

# Computing critical angles between two convex cones

Wellington de Oliveira\*      Valentina Sessa\*      David Sossa†

October 3, 2023

## Abstract.

This paper addresses the numerical computation of critical angles between two convex cones in finite-dimensional Euclidean spaces. We present a novel approach to computing these critical angles by reducing the problem to finding stationary points of a fractional programming problem. To efficiently compute these stationary points, we introduce a partial linearization-like algorithm that offers significant computational advantages. Solving a sequence of strictly convex subproblems with straightforward solutions in several settings gives the proposed algorithm high computational efficiency while delivering reliable results: our theoretical analysis demonstrates that the proposed algorithm asymptotically computes critical angles. Numerical experiments validate the efficiency of our approach, even when dealing with problems of relatively large dimensions: only a few seconds are necessary to compute critical angles between different types of cones in spaces of dimension 1000.

*Mathematics Subject Classification:* 52A40; 90C26; 90C32; 90C46; 17C99

*Keywords:* Critical angle; Convex cone; Fractional Programming; Optimality conditions; Euclidean Jordan algebra.

## 1 Introduction

In this work, we are concerned with the problem of computing critical (principal) angles between two closed convex cones  $P \neq \emptyset$  and  $Q \neq \emptyset$  in Euclidean spaces. This class of problems, investigated from the theoretical point of view in [32] and [33], finds applications in image

---

\*Mines Paris, Université PSL, Centre de Mathématiques Appliquées (CMA), Sophia Antipolis, France. The research of Valentina Sessa and Wellington de Oliveira benefited from the support of the Gaspard-Monge Program for Optimization and Operations Research (PGMO) project “SOLEM - Scalable Optimization for Learning and Energy Management”.

†Universidad de O’Higgins, Instituto de Ciencias de la Ingeniería, Av. Libertador Bernardo O’Higgins 611, Rancagua, Chile (e-mail: david.sossa@uoh.cl). Partially supported by FONDECYT (Chile) through grant 11220268.

classifications and other domains [37]. In the more straightforward case where  $P$  and  $Q$  are linear subspaces, the concept of critical angles has applications in statistics and numerical linear algebra. Such a concept has been abundantly studied in the literature, both from a theoretical and computational point of view [23]. However, to the best of our knowledge, no numerical procedure exists to compute critical angles when  $P$  and  $Q$  are not linear subspaces but convex cones in medium or high dimensional spaces. This work contributes to fill the gap by proposing an implementable algorithm for computing critical angles between two convex cones. Furthermore, our framework is general enough to cover several classes of cones, such as polyhedral, ellipsoidal, Loewnerian cones, and others. Indeed, our methodology applies to the vast family of LISC cones, which are closed convex cones that can be expressed as linear images of the symmetric cones of certain Euclidean Jordan algebras, plus some mild assumptions (see Definition 4 below).

By definition, the maximal angle between  $P$  and  $Q$ , in a Euclidean space  $\mathbb{V}$  equipped with an inner product  $\langle \cdot, \cdot \rangle$ , is the maximum value of  $\arccos \langle u, v \rangle$  on  $(P \cap S_{\mathbb{V}}) \times (Q \cap S_{\mathbb{V}})$ , where  $S_{\mathbb{V}}$  denotes the unit sphere of  $\mathbb{V}$ . Critical angles between  $P$  and  $Q$  are angles formed by vectors  $u \in P$  and  $v \in Q$  satisfying some complementarity conditions (see Definition 1), which are necessary conditions for  $u$  and  $v$  to achieve the maximal angle between  $P$  and  $Q$ .

The study of critical angles for the case  $P = Q$  was first carried out by Iusem and Seeger [19–22]. Later, Seeger and Sossa expanded this study to two different cones [32,33]. See also the recent contributions of Orlitzky [24] and Bauschke et al. [3]. Concerning the computation of critical angles, some methods were given in [33] for specially structured cones such as revolution cones, ellipsoidal, and cones of matrices. These methods are either applicable to specific instances of cones or limited to problems on spaces of small dimensions. In [32], an exhaustive method was developed for computing all critical angles between polyhedral cones. Basically, one must select subsets of generators of the cones, and solve a generalized eigenvalue problem for each selection. This has a combinatorial nature and its numerical implementation is only possible for polyhedral cones with few generators.

In this work, we show that computing critical angles between two convex cones reduces to the problem of finding stationary points of a fractional programming (FP) problem, that is, the optimization of the ratio of two continuous functions over a nonempty, closed convex set [13] (see (5) for the precise definition). When specialized to the family of LISC cones, the feasible set of our FP is the Cartesian product of spectraplexes (sets analogous to the standard simplex) of the involved Euclidean Jordan algebras. Thus, from the numerical point of view, our FP is easier to handle than the original maximal angle problem formulation, whose feasible set has spherical conditions. Notably, most of the theoretical and algorithmic works in fractional programming are applied to the convex case, i.e., convex numerator and concave denominator. Dinkelbach’s method [9, 36] is the classical approach used to solve convex fractional problems by tackling a parametric reformulation of the original FP problem. Several iterative schemes based on Dinkelbach’s idea are proposed for special cases of convex FP [4, 5, 10, 35, 38]. In the recent paper [6], an extrapolated proximal subgradient algorithm is applied to nonconvex and nonsmooth FPs. A disadvantage of that algorithm is that a nonconvex subproblem must be globally solved at every iteration, restricting thus the approach to a few particular FP cases. In theory, that algorithm could be applied to our FP formulation, but the structure of our problem is not (computationally) favorable to that approach.

This study addresses the challenges posed by our FP, a nonconvex FP problem that does not conform to the requirements of existing practical methodologies. Inspired by the recent

paper [14], we introduce a Sequential Regularized Partial Linearization algorithm to compute a stationary point of our FP formulation, thus a critical angle. Our algorithm, that features the main contribution of this work, solves a sequence of two independent convex subproblems, which are projections onto well-structured sets. The essential advantage of our approach lies in the fact that solving these subproblems requires minimal computational efforts as they have straightforward solutions in several settings of practical interest. To demonstrate the effectiveness of our approach, we provide numerical results that focus on computing critical angles between LISC cones. These angles are commonly examined in the relevant literature, but a thorough numerical treatment of these problems has not been previously investigated.

This work is organized as follows. Section 2 presents the main definitions and preliminary results. The concept of LISC cones and relevant properties are presented in Section 3. Section 4 clarifies the connection between critical angles and stationary points of the FP formulation. The proposed algorithm is presented in Section 5 along with a converge analysis. Numerical experiments are reported in Section 6. Finally, some conclusions are given in Section 7.

## 2 Main definitions and prerequisites

Let  $\mathbb{V}$ ,  $\mathbb{X}$  and  $\mathbb{Y}$  be Euclidean spaces. We denote by  $\mathcal{C}(\mathbb{V})$  the set of all nontrivial closed convex cones in  $\mathbb{V}$ , that is,

$$\mathcal{C}(\mathbb{V}) = \{K \text{ closed convex cone in } \mathbb{V} : K \neq \{0\}, K \neq \mathbb{V}\}. \quad (1)$$

The dual cone of  $K \in \mathcal{C}(\mathbb{V})$  is denoted by  $K^*$ :

$$K^* := \{\zeta \in \mathbb{V} : \langle \zeta, u \rangle \geq 0, \forall u \in K\}.$$

Throughout this work,  $\langle \cdot, \cdot \rangle$  refers to the inner product of each Euclidean space that we are considering. The corresponding inner product to its space will be clear from the context. In particular, whenever we consider the Euclidean space  $\mathbb{R}^n$ , we assume that its inner product is  $\langle x, y \rangle = x^\top y$ , for all  $x, y \in \mathbb{R}^n$ . The norm  $\|\cdot\|$  refers to the Euclidean norm induced by  $\langle \cdot, \cdot \rangle$ .

Given two cones  $P, Q \in \mathcal{C}(\mathbb{V})$ , by definition, the maximal angle between  $P$  and  $Q$  is given by

$$\Theta(P, Q) := \begin{cases} \max & \arccos \langle u, v \rangle \\ \text{s.t.} & u \in P, \|u\| = 1 \\ & v \in Q, \|v\| = 1. \end{cases} \quad (2)$$

As in *data analysis*, the cosinus similarity of two vectors  $u, v$  is defined as the cosinus of the angle  $\theta$  between them, i.e.,  $\cos(\theta) = \frac{\langle u, v \rangle}{\|u\| \|v\|}$ . Recall that two proportional vectors have a cosinus similarity of 1, two orthogonal vectors have a similarity of 0, and two opposite vectors have a similarity of -1. Hence, the cosinus of the maximal angle between  $P$  and  $Q$  can be computed by solving the following (nonconvex) optimization problem:

$$\begin{cases} \min & \langle u, v \rangle \\ \text{s.t.} & u \in P, \|u\| = 1, \\ & v \in Q, \|v\| = 1. \end{cases} \quad (3)$$

The following definition is crucial for the remainder of this work.

**Definition 1.** (Maximal and critical angles between two cones).

1. A pair  $(u, v) \in \mathbb{V} \times \mathbb{V}$  solving problem (3) is called an antipodal pair of  $(P, Q)$ . The angle between an antipodal pair is called the maximal angle between the cones  $P$  and  $Q$ .
2. A pair  $(u, v)$  is called a critical pair of  $(P, Q)$  if it satisfies the following conditions:

$$\begin{cases} u \in P, \|u\| = 1, \\ v \in Q, \|v\| = 1, \\ v - \langle u, v \rangle u \in P^*, \\ u - \langle u, v \rangle v \in Q^*, \end{cases} \quad (4)$$

where  $P^*$  and  $Q^*$  are the dual cones of  $P$  and  $Q$ , respectively. The angle between a critical pair is called a critical angle between the cones  $P$  and  $Q$ .

As we will see in Section 3, in many important cases the cones  $P$  and  $Q$  are images of linear mappings applied to more structured/simpler cones  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$ , i.e.,  $P = G(K_{\mathbb{X}})$  and  $Q = H(K_{\mathbb{Y}})$ , with  $G$  and  $H$  linear applications of appropriated dimensions. In these cases, it is computationally advantageous to work with the more structured cones  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$ . As we will see in Section 4, problem (2) can be reformulated as

$$\cos[\Theta(P, Q)] = \begin{cases} \min & \Phi(x, y) := \frac{\langle Gx, Hy \rangle}{\|Gx\| \|Hy\|} \\ \text{s.t.} & x \in K_{\mathbb{X}}, \langle e_{\mathbb{X}}, x \rangle = 1 \\ & y \in K_{\mathbb{Y}}, \langle e_{\mathbb{Y}}, y \rangle = 1. \end{cases} \quad (5)$$

In this case,  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$  are the symmetric cones of certain Euclidean Jordan algebras  $\mathbb{X}$  and  $\mathbb{Y}$ , and  $e_{\mathbb{X}}$  and  $e_{\mathbb{Y}}$  denote the unit elements of  $\mathbb{X}$  and  $\mathbb{Y}$ , respectively.

Observe that the optimization problems (3) and (5) share a common structure. Indeed, they fit into the following more general class of problems:

$$\begin{cases} \min & \Psi(x, y) \\ \text{s.t.} & x \in K_{\mathbb{X}}, \phi_1(x) = 0 \\ & y \in K_{\mathbb{Y}}, \phi_2(y) = 0, \end{cases} \quad (6)$$

where  $\Psi : \mathbb{X} \times \mathbb{Y} \rightarrow \mathbb{R}$ ,  $\phi_1 : \mathbb{X} \rightarrow \mathbb{R}$ , and  $\phi_2 : \mathbb{Y} \rightarrow \mathbb{R}$  are differentiable functions over the feasible set of (6).

**Definition 2.** We say that  $(\bar{x}, \bar{y})$  is a stationary point of (6) if there exist  $\gamma_1, \gamma_2 \in \mathbb{R}$  such that

$$\begin{cases} K_{\mathbb{X}} \ni \bar{x} \perp (\nabla_x \Psi(\bar{x}, \bar{y}) + \gamma_1 \nabla \phi_1(\bar{x})) \in K_{\mathbb{X}}^*, \\ K_{\mathbb{Y}} \ni \bar{y} \perp (\nabla_y \Psi(\bar{x}, \bar{y}) + \gamma_2 \nabla \phi_2(\bar{y})) \in K_{\mathbb{Y}}^*, \\ \phi_1(\bar{x}) = 0, \phi_2(\bar{y}) = 0, \end{cases} \quad (7)$$

where the notation  $\perp$  means orthogonality.

We now show that our concept of stationarity coincides with that of classic stationarity for the nonlinear optimization problem (6).

**Theorem 1.** Let  $K_{\mathbb{X}} \in \mathcal{C}(\mathbb{X})$ ,  $K_{\mathbb{Y}} \in \mathcal{C}(\mathbb{Y})$  and suppose that  $\Psi, \phi_1$  and  $\phi_2$  are differentiable over the feasible set of problem (6). Then the definition of stationarity given in (7) coincides with that of stationarity for the nonlinear programming problem (6).

*Proof.* We can rewrite problem (6) with the help of the indicator function of the convex sets  $C_1 := K_{\mathbb{X}} \times \mathbb{Y}$  and  $C_2 := \mathbb{X} \times K_{\mathbb{Y}}$  as follows:

$$\begin{cases} \min & \Psi(x, y) + i_{C_1}(x, y) + i_{C_2}(x, y) \\ \text{s.t.} & \phi_1(x) = 0, \phi_2(y) = 0. \end{cases}$$

If the above problem has Lagrange multipliers  $\gamma_1, \gamma_2 \in \mathbb{R}$  (e.g., if  $(\nabla\phi_1(\bar{x}), \nabla\phi_2(\bar{y})) \neq (0, 0)$  – satisfying thus the LICQ constraint qualification), then all of its (local) solutions must satisfy the KKT system (see for instance [29, §10])

$$\begin{cases} 0 \in \partial_{(x,y)}^{\mathcal{C}} L(\bar{x}, \bar{y}, \gamma) \\ \bar{x} \in K_{\mathbb{X}}, \phi_1(\bar{x}) = 0 \\ \bar{y} \in K_{\mathbb{Y}}, \phi_2(\bar{y}) = 0, \end{cases}$$

where the Lagrangian function is given by

$$L(x, y, \gamma) := \Psi(x, y) + i_{C_1}(x, y) + i_{C_2}(x, y) + \gamma_1\phi_1(x) + \gamma_2\phi_2(y),$$

and  $\partial^{\mathcal{C}} f$  denotes the Clarke subdifferential of a locally Lipschitz function  $f$ . Observe that

$$\partial_{(x,y)}^{\mathcal{C}} L(x, y, \gamma) = \begin{pmatrix} \nabla_x \Psi(x, y) \\ \nabla_y \Psi(x, y) \end{pmatrix} + \partial^{\mathcal{C}}[i_{C_1}(x, y) + i_{C_2}(x, y)] + \begin{pmatrix} \gamma_1 \nabla \phi_1(x) \\ \gamma_2 \nabla \phi_2(y) \end{pmatrix},$$

and being  $i_{C_1}(x, y) + i_{C_2}(x, y)$  a convex function, its Clarke subdifferential coincides with that of the Convex Analysis. Thus,  $\partial^{\mathcal{C}}[i_{C_1}(x, y) + i_{C_2}(x, y)] = \partial[i_{C_1}(x, y) + i_{C_2}(x, y)] = \partial i_{C_2}(x, y) + \partial i_{C_1}(x, y)$  provided  $\text{ri dom}(i_{C_1}) \cap \text{ri dom}(i_{C_2}) \neq \emptyset$  (see [30, Thm 23.8]). Furthermore,  $\text{ri dom}(i_{C_1}) = \text{ri } K_{\mathbb{X}} \times \mathbb{Y}$  and  $\text{ri dom}(i_{C_2}) = \mathbb{X} \times \text{ri } K_{\mathbb{Y}}$ . As  $K_{\mathbb{X}} \neq \{0\}$  and  $K_{\mathbb{Y}} \neq \{0\}$  (c.f. (1)) then  $\text{ri } K_{\mathbb{X}} \neq \emptyset$  and  $\text{ri } K_{\mathbb{Y}} \neq \emptyset$ . Recall that  $C_1$  is a convex set, and thus  $\partial i_{C_1}(x, y) = N_{C_1}(x, y) = N_{K_{\mathbb{X}}}(x) \times 0$ . Analogously,  $\partial i_{C_2}(x, y) = 0 \times N_{K_{\mathbb{Y}}}(y)$ . Hence, the above KKT system becomes

$$\begin{cases} 0 \in \begin{pmatrix} \nabla_x \Psi(\bar{x}, \bar{y}) \\ \nabla_y \Psi(\bar{x}, \bar{y}) \end{pmatrix} + \begin{pmatrix} N_{K_{\mathbb{X}}}(\bar{x}) \\ N_{K_{\mathbb{Y}}}(\bar{y}) \end{pmatrix} + \begin{pmatrix} \gamma_1 \nabla \phi_1(\bar{x}) \\ \gamma_2 \nabla \phi_2(\bar{y}) \end{pmatrix} \\ \bar{x} \in K_{\mathbb{X}}, \phi_1(\bar{x}) = 0 \\ \bar{y} \in K_{\mathbb{Y}}, \phi_2(\bar{y}) = 0. \end{cases} \quad (8)$$

Then, by using the well-known results in [11, Prop. 1.1.3], we can see that the above system is nothing but the one given in (7). □

It is not surprising that (4) is a particular case of (7), as the following corollary shows.

**Corollary 1.** *Under the setting of Theorem 1, suppose that  $K_{\mathbb{X}} = P$ ,  $K_{\mathbb{Y}} = Q$ ,  $\Psi(x, y) = \langle x, y \rangle$ ,  $\phi_1(x) = \|x\| - 1$ , and  $\phi_2(y) = \|y\| - 1$ , i.e., problem (6) coincides with (3). Furthermore, suppose that  $(\bar{x}, \bar{y})$  is a stationary point of problem (6). Then, the Lagrange multipliers  $\gamma_1, \gamma_2 \in \mathbb{R}$  in (7) are given by  $\gamma_1 = \gamma_2 = -\langle \bar{x}, \bar{y} \rangle$ .*

*Proof.* In this setting, the Lagrange multipliers always exist due to the fact that  $(\nabla\phi_1(\bar{x}), \nabla\phi_2(\bar{y})) = (\bar{x}/\|\bar{x}\|, \bar{y}/\|\bar{y}\|) \neq (0, 0)$ , i.e., the problem satisfies the LICQ constraint qualification. Furthermore, system (8) becomes

$$\begin{cases} 0 \in \begin{pmatrix} \bar{y} \\ \bar{x} \end{pmatrix} + \begin{pmatrix} N_P(\bar{x}) \\ N_Q(\bar{y}) \end{pmatrix} + \begin{pmatrix} \gamma_1 \bar{x}/\|\bar{x}\| \\ \gamma_2 \bar{y}/\|\bar{y}\| \end{pmatrix} \\ \bar{x} \in P, \|\bar{x}\| = 1 \\ \bar{y} \in Q, \|\bar{y}\| = 1. \end{cases}$$

Let  $\zeta \in N_P(\bar{x})$  such that  $0 = \bar{y} + \zeta + \gamma_1 \bar{x} / \|\bar{x}\|$ . By multiplying this equality by  $\bar{x}$  and recalling that  $\|\bar{x}\| = 1$ , we get  $0 = \langle \bar{x}, \bar{y} \rangle + \langle \bar{x}, \zeta \rangle + \gamma_1$ . As  $\langle \bar{x}, \zeta \rangle = 0$ , we have that  $\gamma_1 = -\langle \bar{x}, \bar{y} \rangle$ . Analogously,  $\gamma_2 = -\langle \bar{x}, \bar{y} \rangle$ .  $\square$

It is straightforward to see that, in the setting of the above corollary, the conditions in (7) boils down to (4).

### 3 LISC cones

This section presents the concept of LISC cones which are sets obtained as *Linear Image of Symmetric Cones*. Recall that a *symmetric cone* in a Euclidean space is a cone that is self-dual and homogeneous, see [12, Section I.1]. The symmetric cones possess a reach structure and they are central objects in many optimization problems such as linear programming, second-order cone programming, and semidefinite programming [1, 28, 39]. For those programs, the theory and the numerical resolution are well-developed and nowadays there are many solvers available to deal with them. It turns out that many convex cones arise as the linear image of symmetric ones. As we shall see in the next sections, the computation of critical angles for these cones can be performed by solving a sequence of (convex) symmetric cone programs.

#### 3.1 Main definitions and some examples

It is known that every symmetric cone is obtained as the cone of square elements of a certain Euclidean Jordan Algebra (EJA), see [12, Theorem III.3.1]. Thus, to present our results we shall use the machinery of EJAs. We refer the reader to the book of Faraut and Korányi [12] for an exposition of the definitions and main results of this algebra.

Let  $\mathbb{X}$  be an EJA of rank  $r$  with unit element  $e$ , and equipped with an inner product  $\langle \cdot, \cdot \rangle$  and a Jordan product  $\circ$ . We need to recall some ingredients from EJA concerning the spectral decomposition [12, Theorem III.1.2]. An element  $c \in \mathbb{X}$  is called *idempotent* if  $c \circ c = c$ . An idempotent element  $c$  is called *primitive* if it is non-zero and cannot be written as the sum of two non-zero idempotents. A collection of primitive idempotent elements  $\{c_1, \dots, c_r\}$  is called *Jordan frame* if  $c_i \circ c_j = 0$  for all  $i \neq j$  and  $\sum_{i=1}^r c_i = e$ . Any element  $x$  in  $\mathbb{X}$  admits a spectral decomposition. That is, there exist scalars  $\lambda_1, \dots, \lambda_r$  (called the eigenvalues of  $x$ ) and a Jordan frame  $\{c_1, \dots, c_r\}$  such that  $x = \sum_{i=1}^r \lambda_i c_i$ . The trace of  $x = \sum_{i=1}^r \lambda_i c_i$  is defined as  $\text{tr}(x) = \lambda_1 + \dots + \lambda_r$ .

**Definition 3** (Symmetric cone in  $\mathbb{X}$ ). *We denote by  $K_{\mathbb{X}}$  the symmetric cone in  $\mathbb{X}$ , which corresponds to the cone of square elements of  $\mathbb{X}$ , that is,  $K_{\mathbb{X}} = \{x \circ x : x \in \mathbb{X}\}$ .*

Observe that  $x = \sum_{i=1}^r \lambda_i c_i \in K_{\mathbb{X}}$  if and only if  $\lambda_i \geq 0$  for all  $i = 1, \dots, r$ . Below, we provide some examples of EJAs and corresponding symmetric cones.

**Example 1** (EJAs and corresponding symmetric cones).

- (a) The space  $\mathbb{R}^n$  is an EJA of rank  $n$  with unit element  $\mathbf{1}_n = (1, \dots, 1)^\top$ ,  $\langle x, y \rangle = x^\top y$ , and  $x \circ y = (x_1 y_1, \dots, x_n y_n)^\top$ . Its associated symmetric cone is the nonnegative orthant  $\mathbb{R}_+^n$ .
- (b) Alternatively,  $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$  is an EJA (Jordan spin algebra) of rank 2 with unit element given by  $e = (0, \dots, 0, 1)^\top$ . For  $x = (\xi, t)$ ,  $y = (\eta, s) \in \mathbb{R}^{n-1} \times \mathbb{R}$ , the inner product is  $\langle x, y \rangle = x^\top y$  and  $x \circ y = (s\xi + t\eta, x^\top y)$ . Its corresponding symmetric cone is the *Lorentz cone*, which is given by  $\mathcal{L}_+^n := \{(\xi, t) \in \mathbb{R}^n : \|\xi\| \leq t\}$ .

- (c) The space  $\mathcal{S}^n$  of the symmetric matrices of order  $n$  is an EJA of rank  $n$  with unit element  $I$  (the identity matrix of order  $n$ ), equipped with the Frobenius (or trace) inner product  $\langle X, Y \rangle = \text{Tr}(XY)$  and the Jordan product  $X \circ Y = (XY + YX)/2$ . Its associated symmetric cone is the *positive semidefinite (SDP) cone* which is defined as follows,  $\mathcal{P}_n := \{X \in \mathcal{S}^n : u^\top X u \geq 0, \forall u \in \mathbb{R}^n\}$ .

A cone  $P$  in a Euclidean space  $\mathbb{V}$  is a LISC cone if it can be written as the linear image of some symmetric cone satisfying some mild assumptions. More precisely:

**Definition 4.** *Let  $P$  be a closed convex cone in a Euclidean space  $\mathbb{V}$ . We say that  $P$  is a LISC cone if there exist a Euclidean Jordan algebra  $\mathbb{X}$  and a linear map  $G : \mathbb{X} \rightarrow \mathbb{V}$  such that  $G(K_{\mathbb{X}}) = P$ , satisfying the following assumptions:*

$$G(c) \neq 0 \text{ for every primitive idempotent element } c \in \mathbb{X}. \quad (\text{A1})$$

$$P \text{ is pointed (i.e. } P \cap (-P) = \{0\}\text{)}. \quad (\text{A2})$$

Observe that in the above definition, we are assuming that  $P = G(K_{\mathbb{X}})$  is closed in advance. This is because in general, the linear image of  $K_{\mathbb{X}}$  may fail to be closed (cf. [26]).

There are vast classes of convex cones that are LISC cones. We now list some of them.

**Example 2** (LISC cones).

- (a) *Polyhedral cones* [17]. A cone  $P$  in  $\mathbb{R}^n$  is polyhedral if there exist vectors  $g_1, \dots, g_p \in \mathbb{R}^n$  (generators of  $P$ ) such that

$$P = \text{cone}\{g_1, \dots, g_p\} := \{\alpha_1 g_1 + \dots + \alpha_p g_p : \alpha_i \geq 0, \forall i = 1, \dots, p\}.$$

Then, by taking  $G := [g_1 \ \dots \ g_p]$  as the  $n \times p$  matrix formed by the generators of  $P$ , we have that  $P = G(\mathbb{R}_+^p)$ . It is usual to assume that the generators of  $P$  are positive linear independent (i.e., if  $\sum_{i=1}^p \alpha_i g_i = 0$  and  $\alpha_i \geq 0$  for all  $i$ , then  $\alpha_i = 0$  for all  $i$ ). This property implies (A1) and (A2). Indeed, the primitive idempotent elements of  $\mathbb{R}^p$  are  $e_1, \dots, e_p$  which are the elements of the canonical basis of  $\mathbb{R}^p$ . Observe that  $G(e_i) = g_i \neq 0$  for every  $i$ . Then, (A1) holds. On the other hand, suppose that (A2) is not satisfied. That is, suppose that  $P$  is not pointed. Then, there exists  $u \neq 0$  such that  $u = \sum_{i=1}^p \alpha_i g_i$  and  $-u = \sum_{i=1}^p \beta_i g_i$  for some  $\alpha_i, \beta_i \geq 0$  for all  $i$ . Then,  $\sum_{i=1}^p (\alpha_i + \beta_i) g_i = 0$ . Since we assume that the generators of  $P$  are positive linear independents, we conclude that  $\alpha_i = \beta_i = 0$  for all  $i$ . It means  $u = 0$  which is a contradiction. Hence,  $P$  is a LISC cone with the positive orthant  $\mathbb{R}_+^p$  as its associated symmetric cone.

- (b) *Ellipsoidal cones* [20, 33, 34]. A cone  $P$  in  $\mathbb{R}^n$  is ellipsoidal if there exists a symmetric positive definite matrix  $A$  of order  $n - 1$  such that

$$P = \{(\xi, t) \in \mathbb{R}^{n-1} \times \mathbb{R} : \sqrt{\langle \xi, A\xi \rangle} \leq t\}.$$

By taking the linear map  $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$  given by  $G(\xi, t) = (A^{-1/2}\xi, t)$ , we have that  $P = G(\mathcal{L}_+^n)$ . Since  $G$  is invertible and  $\mathcal{L}_+^n$  is pointed, we have that (A1) and (A2) are satisfied. Thus,  $P$  is a LISC cone with the Lorentz cone  $\mathcal{L}_+^n$  as its associated symmetric cone.

- (c) *Loewnerian cones* [31]. By definition, a cone  $P$  in  $\mathcal{S}^n$  is Loewnerian if there exists an invertible linear map  $G : \mathcal{S}^n \rightarrow \mathcal{S}^n$  such that  $P = G(\mathcal{P}_n)$ . Assumptions (A1) and (A2) hold because  $G$  is invertible and  $\mathcal{P}_n$  pointed. Then,  $P$  is a LISC cone with the SDP cone  $\mathcal{P}_n$  as its associated symmetric cone.

As we will shortly see, our technique to find critical angles between a pair of LISC cones consists of reformulating problem (2) as a minimization problem over a convex set. More precisely, for a pair of LISC cones  $P = G(K_{\mathbb{X}})$  and  $Q = H(K_{\mathbb{Y}})$ , the (bilinear) minimization problem (3) with nonconvex feasible set

$$S_{P,Q} := \{(u, v) \in P \times Q : \|u\| = 1, \|v\| = 1\}$$

will be reformulated as a (fractional) minimization problem (5) with convex set

$$\Delta_{\mathbb{X},\mathbb{Y}} := \{(x, y) \in K_{\mathbb{X}} \times K_{\mathbb{Y}} : \langle e_{\mathbb{X}}, x \rangle = 1, \langle e_{\mathbb{Y}}, y \rangle = 1\},$$

where  $e_{\mathbb{X}}$  and  $e_{\mathbb{Y}}$  are the unit elements of the EJAs  $\mathbb{X}$  and  $\mathbb{Y}$ , respectively. Observe that  $S_{P,Q} = S_P \times S_Q$ , where

$$S_P := \{u \in P : \|u\| = 1\} \quad \text{and} \quad S_Q := \{v \in Q : \|v\| = 1\}.$$

Analogously,  $\Delta_{\mathbb{X},\mathbb{Y}} = \Delta_{\mathbb{X}} \times \Delta_{\mathbb{Y}}$ , where

$$\Delta_{\mathbb{X}} := \{x \in K_{\mathbb{X}} : \langle e_{\mathbb{X}}, x \rangle = 1\} \quad \text{and} \quad \Delta_{\mathbb{Y}} := \{y \in K_{\mathbb{Y}} : \langle e_{\mathbb{Y}}, y \rangle = 1\}.$$

Here,  $\Delta_{\mathbb{X}}$  and  $\Delta_{\mathbb{Y}}$  are called the *spectraplexes* of  $\mathbb{X}$  and  $\mathbb{Y}$ , respectively. A key fact to formulate the maximal angle problem as the minimization problem (3) is that each nonzero  $u \in P$  can be written as a positive multiple of an element of  $S_P$ ; indeed,  $u = \|u\| \frac{u}{\|u\|}$  and  $\frac{u}{\|u\|} \in S_P$  (an analogous property is for a nonzero  $v \in Q$ ). In the same way, we must ensure that each nonzero  $x \in K_{\mathbb{X}}$  can be written as a positive multiple of an element of  $\Delta_{\mathbb{X}}$  (an analogous property for a nonzero  $y \in K_{\mathbb{Y}}$ ). This is proved in the following lemma.

**Lemma 1.** *Let  $\mathbb{X}$  be an EJA of rank  $r$  with unit element  $e_{\mathbb{X}}$ . Let  $K_{\mathbb{X}}$  be the symmetric cone of  $\mathbb{X}$ . Then, the following statements are satisfied:*

1.  $\langle e_{\mathbb{X}}, x \rangle > 0$  for all  $x \in K_{\mathbb{X}} \setminus \{0\}$ .
2.  $\{\alpha x : x \in \Delta_{\mathbb{X}}, \alpha > 0\} = K_{\mathbb{X}} \setminus \{0\}$ .

*Proof.* Let  $x = \sum_{i=1}^r \lambda_i c_i$  be a nonzero element in  $K_{\mathbb{X}}$ . Then,  $\lambda_i \geq 0$  for all  $i = 1, \dots, r$ , and not all of them are zero. Thus,

$$\langle e_{\mathbb{X}}, x \rangle = \left\langle \sum_{i=1}^r c_i, \sum_{i=1}^r \lambda_i c_i \right\rangle = \sum_{i=1}^r \lambda_i \|c_i\|^2 \geq 0.$$

Then, if  $\langle e_{\mathbb{X}}, x \rangle = 0$  we have that  $\lambda_i \|c_i\|^2 = 0$  for every  $i = 1, \dots, r$ . Since any primitive element is nonzero, we deduce that  $\lambda_i = 0$  for all  $i = 1, \dots, r$  which contradicts the fact that  $x$  is nonzero. We conclude  $\langle e_{\mathbb{X}}, x \rangle > 0$ .

Let us prove the second part. Since  $\Delta_{\mathbb{X}}$  is contained in  $K_{\mathbb{X}} \setminus \{0\}$  and because of the cone structure of  $K_{\mathbb{X}} \setminus \{0\}$ , we have that  $\{\alpha x : x \in \Delta_{\mathbb{X}}, \alpha > 0\} \subseteq K_{\mathbb{X}} \setminus \{0\}$ . Let  $u \in K_{\mathbb{X}} \setminus \{0\}$ . From the previous part, we have that  $\langle e_{\mathbb{X}}, u \rangle > 0$ . Then,  $u = \alpha x$  with  $\alpha := \langle e_{\mathbb{X}}, u \rangle > 0$  and  $x := u / \langle e_{\mathbb{X}}, u \rangle \in \Delta_{\mathbb{X}}$ . The proof is complete.  $\square$

Now, to pass from problem (2) (which is equivalent to (3)) to the fractional one (5), we must ensure that  $Gx \neq 0$  for every  $x \in K_{\mathbb{X}} \setminus \{0\}$  because we shall need to normalize  $Gx$  (an analogous observation is for  $Hy$ ). The next lemma shows that assumptions (A1) and (A2) help us to achieve that property. But before, observe that a general LISC cone can also be described by a sort of conic combination of a (possibly uncountable) set of vectors, similar to the polyhedral case. Indeed, let  $\mathcal{O}_{\mathbb{X}}$  denote the set of (ordered) Jordan frames of  $\mathbb{X}$ . That is,  $(c_1, \dots, c_r) \in \mathcal{O}_{\mathbb{X}}$  if and only if  $\{c_1, \dots, c_r\}$  is a Jordan frame of  $\mathbb{X}$ . From the spectral decomposition property, it is not difficult to see that  $K_{\mathbb{X}}$  can be written as

$$K_{\mathbb{X}} = \left\{ \sum_{i=1}^r \alpha_i c_i : \alpha_1, \dots, \alpha_r \geq 0, (c_1, \dots, c_r) \in \mathcal{O}_{\mathbb{X}} \right\} = \bigcup_{(c_1, \dots, c_r) \in \mathcal{O}_{\mathbb{X}}} \text{cone}\{c_1, \dots, c_r\}.$$

Then, the LISC cone  $P = G(K_{\mathbb{X}})$  can be written as

$$P = \bigcup_{(c_1, \dots, c_r) \in \mathcal{O}_{\mathbb{X}}} \text{cone}\{G(c_1), \dots, G(c_r)\}. \quad (9)$$

Observe that in the polyhedral case, the representation (9) coincides with the conic combination of its generators since the only Jordan frame of  $\mathbb{R}^r$  is its canonical basis.

**Lemma 2.** *Let  $P = G(K_{\mathbb{X}})$  be a LISC cone in  $\mathbb{V}$  with  $\mathbb{X}$  being an EJA of rank  $r$ . Then,  $Gx \neq 0$  for every  $x \in K_{\mathbb{X}} \setminus \{0\}$ .*

*Proof.* Reasoning by contradiction, suppose that there is  $x = \sum_{i=1}^r \lambda_i c_i$  in  $K_{\mathbb{X}} \setminus \{0\}$  such that  $G(x) = 0$ . Then,  $\sum_{i=1}^r \lambda_i G(c_i) = 0$ . Since  $x$  is nonzero, we may assume that  $\lambda_1 \neq 0$ . Thus,

$$-G(c_1) = \sum_{i=2}^r \left( \frac{\lambda_i}{\lambda_1} \right) G(c_i).$$

Because of  $x \in K_{\mathbb{X}}$ , we have that  $\lambda_i \geq 0$  for every  $i$ . Then, from representation (9) we deduce that  $-G(c_1) \in P$ . On the other hand, it is clear, also for the representation (9), that  $G(c_1) \in P$ . Because of the assumption (A1), we have that  $G(c_1) \neq 0$ . We have proved that  $P \cap (-P)$  has a nonzero element which contradicts assumption (A2).  $\square$

As a last ingredient of EJA, we now explain how to compute the orthogonal projection of  $b \in \mathbb{X}$  onto the spectraplex  $\Delta_{\mathbb{X}} = \{x \in K_{\mathbb{X}} : \langle e_{\mathbb{X}}, x \rangle = 1\}$ . This result will be used in our algorithmic approach as we shall see in the next sections.

### 3.2 Projection onto spectraplexes

An EJA is said to be *scalarizable* if there exists  $\kappa > 0$  such that  $\langle x, y \rangle = \kappa \text{tr}(x \circ y)$  for every  $x, y \in \mathbb{X}$ . The number  $\kappa$  is called the scaling factor of the EJA. The EJAs in Example 1 are scalarizable: the scaling factors of  $\mathbb{R}^n$ ,  $\mathbb{R}^{n-1} \times \mathbb{R}$  and  $\mathcal{S}^n$  are 1, 1/2 and 1, respectively. From now on, we assume that all the EJAs considered in this work are scalarizable. For  $x \in \mathbb{X}$ ,  $\lambda(x)$  denotes the vector in  $\mathbb{R}^r$  whose components are the eigenvalues of  $x$  arranged in nonincreasing order, i.e.,  $\lambda_1(x) \geq \lambda_2(x) \geq \dots \geq \lambda_r(x)$  are the eigenvalues of  $x$ . A set  $\Omega \subseteq \mathbb{X}$  is called *spectral set* if there exists a permutation invariant set  $\mathcal{Q} \subseteq \mathbb{R}^r$  such that  $\Omega = \lambda^{-1}(\mathcal{Q})$ . Observe that  $\Delta_{\mathbb{X}}$  is a spectral set whose associated permutation invariant set is

$$\Delta_r^{\kappa} := \left\{ \lambda \in \mathbb{R}^r : \lambda \geq 0, \langle \mathbf{1}_r, \lambda \rangle = \frac{1}{\kappa} \right\}.$$

Indeed,  $x \in \Delta_{\mathbb{X}}$  if and only if  $\lambda_i(x) \geq 0$  for all  $i$ , and

$$1 = \langle e_{\mathbb{X}}, x \rangle = \kappa \text{tr}(e_{\mathbb{X}} \circ x) = \kappa \text{tr}(x) = \kappa(\lambda_1(x) + \cdots + \lambda_r(x)).$$

It means  $\lambda(x) \geq 0$  and  $\langle \mathbf{1}_r, \lambda(x) \rangle = \frac{1}{\kappa}$ . Thus,  $\Delta_{\mathbb{X}} = \lambda^{-1}(\Delta_r^\kappa)$ . Observe that when  $\kappa = 1$ , which is the case for the algebras  $\mathbb{R}^n$  and  $\mathcal{S}^n$ ,  $\Delta_r^\kappa$  coincides with the standard simplex

$$\Delta_r := \{x \in \mathbb{R}^r : x \geq 0, \langle \mathbf{1}_r, x \rangle = 1\}.$$

Incidentally, we observe that  $\Delta_{\mathbb{X}}$  is compact. Indeed, it is clear that  $\Delta_r^\kappa$  is compact in  $\mathbb{R}^r$ . Then, the compactness of  $\Delta_r^\kappa$  is transferred to its associated spectral set  $\Delta_{\mathbb{X}}$  [2, Theorem 27]. This result will be used later, so we state it as a proposition.

**Proposition 1.** *The spectraplex  $\Delta_{\mathbb{X}}$  is compact in  $\mathbb{X}$ .*

The next proposition shows that the computation of the orthogonal projection onto  $\Delta_{\mathbb{X}}$  is reduced to the computation of the orthogonal projection onto  $\Delta_r^\kappa$  which is easier to handle.

**Proposition 2.** *Let  $\mathbb{X}$  be a scalarizable EJA with rank  $r$  and scaling factor  $\kappa$ . Let  $b \in \mathbb{X}$  with spectral decomposition  $b = \sum_{i=1}^r \lambda_i(b)c_i$ . Then,  $\bar{x} = \sum_{i=1}^r \bar{\xi}_i c_i$  is the solution of*

$$\begin{aligned} \min \quad & \frac{1}{2} \|x - b\|^2 \\ \text{s.t.} \quad & x \in \Delta_{\mathbb{X}}, \end{aligned} \tag{10}$$

where  $\bar{\xi} := (\bar{\xi}_1, \dots, \bar{\xi}_r)^\top$  is the solution of

$$\begin{aligned} \min \quad & \frac{1}{2} \|\xi - \lambda(b)\|^2 \\ \text{s.t.} \quad & \xi \in \Delta_r^\kappa. \end{aligned} \tag{11}$$

*Proof.* The proof is analogous to the proof of [27, Proposition 10] where the main tool is the commutation principle in EJA developed in [27].  $\square$

**Example 3** (Projections).

- (a) For the EJA  $\mathbb{R}^n$ , problems (10) and (11) are the same. Then, the projection of  $b \in \mathbb{R}^n$  onto the spectraplex of  $\mathbb{R}^n$  is nothing but the projection of  $b$  onto the standard simplex  $\Delta_n$ . That is, the solution to the problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|x - b\|^2 \\ \text{s.t.} \quad & x \geq 0, \langle \mathbf{1}_n, x \rangle = 1. \end{aligned} \tag{12}$$

The solution  $\bar{x}$  of (12) can be characterized as  $\bar{x}_i = \max\{0, b_i + \mu\}$  for all  $i = 1, \dots, n$ , where  $\mu \in \mathbb{R}$  is a solution of  $\sum_{i=1}^n \max\{0, b_i + \mu\} = 1$ . As shown in [7], the point  $\bar{x}$  can be computed efficiently and exactly by specialized algorithms.

- (b) For the Jordan spin algebra  $\mathbb{R}^{n-1} \times \mathbb{R}$ , (11) is formulated in  $\mathbb{R}^2$  (because the rank of the algebra is 2) and the solution can be calculated as follows: for  $b = (\tilde{b}, b_0) \in \mathbb{R}^{n-1} \times \mathbb{R}$ , the solution of (11) is

$$(\bar{\xi}, \bar{t}) = \left( \frac{\tilde{b}}{\max\{1, \|\tilde{b}\|\}}, 1 \right). \tag{13}$$

As a remark for this case, if we work directly (10), we can deduce that  $(\bar{\xi}, \bar{t})$  is the solution of (10) if and only if  $\bar{t} = 1$  and  $\bar{\xi}$  is the orthogonal projection of  $\tilde{b}$  onto the unit ball of  $\mathbb{R}^{n-1}$ . That is,  $\bar{\xi}$  is the solution of

$$\begin{aligned} \min \quad & \frac{1}{2} \|\xi - \tilde{b}\|^2 \\ \text{s.t.} \quad & \|\xi\|^2 \leq 1. \end{aligned}$$

Thus,  $(\bar{\xi}, \bar{t})$  coincides with (13).

(c) For the EJA  $\mathcal{S}^n$ , we have that for  $B \in \mathcal{S}^n$ , problem (10) becomes

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - B\|^2 \\ \text{s.t.} \quad & X \in \mathcal{P}_n, \\ & \text{Tr}(X) = 1. \end{aligned} \tag{14}$$

From Proposition 2, if  $B$  has the spectral decomposition  $B = \sum_{i=1}^n \lambda_i(B) u_i u_i^T$ , where  $\{u_1, \dots, u_n\}$  is an orthonormal basis of  $\mathbb{R}^n$  of eigenvectors of  $B$ , then  $\bar{X} = \sum_{i=1}^n \bar{\xi}_i u_i u_i^T$  is the solution of (14), where  $\bar{\xi}$  is the solution of (12) when  $b = \lambda(B)$ .

## 4 Critical angles between LISC cones

Let  $P, Q \in \mathcal{C}(\mathbb{V})$  be LISC cones represented by  $P = G(K_{\mathbb{X}})$  and  $Q = H(K_{\mathbb{Y}})$ , respectively. With these representations, the cosine of the maximal angle between  $P$  and  $Q$ , formulated in (3), becomes

$$\begin{cases} \min & \langle Gx, Hy \rangle \\ \text{s.t.} & x \in K_{\mathbb{X}}, \|Gx\| = 1, \\ & y \in K_{\mathbb{Y}}, \|Hy\| = 1. \end{cases} \tag{15}$$

**Lemma 3.** *The pair  $(\bar{x}, \bar{y})$  is a stationary point of (15) if and only if*

$$\begin{cases} K_{\mathbb{X}} \ni \bar{x} \perp (G^\top H\bar{y} - \langle G\bar{x}, H\bar{y} \rangle G^\top G\bar{x}) \in K_{\mathbb{X}}, \\ K_{\mathbb{Y}} \ni \bar{y} \perp (H^\top G\bar{x} - \langle G\bar{x}, H\bar{y} \rangle H^\top H\bar{y}) \in K_{\mathbb{Y}}, \\ \|G\bar{x}\| = 1, \|H\bar{y}\| = 1. \end{cases} \tag{16}$$

*Proof.* Observe that for all  $(x, y)$  feasible to problem (15), the gradients of  $\Psi(x, y) = \langle Gx, Hy \rangle$ ,  $\phi_1(x) := \|Gx\| - 1$  and  $\phi_2(y) := \|Hy\| - 1$  are all well defined and given by

$$\nabla \Psi(x, y) = \left( G^\top Hy, H^\top Gx \right), \quad \nabla \phi_1(x) = \frac{G^\top Gx}{\|Gx\|} \quad \text{and} \quad \nabla \phi_2(y) = \frac{H^\top Hy}{\|Hy\|}.$$

It follows from (7) that

$$\begin{aligned} 0 &= \langle \bar{x}, \nabla_x \Psi(\bar{x}, \bar{y}) \rangle + \gamma_1 \langle \bar{x}, \nabla \phi_2(\bar{x}) \rangle = \langle G\bar{x}, H\bar{y} \rangle + \gamma_1 \|G\bar{x}\| = \langle G\bar{x}, H\bar{y} \rangle + \gamma_1 \\ 0 &= \langle \bar{y}, \nabla_y \Psi(\bar{x}, \bar{y}) \rangle + \gamma_2 \langle \bar{y}, \nabla \phi_2(\bar{y}) \rangle = \langle G\bar{x}, H\bar{y} \rangle + \gamma_2 \|H\bar{y}\| = \langle G\bar{x}, H\bar{y} \rangle + \gamma_2 \end{aligned}$$

i.e.,  $\gamma_1 = \gamma_2 = -\langle G\bar{x}, H\bar{y} \rangle$ . The result then follows from (7) by recalling that  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$  are self-dual, then  $K_{\mathbb{X}}^* = K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}^* = K_{\mathbb{Y}}$ . □

**Proposition 3.** *If  $(\bar{x}, \bar{y})$  is a stationary point (resp., solution) of (15), then  $(G\bar{x}, H\bar{y})$  is a critical (resp., antipodal) pair of  $(P, Q)$ . Conversely, if  $(u, v)$  is a critical (resp., antipodal) pair of  $(P, Q)$ , then there exists a stationary point (resp., solution)  $(\bar{x}, \bar{y})$  of (15) such that  $u = G\bar{x}$  and  $v = H\bar{y}$ .*

*Proof.* Let  $(\bar{x}, \bar{y})$  be a stationary point of (15), and set  $u = G\bar{x}$  and  $v = H\bar{y}$ . It follows from (16), and the fact that  $K_{\mathbb{X}}^* = K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}^* = K_{\mathbb{Y}}$ , that

$$\begin{cases} K_{\mathbb{X}} \ni \bar{x} \perp G^{\top}(v - \langle u, v \rangle u) \in K_{\mathbb{X}}^*, \\ K_{\mathbb{Y}} \ni \bar{y} \perp H^{\top}(u - \langle u, v \rangle v) \in K_{\mathbb{Y}}^*, \\ \|u\| = 1, \|v\| = 1. \end{cases}$$

By definition of  $P$  and  $Q$ , the conditions  $\bar{x} \in K_{\mathbb{X}}$  and  $\bar{y} \in K_{\mathbb{Y}}$  imply that  $u = G\bar{x} \in P$  and  $v = H\bar{y} \in Q$ , respectively. These facts and the last line of the above system mean that the first two lines of the definition (4) of critical angles of  $(P, Q)$  are satisfied. Moreover, note that (i)  $\bar{x} \perp G^{\top}(v - \langle u, v \rangle u)$  is equivalent to  $G\bar{x} \perp v - \langle u, v \rangle u$ , and (ii)

$$\begin{aligned} G^{\top}(v - \langle u, v \rangle u) \in K_{\mathbb{X}}^* &\Leftrightarrow \langle G^{\top}(v - \langle u, v \rangle u), x \rangle \geq 0 \quad \forall x \in K_{\mathbb{X}}, \\ &\Leftrightarrow \langle v - \langle u, v \rangle u, Gx \rangle \geq 0 \quad \forall x \in K_{\mathbb{X}}, \\ &\Leftrightarrow \langle v - \langle u, v \rangle u, p \rangle \geq 0 \quad \forall p \in P, \\ &\Leftrightarrow v - \langle u, v \rangle u \in P^*. \end{aligned}$$

Therefore, items (i) and (ii) assert that the first line in the above system is equivalent to  $P \ni u \perp v - \langle u, v \rangle u \in P^*$ . Analogously, the second line in the system reads as  $Q \ni v \perp u - \langle u, v \rangle v \in Q^*$ . Hence,  $(u, v)$  satisfies (4) which means that  $(u, v)$  is a critical pair of  $(P, Q)$ .

Conversely, let  $(u, v)$  be a critical pair of  $(P, Q)$ . Then,  $u \in P = G(K_{\mathbb{X}})$  and  $v \in Q = H(K_{\mathbb{Y}})$  which mean that there exist  $\bar{x} \in K_{\mathbb{X}}$  and  $\bar{y} \in K_{\mathbb{Y}}$  such that  $u = G\bar{x}$  and  $v = H\bar{y}$ . By replacing  $(u, v) = (G\bar{x}, H\bar{y})$  in (4) and by using items (i) and (ii), described above, we deduce that  $(\bar{x}, \bar{y})$  satisfies (16). Thus,  $(\bar{x}, \bar{y})$  is a stationary point of (15).

Now, let  $C_1 \times C_2$  be the feasible set of (15) where  $C_1 := \{x : x \in K_{\mathbb{X}}, \|Gx\| = 1\}$  and  $C_2 := \{y : y \in K_{\mathbb{Y}}, \|Hy\| = 1\}$ . It is straightforward to see that the feasible sets of (3) and (15) are related by

$$\{(Gx, Hy) : (x, y) \in C_1 \times C_2\} = \{(u, v) : u \in P, \|u\| = 1, v \in Q, \|v\| = 1\}. \quad (17)$$

Hence, if  $(\bar{x}, \bar{y})$  is a solution of (15) then  $\langle G\bar{x}, H\bar{y} \rangle \leq \langle Gx, Hy \rangle$  for all  $(x, y) \in C_1 \times C_2$ . From (17) we deduce that  $(\bar{u}, \bar{v}) := (G\bar{x}, H\bar{y})$  is a solution of (3). Then,  $(\bar{u}, \bar{v})$  is an antipodal pair of  $(P, Q)$ . Conversely, if  $(\bar{u}, \bar{v})$  is an antipodal pair of  $(P, Q)$  then there exist  $(\bar{x}, \bar{y}) \in C_1 \times C_2$  such that  $(G\bar{x}, H\bar{y}) = (\bar{u}, \bar{v})$  satisfying  $\langle G\bar{x}, H\bar{y} \rangle \leq \langle u, v \rangle$  for all feasible point  $(u, v)$  of (3). From (17) we conclude that  $(\bar{x}, \bar{y})$  is a solution of (15).  $\square$

The numerical resolution of problem (15) is not an easy task since it has as feasible set

$$\{x \in K_{\mathbb{X}} : \|Gx\| = 1\} \times \{y \in K_{\mathbb{Y}} : \|Hy\| = 1\},$$

which is not convex due to its spherical conditions. Thus, we need to reformulate (15) in such a way that its feasible set can be more conveniently manipulable. Our reformulation will be

a minimization problem with feasible set  $\Delta_{\mathbb{X},\mathbb{Y}} = \Delta_{\mathbb{X}} \times \Delta_{\mathbb{Y}}$ . Recall that  $\Delta_{\mathbb{X}}$  and  $\Delta_{\mathbb{Y}}$  are the spectraplexes of  $\mathbb{X}$  and  $\mathbb{Y}$  which are given respectively by

$$\Delta_{\mathbb{X}} = \{x \in K_{\mathbb{X}} : \langle e_{\mathbb{X}}, x \rangle = 1\} \quad \text{and} \quad \Delta_{\mathbb{Y}} = \{y \in K_{\mathbb{Y}} : \langle e_{\mathbb{Y}}, y \rangle = 1\}.$$

The set  $\Delta_{\mathbb{X},\mathbb{Y}}$  turns out to be convex since we have hyperplanes instead of spheres. As  $Gx \neq 0$  and  $Hy \neq 0$  for all  $(x, y) \in \Delta_{\mathbb{X},\mathbb{Y}}$ , due to Lemma 2, Proposition 3 asserts that problem (15) can be reformulated as (5), fitting thus into the structure of (6). We have that  $\Phi$  (in (5)) is differentiable in  $\Delta_{\mathbb{X},\mathbb{Y}}$ . Its gradient is given  $\nabla\Phi(x, y) = (\nabla_x\Phi(x, y), \nabla_y\Phi(x, y))$ , where

$$\nabla_x\Phi(x, y) = \frac{G^\top Hy}{\|Gx\|\|Hy\|} - \frac{\langle Gx, Hy \rangle}{\|Gx\|\|Hy\|} \frac{G^\top Gx}{\|Gx\|^2}, \quad (18a)$$

$$\nabla_y\Phi(x, y) = \frac{H^\top Gx}{\|Gx\|\|Hy\|} - \frac{\langle Gx, Hy \rangle}{\|Gx\|\|Hy\|} \frac{H^\top Hy}{\|Hy\|^2}. \quad (18b)$$

**Lemma 4.** *The pair  $(\bar{x}, \bar{y})$  is a stationary point of (5) if and only if*

$$\begin{cases} K_{\mathbb{X}} \ni \bar{x} \perp \left( \frac{G^\top H\bar{y}}{\|H\bar{y}\|} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|\|H\bar{y}\|} \frac{G^\top G\bar{x}}{\|G\bar{x}\|^2} \right) \in K_{\mathbb{X}}, \\ K_{\mathbb{Y}} \ni \bar{y} \perp \left( \frac{H^\top G\bar{x}}{\|G\bar{x}\|} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|\|H\bar{y}\|} \frac{H^\top H\bar{y}}{\|H\bar{y}\|^2} \right) \in K_{\mathbb{Y}}, \\ \langle e_{\mathbb{X}}, \bar{x} \rangle = 1, \langle e_{\mathbb{Y}}, \bar{y} \rangle = 1. \end{cases} \quad (19)$$

*Proof.* It follows from Definition 2 and the equations in (18) that  $(\bar{x}, \bar{y})$  is a stationary point of (5) if, and only if, there exist  $\gamma_1, \gamma_2 \in \mathbb{R}$  such that

$$K_{\mathbb{X}} \ni \bar{x} \perp \left( \frac{G^\top H\bar{y}}{\|G\bar{x}\|\|H\bar{y}\|} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|\|H\bar{y}\|} \frac{G^\top G\bar{x}}{\|G\bar{x}\|^2} + \gamma_1 e_{\mathbb{X}} \right) \in K_{\mathbb{X}}^*, \quad (20)$$

$$K_{\mathbb{Y}} \ni \bar{y} \perp \left( \frac{H^\top G\bar{x}}{\|G\bar{x}\|\|H\bar{y}\|} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|\|H\bar{y}\|} \frac{H^\top H\bar{y}}{\|H\bar{y}\|^2} + \gamma_2 e_{\mathbb{Y}} \right) \in K_{\mathbb{Y}}^*, \quad (21)$$

$$\langle e_{\mathbb{X}}, \bar{x} \rangle = 1, \langle e_{\mathbb{Y}}, \bar{y} \rangle = 1. \quad (22)$$

Next we show that  $\gamma_1 = \gamma_2 = 0$ . From the orthogonality conditions (20) and (21) we have that  $\gamma_1 \langle e_{\mathbb{X}}, \bar{x} \rangle = 0$  and  $\gamma_2 \langle e_{\mathbb{Y}}, \bar{y} \rangle = 0$ , respectively. Thus, because of (22) we deduce  $\gamma_1 = \gamma_2 = 0$ . Finally, the system (20)-(22) becomes (19) by recalling that  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$  are self-dual, and by removing w.l.o.g.  $1/\|G\bar{x}\|$  from equation (20) and  $1/\|H\bar{y}\|$  from equation (21).  $\square$

**Theorem 2.** *If  $(\bar{x}, \bar{y})$  is a stationary point (resp., solution) of the problem (15), then  $(\tilde{x}, \tilde{y}) := \left( \frac{\bar{x}}{\langle e_{\mathbb{X}}, \bar{x} \rangle}, \frac{\bar{y}}{\langle e_{\mathbb{Y}}, \bar{y} \rangle} \right)$  is a stationary point (resp., solution) of (5). Conversely, if  $(\tilde{x}, \tilde{y})$  is a stationary point (resp., solution) of (5), then  $(\bar{x}, \bar{y}) := \left( \frac{\tilde{x}}{\|G\tilde{x}\|}, \frac{\tilde{y}}{\|H\tilde{y}\|} \right)$  is a stationary point (resp., solution) of (15).*

*Proof.* Let  $(\bar{x}, \bar{y})$  be a stationary point of (15). Then (16) holds. From Lemma 1 we have that  $\langle e_{\mathbb{X}}, \bar{x} \rangle > 0$  and  $\langle e_{\mathbb{Y}}, \bar{y} \rangle > 0$ . Thus, the pair  $(\tilde{x}, \tilde{y}) := \left( \frac{\bar{x}}{\langle e_{\mathbb{X}}, \bar{x} \rangle}, \frac{\bar{y}}{\langle e_{\mathbb{Y}}, \bar{y} \rangle} \right)$  is well-defined. Observe that because of the unit norm conditions in (16) we get  $\|G\tilde{x}\| = \langle e_{\mathbb{X}}, \bar{x} \rangle^{-1}$  and  $\|H\tilde{y}\| = \langle e_{\mathbb{Y}}, \bar{y} \rangle^{-1}$ .

Then,  $\bar{x} = \frac{\tilde{x}}{\|G\tilde{x}\|}$  and  $\bar{y} = \frac{\tilde{y}}{\|H\tilde{y}\|}$ . By replacing these values in (16) we deduce that  $(\tilde{x}, \tilde{y})$  satisfies (19). Therefore,  $(\tilde{x}, \tilde{y})$  is a stationary point of (5). The converse is proved analogously. In this case, when  $(\tilde{x}, \tilde{y})$  is a stationary point of (5),  $G\tilde{x} \neq 0$  and  $H\tilde{y} \neq 0$  thanks to Lemma 2. Then,  $(\bar{x}, \bar{y}) := \left( \frac{\tilde{x}}{\|G\tilde{x}\|}, \frac{\tilde{y}}{\|H\tilde{y}\|} \right)$  is well-defined, and it turns out to be a stationary point of (15). The details are omitted.

Now, suppose that  $(\bar{x}, \bar{y})$  is a solution of (15). Let  $(x, y)$  be any feasible point of (5). Then,  $\left( \frac{x}{\|Gx\|}, \frac{y}{\|Hy\|} \right)$  is well-defined and it is a feasible point of (15). Therefore,

$$\langle G\bar{x}, H\bar{y} \rangle \leq \Phi(x, y),$$

for all feasible points  $(x, y)$  of (5). In particular,  $(\tilde{x}, \tilde{y}) := \left( \frac{\bar{x}}{\langle e_{\mathbb{X}}, \bar{x} \rangle}, \frac{\bar{y}}{\langle e_{\mathbb{Y}}, \bar{y} \rangle} \right)$  is a feasible point of (5) and it attains the equality in the above displayed inequality. Thus  $(\tilde{x}, \tilde{y})$  is a solution of (5). The converse is analogous.  $\square$

The connection between (15) and the problem of computing a critical angle between the cones  $P = G(K_{\mathbb{X}})$  and  $Q = H(K_{\mathbb{Y}})$  are made explicit in the following corollary.

**Corollary 2.** *If  $(\tilde{x}, \tilde{y})$  is a stationary point (resp., solution) of (5), then  $(u, v) := \left( \frac{G\tilde{x}}{\|G\tilde{x}\|}, \frac{H\tilde{y}}{\|H\tilde{y}\|} \right)$  is a critical (resp., antipodal) pair of  $(P, Q)$ . Conversely, if  $(u, v)$  is a critical (resp., antipodal) pair of  $(P, Q)$ , then there exists a stationary point  $(\tilde{x}, \tilde{y})$  of (5) such that  $(u, v)$  is given as above.*

*Proof.* Combine Theorem 2 with Proposition 3.  $\square$

Hence, computing a critical angle between two cones  $P = G(K_{\mathbb{X}})$  and  $Q = H(K_{\mathbb{Y}})$  amounts to computing a stationary point of the fractional programming problem (5). The next section is dedicated to this task.

## 5 A sequential regularized partial linearization algorithm

In this section, we develop a numerical method to compute critical angles between a pair of LISC cones by computing stationary points of the fractional program (5). By using the well-known approach introduced by Dinkelbach in [9], we consider the following parametric program, for  $\delta \in \mathbb{R}$ :

$$\text{BCP}_{\delta} : \begin{cases} \min & f_{\delta}(x, y) := \langle Gx, Hy \rangle - \delta \|Gx\| \|Hy\| \\ \text{s.t.} & x \in K_{\mathbb{X}}, \langle e_{\mathbb{X}}, x \rangle = 1, \\ & y \in K_{\mathbb{Y}}, \langle e_{\mathbb{Y}}, y \rangle = 1. \end{cases} \quad (23)$$

*Remark 1.* We use the notation  $\text{BCP}_{\delta}$  since that program is constituted by biconvex functions (see [16] for definitions and properties of biconvexity). Indeed,  $\mathcal{B}(x, y) := \langle Gx, Hy \rangle$  is biconvex because it is bilinear. Furthermore,  $\mathcal{D}(x, y) := \|Gx\| \|Hy\|$  is biconvex because its function at each variable is a multiple of the composition of the cartesian norm with a linear map which is a convex function.

Observe that  $f_\delta$  is differentiable at any  $(x, y)$  such that  $Gx \neq 0$  and  $Hy \neq 0$ . As discussed in the previous section, those conditions hold for every  $(x, y) \in \Delta_{\mathbb{X}, \mathbb{Y}}$ . Thus,  $f_\delta$  is differentiable over the set  $\Delta_{\mathbb{X}, \mathbb{Y}}$  of feasible points of  $\text{BCP}_\delta$ , and the gradient is given by  $\nabla f_\delta(x, y) = (\nabla_x f_\delta(x, y), \nabla_y f_\delta(x, y))$ , where

$$\nabla_x f_\delta(x, y) = G^\top Hy - \delta \|Gx\|^{-1} \|Hy\| G^\top Gx, \quad (24a)$$

$$\nabla_y f_\delta(x, y) = H^\top Gx - \delta \|Hy\|^{-1} \|Gx\| H^\top Hy. \quad (24b)$$

The next theorem states that a stationary point of (5) can be computed as a stationary point of  $\text{BCP}_\delta$  where  $f_\delta$  vanishes.

**Theorem 3.** *The pair  $(\bar{x}, \bar{y})$  is a stationary point (resp., solution) of  $\text{BCP}_\delta$  (23) with parameter  $\delta := \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$  if and only if  $(\bar{x}, \bar{y})$  is a stationary point (resp., solution) of FP (5).*

*Proof.* Let  $(\bar{x}, \bar{y})$  be a stationary point of  $\text{BCP}_\delta$  (23). From Definition 2 and because  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$  are self-dual, we get

$$\begin{cases} K_{\mathbb{X}} \ni \bar{x} \perp (G^\top H\bar{y} - \delta \|G\bar{x}\|^{-1} \|H\bar{y}\| G^\top G\bar{x} + \gamma_1 e_{\mathbb{X}}) \in K_{\mathbb{X}}, \\ K_{\mathbb{Y}} \ni \bar{y} \perp (H^\top G\bar{x} - \delta \|H\bar{y}\|^{-1} \|G\bar{x}\| H^\top H\bar{y} + \gamma_2 e_{\mathbb{Y}}) \in K_{\mathbb{Y}}, \\ \langle e_{\mathbb{X}}, \bar{x} \rangle = 1, \langle e_{\mathbb{Y}}, \bar{y} \rangle = 1. \end{cases}$$

From the orthogonality conditions above we deduce that  $0 = \langle G\bar{x}, H\bar{y} \rangle - \delta \|H\bar{y}\| \|G\bar{x}\| + \gamma_1$  and  $0 = \langle G\bar{x}, H\bar{y} \rangle - \delta \|H\bar{y}\| \|G\bar{x}\| + \gamma_2$ . The definition of  $\delta$  then implies that  $\gamma_1 = \gamma_2 = 0$ . The above system then becomes (with  $\delta = \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$ )

$$\begin{cases} K_{\mathbb{X}} \ni \bar{x} \perp \left( G^\top H\bar{y} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|^2} G^\top G\bar{x} \right) \in K_{\mathbb{X}}, \\ K_{\mathbb{Y}} \ni \bar{y} \perp \left( H^\top G\bar{x} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|H\bar{y}\|^2} H^\top H\bar{y} \right) \in K_{\mathbb{Y}}, \\ \langle e_{\mathbb{X}}, \bar{x} \rangle = 1, \langle e_{\mathbb{Y}}, \bar{y} \rangle = 1. \end{cases}$$

As  $\bar{x} \perp \left( G^\top H\bar{y} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|^2} G^\top G\bar{x} \right) \in K_{\mathbb{X}}$ , it is equivalent to

$$\bar{x} \perp \frac{1}{\|H\bar{y}\|} \left( G^\top H\bar{y} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|^2} G^\top G\bar{x} \right) \in K_{\mathbb{X}}.$$

Analogously, we have  $\bar{y} \perp \frac{1}{\|G\bar{x}\|} \left( H^\top G\bar{x} - \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|H\bar{y}\|^2} H^\top H\bar{y} \right) \in K_{\mathbb{Y}}$ . Then, we can deduce that  $(\bar{x}, \bar{y})$  is a stationary point of  $\text{BCP}_\delta$  (23) with  $\delta = \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$  if and only if  $(\bar{x}, \bar{y})$  satisfies (19). Thus, the equivalence follows. Suppose that  $(\bar{x}, \bar{y})$  is a solution of  $\text{BCP}_\delta$  (23) with  $\delta = \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$ . Then,  $f_\delta(\bar{x}, \bar{y}) = 0$ . Furthermore, for any feasible point  $(x, y)$  of  $\text{BCP}_\delta$  (23) we get  $f_\delta(x, y) \geq 0$ , which means

$$\Phi(\bar{x}, \bar{y}) = \delta \leq \Phi(x, y). \quad (25)$$

Then, since problems  $\text{BCP}_\delta$  (23) and FP (5) have the same feasible sets, we conclude from (25) that  $(\bar{x}, \bar{y})$  is a solution of FP (5). The converse follows the same steps but in reverse order.  $\square$

**Corollary 3.** *If  $(\bar{x}, \bar{y})$  is a stationary point (resp., solution) of  $\text{BCP}_\delta$  (23) with parameter  $\delta := \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$ , then  $(u, v) := (\|G\bar{x}\|^{-1}G\bar{x}, \|H\bar{y}\|^{-1}H\bar{y})$  is a critical (resp., antipodal) pair of  $(P, Q)$ . Conversely, if  $(u, v)$  is a critical (resp., antipodal) pair of  $(P, Q)$ , then there exists a stationary point  $(\bar{x}, \bar{y})$  of  $\text{BCP}_\delta$  (23) with  $\delta := \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$  and such that  $(u, v) = (\|G\bar{x}\|^{-1}G\bar{x}, \|H\bar{y}\|^{-1}H\bar{y})$ .*

*Proof.* It follows directly from Theorem 3 and Corollary 2.  $\square$

Next, we formulate our algorithm to compute critical angles between a pair of LISC cones  $(P, Q) = (G(K_{\mathbb{X}}), H(K_{\mathbb{Y}}))$ .

---

**Algorithm 1** Sequential Regularized Partial Linearization

---

▷ **Step 0**

- 1: Choose  $(x^0, y^0) \in \Delta_{\mathbb{X}, \mathbb{Y}}$ ,  $\beta > 0$ ,  $0 < \alpha < 1$ ,  $0 < \rho < 1$ , and prox-parameters  $\mu_1, \mu_2 \geq 0$ .
- 2: Set  $k := 0$ .

▷ **Step 1**

- 3: Set

$$\delta_k := \frac{\langle Gx^k, Hy^k \rangle}{\|Gx^k\| \|Hy^k\|}.$$

▷ **Step 2**

- 4: Let  $L_1^k(x) := \langle Gx, Hy^k - \delta_k \|Gx^k\|^{-1} \|Hy^k\| Gx^k \rangle$ .  
Compute a solution  $\tilde{x}^k$  to the convex program

$$\begin{aligned} \min \quad & L_1^k(x) + \frac{\mu_1}{2} \|x - x^k\|^2 \\ \text{s.t.} \quad & x \in K_{\mathbb{X}}, \langle e_{\mathbb{X}}, x \rangle = 1. \end{aligned} \tag{26}$$

- 5: Let  $L_2^k(y) := \langle Hy, Gx^k - \delta_k \|Gx^k\| \|Hy^k\|^{-1} Hy^k \rangle$ .  
Compute a solution  $\tilde{y}^k$  to the convex program

$$\begin{aligned} \min \quad & L_2^k(y) + \frac{\mu_2}{2} \|y - y^k\|^2 \\ \text{s.t.} \quad & y \in K_{\mathbb{Y}}, \langle e_{\mathbb{Y}}, y \rangle = 1. \end{aligned} \tag{27}$$

- 6: Let  $d_1^k := \tilde{x}^k - x^k$  and  $d_2^k := \tilde{y}^k - y^k$ .

▷ **Step 3**

- 7: If  $L_1^k(d_1^k) = 0$  and  $L_2^k(d_2^k) = 0$ , terminate.

Otherwise, let  $t_k := \beta \rho^{\ell_k}$ , where  $\ell_k$  is the smallest nonnegative integer  $\ell$  such that

$$\Phi(x^k + t^k d_1^k, y^k + t^k d_2^k) \leq \Phi(x^k, y^k) + \alpha t_k \frac{L_1^k(d_1^k) + L_2^k(d_2^k)}{\|Gx^k\| \|Hy^k\|}$$

Set  $(x^{k+1}, y^{k+1}) := (x^k, y^k) + t_k(d_1^k, d_2^k)$ . Go to Step 1.

---

Some comments on Algorithm 1 are in order.

- i) The prox-parameters  $\mu_1$  and  $\mu_2$  can take arbitrary non-negative values. The choice  $\mu_1 = \mu_2 = 0$  is covered by the convergence analysis given below because  $\Delta_{\mathbb{X}, \mathbb{Y}}$  of (5) is compact (see Proposition 1). However, our numerical experiments suggest that

choosing the prox-parameters to be strictly positive is advantageous: it regularizes the iterative process in the sense that the algorithm performs fewer iterations to compute a stationary point within a given tolerance.

- ii) We highlight that Step 2 can be performed in parallel as the two subproblems are independent of each other.
- iii) For LISC cones, the interest of dealing with (5) instead of (3) is now evident: the cones  $K_{\mathbb{X}}$  and  $K_{\mathbb{Y}}$  are in general more structured, permitting (26) and (27) to have straightforward solutions in many settings of interest (c.f. Subsection 3.2). If the considered cones are not LISC, then  $G$  and  $H$  are the identities operators in the corresponding spaces  $\mathbb{X}$  and  $\mathbb{Y}$ , and thus  $P = K_{\mathbb{X}}$  and  $Q = K_{\mathbb{Y}}$ . Algorithm 1 is still applicable: for applications in which the subproblems do not have explicit solutions, we can use of off-the-shelf (linear, quadratic, conic, etc.) solvers to compute  $\tilde{x}^k$  and  $\tilde{y}^k$ .
- iv) In Step 3, we perform the standard Armijo line search rule. Indeed, the term  $\frac{L_1^k(d_1^k) + L_2^k(d_2^k)}{\|Gx^k\| \|Hy^k\|}$  coincides with  $\langle \nabla \Phi(x^k, y^k), d^k \rangle$ . As we show in Proposition 4 below,  $d^k$  is a descend direction for  $\Phi$  at  $(x^k, y^k)$ .

**Lemma 5.** *We have that*

$$L_1^k(d_1^k) \leq -\frac{\mu_1}{2} \|d_1^k\|^2 \quad \text{and} \quad L_2^k(d_2^k) \leq -\frac{\mu_2}{2} \|d_2^k\|^2. \quad (28)$$

Furthermore, if  $x^k$  is not a solution to (26) or  $y^k$  is not a solution to (27), then

$$L_1^k(d_1^k) + L_2^k(d_2^k) < 0. \quad (29)$$

*Proof.* Since  $\tilde{x}^k$  solves (26) and  $x^k$  is feasible, we have that  $L_1^k(\tilde{x}^k) + \frac{\mu_1}{2} \|\tilde{x}^k - x^k\|^2 \leq L_1^k(x^k)$ , with strict inequality if  $x^k$  is not a solution to (26). As  $L_1^k(d_1^k) = L_1^k(\tilde{x}^k) - L_1^k(x^k)$ , we conclude that  $L_1^k(d_1^k) \leq -\frac{\mu_1}{2} \|d_1^k\|^2$ . Analogously, we get  $L_2^k(d_2^k) \leq -\frac{\mu_2}{2} \|d_2^k\|^2$ . Suppose that  $x^k$  is not a solution to (26) or  $y^k$  is not a solution to (27). Then, one of the inequalities in (28) holds strictly. It implies,  $L_1^k(d_1^k) + L_2^k(d_2^k) < -\frac{\mu_1}{2} \|d_1^k\|^2 - \frac{\mu_2}{2} \|d_2^k\|^2 \leq 0$ .  $\square$

**Proposition 4.** *Suppose that  $x^k$  is not a solution to (26) or  $y^k$  is not a solution to (27). Then, the vector  $(d_1^k, d_2^k)$  is a descent direction for the fractional function  $\Phi$  (given in (5)) at  $(x^k, y^k)$ .*

*Proof.* We must prove that

$$\langle \nabla \Phi(x^k, y^k), (d_1^k, d_2^k) \rangle = \langle \nabla_x \Phi(x^k, y^k), d_1^k \rangle + \langle \nabla_y \Phi(x^k, y^k), d_2^k \rangle < 0. \quad (30)$$

From (18), we obtain,

$$\begin{aligned} \nabla_x \Phi(x^k, y^k) &= \frac{G^\top Hy^k}{\|Gx^k\| \|Hy^k\|} - \delta_k \frac{G^\top Gx^k}{\|Gx^k\|^2}, \\ \nabla_y \Phi(x^k, y^k) &= \frac{H^\top Gx^k}{\|Gx^k\| \|Hy^k\|} - \delta_k \frac{H^\top Hy^k}{\|Hy^k\|^2} \end{aligned}$$

Then,

$$\begin{aligned}\langle \nabla_x \Phi(x^k, y^k), d_1^k \rangle &= \|Gx^k\|^{-1} \|Hy^k\|^{-1} L_1^k(d_1^k), \\ \langle \nabla_y \Phi(x^k, y^k), d_2^k \rangle &= \|Gx^k\|^{-1} \|Hy^k\|^{-1} L_2^k(d_2^k).\end{aligned}$$

Since  $x^k$  is not a solution to (26) or  $y^k$  is not a solution to (27), we can use (29) to conclude (30).  $\square$

The following proposition explains why the stop criterium of Algorithm 1 allow us to find solutions for (26) and (27).

**Proposition 5.** *If  $L_1^k(d_1^k) = 0$ , then  $x^k$  solves (26). Moreover, If  $L_2^k(d_2^k) = 0$ , then  $y^k$  solves (27).*

*Proof.* Suppose that  $L_1^k(d_1^k) = 0$ . Then,  $L_1^k(\tilde{x}^k) - L_1^k(x^k) = 0$ . If  $\mu_1 = 0$ , the regularization term in (26) vanishes. Hence, because of  $L_1^k(\tilde{x}^k) = L_1^k(x^k)$  we conclude that  $x^k$  solves (26). If  $\mu_1 > 0$ , from (28) we deduce that  $d_1^k = \tilde{x}^k - x^k = 0$ . Then,  $x^k = \tilde{x}^k$  which implies that  $x^k$  solves (26). The proof of the second part is analogous.  $\square$

Now, the following two results show that Algorithm 1 find stationary points of problem (5).

**Proposition 6.** *Let  $(\bar{x}, \bar{y}) \in \Delta_{\mathbb{X}, \mathbb{Y}}$  and set  $\bar{\delta} := \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\| \|H\bar{y}\|}$ . Then,  $(\bar{x}, \bar{y})$  is a stationary point of (5) if, and only if,  $\bar{x}$  and  $\bar{y}$  are solutions of the respective programs:*

$$\begin{aligned}\min \quad & \langle Gx, H\bar{y} - \bar{\delta} \|G\bar{x}\|^{-1} \|H\bar{y}\| G\bar{x} \rangle + \frac{\mu_1}{2} \|x - \bar{x}\|^2 \\ \text{s.t.} \quad & x \in K_{\mathbb{X}}, \langle e_{\mathbb{X}}, x \rangle = 1,\end{aligned}\tag{31}$$

and

$$\begin{aligned}\min \quad & \langle Hy, G\bar{x} - \bar{\delta} \|G\bar{x}\| \|H\bar{y}\|^{-1} H\bar{y} \rangle + \frac{\mu_2}{2} \|y - \bar{y}\|^2 \\ \text{s.t.} \quad & y \in K_{\mathbb{Y}}, \langle e_{\mathbb{Y}}, y \rangle = 1.\end{aligned}\tag{32}$$

*Proof.* The pair  $(\bar{x}, \bar{y})$  is a stationary point of (5) if, and only if,  $\langle \nabla f_{\bar{\delta}}(\bar{x}, \bar{y}), ((x, y) - (\bar{x}, \bar{y})) \rangle \geq 0$  for all  $(x, y) \in \Delta_{\mathbb{X}, \mathbb{Y}}$ , which is equivalent to

$$\begin{cases} \langle \nabla_x f_{\bar{\delta}}(\bar{x}, \bar{y}), x - \bar{x} \rangle \geq 0 & \forall x \in K_{\mathbb{X}} \text{ s.t. } \langle e_{\mathbb{X}}, x \rangle = 1 \\ \langle \nabla_y f_{\bar{\delta}}(\bar{x}, \bar{y}), y - \bar{y} \rangle \geq 0 & \forall y \in K_{\mathbb{Y}} \text{ s.t. } \langle e_{\mathbb{Y}}, y \rangle = 1.\end{cases}$$

In view of (24a) and (24b), the above system is equivalent to saying that  $\bar{x}$  and  $\bar{y}$  are solutions to (31) and (32), respectively.  $\square$

**Theorem 4.** *Every accumulation point  $(\bar{x}, \bar{y})$  of the sequence  $\{(x^k, y^k)\}$ , generated by Algorithm 1, is a stationary point of problem (5).*

*Proof.* Let  $(\bar{x}, \bar{y})$  be an accumulation point of  $\{(x^k, y^k)\}$  (which is assured to exist because  $\Delta_{\mathbb{X}, \mathbb{Y}}$  is compact, see Proposition 1). By passing to a subsequence if necessary, we may suppose that  $(x^k, y^k) \rightarrow (\bar{x}, \bar{y})$  and  $(\tilde{x}^k, \tilde{y}^k) \rightarrow (\tilde{x}, \tilde{y})$ , as  $k \rightarrow \infty$ , where  $(\tilde{x}, \tilde{y}) \in \Delta_{\mathbb{X}, \mathbb{Y}}$ . Let  $x \in K_{\mathbb{X}}$  s.t.  $\langle e_{\mathbb{X}}, x \rangle = 1$  be arbitrary. Because  $\tilde{x}^k$  solves (26), we have that

$$\begin{aligned}& \left\langle G\tilde{x}^k, Hy^k - \delta_k \|Gx^k\|^{-1} \|Hy^k\| Gx^k \right\rangle + \frac{\mu_1}{2} \|\tilde{x}^k - x^k\|^2 \\ & \leq \left\langle Gx, Hy^k - \delta_k \|Gx^k\|^{-1} \|Hy^k\| Gx^k \right\rangle + \frac{\mu_1}{2} \|x - x^k\|^2.\end{aligned}\tag{33}$$

By letting  $k \rightarrow \infty$  on (33), we obtain

$$\begin{aligned} & \langle G\tilde{x}, H\bar{y} - \bar{\delta}\|G\bar{x}\|^{-1}\|H\bar{y}\|G\bar{x} \rangle + \frac{\mu_1}{2}\|\tilde{x} - \bar{x}\|^2 \\ & \leq \langle Gx, H\bar{y} - \bar{\delta}\|G\bar{x}\|^{-1}\|H\bar{y}\|G\bar{x} \rangle + \frac{\mu_1}{2}\|x - \bar{x}\|^2. \end{aligned}$$

where

$$\bar{\delta} := \frac{\langle G\bar{x}, H\bar{y} \rangle}{\|G\bar{x}\|\|H\bar{y}\|}.$$

Thus, we deduce that  $\tilde{x}$  is a solution to (31). On the other hand, from standard arguments on the Armijo rule we have that  $\langle \nabla\Phi(x^k, y^k), (d_1^k, d_2^k) \rangle \rightarrow 0$  as  $k \rightarrow \infty$ . In Proposition 4, we have shown that  $\langle \nabla_x\Phi(x^k, y^k), d_1^k \rangle \leq 0$  and  $\langle \nabla_y\Phi(x^k, y^k), d_2^k \rangle \leq 0$ . Then,

$$\lim_{k \rightarrow \infty} \langle \nabla_x\Phi(x^k, y^k), d_1^k \rangle = \langle \nabla_x\Phi(\bar{x}, \bar{y}), \tilde{x} - \bar{x} \rangle = 0, \quad (34)$$

$$\lim_{k \rightarrow \infty} \langle \nabla_y\Phi(x^k, y^k), d_2^k \rangle = \langle \nabla_y\Phi(\bar{x}, \bar{y}), \tilde{y} - \bar{y} \rangle = 0. \quad (35)$$

Since  $\langle \nabla_x\Phi(\bar{x}, \bar{y}), \tilde{x} - \bar{x} \rangle = \|G\bar{x}\|^{-1}\|H\bar{y}\|^{-1}\bar{L}_1(\tilde{x} - \bar{x})$ , from (34) we conclude

$$\bar{L}_1(\tilde{x}) - \bar{L}_1(\bar{x}) = \bar{L}_1(\tilde{x} - \bar{x}) = 0.$$

Thus, from Proposition 5 we conclude that not only  $\tilde{x}$  but also  $\bar{x}$  is solution of (31). In the same way, we can prove that both  $\tilde{y}$  and  $\bar{y}$  are solutions of (32). (We highlight that if  $\mu_1, \mu_2 > 0$  in Algorithm 1, then  $\bar{x} = \tilde{x}$  and  $\bar{y} = \tilde{y}$  due to strongly convexity of the objective functions.) Therefore, from Proposition 6 we conclude that  $(\bar{x}, \bar{y})$  is a stationary point of (5).  $\square$

Finally, it follows from Theorem 4 and Corollary 3 that Algorithm 1 asymptotically computes a critical pair  $(u, v) = (\|G\bar{x}\|^{-1}G\bar{x}, \|H\bar{y}\|^{-1}H\bar{y})$  of  $(P, Q)$ .

## 6 Computational experiments

This section reports some numerical experiments for computing critical angles between pairs of LISC cones. For each class of test problems considered below, we describe how the subproblems (26) and (27) take a form that can be efficiently solved.

Indeed, for the pair of LISC cones  $(P, Q) = (G(K_{\mathbb{X}}), H(K_{\mathbb{Y}}))$ , the optimization problems in Step 2 of Algorithm 1 are equivalent to the projection problems onto

$$\Delta_{\mathbb{X}} = \{x \in K_{\mathbb{X}} : \langle e_{\mathbb{X}}, x \rangle = 1\} \quad \text{and} \quad \Delta_{\mathbb{Y}} = \{y \in K_{\mathbb{Y}} : \langle e_{\mathbb{Y}}, y \rangle = 1\},$$

when  $\mu_1 > 0$  and  $\mu_2 > 0$ , respectively (if  $\mu_1 = \mu_2 = 0$ , then subproblems (26) and (27) have straightforward solutions). Let us focus on problem (26), a similar reasoning applies to (27). After some algebraic manipulation on the cost function of (26), we can obtain the solution of (26) by solving the following projection problem:

$$\begin{aligned} \min & \quad \frac{1}{2} \left\| x - \left( x^k - \frac{c_k}{\mu_1} \right) \right\|^2 \\ \text{s.t.} & \quad x \in \Delta_{\mathbb{X}} \end{aligned} \quad \text{with} \quad c_k := G^T(Hy^k - \delta_k\|Gx^k\|^{-1}\|Hy^k\|Gx^k), \quad (36)$$

As benchmark for our analysis, we solved the nonlinear program (5) by using IPOPT [40] distributed in the OPTI toolbox [8]. We use default values for the solver parameters, except for the maximum number of iterations, which is incremented to 5000 and the Absolute Function Tolerance increased to  $10^{-6}$ . The numerical experiments were performed on an Intel Core i7 clocked at 2 GHz (32 GB RAM).

The performance of the two solvers is measured by considering  $10^3$  initial points and by reporting in the tables below the following values:

- $\theta_b(P, Q)$ : the best critical angle, which we compute as the cosine inverse of the best objective function value of (5) obtained by one solver with  $10^3$  initial points.
- $it_b, it_a, it_w$ : minimum, average and maximum number of iterations required by the algorithm to compute a stationary point.
- $CPU_b, CPU_a, CPU_w$ : minimum, average and maximum CPU time in seconds required by the algorithm to compute a stationary point.

In practice, we stop Algorithm 1 when  $|L_1^k(d_1^k)| \leq \epsilon_1$  and  $|L_2^k(d_2^k)| \leq \epsilon_2$ , with  $\epsilon_1, \epsilon_2$  being small positive values. Additionally, we require that the  $\delta_k - \delta_{k+5} \leq \epsilon_3$  for a given small  $\epsilon_3 > 0$ , that is, the algorithm runs until the objective function value of (5) stabilizes over the last five iterations. We also set the maximum number of iterations to 5000.

## 6.1 Critical angles between two polyhedral cones

As in [32], we consider the computation of critical angles of a pair

$$(P, Q) = (G(\mathbb{R}_+^p), H(\mathbb{R}_+^q))$$

of polyhedral cones in  $\mathbb{R}^n$ , where  $G$  and  $H$  are matrices in  $\mathbb{R}^{n \times p}$  and  $\mathbb{R}^{n \times q}$ , respectively, such that  $P$  and  $Q$  are LISC cones. For  $\mu_1, \mu_2 > 0$  and because of (36), we can obtain the solution of (26) by solving the following projection onto the simplex problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \left\| x - \left( x^k - \frac{c_k}{\mu_1} \right) \right\|^2 \\ \text{s.t.} \quad & x \geq 0, \\ & \langle \mathbf{1}_p, x \rangle = 1. \end{aligned}$$

The resolution of this task is explained in Example 3 item (a). It can be efficiently performed by the specialized algorithm of [7].

We compute critical angles between the nonnegative orthant  $P = \mathbb{R}_+^n$  and the Schur cone

$$Q = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^k x_i \geq 0, \text{ for } k \in \{1, \dots, n-1\} \text{ and } x_1 + \dots + x_n = 0 \right\}.$$

We consider  $G = I_n$  and the columns of the matrix  $H$  are in the following form

$$h_1 = \frac{1}{\sqrt{2}}, (1, -1, 0, \dots, 0)^T, \dots, h_{n-1} = \frac{1}{\sqrt{2}}, (0, \dots, 0, 1, -1)^T.$$

We generate several instances by considering  $n \in \{5, 20, 50, 100, 500, 700, 1000\}$  and run Algorithm 1 and IPOPT with  $10^3$  random initial points in the simplex. For these test problems,

we stop Algorithm 1 when the stopping criteria are satisfied with  $\epsilon_1 = 10^{-6}$ ,  $\epsilon_2 = 10^{-6}$  and  $\epsilon_3 = 10^{-5}$ . After some tuning, the prox-parameter values were chosen as  $\mu_1 = 0.01$  and  $\mu_2 = 2.6$ .

The performance of Algorithm 1 and IPOPT for solving these test problems is presented in Tables 1 and 2, respectively. This first set of numerical experiments demonstrates that, unlike IPOPT, the proposed algorithm can calculate a critical angle within the maximum number of iterations for all the instances tested. As expected, the required CPU time increases with the dimension of the problems, but it typically remains in the tens of seconds, even for the largest instance. Furthermore, we observe that, for all but two cases (i.e.,  $n = 100$  and  $n = 200$ ), the best optimal value computed by our algorithm is larger than the one computed by IPOPT, as evidenced by values in the first columns of the corresponding tables. The box-plots in Fig. 1 illustrate the distribution of the critical angles across the different initial points with a digital precision of  $10^{-6}$ . We can see that the interquartile ranges (IQRs) across the box-plots are relatively narrow for both the solvers, indicating consistent optimization performance. For  $n \geq 400$ , our algorithm can converge to better quality points (i.e., corresponding to a large objective function value) than IPOPT for almost all the initial points, confirming its good performance.

n	$\theta_b(P, Q)$	$it_b$	$it_a$	$it_w$	CPU <sub>b</sub>	CPU <sub>a</sub>	CPU <sub>w</sub>
5	8.5242e-01 $\pi$	13	88.64	1101	1.64e-04	1.05e-03	1.11e-02
20	9.2821e-01 $\pi$	48	262.84	1325	7.03e-04	3.63e-03	1.74e-02
50	9.5478e-01 $\pi$	113	519.89	1969	2.40e-03	1.10e-02	4.13e-02
100	9.6794e-01 $\pi$	195	1120.44	4227	8.44e-03	4.80e-02	1.76e-01
200	9.7013e-01 $\pi$	240	1279.29	3216	5.56e-02	3.01e-01	1.26e+00
400	9.6683e-01 $\pi$	347	1257.72	2899	1.89e-01	7.21e-01	3.54e+00
500	9.6689e-01 $\pi$	446	1236.44	2367	3.28e-01	1.00e+00	2.20e+00
700	9.6626e-01 $\pi$	494	1174.94	2403	8.12e-01	2.22e+00	6.05e+00
1000	9.6360e-01 $\pi$	561	1105.09	2009	3.02e+00	6.38e+00	1.20e+01

Table 1: Performance of Algorithm 1 for computing a critical angle between two polyhedral cones with  $10^3$  initial points.

n	$\theta_b(P, Q)$	$it_b$	$it_a$	$it_w$	CPU <sub>b</sub>	CPU <sub>a</sub>	CPU <sub>w</sub>
5	8.5242e-01 $\pi$	11	19.63	79	2.23e-02	3.01e-02	6.97e-02
20	9.2822e-01 $\pi$	60	207.65	1665	6.11e-02	1.96e-01	1.56e+00
50	9.5483e-01 $\pi$	194	1733.40	5000	2.90e-01	2.36e+00	1.08e+01
100	9.6812e-01 $\pi$	729	4517.13	5000	1.21e+00	9.29e+00	2.49e+01
200	9.7352e-01 $\pi$	5000	5000.00	5000	1.35e+01	2.01e+01	5.83e+01
400	9.5666e-01 $\pi$	5000	5000.00	5000	5.23e+01	5.46e+01	8.27e+01
500	9.4717e-01 $\pi$	5000	5000.00	5000	7.85e+01	8.82e+01	1.36e+02
700	9.5286e-01 $\pi$	5000	5000.00	5000	1.66e+02	1.85e+02	2.61e+02
1000	9.3832e-01 $\pi$	5000	5000.00	5000	4.25e+02	4.56e+02	6.32e+02

Table 2: Performance of IPOPT for computing a critical angle between two polyhedral cones with  $10^3$  initial points.

In terms of CPU time, our approach is significantly faster than IPOPT: compare the three

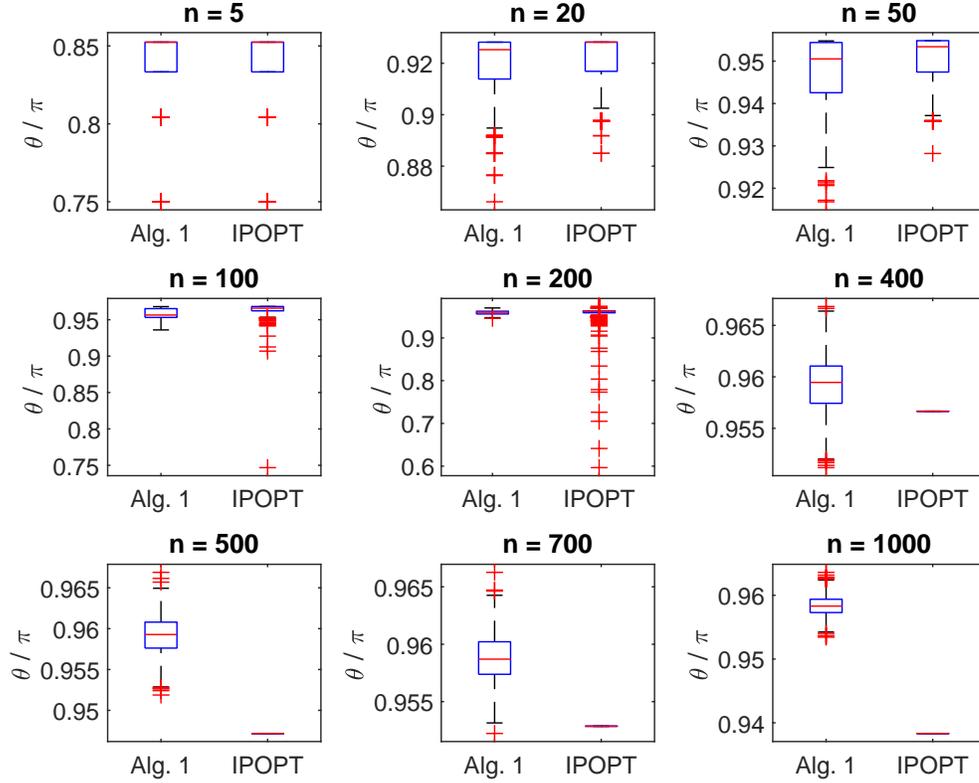


Figure 1: Box-plots of the critical angles between two polyhedral cones computed by Algorithm 1 and IPOPT with  $10^3$  initial points.

last columns of Tables 1 and 2. More precisely, IPOPT took about 226.5 hours (approximately, 9 days) to solve all the 9000 instances, while Algorithm 1 took only 2.9 hours.

In [32, Example 7.4], all the critical angles for the case  $n = 5$  were computed and are shown in Table 3. It is interesting to compare the number of times that our algorithm and IPOPT compute the maximum critical angle. To this end, we ran the two algorithms with  $10^3$  random initial points in the standard simplex, and in Table 3, we report (in percentage) the distribution of the critical angles computed. The results show that both Algorithm 1 and IPOPT compute the maximum angle for more than 60% of the total runs, and their performance is comparable in this case.

Solver / Angle	$0.6476 \pi; 0.6667 \pi; 0.6959 \pi; 0.7180 \pi$	$0.7500 \pi$	$0.7820 \pi$	$0.8041 \pi$	$0.8333 \pi$	$0.8524 \pi$
Algorithm 1	-	1.8%	-	12.6 %	21.6 %	64 %
IPOPT	-	1.7%	-	13.9 %	21.1 %	63.3 %

Table 3: Percentage of critical angles between the nonnegative orthant and the Schur cone in  $\mathbb{R}^5$  found by Algorithm 1 with  $10^3$  initial points.

## 6.2 Critical angles between two ellipsoidal cones

In this section, we evaluate the performance of Algorithm 1 for computing critical angles between a pair of ellipsoidal cones:

$$P = \{(\xi, t) \in \mathbb{R}^{n-1} \times \mathbb{R} : \sqrt{\langle \xi, A\xi \rangle} \leq t\} \quad \text{and} \quad Q = \{(\xi, t) \in \mathbb{R}^{n-1} \times \mathbb{R} : \sqrt{\langle \xi, B\xi \rangle} \leq t\},$$

where  $A$  and  $B$  are symmetric positive definite matrices of order  $n$ . Recall that  $P$  and  $Q$  are LISC cones which are represented by

$$P = G(\mathcal{L}_n^+) \quad \text{and} \quad Q = H(\mathcal{L}_n^+),$$

where  $\mathcal{L}_n^+$  is the Lorentz cone in  $\mathbb{R}^n$  and  $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^n$  are linear maps given by

$$G(\xi, t) = (A^{-1/2}\xi, t) \quad \text{and} \quad H(\xi, t) = (B^{-1/2}\xi, t),$$

for all  $(\xi, t) \in \mathbb{R}^{n-1} \times \mathbb{R}$ . As before, the subproblems in Step 2 of Algorithm 1 are reformulated (when  $\mu_1, \mu_2 > 0$ ) as orthogonal projection problems (36). In this case, from Example 3 item (b), we have that the solution to problem (36) is

$$\tilde{x}^k = \left( \frac{\tilde{b}^k}{\max\{1, \|\tilde{b}^k\|\}}, 1 \right),$$

where  $\tilde{b}^k \in \mathbb{R}^{n-1}$  is the vector consisting of the first  $n-1$  elements of  $(x^k - \frac{c_k}{\mu_1})$  and  $c_k$  given in (36). For this set of test problems, we generate positive definite matrices  $A$  and  $B$  as follows

$$A = C + nI \quad \text{and} \quad B = D + nI$$

where  $C$  and  $D$  are randomly generated sparse symmetric matrices with elements normally distributed, with mean 0 and variance 1. We generate several instances by considering  $n \in \{5, 20, 50, 100, 500, 700, 1000\}$  and density  $d = 0.5$ . The tolerances for the algorithm stopping test are  $\epsilon_1 = 10^{-6}$ ,  $\epsilon_2 = 10^{-6}$  and  $\epsilon_3 = 10^{-7}$ . We consider a maximum number of iterations equal to 5000, and the value of the prox-parameters are chosen as  $\mu_1 = 0.005$  and  $\mu_2 = 0.005$ . Finally, the initial point  $(x^0, y^0)$  is set as  $x^0 = (\xi^0, 1)$  and  $y^0 = (\nu^0, 1)$ , where  $\xi^0$  and  $\nu^0$  are random points in the Euclidean unit ball in  $\mathbb{R}^{n-1}$ .

We report in Tables 4 and 5 the performance of Algorithm 1 and IPOPT, respectively. The computed critical angles by both algorithms are found to be very close for most instances. However, for the last two instances, IPOPT manages to find a slightly better solution. It should be noted that IPOPT achieves this result for only one initial point out of  $10^3$  in the case of  $n = 700$  and  $n = 1000$ , as depicted in Fig 2.

One significant advantage of the proposed algorithm is that it requires significantly less computational effort compared to IPOPT. In fact, our algorithm is observed to be even  $10^2$  times faster for the largest instances. For instance, IPOPT took about 379 hours (15 days) to solve all the 9000 instances, while Algorithm 1 took only 4.5 hours. This substantial reduction in computation time allows for additional runs of our algorithm to explore a larger range of initial points, which could potentially yield improved critical angles.

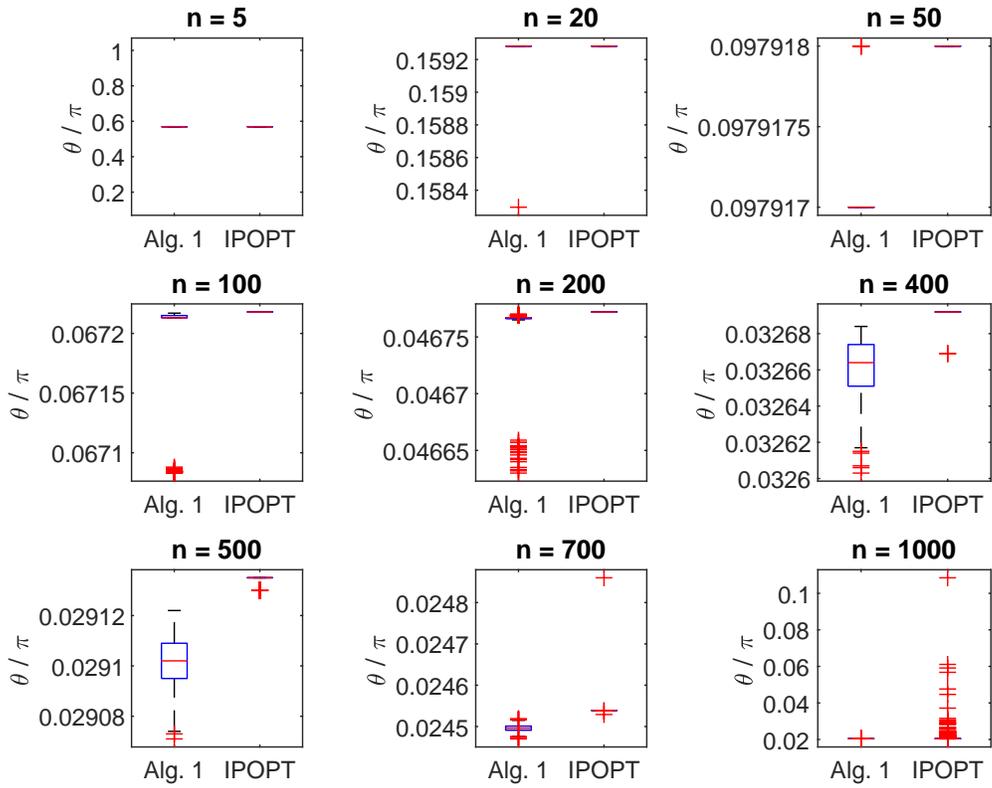


Figure 2: Box-plots of the critical angles between two ellipsoidal cones computed by Algorithm 1 and IPOPT with  $10^3$  initial points.

n	$\theta_b(P, Q)$	$it_b$	$it_a$	$it_w$	CPU <sub>b</sub>	CPU <sub>a</sub>	CPU <sub>w</sub>
5	5.6950e-01 $\pi$	9	10.43	13	1.86e-04	3.93e-04	2.98e-02
20	1.5928e-01 $\pi$	92	286.99	825	1.76e-03	5.68e-03	1.67e-02
50	9.7918e-02 $\pi$	101	266.82	732	5.01e-03	1.16e-02	2.91e-02
100	6.7217e-02 $\pi$	174	598.09	1684	2.26e-02	7.76e-02	2.54e-01
200	4.6770e-02 $\pi$	295	690.63	1912	1.41e-01	3.49e-01	1.17e+00
400	3.2684e-02 $\pi$	381	640.27	1382	7.08e-01	1.34e+00	4.14e+00
500	2.9122e-02 $\pi$	405	556.26	1043	1.48e+00	2.40e+00	4.39e+00
700	2.4519e-02 $\pi$	418	539.18	754	3.19e+00	4.34e+00	7.32e+00
1000	2.0430e-02 $\pi$	433	505.31	610	6.52e+00	7.95e+00	1.57e+01

Table 4: Performance of Algorithm 1 for computing a critical angle between two ellipsoidal cones with  $10^3$  initial points.

n	$\theta_b(P, Q)$	$it_b$	$it_a$	$it_w$	CPU <sub>b</sub>	CPU <sub>a</sub>	CPU <sub>w</sub>
5	5.6950e-01 $\pi$	9	12.86	37	2.09e-02	2.53e-02	4.61e-02
20	1.5928e-01 $\pi$	37	57.04	111	4.63e-02	7.57e-02	4.25e-01
50	9.7918e-02 $\pi$	36	50.20	93	7.61e-02	1.10e-01	2.19e-01
100	6.7218e-02 $\pi$	56	98.32	190	2.79e-01	5.44e-01	1.60e+00
200	4.6773e-02 $\pi$	56	96.89	239	1.59e+00	2.86e+00	8.51e+00
400	3.2693e-02 $\pi$	75	178.35	468	1.36e+01	3.32e+01	8.73e+01
500	2.9136e-02 $\pi$	102	291.72	1051	3.52e+01	1.00e+02	3.54e+02
700	2.4860e-02 $\pi$	99	255.17	973	9.23e+01	2.50e+02	1.00e+03
1000	1.0850e-01 $\pi$	110	314.44	878	3.20e+02	9.79e+02	5.26e+03

Table 5: Performance of IPOPT for computing a critical angle between two polyhedral cones with  $10^3$  initial points.

### 6.3 Critical angles between the SDP cone and the cone of nonnegative matrices

Recall that  $\mathcal{S}^n$  is the space of symmetric matrices of order  $n$  equipped with the trace inner product  $\langle A, B \rangle = \text{Tr}(AB)$ . Furthermore,  $\mathcal{P}_n$  is the SDP cone, and  $\mathcal{N}_n$  denotes the cone of nonnegative matrices in  $\mathcal{S}^n$ . That is,

$$\begin{aligned}\mathcal{P}_n &= \{A \in \mathcal{S}^n : x^\top Ax \geq 0, \forall x \in \mathbb{R}^n\}, \\ \mathcal{N}_n &= \{B \in \mathcal{S}^n : B \text{ is nonnegative entrywise}\}.\end{aligned}$$

Let  $\Theta(\mathcal{P}_n, \mathcal{N}_n)$  denote the maximal angle between  $\mathcal{P}_n$  and  $\mathcal{N}_n$ . The computation of  $\Theta(\mathcal{P}_n, \mathcal{N}_n)$  was exhaustively treated in [15] and [33]. It is known that

$$\Theta(\mathcal{P}_n, \mathcal{N}_n) = (3/4)\pi, \text{ for } n = 2, 3, 4, \quad \text{and} \quad \lim_{n \rightarrow \infty} \Theta(\mathcal{P}_n, \mathcal{N}_n) = \pi.$$

The exact computation of  $\Theta(\mathcal{P}_n, \mathcal{N}_n)$  for  $n \geq 5$  becomes difficult, and only lower bounds for certain values of  $n$  are known. As it is discussed in [18], a motivation to compute  $\Theta(\mathcal{P}_n, \mathcal{N}_n)$  is that it provides a lower bound for the greatest possible angle between two matrices in the copositive cone  $\mathcal{C}_n := \{A \in \mathcal{S}^n : x^\top Ax \geq 0, \forall x \in \mathbb{R}_+^n\}$ , cf. [15]. Our goal in this section is to use Algorithm 1 to improve the knowledge of these lower bounds.

It is clear that  $\mathcal{P}_n$  is a LISC cone since it is the symmetric cone of the algebra  $\mathcal{S}^n$ . In this case, its representation is

$$\mathcal{P}_n = G(\mathcal{P}_n)$$

with  $G : \mathcal{S}^n \rightarrow \mathcal{S}^n$  the identity map.

The cone  $\mathcal{N}_n$  is also LISC. Indeed, it can be expressed as

$$\mathcal{N}_n = H(\mathbb{R}_+^N),$$

where  $N = n(n+1)/2$  and  $H : \mathbb{R}^N \rightarrow \mathcal{S}^n$  is the linear map constructed as follows: Let  $\mathcal{E} := \{e_1, \dots, e_N\}$  be the canonical basis of  $\mathbb{R}^N$  and let  $\{E_{i,j}\}_{i \leq j}$  be the canonical basis of  $\mathcal{S}^n$ . That is,  $E_{i,j} \in \mathcal{S}^n$  is the matrix whose entries are zeros except the entries corresponding to  $(i, j)$  and  $(j, i)$  in which take values equal to one. It is clear that  $H$  is completely defined by knowing the image of each element of its basis  $\mathcal{E}$ . In this way, we take the following assignment: for  $k = 1, \dots, N$ ,

$$H(e_k) = E_{i_k, j_k},$$

where  $i_k := 2k - \ell_k(\ell_k - 1)/2$  and  $j_k := \operatorname{argmin}_{1 \leq \ell \leq n} \{\ell(\ell+1)/2 : \ell(\ell+1)/2 \geq k\}$ . Hence, for  $y \in \mathbb{R}^N$ , the value of  $H(y)$  is nothing but

$$H(y) = \sum_{k=1}^N y_k E_{i_k, j_k}.$$

The mapping  $H$  is invertible since it maps the canonical basis of  $\mathbb{R}^N$  onto the canonical basis of  $\mathcal{S}^n$ . Hence, assumptions (A1) and (A2) are satisfied and  $\mathcal{N}_n$  is a LISC cone.

In the implementation of Algorithm 1, we also need to know the map  $H^T : \mathcal{S}^n \rightarrow \mathbb{R}^N$  (the adjoint map of  $H$ ). It is not difficult to see that it is given by

$$H^T(X) = (\langle X, E_{i_1, j_1} \rangle, \dots, \langle X, E_{i_N, j_N} \rangle)^T.$$

In practice, the numerical evaluation of  $H(y)$  and  $H^T(X)$  is easy. To illustrate it, consider  $n = 3$ . Then  $N = 6$  and we have:

$$H \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} y_1 & y_2 & y_4 \\ y_2 & y_3 & y_5 \\ y_4 & y_5 & y_6 \end{pmatrix} \quad \text{and} \quad H^T \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix} = \begin{pmatrix} x_{11} \\ 2x_{12} \\ x_{22} \\ 2x_{13} \\ 2x_{23} \\ x_{33} \end{pmatrix}.$$

Now, we are ready to implement Algorithm 1 to estimate  $\Theta(\mathcal{P}_n, \mathcal{N}_n)$ . Algorithm 1 generates the points  $(X^k, y^k)$ . Recall that  $X^k$  is a matrix in  $\mathcal{S}^n$  and  $y^k$  is a vector in  $\mathbb{R}^N$ . In Step 2 of Algorithm 1, for  $\mu_1 > 0$ , problem (26) becomes

$$\begin{aligned} \min & \quad \frac{1}{2} \|X - B^k\|^2 \\ \text{s.t.} & \quad X \in \mathcal{P}_n, \\ & \quad \operatorname{Tr}(X) = 1, \end{aligned}$$

where  $B^k := X^k - (1/\mu_1)(Hy^k - \delta_k\|X^k\|^{-1}\|Hy^k\|X^k)$ . If  $B^k$  has the spectral decomposition, the solution to the above problem is  $X^k = Q\text{Diag}(\lambda(B^k)Q^T$ , with  $Q$  being some orthogonal matrix of order  $n$ . From Example 3 we have that  $X^k = Q\text{Diag}(\bar{x})Q^T$ , where  $\bar{x} \in \mathbb{R}^n$  is the solution of

$$\begin{aligned} \min \quad & \frac{1}{2}\|x - \lambda(B^k)\|^2 \\ \text{s.t.} \quad & x \in \mathbb{R}_+^n, \\ & \langle \mathbf{1}_n, x \rangle = 1. \end{aligned}$$

When  $\mu_2 > 0$ , problem (27) becomes

$$\begin{aligned} \min \quad & \frac{1}{2}\|y - b^k\|^2 \\ \text{s.t.} \quad & y \in \mathbb{R}_+^N, \\ & \langle \mathbf{1}_N, y \rangle = 1, \end{aligned}$$

where  $b^k := y^k - (1/\mu_2)H^T(X^k - \delta_k\|X^k\|^{-1}\|Hy^k\|^{-1}Hy^k)$ .

The tolerances for the algorithm stopping test are  $\epsilon_1 = 10^{-6}$ ,  $\epsilon_2 = 10^{-6}$  and  $\epsilon_3 = 10^{-7}$ . We consider a maximum number of iterations equal to 5000, and the value of the prox-parameters are chosen as  $\mu_1 = 0.01$  and  $\mu_2 = 5$ . For the initial point  $(X^0, y^0)$ , we set as  $X^0 = \text{Diag}(x^0)$  and, we generated  $x^0$  and  $y^0$  as two random points in the simplex. The largest dimension for the instances is 60, because the subproblems to be solved at each iteration are more complex. In particular, the first requires the spectral decomposition of a matrix and a subsequent projection on the  $n$ -dimensional standard simplex. The second subproblem also consists in projecting on the standard simplex, but this time the dimension is much higher, that is  $N = n(n+1)/2$  (e.g., with  $n = 60$ ,  $N = 1830$ ).

In Table 6, we report the best critical angle  $\theta_b(\mathcal{P}_n, \mathcal{N}_n)$  obtained by using Algorithm 1 for different values of  $n$ . In the second column of this table, titled ' $\theta_{lb}(\mathcal{P}_n, \mathcal{N}_n)$ ', we additionally show the values computed by using the procedure in [33] which only guarantees a lower bound for the maximal angle between the two considered cones. We solve the optimization problem given in [33] by IPOPT and considering  $10^3$  random initial points. Then, in the table, we provide the best value computed. We can see that our algorithm could find a slightly better solution for  $n = 50$  requiring a very small computational time. However, the reported values improve significantly what was known in [33, Table 2]. We do not compare our algorithm with IPOPT applied to FP (5) because this solver is not applicable in this more complex setting: the Cones  $\mathcal{P}_n$  and  $\mathcal{N}_n$  do not have a favorable structure for off-the-shelf NLP solvers.

## 7 Conclusions

In this work, we have investigated the problem of computing critical angles between pairs of convex cones in finite-dimensional Euclidean spaces. Our contributions are summarized as follows: a) we have shown that the notion of critical angles considered in the literature is equivalent to that of (KKT) stationarity for a class of nonlinear optimization problems; b) we have posed the problem in the fractional programming (FP) setting and proposed a specialized algorithm to compute stationary points, yielding thus critical angles to our original problem; c) we have further specialized our analysis to the broad class of (LISC) convex cones derived as linear images of symmetric cones. Moreover, we have demonstrated that our FP formulation is well-defined and suitable for numerical optimization.

n	$\theta_{lb}(\mathcal{P}_n, \mathcal{N}_n)$	$\theta_b(\mathcal{P}_n, \mathcal{N}_n)$	it <sub>b</sub>	it <sub>a</sub>	it <sub>w</sub>	CPU <sub>b</sub>	CPU <sub>a</sub>	CPU <sub>w</sub>
4	0.7500 $\pi$	0.7500 $\pi$	18	35.96	322	7.29e-04	1.76e-03	3.54e-02
5	0.7575 $\pi$	0.7575 $\pi$	20	102.32	536	8.87e-04	5.84e-03	3.23e-02
6	0.7575 $\pi$	0.7575 $\pi$	19	125.89	479	1.10e-03	8.35e-03	2.78e-02
7	0.7575 $\pi$	0.7575 $\pi$	22	202.35	478	1.16e-03	1.27e-02	3.34e-02
8	0.7608 $\pi$	0.7608 $\pi$	23	237.76	635	1.26e-03	1.59e-02	4.22e-02
9	0.7608 $\pi$	0.7608 $\pi$	20	243.34	530	1.27e-03	1.61e-02	3.95e-02
10	0.7609 $\pi$	0.7609 $\pi$	25	256.06	750	1.57e-03	1.73e-02	4.47e-02
11	0.7627 $\pi$	0.7627 $\pi$	25	260.09	870	1.97e-03	1.80e-02	5.89e-02
12	0.7649 $\pi$	0.7649 $\pi$	141	273.30	903	9.53e-03	1.90e-02	6.14e-02
13	0.7649 $\pi$	0.7649 $\pi$	130	282.42	1114	9.26e-03	2.09e-02	7.03e-02
14	0.7659 $\pi$	0.7659 $\pi$	123	304.45	927	8.73e-03	2.27e-02	7.51e-02
15	0.7678 $\pi$	0.7678 $\pi$	137	327.89	1304	1.04e-02	2.54e-02	9.36e-02
16	0.7699 $\pi$	0.7699 $\pi$	133	345.00	1835	1.06e-02	2.79e-02	1.40e-01
17	0.7699 $\pi$	0.7699 $\pi$	128	350.01	1958	1.17e-02	3.09e-02	1.60e-01
18	0.7699 $\pi$	0.7699 $\pi$	122	350.11	2179	1.23e-02	3.71e-02	1.96e-01
19	0.7703 $\pi$	0.7703 $\pi$	129	355.40	1548	1.24e-02	3.69e-02	1.72e-01
20	0.7719 $\pi$	0.7719 $\pi$	135	390.75	1741	1.37e-02	3.96e-02	2.91e-01
21	0.7719 $\pi$	0.7719 $\pi$	130	396.87	1955	1.52e-02	4.27e-02	1.91e-01
22	0.7719 $\pi$	0.7719 $\pi$	136	420.58	1724	1.57e-02	4.77e-02	1.83e-01
23	0.7722 $\pi$	0.7722 $\pi$	142	412.98	1593	1.73e-02	4.99e-02	1.76e-01
24	0.7735 $\pi$	0.7735 $\pi$	150	420.40	1831	1.81e-02	5.10e-02	2.03e-01
25	0.7735 $\pi$	0.7735 $\pi$	147	425.24	1480	2.00e-02	6.42e-02	2.59e-01
26	0.7735 $\pi$	0.7735 $\pi$	133	458.19	1454	3.25e-02	1.29e-01	4.12e-01
27	0.7739 $\pi$	0.7739 $\pi$	156	456.77	1997	3.93e-02	1.41e-01	5.50e-01
28	0.7750 $\pi$	0.7750 $\pi$	158	461.31	1764	4.14e-02	1.38e-01	5.63e-01
29	0.7750 $\pi$	0.7750 $\pi$	149	492.23	1704	4.21e-02	1.52e-01	5.08e-01
30	0.7757 $\pi$	0.7757 $\pi$	148	504.14	1834	4.04e-02	1.60e-01	6.21e-01
40	0.7789 $\pi$	0.7789 $\pi$	275	787.20	2197	1.27e-01	4.29e-01	1.32e+00
50	0.7809 $\pi$	0.7812 $\pi$	356	1193.38	3108	2.50e-01	9.71e-01	2.65e+00
60	0.7837 $\pi$	0.7837 $\pi$	507	1553.64	4189	5.57e-01	1.78e+00	6.93e+00

Table 6: Performance of Algorithm 1 for computing a critical angle between the SDP cone and the cone of nonnegative matrices.

Our algorithm is the first numerical procedure for computing critical angles between pairs of convex cones. The approach requires solving two independent subproblems at every iteration to compute (with the aid of the Armijo line search) a new iterate. For the class of LISC cones, we have shown that solutions to such subproblems are projections onto well-structured convex sets. Otherwise, for more general cones, we can still apply off-the-shelf (linear, quadratic, conic, etc.) convex solvers to solve such subproblems. Our theoretical analysis demonstrates that our approach asymptotically computes critical angles regardless of initialization.

Numerical experiments on different types of convex cones (polyhedral, ellipsoidal, and Loewnerian cones) in several dimensions illustrate the practical performance of our algorithm. Comparison with IPOPT have shown that exploiting the problem’s structure as done by our approach pays off: for the considered test problems, our algorithm is up to one hundred times faster than IPOPT while maintaining (sometimes even improving) the quality of computed critical angles.

## Data availability

All data generated or analysed during this study are included in this article.

## References

- [1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Program. Ser. B*, 95 (2003), 3–51.
- [2] M. Baes. Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras. *Linear Algebra Appl.*, 422 (2007), 664–700.
- [3] H. Bauschke, H. Ouyang, and X.-Wang. On angles between convex sets in Hilbert spaces. *J. Math. Anal. Appl.*, 502 (2021), 125239.
- [4] H. P. Benson. Fractional programming with convex quadratic forms and functions. *Eur. J. Oper. Res.*, 173 (2006) 351–369.
- [5] R. I. Bot and E. R. Csetnek. Proximal-gradient algorithms for fractional programming. *Optimization*, 66 (2017), 1383-1396.
- [6] R. I. Bot and Minh N. Dao and G. Li. Extrapolated Proximal Subgradient Algorithms for Non-convex and Nonsmooth Fractional Programs. *Math. Oper. Res.*, 47 (2022), 1707–2545.
- [7] L. Condat. Fast projection onto the simplex and the  $l_1$  ball. *Math. Program.*, 158 (2016), 575–585.
- [8] J. Currie, D.I. Wilson. OPTI: Lowering the barrier between open source optimizers and the industrial MATLAB user. Foundations of Computer-Aided Process Operations. Georgia, USA, 2012.
- [9] W. Dinkelbach. On nonlinear fractional programming. *Manag. Sci.*, 13 (1967), 492–498.
- [10] A. Dolatnezhadsomarin and E. Khorram and L. Pourkarimi. Efficient algorithms for solving nonlinear fractional programming problems. *Filomat*, 33 (2019), 2149–2179.
- [11] F. Facchinei, J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems I*. Springer New York, NY, 2011.
- [12] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Clarendon Press, Oxford, UK, 1994.

- [13] H. Frenk and S. Schaible. *Fractional Programming. in: Floudas C., Pardalos P. (eds) Encyclopedia of Optimization.* Springer, Boston, MA, 2008.
- [14] M. Fukushima and J. Judice and W. de Oliveira and V. Sessa. A Sequential Partial Linearization Algorithm for the Symmetric Eigenvalue Complementarity Problem. *Comput. Optim. Appl.*, 77 (2020) 711–728.
- [15] F. Goldberg and N. Shaked-Monderer. On the maximal angle between copositive matrices. *Elec. J. Linear Algebra*, 27 (2014), 837–850.
- [16] J. Gorski, F. Pfeuffer and K. Klamroth. Biconvex sets and optimization with biconvex functions: a survey and extensions. *Math Meth Oper Res*, 66 (2007), 373–407.
- [17] R. Greer. Chapter 2: A tutorial on polyhedral convex cones. *In: Trees and Hills: Methodology for Maximizing Functions of Systems of Linear Relations, in: North-Holland Mathematics Studies*, 96 (1984), 15–81.
- [18] J.-B. Hiriart-Urruty and A. Seeger. A variational approach to copositive matrices. *SIAM Review*, 52 (2010), 593–629.
- [19] A. Iusem and A. Seeger. On pairs of vectors achieving the maximal angle of a convex cone. *Math. Program.*, 104 (2005), 501–523.
- [20] A. Iusem and A. Seeger. Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. *Comput. Applied Math.*, 26 (2007), 191–214.
- [21] A. Iusem and A. Seeger. On convex cones with infinitely many critical angles. *Optimization*, 56 (2007), 115–128.
- [22] A. Iusem and A. Seeger. Searching for critical angles in a convex cone. *Math. Program., Ser B*, 120 (2009), 3–25.
- [23] J. Miao and A. Ben-Israel. On principal angles between subspaces in  $\mathbb{R}^n$ . *Linear Algebra Appl.*, 171 (1992), 81– 98.
- [24] M. Orlitzky. When a maximal angle among cones is nonobtuse. *Comput. Appl. Mathem.*, 39 (2020), 1–14.
- [25] N. Parikh and S. Boyd. Proximal Algorithms. *Foundations and Trends in Optimization*, 1 (2014), 123–231.
- [26] G. Pataki. On the closedness of the linear image of a closed convex cone. *Math. Oper. Res.*, 32 (2007), 395–412.
- [27] H. Ramírez, A. Seeger, and D. Sossa. Commutation principle for variational problems on Euclidean Jordan algebras. *SIAM J. Optim.*, 23 (2013), 687–694.
- [28] H. Ramírez and S. Sossa. On the Central Paths in Symmetric Cone Programming. *J Optim Theory Appl*, 172 (2017), 649–668.
- [29] R.T. Rockafellar, R. J-B. Wets. *Variational Analysis.* Springer Berlin, Heidelberg, 1998.
- [30] R.T. Rockafellar. *Convex analysis.* Princeton University Press, 1970.
- [31] A. Seeger, D. Sossa. Complementarity problems with respect to Loewnerian cones. *J Glob Optim*, 62 (2015), 299–318.

- [32] A. Seeger, D. Sossa. Critical angles between two convex cones I. General theory. *TOP*, 24 (2016), 44–65.
- [33] A. Seeger, D. Sossa. Critical angles between two convex cones II. Special cases. *TOP*, 24 (2016), 66–87.
- [34] A. Seeger, M. Torki. Ellipsoidal cones. *J. Convex Anal.*, 27 (2020), 583–622.
- [35] S. Schaible, Fractional programming: Applications and algorithms. *European Journal of Operational Research*, 7 (1981), 111–120.
- [36] S. Schaible, Fractional programming. II, on Dinkelbach’s algorithm, *Management Science*, 22 (1976), 868–873.
- [37] N. Sogi, R. Zhu, J.-H. Xue and K. Fukui. Constrained mutual convex cone method for image set based recognition. *Pattern Recognition*, 121 (2022), 108190.
- [38] I. M. Stancu-Minasian. *Fractional Programming, Theory, Methods and Applications*. Kluwer Academic Publishers, London, 1997.
- [39] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38 (1996), 49–95.
- [40] A. Wächter, L. Biegler. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math. Program.*, 106 (2006), 25–57.