A D-Induced Duality and Its Applications

Jan Brinkhuis * Shuzhong Zhang †

July 2002; revised May 2003

Abstract

This paper attempts to extend the notion of duality for convex cones, by basing it on a predescribed conic ordering and a fixed bilinear mapping. This is an extension of the standard definition of dual cones, in the sense that the *nonnegativity* of the inner-product is replaced by a pre-specified conic ordering, defined by a convex cone D, and the inner-product itself is replaced by a general multi-dimensional bilinear mapping. This new type of duality is termed the D-*induced duality* in the paper. We further introduce the notion of D-induced polar sets within the same framework, which can be viewed as a generalization of the D-induced dual cones and are convenient to use for some practical applications. Properties of the extended duality, including the extended bi-polar theorem, are proven. Furthermore, attention is paid to the computation and approximation of the D-induced dual objects. We discuss, as examples, applications of the newly introduced D-induced duality concepts in robust conic optimization and the duality theory for multi-objective conic optimization.

Keywords: Convex cones, duality, bi-polar theorem, conic optimization.

AMS subject classification: 46A20, 46N10, 90C25.

*Econometric Institute, Erasmus University, The Netherlands.

†Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported in part by the Foundation 'Vereniging Trustfonds Erasmus Universiteit Rotterdam' in The Netherlands, and in part by Hong Kong RGC Earmarked Grants CUHK4181/00E and CUHK4233/01E.

1 Introduction

Duality plays a central role in the development of the theory as well as the solution methods for optimization. A good example is the success of the so-called primal-dual interior point methods for conic convex optimization; see, e.g., [9].

A dual object, say, the dual of a convex cone, is defined as the set of all nonnegative linear mappings (functionals) over the cone under consideration. This set itself forms a convex cone, which has a lot of intimate and interesting relationships with the original cone. Duality theory is devoted to reveal the nature of the relationship, and beyond any doubt it has become the foundation of optimization. However, there are circumstances where the concept of nonnegativity needs to be extended in order to better suit some new and interesting applications, arising, e.g., from robust analysis of conic optimization. In this paper we introduce a new type of duality for convex cones, where the nonnegativity is induced by an arbitrary given convex cone, which is obviously a generalization of the usual definition of the dual. This usual definition is the special case that the given convex cone is the nonnegative half-line $\Re_+$. We show that under some conditions, important results such as the bi-polar theorem can be carried over. A key issue is the characterization of the new kind of dual cone, termed as the D-*induced dual cone* in this paper. It is linked naturally to other important issues in convex analysis. For instance, it raises questions such as how to compute the tensor product of two convex cones, and what is the calculus rule for the duality operation (in the ordinary sense) for the tensor product of two convex cones. We believe that this triggers interesting research questions to be answered in the future. The concept is further generalized to the setting of D-*induced polar set*. Although the D-induced polar set can be viewed as an inhomogeneous version of the D-induced cone, its direct form is handy to use for practical purposes, and it has interesting properties of its own. Applications arising from robust optimization and multi-objective optimization are discussed. We believe the potentials for the D-induced duality abound. As a key issue we point out that it is crucial to study the approximative computation of a D-induced dual object, as the exact computation is often NP-hard.

This paper is organized as follows. We shall introduce the new type of duality in Section 2. In the same section we show several properties of the new duality operation, including the bi-polar theorem and several calculus rules of the new duality operation. In Section 3 we discuss the related polar operations and their properties. Section 4 is concerned with how to compute and approximate of the new dual objects. In Section 5, we discuss two applications, one from the robust version of conic convex optimization, and the other from multiple objective conic convex optimization. Then, in

Section 6 we discuss how the new type of duality can be characterized and computed from a totally different point of view: the epigraph representation of convex cones. Finally, we conclude the paper in Section 7.

Notations: In most places, letters in calligraphic style, e.g. $\mathcal{X}$, denote vector spaces; $\Re^n$ is n-dimensional Euclidean space; $\Re^n_+$ is the set of all n-dimensional non-negative vectors; $\Re^n_{++}$ is the set of all n-dimensional positive vectors; $\mathcal{S}^m$ is the space of all m by m symmetric real-valued matrices; $\mathcal{S}^m_+$ is the set of all m by m positive semidefinite matrices ($X \in \mathcal{S}^m_+$ is equivalent to $X \succeq 0$); $\mathcal{S}^m_{++}$ is the set of all m by m positive definite matrices; $\| \cdot \|$ is the Euclidean norm with appropriate dimension from the context; cl (S) stands for the closure of the set S; conv (S) stands for the convex hull of the set S; epi (f) stands for the epigraph of the function f; SOC(n) is the standard n-dimensional second order cone, i.e., $\text{SOC}(n) = \left\{ [x_1, x_2, \cdots, x_n]^T \mid x_1 \ge \sqrt{\sum_{i=2}^n x_i^2} \right\}$; $e_i^{(n)}$ indicates the ith unit vector in $\Re^n$, namely the vector whose ith component is 1 while all other components are 0. Finally, vec (A) stands for the vector obtained by stacking together the columns of the matrix A, i.e., vec $(A) = [a_1^T, a_2^T, \cdots, a_m^T]^T$ where $[a_1, a_2, \cdots, a_m] = A$.

2 The D-induced dual cone

Consider three vector spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{W}$.

Let $\mathsf{D} \subseteq \mathcal{W}$ be a certain fixed convex cone. We assume that D is not a linear subspace; this is equivalent to demanding that there is $d \in \mathsf{D}$ such that $-d \notin \mathsf{D}$. This will be called a *non-flat* cone. Due to the convexity of D, by a separation argument this condition further implies the existence of $d \in \mathsf{D}$ such that $-d \notin \text{cl}\, \mathsf{D}$.

Let

$$\langle x, y \rangle : \ (x, y) \in \mathcal{X} \times \mathcal{Y} \to \mathcal{W}$$

be a given bilinear mapping, i.e., for any fixed $x \in \mathcal{X}$, $\langle x, \cdot \rangle : \ \mathcal{Y} \to \mathcal{W}$ is a linear mapping, and for any fixed $y \in \mathcal{Y}$, $\langle \cdot, y \rangle : \ \mathcal{X} \to \mathcal{W}$ is a linear mapping as well.

In this paper, we assume throughout, for the sake of simplicity, that all vector spaces under consideration are finite dimensional, although some of the results can be easily extended to a more general setting. Moreover we choose for each vector space an inner product: we denote the chosen inner product on $\mathcal{X}$ by $\langle \cdot, \cdot \rangle_{\mathcal{X}}$; similarly for $\mathcal{Y}$ and $\mathcal{W}$. In particular, we assume after suitable choices of coordinates that $\mathcal{X} = \Re^n$, $\mathcal{Y} = \Re^m$, and $\mathcal{W} = \Re^k$ and that the inner products are the usual inner product, viz. the sum of coordinate-wise products, e.g. $\langle x, y \rangle_{\mathcal{X}} = x^T y$ for $x, y \in \mathcal{X}$. In our applications we consider often spaces of matrices; here the choice of coordinates means that we stack the matrices into column-vectors as described before.

Let $U \subseteq \mathcal{X}$ be a cone.

The dual of U as induced by D under $\langle \cdot, \cdot \rangle$ is a convex cone in $\mathcal{Y}$, defined by

$$\begin{aligned} U^*_{\mathsf{D}} &= \{ y \in \mathcal{Y} \mid \langle x, y \rangle \in \mathsf{D} \text{ for all } x \in U \} \\ &= \{ y \in \mathcal{Y} \mid \langle U, y \rangle \subseteq \mathsf{D} \}. \end{aligned}$$

In other words, the D-induced dual cone of U is the collection of all linear mappings that take U to D under $\langle x, \cdot \rangle$. Obviously, U^*_{D} is always a closed convex cone provided that D is closed. However, in general D does not have to be closed.

Symmetrically, for a cone V in the space $\mathcal{Y}$, its D-induced dual under the bilinear mapping $\langle \cdot, \cdot \rangle$ is a convex cone in $\mathcal{X}$, defined by

$$\begin{aligned} V^*_{\mathsf{D}} &= \{ x \in \mathcal{X} \mid \langle x, y \rangle \in \mathsf{D} \text{ for all } y \in V \} \\ &= \{ x \in \mathcal{X} \mid \langle x, V \rangle \subseteq \mathsf{D} \}. \end{aligned}$$

Therefore, when we speak of the D-induced dual, it is of importance to specify the space in which the cone in question resides.

It is also evident that there are two key factors in this definition, namely the order-defining cone D and the bilinear mapping $\langle \cdot, \cdot \rangle$. We recall that the dual cone in the ordinary sense of a cone $U \subseteq \mathcal{X}$ is defined to be the set of all $x \in \mathcal{X}$ such that $\langle x, u \rangle_{\mathcal{X}} \geq 0$, $\forall u \in U$. As each linear function on $\mathcal{X}$ can be written as $x \rightarrow \langle a, x \rangle_{\mathcal{X}}$ for a unique $a \in \mathcal{X}$, the dual cone in the ordinary sense is $\Re_+$-induced with the usual inner product $\langle x, y \rangle_{\mathcal{X}} = x^T y$ as the underlying bilinear mapping. Due to the symmetric form of this bilinear mapping, the usual dual cone need not to be further specified as whether it is in the space $\mathcal{X}$ or in the space $\mathcal{Y}$.

In general, using an appropriate coordinate system, any finite-dimensional bilinear mapping can be specified as

$$\langle x, y \rangle = \begin{bmatrix} x^T A_1 y \\ \vdots \\ x^T A_k y \end{bmatrix},$$

where $A_i \in \Re^{n \times m}$, $i = 1, ..., k$.

We note that since a cone contains the origin, any D-induced dual cone must as well contain the origin; thus it is non-empty. Moreover its closure is identical to the (cl D)-induced counter-part. This is formalized in the following proposition.

Proposition 2.1 *Let* D *be a convex cone. Let* $U \subseteq \mathcal{X}$*. It holds that* $\mathrm{cl}\, U^*_{\mathsf{D}} = U^*_{\mathrm{cl}\, \mathsf{D}}$.

Proof. It is obvious that $U^*_{\mathsf{D}} \subseteq U^*_{\mathrm{cl}\, \mathsf{D}}$. Taking closure on both sides yields $\mathrm{cl}\, U^*_{\mathsf{D}} \subseteq U^*_{\mathrm{cl}\, \mathsf{D}}$.

Note that $U^*_{\mathsf{D}} \neq \emptyset$. Take an arbitrary $y \in U^*_{\mathrm{cl}\, \mathsf{D}}$. We have $\langle U, y \rangle \subseteq \mathrm{cl}\, \mathsf{D}$. Suppose by contradiction that $y \notin \mathrm{cl}\, U^*_{\mathsf{D}}$. Let $\hat{y}$ be the projection of y on $\mathrm{cl}\, U^*_{\mathsf{D}}$, and $\|y - \hat{y}\| = \delta > 0$. Due to the first part of the proof, we know that $\hat{y} \in U^*_{\mathrm{cl}\, \mathsf{D}}$; that is, $\langle U, \hat{y} \rangle \subseteq \mathrm{cl}\, \mathsf{D}$. We claim that $(y + \hat{y})/2 \in U^*_{\mathrm{cl}\, \mathsf{D}}$. Let us check this. Choose an infinite sequence $\{y_n \mid n = 1, 2, ...\}$ in U^*_{D} which tends to $\hat{y}$. Then the sequence $\{\frac{y_n + y}{2} \mid n = 1, 2, ...\}$ is contained in U^*_{D} by the convexity of this set and it tends to $\frac{\hat{y}+y}{2}$. This proves the inclusion $(y + \hat{y})/2 \in U^*_{\mathrm{cl}\, \mathsf{D}}$. Moreover, $\|y - (y + \hat{y})/2\| = \delta/2$, contradicting the fact that $\hat{y}$ is the projection of y onto $\mathrm{cl}\, U^*_{\mathsf{D}}$. Thus we must have $y \in \mathrm{cl}\, U^*_{\mathsf{D}}$. Hence, $U^*_{\mathrm{cl}\, \mathsf{D}} \subseteq \mathrm{cl}\, U^*_{\mathsf{D}}$. The proposition is proven. **Q.E.D.**

A key result concerning the D-induced duality is the extended bi-polar theorem. Before presenting this result, let us first introduce the following notion of surjectivity.

Definition 2.2 *Consider the bilinear mapping*

$$\langle x, y \rangle := \begin{bmatrix} x^T A_1 y \\ \vdots \\ x^T A_k y \end{bmatrix} : (x, y) \in \mathcal{X} \times \mathcal{Y} \to \mathcal{W}.$$

We call $\langle \cdot, \cdot \rangle$ *dual surjective with respect to* D *if for any* $a \in \mathcal{X}$ *there is a non-flat direction* $b \in \mathsf{D}$ *(*$-b \notin \mathrm{cl}\, \mathsf{D}$*) and an element* $y \in \mathcal{Y}$ *such that the linear equation*

$$[A_1 y, \cdots, A_k y] = a b^T$$

is satisfied. This concept does not depend on the choice of coordinates as can be seen from the following coordinate-free description: $\langle \cdot, \cdot \rangle$ *is dual surjective with respect to* D *if and only if*

$$\forall a \in \mathcal{X},\ \exists \text{ non-flat } b \in \mathsf{D},\ \exists y \in Y,\ \text{ such that } \langle x, y \rangle = \langle x, a \rangle_{\mathcal{X}} b \text{ for all } x \in \mathcal{X}.$$

Similarly, we call $\langle \cdot, \cdot \rangle$ *primal surjective with respect to* D *if for any* $d \in \mathcal{Y}$ *there is a non-flat direction* $c \in \mathsf{D}$ *(*$-c \notin \mathrm{cl}\, \mathsf{D}$*) and* $x \in \mathcal{X}$ *such that the linear equation*

$$\begin{bmatrix} x^T A_1 \\ \vdots \\ x^T A_k \end{bmatrix} = c d^T$$

is satisfied.

We remark here that the coordinate-free description is handy for the purpose of checking the condition in many applications.

Now we are in a position to state the following extended bi-polar theorem for the D-induced duality.

Theorem 2.3 *Let the bilinear mapping*

$$\langle x, y\rangle := \begin{bmatrix} x^T A_1 y \\ \vdots \\ x^T A_k y \end{bmatrix} : (x,y) \in \mathcal{X} \times \mathcal{Y} \to \mathcal{W}$$

be fixed. Let $\mathsf{D} \subseteq \mathcal{W}$ be a given non-flat convex cone. Suppose that $\langle\cdot,\cdot\rangle$ is dual surjective with respect to D. Let $U \subseteq \mathcal{X}$ be a convex cone. Then it holds that

$$\text{cl } U^{**}_{\mathsf{DD}} = \text{cl } U.$$

Proof. First we prove $U \subseteq U^{**}_{\mathsf{DD}}$.

Take any $x \in U$. Then by definition, $\langle x, y\rangle \in \mathsf{D}$ for all $y \in U^*_{\mathsf{D}}$. Hence, $x \in U^{**}_{\mathsf{DD}}$, and so it follows that $U \subseteq U^{**}_{\mathsf{DD}}$. Consequently, it follows that cl $U \subseteq$ cl U^{**}_{DD}.

Next we shall prove $U^{**}_{\mathsf{DD}} \subseteq$ cl U.

Take any $\hat{x} \in U^{**}_{\mathsf{DD}}$. Thus $\langle \hat{x}, y\rangle \in \mathsf{D}$ for all $y \in U^*_{\mathsf{D}}$.

Let $L(y) := [A_1 y, \cdots, A_k y]$. Observe that

$$\begin{aligned} y \in U^*_{\text{cl } \mathsf{D}} &\iff \langle u, y\rangle \in \text{cl } \mathsf{D} \text{ for all } u \in U \\ &\iff [u^T A_1 y, \cdots, u^T A_k y]v \geq 0 \text{ for all } u \in U \text{ and } v \in \mathsf{D}^* \\ &\iff u^T L(y) v \geq 0 \text{ for all } u \in U \text{ and } v \in \mathsf{D}^*. \end{aligned}$$

In other words,

$$U^*_{\text{cl } \mathsf{D}} = \{y \mid u^T L(y) v \geq 0, \text{ for all } u \in U,\, v \in \mathsf{D}^*\}. \tag{1}$$

Suppose by contradiction that $\hat{x} \notin$ cl U. Then by the separation theorem, there exists $a \in U^*$ such that $a^T \hat{x} < 0$. By the dual surjectivity of the bilinear mapping, we can find $\hat{y}$ such that

$$L(\hat{y}) = ab^T$$

where $b \in \mathsf{D}$ is a non-flat direction. Hence

$$u^T L(\hat{y}) v = (u^T a)(b^T v) \geq 0$$

for all $u \in U$ and $v \in \mathsf{D}^*$ as $a \in U^*$ and $b \in \mathsf{D}$. Consequently, $\hat{y} \in U^*_{\text{cl } \mathsf{D}}$. This leads us to the following contradiction. On the one hand, $\langle \hat{x}, \hat{y} \rangle \in \text{cl } \mathsf{D}$ due to the fact that $\hat{x} \in U^{**}_{\mathsf{DD}}$ and $\hat{y} \in U^*_{\text{cl } \mathsf{D}} = \text{cl } U^*_{\mathsf{D}}$, where we used Proposition 2.1. On the other hand,

$$\langle \hat{x}, \hat{y} \rangle = (\hat{x}^T a b^T)^T = (a^T \hat{x}) b \notin \text{cl } \mathsf{D},$$

due to the fact that $a^T \hat{x} < 0$ and b is a non-flat direction for D. This proves $U^{**}_{\mathsf{DD}} \subseteq \text{cl } U$. The desired result follows by taking closure on both sides. **Q.E.D.**

In the same vein, we have an analogue for the dual space.

Theorem 2.4 *Let the bilinear mapping*

$$\langle x, y \rangle := \begin{bmatrix} x^T A_1 y \\ \vdots \\ x^T A_k y \end{bmatrix} : (x, y) \in \mathcal{X} \times \mathcal{Y} \to \mathcal{W}$$

be fixed. Let $\mathsf{D} \subseteq \mathcal{W}$ be a given non-flat convex cone. Suppose that $\langle \cdot, \cdot \rangle$ is primal surjective with respect to D. Let $V \subseteq \mathcal{Y}$ be a convex cone. Then it holds that

$$\text{cl } V^{**}_{\mathsf{DD}} = \text{cl } V.$$

As a matter of notation, let us introduce the tensor product of two convex cones C and D as follows

$$C \otimes D = \text{conv } \{uv^T \mid u \in C, \, v \in D\}. \tag{2}$$

We call its dual to be the *bi-positive* cone, denoted by

$$\mathcal{B}(C, D) = \{Z \mid u^T Z v \geq 0 \text{ for all } u \in C, \, v \in D\}. \tag{3}$$

Indeed it is elementary to see that

$$(C \otimes D)^* = \mathcal{B}(C, D). \tag{4}$$

A proof for the above equation and other related equations can also be found in [5].

In the proof for Theorem 2.3 we in fact established the following relation; see (1). Let us formalize it as follows, now using the notion of the bi-positive cone.

Proposition 2.5 *Let $U \subseteq \mathcal{X}$ be a cone. Then it holds that*

$$U^*_{\mathrm{cl}\ \mathsf{D}} = \{y \mid L(y) \in \mathcal{B}(U, \mathsf{D}^*)\}.$$

Thus the following result is straightforward.

Proposition 2.6 *Consider convex cones $U_1, \cdots, U_r \subseteq \mathcal{X}$. It holds that*

$$(U_1 + \cdots + U_r)^*_{\mathrm{cl}\ \mathsf{D}} = \bigcap_{i=1}^{r} (U_i)^*_{\mathrm{cl}\ \mathsf{D}}.$$

Proof. According to Proposition 2.5, we have

$$\begin{aligned}
(U_1 + \cdots + U_r)^*_{\mathrm{cl}\ \mathsf{D}} &= \{y \mid L(y) \in \mathcal{B}(U_1 + \cdots + U_r, \mathsf{D}^*)\} \\
&= \{y \mid L(y) \in \bigcap_{i=1}^{r} \mathcal{B}(U_i, \mathsf{D}^*)\} \\
&= \bigcap_{i=1}^{r} (U_i)^*_{\mathrm{cl}\ \mathsf{D}}.
\end{aligned}$$

Q.E.D.

Unlike in the usual duality case, the primal and dual status of the D-induced duality is not symmetric in general. For instance, in case that the bilinear mapping is dual surjective, then the 'dual space' $\mathcal{Y}$ is bigger in some sense. Therefore, it can happen that not all convex cones in $\mathcal{Y}$ can be expressed as the dual of some cone in $\mathcal{X}$. However, if the bilinear mapping is both primal and dual surjective, then one may apply the bi-polar theorem on both sides. As a consequence, the following result follows.

Corollary 2.7 *Suppose that $\langle \cdot, \cdot \rangle$ is both primal and dual surjective with respect to D and that D is closed. Let $U_1, \cdots, U_r$ be arbitrary convex cones in $\mathcal{X}$. Then it holds that*

$$\mathrm{cl}\ \left(\bigcap_{i=1}^{r} \mathrm{cl}\ U_i \right)^*_{\mathsf{D}} = \mathrm{cl}\ \left((U_1)^*_{\mathsf{D}} + \cdots + (U_r)^*_{\mathsf{D}} \right).$$

Proof. The desired result follows immediately if we replace U_i in Proposition 2.6 by $(U_i)^*_{\mathsf{D}}$, $i = 1, ..., r$, and then, on both sides of the resulting identity, take closure and apply Theorem 2.3 from the dual side. **Q.E.D.**

It is in fact quite rare that both the primal and the dual surjectivity conditions are satisfied at the same time. This essentially means that we are dealing with the ordinary duality with the usual inner product and $\mathsf{D} = \Re_+$.

Other simple calculus rules for the cone-induced duality are presented below.

Proposition 2.8 *Let U be a convex cone in $\mathcal{X}$.*

(i) *For any convex cones $\mathsf{D}_1, ..., \mathsf{D}_s$ in $\mathcal{W}$, it holds that*

$$U^*_{\mathsf{D}_1 \cap \cdots \cap \mathsf{D}_s} = U^*_{\mathsf{D}_1} \cap \cdots \cap U^*_{\mathsf{D}_s}$$

and,

(ii)

$$U^*_{\mathsf{D}_1 + \cdots + \mathsf{D}_s} \supseteq U^*_{\mathsf{D}_1} + \cdots + U^*_{\mathsf{D}_s}.$$

(iii) *Suppose that the bilinear mapping is decomposed as*

$$\langle x, y \rangle = \begin{bmatrix} \langle x, y \rangle_1 \\ \vdots \\ \langle x, y \rangle_p \end{bmatrix},$$

where $\langle x, y \rangle_i$ is a bilinear mapping from $\mathcal{X} \times \mathcal{Y}$ to $\mathcal{W}_{k_i}$ where $\mathcal{W}_{k_i} = \Re^{k_i}$ with $\sum_{i=1}^p k_i = k$. Furthermore, suppose that convex cones $\mathsf{D}_i \subseteq \mathcal{W}_{k_i}$ are given, $i = 1, ..., p$. Then it holds that

$$U^*_{\mathsf{D}_1 \times \cdots \times \mathsf{D}_p} = U^*_{\mathsf{D}_1} \cap \cdots \cap U^*_{\mathsf{D}_p}$$

where the set product is defined as

$$\mathsf{D}_1 \times \cdots \times \mathsf{D}_p = \{(z_1, \cdots, z_p) \mid z_i \in \mathsf{D}_i,\ i = 1, ..., p\}.$$

Proof. To show **(i)** we note that

$$\begin{aligned} y \in U^*_{\mathsf{D}_1 \cap \cdots \cap \mathsf{D}_s} &\iff \langle U, y \rangle \in \mathsf{D}_i,\ i = 1, ..., s \\ &\iff y \in U^*_{\mathsf{D}_i},\ i = 1, ..., s \\ &\iff y \in U^*_{\mathsf{D}_1} \cap \cdots \cap U^*_{\mathsf{D}_s}. \end{aligned}$$

For proving **(ii)** we note that if $y \in U^*_{\mathsf{D}_1} + \cdots + U^*_{\mathsf{D}_s}$ then there exist $y_1, \cdots, y_s$ such that $y = y_1 + \cdots + y_s$, and $y_i \in U^*_{\mathsf{D}_i}$, $i = 1, ..., s$. This implies that $\langle U, y_i \rangle \subseteq \mathsf{D}_i$, $i = 1, ..., s$, and so

$$\langle U, y \rangle \subseteq \langle U, y_1 \rangle + \cdots + \langle U, y_s \rangle \subseteq \mathsf{D}_1 + \cdots + \mathsf{D}_s.$$

Thus, $U^*_{\mathsf{D}_1+\cdots+\mathsf{D}_s} \supseteq U^*_{\mathsf{D}_1} + \cdots + U^*_{\mathsf{D}_s}$.

Note that the above inclusion is strict in general. For more discussions on this, see Section 6.

Now we prove **(iii)**. Similarly as in **(i)**,

$$\begin{aligned} y \in U^*_{\mathsf{D}_1\times\cdots\times\mathsf{D}_p} &\iff \langle U, y\rangle_i \in \mathsf{D}_i,\ i = 1, ..., p \\ &\iff y \in U^*_{\mathsf{D}_i},\ i = 1, ..., p \\ &\iff y \in U^*_{\mathsf{D}_1} \cap \cdots \cap U^*_{\mathsf{D}_p}. \end{aligned}$$

Remark here that the statements in **(i)** and **(iii)** are different: in **(i)**, the D_i's are all contained in $\Re^k$, while in **(iii)**, D_i is in $\Re^{k_i}$, $i = 1, ..., p$. **Q.E.D.**

By **(iii)** of Proposition 2.8 it is clear that one needs only to concentrate on the case that the underlying cone D is non-decomposable, for the duality consideration.

Let us see how the D-induced duality actually works.

Example. Consider

$$\mathcal{X} = \mathcal{S}^n,\ \mathcal{Y} = \left\{ \left[\begin{array}{cc} Y_{11} & Y_{12} \\ Y_{12} & Y_{22} \end{array} \right] \middle|\ Y_{12} \in \mathcal{S}^n \right\},\ \mathcal{W} = \mathcal{S}^2.$$

The bilinear mapping is

$$\langle X, Y\rangle = \left[\begin{array}{cc} X \bullet Y_{11} & X \bullet Y_{12} \\ X \bullet Y_{12} & X \bullet Y_{22} \end{array} \right]$$

and $\mathsf{D} = \mathcal{S}^2_+$, where $\bullet$ denotes the usual entry-wise inner product for matrices.

In this case, the bilinear mapping is dual surjective. To see this, let us fix a non-flat direction of D, e.g. $B = \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right]$. Then, for any $A \in \mathcal{X}$ the choice $Y = \left[\begin{array}{cc} A & A \\ A & A \end{array} \right]$ satisfies the dual surjective requirement: $\langle X, Y\rangle = (X \bullet A)B$ for all $X \in \mathcal{X}$.

By Theorem 4.5 of Luo, Sturm and Zhang [5] we have

$$(\mathcal{S}^n_+)^*_{\mathcal{S}^2_+} = \left\{ \left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{array} \right] \succeq 0 \middle|\ Z_{12} \in \mathcal{S}^n \right\}.$$

The extended bi-polar theorem, Theorem 2.3, thus asserts that

$$\left\{ \left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{array} \right] \succeq 0 \middle|\ Z_{12} \in \mathcal{S}^n \right\}^*_{\mathcal{S}^2_+} = \mathcal{S}^n_+.$$

3 The D-induced polar set

It is useful to extend the notion of D-*induced dual cone* to a general D-*induced polar set.* Following the ordinary duality theory, the polar operation can be considered as a truncation operation after the homogenization and the dual cone operations; the exact meaning of this statement will become clear later. However, in our D-induced duality case, the D-induced polar is flexible enough to simply include the D-induced dual cone as a special case. Certainly, we may also view the D-induced polar set as a truncation of the D-induced dual cone in a homogenized space. Hence, in this respect, these two notions are equivalent. For practical purposes however, the D-induced polar set is important on its own right.

Before we discuss the D-induced polar, let us briefly introduce the usual polar duality theory for the sake of making comparisons. For details of duality theory, use Rockafellar [10] as a standard reference.

Let $S \in \mathcal{X}$ be a convex set. Its polar is a closed convex set defined as $S^o = \{y \mid \langle x, y\rangle_{\mathcal{X}} + 1 \geq 0 \text{ for all } x \in S\}$. The standard bi-polar theorem in this context states: if $0 \in S$ then $(S^o)^o = \text{cl } S$. It is also well known that the ordinary conic form bi-polar theorem, i.e. Theorem 2.4 with $\langle x, y\rangle = \langle x, y\rangle_{\mathcal{X}}$ and $\mathsf{D} = \Re_+$, leads to the above statement. To see this, consider first the following homogenization operation. Let $\bar{\mathcal{X}} := \Re \times \mathcal{X}$. For a given set $S \in \mathcal{X}$ define

$$H(S) := \text{cl } \left\{ \left[\begin{array}{c} x_0 \\ x \end{array} \right] \middle| \, x_0 > 0, \, x/x_0 \in S \right\} \subseteq \bar{\mathcal{X}}.$$

Let us also define the following *anti-polar* set:

$$S^a = \{y \mid \langle x, y\rangle_{\mathcal{X}} + 1 \leq 0 \text{ for all } x \in S\}.$$

Obviously, if $0 \in S$ then $S^a = \emptyset$.

We now prove a general relationship between the dual cone and the polar/anti-polar sets.

Theorem 3.1 *It holds that* $H(S)^* = \text{cl } (H(S^o) - H(S^a))$.

Proof. Take any $[y_0, y^T]^T \in (H(S^o) - H(S^a))$ with

$$[y_0, y^T]^T = [y_0', y'^T]^T - [y_0'', y''^T]^T,$$

where $y_0' > 0$, $y_0'' > 0$, $y'/y_0' \in S^o$ and $y''/y_0'' \in S^a$. Then, for any $[x_0, x^T]^T$ with $x_0 > 0$ and $x/x_0 \in S$ we have

$$x_0 y_0 + x^T y \;=\; x_0(y_0' - y_0'') + x^T(y' - y'')$$

$$
\begin{aligned}
&= x_0 y_0' \left[\left(\frac{x}{x_0}\right)^T \left(\frac{y'}{y_0'}\right) + 1 \right] - x_0 y_0'' \left[\left(\frac{x}{x_0}\right)^T \left(\frac{y''}{y_0''}\right) + 1 \right] \\
&\geq 0,
\end{aligned}
$$

leading to the fact that $[y_0, y^T]^T \in H(S)^*$. Taking closure we have $\text{cl } (H(S^o) - H(S^a)) \subseteq H(S)^*$.

Next we consider the other containing relation.

Take any $[y_0, y^T]^T \in H(S)^*$. We have $x_0 y_0 + x^T y \geq 0$ for all $x_0 > 0$ and $x/x_0 \in S$. If $y_0 > 0$ then $x^T(y/y_0) + 1 \geq 0$ for all $x \in S$, thus $[y_0, y^T]^T \in H(S^o)$. If $y_0 = 0$ then we may take an arbitrary $\epsilon > 0$ and conclude that $\epsilon + x^T y \geq x^T y \geq 0$ for all $x \in S$. Therefore, $[\epsilon, y^T]^T \in H(S^o)$. Letting $\epsilon \downarrow 0$ yields $[y_0, y^T]^T = [0, y^T]^T \in H(S^o)$. If $y_0 < 0$ then $x^T \left(\frac{-y}{-y_0}\right) + 1 \leq 0$ for all $x \in S$, implying that $-[y_0, y^T]^T \in H(S^a)$. Therefore, $H(S)^* \subseteq \text{cl } (H(S^o) - H(S^a))$. **Q.E.D.**

Corollary 3.2 *If* $0 \in S$ *then* $H(S)^* = H(S^o)$ *and* $S^{oo} = \text{cl } S$.

Proof. If $0 \in S$ then $S^a = \emptyset$ and so by Theorem 3.1 we have $H(S)^* = H(S^o)$. Now we use the standard bi-polar theorem for convex cones to obtain

$$H(S) = H(S)^{**} = (H(S)^*)^* = (H(S^o))^* = H(S^{oo}).$$

Therefore, $S^{oo} = \text{cl } S$. **Q.E.D.**

Next we proceed to consider the D-induced polar sets. As before, let $\langle x, y \rangle : (x, y) \in \mathcal{X} \times \mathcal{Y} \to \mathcal{W}$ be the bilinear mapping under consideration, and using a suitable coordinate system we may assume

$$\langle x, y \rangle = \begin{bmatrix} x^T A_1 y \\ \vdots \\ x^T A_k y \end{bmatrix}.$$

Let D be a given convex set in $\mathcal{W}$. For simplicity, in this section we assume D to be a closed set. Let $S \in \mathcal{X}$. We call

$$(S)^o_{\mathsf{D}} = \{ y \in \mathcal{Y} \mid \langle x, y \rangle \in \mathsf{D} \text{ for all } x \in S \} \tag{5}$$

the D*-induced polar* of S. Clearly, the D-induced polar set is convex. However it can be empty if no additional assumption is imposed. In most of our analysis, we assume $0 \in \mathsf{D}$ and $0 \in S$, and so $0 \in (S)^o_{\mathsf{D}} \neq \emptyset$. Notice that the ordinary polar corresponds to $\mathsf{D} = [-1, \infty)$. In case D itself is a convex cone, the D-induced polar set of a convex cone U is simply its D-induced dual cone. In this sense, unlike the ordinary polarity, the D-induced polar includes the D-induced dual cone as a special case. We shall see below that these notions are actually equivalent, under the homogenization operation. It is evident that the calculus rules as established in Proposition 2.8 are valid in its original form for the D-induced polar sets as well. However, Proposition 2.6 needs a slight modification to accommodate the possibility that the origin may not be contained in the sets.

Proposition 3.3 *Let $S_i \subseteq \mathcal{X}$ be convex sets, and $a_i \in S_i$, $i = 1, ..., r$. Then,*

$$\left(\sum_{i=1}^{r} S_i\right)^o_{\mathsf{D}} = \bigcap_{i=1}^{r} \left((\sum_{i=1}^{r} a_i - a_i) + S_i\right)^o_{\mathsf{D}}.$$

Let $\bar{\mathcal{Y}} := \Re \times \mathcal{Y}$ and $\bar{\mathcal{W}} := \Re \times \mathcal{W}$. For $\bar{x} = [x_0, x^T]^T \in \bar{\mathcal{X}}$ and $\bar{y} = [y_0, y^T]^T \in \bar{\mathcal{Y}}$, consider the following extended bilinear mapping

$$\langle \bar{x}, \bar{y} \rangle := \left[\begin{array}{c} x_0 y_0 \\ \langle x, y \rangle \end{array} \right]. \tag{6}$$

The order-defining cone is naturally chosen to be $H(\mathsf{D})$. The first result is immediate.

Proposition 3.4 *For any $S \in \mathcal{X}$ with $0 \in S$, we have $H(S)^*_{H(\mathsf{D})} = H(S^o_{\mathsf{D}})$.*

Proof. Take any $[y_0, y^T]^T \in H(S)^*_{H(\mathsf{D})}$. We have

$$\left[\begin{array}{c} x_0 y_0 \\ \langle x, y \rangle \end{array} \right] \in H(\mathsf{D}) \text{ for all } [x_0, x^T]^T \in H(S).$$

Hence, $y_0 \geq 0$. Consider first $y_0 > 0$. In that case, $\langle x, \frac{y}{y_0} \rangle \in \mathsf{D}$ for all $x \in S$, and so $[y_0, y^T]^T \in H(S^o_{\mathsf{D}})$. If $y_0 = 0$ then $[0, y^T]^T \in H(S)^*_{H(\mathsf{D})}$ meaning that there is a sequence $\epsilon_n \downarrow 0$ and $y_n \to y$ such that $\langle x, y_n/\epsilon_n \rangle \in \mathsf{D}$ for all $x \in S$. This implies that $[\epsilon_n, y_n^T]^T \in H(S^o_{\mathsf{D}})$. Taking limit yields $[y_0, y^T]^T \in H(S^o_{\mathsf{D}})$.

Next we show the converse containing relationship. Take $[y_0, y^T]^T \in H(S^o_{\mathsf{D}})$ with $y_0 > 0$. We thus have $\langle x, \frac{y}{y_0} \rangle \in \mathsf{D}$ for all $x \in S$. Now take any $[x_0, x^T]^T \in H(S)$ with $x_0 > 0$. We have $\langle \frac{x}{x_0}, \frac{y}{y_0} \rangle \in \mathsf{D}$, or equivalently,

$$\left[\begin{array}{c} x_0 y_0 \\ \langle x, y \rangle \end{array} \right] \in H(\mathsf{D}).$$

Taking closure it follows that $H(S^o_{\mathsf{D}}) \subseteq H(S)^*_{H(\mathsf{D})}$. The proposition is proven. **Q.E.D.**

For the bipolar theorem (the D-induced polar case), we introduce the following corresponding condition on surjectivity.

Definition 3.5 *We call* $\langle x, y\rangle$ *dual surjective with respect to set* D *if for any* $a \in \mathcal{X}$ *there is* $b \in \mathsf{D}$ *with* $-\lambda b \notin \mathsf{D}$ *for all* $\lambda > 1$*, and an element* $y \in \mathcal{Y}$ *such that the linear equation*

$$[A_1 y, \cdots, A_k y] = ab^T$$

is satisfied. The coordinate-free description is

$$\forall a \in \mathcal{X},\ \exists b \in \mathsf{D}\,(-\lambda b \notin \mathsf{D}\,\forall \lambda > 1),\ \exists y \in Y,\ \text{such that } \langle x, y\rangle = \langle x, a\rangle_{\mathcal{X}} b \text{ for all } x \in \mathcal{X}.$$

The conic bipolar theorem (Theorem 2.4) can be generalized as follows.

Theorem 3.6 *Suppose that the bilinear mapping* $\langle x, y\rangle$ *is dual surjective according to Definition 3.5. Moreover, suppose that* D *is symmetric around the origin. Then for any convex set* $S \subseteq \mathcal{X}$ *which is symmetric around the origin, we have*

$$\left((H(S))^*_{H(\mathsf{D})}\right)^*_{H(\mathsf{D})} = H(S).$$

Proof. That $H(S) \subseteq \left((H(S))^*_{H(\mathsf{D})}\right)^*_{H(\mathsf{D})}$ is obvious. We shall concentrate on proving the converse containing relationship.

We shall first identify the set $(H(S))^*_{H(\mathsf{D})}$. Let $\bar{y} = [y_0, y^T]^T \in (H(S))^*_{H(\mathsf{D})}$. This requires that

$$\begin{bmatrix} x_0 y_0 \\ \langle x, y\rangle \end{bmatrix} \in H(\mathsf{D}) \text{ for all } [x_0, x^T]^T \in H(S).$$

Or, equivalently,

$$x_0 y_0 w_0 + x^T L(y) w \geq 0 \tag{7}$$

for all $[w_0, w^T]^T \in H(\mathsf{D})^* = H(\mathsf{D}^o)$ (Corollary 3.2), and $[x_0, x^T]^T \in H(S)$.

Suppose by contradiction that there is

$$\begin{bmatrix} \hat{x}_0 \\ \hat{x} \end{bmatrix} \in \left((H(S))^*_{H(\mathsf{D})}\right)^*_{H(\mathsf{D})} \setminus H(S).$$

By the separation theorem, there is $[a_0, a^T]^T \in H(S)^* = H(S^o)$ such that

$$a_0 \hat{x}_0 + a^T \hat{x} < 0. \tag{8}$$

Using the surjectivity condition, there is $\hat{y}$ such that $L(\hat{y}) = ab^T$ with $b \in \mathsf{D}$ and $-\lambda b \notin \mathsf{D}$ for all $\lambda > 1$. Now we claim that $[a_0, \hat{y}^T]^T$ satisfies (7). To see this we note that

$$a_0 x_0 + a^T x \geq 0, \, \forall \, [x_0, x^T]^T \in H(S), \text{ (since } [a_0, a^T]^T \in H(S)^*), \tag{9}$$

$$w_0 + w^T b \geq 0, \, \forall \, [w_0, w^T]^T \in H(\mathsf{D})^*, \text{ (since } [1, b^T]^T \in H(\mathsf{D})), \tag{10}$$

$$a_0 \geq 0, x_0 \geq 0, \, w_0 \geq 0 \text{ (due to the homogenization operation).} \tag{11}$$

In fact, since S and D are symmetric around the origin, we also know that S^o and D^o are also symmetric around the origin. Therefore, inequalities (9) and (10) further imply that

$$a_0 x_0 - |a^T x| \geq 0, \text{ for all } [x_0, x^T]^T \in H(S), \tag{12}$$

and

$$w_0 - |w^T b| \geq 0, \text{ for all } [w_0, w^T]^T \in H(\mathsf{D})^*. \tag{13}$$

Using (11), (12) and (13) we have

$$a_0 x_0 w_0 + x^T L(\hat{y}) w = a_0 x_0 w_0 + a^T x b^T w \geq 0$$

for all $[w_0, w^T]^T \in H(\mathsf{D})^*$ and $[x_0, x^T]^T \in H(S)$. Hence $[a_0, \hat{y}^T]^T \in (H(S))^*_{H(\mathsf{D})}$.

Moreover,

$$\left\langle \begin{bmatrix} \hat{x_0} \\ \hat{x} \end{bmatrix}, \begin{bmatrix} a_0 \\ \hat{y} \end{bmatrix} \right\rangle = \begin{bmatrix} a_0 \hat{x}_0 \\ \langle \hat{x}, \hat{y} \rangle \end{bmatrix} = \begin{bmatrix} a_0 \hat{x}_0 \\ (a^T \hat{x}) b \end{bmatrix}.$$

However, using (8) it follows that $[a_0, \hat{y}^T]^T \notin (H(S))^*_{H(\mathsf{D})}$, yielding a contradiction, which in turn shows that

$$\left((H(S))^*_{H(\mathsf{D})} \right)^*_{H(\mathsf{D})} = H(S).$$

Q.E.D.

Using Proposition 3.4 we further obtain:

Corollary 3.7 *Suppose that the bilinear mapping $\langle x, y \rangle$ is dual surjective according to Definition 3.5. Moreover, suppose that D is symmetric around the origin. Then for any convex set $S \subseteq \mathcal{X}$ which is symmetric around the origin, we have $(S^o_{\mathsf{D}})^o_{\mathsf{D}} = \mathrm{cl}\, S$.*

4 Computing the D-induced dual objects

Undoubtedly, theoretical properties such as the bipolar theorems are important for convex analysis. For the practical purposes, however, it is also important to be able to compute the dual objects in a tangible way for all practical purposes. It is well known that the computation of an ordinary dual object is essentially of the same difficulty as the computation of the primal object itself. In particular, it is shown by Grötschel, Lovász, and Schrijver [3] that if there is a polynomial-time procedure to check the membership for the primal convex cone, then there is also a polynomial-time procedure to check the membership for the dual cone.

The situation becomes quite blurred in the D-induced duality. It may happen that both S and D are simple convex sets, but the membership check for S^o_{D} becomes a hard task. Specifically we consider the following result due to Nemirovski and Ben-Tal; see [1]:

Proposition 4.1 *For given symmetric matrices $A_i \in \mathcal{S}^m$, the following decision problem is co-NP complete*

$$\sum_{i=1}^{n} x_i A_i \preceq I, \text{ for all } x \in \Re^n \text{ with } \|x\| \leq 1.$$

In the terminology of D-induced duality, this implies the following. Let us consider $\mathcal{X} = \Re^n$, $\mathcal{Y} = (\mathcal{S}^m)^n$ (the Cartesian product of n copies of the space of $m \times m$ symmetric matrices), and $\mathcal{W} = \mathcal{S}^m$. For $x \in \mathcal{X}$ and $y = (Y_1, ..., Y_n) \in \mathcal{Y}$, let the bilinear product be defined as

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i Y_i.$$

Let $\mathsf{D} = \{W \mid W \preceq I\}$. Then Proposition 4.1 asserts that it is NP-hard to check the membership for the polar set S^o_{D} where S is a unit Euclidean ball in $\mathcal{X}$. Of course, one may homogenize and then reformulate the problem as checking the membership of a $\bar{\mathsf{D}}$-induced dual of a second order cone where $\bar{\mathsf{D}}$ is the $m \times m$ positive semidefinite matrix cone.

However, there are interesting special cases for which the computation can be done satisfactorily. Now let us be more specific about our terminology. As a convention, we shall call the description of a set *satisfactory* if the description is done by a polynomial number of polynomially-sized *Linear Matrix Inequalities* (LMI), where the polynomials are in terms of the input size of the problem. In short, we call such a description *LMI representable*. As is well-known, the usual *linear inequality representation* is a special case of the *linear matrix inequality representation* where all the matrices are diagonal.

Let us first consider the polyhedral case. Note the famous Farkas lemma (the inhomogeneous version): the implication "$Fx \le f$ (assuming it is nonempty) $\Longrightarrow g^T x \le h$" holds if and only if $\exists z \ge 0$ and $z_0 \ge 0$ such that

$$[h, g^T] = [z, z_0] \begin{bmatrix} f & F \\ 1 & 0 \end{bmatrix}.$$

Hence, "$Fx \le f$ (assuming it is nonempty) $\Longrightarrow Gx \le h$" holds if and only if $\exists Z \ge 0$ such that

$$[h, G] = Z \begin{bmatrix} f & F \\ 1 & 0 \end{bmatrix}.$$

Let $\mathcal{X} = \Re^n$, $\mathcal{Y} = \Re^m$, $\mathcal{W} = \Re^k$. The bilinear mapping is given as

$$\langle x, y \rangle = \begin{bmatrix} x^T A_1 y \\ \vdots \\ x^T A_k y \end{bmatrix}. \tag{14}$$

Let $\mathsf{D} = \{w \mid Dw \le d\}$ and $S = \{x \mid Fx \le f\}$. Applying the Farkas lemma one obtains

$$(S)^o_{\mathsf{D}} = \left\{ y \;\middle|\; \exists Z \ge 0 : \; [d, DL(y)] = Z \begin{bmatrix} f & F \\ 1 & 0 \end{bmatrix} \right\},$$

where $L(y) = [A_1 y, A_2 y, ..., A_k y]$.

Consider $S = \{x \mid x^T Q x + 2q^T x + q_0 \le 0\} \subseteq \mathcal{X}$ and $\mathsf{D} = \{w \mid Dw \le d\}$. This is a degenerate case of our next discussion. Let us write $D = [g_1, g_2, ..., g_r]^T$ and $d = [h_1, h_2, ..., h_r]^T$. It follows from (16) that

$$(S)^o_{\mathsf{D}} = \left\{ y \;\middle|\; \exists s_i \ge 0 : \begin{bmatrix} s_i q_0 + h_i & s_i q^T - (L(y) g_i)^T/2 \\ s_i q - L(y) g_i/2 & s_i Q \end{bmatrix} \succeq 0, \; i = 1, ..., r \right\}.$$

If S is a polyhedron and D is an ellipsoid, then checking the membership of $(S)^o_{\mathsf{D}}$ becomes checking whether a polyhedron is contained in an ellipsoid, which is an NP-hard problem.

We now consider the case where both S and D are level sets of quadratic functions.

Let $\mathsf{D} = \{w \mid w^T B w + 2b^T w + b_0 \le 0\} \subseteq \mathcal{W}$ with $B = [b_{ij}]_{k \times k} \in \mathcal{S}^k_{++}$ and $b = [b_i]_{k \times 1} \in \Re^k$. Consider $S = \{x \mid x^T Q x + 2q^T x + q_0 \le 0\} \subseteq \mathcal{X}$. Note that Q may be indefinite. For simplicity we assume that $\operatorname{int} S \neq \emptyset$.

It is obvious that in this case $y \in S^o_{\mathsf{D}}$ if and only if the following implication holds

$$\sum_{i=1}^k \sum_{j=1}^k (x^T A_i y)(x^T A_j y) b_{ij} + 2 \sum_{i=1}^k x^T A_i y b_i + b_0 \le 0$$

for all x satisfying

$$x^T Qx + 2q^T x + q_0 \le 0.$$

By the so-called S-Lemma (for the format that suits this context, cf. Theorem 1 in Sturm and Zhang [11]), the above implication is true if and only if there is $s \ge 0$ such that

$$s \begin{bmatrix} q_0 & q^T \\ q & Q \end{bmatrix} - \begin{bmatrix} b_0 & \left(\sum_{i=1}^k b_i A_i y\right)^T \\ \left(\sum_{i=1}^k b_i A_i y\right) & \sum_{i=1}^k \sum_{j=1}^k (A_i y y^T A_j) b_{ij} \end{bmatrix} \succeq 0. \quad (15)$$

Noting $[A_1 y, A_2 y, ..., A_k y] = L(y)$ we may write (15) as

$$s \begin{bmatrix} q_0 & q^T \\ q & Q \end{bmatrix} - \begin{bmatrix} b_0 & (L(y)b)^T \\ L(y)b & L(y)BL(y)^T \end{bmatrix} \succeq 0,$$

or,

$$\begin{bmatrix} sq_0 - b_0 & sq^T - (L(y)b)^T \\ sq - L(y)b & sQ \end{bmatrix} - \begin{bmatrix} 0 \\ L(y) \end{bmatrix} B \left[0, L(y)^T\right] \succeq 0. \quad (16)$$

Using the Schur complement lemma, (16) can be equivalently written as

$$\begin{bmatrix} \begin{bmatrix} sq_0 - b_0 & sq^T - (L(y)b)^T \\ sq - L(y)b & sQ \end{bmatrix}, & \begin{bmatrix} 0 \\ L(y) \end{bmatrix} \\ \left[0, L(y)^T\right], & B^{-1} \end{bmatrix} \succeq 0. \quad (17)$$

Hence we have the following LMI representation result.

Theorem 4.2 *Suppose that $S = \{x \mid x^T Qx + 2q^T x + q_0 \le 0\} \subseteq \mathcal{X}$ and $\mathsf{D} = \{w \mid w^T Bw + 2b^T w + b_0 \le 0\} \subseteq \mathcal{W}$ with $B \succ 0$, and the bilinear mapping given as (14). Then it holds that*

$$(S)^o_{\mathsf{D}} = \left\{ y \,\middle|\, \exists s \ge 0 : \begin{bmatrix} \begin{bmatrix} sq_0 - b_0 & sq^T - (L(y)b)^T \\ sq - L(y)b & sQ \end{bmatrix}, & \begin{bmatrix} 0 \\ L(y) \end{bmatrix} \\ \left[0, L(y)^T\right], & B^{-1} \end{bmatrix} \succeq 0 \right\}.$$

By Proposition 3.4 we have $(H(S))^o_{H(\mathsf{D})} = H(S^o_{\mathsf{D}})$, provided that the bilinear mapping is given as (6). Hence, by Theorem 4.2, for $S = \{x \mid x^T Qx + 2q^T x + q_0 \le 0\} \subseteq \mathcal{X}$ and $\mathsf{D} = \{w \mid w^T Bw + 2b^T w + b_0 \le 0\} \subseteq \mathcal{W}$,

$$(H(S))^o_{H(\mathsf{D})} = \left\{ \begin{bmatrix} y_0 \\ y \end{bmatrix} \,\middle|\, y_0 \ge 0,\ s \ge 0 : \begin{bmatrix} \begin{bmatrix} sq_0 - b_0 & sq^T - (L(y)b)^T \\ sq - L(y)b & sQ \end{bmatrix}, & \begin{bmatrix} 0 \\ L(y) \end{bmatrix} \\ \left[0, L(y)^T\right], & y_0 B^{-1} \end{bmatrix} \succeq 0 \right\}.$$

Suppose that $B \succeq 0$ is diagonal. Let $\tilde{L}(y) := [\tilde{L}_1(y), ..., \tilde{L}_k(y)]$ where $\tilde{L}_i(y) := A_i y$ if $B_{ii} > 0$ and $\tilde{L}_i(y)$ is the zero vector if $B_{ii} = 0$, $i = 1, ..., k$. Then it is straightforward to obtain from Theorem 4.2 the following result, using the pseudo-inverse.

Corollary 4.3 *Suppose that* $S = \{x \mid x^T Qx + 2q^T x + q_0 \le 0\} \subseteq \mathcal{X}$ *and* $\mathsf{D} = \{w \mid w^T Bw + 2b^T w + b_0 \le 0\} \subseteq \mathcal{W}$ *with* $B = \mathrm{diag}\,(B_{11}, ..., B_{kk}) \succeq 0$*, and the bilinear mapping given as (14). Then it holds that*

$$(S)^o_{\mathsf{D}} = \left\{ y \;\middle|\; \exists s \ge 0 : \begin{bmatrix} \begin{bmatrix} sq_0 - b_0 & sq^T - (L(y)b)^T \\ sq - L(y)b & sQ \end{bmatrix}, & \begin{bmatrix} 0 \\ \tilde{L}(y) \end{bmatrix} \\ \left[0, \tilde{L}(y)^T\right], & B^+ \end{bmatrix} \succeq 0 \right\}$$

where B^+ *stands for the pseudo-inverse of* B*.*

It is unknown whether one can find an LMI representation for the set S^o_{D}, where S is a second order cone in $\mathcal{X}$, D is a second order cone in $\mathcal{W}$, and the bilinear mapping is general. According to Proposition 2.5, this is equivalent to the quest whether the following bi-positive set has an LMI representation or not:

$$\mathcal{B}(C, D) = \{Z \mid u^T Zv \ge 0 \text{ for all } u \in C,\, v \in D\}$$

where C and D are two second order cones.

In general, let $0 \in S \subseteq \Re^n$ be an arbitrary convex set. For $\tau > 0$ we denote

$$\tau S = \{x \mid x/\tau \in S\}.$$

Then, obviously, $(\tau S)^o_{\mathsf{D}} = \frac{1}{\tau}(S)^o_{\mathsf{D}}$. Similarly, if $0 \in \mathsf{D}$ then $(S)^o_{\tau \mathsf{D}} = \tau (S)^o_{\mathsf{D}}$.

Proposition 4.4 *Suppose that* $0 \in \mathrm{int}\,(E_1) \subseteq S \subseteq \tau E_1 \subseteq \Re^n$ *and* $0 \in \mathrm{int}\,(E_2) \subseteq \mathsf{D} \subseteq \kappa E_2 \subseteq \Re^k$*, where* E_1 *and* E_2 *are two convex sets. It holds that*

$$\frac{1}{\tau}(E_1)^o_{E_2} \subseteq (S)^o_{\mathsf{D}} \subseteq \kappa (E_1)^o_{E_2}. \tag{18}$$

One implication of the above proposition is the following. Although it may be difficult to *exactly* compute the D-induced dual sets, it is possible to approximate it with ellipsoidal duals, which have shown to be LMI representable (Theorem 4.2). According to the Löwner-John theorem (see e.g. Theorem 3.1.9 of [3]), any full-dimensional and bounded convex set can be approximated by a pair of co-centered ellipsoids with a factor no more than the dimension of the space. This shows that any D-induced polar of a bounded set can be in principle approximated by LMI representable sets with a finite approximation bound. For some special cases, the approximation bounds can be provably better. We shall consider one example to illustrate our point. First note the following results due to Nemirovski, Roos and Terlaky [6], Nesterov [7, 8], and Ye [12].

Lemma 4.5 *Consider the quadratic optimization problem*

$$\begin{array}{lll} (QP) & \textit{maximize} & x^T P_0 x \\ & \textit{subject to} & x^T P_i x \leq 1,\ i = 1, ..., m \end{array}$$

where $P_i \succeq 0,\ i = 0, 1, ..., m$, *and its semidefinite programming (SDP) relaxation*

$$\begin{array}{lll} (SDP) & \textit{maximize} & P_0 \bullet X \\ & \textit{subject to} & P_i \bullet X \leq 1,\ i = 1, ..., m \\ & & X \succeq 0. \end{array}$$

Suppose that (QP) *has an optimal solution. Let* $v(QP)$ *be the optimal value of* (QP) *and* $v(SDP)$ *be the optimal value of* (SDP)*. Then,*

(1) *([6]) It holds that*

$$v(QP) \leq v(SDP) \leq \lambda v(QP),$$

where $\lambda = 2\log(2m\mu)$ *with* $\mu = \min\{m; \max_{1 \leq i \leq m} \operatorname{rank} P_i\}$.

(2) *([7, 8, 12]) If the matrices* P_i*'s are mutually commute,* $i = 1, ..., m$*, then*

$$v(QP) \leq v(SDP) \leq \frac{\pi}{2} v(QP).$$

Consider $S = \{x \mid x^T P_i x \leq 1,\ i = 1, ..., m\} \subseteq \mathcal{X}$ with $P_i \succeq 0$ for all i and $\sum_{i=1}^m P_i \succ 0$, and $\mathsf{D} = \{w \mid w^T B w \leq 1\} \subseteq \mathcal{W}$ with $B \succ 0$. The bilinear mapping is given as (14). It is clear that

$$(S)^o_{\mathsf{D}} = \left\{ y \,\middle|\, x^T \left(L(y) B L(y)^T \right) x \leq 1 \text{ for all } x^T P_i x \leq 1,\ i = 1, ..., m \right\}.$$

Checking the membership for the above set is NP-hard in general. Consider now an approximation using the SDP relaxation

$$A(S^o_{\mathsf{D}}) := \left\{ y \,\middle|\, \left(L(y) B L(y)^T \right) \bullet X \leq 1 \text{ for all } P_i \bullet X \leq 1,\ i = 1, ..., m,\ X \succeq 0 \right\}.$$

Clearly, $A(S^o_{\mathsf{D}}) \subseteq (S)^o_{\mathsf{D}}$. By the strong duality theorem for SDP under the primal-dual Slater condition, we can rewrite $A(S^o_{\mathsf{D}})$ as

$$\begin{aligned} A(S^o_{\mathsf{D}}) &= \left\{ y \,\middle|\, \sum_{i=1}^m t_i P_i \succeq L(y) B L(y)^T,\ t_i \geq 0,\ i = 1, ..., m,\ \sum_{i=1}^m t_i \leq 1 \right\} \\ &= \left\{ y \,\middle|\, \begin{bmatrix} \sum_{i=1}^m t_i P_i & L(y) \\ L(y)^T & B^{-1} \end{bmatrix} \succeq 0,\ t_i \geq 0,\ i = 1, ..., m,\ \sum_{i=1}^m t_i \leq 1 \right\} \end{aligned}$$

where the second step is due to the Schur complement lemma, and this represents $A(S^o_{\mathsf{D}})$ by LMI. Lemma 4.5 asserts that if $y \in (S)^o_{\mathsf{D}}$ then $\left(L(y) B L(y)^T \right) \bullet X \leq \lambda$ for all $P_i \bullet X \leq 1,\ i = 1, ..., m,\ X \succeq 0,$

where $\lambda = 2\log(2m\mu)$ with $\mu = \min\{m; \max_{1\le i\le m} \text{rank } P_i\}$ for general $P_i \succeq 0$, and $\lambda = \pi/2$ if P_i's are mutually commute. Since $L(y)$ is linear in y, it follows that $y/\sqrt{\lambda} \in A(S^o_{\mathsf{D}})$. Therefore, in this case we have

$$A(S^o_{\mathsf{D}}) \subseteq (S)^o_{\mathsf{D}} \subseteq \sqrt{\lambda} A(S^o_{\mathsf{D}}).$$

5 Applications

To appreciate how the D-induced duality helps to model and solve optimization problems, we shall first discuss two examples of application in this section, before moving on to discuss more theoretical properties of the D-induced dual cones.

5.1 Robust optimization

The notion of robust optimization was studied by Ben Tal and Nemirovski in [1]. Let us now formulate the problem using the newly introduced notion of D-induced duality. Consider a general convex optimization problem as follows

$$\begin{array}{lll} (P) & \text{minimize} & c^T x \\ & \text{subject to} & Ax + b \in \mathsf{D} \end{array}$$

or equivalently,

$$\begin{array}{lll} (P) & \text{minimize} & c^T x \\ & \text{subject to} & Ax + bx_0 \in \mathsf{D} \\ & & x_0 = 1, \end{array}$$

where $\mathsf{D} \subseteq \Re^k$ is a given closed convex set.

In case D is a closed convex cone, (P) is referred to as *convex conic optimization.* The cone D can be the first orthant (linear programming), or the product of second order cones (second order cone programming), or the cone of positive semidefinite matrices (semidefinite programming). In [13], it is shown that a standard convex program can also be formulated in the conic form in a natural way.

The issue of robust optimization arises when the data of the problem, A and b in the above formulation, are uncertain. In other words, they are perturbed by noises, or 'polluted'. Note that the objective c vector can always be assumed to be certain. This is achieved by reformulating the problem as follows. We introduce one additional variable, x_{n+1}, and one additional constraint, $c^T x - x_{n+1} = 0$, the cone is changed to $\mathcal{K} \times \Re$, and the objective is changed to minimize x_{n+1}. For convenience we assume that in the original problem, the objective vector c is already certain.

In many applications, it is crucial to ensure the robustness of the decision x. One way to do this is to guarantee that x should remain feasible for all possible data within an 'uncertainty set'.

Let A_0 and b_0 be the nominal value of the data. The general structure of the 'polluted' data set is assumed to be

$$A = A_0 + \sum_{i=1}^{m} \left(\sum_{j=1}^{k_i} u_j^i A_j^i \right) \text{ and } b = b_0 + \sum_{i=1}^{m} \left(\sum_{j=1}^{k_i} u_j^i b_j^i \right), \tag{19}$$

where $u^i = [u_1^i, ..., u_{k_i}^i]^T \in U^i \subseteq \Re^{k_i}$, and U^i is known as *the uncertainty region*, $i = 1, ..., m$. The matrices $A_j^i \in \Re^{k \times n}$ are used to model the structure of the uncertainty.

A robustly feasible solution for (P), x, will need to satisfy $Ax + b \in \mathsf{D}$ for all data A and b as specified in (19) where $u^i \in U^i$, $i = 1, ..., m$.

Let $\mathcal{X} = \Re^{n+1}$, $\mathcal{Y} = \Re^{k \times (n+1)}$, and $\mathcal{W} = \Re^k$. For $[x_0, x^T]^T \in \mathcal{X}$ and $y = \text{vec}\,([y_0, Y]) \in \mathcal{Y}$ with $y_0 \in \Re^k$ and $Y \in \Re^{k \times n}$, introduce the bilinear mapping as follows

$$\left\langle \begin{bmatrix} x_0 \\ x \end{bmatrix}, \text{vec}\,([y_0, Y]) \right\rangle := [y_0, Y] \begin{bmatrix} x_0 \\ x \end{bmatrix}.$$

Let $k_0 = 1$ and $U^0 = \{1\}$, $A_1^0 = A_0$, and $b_1^0 = b_0$. Let

$$S^i := \left\{ \left[\sum_{j=1}^{k_i} u_j^i b_j^i, \sum_{j=1}^{k_i} u_j^i A_j^i \right] \middle| \, u^i \in U^i \right\}, \; i = 0, 1, ..., m. \tag{20}$$

So, S^0 comprises a singleton, viz. $[b_0, A_0]$. The uncertainty set $S^i \subseteq \mathcal{Y}$ is a linear image of U^i, $i = 1, ..., m$. In particular, if U^i is a polyhedron/ellipsoid/LMI representable set, then S^i is also a polyhedron/ellipsoid/LMI representable set respectively. By this construction we may also assume that $0 \in S^i$, as the set consists of perturbation vectors, $i = 1, ..., m$.

A solution x is feasible for (P) in the robust sense if and only if $\left\langle \begin{bmatrix} x_0 \\ x \end{bmatrix}, \text{vec}\,([y_0, Y]) \right\rangle \in \mathsf{D}$ for all $[y_0, Y] \in \sum_{i=0}^m S^i$. Because S^0 is a singleton and all other S^i's contain the origin, using Proposition 3.3 we obtain

$$\begin{bmatrix} 1 \\ x \end{bmatrix} \in \left(\sum_{i=0}^{m} S^i \right)^o_{\mathsf{D}} = (S^0)^o_{\mathsf{D}} \bigcap \left(\bigcap_{i=1}^{m} \left(S^0 + S^i \right)^o_{\mathsf{D}} \right).$$

The robust optimization version of (P) is

$$\begin{array}{lll} (RP) & \text{minimize} & c^T x \\ & \text{subject to} & \begin{bmatrix} 1 \\ x \end{bmatrix} \in \left(\sum_{i=0}^{m} S^i \right)^o_{\mathsf{D}} = (S^0)^o_{\mathsf{D}} \bigcap \left(\bigcap_{i=1}^{m} \left(S^0 + S^i \right)^o_{\mathsf{D}} \right). \end{array}$$

Note that $\{x \mid [1, x^T]^T \in (S^0)^o_{\mathsf{D}}\}$ is the original non-robust feasible set. The structure of the constraint in (RP) due to the robustness consideration is apparent in this formulation.

Some well-known results in robust optimization (see [1]) follow immediately from this observation.

For robust linear programming, i.e., $\mathsf{D} = \Re^n_+$, if the uncertainty set S^i is a polyhedron/ellipsoid/LMI representable set, then $(S^0 + S^i)^o_{\mathsf{D}}$ is: polyhedron/ellipsoid/LMI representable set respectively, and so the robust counter part (RP) is: a linear program/SOC program (second order cone program)/SDP program (semidefinite program), respectively.

For robust Second Order Cone Programming, i.e., $\mathsf{D} = \mathrm{SOC}(n)$, if the uncertainty set S^i is a polyhedron, then $(S^0 + S^i)^o_{\mathsf{D}}$ is an intersection of second order cones; hence the robust counter part (RP) remains an SOC problem. If the uncertainty set S^i is an ellipsoid, then it remains an open question whether $(S^0 + S^i)^o_{\mathsf{D}}$ is LMI representable or not. Our conjecture is yes. If so, then the corresponding (RP) is an SDP problem. But even if the answer is no, one still has a polynomial time procedure for checking the membership of the convex set $(S^0 + S^i)^o_{\mathsf{D}}$. By the ellipsoid method, the robust counter part (RP) can be solved efficiently in principle. If S^i is a general LMI representable set, then $(S^0 + S^i)^o_{\mathsf{D}}$ is intractable, since this is already so when S^i is given in the form of a polyhedron (linear inequalities) rather then in the form of a polytope (convex combinations of vertices).

For robust Semidefinite Programming, i.e., $\mathsf{D} = \mathcal{S}^m_+$, the robust set $(S^0 + S^i)^o_{\mathsf{D}}$ is intractable when S^i is an ellipsoid or an LMI representable set (see e.g. Proposition 4.1); it is tractable if the uncertainty set S^i is a polytope, when (RP) remains a semidefinite program.

An interesting case is the robust least square problem studied by El Ghaoui and Lebret [2]:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & t \ge (Ax-b)^T(Ax-b),\ [A,b] \in \text{`uncertainty set'}. \end{array}$$

To put the problem into the perspective of the D-induced duality, let us introduce

$$\bar{x} := \begin{bmatrix} x_0 \\ t \\ x \end{bmatrix} \text{ and } \bar{y} := \begin{bmatrix} s \\ y_0 \\ \mathrm{vec}\,(Y) \end{bmatrix}$$

and the bilinear mapping

$$\langle \bar{x}, \bar{y} \rangle := \begin{bmatrix} ts \\ Yx - x_0 y_0 \end{bmatrix} =: \begin{bmatrix} w_0 \\ w \end{bmatrix} \in \mathcal{W}.$$

Let $S := \{\bar{y} \mid s = 1,\, y_0 = b,\, Y = A, \text{ and } [A,b] \in \text{`uncertainty set'}\}$. Let $\mathsf{D} = \{[w_0, w^T]^T \mid w_0 \ge w^T w\}$. According to Corollary 4.3, the set $(S)^o_{\mathsf{D}}$ is LMI representable if the uncertainty set in which

$[A, b]$ resides can be expressed as a sum of ellipsoids. The robust version of the least square problem can thus be formulated as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} 1 \\ t \\ x \end{bmatrix} \in (S)^o_{\mathsf{D}}, \end{array}$$

which turns out to be an SDP problem in that case.

5.2 Multiple objective conic programming

Let us consider the following ordering relation based on a convex cone $\mathsf{D} \subseteq \Re^k$:

$$x \succeq_{\mathsf{D}} y \text{ if and only if } x - y \in \mathsf{D}.$$

The multiple objective convex conic program is now given as

$$\begin{array}{lll} (P)_{\mathsf{D}} & \min_{\mathsf{D}} & Cx \\ & \text{s.t.} & Ax = b \\ & & x \in \mathcal{K} \end{array}$$

where $C \in \Re^{k\times n}$, $A \in \Re^{m\times n}$, $b \in \Re^m$ and $\mathcal{K} \subseteq \Re^n$ is a convex cone. Its dual problem can be derived by means of the Lagrangian multipliers and lower bounding. Namely, we obtain the dual problem by attempting to establish a lower bound for the objective in $(P)_{\mathsf{D}}$. Let Y be a linear mapping $\Re^m \to \Re^k$ to be applied on both sides of $Ax = b$, leading to $YAx = Yb$. In order for YAx to be a lower bound for the objective vector, we need to have $(C - YA)x \in \mathsf{D}$ for all primal feasible x.

Let us consider the bilinear mapping defined as

$$\langle x, S\rangle = Sx \in \Re^k.$$

Clearly, the above bilinear mapping is dual surjective, because $L(S) = S^T$ and so the equation $L(S) = ab^T$ is always solvable.

In the notion of the D-induced duality, the condition $(C - YA)x \in \mathsf{D}$ for all $x \in \mathcal{K}$ is simply $C - YA \in \mathcal{K}^*_{\mathsf{D}}$. Now we wish to optimize over all the bounds obtained this way. This naturally leads to the dual problem given as follows:

$$\begin{array}{lll} (D)_{\mathsf{D}} & \max_{\mathsf{D}} & Yb \\ & \text{s.t.} & YA + S = C \\ & & S \in \mathcal{K}^*_{\mathsf{D}}. \end{array}$$

Denote F_P to be the feasible set for $(P)_{\mathsf{D}}$ and F_D to be the feasible set for $(D)_{\mathsf{D}}$.

By this construction, and also using the extended bipolar theorem, Theorem 2.4, we have the following result, an analog of the weak duality theorem .

Theorem 5.1 *For any $x \in F_P$ and $(Y, S) \in F_D$, we have*

$$Cx \succeq_{\mathsf{D}} Yb.$$

Moreover, if D and $\mathcal{K}$ are closed convex cones then the dual of $(D)_{\mathsf{D}}$ is precisely $(P)_{\mathsf{D}}$ again.

We should note that the D-induced duality is based on the choice of the bilinear mapping. Therefore, if $(D)_{\mathsf{D}}$ would be considered as a primal problem, then its dual is not $(P)_{\mathsf{D}}$ but another problem in a larger space.

It is obvious that if $(P)_{\mathsf{D}}$ is a multi-objective linear program, i.e., both $\mathcal{K} = \Re^n_+$ and $\mathsf{D} = \Re^k_+$, then $(D)_{\mathsf{D}}$ is also a multi-objective linear program. In particular, if the primal problem is

$$\begin{array}{lll} (P)_{\mathsf{D}} & \min_{\Re^k_+} & Cx \\ & \text{s.t.} & Ax = b \\ & & x \geq 0, \end{array}$$

then its dual problem is

$$\begin{array}{lll} (D)_{\mathsf{D}} & \max_{\Re^k_+} & Yb \\ & \text{s.t.} & YA \leq C. \end{array}$$

It is well known that, if the ordering '$\succeq_{\mathsf{D}}$' is incomplete, there might be multiple, incomparable, optimal solutions from either the primal or dual point of view, known as the *Pareto* optimal solutions.

In general, if D is a closed cone, then

$$\mathcal{K}^*_{\mathsf{D}} = \{S \mid Sx \in \mathsf{D}, \forall x \in \mathcal{K}\} = \mathcal{B}(\mathsf{D}^*, \mathcal{K}).$$

If we insist on the strong duality (complementarity), then the optimality can be defined as follows.

Definition 5.2 *We call x^* $((Y^*, S^*))$ a global primal (dual) optimal solution if $Cx - Cx^* \in \mathsf{D}$ for all $x \in F_P$ ($Y^*b - Yb \in \mathsf{D}$ for all $(Y, S) \in F_D$).*

We call x^* and (Y^*, S^*) complementary optimal solutions for $(P)_{\mathsf{D}}$ and $(D)_{\mathsf{D}}$ respectively if they are feasible and $Cx^* = Y^*b$ or equivalently, $S^*x^* = 0 \in \mathsf{D}$.

Obviously, the complementary optimality is a stringent condition: a complementary optimal solution pair must be globally primal and dual optimal solutions respectively.

Consider the following simple multiple objective linear program

$$
\begin{aligned}
(P_1)_{\mathsf{D}} \quad \min\nolimits_{\mathsf{D}} \quad & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
\text{s.t.} \quad & x_1 + x_2 = 1 \\
& \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \Re^2_+ .
\end{aligned}
$$

Let us first consider the case where $\mathsf{D} = \Re^2_+$. Due to the binding relation, no feasible solution of $(P_1)_{\mathsf{D}}$ can be dominated by any other feasible solutions. Therefore, they are all Pareto optimal. However, from the complementary optimality point of view, no dominating optimal solution exists. To see this, consider its dual problem

$$
\begin{aligned}
(D_1)_{\mathsf{D}} \quad \max\nolimits_{\mathsf{D}} \quad & \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
\text{s.t.} \quad & \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} [1, 1] + S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
& S \in (\Re^2_+)^*_{\mathsf{D}},
\end{aligned}
$$

with $\mathsf{D} = \Re^2_+$, or

$$
\begin{aligned}
(D_1)_{\mathsf{D}} \quad \max\nolimits_{\mathsf{D}} \quad & \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
\text{s.t.} \quad & y_1 \le 0,\ y_2 \le 0.
\end{aligned}
$$

Obviously the equation $Sx = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ cannot have a solution among feasible x and S, i.e., no complementary optimal solution exists. However, $(D_1)_{\mathsf{D}}$ does have a globally optimal solution, namely $[y_1^*, y_2^*] = [0, 0]$. It turns out that this is no coincidence, as the following theorem reveals.

Theorem 5.3 *Suppose that a multi-objective linear program $(P_{\Re^k_+})$ is feasible and has a Pareto optimal solution. Then its dual problem, $(D_{\Re^k_+})$, always has a globally optimal solution.*

Proof. Observe that the dual problem can be separated. Namely, by letting $Y = [y_1, \cdots, y_k]^T$ and $C = [c_1, \cdots, c_k]^T$, and solve k linear programs, max $y_i^T b$ subject to $y_i^T A \le c_i^T$, with optimal solution $(y^*)_i^T$, $i = 1, ..., k$, then the optimal solution for $(D_{\Re^k_+})$ is simply $Y^* = [y_1^*, \cdots, y_k^*]^T$. It remains to argue that all these linear programs must have optimal solutions. First, they must all be feasible, because if $y_i^T A \le c_i^T$ is infeasible for some i then by the Farkas lemma there is $x \ge 0$ such that $Ax = 0$ and $c_i^T x < 0$, contradicting with the condition that $(P_{\Re^k_+})$ has a Pareto optimal solution. Second, they must be bounded, for otherwise there should be no $x \ge 0$ such that $Ax = b$, contradicting with the condition that $(P_{\Re^k_+})$ is feasible. This concludes the theorem. **Q.E.D.**

One can easily extend the result to the situation where $\mathsf{D} = \Re^k_+$ and $\mathcal{K}$ is an arbitrary convex cone. In that case, observe that

$$(\mathcal{K})^*_{\mathsf{D}} = \{S = [s_1, ..., s_k]^T \mid s_i^T x \ge 0 \, \forall x \in \mathcal{K}, \, i = 1, ..., k\} = \{S = [s_1, ..., s_k]^T \mid s_i \in \mathcal{K}^*, \, i = 1, ..., k\},$$

and the following extension of Theorem 5.3 can be proved using almost identical arguments, with the extra precaution that one requires the Slater condition for the strong duality; for discussions on the conic duality theory, see [4].

Theorem 5.4 *Suppose that a multi-objective conic convex program* $(P_{\Re^k_+})$ *is strongly feasible, namely there is* $x \in \text{int } \mathcal{K}$ *with* $Ax = b$. *Moreover, suppose that its dual problem,* $(D_{\Re^k_+})$, *is also strongly feasible, namely there is* Y *such that* $C - YA \in \text{int } (\mathcal{K})^*_{\mathsf{D}}$. *Then,* $(D_{\Re^k_+})$ *has a globally optimal solution.*

Theorem 5.4 has an interesting geometric interpretation. Although (P_{D}) may not have any dominating global optimal solution, there is an infeasible point (the dual global optimal solution) that uniformly dominates all the optimal solutions, and yet it is 'closest' to the feasibility. That solution is of special interest and can be viewed as a desirable reference point. Among all Pareto points, one may, for instance, choose the one that is closest to that reference point (projection).

For the problem instance $(P_1)_{\mathsf{D}}$, let us now consider the conic ordering induced by the lexicographic ordering of the coordinates. In this case, since the lexicographic ordering is complete and the feasible set of $(P_1)_{\mathsf{D}}$ is compact, an optimal solution for $(P_1)_{\mathsf{D}}$ exists. In particular, the optimal solution for $(P_1)_{\mathsf{D}}$ is obviously $[x_1^*, x_2^*] = [0, 1]$. Furthermore, we shall see below that in fact a pair of primal-dual complementary optimal solutions exists.

In the two-dimensional case, the lexicographic ordering on (x_1, x_2) corresponds to the conic ordering which is defined by the cone

$$\mathsf{D} = \left\{ \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] \middle| \, x_1 > 0 \right\} \bigcup \left(\Re_+ \left[\begin{array}{c} 0 \\ 1 \end{array} \right] \right) \tag{21}$$

which is a non-flat and non-closed convex cone.

Let us compute the cone $(\Re^2_+)^*_{\mathsf{D}}$. Take an arbitrary $S \in (\Re^2_+)^*_{\mathsf{D}}$, that is $Sx \in \mathsf{D}$ for all $x \in \Re^2_+$. In other words, $s_{11}x_1 + s_{12}x_2 \geq 0$ for all $x_1, x_2 \geq 0$, and if $s_{11}x_1 + s_{12}x_2 = 0$ then $s_{21}x_1 + s_{22}x_2 \geq 0$. This leads to

$$(\Re^2_+)^*_{\mathsf{D}} = \left\{ \left[\begin{array}{cc} s_{11} & s_{12} \\ s_{21} & s_{22} \end{array} \right] \middle| \begin{array}{l} s_{11}, s_{12} > 0, \text{ or } s_{11} = 0, s_{12} > 0, s_{21} \geq 0, \\ \text{or } s_{11} > 0, s_{12} = 0, s_{22} \geq 0, \text{ or } s_{11} = s_{12} = 0, s_{21} \geq 0, s_{22} \geq 0 \end{array} \right\}.$$

The dual problem $(D_1)_{\mathsf{D}}$, where D is given as in (21), has an optimal solution $y_1^* = 0$, $y_2^* = 1$, and $S^* = \left[\begin{array}{cc} 1 & 0 \\ -1 & 0 \end{array} \right]$. Moreover, this optimal solution is complementary to the primal optimal solution $[x_1^*, x_2^*] = [0, 1]$ as $S^* x^* = \left[\begin{array}{c} 0 \\ 0 \end{array} \right]$. A remarkable fact is that $(D_1)_{\mathsf{D}}$ does not even have a closed feasible set.

6 Further properties of the D-induced dual cone

In the previous section we saw some applications of the newly introduced D-induced duality. We also recognized that computing the D-induced dual cone or polar can be a difficult task in general; see Proposition 4.1. In Section 4 we focused on LMI representable D-induced dual subjects, and the approximation of it. The aim of this section is to view the computation of the D-induced dual cones from a totally different angle: sublinear function representations of convex cones and their tensor products.

As Proposition 2.5 asserts, we have for D closed

$$\begin{aligned} U^*_{\mathsf{D}} &= \{y \mid L(y) \in \mathcal{B}(U, \mathsf{D}^*)\} \\ &= \{y \mid L(y) \in (U \otimes \mathsf{D}^*)^*\}. \end{aligned}$$

It is therefore clear that it is crucial to analyze the dual (in the usual sense) of the tensor product (or the Kronecker product in the matrix form) of two convex cones, if we wish to characterize the D-induced dual cone.

Consider two closed and solid - that is, full-dimensional - convex cones, $C \in \mathcal{X}$ and $D \in \mathcal{W}$. First of all, C and D may not be pointed cones. In that case we can decompose them in the following way

$$C = P_c + L_c \text{ and } D = P_d + L_d$$

where P_c and P_d are closed pointed convex cones, and L_c and L_d are linear subspaces.

Then,

$$C \otimes D = P_c \otimes P_d + P_c \otimes L_d + L_c \otimes P_d + L_c \otimes L_d. \tag{22}$$

Lemma 6.1 *Let $K \in \mathcal{X}$ be a solid convex cone, namely,* span $K = \mathcal{X}$, *and L be a linear space. Then, $K \otimes L$ is a linear space, in fact*

$$K \otimes L = (\text{span } K) \otimes L.$$

Proof. Consider any $\sum_i x_i y_i^T \in K \otimes L$ with $x_i \in K$ and $y_i \in L$ for all i. Since L is a linear space, we have

$$-\sum_i x_i y_i^T = \sum_i x_i (-y_i)^T \in K \otimes L.$$

Therefore this concludes that $K \otimes L$ is a linear space, and furthermore $K \otimes L = (\text{span } K) \otimes L$. This proves the lemma. **Q.E.D.**

Lemma 6.1 shows that $C \otimes D$ can be written as the sum of $P_c \otimes P_d$ and a linear subspace. Therefore, in order to analyze the tensor product, without losing generality we may assume C and D to be pointed and solid convex cones.

Now, for a closed, pointed, convex cone C in $\mathcal{X}$, it is convenient to view C, by a suitable choice of coordinates in $\mathcal{X}$, as the epigraph of a sublinear function ϕ, i.e.,

$$C \sim \left\{ \begin{bmatrix} \phi(u) + r \\ u \end{bmatrix} \middle| \, r \in \Re_+, \, u \in \mathcal{U} \right\}$$

where '$\sim$' stands for a certain linear bijective transformation, and $\mathcal{U}$ is a vector space which has dimension equal to the dimension of $\mathcal{X}$ minus one. Remember that by definition a sublinear function $\phi(\cdot)$ is finite-valued and satisfies the following conditions:

$$\phi(tu) = t\phi(u) \text{ and } \phi(u + v) \le \phi(u) + \phi(v)$$

for all $t \ge 0$, and $u, v \in \mathcal{U}$.

For instance, if C is a second order cone, then the corresponding sublinear function can be simply the Euclidean norm $\phi(u) = \|u\|$; if C is an orthant, then $\phi(u) = \max\{\max_i u_i, 0\}$; if C is the cone of positive semidefinite matrices, then $\phi(u) = \max\{\lambda_{\max}(\mathrm{Mat}(u)), 0\}$, where 'Mat$(u)$' stands for the stacking operation to create a symmetric matrix from u, and '$\lambda_{\max}(X)$' stands for the maximum eigenvalue of X.

Similarly, since D is also a closed pointed cone, let us assume that

$$D \sim \left\{ \begin{bmatrix} \psi(w) + s \\ w \end{bmatrix} \,\middle|\, s \in \Re_+,\ w \in \mathcal{W} \right\}$$

where $\mathcal{W}$ is a vector space and ψ is a sublinear function on $\mathcal{W}$.

One may be tempted to conjecture that

$$(C \otimes D)^* = C^* \otimes D^*.$$

Unfortunately, this is not the case, although it is obvious that

$$C^* \otimes D^* \subseteq (C \otimes D)^*.$$

As an example, let us consider $C = D = \mathrm{SOC}(n+1)$. Then, by the Cauchy-Schwartz inequality one easily checks that

$$\begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix} \in (\mathrm{SOC}(n+1) \otimes \mathrm{SOC}(n+1))^*.$$

However,

$$\begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix} \notin \mathrm{SOC}(n+1)^* \otimes \mathrm{SOC}(n+1)^* = \mathrm{SOC}(n+1) \otimes \mathrm{SOC}(n+1).$$

Note that in the argument above we have used the self-duality of the second order cone.

Let us proceed to analyze the structure of $C \otimes D$ using their respective generating sublinear functions ϕ and ψ.

For simplicity, let us assume that the linear transformations '$\sim$' are simply the identity ones. This leads to

$$C = \text{epi } (\phi) \text{ and } D = \text{epi } (\psi).$$

Then, we may explicitly write the cone of the bi-positive mappings as

$$\mathcal{B}(C, D) = \left\{ \begin{bmatrix} t & b^T \\ a & M \end{bmatrix} \,\middle|\, \begin{bmatrix} \phi(u) + r \\ u \end{bmatrix}^T \begin{bmatrix} t & b^T \\ a & M \end{bmatrix} \begin{bmatrix} \psi(w) + s \\ w \end{bmatrix} \ge 0,\ \forall u \in \mathcal{U}, w \in \mathcal{W}, r, s \in \Re_+ \right\}. \quad (23)$$

We now aim at a procedure to check the membership for the cone $\mathcal{B}(C,D)$.

Consider

$$\begin{bmatrix} t & b^T \\ a & M \end{bmatrix} \in \mathcal{B}(C,D).$$

By (23) we know that

$$t(\phi(u)+r)(\psi(w)+s)+(\phi(u)+r)b^Tw+(\psi(w)+s)a^Tu+u^TMw \geq 0 \tag{24}$$

for all $u \in \mathcal{U}$, $w \in \mathcal{W}$, $r \geq 0$ and $s \geq 0$.

It follows immediately from (24) that $t \geq 0$.

If $t = 0$ then

$$(\phi(u)+r)b^Tw+(\psi(w)+s)a^Tu+u^TMw \geq 0$$

for all $u \in \mathcal{U}$, $w \in \mathcal{W}$, $r \geq 0$ and $s \geq 0$. This leads to $a^Tu \geq 0$ for all $u \in \mathcal{U}$, and $b^Tw \geq 0$ for all $w \in \mathcal{W}$. Since $\mathcal{U}$ and $\mathcal{W}$ are vector spaces, we conclude that $a = 0$ and $b = 0$. Therefore the inequality reduces further to:

$$u^TMw \geq 0, \ \forall u \in \mathcal{U} \text{ and } \forall w \in \mathcal{W}.$$

Since $\mathcal{U}$ and $\mathcal{W}$ are vector spaces, we get $M = 0$.

Now we consider the situation $t > 0$. Without losing generality let us scale the value of t and assume $t = 1$. As $\mathcal{B}(C,D)$ is a cone, the scaling does not change the nature of the membership checking procedure.

We now rewrite (24) as

$$\begin{aligned} & (\phi(u)+r)(\psi(w)+s)+(\phi(u)+r)b^Tw+(\psi(w)+s)a^Tu+u^TMw \\ = \ & \phi(u)\psi(w)+u^TMw+\phi(u)b^Tw+\psi(w)a^Tu+(\phi(u)+a^Tu)s+(\psi(w)+b^Tw)r+rs \\ \geq \ & 0 \end{aligned} \tag{25}$$

for all $u \in \mathcal{U}$, $w \in \mathcal{W}$, $r \geq 0$ and $s \geq 0$. It is evident that (25) is equivalent to the following three conditions:

$$\phi(u)+a^Tu \geq 0 \text{ for all } u \in \mathcal{U} \tag{26}$$

$$\psi(w)+b^Tw \geq 0 \text{ for all } w \in \mathcal{W} \tag{27}$$

$$\phi(u)\psi(w)+u^TMw+\phi(u)b^Tw+\psi(w)a^Tu \geq 0 \text{ for all } u \in \mathcal{U}, w \in \mathcal{W}. \tag{28}$$

Conditions (26) and (27) are equivalent to

$$\begin{bmatrix} 1 \\ a \end{bmatrix} \in C^* \tag{29}$$

and

$$\begin{bmatrix} 1 \\ b \end{bmatrix} \in D^*. \tag{30}$$

Finally, condition (28) is equivalent to the following two statements

$$\begin{bmatrix} \psi(w) + b^T w \\ \psi(w) a + M w \end{bmatrix} \in C^* \text{ for all } w \in \mathcal{W} \tag{31}$$

$$\begin{bmatrix} \phi(u) + a^T u \\ \phi(u) b + M^T u \end{bmatrix} \in D^* \text{ for all } u \in \mathcal{U}. \tag{32}$$

We note that it may or may not be a simply task to verify conditions (31) and (32). For instance, when C and D are second order cones, then conditions (31) and (32) can be reduced to verifying

$$1 + b^T w \geq \|Mw + a\| \text{ for all } \|w\| = 1$$

and

$$1 + a^T u \geq \|M^T u + b\| \text{ for all } \|u\| = 1.$$

Using the S-procedure result once again, this can be achieved by checking the following conditions:

$$\begin{cases} b^T b - 1 \leq 0 \\ \exists \tau : \begin{bmatrix} 1 - a^T a & b^T + a^T M \\ b + M^T a & b b^T - M^T M \end{bmatrix} + \tau \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix} \succeq 0 \\ a^T a - 1 \leq 0 \\ \exists \kappa : \begin{bmatrix} 1 - b^T b & a^T + b^T M^T \\ a + M b & a a^T - M M^T \end{bmatrix} + \kappa \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix} \succeq 0. \end{cases}$$

These conditions, though expressed as matrix inequalities, are not linear in a, b and M.

For general convex cones, however, as Proposition 4.1 reveals, the membership checking is a hard task.

7 Concluding remarks

In this paper we extended the definition of duality using a pre-described conic ordering relation and a bilinear mapping. The new type of dual objects are shown to be useful. It also brings up interesting theoretical questions such as how to characterize the bi-positive cones, and how to compute/approximate the relevant dual objects. A good understanding of this subject appears to be important both for the theory and practice of optimization and for convex analysis in general. As immediate research topics we pose the following two questions. (1) Is it possible to describe using LMI's the cone $(U)^*_{\mathsf{D}}$ where U is a second order cone in $\mathcal{X}$ and D is a second order cone in $\mathcal{W}$? As we remarked in Section 4, this amounts to the LMI representation of the bi-positive set $\mathcal{B}(C, D)$ where C and D are second order cones. (2) In Section 5 we showed that the dual of a multi-objective conic optimization problem, denoted by $(D)_{\mathsf{D}}$, always has an attainable global optimal solution under some suitable Slater type regularity conditions, provided that the multi-objectives are ordered by the first orthant. Does this result remain true for a general order-defining cone D or not?

References

[1] A. Ben-Tal, and A. Nemirovski, *Robust Convex Optimization*, Mathematics of Operations Research 23, pp. 769 – 805, 1998.

[2] L. El Ghaoui, and H. Lebret, *Robust solutions to least-squares problems with uncertain data*, SIAM Journal on Matrix Analysis and Applications 18, pp. 1035 – 1064, 1997.

[3] M. Grötschel, L. Lovász, and A. Schrijver, Geometric Algorithms and Combinatorial Optimization, Springer-Verlag, 1993.

[4] Z.Q. Luo, J.F. Sturm, and S. Zhang, *Duality results for conic convex programming*, Technical Report 9719/A, Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands, 1997.

[5] Z.Q. Luo, J.F. Sturm, and S. Zhang, *Multivariate nonnegative quadratic mappings*, Technical Report SEEM2003-02, The Chinese University of Hong Kong, 2003.

[6] A. Nemirovski, C. Roos, and T. Terlaky, *On maximization of quadratic form over intersection of ellipsoids with common center*, Mathematical Programming 86, pp. 463–473, 1999.

[7] Yu. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, Optimization methods and Software 9, pp. 141–160, 1998.

[8] Yu. Nesterov, *Global quadratic optimization on the sets with simple structure*, Working Paper, Université Catholique de Louvain, Louvain-la-Neuve, Belguim, 1999.

[9] Yu. Nesterov and A. Nemirovsky, Interior point polynomial methods in convex programming, Studies in Applied Mathematics 13, SIAM, Philadelphia, PA, 1994.

[10] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, New Jersey, 1970.

[11] J.F. Sturm, and S. Zhang, *On cones of nonnegative quadratic functions*. To appear in *Mathematics of Operations Research*.

[12] Y. Ye, *Approximating quadratic programming with bound and quadratic constraints*, Mathematical programming 84, pp. 219–226, 1999.

[13] S. Zhang, *A new self-dual embedding method for convex programming*. To appear in *Journal of Global Optimization*.