ij}\|^2 \, \|\mathbf{b}_{jk}\|^2 . \end{aligned}$$

Case 2: $1 \leq i \leq k < j \leq r$. Using the Cauchy-Schwarz inequality and the first case, we deduce

$$\begin{aligned} \|\mathbf{a}_{ij}\mathbf{b}_{jk}\|^2 &= \langle \mathbf{a}_{ij}\mathbf{b}_{jk}, \mathbf{a}_{ij}\mathbf{b}_{jk} \rangle \\ &= \left\langle \mathbf{a}_{ij}, (\mathbf{a}_{ij}\mathbf{b}_{jk})\mathbf{b}_{jk}^* \right\rangle && \text{(Remark 5)} \\ &\leq \|\mathbf{a}_{ij}\| \, \left\|(\mathbf{a}_{ij}\mathbf{b}_{jk})\mathbf{b}_{jk}^*\right\| \\ &= \|\mathbf{a}_{ij}\| \, \|\mathbf{a}_{ij}\mathbf{b}_{jk}\| \, \left\|\mathbf{b}_{jk}^*\right\| , \end{aligned}$$

which shows that $\|\mathbf{a}_{ij}\mathbf{b}_{jk}\| \leq \|\mathbf{a}_{ij}\| \, \left\|\mathbf{b}_{jk}^*\right\| = \|\mathbf{a}_{ij}\| \, \|\mathbf{b}_{jk}\|$.

Case 3: $1 \leq j < i \leq k \leq r$. Once again, we use the Cauchy-Schwarz inequality and the first case to deduce

$$\begin{aligned} \|\mathbf{a}_{ij}\mathbf{b}_{jk}\|^2 &= \langle \mathbf{a}_{ij}\mathbf{b}_{jk}, \mathbf{a}_{ij}\mathbf{b}_{jk} \rangle \\ &= \left\langle \mathbf{a}_{ij}^*(\mathbf{a}_{ij}\mathbf{b}_{jk}), \mathbf{b}_{jk} \right\rangle && \text{(Axiom IV)} \\ &\leq \left\|\mathbf{a}_{ij}^*(\mathbf{a}_{ij}\mathbf{b}_{jk})\right\| \, \|\mathbf{b}_{jk}\| \\ &= \left\|\mathbf{a}_{ij}^*\right\| \, \|\mathbf{a}_{ij}\mathbf{b}_{jk}\| \, \|\mathbf{b}_{jk}\| . \end{aligned}$$

By triangle inequality, we have

$$\|\mathbf{ab}\|^2 = \sum_{i,j=1}^{r} \left\| \sum_{k=1}^{r} \mathbf{a}_{ik}\mathbf{b}_{kj} \right\|^2 \leq \sum_{i,j=1}^{r} \left(\sum_{k=1}^{r} \|\mathbf{a}_{ik}\mathbf{b}_{kj}\| \right)^2 .$$

Using (2.1) and Cauchy's inequality, respectively, we further bound

$$\begin{aligned} \sum_{i,j=1}^{r} \left(\sum_{k=1}^{r} \|\mathbf{a}_{ik}\mathbf{b}_{kj}\| \right)^2 &\leq \sum_{i,j=1}^{r} \left(\sum_{k=1}^{r} \|\mathbf{a}_{ik}\| \, \|\mathbf{b}_{kj}\| \right)^2 \\ &\leq \sum_{i,j=1}^{r} \left(\sum_{k=1}^{r} \|\mathbf{a}_{ik}\|^2 \sum_{k=1}^{r} \|\mathbf{b}_{kj}\|^2 \right) \\ &= \|\mathbf{a}\|^2 \, \|\mathbf{b}\|^2 , \end{aligned}$$

hence proving the theorem. □

Let $\mathfrak{T}_*$ (resp., $\mathfrak{T}_+$ and $\mathfrak{T}_{++}$) denote the set of elements of $\mathfrak{T}$ with nonzero (resp., nonnegative and positive) diagonal components.

The homogeneous cone $\mathcal{K}$ can be described as $\{\mathbf{tt}^* : \mathbf{t} \in \mathfrak{T}_{++}\}$. The map $\mathbf{t} \mapsto \mathbf{tt}^*$ is continuous, hence the closure $cl(\mathcal{K})$ of the cone $\mathcal{K}$ contains the set $\{\mathbf{tt}^* : \mathbf{t} \in cl(\mathfrak{T}_{++}) = \mathfrak{T}_+\}$. Conversely every $\mathbf{x} \in cl(\mathcal{K})$ has a sequence $\{\mathbf{t}_k\} \subset \mathfrak{T}_{++}$ with $\mathbf{t}_k\mathbf{t}_k^* \to \mathbf{x}$. The sequence $\mathbf{t}_k$ is bounded since $\|\mathbf{t}_k\|^2 = \operatorname{tr}(\mathbf{t}_k\mathbf{t}_k^*)$ is bounded, hence it has a converging subsequence with limit $\mathbf{t} \in \mathfrak{T}_+$ satisfying $\mathbf{tt}^* = \mathbf{x}$. Consequently

$$cl(\mathcal{K}) = \{\mathbf{tt}^* : \mathbf{t} \in \mathfrak{T}_+\}.$$

Proposition 1. *The sets $\mathfrak{T}_\star$, $\mathfrak{T}_{++}$, $\mathfrak{T}_\star^*$ and $\mathfrak{T}_{++}^*$ are multiplicative groups.*

Proof. We shall prove that $\mathfrak{T}_\star$ and $\mathfrak{T}_{++}$ are multiplicative groups. Similarly result then holds for $\mathfrak{T}_\star^*$ and $\mathfrak{T}_{++}^*$ since involution is anti-homomorphic and $\mathbf{e}^* = \mathbf{e}$. Products of upper triangular elements are upper triangular with diagonal components the products of respective diagonal components. Let $\mathbf{t} \in \mathfrak{T}_\star$ be an arbitrary. Let $\mathbf{u}$ be the upper triangular element whose components are defined recursively by

$$\mathbf{u}_{i,i+j} = \begin{cases} \frac{1}{\mathbf{t}_{ii}} & \text{for } j = 0, \\ -\frac{1}{\mathbf{t}_{i+j,i+j}} \sum_{k=0}^{j-1} \mathbf{u}_{i,i+k}\mathbf{t}_{i+k,i+j} & \text{for } j = 1, \dots, r-i. \end{cases}$$

It is straightforward to check that $\mathbf{u} \in \mathfrak{T}_\star$ and $\mathbf{ut} = \mathbf{tu} = \mathbf{e}$. Moreover if $\mathbf{t} \in \mathfrak{T}_{++}$, then $\mathbf{u} \in \mathfrak{T}_{++}$. Axiom VI states that multiplication of upper triangular elements is associative. □

We shall denote by $\mathbf{t}^{-1}$ the multiplicative inverse of each element $\mathbf{t} \in \mathfrak{T}_\star$. Similarly for elements $\mathbf{l} \in \mathfrak{T}_\star^*$.

Involution is anti-homomorphic, hence the inverse of the involution of an element $\mathbf{t} \in \mathfrak{T}_\star$ is the involution of its inverse, which we shall denote by $\mathbf{t}^{-*}$. Similarly for elements $\mathbf{l} \in \mathfrak{T}_\star^*$.

2.2. **Quadratic representations.**

Definition 8 (Quadratic representation)**.** For each $\mathbf{a} \in \mathfrak{A}$, the *quadratic representation of* $\mathbf{a}$ is the linear operator

$$\mathbf{b} \in \mathfrak{H} \mapsto \frac{1}{2} \left(\mathbf{a}(\mathbf{ba}^*) + \mathbf{a}(\mathbf{ab}) - (\mathbf{aa})\mathbf{b} + (\mathbf{ab})\mathbf{a}^* + (\mathbf{ba}^*)\mathbf{a}^* - \mathbf{b}(\mathbf{a}^*\mathbf{a}^*) \right).$$

It is denoted by $\boldsymbol{\mathcal{Q}}_{\mathbf{a}}$.

Remark 8. Let $\mathfrak{J}$ be a Jordan algebra with binary operator $\circ$. The *quadratic representation* (see [7, Chapter II, §3]) of an element $\mathbf{x} \in \mathfrak{J}$ is defined to be the map

$$\mathbf{y} \in \mathfrak{J} \mapsto 2\mathbf{x} \circ (\mathbf{x} \circ \mathbf{y}) - (\mathbf{x} \circ \mathbf{x}) \circ \mathbf{y}.$$

Recall Examples 1 and 2, where we constructed T-algebras $\mathfrak{A}$ whose space of "hermitian" matrices are simple Euclidean Jordan algebras $\mathfrak{J}$ when equipped with the Jordan product

$$(\mathbf{a}, \mathbf{b}) \mapsto \frac{\mathbf{ab} + \mathbf{ba}}{2}.$$

Not surprisingly, the quadratic representations of "hermitian" elements of $\mathfrak{A}$ coincide with their respective quadratic representations as elements of $\mathfrak{J}$.

Proposition 2. *For each* $\mathbf{t} \in \mathfrak{T}$*, the map* $\boldsymbol{\mathcal{Q}}_{\mathbf{t}}$ *simplifies to*

$$\mathbf{a} \mapsto \mathbf{t}(\langle\!\langle \mathbf{a} \rangle\!\rangle\, \mathbf{t}^*) + (\mathbf{t}\, \langle\!\langle \mathbf{a} \rangle\!\rangle^*)\mathbf{t}^*.$$

If, in addition, $\mathbf{a} = \mathbf{uu}^*$ *for some* $\mathbf{u} \in \mathfrak{T}$*, then*

$$\boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{uu}^*) = (\mathbf{tu})(\mathbf{tu})^*.$$

Proof. Let $\mathbf{t}, \mathbf{u} \in \mathfrak{T}$ and $\mathbf{a} \in \mathfrak{H}$ be arbitrary. Let $\mathbf{w}$ denote the upper triangular element $\langle\!\langle \mathbf{a} \rangle\!\rangle^*$. The first equation follows from

$$\begin{aligned}
&\mathbf{t}(\mathbf{at}^*) + \mathbf{t}(\mathbf{ta}) - (\mathbf{tt})\mathbf{a} \\
&= \mathbf{t}(\mathbf{wt}^*) + \mathbf{t}(\mathbf{w}^*\mathbf{t}^*) + \mathbf{t}(\mathbf{tw}) + \mathbf{t}(\mathbf{tw}^*) - (\mathbf{tt})\mathbf{w} - (\mathbf{tt})\mathbf{w}^* \\
&= \mathbf{t}(\mathbf{wt}^*) + \mathbf{t}(\mathbf{w}^*\mathbf{t}^*) + \mathbf{t}(\mathbf{tw}^*) - (\mathbf{tt})\mathbf{w}^* && \text{(Axiom VI)} \\
&= (\mathbf{tw})\mathbf{t}^* + \mathbf{t}(\mathbf{w}^*\mathbf{t}^*) && \text{(Remark 7)}
\end{aligned}$$

while

$$\begin{aligned}
&\mathbf{t}((\mathbf{uu}^*)\mathbf{t}^*) + \mathbf{t}(\mathbf{t}(\mathbf{uu}^*)) - (\mathbf{tt})(\mathbf{uu}^*) \\
&= \mathbf{t}(\mathbf{u}(\mathbf{u}^*\mathbf{t}^*)) + \mathbf{t}((\mathbf{tu})\mathbf{u}^*) - ((\mathbf{tt})\mathbf{u})\mathbf{u}^* && \text{(Axiom VII)} \\
&= (\mathbf{tu})(\mathbf{tu})^* + (\mathbf{t}(\mathbf{tu}))\mathbf{u}^* - ((\mathbf{tt})\mathbf{u})\mathbf{u}^* && \text{(Remark 7)} \\
&= (\mathbf{tu})(\mathbf{tu})^* && \text{(Axiom VI)}
\end{aligned}$$

proves the second. □

Corollary 1. *The map* $\mathbf{t} \mapsto \boldsymbol{\mathcal{Q}}_{\mathbf{t}}$ *is a group homomorphism on* $\mathfrak{T}_{++}$.

Proof. Let $\mathbf{t}, \mathbf{u} \in \mathfrak{T}$ be arbitrary upper triangular elements with positive diagonal. By the preceding proposition, for any $\mathbf{w} \in \mathfrak{T}$,

$$\begin{aligned}
\boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\boldsymbol{\mathcal{Q}}_{\mathbf{u}}(\mathbf{ww}^*)) &= \boldsymbol{\mathcal{Q}}_{\mathbf{t}}((\mathbf{uw})(\mathbf{uw})^*) = (\mathbf{t}(\mathbf{uw}))(\mathbf{t}(\mathbf{uw}))^* \\
&= ((\mathbf{tu})\mathbf{w}))((\mathbf{tu})\mathbf{w}))^* \quad \text{(Axiom VI)} \\
&= \boldsymbol{\mathcal{Q}}_{\mathbf{tu}}(\mathbf{ww}^*),
\end{aligned}$$

shows that $\boldsymbol{\mathcal{Q}}_{\mathbf{t}}\boldsymbol{\mathcal{Q}}_{\mathbf{u}} = \boldsymbol{\mathcal{Q}}_{\mathbf{tu}}$ on $\mathcal{K}$. Since $\boldsymbol{\mathcal{Q}}_{\mathbf{t}}$ is linear for each $\mathbf{t}$, and the linear span of $\mathcal{K}$ is the space $\mathfrak{H}$, the corollary follows. □

The author [2] proved that the map $\mathbf{t} \mapsto \mathbf{tt}^* = \boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{e})$ is a bijection between $\mathfrak{T}_{++}$ and $\mathcal{K}(\mathfrak{A})$; see also Proposition 2 of [17, Chapter III]. The following proposition shows that it is in fact a diffeomorphism.

Proposition 3. *For each* $\mathbf{u} \in \mathfrak{T}_{++}$*, the map*

$$\mathbf{t} \in \mathfrak{T}_{++} \mapsto \boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{uu}^*)$$

is a diffeomorphism. Moreover its derivative at $\mathbf{t} \in \mathfrak{T}_{++}$ *is*

$$\mathbf{w} \in \mathfrak{T} \mapsto (\mathbf{tu})(\mathbf{wu})^* + (\mathbf{wu})(\mathbf{tu})^*$$

and the derivative of its inverse map at $\boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{uu}^*) \in \mathcal{K}$ *is*

$$\mathbf{a} \in \mathfrak{H} \mapsto \mathbf{tu} \left\langle\!\left\langle \boldsymbol{\mathcal{Q}}_{(\mathbf{tu})^{-1}}(\mathbf{a}) \right\rangle\!\right\rangle^* \mathbf{u}^{-1}.$$

Proof. By Proposition 2, the derivative of $\mathbf{t} \in \mathfrak{T}_{++} \mapsto \boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{uu}^*)$ is the map

$$\mathbf{w} \in \mathfrak{T} \mapsto (\mathbf{tu})(\mathbf{wu})^* + (\mathbf{wu})(\mathbf{tu})^*.$$

By Axiom VI and Proposition 2

$$\begin{aligned}(\mathbf{tu})(\mathbf{wu})^* + (\mathbf{wu})(\mathbf{tu})^* &= (\mathbf{tu})(((\mathbf{wu})^*(\mathbf{tu})^{-*})(\mathbf{tu})^*) + ((\mathbf{tu})((\mathbf{tu})^{-1}(\mathbf{wu})))(\mathbf{tu})^* \\ &= \boldsymbol{\mathcal{Q}}_{\mathbf{tu}}(((\mathbf{tu})^{-1}(\mathbf{wu}))_H),\end{aligned}$$

hence by Corollary 1 the derivative has inverse

$$\mathbf{a} \in \mathfrak{H} \mapsto \mathbf{tu} \left\langle\!\left\langle \boldsymbol{\mathcal{Q}}_{(\mathbf{tu})^{-1}}(\mathbf{a}) \right\rangle\!\right\rangle^* \mathbf{u}^{-1}.$$

The proposition follows from the Inverse Function Theorem. □

An alternative description of the cone $\mathcal{K}(\mathfrak{A})$ based on quadratic representations is

$$\mathcal{K}(\mathfrak{A}) = \{\boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{e}) : \mathbf{t} \in \mathfrak{T}_{++}\}.$$

Corollary 1 shows that we may replace $\mathbf{e}$ in the above description with any $\mathbf{x} \in \mathcal{K}(\mathfrak{A})$, since

$$\boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\mathbf{uu}^*) = \boldsymbol{\mathcal{Q}}_{\mathbf{t}}(\boldsymbol{\mathcal{Q}}_{\mathbf{u}}(\mathbf{e})) = \boldsymbol{\mathcal{Q}}_{\mathbf{tu}}(\mathbf{e})$$

for any $\mathbf{t}, \mathbf{u} \in \mathfrak{T}_{++}$.

The linear map $\boldsymbol{\mathcal{Q}}_{\mathbf{t}}$ and its inverse $\boldsymbol{\mathcal{Q}}_{\mathbf{t}^{-1}}$ take the cone $\mathcal{K}$ into itself, hence $\boldsymbol{\mathcal{Q}}_{\mathbf{t}}$ is an automorphism of the cone $\mathcal{K}$. Moreover $\boldsymbol{\mathcal{Q}}_{\mathbf{ut}^{-1}}$ maps $\mathbf{tt}^*$ to $\mathbf{uu}^*$ for any $\mathbf{t}, \mathbf{u} \in \mathfrak{T}_{++}$. This gives a quick proof that $\mathcal{K} = \{\mathbf{tt}^* : \mathbf{t} \in \mathfrak{T}^*_{++}\}$ is a homogeneous cone. Furthermore $\{\boldsymbol{\mathcal{Q}}_{\mathbf{t}} : \mathbf{t} \in \mathfrak{T}_{++}\}$ acts transitively on $\mathcal{K}$.[2]

2.3. **Dual homogeneous cones.** Using Axiom IV and Remark 5, it is straightforward to check that the adjoint map $\boldsymbol{\mathcal{Q}}^H_{\mathbf{a}}$ of the quadratic representation of an element $\mathbf{a} \in \mathfrak{A}$ is the quadratic representation $\boldsymbol{\mathcal{Q}}_{\mathbf{a}^*}$ of its involution.

In general, if $\boldsymbol{\mathcal{A}}$ is an automorphism of an open cone $\mathcal{C}$, then $\boldsymbol{\mathcal{A}}^H$ is an automorphism of $\mathcal{C}^\sharp$. Moreover if $\mathcal{G}$ is a group of transitive automorphisms of $\mathcal{C}$, then the group $\{\boldsymbol{\mathcal{A}}^H : \boldsymbol{\mathcal{A}} \in \mathcal{G}\}$ also acts transitively on $\mathcal{C}^\sharp$; see Proposition 9 of [17]. Recall that $\{\boldsymbol{\mathcal{Q}}_{\mathbf{t}} : \mathbf{t} \in \mathfrak{T}_{++}\}$ is a transitive subgroup of automorphisms of $\mathcal{K}$. Hence the dual cone $\mathcal{K}^\sharp$ is

$$\{\boldsymbol{\mathcal{Q}}_{\mathbf{t}^*}(\mathbf{x}) : \mathbf{t} \in \mathfrak{T}_{++}\} = \{\boldsymbol{\mathcal{Q}}_{\mathbf{l}}(\mathbf{x}) : \mathbf{l} \in \mathfrak{T}^*_{++}\},$$

where $\mathbf{x} \in \mathcal{K}^\sharp$ arbitrary. The identity $\mathbf{e}$ is a member of $\mathcal{K}^\sharp$ as

$$\langle \mathbf{e}, \mathbf{tt}^* \rangle = \|\mathbf{t}\|^2 > 0$$

for all $\mathbf{t} \in \mathfrak{T}_+ \setminus \{\mathbf{0}\}$. We thus proved the following proposition.

Proposition 4. *The dual cone* $\mathcal{K}^\sharp$ *of* $\mathcal{K}$ *is the homogeneous cone* $\{\mathbf{ll}^* : \mathbf{l} \in \mathfrak{T}^*_{++}\}$*. The group of automorphisms* $\{\boldsymbol{\mathcal{Q}}_{\mathbf{t}} : \mathbf{t} \in \mathfrak{T}^*_{++}\}$ *acts transitively on* $\mathcal{K}^\sharp$*.*

[2] It is in fact a *simply transitive* subgroup; i.e., for every $\mathbf{t}, \mathbf{u} \in \mathfrak{T}_{++}$, there exists a unique $\boldsymbol{\mathcal{Q}}_{\mathbf{w}}$ with $\mathbf{w} \in \mathfrak{T}_{++}$ that maps $\mathbf{tt}^*$ to $\mathbf{uu}^*$.

We now turn our attention to identifying the T-algebra $\mathfrak{B}$ with $\mathcal{K}(\mathfrak{B}) = \mathcal{K}^\sharp$.

It is easy to see that a matrix algebra $\mathfrak{B}$ with involution satisfying $\mathcal{K}(\mathfrak{B}) = \mathcal{K}^\sharp$ can be obtained from $\mathfrak{A}$ by the change in grading

$$\mathfrak{B}_{ij} = \mathfrak{A}_{r+1-i,r+1-j} \ (1 \le i, j \le r)$$

It was pointed out by Vinberg [17, Chapter III, §6] that the matrix algebra $\mathfrak{B}$ is also a T-algebra. Indeed Axioms I–V translates directly from $\mathfrak{A}$ to $\mathfrak{B}$, Axiom VI for $\mathfrak{B}$ is equivalent to that for $\mathfrak{A}$ after taking involution, and Axiom VII for $\mathfrak{B}$ is proved in the following proposition.[3]

Proposition 5. *For each* $\mathbf{t}, \mathbf{u} \in \mathfrak{T}$*,*

$$\mathbf{t}^*(\mathbf{u}^*\mathbf{u}) = (\mathbf{t}^*\mathbf{u}^*)\mathbf{u}.$$

Equivalently, taking involution and/or polarization gives

(1) $(\mathbf{u}\mathbf{u}^*)\mathbf{t}^* = \mathbf{u}(\mathbf{u}^*\mathbf{t}^*)$ *for all* $\mathbf{t}, \mathbf{u} \in \mathfrak{T}^*$*,*
(2) $\mathbf{t}^*(\mathbf{u}^*\mathbf{w}) + \mathbf{t}^*(\mathbf{w}^*\mathbf{u}) = (\mathbf{t}^*\mathbf{u}^*)\mathbf{w} + (\mathbf{t}^*\mathbf{w}^*)\mathbf{u}$ *for all* $\mathbf{t}, \mathbf{u}, \mathbf{w} \in \mathfrak{T}$*, and*
(3) $(\mathbf{u}\mathbf{w}^*)\mathbf{t}^* + (\mathbf{w}\mathbf{u}^*)\mathbf{t}^* = \mathbf{u}(\mathbf{w}^*\mathbf{t}^*) + \mathbf{w}(\mathbf{u}^*\mathbf{t}^*)$ *for all* $\mathbf{t}, \mathbf{u}, \mathbf{w} \in \mathfrak{T}^*$*.*

Proof. Let $\mathbf{t}, \mathbf{u} \in \mathfrak{T}$ be arbitrary. Let $\mathbf{a}$ denote the difference $\mathbf{t}^*(\mathbf{u}^*\mathbf{u}) - (\mathbf{t}^*\mathbf{u}^*)\mathbf{u}$. We shall prove the proposition by showing that $\|\mathbf{a}\| = 0$, whence $\mathbf{a} = \mathbf{0}$ by Axiom V. We evaluate $\|\mathbf{a}\|^2$ as the sum of $\langle \mathbf{w}, \mathbf{a} \rangle$ and $\langle \mathbf{w}^*, \mathbf{a} \rangle$, where $\mathbf{w}$ denotes the upper triangular element $\langle\!\langle \mathbf{a} \rangle\!\rangle^*$. Now

$$\begin{aligned} \langle \mathbf{w}, \mathbf{t}^*(\mathbf{u}^*\mathbf{u}) \rangle &= \langle \mathbf{u}(\mathbf{t}\mathbf{w}), \mathbf{u} \rangle && \text{(Axiom IV)} \\ &= \langle (\mathbf{u}\mathbf{t})\mathbf{w}, \mathbf{u} \rangle && \text{(Axiom VI)} \\ &= \langle \mathbf{w}, (\mathbf{t}^*\mathbf{u}^*)\mathbf{u} \rangle && \text{(Axiom IV)} \end{aligned}$$

implies $\langle \mathbf{w}, \mathbf{a} \rangle = 0$. Also

$$\begin{aligned} \langle \mathbf{u}, \mathbf{u}(\mathbf{w}\mathbf{t}^*) \rangle &= \langle \mathbf{w}^*(\mathbf{u}^*\mathbf{u}), \mathbf{t}^* \rangle && \text{(Axiom IV)} \\ &= \langle \mathbf{w}^*, \mathbf{t}^*(\mathbf{u}^*\mathbf{u}) \rangle && \text{(Remark 5)} \\ &= \langle \mathbf{u}(\mathbf{t}\mathbf{w}^*), \mathbf{u} \rangle && \text{(Axiom IV)} \end{aligned}$$

and

$$\begin{aligned} \langle \mathbf{u}, (\mathbf{u}\mathbf{w})\mathbf{t}^* \rangle &= \langle (\mathbf{w}^*\mathbf{u}^*)\mathbf{u}, \mathbf{t}^* \rangle && \text{(Axiom IV)} \\ &= \langle \mathbf{w}^*, (\mathbf{t}^*\mathbf{u}^*)\mathbf{u} \rangle && \text{(Remark 5)} \\ &= \langle (\mathbf{u}\mathbf{t})\mathbf{w}^*, \mathbf{u} \rangle, && \text{(Axiom IV)} \end{aligned}$$

whence

$$\begin{aligned} 2\,\langle \mathbf{w}^*, \mathbf{t}^*(\mathbf{u}^*\mathbf{u}) \rangle &= \langle \mathbf{u}(\mathbf{w}\mathbf{t}^*) + \mathbf{u}(\mathbf{t}\mathbf{w}^*), \mathbf{u} \rangle \\ &= \langle (\mathbf{u}\mathbf{w})\mathbf{t}^* + (\mathbf{u}\mathbf{t})\mathbf{w}^*, \mathbf{u} \rangle && \text{(Remark 7)} \\ &= 2\,\langle \mathbf{w}^*, (\mathbf{t}^*\mathbf{u}^*)\mathbf{u} \rangle \end{aligned}$$

shows that $\langle \mathbf{w}^*, \mathbf{a} \rangle = 0$. □

The following propositions are the counterparts of Propositions 2 and 3 for $\mathfrak{B}$, stated in terms of the quadratic representations in $\mathfrak{A}$ instead of those in $\mathfrak{B}$.

[3]It was noted by Vinberg that by observing that the subspace of upper triangular elements of $\mathfrak{B}$ is precisely $\mathfrak{T}^*$, both Axioms VI and VII for $\mathfrak{B}$ follows from those for $\mathfrak{A}$ by taking involution. While this is certainly true for Axiom VI, it does not hold for Axiom VII.

Proposition 6. *For each* $\mathbf{l} \in \mathfrak{T}^*$*, the map* $\boldsymbol{\mathcal{Q}}_{\mathbf{l}}$ *simplifies to*

$$\mathbf{a} \mapsto \mathbf{l}(\langle\!\langle \mathbf{a} \rangle\!\rangle^* \, \mathbf{l}^*) + (\mathbf{l} \, \langle\!\langle \mathbf{a} \rangle\!\rangle)\mathbf{l}^*.$$

If, in addition, $\mathbf{a} = \mathbf{k}\mathbf{k}^*$ *for some* $\mathbf{k} \in \mathfrak{T}^*$*, then*

$$\boldsymbol{\mathcal{Q}}_{\mathbf{l}}(\mathbf{k}\mathbf{k}^*) = (\mathbf{l}\mathbf{k})(\mathbf{l}\mathbf{k})^*.$$

Proposition 7. *For each* $\mathbf{k} \in \mathfrak{T}^*_{++}$*, the map*

$$\mathbf{l} \in \mathfrak{T}^*_{++} \mapsto \boldsymbol{\mathcal{Q}}_{\mathbf{l}}(\mathbf{k}\mathbf{k}^*)$$

is a diffeomorphism. Moreover its derivative at $\mathbf{l} \in \mathfrak{T}^*_{++}$ *is*

$$\mathbf{m} \in \mathfrak{T}^* \mapsto (\mathbf{l}\mathbf{k})(\mathbf{m}\mathbf{k})^* + (\mathbf{m}\mathbf{k})(\mathbf{l}\mathbf{k})^*$$

and the derivative of its inverse map at $\boldsymbol{\mathcal{Q}}_{\mathbf{l}}(\mathbf{k}\mathbf{k}^*) \in \mathcal{K}^\sharp$ *is*

$$\mathbf{a} \in \mathfrak{H} \mapsto \mathbf{l}\mathbf{k} \left\langle\!\!\left\langle \boldsymbol{\mathcal{Q}}_{(\mathbf{l}\mathbf{k})^{-1}}(\mathbf{a}) \right\rangle\!\!\right\rangle \mathbf{k}^{-1}.$$

3. Primal-Dual Central Paths

In this section, we define a primal-dual central path for the pair of primal-dual homogeneous cone programming problems HCP and HCD based on a pair of optimal logarithmically homogeneous self-concordant barriers.

3.1. **Optimal barriers.**

We begin by recalling the definition of ϑ-logarithmically homogeneous self-concordant barriers.

Definition 9 (Self-concordant barrier)**.** A function $f : \mathcal{C} \to \mathbb{R}$ on an open convex cone $\mathcal{C}$ is said to be a *(non-degenerate, strongly)* ϑ*-self-concordant barrier* (for $\mathcal{C}$) if it is a strictly convex, three-times continuously differentiable function that diverges to infinity as its argument approaches a point on the boundary $\partial\mathcal{C}$ of $\mathcal{C}$, and satisfies

$$f'''(\mathbf{x}; \mathbf{h}) \leq 2(f''(\mathbf{x}; \mathbf{h}))^{3/2}$$

and

$$f'(\mathbf{x}; \mathbf{h}) \leq (\vartheta f''(\mathbf{x}; \mathbf{h}))^{1/2}$$

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

Definition 10 (Logarithmically homogeneous function)**.** A function $f : \mathcal{C} \to \mathbb{R}$ on an open convex cone $\mathcal{C}$ is said to be ϑ*-logarithmically homogeneous* if for every $t \in \mathbb{R}_{++}$ and every $\mathbf{x} \in \mathcal{C}$, it holds

$$f(t\mathbf{x}) = f(\mathbf{x}) - \vartheta \log t.$$

When a self-concordant barrier on $\mathcal{C}$ is ϑ-logarithmically homogeneous, it has complexity parameter ϑ. We thus called such function a ϑ*-logarithmically homogeneous self-concordant barrier* on $\mathcal{C}$. Every open convex cone $\mathcal{C}$ admits logarithmically homogeneous self-concordant barriers; see [12, Section 2.5]. Among all logarithmically homogeneous self-concordant barriers for $\mathcal{C}$, a barrier with the least complexity parameter is called an *optimal barrier.*

Optimal barriers for a homogeneous cone $\mathcal{K}$ of rank r has complexity parameter r. Moreover the function

$$f : \mathbf{x} \in \mathcal{K} \mapsto -\sum_{i=1}^{r} \log(\rho_i(\mathbf{u_x})^2) \tag{3.1}$$

is a r-logarithmically homogeneous self-concordant barrier for $\mathcal{K}$, where $\mathbf{u_x}$ denotes the unique element in $\mathfrak{T}_{++}$ satisfying $\mathbf{u_x}\mathbf{u_x}^* = \mathbf{x}$; see [9]. To the best of the author's knowledge, this is the only known optimal barrier for the homogeneous cone. This barrier is a member of the family of logarithmically homogeneous self-concordant barriers

$$\left\{\mathbf{x} \in \mathcal{K} \mapsto -\sum_{i=1}^{r} w_i \log(\rho_i(\mathbf{u_x})^2) : w_i \geq 1\right\};$$

see [2, 5] for more details.

By writing the barrier f as the composition of $\mathbf{x} \in \mathcal{K} \mapsto \mathbf{u_x}$ and

$$\mathbf{t} \in \mathfrak{T}_{++} \mapsto -\sum_{i=1}^{r} \log(\rho_i(\mathbf{t})^2),$$

it is straightforward to compute the gradient of f using Proposition 3. This gives the gradient at each $\mathbf{x} \in \mathcal{K}$ as

$$\nabla f(\mathbf{x}) = \mathbf{u_x}^{-*}\mathbf{u_x}^{-1}.$$

A simple modification of the Fenchel conjugate function of f gives the function

$$f^\sharp : \mathbf{s} \in \mathcal{K}^\sharp \mapsto -\sum_{i=1}^{r} \log(\rho_i(\mathbf{l_s})^2)$$

on the dual cone, where $\mathbf{l_s}$ denotes the unique element in $\mathfrak{T}_{++}^*$ satisfying $\mathbf{l_s}\mathbf{l_s}^* = \mathbf{s}$. This is in fact the optimal barrier given by (3.1) if we replace $\mathcal{K}$ with the homogeneous cone $\mathcal{K}^\sharp$. The gradient of this barrier at each $\mathbf{s} \in \mathcal{K}^\sharp$ is

$$\nabla f^\sharp(\mathbf{s}) = \mathbf{l_s}^{-*}\mathbf{l_s}^{-1}.$$

3.2. **The central path.** The central path of the primal-dual pair of homogeneous cone programming problems (HCP) and (HCD) is the set $\{(\mathbf{x}(\mu), \mathbf{s}(\mu)) : \mu > 0\}$ of pairs of solutions to the primal and dual barrier problems

$$\inf_{\mathbf{x}} \left\{\langle \widehat{\mathbf{s}}, \mathbf{x}\rangle - \mu \sum_{i=1}^{r} \log(\rho_i(\mathbf{u_x})^2) : \mathbf{x} \in \mathcal{L} + \{\widehat{\mathbf{x}}\}\right\},$$

and

$$\inf_{\mathbf{s}} \left\{\langle \widehat{\mathbf{x}}, \mathbf{s}\rangle - \mu \sum_{i=1}^{r} \log(\rho_i(\mathbf{l_s})^2) : \mathbf{s} \in \mathcal{L}^\perp + \{\widehat{\mathbf{s}}\}\right\},$$

respectively. We deduce from the optimality conditions for the barrier problems that the pair $(\mathbf{x}(\mu), \mathbf{s}(\mu))$ solves the *central path equations*

$$\begin{aligned} \mathbf{x} \in \mathcal{L} + \{\widehat{\mathbf{x}}\}, \quad & \mathbf{x} \in \mathcal{K}, \\ \mathbf{s} \in \mathcal{L}^\perp + \{\widehat{\mathbf{s}}\}, \quad & \mathbf{s} \in \mathcal{K}^\sharp, \\ \boldsymbol{\mathcal{Q}}_{\mathbf{l_s}^*}(\mathbf{x}) = \mu \mathbf{e}. & \end{aligned} \tag{CP_μ}$$

3.3. **Tracing the central path.** We now consider the problem of estimating the pair $(\mathbf{x}(\mu_{++}), \mathbf{s}(\mu_{++}))$ that solves the central path equations

$$\begin{gathered} \mathbf{x}(\mu_{++}) \in \mathcal{L} + \{\widehat{\mathbf{x}}\}, \quad \mathbf{x}(\mu_{++}) \in \mathcal{K}, \\ \mathbf{s}(\mu_{++}) \in \mathcal{L}^{\perp} + \{\widehat{\mathbf{s}}\}, \quad \mathbf{s}(\mu_{++}) \in \mathcal{K}^{\sharp}, \\ \mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}(\mu_{++})}}(\mathbf{x}(\mu_{++})) = \mu_{++}\mathbf{e}, \end{gathered} \tag{$CP_{\mu_{++}}$}$$

where $\mu_{++} \in \mathbb{R}_+$ is given. Note that we allow μ_{++} to be zero, in which case we are solving for optimal solutions. Given a primal-dual strictly feasible pair, we measure its proximity to the pair $(\mathbf{x}(\mu), \mathbf{s}(\mu))$, where $\mu \in \mathbb{R}_{++}$, with the proximity measure $d : \mathcal{K} \oplus \mathcal{K}^{\sharp} \oplus \mathbb{R}_{++} \to \mathbb{R}_+$ defined by

$$(\mathbf{x}, \mathbf{s}; \mu) \mapsto \frac{1}{\mu} \left\| \mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}}}(\mathbf{x}) - \mu\mathbf{e} \right\|.$$

Suppose we have a pair of estimates $(\mathbf{x}_+, \mathbf{s}_+)$ satisfying $d(\mathbf{x}_+, \mathbf{s}_+; \mu_+) \le \beta$ for some $\mu_+ := \frac{1}{r} \langle \mathbf{x}_+, \mathbf{s}_+ \rangle > \mu_{++}$ and some $\beta < 1$. We shall estimate $(\mathbf{x}(\mu_{++}), \mathbf{s}(\mu_{++}))$ with Newton's Method. This calls for the linearization of (CP_μ).

Proposition 8. *Suppose* $\mathbf{x} \in \mathcal{K}$, $\mathbf{s} \in \mathcal{K}^{\sharp}$ *and* $\mathbf{a}, \mathbf{b} \in \mathfrak{H}$. *Then the directional derivative of*

$$(\mathbf{tt}^*, \mathbf{ll}^*) \in \mathcal{K} \oplus \mathcal{K}^{\sharp} \mapsto \mathcal{Q}_{\mathbf{l}^*}(\mathbf{tt}^*) = (\mathbf{l}^*\mathbf{t})(\mathbf{l}^*\mathbf{t})^*$$

at $(\mathbf{x}, \mathbf{s})$ *along* $(\mathbf{a}, \mathbf{b})$ *is*

$$\mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}}}(\mathbf{a}) + \left(\mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}}}(\mathbf{x}) \left\langle\!\left\langle \mathcal{Q}_{\mathbf{l}^{-1}_{\mathbf{s}}}(\mathbf{b}) \right\rangle\!\right\rangle\right)_H.$$

Proof. Write the map

$$(\mathbf{tt}^*, \mathbf{ll}^*) \in \mathcal{K} \oplus \mathcal{K}^{\sharp} \mapsto \mathcal{Q}_{\mathbf{l}^*}(\mathbf{tt}^*) = (\mathbf{l}^*\mathbf{t})(\mathbf{l}^*\mathbf{t})^*$$

as the composition of

$$(\mathbf{tt}^*, \mathbf{ll}^*) \in \mathcal{K} \oplus \mathcal{K}^{\sharp} \mapsto (\mathbf{tt}^*, \mathbf{l}^*) \quad \text{and} \quad (\mathbf{tt}^*, \mathbf{l}^*) \in \mathcal{K} \oplus \mathfrak{T}_{++} \mapsto \mathcal{Q}_{\mathbf{l}^*}(\mathbf{tt}^*).$$

By Proposition 7 the derivative of the first map at $(\mathbf{tt}^*, \mathbf{ll}^*)$ is

$$(\mathbf{c}, \mathbf{d}) \in \mathfrak{H} \oplus \mathfrak{H} \mapsto (\mathbf{c}, \langle\!\langle \mathcal{Q}_{\mathbf{l}^{-1}}(\mathbf{d}) \rangle\!\rangle^* \mathbf{l}^*)$$

and by Proposition 3, the derivative of the second map at $(\mathbf{tt}^*, \mathbf{l}^*)$ is

$$(\mathbf{c}, \mathbf{k}^*) \in \mathfrak{H} \oplus \mathfrak{T} \mapsto \mathcal{Q}_{\mathbf{l}^*}(\mathbf{c}) + ((\mathbf{l}^*\mathbf{t})(\mathbf{k}^*\mathbf{t})^*)_H.$$

Hence Chain Rule gives the required directional derivative as

$$\mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}}}(\mathbf{a}) + ((\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})(\left\langle\!\left\langle \mathcal{Q}_{\mathbf{l}^{-1}_{\mathbf{s}}}(\mathbf{b}) \right\rangle\!\right\rangle^* \mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})^*)_H = \mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}}}(\mathbf{a}) + ((\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})((\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})^* \left\langle\!\left\langle \mathcal{Q}_{\mathbf{l}^{-1}_{\mathbf{s}}}(\mathbf{b}) \right\rangle\!\right\rangle))_H$$

By Axiom VII,

$$(\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})((\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})^* \left\langle\!\left\langle \mathcal{Q}_{\mathbf{l}^{-1}_{\mathbf{s}}}(\mathbf{b}) \right\rangle\!\right\rangle) = ((\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})(\mathbf{l}^*_{\mathbf{s}}\mathbf{u}_{\mathbf{x}})^*) \left\langle\!\left\langle \mathcal{Q}_{\mathbf{l}^{-1}_{\mathbf{s}}}(\mathbf{b}) \right\rangle\!\right\rangle,$$

hence the proposition follows from Proposition 2. □

The above proposition gives the linearization of (CP_μ) at $(\mathbf{x}_+, \mathbf{s}_+)$ as

$$\begin{gathered} \mathbf{\Delta}_{\mathbf{x}} \in \mathcal{L}, \quad \mathbf{\Delta}_{\mathbf{s}} \in \mathcal{L}^{\perp}, \\ \mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}_+}}(\mathbf{\Delta}_{\mathbf{x}}) + \left(\mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}_+}}(\mathbf{x}_+) \left\langle\!\left\langle \mathcal{Q}_{\mathbf{l}^{-1}_{\mathbf{s}_+}}(\mathbf{\Delta}_{\mathbf{s}}) \right\rangle\!\right\rangle\right)_H = \mu_{++}\mathbf{e} - \mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}_+}}(\mathbf{x}_+). \end{gathered} \tag{3.2}$$

One way to ensure that $(\mathbf{x}_+ + \mathbf{\Delta}_{\mathbf{x}}, \mathbf{s}_+ + \mathbf{\Delta}_{\mathbf{s}})$ is a good estimate of $(\mathbf{x}(\mu_{++}), \mathbf{s}(\mu_{++}))$ is to get a good upper bound on the sizes of $\mathbf{\Delta}_{\mathbf{x}}$ and $\mathbf{\Delta}_{\mathbf{s}}$.

The analysis can be greatly simplified by applying the primal-dual transformations

$$(\mathbf{x},\mathbf{s}) \mapsto (\boldsymbol{\mathcal{Q}}_{\mathbf{l}^*_{\mathbf{s}_+}}(\mathbf{x}), \boldsymbol{\mathcal{Q}}_{\mathbf{l}^{-1}_{\mathbf{s}_+}}(\mathbf{s})).$$

Let $\mathbf{v}$, $(\boldsymbol{\Delta}_{\widetilde{\mathbf{x}}}, \boldsymbol{\Delta}_{\widetilde{\mathbf{s}}})$ and $\widetilde{\mathcal{L}}$ denote, respectively, the images of $\mathbf{x}_+$, $(\boldsymbol{\Delta}_{\mathbf{x}}, \boldsymbol{\Delta}_{\mathbf{s}})$ and $\mathcal{L}$ under these transformations. Then $(\boldsymbol{\Delta}_{\widetilde{\mathbf{x}}}, \boldsymbol{\Delta}_{\widetilde{\mathbf{s}}})$ solves

$$\begin{aligned} \boldsymbol{\Delta}_{\widetilde{\mathbf{x}}} \in \widetilde{\mathcal{L}}, \quad \boldsymbol{\Delta}_{\widetilde{\mathbf{s}}} \in \widetilde{\mathcal{L}}^{\perp}, \\ \boldsymbol{\Delta}_{\widetilde{\mathbf{x}}} + (\mathbf{v}\,\langle\!\langle \boldsymbol{\Delta}_{\widetilde{\mathbf{s}}} \rangle\!\rangle)_H = \mu_{++}\mathbf{e} - \mathbf{v}. \end{aligned} \tag{3.3}$$

Note that the above linear system has a unique pair of solutions if and only if the same holds for (3.2). Since the linear system is square, the following proposition shows that this holds when $\beta < 1/\sqrt{2}$.

Proposition 9. *Suppose* $\mathbf{v} \in \mathcal{K}$ *satisfies*

$$\|\mathbf{v} - \mu\mathbf{e}\| \le \beta\mu \tag{3.4}$$

for some $\beta \in (0, 1/\sqrt{2})$ *and some* $\mu > 0$*. If* $(\boldsymbol{\Delta}_{\mathbf{x}}, \boldsymbol{\Delta}_{\mathbf{s}})$ *satisfies* $\langle \boldsymbol{\Delta}_{\mathbf{x}}, \boldsymbol{\Delta}_{\mathbf{s}} \rangle \ge 0$ *and*

$$\boldsymbol{\Delta}_{\mathbf{x}} + (\mathbf{v}\,\langle\!\langle \boldsymbol{\Delta}_{\mathbf{s}} \rangle\!\rangle)_H = \mathbf{h} \tag{3.5}$$

for some $\mathbf{h} \in \mathfrak{H}$*, then*

$$\max\{\|\boldsymbol{\Delta}_{\mathbf{x}}\|, \mu\|\boldsymbol{\Delta}_{\mathbf{s}}\|\} \le \frac{\|\mathbf{h}\|}{1-\sqrt{2}\beta}$$

Proof. If $\langle \boldsymbol{\Delta}_{\mathbf{x}}, \boldsymbol{\Delta}_{\mathbf{s}} \rangle \ge 0$, then

$$\max\{\|\boldsymbol{\Delta}_{\mathbf{x}}\|, \mu\|\boldsymbol{\Delta}_{\mathbf{s}}\|\} \le \sqrt{\|\boldsymbol{\Delta}_{\mathbf{x}}\|^2 + \mu\|\boldsymbol{\Delta}_{\mathbf{s}}\|^2} \le \|\boldsymbol{\Delta}_{\mathbf{x}} + \mu\boldsymbol{\Delta}_{\mathbf{s}}\|.$$

By triangle inequality, hypotheses (3.4) and (3.5), and the sub-multiplicativity of the norm $\|\cdot\|$, we have

$$\begin{aligned} \|\boldsymbol{\Delta}_{\mathbf{x}} + \mu\boldsymbol{\Delta}_{\mathbf{s}}\| &\le \|\boldsymbol{\Delta}_{\mathbf{x}} + (\mathbf{v}\,\langle\!\langle \boldsymbol{\Delta}_{\mathbf{s}} \rangle\!\rangle)_H\| + \|((\mathbf{v}-\mu\mathbf{e})\,\langle\!\langle \boldsymbol{\Delta}_{\mathbf{s}} \rangle\!\rangle)_H\| \\ &\le \|\mathbf{h}\| + 2\,\|(\mathbf{v}-\mu\mathbf{e})\,\langle\!\langle \boldsymbol{\Delta}_{\mathbf{s}} \rangle\!\rangle\| \\ &\le \|\mathbf{h}\| + 2\beta\mu\,\|\langle\!\langle \boldsymbol{\Delta}_{\mathbf{s}} \rangle\!\rangle\| \\ &\le \|\mathbf{h}\| + \sqrt{2}\beta\mu\,\|\boldsymbol{\Delta}_{\mathbf{s}}\| \\ &\le \|\mathbf{h}\| + \sqrt{2}\beta \max\{\|\boldsymbol{\Delta}_{\mathbf{x}}\|, \mu\|\boldsymbol{\Delta}_{\mathbf{s}}\|\}. \end{aligned}$$

Consequently under the hypotheses of the proposition,

$$(1-\sqrt{2}\beta)\max\{\|\boldsymbol{\Delta}_{\mathbf{x}}\|, \mu\|\boldsymbol{\Delta}_{\mathbf{s}}\|\} \le \|\mathbf{h}\|$$

as required. □

Let $\mathbf{x}_\alpha$ and $\mathbf{s}_\alpha$ denote, respectively, $\mathbf{x}_+ + \alpha\boldsymbol{\Delta}_{\mathbf{x}}$ and $\mathbf{s}_+ + \alpha\boldsymbol{\Delta}_{\mathbf{s}}$. Similarly, let $\widetilde{\mathbf{x}}_\alpha$ and $\widetilde{\mathbf{s}}_\alpha$ denote, respectively, $\mathbf{v} + \alpha\boldsymbol{\Delta}_{\widetilde{\mathbf{x}}} = \boldsymbol{\mathcal{Q}}_{\mathbf{l}^*_{\mathbf{s}_+}}(\mathbf{x}_\alpha)$ and $\mathbf{e} + \alpha\boldsymbol{\Delta}_{\widetilde{\mathbf{s}}} = \boldsymbol{\mathcal{Q}}_{\mathbf{l}^{-1}_{\mathbf{s}_+}}(\mathbf{s}_\alpha)$. Let μ_α denote the scaled duality gap $\frac{1}{r}\langle \mathbf{x}_\alpha, \mathbf{s}_\alpha \rangle = \frac{1}{r}\langle \widetilde{\mathbf{x}}_\alpha, \widetilde{\mathbf{s}}_\alpha \rangle$. It follows straightforwardly from the linearization (3.2) that

$$\mu_\alpha = (1-\alpha+\alpha\sigma)\mu_+,$$

where $\sigma = \mu_{++}/\mu_+ < 1$.

We shall give an upper bound on $d(\mathbf{x}_\alpha, \mathbf{s}_\alpha; \mu_\alpha)$ using the following lemma, which generalizes the local Lipschitz constant of Cholesky factorization of real symmetric matrices near the identity matrix.

Lemma 1 (c.f. Lemma 11 of [6]). *If* $\mathbf{h} \in \mathfrak{H}$ *satisfies* $\|\mathbf{h}\| \le 1/2$, *then*

$$\|\mathbf{l}_{\mathbf{e}+\mathbf{h}} - \mathbf{e}\| \le \sqrt{2}\,\|\mathbf{h}\|.$$

Proof. Let $\boldsymbol{\Delta}_{\mathbf{l}}(t)$ denote the lower triangular element $\mathbf{l}_{\mathbf{e}+t\mathbf{h}} - \mathbf{e}$. Note that for all $t \in \mathbb{R}$,

$$\boldsymbol{\Delta}_{\mathbf{l}}(t)_H + \boldsymbol{\Delta}_{\mathbf{l}}(t)\boldsymbol{\Delta}_{\mathbf{l}}(t)^* = t\mathbf{h}.$$

For $t \in [0,1]$, we have

$$\begin{aligned} t\,\|\mathbf{h}\| &= \|\boldsymbol{\Delta}_{\mathbf{l}}(t)_H + \boldsymbol{\Delta}_{\mathbf{l}}(t)\boldsymbol{\Delta}_{\mathbf{l}}(t)^*\| \\ &\ge \|\boldsymbol{\Delta}_{\mathbf{l}}(t)_H\| - \|\boldsymbol{\Delta}_{\mathbf{l}}(t)\boldsymbol{\Delta}_{\mathbf{l}}(t)^*\| \\ &\ge \sqrt{2}\,\|\boldsymbol{\Delta}_{\mathbf{l}}(t)\| - \|\boldsymbol{\Delta}_{\mathbf{l}}(t)\|^2. \end{aligned} \tag{3.6}$$

Solving this quadratic in $\|\boldsymbol{\Delta}_{\mathbf{l}}(t)\|$ gives

$$\|\boldsymbol{\Delta}_{\mathbf{l}}(t)\| \le \frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - t\,\|\mathbf{h}\|} \quad \text{or} \quad \|\boldsymbol{\Delta}_{\mathbf{l}}(t)\| \ge \frac{1}{\sqrt{2}} + \sqrt{\frac{1}{2} - t\,\|\mathbf{h}\|}.$$

Since $\boldsymbol{\Delta}_{\mathbf{l}}(t)$, whence $\|\boldsymbol{\Delta}_{\mathbf{l}}(t)\|$, is continuous in t, it follows that

$$\|\boldsymbol{\Delta}_{\mathbf{l}}(t)\| \le \frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - t\,\|\mathbf{h}\|}$$

whenever $t\,\|\mathbf{h}\| \le 1/2$. Under the hypothesis $\|\mathbf{h}\| \le 1/2$, this indeed hold for $t = 1$, thus

$$\|\mathbf{l}_{\mathbf{e}+\mathbf{h}} - \mathbf{e}\| \le \frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - \|\mathbf{h}\|} \le \frac{1}{\sqrt{2}}.$$

Finally, applying this upper bound in (3.6) with $t = 1$ gives

$$\|\mathbf{h}\| \ge \sqrt{2}\,\|\mathbf{l}_{\mathbf{e}+\mathbf{h}} - \mathbf{e}\| - \frac{1}{\sqrt{2}}\,\|\mathbf{l}_{\mathbf{e}+\mathbf{h}} - \mathbf{e}\| = \frac{1}{\sqrt{2}}\,\|\mathbf{l}_{\mathbf{e}+\mathbf{h}} - \mathbf{e}\|$$

as required. □

Lemma 2. *Suppose* $\beta < 1/\sqrt{2}$ *and* $\chi = \beta + (1-\sigma)\sqrt{r}$. *Then* $\mathbf{s}_\alpha \in \mathcal{K}^\sharp$ *and*

$$\mu_+^{-1}\left\|\mathcal{Q}_{\mathbf{l}^*_{\mathbf{s}_\alpha}}(\mathbf{x}_\alpha) - \mu_\alpha\mathbf{e}\right\| \le (1-\alpha)\beta + \alpha^2\frac{\chi^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \alpha^3\frac{3\chi^3}{(1-\sqrt{2}\beta)^3} \tag{3.7}$$

whenever $0 \le \alpha \le \min\{1, (1-\sqrt{2}\beta)/(2\chi)\}$.

Moreover, if $\beta \le \overline{\beta} < 1$ *with* $\sigma(\overline{\beta} - \beta) \ne 0$, *then there exists* $\overline{\alpha} > 0$ *such that*

(i) *the cubic polynomial*

$$p : \alpha \mapsto (1-\alpha)\beta + \alpha^2\frac{\chi^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \alpha^3\frac{3\chi^3}{(1-\sqrt{2}\beta)^3} - (1-\alpha+\sigma\alpha)\overline{\beta} \tag{3.8}$$

satisfies $p(\alpha) \le 0$ *for all* $\alpha \in [0, \overline{\alpha}]$; *and*

(ii) *for all* $0 \le \alpha \le \min\{1, \overline{\alpha}\}$, *the pair* $(\mathbf{x}_\alpha, \mathbf{s}_\alpha)$ *lies in* $\mathcal{K} \oplus \mathcal{K}^\sharp$ *and*

$$d(\mathbf{x}_\alpha, \mathbf{s}_\alpha; \mu_\alpha) \le \frac{1}{1-\alpha+\alpha\sigma}\left((1-\alpha)\beta + \alpha^2\frac{\chi^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \alpha^3\frac{3\chi^3}{(1-\sqrt{2}\beta)^3}\right) \le \overline{\beta} \tag{3.9}$$

Proof. According to Proposition 9, if $\beta < 1/\sqrt{2}$, then

$$\max\{\|\boldsymbol{\Delta}_{\tilde{\mathbf{x}}}\|, \mu_+\|\boldsymbol{\Delta}_{\tilde{\mathbf{s}}}\|\} \le \frac{\|\mu_{++}\mathbf{e} - \mathbf{v}\|}{1-\sqrt{2}\beta} \le \mu_+\frac{\beta + (1-\sigma)\sqrt{r}}{1-\sqrt{2}\beta} \tag{3.10}$$

by triangle inequality. Thus for every $\mathbf{u} \in \mathfrak{T}_+ \setminus \{\mathbf{0}\}$,

$$\langle \widetilde{\mathbf{s}}_\alpha, \mathbf{u}\mathbf{u}^* \rangle = \langle \mathbf{e}, \mathbf{u}\mathbf{u}^* \rangle + \alpha \langle \mathbf{\Delta}_{\widetilde{\mathbf{s}}}, \mathbf{u}\mathbf{u}^* \rangle \geq \|\mathbf{u}\|^2 - |\alpha| \, \|\mathbf{\Delta}_{\widetilde{\mathbf{s}}}\| \, \|\mathbf{u}\|^2 = (1 - |\alpha| \, \|\mathbf{\Delta}_{\widetilde{\mathbf{s}}}\|) \, \|\mathbf{u}\|^2 > 0$$

whenever $|\alpha| < (1 - \sqrt{2}\beta)/\chi$. Thus $\widetilde{\mathbf{s}}_\alpha \in \mathcal{K}^\sharp$, whence $\mathbf{s}_\alpha = \mathcal{Q}_{\mathbf{l}_{\mathbf{s}_+}}(\widetilde{\mathbf{s}}_\alpha) \in \mathcal{K}^\sharp$, for every $\alpha \in [0, (1 - \sqrt{2}\beta)/(2\chi)]$.

Note that whenever $\mathbf{s}_\alpha \in \mathcal{K}^\sharp$, we have

$$\widetilde{\mathbf{s}}_\alpha = \mathcal{Q}_{\mathbf{l}_{\mathbf{s}_+}^{-1}}(\mathbf{s}_\alpha) = (\mathbf{l}_{\mathbf{s}_+}^{-1}\mathbf{l}_{\mathbf{s}_\alpha})(\mathbf{l}_{\mathbf{s}_+}^{-1}\mathbf{l}_{\mathbf{s}_\alpha})^*,$$

hence $\mathbf{l}_{\widetilde{\mathbf{s}}_\alpha} = \mathbf{l}_{\mathbf{s}_+}^{-1}\mathbf{l}_{\mathbf{s}_\alpha}$. In this case we deduce, using Corollary 1, that

$$\mathcal{Q}_{\mathbf{l}_{\widetilde{\mathbf{s}}_\alpha}^*}(\widetilde{\mathbf{x}}_\alpha) = \mathcal{Q}_{\mathbf{l}_{\mathbf{s}_\alpha}^*\mathbf{l}_{\mathbf{s}_+}^{-*}}(\mathcal{Q}_{\mathbf{l}_{\mathbf{s}_+}^*}(\mathbf{x}_\alpha)) = \mathcal{Q}_{\mathbf{l}_{\mathbf{s}_\alpha}^*}(\mathbf{x}_\alpha).$$

Thus we may equivalently prove (3.7) for the pair $(\widetilde{\mathbf{x}}_\alpha, \widetilde{\mathbf{s}}_\alpha)$ instead.

Let $\mathbf{l}_\alpha$ denote the lower triangular element $\mathbf{l}_{\widetilde{\mathbf{s}}_\alpha} - \mathbf{e} = \mathbf{l}_{\mathbf{e}+\alpha\mathbf{\Delta}_{\widetilde{\mathbf{s}}}} - \mathbf{e}$, which is well-defined when $\alpha \in [0, (1 - \sqrt{2}\beta)/(2\chi)]$. By Proposition 2,

$$\begin{aligned}
\mathcal{Q}_{\mathbf{l}_{\widetilde{\mathbf{s}}_\alpha}^*}(\widetilde{\mathbf{x}}_\alpha) &= \mathcal{Q}_{\mathbf{e}+\mathbf{l}_\alpha^*}(\widetilde{\mathbf{x}}_\alpha) \\
&= ((\mathbf{e} + \mathbf{l}_\alpha^*)\,(\langle\!\langle \widetilde{\mathbf{x}}_\alpha \rangle\!\rangle\,(\mathbf{e} + \mathbf{l}_\alpha)))_H \\
&= (\langle\!\langle \widetilde{\mathbf{x}}_\alpha \rangle\!\rangle + \langle\!\langle \widetilde{\mathbf{x}}_\alpha \rangle\!\rangle\,\mathbf{l}_\alpha + \mathbf{l}_\alpha^*\,\langle\!\langle \widetilde{\mathbf{x}}_\alpha \rangle\!\rangle + \mathbf{l}_\alpha^*\,(\langle\!\langle \widetilde{\mathbf{x}}_\alpha \rangle\!\rangle\,\mathbf{l}_\alpha))_H \\
&= \widetilde{\mathbf{x}}_\alpha + (\widetilde{\mathbf{x}}_\alpha\mathbf{l}_\alpha)_H + \mathcal{Q}_{\mathbf{l}_\alpha^*}(\widetilde{\mathbf{x}}_\alpha) \\
&= \mathbf{v} + \alpha\mathbf{\Delta}_{\widetilde{\mathbf{x}}} + (\mathbf{v}\mathbf{l}_\alpha)_H + \alpha(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}\mathbf{l}_\alpha)_H + \mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{v}) + \alpha\mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}),
\end{aligned}$$

whence the difference $(\mathcal{Q}_{\mathbf{l}_{\widetilde{\mathbf{s}}_\alpha}^*}(\widetilde{\mathbf{x}}_\alpha) - \mu_\alpha\mathbf{e})$ is

$$\mathbf{v} - (1-\alpha)\mu_+\mathbf{e} - \alpha\sigma\mu_+\mathbf{e} + \alpha\mathbf{\Delta}_{\widetilde{\mathbf{x}}} + (\mathbf{v}\mathbf{l}_\alpha)_H + \alpha(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}\mathbf{l}_\alpha)_H + \mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{v}) + \alpha\mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}).$$

Using (3.3), we re-express this as

$$\begin{aligned}
&(1-\alpha)(\mathbf{v} - \mu_+\mathbf{e}) + (\mathbf{v}(\mathbf{l}_\alpha - \alpha\,\langle\!\langle \mathbf{\Delta}_{\widetilde{\mathbf{s}}} \rangle\!\rangle))_H + \alpha(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}\mathbf{l}_\alpha)_H \\
&\ + \mu_+\mathbf{l}_\alpha^*\mathbf{l}_\alpha + \mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{v} - \mu_+\mathbf{e}) + \alpha\mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}) \\
&= (1-\alpha)(\mathbf{v} - \mu_+\mathbf{e}) - \mu_+\mathbf{l}_\alpha\mathbf{l}_\alpha^* - ((\mathbf{v} - \mu_+\mathbf{e})\,\langle\!\langle \mathbf{l}_\alpha\mathbf{l}_\alpha^* \rangle\!\rangle)_H + \alpha(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}\mathbf{l}_\alpha)_H \\
&\quad + \mu_+\mathbf{l}_\alpha^*\mathbf{l}_\alpha + \mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{v} - \mu_+\mathbf{e}) + \alpha\mathcal{Q}_{\mathbf{l}_\alpha^*}(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}).
\end{aligned}$$

Using the sub-multiplicativity of $\|\cdot\|$, Lemma 1, and (3.10), we bound for each $\alpha \in [0, (1 - \sqrt{2}\beta)/(2\chi)]$,

$$\|\mathbf{l}_\alpha\mathbf{l}_\alpha^*\|, \|\mathbf{l}_\alpha^*\mathbf{l}_\alpha\| \leq \|\mathbf{l}_\alpha\|^2 \leq 2\alpha^2\,\|\mathbf{\Delta}_{\widetilde{\mathbf{s}}}\|^2 \leq 2\alpha^2\frac{\chi^2}{(1-\sqrt{2}\beta)^2},$$

$$\begin{aligned}
\|((\mathbf{v} - \mu_+\mathbf{e})\,\langle\!\langle \mathbf{l}_\alpha\mathbf{l}_\alpha^* \rangle\!\rangle)_H\| &\leq 2\,\|\mathbf{v} - \mu_+\mathbf{e}\|\,\|\langle\!\langle \mathbf{l}_\alpha\mathbf{l}_\alpha^* \rangle\!\rangle\| \\
&\leq \sqrt{2}\,\|\mathbf{v} - \mu_+\mathbf{e}\|\,\|\mathbf{l}_\alpha\mathbf{l}_\alpha^*\| \\
&\leq 2\sqrt{2}\alpha^2\mu_+\beta\frac{\chi^2}{(1-\sqrt{2}\beta)^2} \leq 3\alpha^2\mu_+\beta\frac{\chi^2}{(1-\sqrt{2}\beta)^2},
\end{aligned}$$

$$\|\alpha(\mathbf{\Delta}_{\widetilde{\mathbf{x}}}\mathbf{l}_\alpha)_H\| \leq 2\alpha\,\|\mathbf{\Delta}_{\widetilde{\mathbf{x}}}\|\,\|\mathbf{l}_\alpha\| \leq 2\sqrt{2}\alpha^2\mu_+\frac{\chi^2}{(1-\sqrt{2}\beta)^2} \leq 3\alpha^2\mu_+\frac{\chi^2}{(1-\sqrt{2}\beta)^2},$$

$$\begin{aligned}\left\|\boldsymbol{\mathcal{Q}}_{\mathbf{l}_\alpha^*}(\mathbf{v}-\mu_+\mathbf{e})\right\| &\le 2\left\|\langle\!\langle \mathbf{v}-\mu_+\mathbf{e}\rangle\!\rangle\right\|\left\|\mathbf{l}_\alpha\right\|^2 \\ &\le 2\sqrt{2}\alpha^2\mu_+\beta\frac{\chi^2}{(1-\sqrt{2}\beta)^2} \le 3\alpha^2\mu_+\beta\frac{\chi^2}{(1-\sqrt{2}\beta)^2},\end{aligned}$$

and

$$\left\|\alpha\boldsymbol{\mathcal{Q}}_{\mathbf{l}_\alpha^*}(\boldsymbol{\Delta}_{\widetilde{\mathbf{x}}})\right\| \le 2\alpha\left\|\langle\!\langle \boldsymbol{\Delta}_{\widetilde{\mathbf{x}}}\rangle\!\rangle\right\|\left\|\mathbf{l}_\alpha\right\|^2 \le 2\sqrt{2}\alpha^3\mu_+\frac{\chi^3}{(1-\sqrt{2}\beta)^3} \le 3\alpha^3\mu_+\frac{\chi^3}{(1-\sqrt{2}\beta)^3}.$$

The required inequality (3.7) then follows from the triangle inequality.

(i) If $\overline{\beta} > \beta$, then $p(0) = \beta - \overline{\beta} < 0$ shows that p has at least one positive real root. If not, then $\sigma > 0$, $p(0) = 0$ and $p'(0) = \overline{\beta} - \sigma\overline{\beta} - \beta = -\sigma\overline{\beta} < 0$ also shows the same. Hence p has a smallest positive real root $\overline{\alpha}$ and $p(\alpha) \le 0$ for all $\alpha \in [0, \overline{\alpha}]$.

(ii) If $(1-\sqrt{2}\beta)/(2\chi) < 1$, then for $\alpha = (1-\sqrt{2}\beta)/(2\chi) \in (0,1)$, we have

$$p(\alpha) > 7\alpha^2\frac{\chi^2}{(1-\sqrt{2}\beta)^2} - \overline{\beta} = \frac{7}{4} - \overline{\beta} > 0.$$

Hence $0 \le \alpha \le \min\{1, (1-\sqrt{2}\beta)/(2\chi)\}$ whenever $0 \le \alpha \le \min\{1, \overline{\alpha}\}$. Subsequently $\mathbf{s}_\alpha \in \mathcal{K}^\sharp$ and (3.7) holds whenever $0 \le \alpha \le \min\{1, \overline{\alpha}\}$. This leads to

$$\begin{aligned}\left\|\boldsymbol{\mathcal{Q}}_{\mathbf{l}_{\mathbf{s}_\alpha}^*}(\mathbf{x}_\alpha) - \mu_\alpha\mathbf{e}\right\| &\le p(\alpha) + (1-\alpha+\alpha\sigma)\mu_+\overline{\beta} \\ &\le (1-\alpha+\alpha\sigma)\mu_+\overline{\beta} = \mu_\alpha\overline{\beta},\end{aligned}$$

whence

$$\begin{aligned}\left\langle\boldsymbol{\mathcal{Q}}_{\mathbf{l}_{\mathbf{s}_\alpha}^*}(\mathbf{x}_\alpha), \mathbf{l}\mathbf{l}^*\right\rangle &= \mu_\alpha\left\langle\mathbf{e}, \mathbf{l}\mathbf{l}^*\right\rangle + \left\langle\boldsymbol{\mathcal{Q}}_{\mathbf{l}_{\mathbf{s}_\alpha}^*}(\mathbf{x}_\alpha) - \mu_\alpha\mathbf{e}, \mathbf{l}\mathbf{l}^*\right\rangle \\ &\ge \mu_\alpha\left\|\mathbf{l}\right\|^2 - \left\|\boldsymbol{\mathcal{Q}}_{\mathbf{l}_{\mathbf{s}_\alpha}^*}(\mathbf{x}_\alpha) - \mu\mathbf{e}\right\|\left\|\mathbf{l}\right\|^2 \\ &\ge \mu_\alpha(1-\overline{\beta})\left\|\mathbf{l}\right\|^2 > 0.\end{aligned}$$

for all $\mathbf{l} \in \mathfrak{T}_+^* \setminus \{\mathbf{0}\}$. We then conclude that $\boldsymbol{\mathcal{Q}}_{\mathbf{l}_{\mathbf{s}_\alpha}^*}(\mathbf{x}_\alpha) \in \mathcal{K}$, whence $\mathbf{x}_\alpha \in \mathcal{K}$. Finally (3.9) follows directly from (3.7).

□

4. Primal-Dual Path-Following Algorithms

The three algorithms in this paper are based on the following generic primal-dual path-following framework.

Algorithm 1. (Primal-dual path-following framework)
Given a pair of primal-dual strictly feasible solutions $(\mathbf{x}_{in}, \mathbf{s}_{in}) \in \mathcal{K} \oplus \mathcal{K}^\sharp$ *and* $\mu_{in} \in \mathbb{R}_{++}$ *with* $d(\mathbf{x}_{in}, \mathbf{s}_{in}; \mu_{in}) \le \beta$ *for some* $\beta \in (0, 1/\sqrt{2})$, *and the required accuracy* $\varepsilon > 0$.

(1) *Set* $(\mathbf{x}_+, \mathbf{s}_+) = (\mathbf{x}_{in}, \mathbf{s}_{in})$ *and* $\mu_+ = \mu_{in}$.
(2) *While* $\langle\mathbf{x}_+, \mathbf{s}_+\rangle > \varepsilon\,\langle\mathbf{x}_{in}, \mathbf{s}_{in}\rangle$,
  (a) *Pick* $\widetilde{\beta} \in (0, 1/\sqrt{2})$ *and* $\sigma \in [0,1]$.
  (b) *Solve* (3.2) *with* μ_{++} *replaced by* $\sigma\mu_+$. *For each* $\alpha \in [0,1]$, *let* $(\mathbf{x}_\alpha, \mathbf{s}_\alpha) = (\mathbf{x}_+ + \alpha\boldsymbol{\Delta}_{\mathbf{x}}, \mathbf{s}_+ + \alpha\boldsymbol{\Delta}_{\mathbf{s}})$, *and let* $\mu_\alpha = (1-\alpha+\alpha\sigma)\mu_+$. *Pick* $\widehat{\alpha} \in [0,1]$ *such that* $(\mathbf{x}_{\widehat{\alpha}}, \mathbf{s}_{\widehat{\alpha}}) \in \mathcal{K} \oplus \mathcal{K}^\sharp$ *and*

$$d_2(\mathbf{x}_{\widehat{\alpha}}, \mathbf{s}_{\widehat{\alpha}}; \mu_{\widehat{\alpha}}) \le \widetilde{\beta}.$$

  (c) *Update* $(\mathbf{x}_+, \mathbf{s}_+) \leftarrow (\mathbf{x}_{\widehat{\alpha}}, \mathbf{s}_{\widehat{\alpha}})$ *and* $\mu_+ \leftarrow \mu_{\widehat{\alpha}}$.

(3) *Output* $(\mathbf{x}_{out}, \mathbf{s}_{out}) = (\mathbf{x}_+, \mathbf{s}_+)$.

4.1. **Short-step algorithm.** In our short-step algorithm, conservative updates of the parameter μ are used and full Newton steps are taken. The target parameter μ_{++} is chosen so that the ratio $\sigma = \mu_{++}/\mu_+$ is $(1-\delta/\sqrt{r})$, where $\delta \in (0,1)$ is a constant. We shall show that with appropriate choices of β and δ, Algorithm 1 requires at most $O(\sqrt{r})$ iterations to reduce the duality gap by a constant factor.

Theorem 3. *If $\beta \in (0, 1/\sqrt{2})$ and $\delta \in (0,1)$ satisfy*

$$\frac{(\beta+\delta)^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \frac{3(\beta+\delta)^3}{(1-\sqrt{2}\beta)^3} < \left(1 - \frac{\delta}{\sqrt{r}}\right)\beta, \tag{4.1}$$

then we may use $\widehat{\alpha} = 1$ in each iteration of Algorithm 1 with $\sigma = 1 - \delta/\sqrt{r}$ and $\widetilde{\beta} = \beta$. Moreover the algorithm terminates after $O(\sqrt{r}\log\frac{1}{\varepsilon})$ iterations.

Proof. Consider an iteration of the algorithm. By choice of $\widetilde{\beta}$, we have $d(\mathbf{x}_+, \mathbf{s}_+; \mu_+) \le \beta$ at the beginning of the iteration. Let $\overline{\alpha}$ be the smallest positive root of the cubic polynomial (3.8) with $\sigma = (1-\delta/\sqrt{r})$, $\chi = \beta + (1-\sigma)\sqrt{r} = \beta + \delta$, and $\overline{\beta} = \beta$. Under the hypothesis (4.1), we have

$$\begin{aligned} &(1-\alpha)\beta + \alpha^2\frac{\chi^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \alpha^3\frac{3\chi^3}{(1-\sqrt{2}\beta)^3} - (1-\alpha+\sigma\alpha)\overline{\beta} \\ &\le \alpha\left[\frac{(\beta+\delta)^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \frac{3(\beta+\delta)^3}{(1-\sqrt{2}\beta)^3} - \left(1-\frac{\delta}{\sqrt{r}}\right)\beta\right] < 0 \end{aligned}$$

for all $0 < \alpha \le 1$, whence $\overline{\alpha} > 1$. Thus by Lemma 2, we have $(\mathbf{x}_\alpha, \mathbf{s}_\alpha) \in \mathcal{K} \oplus \mathcal{K}^\sharp$ and $d(\mathbf{x}_\alpha, \mathbf{s}_\alpha; \mu_\alpha) \le \beta$ for all $\alpha \in [0,1]$. This means we may take $\widehat{\alpha} = 1$ in the iteration. Finally, the duality gap reduces by the factor $(1-\delta/\sqrt{r})$ in each iteration, whence by a factor of ε in $O(\sqrt{r}\log\frac{1}{\varepsilon})$ iterations. □

4.2. **Large-update algorithm.** For the large-update algorithm, large updates $\mu_{++} = \sigma\mu_+$, where $\sigma \in (0,1)$ is a constant independent of r, are taken. We shall show that as long as $\beta \in (0, 1/(2\sqrt{2}))$, Algorithm 1 requires at most $O(r)$ iterations to reduce the duality gap by a constant factor.

Theorem 4. *If $\beta \in (0, 1/(2\sqrt{2}))$, then we may use $\widetilde{\alpha} = \Omega(1/r)$ in Algorithm 1 with $\widetilde{\beta} = \beta$ and $\sigma \in (0,1)$ arbitrary but fixed. Moreover the algorithm terminates after $O(r\log\frac{1}{\varepsilon})$ iterations.*

Proof. Consider an iteration of the algorithm. By choice of $\widetilde{\beta}$, we have $d(\mathbf{x}_+, \mathbf{s}_+; \mu_+) \le \beta$ at the beginning of the iteration. Let $\overline{\alpha}$ be the smallest positive root of the cubic polynomial (3.8) with $\chi = \beta + (1-\sigma)\sqrt{r}$ and $\overline{\beta} = \beta$; i.e., $\overline{\alpha}$ is the smallest positive root of

$$\begin{aligned} \alpha \mapsto &-\sigma\alpha\beta + \alpha^2\frac{(\beta+(1-\sigma)\sqrt{r})^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \alpha^3\frac{3(\beta+(1-\sigma)\sqrt{r})^3}{(1-\sqrt{2}\beta)^3} \\ &= \left(-\sigma\beta + \alpha\frac{(\beta+(1-\sigma)\sqrt{r})^2(7+6\beta)}{(1-\sqrt{2}\beta)^2} + \alpha^2\frac{3(\beta+(1-\sigma)\sqrt{r})^3}{(1-\sqrt{2}\beta)^3}\right)\alpha. \end{aligned}$$

By Lemma 2, we have $(\mathbf{x}_\alpha, \mathbf{s}_\alpha) \in \mathcal{K} \oplus \mathcal{K}^\sharp$ and

$$d(\mathbf{x}_\alpha, \mathbf{s}_\alpha; \mu_\alpha) \le \beta = \widetilde{\beta}$$

whenever $0 \le \alpha \le \min\{1, \overline{\alpha}\}$. Since $\overline{\alpha} = \Omega(1/r)$, we may use $\widetilde{\alpha} = \min\{1, \overline{\alpha}\} = \Omega(1/r)$. Finally, the duality gap reduces by the constant factor $(1 - \Omega(1/r))$ every two iterations, whence by a factor of ε in $O(r \log \frac{1}{\varepsilon})$ iterations. □

4.3. **Predictor-corrector algorithm.** In our predictor-corrector algorithm, the most aggressive updates $\mu_{++} = 0$ are taken, each of which is immediately followed by a zero update step $\mu_{++} = \mu_{+}$. This is a direct extension of the Mizuno-Todd-Ye predictor-corrector algorithm for linear programming [10]. The aggressive update steps are called *predictor steps* and the zero update steps are *corrector steps*. We shall show that with appropriate choices of β and $\widetilde{\beta}$, Algorithm 1 requires at most $O(\sqrt{r})$ iterations to reduce the duality gap by a constant factor.

Theorem 5. *If $\beta \in (0, 1/(2\sqrt{2}))$ satisfies*

$$\frac{4\beta^2(7 + 12\beta)}{(1 - 2\sqrt{2}\beta)^2} + \frac{24\beta^3}{(1 - 2\sqrt{2}\beta)^3} \le \beta, \tag{4.2}$$

then by alternating between $(\sigma, \widetilde{\beta}) = (0, 2\beta)$ and $(\sigma, \widetilde{\beta}) = (1, \beta)$ in Algorithm 1, we may use $\widetilde{\alpha} = \Omega(1/\sqrt{r})$ and $\widetilde{\alpha} = 1$, respectively. Moreover the algorithm terminates after $O(\sqrt{r} \log \frac{1}{\varepsilon})$ iterations.

Proof. Consider an iteration of the algorithm where a predictor step is to be taken. By choice of $\widetilde{\beta}$ in the previous iteration, we have $d(\mathbf{x}_+, \mathbf{s}_+; \mu_+) \le \beta$ at the beginning of the iteration. Let $\overline{\alpha}$ be the smallest positive root of the cubic polynomial (3.8) with $\sigma = 0$, $\chi = \beta + (1 - \sigma)\sqrt{r} = \beta + \sqrt{r}$, and $\overline{\beta} = 2\beta \in (\beta, 1)$; i.e., $\overline{\alpha}$ is the smallest positive root of

$$\alpha \mapsto -\beta + \alpha\beta + \alpha^2 \frac{(\beta + \sqrt{r})^2(7 + 6\beta)}{(1 - \sqrt{2}\beta)^2} + \alpha^3 \frac{3(\beta + \sqrt{r})^3}{(1 - \sqrt{2}\beta)^3}.$$

By Lemma 2, we have $(\mathbf{x}_\alpha, \mathbf{s}_\alpha) \in \mathcal{K} \oplus \mathcal{K}^\sharp$ and

$$d(\mathbf{x}_\alpha, \mathbf{s}_\alpha; \mu_\alpha) \le 2\beta = \widetilde{\beta}$$

whenever $0 \le \alpha \le \min\{1, \overline{\alpha}\}$. Since $\overline{\alpha} = \Omega(1/\sqrt{r})$, we may use $\widetilde{\alpha} = \min\{1, \overline{\alpha}\} = \Omega(1/\sqrt{r})$.

Now consider an iteration of the algorithm where a corrector step is to be taken. By choice of $\widetilde{\beta}$ in the previous iteration, we have $d(\mathbf{x}_+, \mathbf{s}_+; \mu_+) \le 2\beta$ at the beginning of the iteration. Under the hypothesis (4.2), for $\sigma = 1$, $\chi = (2\beta) + (1 - \sigma)\sqrt{r} = 2\beta$ and $\overline{\beta} = 2\beta$, we have

$$\begin{aligned}
&(1 - \alpha)(2\beta) + \alpha^2 \frac{\chi^2(7 + 6(2\beta))}{(1 - \sqrt{2}(2\beta))^2} + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2}(2\beta))^3} - (1 - \alpha + \sigma\alpha)\overline{\beta} \\
&\le \alpha \left[\frac{4\beta^2(7 + 12\beta)}{(1 - 2\sqrt{2}\beta)^2} + \frac{24\beta^3}{(1 - 2\sqrt{2}\beta)^3} - 2\beta \right] < 0
\end{aligned}$$

for all $0 < \alpha \le 1$, whence by Lemma 2, we have $(\mathbf{x}_\alpha, \mathbf{s}_\alpha) \in \mathcal{K} \oplus \mathcal{K}^\sharp$ and

$$\begin{aligned}
d(\mathbf{x}_\alpha, \mathbf{s}_\alpha; \mu_\alpha) &\le \frac{1}{1 - \alpha + \alpha\sigma} \left((1 - \alpha)(2\beta) + \alpha^2 \frac{\chi^2(7 + 6(2\beta))}{(1 - \sqrt{2}(2\beta))^2} + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2}(2\beta))^3} \right) \\
&\le \left(2(1 - \alpha)\beta + \alpha \left(\frac{4\beta^2(7 + 12\beta)}{(1 - 2\sqrt{2}\beta)^2} + \frac{24\beta^3}{(1 - 2\sqrt{2}\beta)^3} \right) \right) \\
&\le 2\beta - \alpha\beta
\end{aligned}$$

for all $\alpha \in [0, 1]$, in particular $d(\mathbf{x}_1, \mathbf{s}_1; \mu_1) \leq \beta$. This means we may take $\widehat{\alpha} = 1$ in the iteration.

Finally, the duality gap reduces by the constant factor $(1 - \Omega(1/\sqrt{r}))$ every two iterations, whence by a factor of ε in $O(\sqrt{r} \log \frac{1}{\varepsilon})$ iterations. □

5. Conclusion

In this paper, a generic primal-dual path-following framework for homogeneous cone programming is presented. With it, we can design a short-step algorithm, a large-update algorithm, and a predictor-corrector algorithm. Each algorithm is shown to have polynomial iteration complexity bounds that match existing bounds for their counter-parts in semidefinite programming.

There are two rather unsatisfactory features of this generic primal-dual framework worth further investigations:

(1) There is a lack of primal-dual symmetry. In another words, when we swap the primal problem with the dual problem, together with the initial primal-dual pair of solutions, and apply the same algorithm using the same respective T-algebras for the primal and dual homogeneous cones, we may not get the same iterates.
(2) In order for the primal-dual search direction to be uniquely defined, it is necessary for the primal-dual framework to use the narrow ℓ_2-neighbourhood. A first step towards designing a wide-neighbourhood primal-dual algorithm for homogeneous cone programming is to find a search direction that is defined at all pairs of primal-dual strictly feasible solutions.

References

[1] F. Alizadeh and S. H. Schmieta, *Extension of primal-dual interior point algorithms to symmetric cones*, Math. Program. **96** (2003), 409–438.
[2] C. B. Chua, *An algebraic perspective on homogeneous cone programming, and the primal-dual second order cone approximations algorithm for symmetric cone programming*, Ph.D. thesis, Cornell University, Ithaca, NY, USA, May 2003.
[3] ———, *The primal-dual second-order cone approximations algorithm for symmetric cone programming*, manuscript, Cornell University, Feburary 2003, accepted by Found. Comput. Math.
[4] ———, *Relating homogeneous cones and positive definite cones via T-algebras*, SIAM J. Optim. **14** (2003), no. 2, 500–506.
[5] ———, *A new notion of weighted centers for semidefinite programming*, SIAM J. Optim. **16** (2006), no. 4, 1092–1109.
[6] ———, *Target following algorithms for semidefinite programming*, Research Report CORR 2006-10, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Canada, May 2006, 40 pages.
[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** (1997), 149–175.
[9] O. Güler and L. Tunçel, *Characterization of the barrier parameter of homogeneous convex cones*, Math. Program. **81** (1998), 55–76.
[10] S. Mizuno, M. J. Todd, and Y. Ye, *An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm*, Math. Oper. Res. **19** (1994), 53–67.
[11] R. D. C. Monteiro and P. Zanjácomo, *Implementation of primal-dual methods for semidefinite programming based on Monteiro and Tsuchiya Newton directions and their variants*, Optim. Methods Softw. **11/12** (1999), 91–140.
[12] Yu. E. Nesterov and A. S. Nemirovski, *Interior point polynomial algorithms in convex programming*, SIAM Stud. Appl. Math., vol. 13, SIAM Publication, Philadelphia, PA, USA, 1994.

[13] Yu. E. Nesterov and M. J. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM J. Optim. **8** (1998), 324–364.
[14] B. K. Rangarajan, *Polynomial convergence of infeasible-interior-point methods over symmetric cones*, SIAM J. Optim. **16** (2006), 1211–1229.
[15] L. Tunçel, *Primal-dual symmetry and scale invariance of interior-point algorithms for convex optimization*, Math. Oper. Res. **23** (1998), 708–718.
[16] ———, *Generalization of primal-dual interior-point methods to convex optimization problems in conic form*, Found. Comput. Math. **1** (2001), 229–254.
[17] È. B. Vinberg, *The theory of convex homogeneous cones*, Tr. Mosk. Mat. O.-va **12** (1963), 303–358.
[18] H. Wolkowicz, R. Saigal, and L. Vandenberghe (eds.), *Handbook of semidefinite programming: Theory, algorithms, and applications*, Springer-Verlag, Berlin-Heidelberg-New York, 2000.
[19] S. J. Wright, *Primal-dual interior-point methods*, SIAM Publication, Philadelphia, PA, USA, 1997.

Division of Mathematical Sciences, Nanyang Technological University, 1 Nanyang Walk, Blk 5 Level 3, Singapore 637616, Singapore
E-mail address: cbchua@ntu.edu.sg