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 cones via $T$-algebras. The analysis shows that the algorithms have polynomial iteration complexity.

Key words. homogeneous cone programming, $T$-algebra, primal-dual interior-point algorithm

AMS subject classifications. 90C25, 90C51

DOI. 10.1137/060677343

1. Introduction. Primal-dual interior-point algorithms—first designed for linear programming (see, e.g., [20]), and subsequently extended to semidefinite programming (see, e.g., [19, 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 facts strongly motivate 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 for all convex optimization problems in conic form 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 subclass, symmetric cone programming. While there is a finite number of nonisomorphic 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 [3] 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 determine the best known worst-
case 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, 6, 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 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 $E$ denote a Euclidean space with inner product $\langle \cdot , \cdot \rangle$.

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

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

We consider the following homogeneous cone programming problem

\[
(HCP) \quad \inf_{x} \langle \hat{s}, x \rangle \quad \text{subject to} \quad x \in L + \{\hat{x}\}, \quad x \in cl(K)
\]

and its Lagrangian dual problem

\[
(HCD) \quad \inf_{s} \langle \hat{x}, s \rangle \quad \text{subject to} \quad s \in L^\perp + \{\hat{s}\}, \quad s \in cl(K^\sharp),
\]

where $L^\perp$ denotes the orthogonal complement of $L$ in $E$ and $K^\sharp$ denotes the (open) dual cone $\{s \in E : \langle x, s \rangle > 0 \quad \forall x \in cl(K) \setminus \{0\}\}$. Since $K$ is pointed, the dual cone $K^\sharp$ is nonempty.
2.1. T-algebras. We begin with some basics of $T$-algebras; see [17], [18], and [2, Chapter 2] for more details.

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

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

The positive integer $r$ is called the rank of the matrix algebra. For each $a \in \mathfrak{A}$, we denote by $a_{ij}$ 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., $(a^*)^* = a$ for all $a \in \mathfrak{A}$),
2. antihomomorphic (i.e., $(ab)^* = b^*a^*$ for all $a, b \in \mathfrak{A}$),

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 $i \in \{1, \ldots, r\}$, the subalgebra $\mathfrak{A}_{ii}$ is isomorphic to the reals.

For the remaining axioms, we shall use $\rho_i$ to denote the isomorphism from $\mathfrak{A}_{ii}$ to $\mathbb{R}$, and use $e_i$ to denote the unit of $\mathfrak{A}_{ii}$.

II. For each $a \in \mathfrak{A}$ and each $i, j \in \{1, \ldots, r\}$,

$$a_{ji}e_i = a_{jj} \quad \text{and} \quad e_i a_{ij} = a_{ij}.$$

III. For each $a, b \in \mathfrak{A}$ and each $i, j \in \{1, \ldots, r\}$,

$$\rho_i(a_{ij}b_{ji}) = \rho_j(b_{ji}a_{ij}).$$

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

$$a_{ij}(b_{jk}c_{ki}) = (a_{ij}b_{jk})c_{ki}.$$  

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

$$\rho_i(a_{ij}^*a_{ij}) \geq 0$$

with equality when and only when $a_{ij} = 0$.

VI. For each $a, b, c \in \mathfrak{A}$ and each $i, j, k, l \in \{1, \ldots, r\}$ with $i \leq j \leq k \leq l$,

$$a_{ij}(b_{jk}c_{kl}) = (a_{ij}b_{jk})c_{kl}.$$

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

$$a_{ij}(b_{jk}b_{lk}^*) = (a_{ij}b_{jk})b_{lk}^*.$$

**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 $\pi$ is a permutation of $\{1, \ldots, r\}$. 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 other words, the permutation \( \pi \) in an inessential change satisfies
\[
(i < j) \land (\pi(i) > \pi(j)) \implies \mathcal{A}_{ij} = \{0\}.
\]

**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 bigraded algebras with involution.

**Definition 7** (associated cone). The cone associated with a \( T \)-algebra \( \mathfrak{A} \) of rank \( r \) is
\[
\{ tt^*: t \in \mathfrak{A}, \ t_{ij} = 0 \ \forall \ 1 \leq j < i \leq r, \ t_{ii} > 0 \ \forall \ 1 \leq i \leq r \}.
\]
It is denoted by \( K(\mathfrak{A}) \).

**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
\[
(ab)_{ij} = a_{ij}^T b_{ji} \quad (i, j \in \{1, 2\}, \ i \neq j).
\]
By Corollary IV.1.5 of [7], the subset of elements of \( \mathfrak{A} \) satisfying \( a^* = a \), when equipped with the Jordan product
\[
(a, b) \mapsto \frac{ab + ba}{2},
\]
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 \( i \neq j \), \( \mathfrak{A}_{ii} = \mathbb{R} \) for \( i \in \{1, \ldots, r\} \), and define the matrix algebra \( \mathfrak{A} = \bigoplus_{i,j=1}^r \mathfrak{A}_{ij} \) by
\[
(ab)_{ij} = \begin{cases} 
\sum_{k=1}^r \Re(a_{ik} b_{kj}) & \text{if } i = j, \\
\sum_{k=1}^r a_{ik} b_{kj} & \text{if } i \neq j,
\end{cases}
\]
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
\[
a_{ij}^* = a_{ij}
\]
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 \( a_{ij} \), \( b_{ik} \), and \( c_{kl} \) 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 subalgebra generated by any two elements is associative), and both \( x \) and \( \bar{x} \) are in the subalgebra 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 \( a^* = a \), when equipped with the Jordan product, is a Euclidean Jordan algebra of rank 3.

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

**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 $K(\mathfrak{A}) = K$;
2. $e_i$ shall denote the unit of the subalgebra $\mathfrak{A}_{ii}$ and $\rho_i$ shall denote the linear function on $\mathfrak{A}$ satisfying
   
   $$a_{ii} = \rho_i(a) e_i \quad (a \in \mathfrak{A})$$

   for each $i \in \{1, \ldots, r\}$;
3. $e$ shall denote the element in $\mathfrak{A}$ satisfying
   
   $$e_{ii} = e_i \quad (1 \leq i \leq r) \quad \text{and} \quad e_{ij} = e_{ji} = 0 \quad (1 \leq i < j \leq r);$$
4. $\text{tr}(\cdot)$ shall denote the linear function
   
   $$a \in \mathfrak{A} \mapsto \sum_{i=1}^{r} \rho_i(a);$$
5. $(\cdot, \cdot)$ shall denote the bilinear function
   
   $$\langle a, b \rangle \in \mathfrak{A} \otimes \mathfrak{A} \mapsto \text{tr}(a^* b);$$
6. $\mathfrak{T}$ shall denote the subalgebra
   
   $$\{a \in \mathfrak{A} : a_{ij} = 0 \quad (1 \leq j < i \leq r)\}$$

   of upper triangular elements (so that $K = K(\mathfrak{A}) = \{tt^* : t \in \mathfrak{T}, \rho_i(t) > 0 \ \forall i\}$);
7. $\mathfrak{H}$ shall denote the subspace
   
   $$\{a \in \mathfrak{A} : a_{ij} = a_{ji}^* \quad (1 \leq j < i \leq r)\}$$

   of “hermitian” elements (so that the linear span of $K$ is $\mathfrak{H}$);
8. for any $a \in \mathfrak{A}$, the unique lower triangular element $l$ satisfying $a - l \in \mathfrak{T}$ and
   
   $$l_{ii} = a_{ii}/2 \quad \text{for} \quad i = 1, \ldots, r,$$

   shall be denoted by $\langle a \rangle$; and
9. for any $a \in \mathfrak{A}$, the element $a + a^* \in \mathfrak{H}$ shall be denoted by $a_H$.

**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 $\text{tr}(\cdot)$ is invariant under involution. Subsequently $(\cdot, \cdot)$ is symmetric.

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

**Remark 3.** Axiom III is equivalent to $\text{tr}(ab) = \text{tr}(ba)$ for all $a, b \in \mathfrak{A}$. This is further equivalent to $\langle a, b \rangle = \langle a^*, b^* \rangle$; see also, [18, Chapter I, section 3].
Remark 4. Axiom V is equivalent to \( \text{tr}(ab^*) \geq 0 \) for all \( a \in \mathfrak{A} \) with equality when and only when \( a = 0 \). 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 \( a \mapsto \sqrt{\langle a, a \rangle} \). We shall view \( \mathcal{K} \) as a cone in the Euclidean space \( (\mathfrak{K}, \langle \cdot, \cdot \rangle) \).

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

\[
\langle ab, c \rangle = \langle b^*a^*, c^* \rangle = \langle a^*, bc^* \rangle = \langle a, cb^* \rangle \quad (a, b, c \in \mathfrak{A});
\]
i.e., the adjoint of the linear map \( b \in \mathfrak{A} \mapsto ba \) under the inner product \( \langle \cdot, \cdot \rangle \) is the linear map \( \text{adj} \) for each \( a \in \mathfrak{A} \); see, also, [18, Chapter I, section 3].

Remark 6. Axiom VI is equivalent to \( \text{tr}(uw^*) = \text{tr}(wu^*) \) for all \( t, u, w \in \mathfrak{K} \). Taking involution gives the equivalent statement: \( \text{tr}(uw) = \text{tr}(wu) \) for all \( t, u, w \in \mathfrak{K}^* \).

Remark 7. Axiom VII is equivalent to \( \text{tr}(uu^*) = \text{tr}(u^*u) \) for all \( u \in \mathfrak{K} \). Taking involution gives the equivalent statement: \( \text{tr}(u^*t) = \text{tr}(ut) \) for all \( t, u \in \mathfrak{K}^* \).

Polarization gives the equivalent statements:

1. \( \text{tr}(wu^*) + \text{tr}(wu^*) = \text{tr}(wu^*) + \text{tr}(wu^*) \) for all \( t, u, w \in \mathfrak{K} \), and
2. \( \text{tr}(u^*w^*) = \text{tr}(u^*w^*) + \text{tr}(w^*u^*) \) for all \( t, u, w \in \mathfrak{K}^* \).

Theorem 2 (sub-multiplicativity of Euclidean norm; c.f. Proposition 2.10 of [2]). The norm \( \| \cdot \| \) is sub-multiplicative; i.e., \( \| ab \| \leq \| a \| \| b \| \) for all \( a, b \in \mathfrak{A} \).

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

\[
(2.1) \quad \| a_i b_{jk} \| \leq \| a_i \| \| b_{jk} \| \quad \forall 1 \leq i, j, k \leq r.
\]

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

\[
\| a_i b_{jk} \|^2 = \langle a_i b_{jk}, a_i b_{jk} \rangle = \langle a_i, (a_j b_{jk})b_{jk}^* \rangle = \langle a_i, a_j (b_{jk} b_{jk}^*) \rangle = \langle a_i^* a_j, b_{jk} b_{jk}^* \rangle = \rho_j (a_i^* a_j) \rho_j (b_{jk} b_{jk}^*).
\]

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

\[
\| a_i b_{jk} \|^2 = \langle a_i b_{jk}, a_i b_{jk} \rangle = \langle a_i, (a_j b_{jk})b_{jk}^* \rangle = \langle a_i, a_j (b_{jk} b_{jk}^*) \rangle 
\leq \| a_i \| \| (a_j b_{jk}) b_{jk}^* \| 
= \| a_i \| \| a_j b_{jk} \| \| b_{jk}^* \|
\]

which shows that \( \| a_i b_{jk} \| \leq \| a_i \| \| b_{jk}^* \| = \| a_i \| \| b_{jk} \| \).
Case 3: $1 \leq j < i \leq k \leq r$. Once again, we use the Cauchy–Schwarz inequality and the first case to deduce
\[
\|a_ib_{jk}\|^2 = \langle a_ib_{jk}, a_ib_{jk} \rangle \\
= \langle a_i^*(a_ib_{jk}), b_{jk} \rangle \quad \text{(Axiom IV)} \\
\leq \|a_i^*(a_ib_{jk})\| \|b_{jk}\| \\
= \|a_i^*\| \|a_ib_{jk}\| \|b_{jk}\| .
\]

By triangle inequality, we have
\[
\|ab\|^2 = \sum_{i,j=1}^{r} \left( \sum_{k=1}^{r} a_{ik}b_{kj} \right)^2 \leq \sum_{i,j=1}^{r} \left( \sum_{k=1}^{r} \|a_{ik}\| \|b_{kj}\| \right)^2 \\
\leq \sum_{i,j=1}^{r} \left( \sum_{k=1}^{r} \|a_{ik}\|^2 \sum_{k=1}^{r} \|b_{kj}\|^2 \right) \\
= \|a\|^2 \|b\|^2 ,
\]
hence proving the theorem. □

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

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

**Proposition 1** (cf. first paragraph of Chapter II, section 2 of [17]). The sets $\Sigma_*$, $\Sigma_{++}$, $\Sigma_+^*$, and $\Sigma_{++}^*$ are multiplicative groups.

**Proof.** We shall prove that $\Sigma_*$ and $\Sigma_{++}$ are multiplicative groups. A similar result then holds for $\Sigma_+^*$ and $\Sigma_{++}^*$ since involution is antihomomorphic and $e^* = e$.

Products of upper triangular elements are upper triangular with diagonal components the products of respective diagonal components. Let $t \in \Sigma_*$ be an arbitrary. Let $u$ be the upper triangular element whose components are defined recursively by
\[
u_i,i+j = \begin{cases} \frac{1}{\rho_{i}(t)} e_i & \text{for } j = 0, \\ \frac{1}{\rho_{i+j}(t)} \sum_{k=0}^{j-1} u_{i,i+k} \rho_{i+k,i+j} & \text{for } j = 1, \ldots, r - i. \end{cases}
\]

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

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

Involution is antihomomorphic, and hence the inverse of the involution of an element $t \in \Sigma_*$ is the involution of its inverse, which we shall denote by $t^{-*}$. Similarly for elements $l \in \Sigma_*^*$. 
2.2. Quadratic representations. Consider, for each $t \in \mathbb{T}_+$, the map

$$\tilde{Q}_t : uu^* \in K \mapsto (tu)(tu)^*$$
on $K$. We shall demonstrate that the collection $\{\tilde{Q}_t : t \in \mathbb{T}_+\}$ gives rise to a transitive subgroup of $\text{Aut}(K)$.

We first show that for each $t \in \mathbb{T}_+$, the map $\tilde{Q}_t$ can be extended to a linear map on $E$.

Proposition 2. For each $t, u \in \mathbb{T}$,

$$(tu)(tu)^* = t((uu^*)t^*) + t(t(uu^*)) - (tt)(uu^*).$$

Proof. For each $t, u \in \mathbb{T}$,

$$t((uu^*)t^*) + t(t(uu^*)) - (tt)(uu^*)$$
$$= t(u(u^*)t^*) + t((tu)u^*) - ((tt)u)u^* \quad (\text{Axiom VII})$$
$$= (tu)(tu)^* + (tt)u^* - ((tt)u)u^* \quad (\text{Remark 7})$$
$$= (tu)(tu)^* \quad (\text{Axiom VI})$$

proves the proposition.  \[\square\]

This proposition suggests the following extension of the maps $\tilde{Q}_t$.

Definition 8 (quadratic representation). For each $a \in A$, the quadratic representation of $a$ is the linear map

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

It is denoted by $Q_a$.

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

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

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

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

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

Next we show that the collection $\{Q_t : t \in \mathbb{T}_+\}$ of extensions forms a group.

Corollary 1. The map $t \mapsto Q_t$ is a group homomorphism on $\mathbb{T}_+$.  

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

$$Q_t(Q_u(ww^*)) = Q_t((uw)(uw)^*) = (t(uw))(t(uw))^*$$
$$= ((tu)w)((tu)w)^* \quad (\text{Axiom VI})$$
$$= Q_{tu}(ww^*),$$
shows that $Q_t Q_u = Q_{tu}$ on $K$. Since $Q_t$ is linear for each $t$, and the linear span of $K$ is the space $H$, the corollary follows.

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

Besides giving us a transitive subgroup of $\text{Aut}(K)$, the maps $Q_t$, where $t \in T_+$, play a key role in the design and analysis of the primal-dual algorithms in this paper.

We end this section with several useful properties of these maps.

**Proposition 3.** For each $t \in T$, the map $Q_t$ simplifies to

$$a \mapsto t(\langle a \rangle t^*) + (t \langle a \rangle^*) t^*.$$

**Proof.** Let $t \in T$ and $a \in H$ be arbitrary. Let $w$ denote the upper triangular element $\langle a \rangle^*$. Then

$$t(at^*) + t(ta) - (tt)a = t(wt^*) + t(w^*t^*) + t(tw) + t(tw^*) - (tt)w - (tt)w^*$$

$$= t(wt^*) + t(w^*t^*) + t(tw^*) - (tt)w^*$$

$$= (tw)t^* + (w^*t^*)$$

(Axiom VI)

(Remark 7)

proves the proposition. □

The author [2] proved that the map $t \mapsto tt^* = Q_t(e)$ is a bijection between $T_+$ and $K(H)$; see also Proposition 2 of [17, Chapter III]. The following proposition shows that it is in fact a diffeomorphism.

**Proposition 4.** For each $u \in T_+$, the map

$$t \in T_+ \mapsto Q_t(uu^*)$$

is a diffeomorphism. Moreover, its derivative at $t \in T_+$ is

$$w \in T \mapsto (tu)(wu)^* + (wu)(tu)^*,$$

and the derivative of its inverse map at $Q_t(uu^*) \in K$ is

$$a \in H \mapsto tu \langle (Q_{tu^{-1}}(a))^* u^{-1}. $$

**Proof.** By Proposition 2, the derivative of $t \in T_+ \mapsto Q_t(uu^*)$ is the map

$$w \in T \mapsto (tu)(wu)^* + (wu)(tu)^*.$$

By Axiom VI and Proposition 3

$$(tu)(wu)^* + (wu)(tu)^* = (tu)((wu)^*(tu)^{-1})(tu)^* + ((tu)((tu)^{-1}(wu))(tu)^*$$

$$= Q_{tu}(((tu)^{-1}(wu))t);$$

hence by Corollary 1 the derivative has inverse

$$a \in H \mapsto tu \langle (Q_{tu^{-1}}(a))^* u^{-1}. $$

The proposition follows from the inverse function theorem. □

---

2It is, in fact, a simply transitive subgroup; i.e., for every $t, u \in T_+$, there exists a unique $Q_w$ with $w \in T_+$ that maps $tt^*$ to $uu^*$. 

---
2.3. Dual homogeneous cones. Using Axiom IV and Remark 5, it is straightforward to check that the adjoint map \( Q^H_a \) of the quadratic representation of an element \( a \in \mathfrak{g} \) is the quadratic representation \( Q^*_a \) of its involution.

In general, if \( \mathfrak{g} \) is an automorphism of an open cone \( \mathcal{C} \), then \( \mathfrak{g}^H \) is an automorphism of \( \mathcal{C}^\sharp \). Moreover, if \( \mathcal{G} \) is a group of transitive automorphisms of \( \mathcal{C} \), then the group \( \{ \mathfrak{g}^H : \mathfrak{g} \in \mathcal{G} \} \) also acts transitively on \( \mathcal{C}^\sharp \); see Proposition 9 of [17, Chapter I].

Recall that \( \{ Q_t : t \in \mathfrak{T}^+ \} \) is a transitive subgroup of automorphisms of \( \mathcal{K} \). Hence the dual cone \( \mathcal{K}^\sharp \) is

\[
\{ Q_t(x) : t \in \mathfrak{T}^+ \} = \{ Q_t(x) : t \in \mathfrak{T}^+_+ \},
\]

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

\[
\langle e, tt^* \rangle = \| t \|^2 > 0
\]

for all \( t \in \mathfrak{T} \setminus \{ 0 \} \). We thus proved the following proposition; see, also, [17, Chapter III, section 6].

**Proposition 5.** The dual cone \( \mathcal{K}^\sharp \) of \( \mathcal{K} \) is the homogeneous cone \( \{ \mathfrak{I}^* : t \in \mathfrak{T}^+_+ \} \).

The group of automorphisms \( \{ Q_t : t \in \mathfrak{T}^+_+ \} \) acts transitively on \( \mathcal{K}^\sharp \).

We now turn our attention to identifying the automorphism \( \mathfrak{T}^* \). It was pointed out by Vinberg [17, Chapter III, section 6] that the matrix algebra \( \mathfrak{B} \) is also a \( T \)-algebra. Indeed, Axioms I–V translate 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.

**Proposition 6** (cf. Proposition 1 of Chapter I of [18]). For each \( t, u \in \mathfrak{T} \),

\[
t^*(u^*u) = (t^*u^*)u.
\]

Equivalently, taking involution and/or polarization gives

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

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

\[
\langle w, t^*(u^*u) \rangle = \langle u(tw), u \rangle = \langle (ut)w, u \rangle (\text{Axiom IV})
\]

\[
= \langle w, (t^*u^*)u \rangle (\text{Axiom IV})
\]

implies \( \langle w, a \rangle = 0 \). Also

\[
\langle u, u(wt^*) \rangle = \langle w^*(u^*u), t^* \rangle (\text{Axiom IV})
\]

\[
= \langle w^*, t^*(u^*u) \rangle (\text{Remark 5})
\]

\[
= \langle (u(tw^*), u \rangle (\text{Axiom IV})
\]

\footnote{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 \( \mathfrak{B} \) follow from those for \( \mathfrak{A} \) by taking involution. While this is certainly true for Axiom VI, it does not hold for Axiom VII.}
and
\[
\langle u, (uw)t^* \rangle = ((w^*u^*)u, t^*) = (w^*, (t^*u^*)u) = ((ut)w^*, u),
\]
whence
\[
2 \langle w^*, t^*(u^*u) \rangle = \langle u(wt^*) + u(tw^*), u \rangle = 2 \langle w^*, (t^*u^*)u \rangle
\]
shows that \( \langle w^*, a \rangle = 0 \).

The following propositions are the counterparts of Propositions 2, 3, and 4 for \( B \), stated in terms of the quadratic representations in \( A \) instead of those in \( B \).

**Proposition 7.** For each \( l \in \mathfrak{T}^* \), the map \( Q_l \) simplifies to
\[
a \mapsto l(\langle a^* \rangle l^* + (l \langle a \rangle)l^*).
\]
If, in addition, \( a = kk^* \) for some \( k \in \mathfrak{T}^* \), then
\[
Q_l(kk^*) = (lk)(lk)^*.
\]

**Proposition 8.** For each \( k \in \mathfrak{T}_+^* \), the map
\[
l \in \mathfrak{T}_+^* \mapsto Q_l(kk^*)
\]
is a diffeomorphism. Moreover, its derivative at \( l \in \mathfrak{T}_+^* \) is
\[
m \in \mathfrak{T}^* \mapsto (lk)(mk)^* + (mk)(lk)^*,
\]
and the derivative of its inverse map at \( Q_l(kk^*) \in \mathfrak{K}^* \) is
\[
a \in \mathfrak{F} \mapsto lk \langle Q_{lk}^{-1}(a) \rangle k^{-1}.
\]

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 \( \theta \)-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 (nondegenerate, strongly) \( \theta \)-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
\[
f'''(x; h) \leq 2(f''(x; h))^{3/2}
\]
and
\[
f'(x; h) \leq (\theta f''(x; h))^{1/2}
\]
for all $x \in K$ and $h \in E$. Here, $f^{(k)}(x; h)$ denotes the $k$th directional derivative of $f$ along $h$ (i.e., $\frac{d^k f(x+th)}{dt^k}|_{t=0}$). The parameter $\vartheta$ is called the complexity parameter.

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

$$f(tx) = f(x) - \vartheta \log t.$$

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

Optimal barriers for a homogeneous cone $K$ of rank $r$ has complexity parameter $r$. Moreover, the function

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

is a $r$-logarithmically homogeneous self-concordant barrier for $K$, where $u_x$ denotes the unique element in $T^{++}_+$ satisfying $u_x u_x^* = x$; 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

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

see [2, 4] for more details.

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

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

it is straightforward to compute the gradient of $f$ using Proposition 4. This gives the gradient at each $x \in K$ as

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

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

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

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

$$\nabla f^\sharp(s) = -l_s^{-*} l_s^{-1}.$$
3.2. The central path. The central path of the primal-dual pair of homogeneous cone programming problems \((HCP)\) and \((HCD)\) is the set \(\{(x(\mu), s(\mu)) : \mu > 0\}\) of pairs of solutions to the primal and dual barrier problems

\[
\inf \left\{ \langle \bar{s}, x \rangle - \mu \sum_{i=1}^{r} \log (\rho_i(u^x_i)^2) : x \in \mathcal{L} + \{\bar{x}\} \right\},
\]

and

\[
\inf \left\{ \langle \bar{x}, s \rangle - \mu \sum_{i=1}^{r} \log (\rho_i(l^s_i)^2) : s \in \mathcal{L}^\perp + \{\bar{s}\} \right\},
\]

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

\[
(CP_\mu)
\]

\[
\begin{align*}
x & \in \mathcal{L} + \{\bar{x}\}, \quad x \in \mathcal{K}, \\
s & \in \mathcal{L}^\perp + \{\bar{s}\}, \quad s \in \mathcal{K}^2, \\
Q^{\ast}_r(x) &= \mu e.
\end{align*}
\]

3.3. Tracing the central path. We now consider the problem of estimating the pair \((x(\mu_\ast), s(\mu_\ast))\) that solves the central path equations

\[
(CP_{\mu_\ast})
\]

\[
\begin{align*}
x(\mu_\ast) & \in \mathcal{L} + \{\bar{x}\}, \quad x(\mu_\ast) \in \mathcal{K}, \\
s(\mu_\ast) & \in \mathcal{L}^\perp + \{\bar{s}\}, \quad s(\mu_\ast) \in \mathcal{K}^2, \\
Q^{\ast}_r(x(\mu_\ast)) &= \mu_\ast e,
\end{align*}
\]

where \(\mu_\ast \in \mathbb{R}_+\) is given. Note that we allow \(\mu_\ast\) 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 \((x(\mu), s(\mu))\), where \(\mu \in \mathbb{R}_+\), with the proximity measure

\[
d : \mathcal{K} \oplus \mathcal{K}^2 \oplus \mathbb{R}_+ \to \mathbb{R}_+\]

defined by

\[
(x, s, \mu) \mapsto \frac{1}{\mu} \|Q^{\ast}_r(x) - \mu e\|.
\]

Suppose we have a pair of estimates \((x_+, s_+)\) satisfying \(d(x_+, s_+; \mu_+) \leq \beta\) for some \(\mu_+ := \frac{1}{r} \langle x_+, s_+ \rangle > \mu_\ast\) and some \(\beta < 1\). We shall estimate \((x(\mu_\ast), s(\mu_\ast))\) with Newton’s method. This calls for the linearization of \((CP_\mu)\).

Proposition 9. Suppose \(x \in \mathcal{K}, s \in \mathcal{K}^2\), and \(a, b \in \mathcal{L}\). Then the directional derivative of

\[
(tt^\ast, ll^\ast) \in \mathcal{K} \oplus \mathcal{K}^2 \mapsto Q^{\ast}_r(tt^\ast) = (l^t)(l^t)^*\]

at \((x, s)\) along \((a, b)\) is

\[
Q^{\ast}_r(a) + \left( Q^{\ast}_r(x) \left\langle Q^{\ast}_{r^*}(b) \right\rangle \right)^H.
\]

Proof. Write the map

\[
(tt^\ast, ll^\ast) \in \mathcal{K} \oplus \mathcal{K}^2 \mapsto Q^{\ast}_r(tt^\ast) = (l^t)(l^t)^*
\]

as the composition of

\[
(tt^\ast, ll^\ast) \in \mathcal{K} \oplus \mathcal{K}^2 \mapsto (tt^\ast, l^\ast) \quad \text{and} \quad (tt^\ast, l^\ast) \in \mathcal{K} \oplus \mathcal{K}^2 \mapsto Q^{\ast}_r(tt^\ast).
\]
By Proposition 8 the derivative of the first map at \((c, d)\) is
\[(c, d) \in \mathcal{H} \oplus \mathcal{H} \mapsto (c, \langle \mathcal{Q}_{1}^{-1}(d) \rangle^{*} 1')\],
and by Proposition 4, the derivative of the second map at \((tt^*, 1')\) is
\[(c, k') \in \mathcal{H} \oplus \mathcal{H} \mapsto \mathcal{Q}_{1'}(c) + ((l^* t)(k^* t)^*)_{H}\].
Hence, chain rule gives the required directional derivative as
\[
\mathcal{Q}_{1'}(a) + (l^*_x u_x) \left( \left( \left\langle \mathcal{Q}_{1}^{-1}(b) \right\rangle^{*} l^*_x u_x \right)^* \right)_{H} = \mathcal{Q}_{1'}(a) + \left( (l^*_x u_x) \left( (l^*_x u_x)^* \left\langle \mathcal{Q}_{1}^{-1}(b) \right\rangle \right) \right)_{H}.
\]
By Axiom VII,
\[
(l^*_x u_x) \left( (l^*_x u_x)^* \left\langle \mathcal{Q}_{1}^{-1}(b) \right\rangle \right) = (l^*_x u_x)(l^*_x u_x)^* \left\langle \mathcal{Q}_{1}^{-1}(b) \right\rangle.
\]
and hence the proposition follows from Proposition 2. □

The above proposition gives the linearization of \((CP^*)\) at \((x_+, s_+)\) as
\[
\Delta_x \in \mathcal{L}, \quad \Delta_s \in \mathcal{L}^l,
\]
\[
\mathcal{Q}_{l^*_x}(\Delta_x) + \left( \mathcal{Q}_{l^*_x}(x_+) \left\langle \mathcal{Q}_{l^*_x}^{-1}(\Delta_s) \right\rangle \right)_{H} = \mu_{++} e - \mathcal{Q}_{l^*_x}(x_+).
\]
One way to ensure that \((x_+ + \Delta_x, s_+ + \Delta_s)\) is a good estimate of \((x(\mu_{++}), s(\mu_{++}))\) is to get a good upper bound on the sizes of \(\Delta_x\) and \(\Delta_s\).

The analysis can be greatly simplified by applying the primal-dual transformations
\[
(x, s) \mapsto (\mathcal{Q}_{l^*_x}(x), \mathcal{Q}_{l^*_s}^{-1}(s)).
\]
Let \(v, (\Delta_x, \Delta_s)\) and \(\tilde{\mathcal{L}}\) denote, respectively, the images of \(x_+, (\Delta_x, \Delta_s)\) and \(\mathcal{L}\) under these transformations. Then \((\Delta_x, \Delta_s)\) solves
\[
\Delta_x \in \tilde{\mathcal{L}}, \quad \Delta_s \in \tilde{\mathcal{L}}^l,
\]
\[
\Delta_x + (v \left\langle \Delta_s \right\rangle)_{H} = \mu_{++} e - v.
\]
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 10. Suppose \(v \in \mathcal{K}\) satisfies
\[
\|v - \mu e\| \leq \beta \mu
\]
for some \(\beta \in (0, 1/\sqrt{2})\) and some \(\mu > 0\). If \((\Delta_x, \Delta_s)\) satisfies \((\Delta_x, \Delta_s) \geq 0\) and
\[
\Delta_x + (v \left\langle \Delta_s \right\rangle)_{H} = h
\]
for some \(h \in \mathcal{H}\), then
\[
\max\{\|\Delta_x\|, \mu \|\Delta_s\|\} \leq \frac{\|h\|}{1 - \sqrt{2} \beta}.
\]
Proof. If \( \langle \Delta_x, \Delta_s \rangle \geq 0 \), then
\[
\max\{\|\Delta_x\|, \mu \|\Delta_s\|\} \leq \sqrt{\|\Delta_x\|^2 + \mu \|\Delta_s\|^2} \leq \|\Delta_x + \mu \Delta_s\|.
\]
By triangle inequality, hypotheses (3.4) and (3.5), and the submultiplicativity of the norm \( \| \cdot \| \), we have
\[
\|\Delta_x + \mu \Delta_s\| \leq \|\Delta_x + (v \langle \Delta_s \rangle)_H\| + \|((v - \mu e) \langle \Delta_s \rangle)_H\|
\leq \|h\| + 2 \|(v - \mu e) \langle \Delta_s \rangle\|
\leq \|h\| + 2\beta \mu \|\langle \Delta_s \rangle\|
\leq \|h\| + \sqrt{2}\beta \max\{\|\Delta_x\|, \mu \|\Delta_s\|\}.
\]
Consequently, under the hypotheses of the proposition,
\[
(1 - \sqrt{2}\beta) \max\{\|\Delta_x\|, \mu \|\Delta_s\|\} \leq \|h\|
\]
as required. \( \square \)

Let \( x_+ \) and \( s_+ \) denote, respectively, \( x_+ = \alpha \Delta_x \) and \( s_+ = \alpha \Delta_s \). Similarly, let \( \bar{x}_\alpha \) and \( \bar{s}_\alpha \) denote, respectively, \( v + \alpha \Delta_x = \mathcal{Q}_{\mathcal{L}^2} (x_\alpha) \) and \( e + \alpha \Delta_s = \mathcal{Q}_{\mathcal{L}^2} (s_\alpha) \). Let \( \mu_\alpha \) denote the scaled duality gap \( \frac{1}{\gamma} \langle x_\alpha, s_\alpha \rangle = \frac{1}{\gamma} \langle \bar{x}_\alpha, \bar{s}_\alpha \rangle \). 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 \( d(x_\alpha, s_\alpha; \mu_\alpha) \) 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 [5]). If \( h \in \mathcal{H} \) satisfies \( \|h\| \leq 1/2 \), then
\[
\|l_{e+h} - e\| \leq \sqrt{2} \|h\|.
\]

**Proof.** Let \( \Delta_1(t) \) denote the lower triangular element \( l_{e+th} - e \). Note that for all \( t \in \mathbb{R} \),
\[
\Delta_1(t)_H + \Delta_1(t) \Delta_1(t)^* = th.
\]
For \( t \in [0, 1] \), we have
\[
t \|h\| = \|\Delta_1(t)_H + \Delta_1(t) \Delta_1(t)^*\|
\geq \|\Delta_1(t)_H\| - \|\Delta_1(t) \Delta_1(t)^*\|
\geq \sqrt{2} \|\Delta_1(t)\| - \|\Delta_1(t)\|^2.
\]
Solving this quadratic in \( \|\Delta_1(t)\| \) gives
\[
\|\Delta_1(t)\| \leq \frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - t \|h\|} \quad \text{or} \quad \|\Delta_1(t)\| \geq \frac{1}{\sqrt{2}} + \sqrt{\frac{1}{2} - t \|h\|}.
\]
Since \( \Delta_1(t) \), whence \( \|\Delta_1(t)\| \) is continuous in \( t \), it follows that
\[
\|\Delta_1(t)\| \leq \frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - t \|h\|}
\]
whenever \( t \| h \| \leq 1/2 \). Under the hypothesis \( \| h \| \leq 1/2 \), this, indeed, holds for \( t = 1 \); thus
\[
\| l_{e+} - e \| \leq \frac{1}{\sqrt{2}} - \sqrt{\frac{1}{2} - \| h \|} \leq \frac{1}{\sqrt{2}}.
\]
Finally, applying this upper bound in (3.6) with \( t = 1 \) gives
\[
\| h \| \geq \sqrt{2} \| l_{e+} - e \| - \frac{1}{\sqrt{2}} \| l_{e+} - e \| = \frac{1}{\sqrt{2}} \| l_{e+} - e \|
\]
as required. \( \square \)

**Lemma 2.** Suppose \( \beta < 1/\sqrt{2} \) and \( \chi = \beta + (1 - \sigma)\sqrt{T} \). Then \( s_\alpha \in K^\perp \) and
\[
(3.7) \quad \mu_{+}^{-1} \left\| Q_{l_{s_\alpha}} (x_\alpha) - \mu_\alpha e \right\| \leq (1 - \alpha)\beta + \alpha^2 \chi^2(7 + 6\beta) + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2}\beta)^3}
\]
whenever \( 0 \leq \alpha \leq \min\{1, (1 - \sqrt{2}\beta)/(2\chi)\} \).
Moreover, if \( \beta \leq \beta < 1 \) with \( \sigma(\beta - \beta) \neq 0 \), then there exists \( \overline{\alpha} > 0 \) such that
(i) the cubic polynomial
\[
(3.8) \quad p : \alpha \mapsto (1 - \alpha)\beta + \alpha^2 \chi^2(7 + 6\beta) + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2}\beta)^3} - (1 - \alpha + \sigma\alpha)\overline{\beta}
\]
satisfies \( p(\alpha) \leq 0 \) for all \( \alpha \in [0, \overline{\alpha}] \); and
(ii) for all \( 0 \leq \alpha \leq \min\{1, \overline{\alpha}\} \), the pair \( (x_\alpha, s_\alpha) \) lies in \( K \oplus K^\perp \) and
\[
(3.9) \quad d(x_\alpha, s_\alpha; \mu_\alpha) \leq \frac{1}{1 - \alpha + \sigma} \left( (1 - \alpha)\beta + \alpha^2 \chi^2(7 + 6\beta) + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2}\beta)^3} \right)
\]
\[
\leq \overline{\beta}.
\]

**Proof.** According to Proposition 10, if \( \beta < 1/\sqrt{2} \), then
\[
(3.10) \quad \max \{ \| \Delta_{\bar{x}} \|, \mu_{+} \| \Delta_{\bar{s}} \| \} \leq \| \mu_{+} e - v \| < \frac{\mu_{+} e - v \|}{1 - \sqrt{2}\beta} \leq \mu_{+} \\frac{\beta + (1 - \sigma)\sqrt{T}}{1 - \sqrt{2}\beta}
\]
by triangle inequality. Thus for every \( u \in \mathbb{F}_+ \setminus \{0\} \),
\[
\langle \bar{s}_\alpha, uu^* \rangle - \langle e, uu^* \rangle + \alpha \langle \Delta_{\bar{x}}, uu^* \rangle \geq \| u \|^2 - |\alpha| \| \Delta_{\bar{x}} \| \| u \|^2 = (1 - |\alpha| \| \Delta_{\bar{x}} \|) \| u \|^2 > 0
\]
whenever \( |\alpha| < (1 - \sqrt{2}\beta)/\chi \). Thus \( \bar{s}_\alpha \in K^\perp \), whence \( s_\alpha = Q_{l_{s_\alpha}} (\bar{s}_\alpha) \in K^\perp \), for every \( \alpha \in [0, (1 - \sqrt{2}\beta)/(2\chi)] \).

Note that whenever \( s_\alpha \in K^\perp \), we have
\[
\bar{s}_\alpha = Q_{l_{s_\alpha}} (s_\alpha) = \left( l_{s_\alpha}^{-1} l_{s_\alpha} \right) \left( l_{s_\alpha}^{-1} l_{s_\alpha} \right)^*;
\]
and hence \( l_{\bar{s}_\alpha} = l_{s_\alpha}^{-1} l_{s_\alpha} \). In this case we deduce, using Corollary 1, that
\[
Q_{l_{s_\alpha}} (\bar{s}_\alpha) = Q_{l_{s_\alpha}} l_{s_\alpha} (Q_{l_{s_\alpha}} (x_\alpha)) = Q_{l_{s_\alpha}} (x_\alpha).
\]
Thus we may equivalently prove (3.7) for the pair \( (\bar{x}_\alpha, \bar{s}_\alpha) \) instead.
Let \( l_\alpha \) denote the lower triangular element \( l_{\alpha\alpha} - e = l_{\alpha\alpha} + \alpha \Delta_{\alpha} - e \), which is well-defined when \( \alpha \in [0, (1 - \sqrt{2}/(2\chi))] \). By Proposition 3,

\[
Q_{l_\alpha}^\top (\tilde{x}_\alpha) = Q_{e + l_\alpha^\top} (\tilde{x}_\alpha) \\
= ((e + l_\alpha^\top) \langle \langle \tilde{x}_\alpha \rangle \rangle (e + l_\alpha))_H \\
= \langle \langle \tilde{x}_\alpha \rangle \rangle l_\alpha + 1^\top_\alpha \langle \langle \tilde{x}_\alpha \rangle \rangle l_\alpha \\
= v + \alpha \Delta_{\tilde{x}} + (v l_\alpha)_H + \alpha(\Delta_{\tilde{x}} l_\alpha)_H + Q_{l_\alpha^\top}^\top (v) + \alpha Q_{l_\alpha^\top}^\top (\Delta_{\tilde{x}}),
\]

whence the difference \( (Q_{l_\alpha}^\top (\tilde{x}_\alpha) - \mu_\alpha e) \) is

\[
v - (1 - \alpha)\mu_\alpha e - \alpha \sigma \mu_\alpha e + \alpha \Delta_{\tilde{x}} + (v l_\alpha)_H + \alpha(\Delta_{\tilde{x}} l_\alpha)_H + Q_{l_\alpha^\top}^\top (v) + \alpha Q_{l_\alpha^\top}^\top (\Delta_{\tilde{x}}).
\]

Using (3.3), we re-express this as

\[
(1 - \alpha)(v - \mu_\alpha e) + (v l_\alpha - \alpha \Delta_{\tilde{x}})_H + \alpha(\Delta_{\tilde{x}} l_\alpha)_H \\
+ \mu_\alpha 1^\top l_\alpha + Q_{l_\alpha^\top}^\top (v - \mu_\alpha e) + \alpha Q_{l_\alpha^\top}^\top (\Delta_{\tilde{x}}) \\
= (1 - \alpha)(v - \mu_\alpha e) - \mu_\alpha 1^\top l_\alpha - (v - \mu_\alpha e)_H + \alpha(\Delta_{\tilde{x}} l_\alpha)_H \\
+ \mu_\alpha 1^\top l_\alpha + Q_{l_\alpha^\top}^\top (v - \mu_\alpha e) + \alpha Q_{l_\alpha^\top}^\top (\Delta_{\tilde{x}}).
\]

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

\[
\| l_\alpha^\top l_\alpha^\top \|, \| 1^\top_\alpha l_\alpha^\top \| \leq \| l_\alpha \| \leq 2\alpha^2 \| \Delta_{\tilde{x}} \|^2 \leq 2\alpha^2 \frac{\chi^2}{(1 - \sqrt{2}\beta)^2},
\]

\[
\| (v - \mu_\alpha e) \|_H \leq 2 \| v - \mu_\alpha e \| \| l_\alpha^\top l_\alpha^\top \| \leq \sqrt{2} \| v - \mu_\alpha e \| \| l_\alpha^\top l_\alpha^\top \| \leq 2\sqrt{2}\alpha^2 \mu_\alpha \beta \frac{\chi^2}{(1 - \sqrt{2}\beta)^2} \leq 3\alpha^2 \mu_\alpha \beta \frac{\chi^2}{(1 - \sqrt{2}\beta)^2}.
\]

\[
\| \alpha(\Delta_{\tilde{x}} l_\alpha)_H \| \leq 2\alpha \| \Delta_{\tilde{x}} \| \| l_\alpha \| \leq 2\sqrt{2}\alpha^2 \mu_\alpha \beta \frac{\chi^2}{(1 - \sqrt{2}\beta)^2} \leq 3\alpha^2 \mu_\alpha \beta \frac{\chi^2}{(1 - \sqrt{2}\beta)^2},
\]

\[
\| Q_{l_\alpha^\top}^\top (v - \mu_\alpha e) \| \leq 2 \| (v - \mu_\alpha e) \| \| l_\alpha \| \leq 2\sqrt{2}\alpha^2 \mu_\alpha \beta \frac{\chi^2}{(1 - \sqrt{2}\beta)^2} \leq 3\alpha^2 \mu_\alpha \beta \frac{\chi^2}{(1 - \sqrt{2}\beta)^2},
\]

and

\[
\| \alpha Q_{l_\alpha^\top}^\top (\Delta_{\tilde{x}}) \| \leq 2\alpha \| (\Delta_{\tilde{x}}) \| \| l_\alpha \| \leq 2\sqrt{2}\alpha^2 \mu_\alpha \beta \frac{\chi^3}{(1 - \sqrt{2}\beta)^3} \leq 3\alpha^2 \mu_\alpha \beta \frac{\chi^3}{(1 - \sqrt{2}\beta)^3}.
\]

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

(i) If \( \beta > \beta_H \), then \( p(0) = \beta - \beta_H < 0 \) shows that \( p \) has at least one positive real root. If not, then \( \sigma > 0 \), \( p(0) = 0 \) and \( p'(0) = \beta - \sigma \beta - \beta = -\sigma \beta < 0 \) also shows the same. Hence \( p \) has a smallest positive real root \( \beta_H \) and \( p(\alpha) \leq 0 \) for all \( \alpha \in [0, \beta_H] \).
(ii) If \((1 - \sqrt{2}/\sqrt{2})/2\chi < 1\), then for \(\alpha = (1 - \sqrt{2}/\sqrt{2})/2\chi \in (0, 1)\), we have
\[
p(\alpha) > 7a^2 \frac{\chi^2}{(1 - \sqrt{2}/\sqrt{2})^2} - \beta = \frac{7}{4} - \beta > 0.
\]
Hence \(0 \leq \alpha \leq \min\{1, (1 - \sqrt{2}/\sqrt{2})/2\chi\}\) whenever \(0 \leq \alpha \leq \min\{1, \nu\}\). Subsequently \(s_\alpha \in K^2\) and (3.7) holds whenever \(0 \leq \alpha \leq \min\{1, \nu\}\). This leads to
\[
\|Q_{\nu s_\alpha}(x_\alpha) - \mu_\alpha e\| \leq p(\alpha) + (1 - \alpha + \alpha\sigma)\mu_+ \beta
\leq (1 - \alpha + \alpha\sigma)\mu_+ \beta = \mu_\alpha \beta,
\]
whence
\[
\langle Q_{\nu s_\alpha}(x_\alpha), l^+ \rangle = \mu_\alpha \langle e, l^+ \rangle + \langle Q_{\nu s_\alpha}(x_\alpha) - \mu_\alpha e, l^+ \rangle
\geq \mu_\alpha \|l^+\|^2 - \|Q_{\nu s_\alpha}(x_\alpha) - \mu_\alpha e\| \|l^+\|^2
\geq \mu_\alpha (1 - \beta) \|l^+\|^2 > 0,
\]
for all \(l \in \mathcal{T}_+ \setminus \{0\}\). We then conclude that \(Q_{\nu s_\alpha}(x_\alpha) \in K\), whence \(x_\alpha \in K\). Finally (3.9) follows directly from (3.7). \(\Box\)

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 \((x_{in}, s_{in}) \in K \oplus K^2\) and \(\mu_{in} \in R^+\) with \(d(x_{in}, s_{in}, \mu_{in}) \leq \beta\) for some \(\beta \in (0, 1/\sqrt{2})\), and the required accuracy \(\varepsilon > 0\).*

1. Set \((x_+, s_+) = (x_{in}, s_{in})\) and \(\mu_+ = \mu_{in}\).
2. While \(\langle x_+, s_+ \rangle > \varepsilon \langle x_{in}, s_{in} \rangle\),
   (a) Pick \(\beta \in (0, 1/\sqrt{2})\) and \(\sigma \in [0, 1]\).
   (b) Solve (3.2) with \(\mu_{++}\) replaced by \(\sigma \mu_+\). For each \(\alpha \in [0, 1]\), let \((x_\alpha, s_\alpha) = (x_+, \alpha \Delta x, s_+ + \alpha \Delta s)\), and let \(\mu_\alpha = (1 - \alpha + \alpha \sigma)\mu_+\). Pick \(\alpha \in [0, 1]\) such that \((x_\alpha, s_\alpha) \in K \oplus K^2\) and
   \[
d_2(x_\alpha, s_\alpha; \mu_\alpha) \leq \beta.
\]
   (c) Update \((x_+, s_+) \leftarrow (x_\alpha, s_\alpha)\) and \(\mu_+ \leftarrow \mu_\alpha\).
3. Output \((x_{out}, s_{out}) = (x_+, s_+)\).

4.1. Short-step algorithm. In our short-step algorithm, conservative updates of the parameter \(\mu\) are used and full Newton steps are taken. The target parameter \(\mu_{++}\) is chosen so that the ratio \(\sigma = \mu_{++}/\mu_+\) is \((1 - \delta/\sqrt{2})\), where \(\delta \in (0, 1)\) is a constant. We shall show that with appropriate choices of \(\beta\) and \(\delta\), 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
\[
\frac{(\beta + \delta)^2(7 + 6\beta)}{(1 - \sqrt{2}/\sqrt{2})^2} + \frac{3(\beta + \delta)^3}{(1 - \sqrt{2}/\sqrt{2})^3} < \left(1 - \frac{\delta}{\sqrt{2}}\right) \beta,
\]
then we may use \(\alpha = 1\) in each iteration of Algorithm 1 with \(\sigma = 1 - \delta/\sqrt{2}\) and \(\beta = \beta\). Moreover, the algorithm terminates after \(O(\sqrt{r} \log \frac{1}{\delta})\) iterations.
Proof. Consider an iteration of the algorithm. By the choice of \( \tilde{\beta} \), we have 
\[
d(x_+, s_+; \mu_+) \leq \beta \text{ at the beginning of the iteration. Let } \pi \text{ be the smallest positive root of the cubic polynomial (3.8) with } \sigma = (1 - \delta/\sqrt{r}), \chi = \beta + (1 - \sigma)\sqrt{r} = \beta + \delta, \text{ and } \bar{\beta} = \beta. \]
Under the hypothesis (4.1), we have
\[
(1 - \alpha)\beta + \alpha^2 \frac{\beta^2(7 + 6\beta)}{(1 - \sqrt{2}\beta)^2} + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2}\beta)^3} - (1 - \alpha + \sigma\alpha)\bar{\beta}
\leq \alpha \left[ \frac{(\beta + \delta)^2(7 + 6\beta)}{(1 - \sqrt{2}\beta)^2} + \frac{3(\beta + \delta)^3}{(1 - \sqrt{2}\beta)^3} - \left(1 - \frac{\delta}{\sqrt{r}}\right)\bar{\beta} \right] < 0
\]
for all \( 0 < \alpha \leq 1 \), whence \( \pi > 1 \). Thus by Lemma 2, we have \((x_\alpha, s_\alpha) \in K \oplus K^2\) and 
\[
d(x_\alpha, s_\alpha; \mu_\alpha) \leq \beta \text{ for all } \alpha \in [0, 1]. \]
This means we may take \( \tilde{\alpha} = 1 \) in the iteration. Finally, the duality gap reduces by the factor \((1 - \delta/\sqrt{r})\) in each iteration, whence by a factor of \( \varepsilon \) in \( O(\sqrt{r} \log \frac{1}{\varepsilon}) \) iterations. \( \square \)

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 \( \tilde{\alpha} = \Omega(1/r) \) in Algorithm 1 with \( \beta = \beta \) 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 the choice of \( \tilde{\beta} \), we have 
\[
d(x_+, s_+; \mu_+) \leq \beta \text{ at the beginning of the iteration. Let } \pi \text{ be the smallest positive root of } \alpha \mapsto -\sigma\alpha + \alpha^2 \frac{(\beta + (1 - \sigma)\sqrt{r})^2(7 + 6\beta)}{(1 - \sqrt{2}\beta)^2} + \alpha^3 \frac{3(\beta + (1 - \sigma)\sqrt{r})^3}{(1 - \sqrt{2}\beta)^3}
\]
\[
= \left( -\sigma\beta + \alpha \frac{(\beta + (1 - \sigma)\sqrt{r})^2(7 + 6\beta)}{(1 - \sqrt{2}\beta)^2} + \alpha^2 \frac{3(\beta + (1 - \sigma)\sqrt{r})^3}{(1 - \sqrt{2}\beta)^3} \right) \alpha.
\]
By Lemma 2, we have \((x_\alpha, s_\alpha) \in K \oplus K^2\) and 
\[
d(x_\alpha, s_\alpha; \mu_\alpha) \leq \beta = \bar{\beta}
\]
whenever \( 0 \leq \alpha \leq \min\{1, \pi\} \). Since \( \pi = \Omega(1/r) \), we may use \( \tilde{\alpha} = \min\{1, \pi\} = \Omega(1/r) \). Finally, the duality gap reduces by the constant factor \((1 - \Omega(1/r))\) every two iterations, whence by a factor of \( \varepsilon \) in \( O(r \log \frac{1}{\varepsilon}) \) iterations. \( \square \)

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 \( \beta \) and \( \bar{\beta} \), 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
\[
4\beta^2(7 + 12\beta) + \frac{24\beta^3}{(1 - 2\sqrt{2}\beta)^2} \leq \beta,
\]

(A T-ALGEBRAIC APPROACH TO PRIMAL-DUAL ALGORITHMS)
then by alternating between \((\sigma, \tilde{\beta}) = (0, 2\beta)\) and \((\sigma, \tilde{\beta}) = (1, \beta)\) in Algorithm 1, we may use \(\alpha = \Omega(1/\sqrt{\tau})\) and \(\alpha = 1\), respectively. Moreover, the algorithm terminates after \(O(\sqrt{\tau} \log \frac{1}{\varepsilon})\) iterations.

Proof. Consider an iteration of the algorithm where a predictor step is to be taken. By choice of \(\beta\) in the previous iteration, we have \(d(x_+, s_+; \mu_+) \leq \beta\) at the beginning of the iteration. Let \(\tau\) be the smallest positive root of the cubic polynomial (3.8) with \(\sigma = 0\), \(\chi = \beta + (1 - \sigma)\sqrt{\tau} = \beta + \sqrt{\tau}\), and \(\beta = 2\beta = \Omega(\beta, 1)\); i.e., \(\tau\) is the smallest positive root of

\[
\alpha \mapsto -\beta + \alpha \beta + \alpha^2 \frac{(\beta + \sqrt{\tau})^2(7 + 6\beta)}{(1 - \sqrt{2\beta})^2} + \alpha^3 \frac{3(\beta + \sqrt{\tau})^3}{(1 - \sqrt{2\beta})^3}.
\]

By Lemma 2, we have \((x_\alpha, s_\alpha) \in K \oplus K^2\) and

\[
d(x_\alpha, s_\alpha; \mu_\alpha) \leq 2\beta = \tilde{\beta}
\]

whenever \(0 \leq \alpha \leq \min\{1, \tau\}\). Since \(\tau = \Omega(1/\sqrt{\tau})\), we may use \(\alpha = \min\{1, \tau\} = \Omega(1/\sqrt{\tau})\).

Now consider an iteration of the algorithm where a corrector step is to be taken. By choice of \(\beta\) in the previous iteration, we have \(d(x_+, s_+; \mu_+) \leq \beta\) at the beginning of the iteration. Under the hypothesis (4.2), for \(\sigma = 1\), \(\chi = (2\beta) + (1 - \sigma)\sqrt{\tau} = 2\beta\), and \(\beta = 2\beta\), we have

\[
(1 - \alpha)(2\beta) + \alpha^2 \frac{\chi^2(7 + 6(2\beta))}{(1 - \sqrt{2\beta})^2} + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2\beta})^3} - (1 - \alpha + \sigma\alpha)\beta
\]

\[
\leq \alpha \left[ \frac{4\beta^2(7 + 12\beta)}{(1 - 2\sqrt{2\beta})^2} + \frac{24\beta^3}{(1 - 2\sqrt{2\beta})^3} - 2\beta \right] < 0
\]

for all \(0 < \alpha \leq 1\), whence by Lemma 2, we have \((x_\alpha, s_\alpha) \in K \oplus K^2\) and

\[
d(x_\alpha, s_\alpha; \mu_\alpha) \leq \frac{1}{1 - \alpha + \alpha\sigma} \left( (1 - \alpha)(2\beta) + \alpha^2 \frac{\chi^2(7 + 6(2\beta))}{(1 - \sqrt{2\beta})^2} + \alpha^3 \frac{3\chi^3}{(1 - \sqrt{2\beta})^3} \right)
\]

\[
\leq \left( 2(1 - \alpha)\beta + \alpha \left( \frac{4\beta^2(7 + 12\beta)}{(1 - 2\sqrt{2\beta})^2} + \frac{24\beta^3}{(1 - 2\sqrt{2\beta})^3} \right) \right)
\]

\[
\leq 2\beta - \alpha\beta
\]

for all \(\alpha \in [0, 1]\), in particular \(d(x_1, s_1; \mu_1) \leq \beta\). This means we may take \(\hat{\alpha} = 1\) in the iteration.

Finally, the duality gap reduces by the constant factor \((1 - \Omega(1/\sqrt{\tau}))\) every two iterations, whence by a factor of \(\varepsilon\) in \(O(\sqrt{\tau} \log \frac{1}{\varepsilon})\) 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 the existing bounds for their counterparts 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 other 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 $ℓ_2$-neighborhood. A first step towards designing a wide-neighborhood 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.

Acknowledgment. The author thanks the anonymous referees for their invaluable suggestions on the overall presentation of this paper.

REFERENCES


