TOP (2016) 24:44–65
DOI 10.1007/s11750-015-0375-y

ORIGINAL ARTICLE

Critical angles between two convex cones I. General theory

Alberto Seeger · **David Sossa**

Received: 27 March 2014 / Accepted: 19 February 2015 / Published online: 5 March 2015

Abstract The concept of critical (or principal) angle between two linear subspaces has applications in statistics, numerical linear algebra, and other areas. Such concept has been abundantly studied in the literature, both from a theoretical and computational point of view. Part I of this work is an attempt to build a general theory of critical angles for a pair of closed convex cones. The need of such theory is motivated, among other reasons, by some specific problems arising in regression analysis of cone-constrained data, see Tenenhaus (Psychometrika 53:503–524, 1988). Angle maximization and/or angle minimization problems involving a pair of convex cones are at the core of our discussion. Such optimization problems are nonconvex in general and their numerical resolution offer a number of challenges. Part II of this work focusses on the practical computation of the maximal and/or minimal angle between specially structured cones.

Keywords Maximal angle · Critical angle · Principal angle · Convex cone · Canonical analysis · Nonconvex optimization · Optimality conditions

Mathematics Subject Classification 15A18 · 15A48 · 52A40 · 90C26 · 90C33

D. Sossa was supported by CONICYT, Chile.

A. Seeger (✉)
Département de Mathématiques, Université d'Avignon, 33 rue Louis Pasteur, 84000 Avignon, France
e-mail: alberto.seeger@univ-avignon.fr

D. Sossa
Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada, 2120 Santiago, Chile
e-mail: dsossa@dim.uchile.cl

1 Introduction

This work is the merging point of two independent sources: the recent theory of critical angles for a closed convex cone, as developed by Iusem and Seeger (2005), and the old theory of principal angles for a pair of linear subspaces. Let $(\mathbb{X}, \langle\cdot,\cdot\rangle)$ be a Euclidian space of dimension at least two and let $\mathcal{C}(\mathbb{X})$ be the set of nontrivial closed convex cones in $\mathbb{X}$. That a closed convex cone is nontrivial means that it is different from the zero cone and different from the whole space. Computing the maximal angle of a closed convex cone is an issue of importance in a number of applications, see for instance Clarke et al. (1997, 1999) and Peña and Renegar (2000). By definition, the maximal angle of $K \in \mathcal{C}(\mathbb{X})$ is the number

$$\theta_{\max}(K) := \max_{u,v\in K\cap S_{\mathbb{X}}} \arccos\langle u, v\rangle, \tag{1}$$

where $S_{\mathbb{X}}$ is the unit sphere of $\mathbb{X}$. By writing down the necessary optimality conditions for the nonconvex optimization problem (24), one gets

$$u, v \in K \cap S_{\mathbb{X}}, \quad v - \langle u, v\rangle u \in K^*, \quad u - \langle u, v\rangle v \in K^*, \tag{2}$$

where K^* denotes the positive dual cone of K. If the system (2) holds, then (u, v) is called a critical pair of K and $\arccos\langle u, v\rangle$ is called a critical angle of K. The study of critical angles in a convex cone was initiated in Iusem and Seeger (2005) and further continued in Gourion and Seeger (2010) and Iusem and Seeger (2007a, b, 2008a, 2009). The purpose of the present work is to introduce the first ingredients for a theory of critical angles in a pair of convex cones. Such a theory is to encompass, as particular cases, the theory of principal angles in a pair of subspaces and the theory of critical angles in a single cone. The starting point of our analysis is the formulation of the optimization problem that defines the maximal angle between two convex cones.

Definition 1.1 Let $P, Q \in \mathcal{C}(\mathbb{X})$. The *maximal angle* of (P, Q) or, more precisely, the maximal angle between P and Q, is given by

$$\Theta(P, Q) := \max_{u\in P\cap S_{\mathbb{X}},\, v\in Q\cap S_{\mathbb{X}}} \arccos\langle u, v\rangle. \tag{3}$$

An *antipodal pair* of (P, Q) is any pair $(u, v) \in \mathbb{X}^2$ solving the above maximization problem.

Antipodal pairs always exist, but they are not unique in general. On the other hand, it is important to observe that (3) is a nonconvex optimization problem. As we shall see in Proposition 1.3, any antipodal pair is a critical pair in the following sense:

Definition 1.2 Let $P, Q \in \mathcal{C}(\mathbb{X})$. A *critical pair* of (P, Q) is a pair $(u, v) \in \mathbb{X}^2$ satisfying

$$\begin{cases} u \in P \cap S_{\mathbb{X}}, \\ v \in Q \cap S_{\mathbb{X}}, \\ v - \langle u, v\rangle u \in P^*, \\ u - \langle u, v\rangle v \in Q^*. \end{cases} \tag{4}$$

The corresponding angle $\arccos\langle u, v\rangle$ is called a *critical angle* of (P, Q). A critical pair (u, v) of (P, Q) is called *proper* if u and v are not collinear. The corresponding angle is called a *proper critical angle* of (P, Q).

Definition 1.2 is directly inspired from the particular case (2) relative to a single cone. The set of critical angles of (P, Q) is denoted by $\Gamma(P, Q)$ and it is called the angular spectrum of (P, Q). By convention, one writes $\Gamma(P, Q) = \{0, \pi\}$ if either P or Q is the whole space $\mathbb{X}$. In general, $\Gamma(P, Q)$ is a nonempty closed subset of the interval $[0, \pi]$. Beware that the cardinality of $\Gamma(P, Q)$ is not necessarily finite. In other words, a pair (P, Q) may have infinitely many critical angles.

Proposition 1.3 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. A necessary condition for $(u, v) \in \mathbb{X}^2$ to be an antipodal pair of (P, Q) is to satisfy the system (4).*

Proof Let (u, v) be an antipodal pair of (P, Q). In particular, the component u minimizes the linear form $\langle \cdot, v\rangle$ on $P \cap S_{\mathbb{X}}$. Consider an arbitrary nonzero vector $d \in P$. Clearly,

$$\mathbf{u}(t) := \|u + td\|^{-1}(u + td) \in P \cap S_{\mathbb{X}}$$

for all t in an interval $[0, \varepsilon[$. Furthermore, $t = 0$ is a minimum of $t \in [0, \varepsilon[\mapsto f(t) := \langle \mathbf{u}(t), v\rangle$. Hence, the right-derivative

$$f'_+(0) = \langle v, d\rangle - \langle u, v\rangle\langle u, d\rangle$$

is nonnegative. This proves the third condition in (4). Analogously, the last condition in (4) is obtained by using the fact v minimizes $\langle u, \cdot\rangle$ on $Q \cap S_{\mathbb{X}}$. □

In a similar way one can treat the angle minimization problem

$$\Psi(P, Q) := \min_{u \in P \cap S_{\mathbb{X}}, v \in Q \cap S_{\mathbb{X}}} \arccos\langle u, v\rangle. \tag{5}$$

Angle minimization problems like (5) arise in a number of applications, for instance in the theory of exponential dichotomies for linear ODEs (cf. Obert 1991) and in regression analysis of ordinal data (cf. Tenenhaus 1988). The necessary optimality conditions for the angle minimization problem (5) are similar to (4), but one must change dual cones by polar cones. On the other hand, one readily sees that

$$\cos[\Psi(P, Q)] = -\cos[\Theta(P, -Q)], \quad \Psi(P, Q) = \pi - \Theta(P, -Q).$$

So, there is no loss of generality in focussing the attention just on angle maximization. We give the priority to angle maximization over angle minimization. Among other reasons, our choice is motivated by the following facts:

1. When P and Q are equal, the minimization problem (5) is of no interest, whereas the maximization problem (3) serves to define the maximal angle of a single cone. One of the goals of this work is to extend Iusem–Seeger's theory of critical angles from one convex cone to a pair of convex cones.

2. Many angle optimization problems come originally in the maximization format, see examples in Sect. 1.1.
3. As we shall see in Sect. 4, angle maximality has a bearing with the issue of measuring the degree of pointedness and the degree of reproducibility of a pair of convex cones.

Our work programme covers many different themes and for this reason the paper may give the impression of lack of unity. The overall structure of the paper is as follows:

- Section 2 addresses duality issues. We show that there is a simple link between the critical pairs of (P, Q) and the critical pairs of (P^*, Q^*). Section 2 establishes also a certain boundary principle, according which the components of a proper critical pair of (P, Q) should be sough on the boundaries of P and Q, respectively.
- Section 3 provides alternative characterizations of criticality and duality.
- Section 4 shows that the degree of pointedness of a pair (P, Q) can be expressed as function of the maximal angle between P and Q. Similarly, the degree of reproducibility of (P, Q) can be expressed as function of the maximal angle between P^* and Q^*.
- Section 5 addresses continuity issues. One equips the set $\mathcal{C}(\mathbb{X})$ with a suitable metric and one shows that $\Theta(P, Q)$ behaves in a Lipschitz continuous manner with respect to perturbations in P and Q.
- Section 6 is devoted to the analysis of critical angles for a pair of subspaces. One shows that, in such a particular context, the concept of critical angle coincides with the classical concept of principal angle.
- Section 7 is devoted to the analysis of critical angles for a pair (P, Q) of polyhedral cones. In such a context, the angular spectrum $\Gamma(P, Q)$ is a finite set. We provide easily computable upper bounds for the cardinality of $\Gamma(P, Q)$ and, what is more important, we explain how to compute all the elements of $\Gamma(P, Q)$.

1.1 Some motivating examples

The formulation of the maximization problem (3) is motivated by theoretical and practical considerations.

Example 1.4 Consider the problem of finding a point z in the intersection of two closed convex sets $C, D \subseteq \mathbb{X}$. Von Neuman's alternating projection algorithm produces a sequence $\{z^k\}_{k\in\mathbb{N}}$ obtained by successive projections $z^{2k+1} = \Pi_C(z^{2k})$, $z^{2k+2} = \Pi_D(z^{2k+1})$ onto C and D, respectively. If the algorithm is initialized at a point z^0 near z, then it is reasonable to expect convergence toward z. As explained in Lewis et al. (2009, Theorem 5.16), the rate of convergence depends on the coefficient

$$\tau_z := \max_{\substack{u\in N_C(z),\ \|u\|\leq 1\\ v\in -N_D(z),\ \|v\|\leq 1}} \langle u, v\rangle, \tag{6}$$

where $N_C(\cdot)$ stands for the normal cone map associated to C. Note that (6) can be rewritten as

$$\tau_z = \max\{0, -\cos[\Theta(N_C(z), N_D(z))]\},$$

so one must evaluate the maximal angle between the cones $N_C(z)$ and $N_D(z)$. In a similar context, Drusvyatkiy (2013, Proposition 3.2.2) uses the term $\Theta(N_C(z), N_D(z))$ for estimating a coefficient called modulus of intrinsic transversality of the pair (C, D) at the point z.

Example 1.5 Consider the space $\mathrm{Sym}(n)$ of symmetric matrices of order n equipped with the trace inner product $\langle A, B\rangle = \mathrm{tr}(AB)$. An interesting question of linear algebra is to compute the maximal angle between the cones

$$\mathcal{P}_n := \{A \in \mathrm{Sym}(n) : A \text{ is positive semidefinite}\},$$
$$\mathcal{N}_n := \{B \in \mathrm{Sym}(n) : B \text{ is nonnegative entrywise}\}.$$

One usually refers to $\mathcal{P}_n$ as the SDP cone in $\mathrm{Sym}(n)$. By relying on graph theory arguments, Goldberg and Shaked-Monderer (2014) obtained a lower bound for the maximal angle $\Theta(\mathcal{P}_n, \mathcal{N}_n)$ and proved the asymptotic formula $\lim_{n\to\infty}\Theta(\mathcal{P}_n, \mathcal{N}_n) = \pi$. It remains an open question to compute the exact value of $\Theta(\mathcal{P}_n, \mathcal{N}_n)$.

2 Duality and boundary principles for critical pairs

Let $\Omega_{\mathbb{X}}$ denote the set of all pairs of unit vectors in $\mathbb{X}$ that are not collinear, i.e.,

$$\Omega_{\mathbb{X}} := \{(u, v) \in S_{\mathbb{X}}^2 : |\langle u, v\rangle| \neq 1\}.$$

To each $(u, v) \in \Omega_{\mathbb{X}}$, one can associate its *conjugate pair*

$$g(u, v) = \left(\frac{v - \langle u, v\rangle u}{\sqrt{1 - \langle u, v\rangle^2}}, \frac{u - \langle u, v\rangle v}{\sqrt{1 - \langle u, v\rangle^2}}\right).$$

It is not difficult to check that $g : \Omega_{\mathbb{X}} \to \Omega_{\mathbb{X}}$ is a bijection with $g^{-1} = g$. In other words, g is an involution on $\Omega_{\mathbb{X}}$. The following duality principle is an extension of Iusem and Seeger (2009, Theorem 2).

Theorem 2.1 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Let $(u, v) \in \Omega_{\mathbb{X}}$ and $(y, z) \in \Omega_{\mathbb{X}}$ be conjugate pairs. Then (u, v) is a critical pair of (P, Q) if and only if (y, z) is a critical of (P^*, Q^*).*

Proof Theorem 2 in Iusem and Seeger (2009) takes care of the particular case in which P is equal to Q. The proof of the general case follows the same pattern. Assume that (u, v) is critical for (P, Q) and write $\lambda := \langle u, v\rangle$. Clearly,

$$y = [1 - \lambda^2]^{-1/2}(v - \lambda u) \in P^* \cap S_{\mathbb{X}},$$
$$z = [1 - \lambda^2]^{-1/2}(u - \lambda v) \in Q^* \cap S_{\mathbb{X}}.$$

Furthermore, $\mu := \langle y, z \rangle = -\lambda$ and

$$z - \mu y = [1 - \mu^2]^{1/2}\, u \in P,$$
$$y - \mu z = [1 - \mu^2]^{1/2}\, v \in Q.$$

Hence, (y, z) is critical for (P^*, Q^*). The reverse implication is proven in a similar way. □

Corollary 2.2 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Let $\theta, \psi \in \,]0, \pi[$ be conjugate angles, i.e., $\theta + \psi = \pi$. Then θ is a critical angle of (P, Q) if and only if ψ is a critical angle of (P^*, Q^*).*

Proof Suppose that θ is a critical angle of (P, Q). Let (u, v) be any proper critical pair of (P, Q) such that $\cos\theta = \langle u, v \rangle$. Thanks to the duality principle established in Theorem 2.1, the conjugate pair $(y, z) = g(u, v)$ is critical for (P^*, Q^*). Since

$$\langle y, z \rangle = -\langle u, v \rangle = -\cos\theta = \cos(\pi - \theta) = \cos\psi,$$

one deduces that ψ is a critical angle of (P^*, Q^*). The proof of the reverse implication is similar. □

Intuitively speaking, the components u and v of a proper critical pair of (P, Q) should be on the boundaries of P and Q, respectively. The next theorem clarifies this point. For a set C contained in a linear subspace L of $\mathbb{X}$, the symbol $\mathrm{bd}_L(C)$ refers to the boundary of C relative to L.

Theorem 2.3 *Let $L \subseteq \mathbb{X}$ be the smallest linear subspace containing both $P \in \mathcal{C}(\mathbb{X})$ and $Q \in \mathcal{C}(\mathbb{X})$. Suppose that (u, v) is a proper critical pair of (P, Q). Then $u \in \mathrm{bd}_L(P)$ and $v \in \mathrm{bd}_L(Q)$.*

Proof Since (u, v) is proper, one has $\lambda := \langle u, v \rangle \notin \{-1, 1\}$. Suppose, to the contrary, that u belongs to the interior of P relative to L, i.e., there exists a positive ε such that

$$u + \varepsilon(B_{\mathbb{X}} \cap L) \subseteq P,$$

where $B_{\mathbb{X}}$ is the closed unit ball of $\mathbb{X}$. It follows that

$$0 \leq \langle v - \lambda u, u + \varepsilon w \rangle = \varepsilon \langle v - \lambda u, w \rangle$$

for all $w \in B_{\mathbb{X}} \cap L$. The particular choice $w = \|\lambda u - v\|^{-1}(\lambda u - v)$ leads to

$$0 \leq \varepsilon \|\lambda u - v\|^{-1} \langle v - \lambda u, \lambda u - v \rangle = -\varepsilon \|v - \lambda u\| < 0,$$

a clear contradiction. This shows that $u \in \mathrm{bd}_L(P)$. The proof of $v \in \mathrm{bd}_L(Q)$ is similar. □

The next corollary follows straightforwardly by combining the duality principle stated in Theorem 2.1 and the boundary principle stated in Theorem 2.3.

Corollary 2.4 *Let $P, Q \in \mathcal{C}(\mathbb{X})$ and (u, v) be a proper critical pair of (P, Q). Then*

$$\begin{cases} v - \langle u, v\rangle u \in \mathrm{bd}_M(P^*), \\ u - \langle u, v\rangle v \in \mathrm{bd}_M(Q^*), \end{cases}$$

where $M \subseteq \mathbb{X}$ is the smallest linear subspace containing both P^* and Q^*.

3 Further characterization of criticality and antipodality

Let $\Pi_C(x)$ denote the projection of a point $x \in \mathbb{X}$ onto a nonempty closed convex set $C \subseteq \mathbb{X}$. The next proposition expresses criticality for a pair (P, Q) in terms of the projection maps Π_P and Π_Q.

Proposition 3.1 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Let u, v be distinct points on the sphere $S_{\mathbb{X}}$. Then (u, v) is a critical pair of (P, Q) if and only if*

$$\begin{cases} \Pi_P(u - v) = (1 - \langle u, v\rangle)\, u, \\ \Pi_Q(v - u) = (1 - \langle u, v\rangle)\, v. \end{cases} \tag{7}$$

In particular, $\mathrm{dist}(u - v, P) = \mathrm{dist}(v - u, Q)$ *is a necessary condition for (u, v) to be a critical pair of (P, Q).*

Proof Let $\lambda := \langle u, v\rangle$. Let $N_P(u)$ denote the normal cone to P at u. Note that

$$\begin{aligned} \Pi_P(u - v) = (1 - \lambda)u &\Leftrightarrow \Pi_P((1 - \lambda)^{-1}(u - v)) = u \\ &\Leftrightarrow (1 - \lambda)^{-1}(u - v) \in u + N_P(u) \\ &\Leftrightarrow -(v - \lambda u) \in N_P(u) \\ &\Leftrightarrow u \in P,\ \ v - \lambda u \in P^*. \end{aligned}$$

Similarly, the second condition in (7) amounts to saying that $v \in Q$ and $u - \lambda v \in Q^*$. □

Sometimes it is helpful to write the angle maximization problem (3) in any of the following equivalent forms

$$\cos[\Theta(P, Q)] = \min_{u \in P \cap S_{\mathbb{X}}, v \in Q \cap S_{\mathbb{X}}} \langle u, v\rangle, \tag{8}$$

$$\kappa(P, Q) := \min_{u \in P \cap S_{\mathbb{X}}, v \in Q \cap S_{\mathbb{X}}} (1/2)\, \|u + v\|\,. \tag{9}$$

The problems (3), (8), and (9) have clearly the same solution set. Furthermore,

$$\kappa(P, Q) = \cos\left[\Theta(P, Q)/2\,\right]. \tag{10}$$

The next theorem relates (3) to the minimization problem

$$\chi(P, Q) := \min_{z \in S_{\mathbb{X}}} \max\{\mathrm{dist}(z, P), \mathrm{dist}(z, -Q)\}\,. \tag{11}$$

Although it is not clear at first sight, it turns out that solving (3) is equivalent to solving (11).

Theorem 3.2 *Let* $P, Q \in \mathcal{C}(\mathbb{X})$. *Then*

$$\Theta(P, Q) = 2 \arccos[\chi(P, Q)]. \tag{12}$$

Suppose, in addition, that P, Q *are not equal to a common ray. In such a case, the solution set* $\mathcal{A}(P, Q)$ *to the angle maximization problem* (3) *and the solution set* $\mathcal{R}(P, Q)$ *to the problem* (11) *are related as follows:*

$$\mathcal{R}(P, Q) = \left\{ \frac{u - v}{\|u - v\|} : (u, v) \in \mathcal{A}(P, Q) \right\}, \tag{13}$$

$$\mathcal{A}(P, Q) = \left\{ \left(\frac{\Pi_P(z)}{\|\Pi_P(z)\|}, \frac{\Pi_Q(-z)}{\|\Pi_Q(-z)\|} \right) : z \in \mathcal{R}(P, Q) \right\}. \tag{14}$$

Proof If P and Q are equal to a common ray, then both sides of (12) are equal to 0. Suppose now that P and Q are not equal to a common ray. Suppose also that $P \cap -Q = \{0\}$, otherwise both sides of (12) are equal to π and the proof of (13)–(14) is immediate. Let z_0 be a solution to (11). Hence,

$$\operatorname{dist}(z_0, P) = \operatorname{dist}(z_0, -Q) = \chi(P, Q),$$
$$\|\Pi_P(z_0)\| = \|\Pi_Q(-z_0)\| = s,$$

with $s := \left(1 - [\chi(P, Q)]^2\right)^{1/2}$ belonging to $]0, 1[$. The pair

$$\begin{aligned}(u_0, v_0) &:= \left(\|\Pi_P(z_0)\|^{-1} \Pi_P(z_0), \|\Pi_Q(-z_0)\|^{-1} \Pi_Q(-z_0) \right) \\ &= (1/s) \left(\Pi_P(z_0), \Pi_Q(-z_0) \right)\end{aligned}$$

is then well defined. We claim that

$$z_0 = \|u_0 - v_0\|^{-1}(u_0 - v_0), \tag{15}$$

$$\kappa(P, Q) \leq (1/2) \|u_0 + v_0\| = \chi(P, Q). \tag{16}$$

The inequality in (16) is obvious, but it is added for convenience. Let $c : \mathbb{X} \to \mathbb{R}$ be the cost function of the minimization problem (11), i.e.,

$$c(z) := \max\{\operatorname{dist}(z, P), \operatorname{dist}(z, -Q)\}.$$

Since z_0 minimizes $(1/2)c^2(\cdot)$ on $S_{\mathbb{X}}$, it satisfies the optimality condition

$$\lambda_1 (z_0 - \Pi_P(z_0)) + \lambda_2 \left(z_0 - \Pi_{-Q}(z_0)\right) + \mu z_0 = 0, \tag{17}$$

where $\mu \in \mathbb{R}$ is a Lagrange multiplier and $\lambda_1, \lambda_2 \in \mathbb{R}$ are nonnegative scalars adding up to 1. Since $z_0 - \Pi_P(z_0)$ is orthogonal to $\Pi_P(z_0)$ and $z_0 - \Pi_{-Q}(z_0)$ is orthogonal to $\Pi_{-Q}(z_0)$, one gets

$$\langle z_0, \Pi_P(z_0)\rangle = \langle z_0, \Pi_{-Q}(z_0)\rangle = s^2.$$

This and (17) yield $\mu + 1 = s^2$ and $\lambda_1 = \lambda_2 = 1/2$. Hence,

$$z_0 = (1/2s^2)\left(\Pi_P(z_0) + \Pi_{-Q}(z_0)\right) = (1/2s)(u_0 - v_0).$$

This proves (15) and shows that

$$z_0 \in \operatorname{cone}\left\{\Pi_P(z_0), \Pi_{-Q}(z_0)\right\} \subseteq L := \operatorname{span}\{u_0, v_0\}.$$

In the plane L, consider the rectangular triangles $\operatorname{co}\{0, z_0, \Pi_P(z_0)\}$ and $\operatorname{co}\{0, z_0, \Pi_{-Q}(z_0)\}$. Both triangles have the same angle at the vertex 0, namely, $\psi = \arcsin[\chi(P, Q)]$. It is plain to see that

$$\langle u_0, -v_0\rangle = \cos(2\psi) = 1 - 2\sin^2\psi = 1 - 2[\chi(P, Q)]^2.$$

This leads directly to the equality (16). Next, let (u_1, v_1) be an antipodal pair of (P, Q). Then $z_1 := \|u_1 - v_1\|^{-1}(u_1 - v_1)$ is well defined. From Proposition 3.1 one knows that

$$\begin{cases} \Pi_P(z_1) = \|u_1 - v_1\|^{-1}(1 - \langle u_1, v_1\rangle)\, u_1, \\ \Pi_Q(-z_1) = \|u_1 - v_1\|(1 - \langle u_1, v_1\rangle)\, v_1. \end{cases} \tag{18}$$

Hence,

$$\operatorname{dist}(z_1, P) = \|z_1 - \Pi_P(z_1)\| = \frac{\|v_1 - \langle u_1, v_1\rangle u_1\|}{\|u_1 - v_1\|} = (1/2)\,\|u_1 + v_1\| = \kappa(P, Q)$$

and, similarly, $\operatorname{dist}(z_1, -Q) = \operatorname{dist}(-z_1, Q) = \kappa(P, Q)$. It follows that

$$\chi(P, Q) \le c(z_1) = \kappa(P, Q). \tag{19}$$

From (18) one deduces also that

$$(u_1, v_1) = \left(\|\Pi_P(z_1)\|^{-1}\Pi_P(z_1), \|\Pi_Q(-z_1)\|^{-1}\Pi_Q(-z_1)\right). \tag{20}$$

The combination of (15)–(16) and (19)–(20) completes the proof of the theorem. □

4 Antipodality, pointedness and reproducibility

The sum of two closed convex cones may not be closed. The next proposition is part of the folklore on convex cones, cf. Beutner (2007, Theorem 3.2). A pair (P, Q) of elements in $\mathcal{C}(\mathbb{X})$ is said to be *pointed* if $P \cap -Q = \{0\}$. A single cone $K \in \mathcal{C}(\mathbb{X})$ is declared pointed if the pair (K, K) is pointed.

Proposition 4.1 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. The following conditions are equivalent and imply that $P + Q$ is closed:*

(a) *(P, Q) is pointed.*
(b) *There exists a positive constant β such that*

$$\beta(\|u\| + \|v\|) \leq \|u + v\| \quad \text{for all} \quad u \in P, v \in Q. \tag{21}$$

The reverse triangular inequality (21) holds of course with $\beta = 0$, but such choice is useless. What is interesting to know is the best constant

$$\beta(P, Q) := \max\{\beta \in [0, 1] : \beta \text{ satisfies } (21)\}.$$

Such a coefficient measures to which extent the pair (P, Q) is pointed. The next proposition relates $\beta(P, Q)$ to the maximal angle of (P, Q).

Proposition 4.2 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then $\beta(P, Q) = \cos[\Theta(P, Q)/2]$.*

Proof We need to prove that $\beta(P, Q) = \kappa(P, Q)$. Assume that P and Q are not equal to a common ray, otherwise we are done. Clearly,

$$\beta(P, Q) = \min_{\substack{u \in P,\, v \in Q \\ (u,v) \neq (0,0)}} \frac{\|u + v\|}{\|u\| + \|v\|}.$$

But the Dunkl–Williams inequality implies that

$$\frac{1}{2}\left\| \frac{u}{\|u\|} + \frac{v}{\|v\|} \right\| \leq \frac{\|u + v\|}{\|u\| + \|v\|}$$

whenever $u, v \in \mathbb{X}$ are nonzero vectors. Hence, $\beta(P, Q)$ is greater than or equal to $\kappa(P, Q)$. The reverse inequality is obvious. □

A pair (P, Q) of elements in $\mathcal{C}(\mathbb{X})$ is said to be *reproducing* if $P - Q = \mathbb{X}$. A single cone $K \in \mathcal{C}(\mathbb{X})$ is called reproducing if the pair (K, K) is reproducing. Clearly, (P, Q) is reproducing if and only if (P^*, Q^*) is pointed. The next result comes then without surprise.

Proposition 4.3 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. The following conditions are equivalent and imply that $P^* + Q^*$ is closed:*

(a) *(P, Q) is reproducing.*
(b) *There exists a positive constant α such that $\alpha B_{\mathbb{X}} \subseteq \mathrm{co}((P \cup -Q) \cap B_{\mathbb{X}})$.*

The notation "co" refers of course to the convex hull operation. One can see Proposition 4.3 as a sort of dual version of Proposition 4.1. The coefficient

$$\alpha(P, Q) := \max\{\alpha \in [0, 1] : \alpha B_{\mathbb{X}} \subseteq \mathrm{co}((P \cup -Q) \cap B_{\mathbb{X}})\} \tag{22}$$

measures to which extent the pair (P, Q) is reproducing. The next proposition shows that evaluating the reproducibility coefficient of (P, Q) amounts to compute the maximal angle of (P^*, Q^*).

Proposition 4.4 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then $\alpha(P, Q) = \cos\left[\Theta(P^*, Q^*)/2\right]$.*

Proof By using duality arguments (namely, calculus rules for polar sets) one can show that the inclusion in (22) can be written in the equivalent form

$$\alpha\left[(B_{\mathbb{X}} + P^*) \cap (B_{\mathbb{X}} - Q^*)\right] \subseteq B_{\mathbb{X}}.$$

Hence, $\alpha(P^*, Q^*)$ is equal to the coefficient

$$\nu(P, Q) := \max\left\{r \in [0, 1] : r\left[(B_{\mathbb{X}} + P) \cap (B_{\mathbb{X}} - Q)\right] \subseteq B_{\mathbb{X}}\right\}.$$

By proceeding as in Iusem and Seeger (2008b, Theorem 2), one can check that $\nu(P, Q) = \chi(P, Q)$. Theorem 3.2 does the rest of the job. □

5 Lipschitzness of the maximal angle function

Topological issues on $\mathcal{C}(\mathbb{X})$ are relative to the spherical metric δ, which is defined by

$$\delta(K_1, K_2) := \mathrm{haus}(K_1 \cap S_{\mathbb{X}}, K_2 \cap S_{\mathbb{X}}).$$

Here,

$$\mathrm{haus}(C_1, C_2) := \max\left\{\max_{x \in C_1} \mathrm{dist}(x, C_2), \max_{x \in C_2} \mathrm{dist}(x, C_1)\right\}$$

stands for the classical Pompeiu–Hausdorff distance between a pair C_1, C_2 of nonempty compact subsets of $\mathbb{X}$. Convergence with respect to the spherical metric is equivalent to convergence in the Painlevé–Kuratowski sense. Topological issues on the product space $\mathcal{C}^2(\mathbb{X}) := \mathcal{C}(\mathbb{X}) \times \mathcal{C}(\mathbb{X})$ refer to the metric

$$\Delta((P_1, Q_1), (P_2, Q_2)) := \max\left\{\delta(P_1, P_2), \delta(Q_1, Q_2)\right\}.$$

The next proposition concerns the continuity behavior of the multivalued map Γ. Upper and lower-semicontinuity of multivalued maps between metric spaces are understood in the classical sense, cf. Aubin and Frankowska (1990, Section 1.4).

Proposition 5.1 *The multivalued map $\Gamma : \mathcal{C}^2(\mathbb{X}) \rightrightarrows \mathbb{R}$ is upper-semicontinuous, but not lower-semicontinuous.*

Proof The values of Γ are closed subsets of the compact interval $[0, \pi]$. For proving that Γ is upper-semicontinuous, it is enough to check that

$$\mathrm{gr}\,\Gamma := \{(P, Q, \theta) \in \mathcal{C}^2(\mathbb{X}) \times \mathbb{R} : \theta \in \Gamma(P, Q)\}$$

is a closed set. Let $\{(P_k, Q_k, \theta_k)\}_{k\in\mathbb{N}}$ be a sequence in $\mathrm{gr}\,\Gamma$ converging to some $(P, Q, \theta) \in \mathcal{C}^2(\mathbb{X}) \times \mathbb{R}$. For each $k \in \mathbb{N}$, there exists a pair $(u_k, v_k) \in \mathbb{X}^2$ such that

$$\begin{cases} \theta_k = \arccos\langle u_k, v_k\rangle, \\ u_k \in P_k \cap S_{\mathbb{X}}, \quad v_k \in Q_k \cap S_{\mathbb{X}}, \\ v_k - \langle u_k, v_k\rangle u_k \in P_k^*, \\ u_k - \langle u_k, v_k\rangle v_k \in Q_k^*. \end{cases} \tag{23}$$

Let (u, v) be the limit of some subsequence $\{(u_{\varphi(k)}, v_{\varphi(k)})\}_{k\in\mathbb{N}}$. We write (23) with $\varphi(k)$ instead of k. By passing then to the limit, one deduces that $(P, Q, \theta) \in \mathrm{gr}\Gamma$. We now prove that Γ is not lower-semicontinuous. Let $e_1, e_2 \in S_{\mathbb{X}}$ be orthogonal. For each integer $k \geq 1$, let

$$\begin{aligned} u_k &:= -k^{-1}\, e_1 + \left(1 - k^{-2}\right)^{1/2} e_2, \\ P_k &= \vec{u}_k := \{t u_k : t \geq 0\}, \\ Q_k &= H_{e_1} := \{x \in \mathbb{X} : \langle e_1, x\rangle \geq 0\}. \end{aligned}$$

A matter of computation shows that $(u_k, -u_k)$ is the unique critical pair of (P_k, Q_k). Thus, $\Gamma(P_k, Q_k) = \{\pi\}$. On the other hand,

$$\Gamma\left(\lim_{k\to\infty} P_k, \lim_{k\to\infty} Q_k\right) = \Gamma(\vec{e}_2, H_{e_1}) = \{0, \pi\}.$$

This proves that Γ is not lower-semicontinuous. □

As shown in the next theorem, the maximal angle function $\Theta : \mathcal{C}^2(\mathbb{X}) \to \mathbb{R}$ is not merely continuous, but it is also Lipschitzian.

Theorem 5.2 *There exists a constant $\ell_{\mathbb{X}}$ such that*

$$|\Theta(P_1, Q_1) - \Theta(P_2, Q_2)| \leq \ell_{\mathbb{X}} \Delta((P_1, Q_1), (P_2, Q_2))$$

for all $(P_1, Q_1), (P_2, Q_2) \in \mathcal{C}^2(\mathbb{X})$.

Proof One knows from (10) that $\Theta : \mathcal{C}^2(\mathbb{X}) \to \mathbb{R}$ admits the characterization

$$\Theta(P, Q) = 2 \arccos\left(\kappa(P, Q)\right). \tag{24}$$

We claim that κ satisfies the Lipschitz condition

$$|\kappa(P_1, Q_1) - \kappa(P_2, Q_2)| \leq \Delta((P_1, Q_1), (P_2, Q_2)). \tag{25}$$

The proof of (25) follows the same pattern as in Iusem and Seeger (2008b, Lemma 1). Let $u_2 \in P_2 \cap S_{\mathbb{X}}$ and $v_2 \in Q_2 \cap S_{\mathbb{X}}$ be such that $2\kappa(P_2, Q_2) = \|u_2 + v_2\|$. Let u_1, v_1

be projections of u_2, v_2 onto $P_1 \cap S_{\mathbb{X}}$ and $Q_1 \cap S_{\mathbb{X}}$, respectively. Hence,

$$\begin{aligned} 2(\kappa(P_1, Q_1) - \kappa(P_2, Q_2)) &\leq \|u_1 + v_1\| - \|u_2 + v_2\| \\ &\leq \|u_1 - u_2\| + \|v_1 - v_2\| \\ &= \mathrm{dist}(u_2, P_1 \cap S_{\mathbb{X}}) + \mathrm{dist}(v_2, Q_1 \cap S_{\mathbb{X}}) \\ &\leq \mathrm{e}(P_2, P_1) + \mathrm{e}(Q_2, Q_1), \end{aligned}$$

where one uses the notation

$$\mathrm{e}(K_2, K_1) := \sup_{u \in K_2 \cap S_{\mathbb{X}}} \mathrm{dist}(u, K_1 \cap S_{\mathbb{X}}).$$

In a similar way one gets

$$2(\kappa(P_2, Q_2) - \kappa(P_1, Q_1)) \leq \mathrm{e}(P_1, P_2) + \mathrm{e}(Q_1, Q_2).$$

Thus,

$$\begin{aligned} 2|\kappa(P_1, Q_1) - \kappa(P_2, Q_2)| &\leq \max\{\mathrm{e}(P_2, P_1) + \mathrm{e}(Q_2, Q_1), \mathrm{e}(P_1, P_2) + \mathrm{e}(Q_1, Q_2)\} \\ &\leq \underbrace{\max\{\mathrm{e}(P_2, P_1), \mathrm{e}(P_1, P_2)\}}_{\delta(P_1, P_2)} + \underbrace{\max\{\mathrm{e}(Q_2, Q_1), \mathrm{e}(Q_1, Q_2)\}}_{\delta(Q_1, Q_2)}. \end{aligned}$$

This leads to (25). Next we observe that

$$\Theta(P, Q) = \arccos\left(1 - (1/2)[\mathrm{diam}(P, Q)]^2\right), \tag{26}$$

where

$$\mathrm{diam}(P, Q) := \max_{u \in P \cap S_{\mathbb{X}},\, v \in Q \cap S_{\mathbb{X}}} \|u - v\|.$$

It is not difficult to check that

$$\begin{aligned} |\mathrm{diam}(P_1, Q_1) - \mathrm{diam}(P_2, Q_2)| &\leq \delta(P_1, P_2) + \delta(Q_1, Q_2) \\ &\leq 2\,\Delta((P_1, Q_1), (P_2, Q_2)). \end{aligned} \tag{27}$$

The Lipschitzness of Θ is obtained by combining (24)–(27). To see this, one can follow the same procedure as in Seeger (2014, Theorem 2). The details are omitted. □

6 Critical angles in a pair of linear subspaces

Which is the minimal angle between a pair of linear subspaces? And which one is the maximal angle? Are they other interesting angles, besides the minimal and the maximal one? This sort of questions has lead to develop the classical theory of principal angles. Recall that the principal angles $\theta_1, \dots, \theta_m$ of a pair (P, Q) of nontrivial linear subspaces of $\mathbb{X}$ are defined recursively by

$$\cos\theta_k = \max_{u \in P_k \cap S_{\mathbb{X}}, v \in Q_k \cap S_{\mathbb{X}}} \langle u, v \rangle, \tag{28}$$

where $m := \min\{\dim P, \dim Q\}$ and

$$\begin{cases} P_1 := P, \quad Q_1 := Q, \\ P_{k+1} := \{x \in P_k : \langle u_k, x \rangle = 0\}, \\ Q_{k+1} := \{x \in Q_k : \langle v_k, x \rangle = 0\}, \\ (u_k, v_k) \text{ solution to (28).} \end{cases}$$

The vectors u_k and v_k are not uniquely defined, but the θ_k are unique. Interesting material on principal angles can be found in the linear algebra book by Meyer (2000, Section 5.15), see also the references Björck and Golub (1973), Miao and Ben-Israel (1992) and Roy (1947). When P and Q are nontrivial linear subspaces of $\mathbb{X}$, the system (4) becomes

$$u \in P \cap S_{\mathbb{X}}, \quad v \in Q \cap S_{\mathbb{X}}, \quad v - \langle u, v \rangle u \in P^{\perp}, \quad u - \langle u, v \rangle v \in Q^{\perp}, \tag{29}$$

where $\perp$ indicates orthogonal complementation relative to $\mathbb{X}$. In this special context, there is no distinction between criticality for angle maximization and criticality for angle minimization. As a first elementary observation, we mention the following conjugacy principle.

Proposition 6.1 *Let P and Q be nontrivial linear subspaces of $\mathbb{X}$. Let $\theta, \psi \in [0, \pi]$ be conjugate angles. Then θ is a critical angle of (P, Q) if and only if ψ is a critical angle of (P, Q).*

Proof Clearly, (u, v) satisfies (29) if and only if $(u, -v)$ satisfies (29). It suffices now to observe that $\arccos\langle u, -v \rangle$ and $\arccos\langle u, v \rangle$ are conjugate angles. □

The combination of Corollary 2.2 and Proposition 6.1 yields a duality result established by Miao and Ben-Israel (1992, Theorem 3). In view of Proposition 6.1, it is enough to compute the critical angles of (P, Q) that are in the subinterval $[0, \pi/2]$. The remaining critical angles are obtained by conjugation. The next theorem shows that the principal angles of (P, Q) are equal to the critical angles of (P, Q) that are in $[0, \pi/2]$. In what follows, we use the notation

$$\mathcal{O}(\mathbb{R}^n, \mathbb{X}) := \{W \in \mathcal{L}(\mathbb{R}^n, \mathbb{X}) : W^T W = I_n\},$$

where I_n is the identity matrix of order n and $\mathcal{L}(\mathbb{R}^n, \mathbb{X})$ is the vector space of linear maps from $\mathbb{R}^n$ to $\mathbb{X}$. The symbol $\operatorname{Im} W$ refers to the image space of W.

Theorem 6.2 *Let $P = \operatorname{Im} U$ and $Q = \operatorname{Im} V$ be nontrivial linear subspaces of $\mathbb{X}$ represented by $U \in \mathcal{O}(\mathbb{R}^p, \mathbb{X})$ and $V \in \mathcal{O}(\mathbb{R}^q, \mathbb{X})$, respectively. For $\theta \in [0, \pi/2]$, the following four conditions are equivalent:*

(a) *θ is a critical angle of (P, Q).*
(b) *θ is a principal angle of (P, Q).*
(c) *$\cos\theta$ is a singular value of the rectangular matrix $E := V^T U$.*
(d) *There are unit vectors $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ such that*

$$\begin{bmatrix} 0 & E^T \\ E & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \cos\theta \begin{bmatrix} x \\ y \end{bmatrix}. \tag{30}$$

Furthermore, if x and y are as in (d)*, then (Ux, Vy) is a critical pair of (P, Q) and $\cos\theta = \langle Ux, Vy\rangle$.*

Proof (b) $\Leftrightarrow$ (c). This equivalence is stated in Björck and Golub (1973, Theorem 1).
(a) $\Rightarrow$ (d). Let the angle θ be formed with some pair (u, v) satisfying the system (29). There are unit vectors $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ such that $(u, v) = (Ux, Vy)$. Hence, $\cos\theta = \langle u, v\rangle = \langle Ux, Vy\rangle$ and

$$Vy - (\cos\theta)Ux \in P^\perp, \quad Ux - (\cos\theta)Vy \in Q^\perp.$$

Since $P^\perp = \mathrm{Ker}(U^T)$ and $Q^\perp = \mathrm{Ker}(V^T)$, one gets

$$E^T y = (\cos\theta)x, \quad Ex = (\cos\theta)y. \tag{31}$$

which is an equivalent way of writing (30).
(d) $\Rightarrow$ (c). By exchanging the roles of P and Q if necessary, one may suppose that $\min\{p, q\} = p$. Let x and y be as in (d). From (31), one gets $E^T Ex = (\cos\theta)^2 x$. Hence, $(\cos\theta)^2$ is an eigenvalue of $E^T E$ and $\cos\theta$ is a singular value of E.
(c) $\Rightarrow$ (a). One can write $E = F\Sigma G^T$, where $G = [g_1 \ldots g_p]$ and $F = [f_1 \ldots f_q]$ are orthogonal matrices of order p and q, respectively, and Σ is a $q \times p$ diagonal matrix with the singular values of E placed on the diagonal entries. Let $\cos\theta$ be a singular value of E. Suppose that $\cos\theta$ is placed on kth diagonal entry of Σ. One gets $Eg_k = (\cos\theta) f_k$ and $E^T f_k = (\cos\theta)g_k$. Hence, the system (30) holds with $x = g_k$ and $y = f_k$. One deduces that $(u, v) = (Ug_k, Vf_k)$ is a critical pair of (P, Q) producing the angle θ. □

Let P and Q be as in Theorem 6.2. By combining Theorem 6.2 and Proposition 6.1, one obtains

$$\Gamma(P, Q) = \bigcup_{\sigma\in\Sigma(E)} \{\arccos\sigma, \pi - \arccos\sigma\},$$

where $\Sigma(E)$ is the set of singular values of E. In particular, the critical angles of (P, Q) are at most $2\min\{p, q\}$ and they come in conjugate pairs. Such upper bound can be sharpened as follows.

Proposition 6.3 *Let P and Q be nontrivial linear subspaces of $\mathbb{X}$ of dimensions p and q, respectively. Let r be the dimension of $P \cap Q$. Then*

$$\mathrm{card}\,[\Gamma(P, Q)] \le 2\min\{p, q\} - 2\max\{0, r - 1\}. \tag{32}$$

Proof We assume that $1 \leq r < \min\{p, q\}$, otherwise we are done. One can represent P and Q as in Theorem 6.2, with the additional feature that $U = [W, U_0]$ and $V = [W, V_0]$ have a portion in common. The common part $W \in \mathcal{O}(\mathbb{R}^r, \mathbb{X})$ serves to represent the intersection of P and Q, i.e., $P \cap Q = \operatorname{Im} W$. One has

$$\begin{aligned} W^T W = I_r, \quad U_0^T U_0 = I_{p-r}, \quad V_0^T V_0 = I_{q-r}, \\ W^T U_0 = 0, \quad U_0^T W = 0, W^T V_0 = 0, \quad V_0^T W = 0. \end{aligned}$$

Let the angle θ be formed with some pair (u, v) satisfying the system (29). There are vectors $\xi, \eta \in \mathbb{R}^r$ and $x \in \mathbb{R}^{p-r}$, $y \in \mathbb{R}^{q-r}$ such that $\|\xi\|^2 + \|x\|^2 = 1$, $\|\eta\|^2 + \|y\|^2 = 1$, and

$$(u, v) = (W\xi + U_0 x, W\eta + V_0 y). \tag{33}$$

By substituting (33) into (29), one gets

$$\begin{cases} W^T[W\eta + V_0 y - \lambda(W\xi + U_0 x)] = 0, \\ U_0^T[W\eta + V_0 y - \lambda(W\xi + U_0 x)] = 0, \\ W^T[W\xi + U_0 x - \lambda(W\eta + V_0 y)] = 0, \\ V_0^T[W\xi + U_0 x - \lambda(W\eta + V_0 y)] = 0, \end{cases}$$

with $\lambda = \langle W\xi + U_0 x, W\eta + V_0 y\rangle$. After simplification, one obtains $\eta = \lambda\xi$, $\xi = \lambda\eta$, and

$$U_0^T V_0 y = \lambda x, \quad V_0^T U_0 x = \lambda y. \tag{34}$$

If $\lambda \notin \{-1, 1\}$, then $\xi = 0$, $\eta = 0$ and $x \in \mathbb{R}^{p-q}$, $y \in \mathbb{R}^{q-r}$ are unit vectors satisfying (34). Hence, λ may take at most $2\min\{p - r, q - r\}$ different values. To this count one should add the potential candidates $\lambda = -1$ and $\lambda = 1$. One gets in this way the upper estimate

$$\operatorname{card}[\Gamma(P, Q)] \leq 2\min\{p - r, q - r\} + 2.$$

This proves (32). $\square$

7 Critical angles in a pair of polyhedral cones

This section is devoted to the analysis of critical angles in a pair (P, Q) of polyhedral cones. We suppose that the reader is acquainted with the theory of faces of convex polyhedra. The notation that we use is as follows:

$$\begin{aligned} \mathcal{F}(P) &:= \{F \subseteq \mathbb{X} : F \text{ is a nonzero face of } P\}, \\ \operatorname{span} F &:= \text{linear subspace spanned by } F, \\ \operatorname{ri} F &:= \text{relative interior of } F, \\ \dim F &:= \text{dimension of } \operatorname{span} F, \\ \Pi^F &:= \text{orthogonal projector onto } \operatorname{span} F. \end{aligned}$$

For a nonzero vector u in a polyhedral cone $P \in \mathcal{C}(\mathbb{X})$, there exists a unique $F \in \mathcal{F}(P)$ such that $u \in \mathrm{ri}\, F$. Such F is called the face associated to u.

Theorem 7.1 *Let $P, Q \in \mathcal{C}(\mathbb{X})$ be polyhedral cones. If (u, v) is a critical pair of (P, Q), then*

$$\Pi^F v = \langle u, v\rangle u, \quad \Pi^E u = \langle u, v\rangle v, \tag{35}$$

where F is the face of P associated to u and E is the face of Q associated to v. In particular,

$$\Gamma(P, Q) \subseteq \bigcup_{F\in\mathcal{F}(P)} \bigcup_{E\in\mathcal{F}(Q)} \Gamma(\mathrm{span}\, F, \mathrm{span}\, E). \tag{36}$$

Proof By assumption, $u \in \mathrm{ri}\, F$ and $v \in \mathrm{ri}\, E$ satisfy the criticality conditions stated in (4). By proceeding as Seeger and Torki (2003, Theorem 3.4), one can check that

$$\begin{cases} v - \langle u, v\rangle u \in (\mathrm{span}\, F)^{\perp}, \\ u - \langle u, v\rangle v \in (\mathrm{span}\, E)^{\perp}. \end{cases}$$

But this is clearly equivalent to (35). □

By using the inclusion (36), one gets the upper bound

$$\mathrm{card}[\Gamma(P, Q)] \leq \sum_{F\in\mathcal{F}(P)} \sum_{E\in\mathcal{F}(Q)} \mathrm{card}[\Gamma(\mathrm{span}\, F, \mathrm{span}\, E)]. \tag{37}$$

The above inequality becomes an equality, for instance, if P and Q are nontrivial linear subspaces. Since the double sum in (37) is finite, any pair of polyhedral cones has finitely many critical angles. By combining (37) and the estimate

$$\mathrm{card}[\Gamma(\mathrm{span}\, F, \mathrm{span}\, E)] \leq 2 \min\{\dim F, \dim E\},$$

one gets in particular

$$\mathrm{card}\,[\Gamma(P, Q)] \leq 2 \sum_{k=1}^{\dim P} \sum_{\ell=1}^{\dim Q} c_P(k) c_Q(\ell) \min\{k, \ell\}, \tag{38}$$

where $c_P(k)$ stands for the number of k-dimensional faces of P. The upper bound (38) is coarse in general, so we shall not elaborate further on the practical evaluation of such an expression.

Corollary 7.2 *Let $P, Q \in \mathcal{C}(\mathbb{X})$ be polyhedral cones. Let (u_1, v) and (u_2, v) be critical pairs of (P, Q). If u_1 and u_2 have the same associated face, then the critical angles $\theta_1 := \arccos\langle u_1, v\rangle$ and $\theta_2 := \arccos\langle u_2, v\rangle$ are equal or conjugate.*

Proof Suppose that u_1 and u_2 have $F \in \mathcal{F}(P)$ as common associated face. In such a case, Theorem 7.1 yields $\langle u_1, v\rangle u_1 = \Pi^F v = \langle u_2, v\rangle u_2$. By taking norms, one sees that $\langle u_1, v\rangle$ and $\langle u_2, v\rangle$ have the same absolute value. This proves that θ_1 and θ_2 are equal or conjugate. □

We now concentrate on the numerical computation of the critical angles of a pair

$$(P, Q) = (G(\mathbb{R}^p_+), H(\mathbb{R}^q_+)) \tag{39}$$

of polyhedral cones in $\mathbb{R}^n$. Here, $G = [g_1, \ldots, g_p]$ and $H = [h_1, \ldots, h_q]$ are matrices of size $n \times p$ and $n \times q$, respectively. Without loss of generality, we assume that

$$\begin{cases} g_1, \ldots, g_p \text{ are conically independent unit vectors,} \\ h_1, \ldots, h_q \text{ are conically independent unit vectors.} \end{cases} \tag{40}$$

That a collection of vectors is conically independent simply means that no element from the collection can be expressed as positive linear combination of those remaining. For notational convenience, we introduce the index sets

$$\mathcal{I}(G) := \{I \subseteq \{1, \ldots, p\} : I \neq \emptyset \quad \text{and} \quad \{g_i : i \in I\} \text{ is linearly independent}\},$$
$$\mathcal{J}(H) := \big\{J \subseteq \{1, \ldots, q\} : J \neq \emptyset \quad \text{and} \quad \{h_j : j \in J\} \text{ is linearly independent}\big\}.$$

The cardinality of an index set, say I, is denoted by $|I|$. We write G_I to indicate the submatrix of G with columns indexed by I. The definition of H_J is similar. Without further ado, we state the next theorem.

Theorem 7.3 *Let (P, Q) be as in (39)–(40). Then the following statements are equivalent:*

(a) $\theta \in \Gamma(P, Q)$,
(b) *there are index sets $I \in \mathcal{I}(G)$, $J \in \mathcal{J}(H)$ and vectors $\xi \in \mathbb{R}^{|I|}$, $\eta \in \mathbb{R}^{|J|}$ such that*

$$\begin{bmatrix} 0 & G_I^T H_J \\ H_J^T G_I & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \cos\theta \begin{bmatrix} G_I^T G_I & 0 \\ 0 & H_J^T H_J \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \tag{41}$$
$$\langle g_k, H_J\eta - (\cos\theta) G_I \xi\rangle \geq 0 \quad \text{for all} \quad k \notin I, \tag{42}$$
$$\langle h_\ell, G_I\xi - (\cos\theta) H_J \eta\rangle \geq 0 \quad \text{for all} \quad \ell \notin J, \tag{43}$$
$$\langle \xi, G_I^T G_I \xi\rangle = 1, \quad \xi \in \mathrm{int}(\mathbb{R}^{|I|}_+), \tag{44}$$
$$\langle \eta, H_J^T H_J \eta\rangle = 1, \quad \eta \in \mathrm{int}(\mathbb{R}^{|J|}_+). \tag{45}$$

Furthermore, when these equivalent statements hold, the critical angle θ is formed with the critical pair $(u, v) = (G_I\xi, H_J\eta)$.

Proof We follow similar steps as in Iusem and Seeger (2008b, Theorem 3), except that now the polyhedral cones P and Q are not necessarily equal. Besides, we do not restrict the attention to proper critical angles. For the sake of completeness, we give a sketch of the proof:
(a) $\Rightarrow$ (b). Let (u, v) be a critical pair of (P, Q) such that $\lambda := \langle u, v\rangle = \cos\theta$. The cone version of Caratheodory's theorem ensures the existence of index sets $I \in \mathcal{I}(G)$, $J \in \mathcal{J}(H)$, and vectors $\xi \in \mathbb{R}^{|I|}$, $\eta \in \mathbb{R}^{|J|}$ with positive components, such that

$u = G_I\xi$ and $v = H_J\eta$. The normalization conditions in (44)–(45) are obtained from the fact that u and v are unit vectors. Criticality of (u, v) leads to the system

$$\begin{cases} H_J\eta - \lambda G_I\xi \in P^*, \\ G_I\xi - \lambda H_J\eta \in Q^*, \end{cases}$$

or, equivalently,

$$\begin{aligned} \langle g_k, H_J\eta - \lambda G_I\xi\rangle &\geq 0 \quad \text{for all} \quad k = 1, \dots, p, \\ \langle h_\ell, G_I\xi - \lambda H_J\eta\rangle &\geq 0 \quad \text{for all} \quad \ell = 1, \dots, q. \end{aligned}$$

This yields (42) and (43). Furthermore, since

$$\begin{aligned} 0 = \langle u, v - \lambda u\rangle &= \langle \xi, G_I^T H_J\eta - \lambda G_I^T G_I\xi\rangle, \\ 0 = \langle v, u - \lambda v\rangle &= \langle \eta, H_J^T G_I\xi - \lambda H_J^T H_J\eta\rangle, \end{aligned}$$

one gets

$$\begin{aligned} G_I^T H_J\eta - \lambda G_I^T G_I\xi &= 0, \\ H_J^T G_I\xi - \lambda H_J^T H_J\eta &= 0, \end{aligned}$$

which is nothing but (41).
(b) $\Rightarrow$ (a). If one sets $(u, v) := (G_I\xi, H_J\eta)$, then one can check that (u, v) is a critical pair of (P, Q) with $\cos\theta = \langle u, v\rangle$. □

The index sets I, J and the vectors ξ, η in Theorem 7.3(b) are not necessarily unique. Anyway, one can write

$$\Gamma(P, Q) = \bigcup_{I\in\mathcal{I}(G)} \bigcup_{J\in\mathcal{J}(H)} \Gamma_{I,J}(P, Q),$$

where $\Gamma_{I,J}(P, Q)$ captures the critical angles produced by (I, J), that is,

$$\Gamma_{I,J}(P, Q) := \{\arccos\langle G_I\xi, H_J\eta\rangle : (\xi, \eta) \text{ as in (41)–(45)}\}.$$

One refers to (I, J) as a *successful configuration of index sets* if $\Gamma_{I,J}(P, Q)$ is nonempty. For each pair (I, J), we construct $\Gamma_{I,J}(P, Q)$ by using the following algorithm:

– Step 1: Solve the generalized eigenvalue problem $A^{I,J}z = \lambda\, B^{I,J}z$, where

$$A^{I,J} := \begin{bmatrix} 0 & G_I^T H_J \\ H_J^T G_I & 0 \end{bmatrix}, \ B^{I,J} := \begin{bmatrix} G_I^T G_I & 0 \\ 0 & H_J^T H_J \end{bmatrix}, \ z := \begin{bmatrix} \xi \\ \eta \end{bmatrix}.$$

– Step 2: Declare "acceptable" each eigenvalue that admits an associated eigenvector satisfying the conditions (42)–(45), where one identifies λ with $\cos\theta$. Take the arccosinus of each acceptable eigenvalue and put it in the set $\Gamma_{I,J}(P, Q)$.

Table 1 Critical angles between the nonnegative orthant and the Schur cone in $\mathbb{R}^5$

I	J	$\cos\theta$	θ
$\{2,3,4,5\}$	$\{1,2,3,4\}$	$-1/\sqrt{5}$	$0.6476\,\pi$
$\{2,3,4\}$	$\{1,2,3\}$	$-1/2$	$0.6667\,\pi$
$\{2,3\}$	$\{1,2\}$	$-1/\sqrt{3}$	$0.6959\,\pi$
$\{2,4,5\}$	$\{1,2,3,4\}$	$-\sqrt{2}/\sqrt{5}$	$0.7180\,\pi$
$\{3,4,5\}$	$\{1,2,3,4\}$	$-\sqrt{2}/\sqrt{5}$	$0.7180\,\pi$
$\{2\}$	$\{1\}$	$-1/\sqrt{2}$	$0.7500\,\pi$
$\{3,4\}$	$\{1,2,3\}$	$-1/\sqrt{2}$	$0.7500\,\pi$
$\{3,5\}$	$\{1,2,3,4\}$	$-\sqrt{3}/\sqrt{5}$	$0.7820\,\pi$
$\{4,5\}$	$\{1,2,3,4\}$	$-\sqrt{3}/\sqrt{5}$	$0.7820\,\pi$
$\{3\}$	$\{1,2\}$	$-\sqrt{2}/\sqrt{3}$	$0.8041\,\pi$
$\{4\}$	$\{1,2,3\}$	$-\sqrt{3}/2$	$0.8333\,\pi$
$\{5\}$	$\{1,2,3,4\}$	$-2/\sqrt{5}$	$0.8524\,\pi$

Example 7.4 By way of example, consider the nonnegative orthant $P = \mathbb{R}^n_+$ and the Schur cone

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

In this case, $G = I_n$ and H is formed with the n-dimensional vectors

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

Table 1 concerns the particular case $n = 5$. It displays the successful configurations (I, J) and the critical angles produced by each one of these configurations. As one can see from Table 1, different configurations (I, J) may produce the same critical angle. There are 465 configurations (I, J) in all, but only 12 are successful. It is not surprising that all the critical angles reported in Table 1 are obtuse. This corresponds to a general fact concerning critical angles between nonnegative orthants and Schur cones.

Remark 7.5 Let (P, Q) be as in (39)–(40). Theorem 7.3 yields the following cardinality estimate

$$\mathrm{card}[\Gamma(P, Q)] \leq \sum_{k=1}^{n} \sum_{\ell=1}^{n} C_k^p C_\ell^q (k + \ell),$$

where

$$C_k^p := \begin{cases} \frac{p!}{k!(p-k)!} & \text{if} \quad k \leq p, \\ 0 & \text{if} \quad k > p. \end{cases} \tag{46}$$

We end this section by considering the critical angles between a polyhedral cone P and a ray $\vec{v} := \{tv : t \geq 0\}$ generated by some unit vector $v \in \mathbb{R}^n$. By adapting Theorem 7.3 to this special setting, one gets the following result.

Proposition 7.6 *Let P be as in (39)–(40) and v be a unit vector in $\mathbb{R}^n$. Then, $\theta \in \Gamma(P, \vec{v})$ if and only if there are an index set $I \in \mathcal{I}(G)$ and a vector $\xi \in \mathbb{R}^{|I|}$ such that*

$$
\begin{aligned}
(\cos\theta)\xi &= (G_I^T G_I)^{-1} G_I^T v, \\
\langle g_k, v - (\cos\theta) G_I \xi \rangle &\geq 0 \quad \text{for all} \quad k \notin I, \\
\langle \xi, G_I^T G_I \xi \rangle &= 1, \quad \xi \in \operatorname{int}(\mathbb{R}_+^{|I|})
\end{aligned}
\tag{47}
$$

Furthermore, the critical pairs of $(P, \vec{v})$ are exactly those of the form $(G_I \xi, v)$, with I and ξ as above.

Proof The conditions (43) and (45) are here superfluous. The generalized eigenvalue problem (41) becomes

$$
\begin{bmatrix} 0 & G_I^T v \\ v^T G_I & 0 \end{bmatrix} \begin{bmatrix} \xi \\ 1 \end{bmatrix} = \cos\theta \begin{bmatrix} G_I^T G_I & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi \\ 1 \end{bmatrix},
$$

but this is equivalent to (47). □

The next corollary shows that, for each $I \in \mathcal{I}(G)$, the set $\Gamma_I(P, \vec{v})$ is empty or a singleton. The proof of such result is omitted, because it follows straightforwardly from Proposition 7.6.

Corollary 7.7 *Let P be as in (39)–(40) and v be a unit vector in $\mathbb{R}^n$. For each $I \in \mathcal{I}(G)$, consider the condition*

$$
\langle g_k, v - v^I \rangle \geq 0 \quad \textit{for all} \quad k \notin I,
\tag{48}
$$

where $v^I := G_I (G_I^T G_I)^{-1} G_I^T v$ is the orthogonal projection of v onto $\operatorname{Im} G_I$. *Then*

$$
\Gamma_I(P, \vec{v}) = \begin{cases} \{\pi/2\} & \text{if } G_I^T v = 0 \text{ and (48) holds}, \\ \left\{\arccos \|v^I\|\right\} & \text{if } (G_I^T G_I)^{-1} G_I^T v \in \operatorname{int}(\mathbb{R}_+^{|I|}) \text{ and (48) holds}, \\ \left\{\pi - \arccos \|v^I\|\right\} & \text{if } -(G_I^T G_I)^{-1} G_I^T v \in \operatorname{int}(\mathbb{R}_+^{|I|}) \text{ and (48) holds}, \\ \emptyset & \text{otherwise}. \end{cases}
$$

In particular, $\operatorname{card}[\Gamma(P, \vec{v})] \leq \sum_{k=1}^{n} C_k^p$ *with C_k^p given by (46).*

References

Aubin J-P, Frankowska H (1990) Set-valued analysis. Birkhäuser, Boston

Beutner E (2007) On the closedness of the sum of closed convex cones in reflexive Banach spaces. J Conv Anal 14:99–102

Björck A, Golub GH (1973) Numerical methods for computing angles between linear subspaces. Math Comput 27:579–594

Clarke FH, Ledyaev YS, Stern RJ (1997) Complements, approximations, smoothings and invariance properties. J Conv Anal 4:189–219

Clarke FH, Ledyaev YS, Stern RJ (1999) Invariance, monotonicity, and applications. In: Nonlinear analysis, differential equations and control. NATO Sci Ser C Math Phys Sci, vol 528. Kluwer, Dordrecht, pp 207–305

Drusvyatskiy D (2013) Slope and geometry in variational mathematics. Ph.D. thesis, Cornell University

Goldberg F, Shaked-Monderer N (2014) On the maximal angle between copositive matrices. Electron J Linear Algebra 27:837–850

Gourion D, Seeger A (2010) Critical angles in polyhedral convex cones: numerical and statistical considerations. Math Program 123:173–198

Iusem A, Seeger A (2005) On pairs of vectors achieving the maximal angle of a convex cone. Math Program 104:501–523

Iusem A, Seeger A (2007) Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. Comput Appl Math 26:191–214

Iusem A, Seeger A (2007) On convex cones with infinitely many critical angles. Optimization 56:115–128

Iusem A, Seeger A (2008) Antipodal pairs, critical pairs, and Nash angular equilibria in convex cones. Optim Meth Softw 23:73–93

Iusem A, Seeger A (2008) Normality and modulability indices. II. Convex cones in Hilbert spaces. J Math Anal Appl 338:392–406

Iusem A, Seeger A (2009) Searching for critical angles in a convex cone. Math Program 120:3–25

Lewis AS, Luke DR, Malick J (2009) Local linear convergence for alternating and averaged nonconvex projections. Found Comput Math 9:485–513

Meyer C (2000) Matrix analysis and applied linear algebra. SIAM Publications, Philadelphia

Miao JM, Ben-Israel A (1992) On principal angles between subspaces in R^n. Linear Algebra Appl 171:81–98

Obert DG (1991) The angle between two cones. Linear Algebra Appl 144:63–70

Peña J, Renegar J (2000) Computing approximate solutions for convex conic systems of constraints. Math Program Ser A 87:351–383

Roy SN (1947) A note on critical angles between two flats in hyperspace with certain statistical applications. Sankhya 8:177–194

Seeger A (2014) Lipschitz and Hölder continuity results for some functions of cones. Positivity 18:505–517

Seeger A, Torki M (2003) On eigenvalues induced by a cone constraint. Linear Algebra Appl 372:181–206

Tenenhaus M (1988) Canonical analysis of two convex polyhedral cones and applications. Psychometrika 53:503–524