McMaster University

Advanced Optimization Laboratory

Title:

Exact duality for
optimization over symmetric cones

Author:

Imre Pólik, Tamás Terlaky

AdvOL-Report No. 2007/10

August, 2007
Hamilton, Ontario, Canada

Noname manuscript No.
(will be inserted by the editor)

Exact duality for optimization over symmetric cones

Imre Pólik · **Tamás Terlaky**

Received: August 22, 2007 / Accepted:

Abstract We present a strong duality theory for optimization problems over symmetric cones without assuming any constraint qualification. We show important complexity implications of the result to semidefinite and second order conic optimization. The result is an application of Borwein and Wolkowicz's facial reduction procedure to express the minimal cone. We use Pataki's simplified analysis and provide an explicit formulation for the minimal cone of a symmetric cone optimization problem. In the special case of semidefinite optimization the dual has better complexity than Ramana's strong semidefinite dual. After specializing the dual for second order cone optimization we argue that new software for homogeneous cone optimization problems should be developed.

Keywords Ramana-dual · symmetric cone optimization · exact duality · minimal cone · homogeneous cones

1 Introduction

1.1 Historical background

In his seminal paper [21], Ramana presented an exact duality theory for semidefinite optimization without any constraint qualification. Later, in a joint paper [23] with Tunçel and Wolkowicz, they showed that the result can be derived from a more general theorem for convex problems [6]. We had hoped to construct a purely second order conic strong dual for second order conc optimization problems. Our continued failure to do so motivated us to look for more general classes of cones on the dual side, this is how we found the class of homogeneous cones. They are quite general, yet possess all the properties we need for the construction of the dual. The primal problem is defined over a symmetric (i.e., homogeneous and self dual) cone, but the strong dual problem will involve non self dual homogeneous cones. We use the general theory of homogeneous cones during the construction.

The discussion uses Pataki's simplified analysis of the facial reduction algorithm for general convex conic optimization [18].

The authors were supported by the Canada Research Chairs Program, NSERC (Discovery Grant #5-48923), and MITACS.

McMaster University, School of Computational Engineering and Science, Hamilton, ON, Canada

1.2 Notations

Let us briefly review the notation we are using in the paper. The set of nonnegative numbers is denoted by $\mathbb{R}_+$. Vectors throughout the paper are denoted by lowercase letters and are assumed to be column vectors. If v is a vector then v_i is its i^{th} component, and $v_{i:j}$ is the subvector consisting of the components from the i^{th} to (and including) the j^{th}. If $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator then its null space (kernel) is denoted by $\operatorname{Ker}(A)$. We will use $\langle u, v \rangle$ to denote the scalar product of u and v, which in some special cases will be $u^T v$. If $V \subseteq \mathbb{R}^n$ is a subspace then its orthogonal complement is $V^\perp = \{u \in \mathbb{R}^n : \langle u, v \rangle = 0, \forall v \in V\}$. For a set $\mathscr{H}$, $\operatorname{int}(\mathscr{H})$ denotes the interior of the set, $\operatorname{rel\,int}(\mathscr{H})$ its relative interior and $\operatorname{cl}(\mathscr{H})$ its closure.

Let $\mathbb{S}^{n\times n}$ denote the set of real symmetric matrices, $\mathbb{S}_+^{n\times n}$ the set of symmetric positive semidefinite matrices and $\mathbb{S}_{++}^{n\times n}$ the set of symmetric positive definite matrices. We will write $x \succeq y$ ($x \succ y$) to denote that $x - y$ is positive semidefinite (definite). The Lorentz or second order cone in standard form is defined as $\mathbb{L} = \{x \in \mathbb{R}^n : x_1 \geq \|x_{2:n}\|_2\}$ and in rotated form it is $\mathbb{L}_\mathrm{r} = \left\{x \in \mathbb{R}^n : x_1 x_2 \geq \|x_{3:n}\|_2^2, x_1, x_2 \geq 0\right\}$.

1.3 The structure of the paper

First, in §2 we give a quick overview of classical Lagrange duality theory for convex optimization and discuss the need for constraint qualifications. We then present the facial reduction algorithm of Borwein and Wolkowicz with Pataki's simplified analysis. In §3 we introduce homogeneous cones, discuss their constructions and basic properties. Section 4 contains our new results, the application of the facial reduction algorithm to the symmetric cone optimization problem. We prove that symmetric cones satisfy Pataki's sufficient conditions and give an explicit formula for the cones in the dual problem. In §5 we specialize the result to semidefinite and second order conic optimization. We show that in the semidefinite case our dual has better complexity than Ramana's strong semidefinite dual. Final conclusions are drawn and new research directions are presented in §6.

2 Classical Lagrange duality in conic optimization

In this section we present a brief overview of some important concepts in convex optimization. See [15, 25, 27] for more details.

2.1 Foundations

A set $\mathscr{K}$ is a cone if for any $\lambda \geq 0$, $\lambda \mathscr{K} \subseteq \mathscr{K}$. A cone is convex, if $\mathscr{K} + \mathscr{K} = \mathscr{K}$. A cone is pointed if it does not contain a line, and it is solid if its interior is not empty. If $\mathscr{K} \subseteq \mathbb{R}^n$ is a closed convex cone then we can define its dual cone:

$$\mathscr{K}^* = \{x \in \mathbb{R}^n : \langle x, y \rangle \geq 0, \forall y \in \mathscr{K}\}. \tag{2.1}$$

If $\mathscr{K}$ is a pointed cone then its dual is solid, and vice versa. If $\mathscr{K}$ is a closed cone then $\mathscr{K}^{**} = \mathscr{K}$.

Let $\mathscr{K} \subset \mathbb{R}^n$ be a closed, convex cone. The optimization problem in the focus of our paper is defined as follows:

$$\begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\in \mathscr{K}, \end{aligned} \tag{P}$$

where $A \in \mathbb{R}^{m\times n}$, $c,s \in \mathbb{R}^n$, $b,y \in \mathbb{R}^m$. The importance of this conic form comes from the result that under mild assumptions any convex optimization problem can be written this way, see [16] for details. The primal problem (P) has a corresponding Lagrange-Slater dual (see [25,27]):

$$\begin{aligned} \min\ & \langle c,x\rangle \\ Ax &= b \\ x &\in \mathscr{K}^*, \end{aligned} \tag{D}$$

where $x \in \mathbb{R}^n$. This dual problem supplies critical information about (P). The weak duality theorem holds without any additional assumption:

Theorem 2.1 (Weak duality) *If y, s and x are feasible solutions of* (P) *and* (D) *then* $\langle x,s\rangle = \langle c,x\rangle - \langle b,y\rangle \geq 0$, *this quantity is the duality gap. Moreover, if the duality gap is zero then y, s and x are optimal solutions of* (P) *and* (D).

The most natural question about the weak duality theorem is whether the last condition is necessary for optimality. The answer in general is negative, we need a regularity condition, also called a constraint qualification (CQ) for a stronger result. In this paper we use the most common CQ, the Slater-condition:[1]

Definition 2.1 (Slater condition) Problem (P) satisfies the Slater condition if it has a relatively strictly feasible solution, i.e., $s \in \operatorname{rel\,int}(\mathscr{K})$, $A^T y + s = c$. Similarly, for the dual problem (D) this translates to $x \in \operatorname{rel\,int}(\mathscr{K}^*)$, $Ax = b$.

Theorem 2.2 (Strong duality) *If either* (P) *or* (D) *satisfies the Slater condition then the other problem is solvable and the optimal objective values are equal. If both problems are solvable, then the duality gap is zero for y, s and x if and only if they are optimal solutions for* (P) *and* (D).

Example 2.1 (Lack of strong duality) The following problem is from [16]:

$$\begin{aligned} \max\ & u_2 \\ \begin{pmatrix} u_2 & 0 & 0 \\ 0 & u_1 & u_2 \\ 0 & u_2 & 0 \end{pmatrix} &\preceq \begin{pmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{2.2}$$

[1] The classical definition of the Slater condition requires a point from the interior of the feasible set thus excludes cones with empty interior. This weaker condition is from [15].

For any feasible solution of this problem we have $u_2 = 0$ so the optimal objective value is 0. However, the dual problem

$$\begin{aligned} \min\ \alpha v_{11} \\ v_{22} &= 0 \\ v_{11} + 2v_{23} &= 1 \\ \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} &\succeq 0 \end{aligned} \tag{2.3}$$

has $v_{22} = v_{23} = 0$ and $v_{11} = 1$, thus the optimal value is α. This shows that the duality gap can be arbitrarily large. The set of feasible solutions for the primal problem is $u_1 \leq 0, u_2 = 0$, so there is no strictly feasible primal solution.

Intuitively, this example shows that if the set of feasible solutions is entirely on the boundary of the cone $\mathscr{K}$ then the problem might not have a strong Lagrange dual. One possible solution is to replace $\mathscr{K}$ in (P) with a smaller cone without changing the feasible set. This is the topic of the next section.

2.2 The minimal cone and the facial reduction algorithm

Let $\mathscr{K}$ be a closed, convex cone, then a closed set $\mathscr{F} \subseteq \mathscr{K}$ is by definition a *face* of $\mathscr{K}$ if for every $x, y \in \mathscr{K}$, $x + y \in \mathscr{F}$ implies $x, y \in \mathscr{F}$. This relation is denoted by $\mathscr{F} \trianglelefteq \mathscr{K}$. If $\mathscr{C} \subseteq \mathscr{K}$ is a subset of $\mathscr{K}$ then $\mathscr{F}(\mathscr{C})$ denotes the face generated by $\mathscr{C}$ in $\mathscr{K}$, i.e., the smallest face of $\mathscr{K}$ containing $\mathscr{C}$. This is given as

$$\mathscr{F}(\mathscr{C}) = \bigcap \{\mathscr{F} : \mathscr{C} \subseteq \mathscr{F} \trianglelefteq \mathscr{K}\}. \tag{2.4}$$

Since the intersection of faces is a face, $\mathscr{F}(\mathscr{C})$ is indeed a face and its minimality is due to the construction. Let $\mathscr{F}$ be a face of $\mathscr{K}$, then its complementary (or conjugate) face is defined as $\mathscr{F}^{\mathrm{c}} = \mathscr{F}^{\perp} \cap \mathscr{K}^* \trianglelefteq \mathscr{K}^*$. A face and its complementary face belong to the primal and dual cones and are orthogonal. If $\mathscr{F} \trianglelefteq \mathscr{K}^*$ then we use the same notation for $\mathscr{F}^{\mathrm{c}} = \mathscr{F}^{\perp} \cap \mathscr{K}$. This ambiguity will not cause a problem as $\mathscr{F}$ always determines the appropriate cone.

For the optimization problem (P) we can define a corresponding *minimal cone* (see [6, 21,23]), which is the smallest face of $\mathscr{K}$ containing all the feasible solutions s of (P).

$$\mathscr{K}_{\min} = \mathscr{F}\left(\left\{c - A^T y : y \in \mathbb{R}^m\right\} \cap \mathscr{K}\right). \tag{2.5}$$

Replacing cone $\mathscr{K}$ with $\mathscr{K}_{\min}$ in (P) we get an equivalent problem:

$$\begin{aligned} \max\ b^T y \\ A^T y + s &= c \\ s &\in \mathscr{K}_{\min}. \end{aligned} \tag{P$_{\min}$}$$

Moreover, due to the minimality of $\mathscr{K}_{\min}$ this new problem satisfies the Slater condition,[2] thus it is in strong duality with its Lagrange dual:

$$\begin{aligned} \min\ & \langle c,x\rangle \\ Ax &= b \\ x &\in \mathscr{K}_{\min}^*. \end{aligned} \tag{D$_{\min}$}$$

This abstract dual is quite useless unless we can express either the primal minimal cone or its dual explicitly. One algorithmic description is given by Borwein and Wolkowicz in [6]. The following simplified form is due to Pataki [18]. First let us define an interesting set of cones:

Definition 2.2 (Nice cone) A convex cone $\mathscr{K}$ is *nice* if $\mathscr{F}^* = \mathscr{K}^* \oplus \mathscr{F}^\perp$ for all faces $\mathscr{F} \trianglelefteq \mathscr{K}$. Equivalently, since $\mathscr{F}^* \subseteq \operatorname{cl}\left(\mathscr{K}^* \oplus \mathscr{F}^\perp\right)$, $\mathscr{K}$ is nice if $\mathscr{K}^* \oplus \mathscr{F}^\perp$ is closed.

Most of the cones arising in practical applications (polyhedral, semidefinite and Lorentz cones) are nice. Now consider the following system:

$$\begin{aligned} A(u+v) &= 0 \\ \langle c, u+v\rangle &= 0 \\ (u,v) &\in \mathscr{K}^* \times \mathscr{F}^\perp. \end{aligned} \tag{FR}$$

The next lemma was proved by Borwein and Wolkowicz in [6].

Lemma 2.1 (Facial reduction) *If $\mathscr{K}_{\min} \trianglelefteq \mathscr{F} \trianglelefteq \mathscr{K}$ then for any solution u,v of system* (FR) *we have $\mathscr{K}_{\min} \subseteq \{u\}^\perp \cap \mathscr{F} \trianglelefteq \mathscr{F}$. If $\mathscr{K}_{\min} \subsetneq \mathscr{F}$ then there is a solution (u,v) such that $\mathscr{F} \cap \{u\}^\perp \subsetneq \mathscr{F}$. In this case we say that u is a reducing certificate for the system $c - A^T y \in \mathscr{F}$.*

Proof As $0 = \left\langle u+v, (A^T y - c)\right\rangle = \left\langle u, (A^T y - c)\right\rangle$ for all values of y, we get that $\{u\}^\perp \supseteq \mathscr{K}_{\min}$. Moreover, if $x,y \in \mathscr{F}$ and $x+y \in \{u\}^\perp \cap \mathscr{F}$ then $\langle u,(x+y)\rangle = 0$, but since $u \in \mathscr{F}^*$, both $\langle u,x\rangle$ and $\langle u,y\rangle$ are nonnegative, thus they are both 0. This shows that $x,y \in \{u\}^\perp \cap \mathscr{F}$, thus $\{u\}^\perp \cap \mathscr{F}$ is a face of $\mathscr{F}$.

To prove the second statement let us choose an $f \in \operatorname{rel\,int}(\mathscr{F})$, then $\mathscr{K}_{\min} \neq \mathscr{F}$ if and only if $c - A^T y - \alpha f \in \mathscr{F}$ implies $\alpha \leq 0$. This system is strictly feasible, thus there is a certificate for this implication, i.e., $\mathscr{K}_{\min} \neq \mathscr{F}$ if and only if

$$\begin{aligned} \exists x &\in \mathscr{F}^* \\ Ax &= 0 \\ \langle c,x\rangle &\leq 0 \\ \langle f,x\rangle &= 1, \end{aligned} \tag{2.6}$$

which (since $\mathscr{K}$ is a nice cone) is further equivalent to

$$\begin{aligned} \exists (u,v) &\in \mathscr{K}^* \times \mathscr{F}^\perp \\ A(u+v) &= 0 \\ \langle c, u+v\rangle &\leq 0 \\ \langle f, u+v\rangle &= 1. \end{aligned} \tag{2.7}$$

[2] This is true because we use a relaxed version of the Slater condition, see Definition 2.1.

Notice that we must have $\langle c,x\rangle = 0$ otherwise (P) would be infeasible by the convex Farkas theorem ([27], §6.10). Further, $\langle f,v\rangle = 0$, since $f \in \mathscr{F}$ and $v \in \mathscr{F}^{\perp}$. This implies that $\langle f,u\rangle = 1$, thus $u \notin \mathscr{F}^{\perp}$, which translates to $\{u\}^{\perp} \not\supseteq \mathscr{F}$, proving that $\mathscr{F} \cap \{u\}^{\perp} \subsetneq \mathscr{F}$. □

We can turn the result into an algorithm to construct $\mathscr{K}_{\min}$ by repeatedly intersecting $\mathscr{K}$ with subspaces until we arrive at $\mathscr{K}_{\min}$. This is the idea behind the dual construction of Borwein and Wolkowicz in [6]. We use a simplified form, see Algorithm 1.

Algorithm 1 The facial reduction algorithm

Input: A, c and $\mathscr{K}$
  Set $(u^0,v^0) = (0,0)$, $\mathscr{F}_0 = \mathscr{K}$, $i = 0$
  while $\mathscr{K}_{\min} \neq \mathscr{F}_i$
    Find (u^{i+1},v^{i+1}) satisfying

$$\begin{aligned} A(u+v) &= 0 \\ \langle c,u+v\rangle &= 0 \\ (u,v) &\in \mathscr{K}^* \times (\mathscr{F}_i)^{\perp}. \end{aligned} \tag{FR$_i$}$$

    Set $\mathscr{F}_{i+1} = \mathscr{F}_i \cap \left\{u^{i+1}\right\}^{\perp}$, $i = i+1$
  end while
Output: $\ell \geq 0$, $u^0, u^1, \ldots, u^{\ell} \in \mathscr{K}^*$ such that $\mathscr{K}_{\min} = \mathscr{K} \cap \left\{u^0 + u^1 + \cdots + u^{\ell}\right\}^{\perp}$

The correctness of this algorithm was proved by Pataki in [18], see also [19,20] for details.

Theorem 2.3 (Facial reduction algorithm for nice cones) *If $\mathscr{K}$ is a nice cone then*

1. *During the algorithm ($i = 0,\ldots,\ell$):*
   (a) *$\mathscr{K}_{\min} \subseteq \mathscr{F}_i$ and*
   (b) *$\mathscr{F}_i = \mathscr{F}_{i-1} \cap \left\{u^i\right\}^{\perp} = \mathscr{K} \cap \left\{u^0 + \cdots + u^i\right\}^{\perp} = \left(\mathscr{F}\left(u^0 + \cdots + u^i\right)\right)^{\mathrm{c}}$.*
2. *The facial reduction algorithm is finite and it returns*

$$\mathscr{K}_{\min} = \mathscr{F}_{\ell} = \mathscr{K} \cap \left\{u^0 + \cdots + u^{\ell}\right\}^{\perp} = \left(\mathscr{F}\left(u^0 + \cdots + u^{\ell}\right)\right)^{\mathrm{c}}. \tag{2.8}$$

3. *The number of iterations is $\ell \leq L$, where*

$$L = \min\left\{\dim\left(\operatorname{Ker}(A) \cap \{c\}^{\perp}\right), \textit{ length of the longest chain of faces in } \mathscr{K}\right\}. \tag{2.9}$$

In general, this algorithm is not a viable method to express the minimal cone, since the solution of the auxiliary system is comparable to solving the original problem. The significance of the algorithm is its finiteness: this implies that by combining all the auxiliary systems (FR$_i$) from the iterations we can construct an extended strong dual problem for (P), assuming $\mathscr{K}$ is a nice cone.

$$
\begin{aligned}
\min\ & \left\langle c, \left(u^{L+1}+v^{L+1}\right)\right\rangle \\
A\left(u^{L+1}+v^{L+1}\right) &= b \\
A\left(u^{i}+v^{i}\right) &= 0,\ i=1,\dots,L \\
\left\langle c, \left(u^{i}+v^{i}\right)\right\rangle &= 0,\ i=1,\dots,L \\
u^0, v^0 &= 0 \\
u^i &\in \mathscr{K}^*,\ i=1,\dots,L+1 \\
v^i &\in \mathscr{F}_{i-1}^{\perp},\ i=1,\dots,L+1.
\end{aligned}
\tag{ED$_{\text{nice}}$}
$$

Our original dual variable would only be $x = u^{L+1}+v^{L+1}$, the additional variables are needed to describe the dual of the minimal cone. Unfortunately, this dual is useful only if we can give an explicit expression for $\mathscr{F}_i^{\perp}$. For the case when $\mathscr{K}$ is the cone of positive semidefinite matrices then this description is given in [23]. In [18], Pataki poses the open problem to find a broader class of cones for which such a description is possible. Later we will show that the theory extends to symmetric cones and we also show a way to extend the algorithm to general homogeneous cones.

3 The geometry of homogeneous cones

3.1 Definition, basic properties

The central objects of this paper are the homogeneous cones:

Definition 3.1 A closed, convex cone $\mathscr{K} \subseteq \mathbb{R}^n$ with nonempty interior is *homogeneous* if for any $u, v \in \operatorname{int}(\mathscr{K})$ there exists an invertible linear map $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ such that

1. $\varphi(\mathscr{K}) = \mathscr{K}$, i.e., φ is an automorphism of $\mathscr{K}$, and
2. $\varphi(u) = v$.

In other words, the group of automorphisms of $\mathscr{K}$ acts transitively on the interior of $\mathscr{K}$.

Typical examples of homogeneous cones are polyhedral cones in $\mathbb{R}^n$ with exactly n extreme rays (e.g., the nonnegative orthant), the set of symmetric positive semidefinite matrices and the second order or Lorentz cone.

It is important to note here that if $\mathscr{K}$ is homogeneous then so is its dual, $\mathscr{K}^*$. A cone is self-dual if $\mathscr{K} = \mathscr{K}^*$; a typical homogeneous cone is not self-dual. Self-dual homogeneous cones are called symmetric, and they are very special, see [10] for more details. We have to note here that the usual approach to symmetric cones is to use Jordan algebras. In our case it will be very important that the cone is homogeneous and therefore T-algebras provide a more powerful analysis. For a classical text on symmetric cones and T-algebras see [29].

3.2 Constructing homogeneous cones from T-algebras

Definition 3.1 does not give too much insight into the structure of homogeneous cones. Fortunately, there is a constructive way to build homogeneous cones using generalized matrices. We briefly summarize the necessary properties and techniques based on [8].

A generalized matrix is a matrix-like structure whose elements are vectors. Formally, let r be a positive integer, $1 \le i, j \le r$, and let $\mathscr{A}_{ij} \subseteq \mathbb{R}^{n_{ij}}$ be a vector space of dimension n_{ij}. Let us assume that for every $1 \le i, j, k, \ell \le r$ we have a bilinear product such that if $a_{ij} \in \mathscr{A}_{ij}$ and $a_{k\ell} \in \mathscr{A}_{k\ell}$ then

$$a_{ij} a_{k\ell} \in \begin{cases} \mathscr{A}_{i\ell}, & \text{if } j = k, \\ \{0\}, & \text{if } j \neq k. \end{cases} \tag{3.1}$$

Now consider $\mathscr{A} = \bigoplus_{i,j=1}^{r} \mathscr{A}_{ij}$, an algebra of rank r. Every element $a \in \mathscr{A}$ is a generalized matrix, while a_{ij} is an n_{ij}-dimensional generalized element of the matrix, $a_{ij} \in \mathscr{A}_{ij}$. In other words, a_{ij} is the projection of a onto $\mathscr{A}_{ij}$. The multiplication of two elements $a, b \in \mathscr{A}$ is analogous to the multiplication of matrices, i.e.,

$$(ab)_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}. \tag{3.2}$$

An involution * of the matrix algebra $\mathscr{A}$ is a linear mapping such that for every $a, b \in \mathscr{A}$ and $1 \le i, j \le r$

1. $^* : \mathscr{A} \to \mathscr{A}$,
2. $a^{**} = a$,
3. $(ab)^* = b^* a^*$,
4. $\mathscr{A}_{ij}^* = \mathscr{A}_{ji}$,
5. $(a^*)_{ij} = (a_{ij})^*$.

This generalizes the classical notion of the (conjugate) transpose. From now on we assume that our matrix algebra is equipped with an involution. Let

$$\mathscr{T} = \bigoplus_{i \le j} \mathscr{A}_{ij} \tag{3.3}$$

be the set of upper triangular elements, and let

$$\mathscr{H} = \{a \in \mathscr{A} : a = a^*\} \tag{3.4}$$

be the set of Hermitian elements. Assume that for every i, $\mathscr{A}_{ii}$ is isomorphic to $\mathbb{R}$, let ρ_i be the isomorphism and let e_i denote the unit element of $\mathscr{A}_{ii}$. We define the trace of an element[3] as

$$\operatorname{tr}(a) = \sum_{i=1}^{r} \rho_i(a_{ii}). \tag{3.5}$$

We will need some technical conditions about $\mathscr{A}$, these are summarized in the following definition, originally introduced by Vinberg in [28].

Definition 3.2 (T-algebra) The set of generalized matrices $\mathscr{A}$ is a T-algebra if for all $a, b, c \in \mathscr{A}$, $t, u, v \in \mathscr{T}$ and $1 \le i, j \le r$

1. $\mathscr{A}_{ii}$ is isomorphic to $\mathbb{R}$,
2. $e_i a_{ij} = a_{ij} e_j = a_{ij}$ for all $a_{ij} \in \mathscr{A}_{ij}$,
3. $\operatorname{tr}(ab) = \operatorname{tr}(ba)$,
4. $\operatorname{tr}(a(bc)) = \operatorname{tr}((ab)c)$,
5. $\operatorname{tr}(a^* a) \ge 0$, and $\operatorname{tr}(a^* a) = 0$ implies $a = 0$,

[3] Actually, we will not make use of the explicit form of this definition. Any formula satisfying Definition 3.2 would work.

6. $t(uv) = (tu)v$,
7. $t(uu^*) = (tu)u^*$.

It is important to note that multiplication in a T-algebra is neither commutative nor associative.

Based on the properties of the trace we can define a natural inner product on $\mathscr{A}$, namely $\langle a, b\rangle = \mathrm{tr}(a^* b)$, which will provide us a Hilbert-space structure. Let now

$$\mathscr{I} = \{t \in \mathscr{T} : \rho_i(t_{ii}) > 0, 1 \leq i \leq r\} \tag{3.6}$$

be the set of upper triangular matrices with positive diagonal elements, and define

$$\mathscr{K}(\mathscr{A}) = \{tt^* : t \in \mathscr{I}\} \subseteq \mathscr{H}. \tag{3.7}$$

The fundamental representation theorem of homogeneous cones was proved by Vinberg [28]:

Theorem 3.1 (Representation of homogeneous cones) *A cone $\mathscr{K}$ is homogeneous if and only if there is a T-algebra $\mathscr{A}$ such that* $\mathrm{int}(\mathscr{K})$ *is isomorphic to $\mathscr{K}(\mathscr{A})$. Moreover, given $\mathscr{K}(\mathscr{A})$, the representation of an element from* $\mathrm{int}(\mathscr{K})$ *in the form tt^* is unique. Finally, the interior of the dual cone $\mathscr{K}^*$ can be represented as $\{t^*t : t \in \mathscr{I}\}$.*

The rank of the homogeneous cone $\mathscr{K}$ is the rank of the generalized matrix algebra $\mathscr{A}$. It is denoted by $\mathrm{rank}(\mathscr{K})$ and it is the size of the generalized matrices used in the construction.

Remark 3.1 This theorem is analogous to the representation of symmetric cones as the set of squares over a Jordan algebra. However, for T-algebras and homogeneous cones it is not true that for a given $a \in \mathscr{A}$ we have $aa^* \in \mathscr{K}$.

Remark 3.2 The positivity of the diagonal elements ($t \in \mathscr{I}$) is only required to ensure uniqueness. In general every $x \in \mathscr{K}$ can be expressed as $x = tt^*$ for some $t \in \mathscr{T}$.

Example 3.1 The cone of positive semidefinite matrices is homogeneous, since every positive definite matrix admits a unique Cholesky factorization of the form UU^T with positive diagonal elements in U.

Example 3.2 (Rotated Lorentz cone) By the classical definition, a rotated Lorentz cone is the following set:

$$\mathbb{L}_{\mathrm{r}} = \left\{(x_0, x_1, x) \in \mathbb{R}^{n+2} : x_0 \geq 0, x_0 x_1 - \|x\|^2 \geq 0\right\}. \tag{3.8}$$

Rotated Lorentz cones can be represented as generalized matrices in the following way. Let the T-algebra be defined as

$$\mathscr{A}_{\mathbb{L}_{\mathrm{r}}} = \left\{\begin{pmatrix} v_0 & v^T \\ u & u_0 \end{pmatrix} : u_0, v_0 \in \mathbb{R}, u, v \in \mathbb{R}^n\right\}, \tag{3.9}$$

with the product

$$\begin{pmatrix} v_0 & v^T \\ u & u_0 \end{pmatrix}\begin{pmatrix} q_0 & q^T \\ p & p_0 \end{pmatrix} = \begin{pmatrix} v_0 q_0 + v^T q & v_0 q^T + p_0 v^T \\ q_0 u + u_0 p & u_0 p_0 + q^T u \end{pmatrix}. \tag{3.10}$$

The involution is the transpose. It is straightforward to verify that this algebra does satisfy all the axioms of Definition 3.2, therefore it is a T-algebra. This implies that the set

$$\mathscr{K}\left(\mathscr{A}_{\mathbb{L}_\mathrm{r}}\right) = \left\{ \begin{pmatrix} t_1 & t^T \\ 0 & t_0 \end{pmatrix} \begin{pmatrix} t_1 & 0 \\ t & t_0 \end{pmatrix} : t_0, t_1 > 0 \right\} \tag{3.11}$$

is a homogeneous cone. Now considering the element

$$\begin{pmatrix} x_1 & x^T \\ x & x_0 \end{pmatrix} \in \mathscr{A}_{\mathbb{L}} \tag{3.12}$$

with $x_0, x_1 > 0$ and $(x_0, x_1, x) \in \mathbb{L}_\mathrm{r}$. We have the factorization

$$\begin{pmatrix} x_1 & x^T \\ x & x_0 \end{pmatrix} = \begin{pmatrix} \sqrt{x_1 - \frac{\|x\|^2}{x_0}} & \frac{x^T}{\sqrt{x_0}} \\ 0 & \sqrt{x_0} \end{pmatrix} \begin{pmatrix} \sqrt{x_1 - \frac{\|x\|^2}{x_0}} & 0 \\ \frac{x}{\sqrt{x_0}} & \sqrt{x_0} \end{pmatrix}, \tag{3.13}$$

establishing an isomorphism between $\operatorname{int}(\mathbb{L}_\mathrm{r})$ and $\mathscr{K}\left(\mathscr{A}_{\mathbb{L}_\mathrm{r}}\right)$. It is interesting to note that the rotated Lorentz cone arises more naturally in this framework, while standard Lorentz cones are easier to construct in Jordan algebras. Naturally, they are isomorphic.

3.3 Recursive construction of homogeneous cones

Homogeneous cones can also be constructed in a recursive way, due to Vinberg [28]. Here we follow [13].

Definition 3.3 (Homogeneous symmetric form on a cone)
Let $\mathscr{K} \subseteq \mathbb{R}^n$ be a homogeneous cone, and consider a symmetric bilinear mapping $B : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^n$ such that for every $u, v \in \mathbb{R}^p$ and $\lambda_1, \lambda_2 \in \mathbb{R}$

1. $B(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 B(u_1, v) + \lambda_2 B(u_2, v)$,
2. $B(u, v) = B(v, u)$,
3. $B(u, u) \in \mathscr{K}$,
4. $B(u, u) = 0$ implies $u = 0$.

A symmetric bilinear form B is called homogeneous if $\mathscr{K}$ is a homogeneous cone and there is a transitive subgroup $G \subseteq \operatorname{Aut}(\mathscr{K})$ such that for every $g \in G$ there is a linear transformation $\tilde{g}$ on $\mathbb{R}^p$ such that

$$g\left(B(u, v)\right) = B(\tilde{g}(u), \tilde{g}(v)), \tag{3.14}$$

in other words, the diagram

$$\begin{array}{ccc} \mathbb{R}^p \times \mathbb{R}^p & \xrightarrow{\tilde{g} \times \tilde{g}} & \mathbb{R}^p \times \mathbb{R}^p \\ {\scriptstyle B}\big\downarrow & & \big\downarrow{\scriptstyle B} \\ \mathbb{R}^n & \xrightarrow{g} & \mathbb{R}^n \end{array} \tag{3.15}$$

is commutative.

Having such a bilinear function we can define a new set:

Definition 3.4 (Siegel cone) The Siegel cone $\mathrm{SC}(\mathscr{K},B)$ of $\mathscr{K}$ and B is defined as

$$\mathrm{SC}(\mathscr{K},B) = \mathrm{cl}\left(\{(x,u,t) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R} : t > 0, tx - B(u,u) \in \mathscr{K}\}\right). \tag{3.16}$$

Theorem 3.2 (Properties of the Siegel cone) *The Siegel cone has the following properties:*

1. *If $\mathscr{K}$ is a homogeneous cone and B is a homogeneous symmetric bilinear form then $\mathrm{SC}(\mathscr{K},B)$ is a homogeneous cone.*
2. *Every homogeneous cone can be obtained as the Siegel cone of another homogeneous cone using an appropriate bilinear function B.*
3. *For the rank of the Siegel cone we have*

$$\mathrm{rank}(\mathrm{SC}(\mathscr{K},B)) = \mathrm{rank}(\mathscr{K}) + 1, \tag{3.17}$$

i.e., the rank of a homogeneous cone is exactly the number of Siegel extension steps needed to construct the cone starting from $\{0\}$, the cone of rank 0.

This result is analogous to the concept of the Schur complement for semidefinite matrices. It expresses a quadratic relation with a linear one of higher rank and dimension.

Example 3.3 (The rotated Lorentz cone as a Siegel cone) Let us see how the rotated Lorentz cone $\mathbb{L}_\mathrm{r}$ can be constructed this way. Starting with $\mathscr{K} = \mathbb{R}_+$ as a homogeneous cone of rank 1, we choose the bilinear function $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ to be the usual scalar product, i.e., $B(u,v) = u^T v$. All the automorphisms of $\mathbb{R}_+$ are multiplications with a positive number, thus if $g(u) = \alpha u$ then with $\tilde{g}(u) = \sqrt{\alpha} u$ we have

$$g(B(u,v)) = \alpha u^T v = \left(\sqrt{\alpha} u\right)^T \left(\sqrt{\alpha} v\right) = B(\tilde{g}(u), \tilde{g}(v)). \tag{3.18}$$

This shows that B is a homogeneous symmetric form. Now the resulting Siegel cone is

$$\mathrm{SC}(\mathscr{K},B) = \left\{(x,u,t) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} : t \geq 0, tx \geq \|u\|^2\right\}, \tag{3.19}$$

which is exactly the rotated Lorentz cone. This shows that the rank of this cone is indeed 2.

3.4 Another representation for homogeneous cones

Despite the quite abstract definition, it turns out that homogeneous cones are slices of the positive semidefinite cone. More precisely (see [8]):

Proposition 3.1 (Homogeneous cones as semidefinite slices) *If $\mathscr{K} \subset \mathbb{R}^n$ is a homogeneous cone then there exist an $m \leq n$ and an injective linear map $M : \mathbb{R}^n \to \mathbb{S}^{m\times m}$ such that*

$$M(\mathscr{K}) = \mathbb{S}_+^{m\times m} \cap M(\mathbb{R}^n). \tag{3.20}$$

A similar result was obtained independently by Faybusovich, [11]. We have to note here that not all slices of the positive semidefinite cone are homogeneous, see [8] for a counterexample.

3.5 Self dual homogeneous (symmetric) cones

Throughout the rest of this chapter we will assume that $\mathscr{K}$ is a symmetric cone, i.e., it is homogeneous and self dual. In this case the theory can be simplified a bit, see [29]. In particular, we will use the following proposition:

Proposition 3.2 (Special properties of symmetric cones) *If $\mathscr{A}$ is a T-algebra and $\mathscr{K}$ is the corresponding self dual homogeneous cone, then for every $a,b \in \mathscr{A}$ we have*

$$\operatorname{tr}\left((aa^*)(bb^*)\right) = \operatorname{tr}\left((ab)(ab)^*\right). \tag{3.21}$$

Consequently, $aa^* \in \mathscr{K}$ for every $a \in \mathscr{A}$. Moreover, the mapping

$$g_w : u \mapsto wuw^* \tag{3.22}$$

is well-defined for every $u \in \mathscr{H}$.

Remark 3.3 Remember that in the case of a general T-algebra the identity (3.21) holds typically only if $a, b^* \in \mathscr{T}$, and in that case $aa^* \in \mathscr{K}$ and $bb^* \in \mathscr{K}^*$.

4 An exact duality theory for symmetric cones

Now we are ready to present our results about the strong dual for the symmetric cone optimization problems. First let us repeat the primal problem:

$$\begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\in \mathscr{K}, \end{aligned} \tag{P}$$

where $A \in \mathbb{R}^{m\times n}$, $c, s \in \mathbb{R}^n$, $b, y \in \mathbb{R}^m$. We assume that $\mathscr{K} = \mathscr{K}^*$ is a homogeneous cone given in the form of $\mathscr{K}(\mathscr{A})$, i.e., every element in $\operatorname{int}(\mathscr{K})$ is represented as tt^* with $t \in \mathscr{I}$. Our goal is to derive a dual for (P) satisfying the following requirements (cf. [21, §1.4]):

1. The dual problem is a homogeneous cone optimization problem that can be generated easily from the primal input data.
2. If the primal problem is feasible and bounded then the duality gap (the difference of the optimal primal and dual objective values) is 0, and the optimum is attained on the dual side.
3. It yields a theorem of the alternative for symmetric conic feasibility systems, i.e., the infeasibility of a symmetric conic system can be characterized by the feasibility of a homogeneous conic system.

4.1 The facial reduction algorithm for symmetric cones

We will use the facial reduction algorithm presented in §2.2. In light of Theorem 2.3 we have to establish two things:

1. prove that symmetric cones are nice, therefore the facial reduction algorithm can be applied;

2. find an explicit expression for the space $(\mathscr{F}_i)^{\perp}$ and show that it can be expressed with homogeneous cones.

For the proof of the first part we need some lemmas.

Lemma 4.1 (Faces of intersections of cones) *If $\mathscr{K} = \mathscr{K}_1 \cap \mathscr{K}_2$ then $\mathscr{F} \trianglelefteq \mathscr{K}$ if and only if $\mathscr{F} = \mathscr{F}_1 \cap \mathscr{F}_2$ where $\mathscr{F}_1 \trianglelefteq \mathscr{K}_1$ and $\mathscr{F}_2 \trianglelefteq \mathscr{K}_2$. In other words, a face of the intersection of cones is the intersection of faces of the cones.*

Remark 4.1 This classical result, commonly attributed to Dubins [9], first appeared in [5]. An easily accessible source is [27, Theorem 3.6.19].

Pataki [17] proves the following lemma about the intersection of nice cones. We include a short proof for completeness.

Lemma 4.2 *The intersection of nice cones is nice.*

Proof Let $\mathscr{K}_1$ and $\mathscr{K}_2$ be nice cones, $\mathscr{K} = \mathscr{K}_1 \cap \mathscr{K}_2$, $\mathscr{F} \trianglelefteq \mathscr{K}$. From the previous lemma we know that $\mathscr{F}$ can be expressed as $\mathscr{F} = \mathscr{F}_1 \cap \mathscr{F}_2$, with $\mathscr{F}_i \trianglelefteq \mathscr{K}_i$, $i = 1,2$.

$$\begin{aligned}
\mathscr{F}^* &= (\mathscr{F}_1 \cap \mathscr{F}_2)^* = \operatorname{cl}(\mathscr{F}_1^* \oplus \mathscr{F}_2^*) = \operatorname{cl}\left(\left(\mathscr{K}_1^* \oplus \mathscr{F}_1^{\perp}\right) \oplus \left(\mathscr{K}_2^* \oplus \mathscr{F}_2^{\perp}\right)\right) \\
&= \operatorname{cl}\left((\mathscr{K}_1^* \oplus \mathscr{K}_2^*) \oplus \left(\mathscr{F}_1^{\perp} \oplus \mathscr{F}_2^{\perp}\right)\right) = (\mathscr{K}_1 \cap \mathscr{K}_2)^* \oplus (\mathscr{F}_1 \cap \mathscr{F}_2)^{\perp} \\
&= \mathscr{K}^* \oplus \mathscr{F}^{\perp},
\end{aligned} \tag{4.1}$$

where we used the basic properties of the dual cones. □

These two lemmas lead us to the following theorem:

Theorem 4.1 *Homogeneous cones (and thus symmetric cones) are nice.*

Proof By Proposition 3.1 homogeneous cones are slices of the positive semidefinite cone. The semidefinite cone and all the subspaces are nice, thus their intersection is nice. □

Remark 4.2 So far we have not used the fact the $\mathscr{K}$ is self dual.

For the second part we will consider the following bilinear function:

$$B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \tag{4.2a}$$

$$B(u,v) = \frac{1}{2}(uv^* + vu^*) \tag{4.2b}$$

Our first result ensures that B is an appropriate bilinear form for our construction:

Lemma 4.3 *The bilinear function B satisfies all the requirements of Definition 3.3 thus it is a homogeneous symmetric form for $\mathscr{K}$ and consequently $\operatorname{SC}(\mathscr{K}, B)$ is a valid Siegel cone.*

Proof Bilinearity and symmetry are obvious. Further,

$$B(u,u) = uu^* \in \mathscr{K} \tag{4.3}$$

by Proposition 3.2. If $B(u,u) = 0$ then necessarily $u = 0$.

To prove the homogeneity consider the group (see [28,29])

$$G = \{g_w \in \operatorname{Aut}(\mathscr{K}) : g_w(u) = wuw^* \text{ for an invertible element } w \in \mathscr{T}\}. \tag{4.4}$$

The mapping g_w is well defined by Proposition 3.2 and G is indeed a group. Its transitivity is due to the fact that $g_{w^{-1}}(ww^*) = ee^* = e$. For $g_w \in G$ let us define $\tilde{g}_w(u) = wu$. We then have

$$g_w(B(u,v)) = \frac{1}{2}w(uv^* + vu^*)w^* = \frac{1}{2}\left((wu)(wv)^* + (wv)(wu)^*\right) = B\left(\tilde{g}_w(u), \tilde{g}_w(v)\right). \quad (4.5)$$

This completes the proof of the homogeneity of B. □

Remember that if $\mathscr{C} \subseteq \mathscr{K}$ is a convex subset of $\mathscr{K}$, then $\mathscr{F}(\mathscr{C})$ denotes the smallest face of $\mathscr{K}$ containing $\mathscr{C}$. If $\mathscr{F} \trianglelefteq \mathscr{K}$ is a face of $\mathscr{K}$ then its conjugate face is defined as $\mathscr{F}^{\mathrm{c}} = \mathscr{F}^{\perp} \cap \mathscr{K}^*$, where the orthogonal complement is now defined relative to the set $\mathscr{H}$ of self-adjoint elements. The following result is a generalization of Lemma 2.1 in [23] and it plays a key role in the construction of the dual problem:

Theorem 4.2 (The description of $((\mathscr{F}(\mathscr{C}))^{\mathrm{c}})^{\perp}$) *If $\mathscr{C} \subseteq \mathscr{K}$ is a convex subset of the symmetric cone $\mathscr{K}$ then*

$$((\mathscr{F}(\mathscr{C}))^{\mathrm{c}})^{\perp} = \{w + w^* : \exists u \in \mathscr{C}, u - B(w,w) \in \mathscr{K}\}. \quad (4.6)$$

Proof First, let $u \in \mathscr{C}$ and $w \in \mathscr{A}$ be such that $u - B(w,w) \in \mathscr{K}$ and let us choose an arbitrary element $v \in (\mathscr{F}(\mathscr{C}))^{\mathrm{c}} \subseteq \mathscr{K}^*$. By the properties of the dual cone we have

$$\langle v, u - B(w,w) \rangle \geq 0. \quad (4.7)$$

Since $u \in \mathscr{F}(\mathscr{C})$ and $v \in \mathscr{F}(\mathscr{C})^{\perp}$, we have $\langle v, u \rangle = 0$. Now $B(w,w) \in \mathscr{K}$ implies that

$$0 = \langle v, B(w,w) \rangle = \langle v, ww^* \rangle. \quad (4.8)$$

Now as $v \in \mathscr{K}^*$ it can be written as $v = t^*t$ for some upper triangular element $t \in \mathscr{T}$. This gives

$$0 = \langle v, ww^* \rangle = \mathrm{tr}\left((tw)(tw)^*\right), \quad (4.9)$$

which implies that $tw = 0$ thus

$$\langle v, w + w^* \rangle = \mathrm{tr}\left(v(w + w^*)\right) = \mathrm{tr}\left(t^*t(w + w^*)\right) = \mathrm{tr}\left(t^*(tw) + (w^*t^*)t\right) = 0. \quad (4.10)$$

This shows that $w + w^* \in ((\mathscr{F}(\mathscr{C}))^{\mathrm{c}})^{\perp}$.

For the converse direction let us take $w + w^* \in ((\mathscr{F}(\mathscr{C}))^{\mathrm{c}})^{\perp}$ and $u \in \mathscr{C} \cap \mathrm{rel\,int}\left(\mathscr{F}(\mathscr{C})\right)$. At this point we do not specify w, we only fix the sum $w + w^*$. The exact form of w will be specified later. We will show that there exists an $\alpha \geq 0$ such that $\alpha u - B(w,w) \in \mathscr{K}$, for this we need to show that $B(w,w) \in \mathscr{F}(\mathscr{C})$.

Due to the homogeneity of $\mathscr{K}$ there is an automorphism $\varphi \in \mathrm{Aut}\left(\mathscr{K}\right)$ such that $\varphi(u)$ is a diagonal element,[4] or in other words,

$$\varphi(u) = \begin{pmatrix} \mathrm{Diag}\left(d_{11}, \ldots, d_{kk}\right) & 0 \\ 0 & 0 \end{pmatrix}, \quad (4.11)$$

where the diagonal elements are positive, i.e., $\rho(d_{11}), \ldots, \rho(d_{kk}) > 0$. The generated face is

$$\mathscr{F}(\mathscr{C}) = \mathscr{F}(u) = \left\{ \varphi^{-1}\left[\begin{pmatrix} s & 0 \\ 0 & 0 \end{pmatrix}\right] : \begin{pmatrix} s & 0 \\ 0 & 0 \end{pmatrix} \in \mathscr{K} \right\}, \quad (4.12)$$

[4] Remember that we consider the elements of the homogeneous cone in matrix form.

and the complementary face is

$$\mathscr{F}(\mathscr{C})^{\mathrm{c}} = \mathscr{F}(u)^{\mathrm{c}} = \left\{ \varphi^{-1}\left[\begin{pmatrix} 0 & 0 \\ 0 & z \end{pmatrix}\right] : \begin{pmatrix} 0 & 0 \\ 0 & z \end{pmatrix} \in \mathscr{K}^* \right\}. \tag{4.13}$$

Now if $w + w^* \in (\mathscr{F}(\mathscr{C})^{\mathrm{c}})^{\perp}$ then it can be expressed as

$$w + w^* = \varphi^{-1}\left[\begin{pmatrix} p_1 & p_2 \\ p_2^* & 0 \end{pmatrix}\right] \tag{4.14}$$

with a symmetric p_1. Defining

$$w = \varphi^{-1}\left[\begin{pmatrix} \frac{1}{2}p_1 & p_2 \\ 0 & 0 \end{pmatrix}\right] \tag{4.15}$$

we have

$$\begin{aligned} B(w,w) = ww^* &= \varphi^{-1}\left[\begin{pmatrix} \frac{1}{2}p_1 & p_2 \\ 0 & 0 \end{pmatrix}\right] \varphi^{-1}\left[\begin{pmatrix} \frac{1}{2}p_1 & 0 \\ p_2 & 0 \end{pmatrix}\right] \\ &= \varphi^{-1}\left[\begin{pmatrix} \frac{1}{2}p_1 & p_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{2}p_1 & 0 \\ p_2^* & 0 \end{pmatrix}\right] = \varphi^{-1}\left[\begin{pmatrix} \frac{1}{4}p_1^2 + p_2 p_2^* & 0 \\ 0 & 0 \end{pmatrix}\right] \in \mathscr{F}(\mathscr{C}). \end{aligned} \tag{4.16}$$

Finally, as $B(w,w) \in \mathscr{F}(\mathscr{C})$ and $u \in \operatorname{rel\,int}(\mathscr{F}(\mathscr{C}))$ we have an $\alpha \geq 0$ such that $\alpha u - B(w,w) \in \mathscr{F}(\mathscr{C}) \subseteq \mathscr{K}$, see [2,3] for the proof of this classical but simple result. □

Now we can give an explicit expression for $(\mathscr{F}_i)^{\perp}$:

Theorem 4.3 *The space* $(\mathscr{F}_i)^{\perp}$, $(i = 1, \dots, L+1)$ *can be expressed as*

$$(\mathscr{F}_i)^{\perp} = \left\{ w + w^* : \exists \alpha \geq 0, \left(u^1 + \dots + u^i, w, \alpha\right) \in \mathrm{SC}(\mathscr{K}^*, B) \right\}, \tag{4.17}$$

where $\mathrm{SC}(\mathscr{K}^*, B)$ *is the Siegel cone built on* $\mathscr{K}^*$ *with the bilinear form* B.

Proof Theorem 2.3 about the facial reduction algorithm establishes that

$$(\mathscr{F}_i)^{\perp} = \left(\left(\mathscr{F}\left(u^1 + \dots + u^i\right)\right)^{\mathrm{c}}\right)^{\perp}, \tag{4.18}$$

where $u^1 + \dots + u^i \in \mathscr{K}^*$, and the complementary face is taken in $\mathscr{K}$. We can now apply the previous proposition with $\mathscr{C} = u^1 + \dots + u^i$ and $\mathscr{K}^*$ instead of $\mathscr{K}$ to get

$$(\mathscr{F}_i)^{\perp} = \left\{ w + w^* : \exists \alpha \geq 0, \alpha\left(u^1 + \dots + u^i\right) - B(w,w) \in \mathscr{K}^* \right\}. \tag{4.19}$$

Lemma 4.3 enables us to rewrite the condition $\alpha\left(u^1 + \dots + u^i\right) - B(w,w) \in \mathscr{K}^*$ using the Siegel cone, this yields the final form for $(\mathscr{F}_i)^{\perp}$:

$$(\mathscr{F}_i)^{\perp} = \left\{ w + w^* : \exists \alpha \geq 0, (u^1 + \dots + u^i, w, \alpha) \in \mathrm{SC}(\mathscr{K}^*, B) \right\}. \tag{4.20}$$

This formulation suits our needs as it is expressed with homogeneous cones using dual information only. □

Corollary 4.1 *Now we can apply this theorem along with the facial reduction algorithm described in §2.2 to obtain an explicit strong dual for* (P)*. This provides the following extended dual problem:*

$$\begin{aligned}
&\min \left\langle c, \left(x^{L+1}+z^{L+1}\right)\right\rangle \\
A\left(x^{L+1}+z^{L+1}\right) &= b \\
A\left(x^{i}+z^{i}\right) &= 0,\, i=1,\dots,L \\
\left\langle c, \left(x^{i}+z^{i}\right)\right\rangle &= 0,\, i=1,\dots,L \\
x^0, z^1 &= 0 \\
\left(x^1+\dots+x^{i-1}, z^i, \alpha^i\right) &\in \mathrm{SC}\left(\mathscr{K}^*, B\right),\, i=2,\dots,L+1 \\
x^i &\in \mathscr{K}^*,\, i=1,\dots,L+1,
\end{aligned} \tag{ED}$$

where L is the lesser of $\dim\left(\mathrm{Ker}\,(A)\cap\{c\}^{\perp}\right)$ *and the length of the longest chain of faces in $\mathscr{K}^*$. We used the fact that $\langle g, w+w^*\rangle = 2\langle g, w\rangle$ if $g=g^*$ to simplify the expressions.*

The dual problem is a homogeneous cone optimization problem as the dual cone $\mathscr{K}^* = \mathscr{K}$ and its Siegel cone $\mathrm{SC}\left(\mathscr{K}^*, B\right)$ are both homogeneous cones.

Although this extended dual problem satisfies all the requirements we set up at the beginning it can still be improved using a reduction technique due to Pataki [18]. The idea is to disaggregate the LHS of the Siegel cone constraint. This provides the following dual problem:

$$\begin{aligned}
&\min \left\langle c, \left(x^{L+1}+z^{L+1}\right)\right\rangle \\
A\left(x^{L+1}+z^{L+1}\right) &= b \\
A\left(x^{i}+z^{i}\right) &= 0,\, i=1,\dots,L \\
\left\langle c, \left(x^{i}+z^{i}\right)\right\rangle &= 0,\, i=1,\dots,L \\
x^0, z^1 &= 0 \\
\left(x^{i-1}, z^i, 1\right) &\in \mathrm{SC}\left(\mathscr{K}^*, B\right),\, i=2,\dots,L+1 \\
x^{L+1} &\in \mathscr{K}^*.
\end{aligned} \tag{ED$_{\mathrm{disagg}}$}$$

4.2 Complexity of the dual problem

Let us examine system (ED) from an algorithmic point of view.

Optimization problems over convex cones are generally solved with interior point methods. From the viewpoint of this paper the internal workings of these algorithms are not so important, the interested reader is directed to the literature, most importantly [16, 24]. These methods solve problems (P) and (D) to precision ε in at most $\mathscr{O}\left(\sqrt{\vartheta}\log(1/\varepsilon)\right)$ iterations, where ϑ is a complexity parameter depending only on the cone $\mathscr{K}$. It is important to note that the iteration complexity does not depend on the dimension of $\mathscr{K}$ or the number of linear equalities, but of course the cost of one iteration is determined by these quantities. There are several estimates for ϑ depending on the geometric and algebraic properties of the cone, for the Lorentz cone in any dimension $\vartheta = 2$, for the cone of $n\times n$ positive semidefinite matrices it is n. For homogeneous cones the two most important results are (see [12, 13]):

1. $\vartheta(\mathscr{K}) = \vartheta(\mathscr{K}^*)$, and
2. $\vartheta(\mathscr{K}) = \operatorname{rank}(\mathscr{K})$.

Moreover, we know that a Siegel extension increases the rank of the cone and thus the complexity by 1. We can also give an explicit barrier function for the Siegel cones, so we can solve the extended dual problems with interior point methods.

The complexity of the dual problem can be computed easily from the rank of the cones. Let $r = \vartheta(\mathscr{K}) = \operatorname{rank}(\mathscr{K})$ be the complexity parameter (and also the rank) of $\mathscr{K}$, then the complexity of the dual problem is $r + L(r+1)$. This seems to contradict the fact that $\operatorname{rank}(\mathscr{K}) \geq \operatorname{rank}(\mathscr{K}_{\min}) = \operatorname{rank}(\mathscr{K}^*_{\min})$, thus we might expect to have a lower complexity. The contradiction is easily removed by noticing that the dual problem (ED) uses only one possible representation of $\mathscr{K}_{\min}$, with several other cones linked to each other through linear and conic constraints. In this construction $\mathscr{K}_{\min}$ is represented as an intersection of L cones therefore its dual will be a sum of L cones. Better representations may also exist. The following is an interesting open question: what is the best possible complexity for an explicitly formed, easily computable strong dual problem to (P)?

These complexity results are constructive in the sense that the appropriate barrier function of $\operatorname{SC}(\mathscr{K}, B)$ can be constructed from the barrier function of $\mathscr{K}$. The only practical difficulty can be the actual computability of these functions, but we can always use an approximation [26] to drive the algorithm.

5 Special cases

In this section we specialize the new dual for semidefinite and second-order conic optimization.

5.1 Semidefinite optimization

For the case when $\mathscr{K} = \mathbb{S}_+^{n\times n}$ is the set of $n \times n$ positive semidefinite matrices, Ramana's original paper [21] proposes the following extended dual problem:

$$
\begin{aligned}
&\min\ \operatorname{Tr}\left(c^T\left(x^{L+1}+z^{L+1}\right)\right)\\
&A\left(x^{L+1}+z^{L+1}\right) = b\\
&\operatorname{Tr}\left(c^T\left(x^i+z^i\right)\right) = 0,\ i = 1,\dots,L\\
&A\left(x^i+z^i\right) = 0,\ i = 1,\dots,L\\
&x^0, z^1 = 0\\
&\begin{pmatrix} x^{i-1} & z^{iT} \\ z^i & I \end{pmatrix} \in \mathbb{S}_+^{2n\times 2n},\ i = 2,\dots,L+1\\
&x^{L+1} \in \mathbb{S}_+^{n\times n},
\end{aligned}
\tag{ED$_{\text{Ramana}}$}
$$

were c and all the variables are $n \times n$ matrices, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times(n\times n)}$.

Our dual construction specialized for semidefinite optimization is only slightly different:

$$\begin{aligned} \min\ & \operatorname{Tr}\left(c^T\left(x^{L+1}+z^{L+1}\right)\right) \\ A\left(x^{L+1}+z^{L+1}\right) &= b \\ \operatorname{Tr}\left(c^T\left(x^i+z^i\right)\right) &= 0,\ i=1,\dots,L \\ A\left(x^i+z^i\right) &= 0,\ i=1,\dots,L \\ x^0,z^1 &= 0 \\ \left(x^{i-1},z^i,1\right) &\in \operatorname{SC}\left(\mathbb{S}_+^{n\times n},B\right),\ i=2,\dots,L+1 \\ x^{L+1} &\in \mathbb{S}_+^{n\times n}, \end{aligned} \tag{$\mathrm{ED}_{\mathrm{HomPSD}}$}$$

with the bilinear form $B(u,v)=(uv^T+vu^T)/2$. The difference between the two duals is that the Ramana dual uses the Schur complement and thus the dual is a semidefinite problem, while our dual uses the Siegel cone, and arrives at a general homogeneous conic problem. The Schur complement gives the semidefinite representation of the resulting Siegel cone.

This shows that we actually derived a new strong dual for the semidefinite optimization problem, which is not a semidefinite problem. The benefit of our dual is its improved complexity. The complexity parameter (see §4.2) of ($\mathrm{ED}_{\mathrm{Ramana}}$) is $2nL+n$, since the Schur complement results in $2n\times 2n$ matrices. The complexity parameter of ($\mathrm{ED}_{\mathrm{HomPSD}}$) is only $(n+1)L+n$, as the Siegel cone construction increases the complexity only by 1. The difference is explained by the fact that our dual is not a semidefinite problem.

The difference in complexity answers another question. Since homogeneous cones are slices of an appropriate dimensional positive semidefinite cone (see Proposition 3.1, we can rewrite the original problem (P) as a semidefinite problem, apply the classical Ramana dual and then transform the problem back to homogeneous cones. But as our dual for the semidefinite problem has better complexity than the Ramana dual, we cannot hope to recover our dual from this process.

5.2 Second order conic optimization

Originally, our research was sparked by the idea to construct a purely second order conic strong dual to the second order conic optimization problem. Our continued failure to construct such a problem led us to believe that the strong dual cannot be a second order conic optimization problem. After developing the theory for symmetric cones we can partially answer this question.

For notational simplicity let us consider the problem with only one cone:[5]

$$\begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\in \mathbb{L}, \end{aligned} \tag{$\mathrm{P}_{\mathrm{Lor}}$}$$

[5] There is a closed form solution to this problem due to Alizadeh and Goldfarb [1], but even this simplest form is enough to illustrate our point.

where $\mathbb{L} = \{u \in \mathbb{R}^n : u_1 \geq \|u_{2:n}\|_2\}$. The extended dual for problem (P_{Lor}) is ($L \leq 2$, the length of the longest chain of faces in $\mathbb{L}$):

$$
\begin{aligned}
\min\ & \left\langle c, \left(x^3 + z^3\right)\right\rangle \\
A\left(x^3 + z^3\right) &= b \\
A\left(x^i + z^i\right) &= 0,\ i = 1,2 \\
\left\langle c, \left(x^i + z^i\right)\right\rangle &= 0,\ i = 1,2 \\
x^0, z^1 &= 0 \\
\left(x^{i-1}, z^i, 1\right) &\in \mathrm{SC}\left(\mathbb{L}, B\right),\ i = 2,3 \\
x^3 &\in \mathbb{L}.
\end{aligned}
\tag{$\mathrm{ED}_{\mathrm{Lor}}$}
$$

The rank of the Siegel cone built over the Lorentz cone is 3 and thus it cannot be a Lorentz cone itself, see the classification of lower dimensional homogeneous cones in [14]. It is still an open question whether a purely Lorentz cone strong dual exists to (P_{Lor}). For duals based on the facial reduction algorithm the answer is negative.

Actually, for second order cones the bilinear form simplifies to $B(u,v) = (u^T v, u_1 v_{2:n} + v_1 u_{2:n})$, i.e., the usual product in the Jordan algebra of the Lorentz cone. This product is commutative, which simplifies the proofs. The Siegel cone corresponding to this form will be

$$
\begin{aligned}
\mathrm{SC}\left(\mathbb{L}, B\right) &= \left\{(x,z,t) \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \times \mathbb{R}_+ : tx - B(z,z) \in \mathbb{L}\right\} \\
&= \left\{(x,z,t) \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \times \mathbb{R}_+ : tx_1 - \|z\|^2 \geq \|tx_{2:n} - 2z_1 z_{2:n}\|\right\}.
\end{aligned}
\tag{5.1}
$$

One might wonder how to express the Siegel cone $\mathrm{SC}(\mathbb{L}, B)$ with a semidefinite constraint. An n dimensional Lorentz cone can be represented by $n \times n$ semidefinite matrices, then the direct application of the Schur complement for the Siegel extension yields a representation of $\mathrm{SC}(\mathbb{L}, B)$ with $2n \times 2n$ matrices. Moreover, Proposition 3.1 guarantees the existence of a representation with $(2n-1) \times (2n-1)$ matrices. Both are much worse than the actual complexity of the cone, which is 3. It is unknown if there is a better construction, the theoretical lower bound (by a simple dimension argument) is in the order of $2\sqrt{n}$.

6 Conclusions and further questions

The central result of our paper was to provide an explicit strong dual for optimization problems over symmetric cones. The construction is based on the facial reduction algorithm and it uses the machinery of homogeneous cones quite heavily. Using the dual we showed a new strong dual for semidefinite problems, which has better complexity than the classical Ramana dual. We have also specialized the dual for second order cone optimization and found that the dual problem cannot be expressed with second order cones. This answers an open question.

There is much more to explore around this topic. First of all, as we discussed it in §4.2, the best possible complexity of an easily constructible strong dual to (P) is still an open question. Moreover, the extended dual is deficient in the sense that it only guarantees zero duality gap and dual (but not primal) solvability. This follows from the properties of the construction in [6]. Another idea, stemming from [22] is to investigate the Lagrange dual of the extended dual (called the corrected primal).

A natural question is whether the results of the present paper extend to more general cones. We proved that homogeneous cones are nice, thus the facial reduction algorithm can be applied. To obtain a description of $(\mathscr{F}(\mathscr{C})^{c})^{\perp}$ all we need is a bilinear function B that satisfies the requirements of Definition 3.3 and works in the proof of Theorem 4.2. The existence and possible construction of such a function is an open question.[6]

Generalizations to more general cones are probably not possible. Of course, if a cone is a slice of a higher dimensional semidefinite cone then the problem can be rewritten as a semidefinite optimization problem and Ramana's strong dual can be applied. However, the complexity of the dual problem will be much worse than the complexity of the original problem, see §5.2 about the complexity of the strong dual of a second order conic problem. The bottleneck seems to be some Schur complement-like result, which can convert a quadratic conic constraint into a linear one of good complexity. For homogeneous cones the Siegel cone construction of §3.3 gives the answer, but for more general cones there are no such results. Homogeneous cones possess some remarkable properties, which are essential for this theory.

Homogeneous cones also provide an efficient modeling tool for optimization problems currently modeled directly as semidefinite problems. Siegel cones give rise to new modeling formulations, especially since the Schur complement is a fundamental tool wherever semidefinite optimization is applied, see [4, 7, 30] for some examples. Using Siegel cones instead of the Schur complement would greatly improve the complexity of the models. This should serve as an incentive to develop software for homogeneous conic optimization problems. Another approach to solve these problems is to make use of specialized linear algebra to solve system ($\mathrm{ED}_{\mathrm{Lor}}$) effectively.

Acknowledgements The idea to consider homogeneous cones in the context of strong conic duality was inspired by a seminar talk by Chek Beng Chua at McMaster University in May 2006. Discussions with Gábor Pataki were quite influential during the preparation of the manuscript.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Mathematical Programming **95**(1), 3–51 (2003)
2. Barker, G.P.: The lattice of faces of a finite dimensional cone. Linear Algebra and Its Applications **7**, 71–82 (1973)
3. Barker, G.P., Carlson, D.: Cones of diagonally dominant matrices. Pacific Journal of Mathematics **57**(1), 15–32 (1975)
4. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. MPS-SIAM Series on Optimization. SIAM, Philadelphia, PA (2001)
5. Bonnesen, T., Fenchel, W.: Theorie der konvexen Körper. Chelsea Publishing Company, New York (1948)
6. Borwein, J., Wolkowicz, H.: Regularizing the abstract convex program. Journal of Mathematical Analysis and Applications **83**, 495–530 (1981)
7. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V.: Linear Matrix Inequalities in System and Control Theory, *SIAM Studies in Applied Mathematics*, vol. 15. SIAM, Philadelphia, PA (1994)
8. Chua, C.B.: Relating homogeneous cones and positive cones via T-algebras. SIAM Journal on Optimization **14**(2), 500–506 (2003)
9. Dubins, L.E.: On extreme points of convex sets. J. Math. Anal. Appl. **5**, 237–244 (1962)
10. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs. Oxford University Press (1994)
11. Faybusovich, L.: On Nesterov's approach to semi-infinite programming. Acta Applicandae Mathematicae **74**, 195–215 (2002)

[6] We strongly believe such a function exists.

12. Gindikin, S.G.: Tube Domains and the Cauchy Problem, *Translation of Mathematical Monographs*, vol. 111. AMS, Providence, RI (1992)
13. Güler, O., Tunçel, L.: Characterization of the barrier parameter of homogeneous convex cones. Math. Programming **81**, 55–76 (1998)
14. Kaneyuki, S., Tsuyi, T.: Classification of homogeneous bounded domains of lower dimension. Nagoya Math. Journal **53**, 1–46 (1974)
15. de Klerk, E., Roos, C., Terlaky, T.: Nonlinear Optimization. Lecture notes, Delft University of Technology, Delft, The Netherlands (2003)
16. Nesterov, Y., Nemirovski, A.: Interior-Point Polynomial Algorithms in Convex Programming. SIAM Publications, Philadelphia, PA (1994)
17. Pataki, G.: A partial characterization of nice cones. Forthcoming.
18. Pataki, G.: A simple derivation of a facial reduction algorithm and extended dual systems. Technical report, Columbia University (1996). In preparation.
19. Pataki, G.: On the geometry of cone LPs. In: H. Wolkowicz, L. Vandenberghe, R. Saigal (eds.) Handbook of Semidefinite Programming: Theory, Algorithms, and Applications. Kluwer (2000)
20. Pataki, G.: On the closedness of the linear image of a closed convex cone. Mathematics of Operations Research (2007). Accepted for publication.
21. Ramana, M.V.: An exact duality theory for semidefinite programming and its complexity implications. Math. Programming (77), 129–162 (1997)
22. Ramana, M.V., Freund, R.M.: On the ELSD duality theory for SDP (1996). Unpublished manuscript.
23. Ramana, M.V., Tunçel, L., Wolkowicz, H.: Strong duality for semidefinite programming. SIAM Journal on Optimization **7**(3), 641–662 (1997)
24. Renegar, J.: A Mathematical View of Interior-Point Methods in Convex Optimization. MPS/SIAM series on optimization. SIAM (2001)
25. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton, N. J. (1970)
26. Schurr, S.P., O'Leary, D.P., Tits, A.: A polynomial-time interior-point method for conic optimization, with inexact barrier evaluations. Technical report, University of Maryland (2007)
27. Stoer, J., Witzgall, C.: Convexity and Optimization in Finite Dimensions, vol. I. Springer-Verlag, Heidelberg (1970)
28. Vinberg, E.B.: The theory of convex homogeneous cones. Transactions of the Moscow Mathematical Society **12**, 340–403 (1963)
29. Vinberg, E.B.: The structure of the group of automorphisms of a homogeneous convex cone. Transactions of the Moscow Mathematical Society **13**, 63–93 (1965)
30. Zhang, F. (ed.): The Schur Complement and its Applications, *Numerical Methods and Algorithms*, vol. 4. Springer (2005)