Facial reduction for nice (and non-nice) convex programs

Ying Lin[*] Tianxiang Liu[†] Bruno F. Lourenço[‡]

December 17, 2025

Abstract

Consider the primal problem of minimizing the sum of two closed proper convex functions f and g. If the relative interiors of the domains of f and g intersect, then the primal problem and its corresponding Fenchel dual satisfy strong duality. When these relative interiors fail to intersect, pathologies and numerical difficulties may occur. In this paper, we propose a facial reduction algorithm that outputs minimal faces of $\operatorname{dom} f$ and $\operatorname{dom} g$ containing the feasible region so that we may reformulate the primal problem in such a way that Slater's condition is ensured and strong duality is always satisfied. More generally, given two nonempty convex sets C_1, C_2, our facial reduction algorithm either finds minimal faces of C_1 and C_2 containing the intersection $C_1 \cap C_2$ or certifies that the intersection is empty. Along the way, we extend the notion of niceness from convex cones to general convex set and consider a suitable notion of niceness for convex functions called *vertical niceness*. Under these niceness assumptions, Fenchel duals arising from the application of our facial reduction algorithm have simplified expressions. Finally, inspired by Ramana's extended dual for semidefinite programming, we then use these tools to develop extended duals for general convex problems.

1 Introduction

Let $\mathcal{E}$ be a finite dimensional Euclidean space and consider the problem of minimizing the sum of two closed proper convex function $f, g : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$:

$$\inf f(x) + g(x). \tag{1.1}$$

The Fenchel dual of (1.1) is given by

$$\sup_{\lambda} -f^*(-\lambda) - g^*(\lambda), \tag{1.2}$$

where f^* and g^* denote the conjugate function of f and g respectively.

[*]Department of Data and Systems Engineering, The University of Hong Kong, Hong Kong, China. (ylin95@hku.hk)

[†]Institute of Systems and Information Engineering, University of Tsukuba, Japan. This author was supported partly by JSPS KAKENHI (Grant Number: 25K15000). (liutx@sk.tsukuba.ac.jp)

[‡]Department of Fundamental Statistical Mathematics, Institute of Statistical Mathematics, Japan. This author was supported partly by the JSPS Grant-in-Aid for Early-Career Scientists 23K16844. (bruno@ism.ac.jp)

When the relative interiors of the domains of f and g intersect (i.e., $\mathrm{ri}\,(\mathrm{dom}\, f) \cap \mathrm{ri}\,(\mathrm{dom}\, g) \neq \emptyset$), the optimal values of (1.1) and (1.2) coincide and the latter is *attained* in the sense that there exists λ^* such that $-f^*(-\lambda^*) - g^*(\lambda^*)$ equals the common optimal value (even if it is $-\infty$), see [35, Theorem 31.1].

The case where $\mathrm{ri}\,(\mathrm{dom}\, f) \cap \mathrm{ri}\,(\mathrm{dom}\, g)$ is empty is typically less stable numerically. Furthermore, the primal and dual optimal value may differ, which is undesirable and may lead to further difficulties both theoretical and numerical.

Motivated by the facial reduction algorithm in conic convex optimization [4, 3], the main goal of this paper is to develop a new duality theory for (1.1) that has desirable properties even when the intersection $\mathrm{ri}\,(\mathrm{dom}\, f) \cap \mathrm{ri}\,(\mathrm{dom}\, g)$ is empty. Along the way, we will introduce a facial reduction algorithm appropriate for handling (1.1) and new notions such as *nice convex sets* and *vertically nice functions* that will help us achieve our goal.

1.1 Previous works

Let $\mathcal{K} \subseteq \mathcal{E}$ be a closed convex cone. The case where the problem under consideration is minimizing a convex function under a constraint of the form $\{x \in \Omega \mid h(x) \in -\mathcal{K}\}$, where Ω is a convex set and h is a $\mathcal{K}$-convex function was discussed extensively by Borwein and Wolkowicz in [4, 3], which are foundational papers behind what is known as *facial reduction*.

Now, let $c \in \mathcal{E}$ and let $\mathcal{V} \subseteq \mathcal{E}$ be a nonempty affine space. Denoting by $\delta_{\mathcal{V}}$ the indicator function of $\mathcal{V}$, suppose we let f be such that $f(x) = \langle c, x\rangle + \delta_{\mathcal{V}}$ and $g = \delta_{\mathcal{K}}$. Then, the pair (1.1) and (1.2) becomes primal and dual *conic linear programs* (CLPs). The condition that $\mathrm{ri}\,(\mathrm{dom}\, f) \cap \mathrm{ri}\,(\mathrm{dom}\, g)$ be nonempty becomes the classical *Slater's condition* (i.e., $\mathcal{V} \cap \mathrm{ri}\,(\mathcal{K}) \neq \emptyset$). Slater's condition play an important role in ensuring good duality properties and stability of numerical methods for conic linear programs. If Slater's condition is not satisfied, numerical methods may misbehave, e.g., see [46].

The goal of facial reduction is to reformulate a problem into a new equivalent one for which Slater's condition is satisfied, which, hopefully, improves the numerical and theoretical properties of the problem. In the context of conic linear programming, several works revisited Borwein and Wolkowicz facial reduction algorithm [30, 45, 19, 20]. The basic idea is that whenever a problem does not satisfy Slater's condition, it is because the feasible region is contained in a proper face of the underlying cone. If this proper face is chosen to be *minimal*, then Slater's condition is restored and a new equivalent problem is obtained that has improved duality properties. Facial reduction has inspired many approaches for preprocessing CLPs [32, 47] and is an important tool for simplifying CLPs that arise as relaxations of combinatorial problems.

In order to apply facial reduction, given $\mathcal{V}$ and $\mathcal{K}$, we need to *explicitly* compute such a minimal face (either analytically or numerically) and, based on that, we obtain a new dual problem that has better duality properties. However, there are *implicit* approaches in the sense that based purely on $\mathcal{V}$ and $\mathcal{K}$, one is able to write down a so-called extended dual that has good duality properties. A pioneering approach along this line was Ramana's famous extended dual for semidefinite programming [33]. If, say, $\mathcal{K}$ is the cone of $n \times n$ real symmetric positive semidefinite matrices $\mathcal{S}^n_+$ and $\mathcal{V}$ is represented as the solution set of some linear system $\mathcal{A}x = b$ then Ramana's dual can be written down purely in terms of $\mathcal{A}, b, c$ and n, without the need of explicitly determining the minimal face beforehand. It turns out that Ramana's dual is strongly connected to facial reduction as discussed in [34, 30, 30, 19], see also related discussions in [13, 18, 21, 31].

Another related topic is the study of classes of cones that are particularly suitable for facial reduction. Two such classes are the so-called *nice* cones and *projectionally exposed* cones. Both conditions were considered already in [4], but the name "nice" was coined later in [28]. It turns out that nice closed convex cones are facially exposed [29], but the converse is not true in general [37]. Sitting between nice and projectionally exposed cones is the class of *amenable* cones [23]. Explorations on these cone classes and beyond can be found, e.g., in [38, 22].

1.2 Our contributions

In principle, any convex optimization problem (1.1) can be reformulated as a conic linear program over an appropriate cone, which can then be solved numerically if such a cone is among those supported by some solver [43, 44, 41, 8, 5, 9, 25, 12, 26]. However conic reformulations may increase the dimension of the problem. For example, the Ky-Fan k-norm which corresponds to the sum of the largest k singular values of a $n \times m$ matrix is SDP-representable but one of the most common SDP representations is not efficient in the sense that it requires SDP constraints over $(n+m) \times (n+m)$ matrices [2, Section 4.2, 18c and 19]. There are also other inconveniences, e.g., we may have closed form expressions for the proximal maps of f and g but this may be lost if we go to a conic linear representation of (1.1).

In this way, while the CLP paradigm is powerful and flexible, depending on the situation, it may be desirable to solve a given problem in its "natural" non-reformulated version. This motivates the development of a facial reduction algorithm directly applicable to general convex programs, without any reformulation. Our contributions are as follows.

(a) Given two convex sets C_1, C_2, our facial reduction algorithm (Algorithm 1) either detects that $C_1 \cap C_2 = \varnothing$, or outputs two faces $F^1_{\min} \subseteq C_1$ and $F^2_{\min} \subseteq C_2$ satisfying Slater's condition (i.e., $\operatorname{ri}(F^1_{\min}) \cap \operatorname{ri}(F^2_{\min}) \neq \varnothing$) while keeping the feasible region unchanged, that is,

$$\operatorname{ri}(F^1_{\min}) \cap \operatorname{ri}(F^2_{\min}) \neq \varnothing, \quad F^1_{\min} \cap F^2_{\min} = C_1 \cap C_2,$$

see (3.5). Throughout the paper we discuss the application of this approach to case where $C_1 = \operatorname{dom} f$, $C_2 = \operatorname{dom} g$ with f and g as in (1.1).

(b) Using the infimal convolution we extend the definition of niceness from convex cones to arbitrary convex sets. We also develop an analogous notion for convex functions called *vertical niceness*, which is weaker than requiring the niceness of the whole epigraph. Nevertheless, niceness and vertical niceness are enough for our purposes and we will show in Section 4 that the Fenchel dual of a regularized problem can be further simplified under vertical niceness, see Theorem 4.13. We also develop some basic calculus rules for niceness/vertical niceness and, later in the paper, we show that nice closed convex sets must be facially exposed (Corollary 6.6).

(c) The approach described in item (a) and (b) require that the minimal faces $F^1_{\min}$ and $F^2_{\min}$ be explicitly determined. Inspired by Ramana's dual and later extensions to conic linear programming, we present extended duals for (1.1) that do not require explicit expressions for $F^1_{\min}$ and $F^2_{\min}$, see Section 5 and 6. The most general version of this theory is contained in Section 5 and in Section 6 we present a simplified version that holds under the assumption of niceness and the closedness of the domains of f and g.

This paper is organized as follows. In Section 2, we fix notation used throughout the paper, review definitions, useful preliminary results and literature. Our facial reduction algorithm for convex sets is discussed in Section 3. Then, niceness of convex sets and vertical niceness of proper convex functions are defined and discussed in Section 4. Extended duals without and with niceness are presented in Section 5 and Section 6, respectively. A detailed example illustrating our approach is given in Section 7. We conclude this paper in Section 8.

2 Notation and Preliminaries

In this section, we review some notation, definitions and useful results. All the proofs in this section are deferred to Appendix A. Throughout this paper, we let $\mathcal{E}$ denote a finite dimensional real vector space equipped with inner product $\langle \cdot, \cdot \rangle$.

Given $C \subseteq \mathcal{E}$, we let $\dim(C)$, $\mathrm{ri}\,(C)$, $\mathrm{cl}(C)$, $\mathrm{aff}(C)$, $\mathrm{span}\,(C)$ $\mathrm{cone}(C)$, $\mathrm{rec}(C)$, δ_C and $C^\perp$ denote its *dimension*, *relative interior*, *closure*, *affine hull*, *span*, *conic hull*, *recession cone*, *indicator function* and *orthogonal complement*, respectively. For simplicity, we denote the closure of the conic hull of C by $\overline{\mathrm{cone}}(C)$, i.e., $\overline{\mathrm{cone}}(C) = \mathrm{cl}(\mathrm{cone}(C))$.

The following lemma contains well-known properties of affine hulls for convex sets.

Lemma 2.1 (Affine hull properties of convex sets)**.** *Let $C, C_1, C_2 \subseteq \mathcal{E}$ be nonempty convex sets.*

(i) For every $s \in \mathrm{ri}\,(C)$, we have

$$\mathrm{aff}(C) = \{s + z \mid \exists t \in (0,1), s + tz \in C\}. \tag{2.1}$$

(ii) If $\mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2) \neq \varnothing$, then $\mathrm{aff}(C_1) \cap \mathrm{aff}(C_2) = \mathrm{aff}(C_1 \cap C_2)$.

(iii) If $\mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2) \neq \varnothing$, then $\mathrm{aff}(C_1 - C_2) = \mathrm{aff}(C_1) - \mathrm{aff}(C_2)$.

For a convex set $C \subseteq \mathcal{E}$, its *dual cone* and *polar cone* are denoted by C^* and C°, respectively, defined as

$$C^* \coloneqq \{x \in \mathcal{E} \mid \langle x, y \rangle \geqslant 0, \forall\, y \in C\} \quad \text{and} \quad C^\circ \coloneqq -C^*.$$

A note of caution is that, in the literature (e.g., [35, 10]), C° is typically used to denote the polar set $\{x \in \mathcal{E} \mid \langle x, y \rangle \leqslant 1, \forall\, y \in C\}$. Here, however, C° will always denote the polar cone.

The following lemma establishes an important relationship between the dual of the difference of a closed convex set and a closed convex cone, and the intersection of their respective dual cone and polar cone.

Lemma 2.2 (Dual of difference and intersection of duals)**.** *Let $C_1 \subseteq \mathcal{E}$ be a closed convex cone and $C_2 \subseteq \mathcal{E}$ be a closed convex set such that $C_1 \cap C_2 \neq \varnothing$. Then $(C_2 - C_1)^* = C_2^* \cap C_1^\circ$.*

Faces and tangent cones For a convex set $C \subseteq \mathcal{E}$, a nonempty convex subset F of C is said to be a *face* of C, denoted by $F \trianglelefteq C$, if for $x, y \in C$ and some $\alpha \in (0,1)$, $\alpha x + (1-\alpha)y \in F$ implies $x, y \in F$. Faces satisfy the following important well-known relation:

$$F = C \cap \operatorname{aff}(F), \quad \forall F \trianglelefteq C, \tag{2.2}$$

which follows, for example, from Lemma 2.1(*i*).

A face $F \trianglelefteq C$ is said to be *proper* if $F \neq C$, and we write $F \ntrianglelefteq C$ in this case. A face $F \trianglelefteq C$ is an *exposed face* if $F = C \cap H$ for some supporting hyperplane H of C, in this case we also say that H *exposes* F. If all faces of C are exposed, then C is said to be *facially exposed*. For any $x \in C$, we use $\operatorname{minFace}(x, C)$ to denote the *minimal face* of C that contains x, which always exists and is unique. We have

$$x \in \operatorname{ri}(\operatorname{minFace}(x, C)) \quad \forall\, x \in C, \tag{2.3}$$

e.g., see [27, Proposition 3.2.2].

The following proposition is restated from [22, Proposition 2.2] without assuming the closedness of C_1 and C_2. This proposition can be shown by simply repeating the proof of [22, Proposition 2.2]; indeed, the original statements in [7, Section IV] do not require closedness.

Proposition 2.3 (Faces of intersection)**.** *Let C_1, C_2 be two convex sets such that $C := C_1 \cap C_2$ is nonempty. Let $F \trianglelefteq C$. Then, there are $F_1 \trianglelefteq C_1$, $F_2 \trianglelefteq C_2$ such that $F = F_1 \cap F_2$ and $\operatorname{ri}(F) = \operatorname{ri}(F_1) \cap \operatorname{ri}(F_2)$.*

For a closed convex set $C \subseteq \mathcal{E}$ and $x \in C$, the feasible directions $\operatorname{dir}(x, C)$ and the tangent space $\tan(x, C)$ at x in C are defined respectively as

$$\begin{aligned} \operatorname{dir}(x, C) &:= \{y \mid x + ty \in C \text{ for some } t > 0\}, \\ \operatorname{tanCone}(x, C) &:= \operatorname{cl}(\operatorname{dir}(x, C)), \\ \tan(x, C) &:= \operatorname{tanCone}(x, C) \cap -\operatorname{tanCone}(x, C). \end{aligned} \tag{2.4}$$

Given a face F of a closed convex cone K and an arbitrary $x \in \operatorname{ri}(F)$, the *conjugate face* of F is defined as $F^{\triangle} := K^* \cap F^{\perp} = K^* \cap \{x\}^{\perp}$. We have

$$\begin{aligned} \operatorname{tanCone}(x, K) &= \operatorname{minFace}(x, K)^{\triangle *}, \\ \tan(x, K) &= \operatorname{minFace}(x, K)^{\triangle \perp}, \end{aligned} \tag{2.5}$$

which follows, for example, from [27, Lemma 3.2.1]. Combining (2.3) and (2.5) with the definition of the conjugate face of F, we obtain

$$F = K \cap \{x\}^{\perp} \text{ with } x \in K^* \implies F = \operatorname{minFace}(x, K^*)^{\triangle}. \tag{2.6}$$

Lemma 2.4. *Suppose $C_1, C_2 \subseteq \mathcal{E}$ are nonempty convex sets. Let $F_1 \trianglelefteq C_1$ and $F_2 \trianglelefteq C_2$. Then, $s \in (F_2 - F_1)^*$ if and only if $\sigma_{F_1}(s) + \sigma_{F_2}(-s) \leqslant 0$. If in addition $F_1 \cap F_2 \neq \emptyset$, then $s \in (F_2 - F_1)^*$ if and only if $\sigma_{F_1}(s) + \sigma_{F_2}(-s) = 0$.*

Proof. We have

$$s \in (F_2 - F_1)^* \iff -s \in (F_2 - F_1)^\circ \iff \sigma_{F_2-F_1}(-s) \leqslant 0 \iff \sigma_{F_1}(s) + \sigma_{F_2}(-s) \leqslant 0, \tag{2.7}$$

where the last equivalence comes from the fact that $\sigma_{A+B} = \sigma_A + \sigma_B$ for two convex sets A and B. This proves the first part. Suppose that $F_1 \cap F_2 \neq \varnothing$. For the second part, we see from (2.7) that it suffices to prove that $\sigma_{F_1}(s) + \sigma_{F_2}(-s) \geqslant 0$. Indeed,

$$\sigma_{F_1}(s) = \sup_{x \in F_1} \langle s,\, x\rangle \geqslant \sup_{x \in F_1 \cap F_2} \langle s,\, x\rangle \geqslant \inf_{x \in F_1 \cap F_2} \langle s,\, x\rangle \geqslant \inf_{x \in F_2} \langle s,\, x\rangle = -\sigma_{F_2}(-s),$$

which completes the proof. $\square$

Miscellanea on convex functions For a function $f : \mathcal{E} \to \mathbb{R} \cup \{-\infty, +\infty\}$, we use $\operatorname{dom} f$, $\operatorname{gra} f$ and $\operatorname{epi} f$ to represent the *domain*, *graph* and *epigraph* of f, respectively. That is,

$$\begin{aligned} \operatorname{dom} f &\coloneqq \{x \in \mathcal{E} \mid f(x) < +\infty\}; \\ \operatorname{gra} f &\coloneqq \{(x, \mu) \mid x \in \operatorname{dom} f, f(x) = \mu\}; \\ \operatorname{epi} f &\coloneqq \{(x, \mu) \mid x \in \operatorname{dom} f, f(x) \leqslant \mu\}. \end{aligned}$$

The *convex conjugate function* of f is defined by

$$f^*(x) \coloneqq \sup_{y \in \mathcal{E}} \{\langle x,\, y\rangle - f(y)\}.$$

A function is said to be *closed* if its epigraph is closed. The *closure* of f, denoted by $\operatorname{cl} f$ is the lower semi-continuous hull of f, so that

$$(\operatorname{cl} f)(x) = \liminf_{y \to x} f(y) \quad \text{and} \quad \operatorname{epi} \operatorname{cl} f = \operatorname{cl} \operatorname{epi} f.$$

holds, see [36, Chapter 1.D][1]. Moreover, for any functions $f_1, f_2 : \mathcal{E} \to \mathbb{R} \cup \{-\infty, +\infty\}$, it holds that $\operatorname{cl} f_1 \leqslant f_1$, and $f_1 \leqslant f_2$ implies $\operatorname{cl} f_1 \leqslant \operatorname{cl} f_2$, i.e., the closure operation is order-preserving.

Let $F \subseteq \operatorname{dom} f$. We denote by $f_{|F}$ the restriction of f to F which, by definition, is $f_{|F} \coloneqq f + \delta_F$.

The following lemma concerns the closure and infimum over a family of proper convex functions.

Lemma 2.5 (Interchangeability of closure and infimum operations)**.** *Let $\mathcal{I}$ be an index set. The following items hold.*

(i) Suppose that $f_i : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ is a proper convex function for every $i \in \mathcal{I}$. Then $\operatorname{cl} \inf_{i \in \mathcal{I}} \operatorname{cl} f_i = \operatorname{cl} \inf_{i \in \mathcal{I}} f_i$.

(ii) Let $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function and let $\{C_i \subseteq \mathcal{E}\}_{i \in \mathcal{I}}$ be such that $\bigcup_{i \in \mathcal{I}} C_i = C \subseteq \mathcal{E}$. Then it holds that $\inf_{i \in \mathcal{I}} \inf_{x \in C_i} f(x) = \inf_{x \in C} f(x)$.

[1] In [35, Section 7], the definition of closure of a convex function g is different in that the closure is the lower semicontinuous hull only if g is never $-\infty$. Naturally, for proper convex functions both definitions coincide.

The *support function* of a given set $C \subseteq \mathcal{E}$ is defined as $\sigma_C(x) := \sup_{y\in C}\langle x,\, y\rangle$. The next lemma concerns the support function of the affine hull of a set; throughout this paper, for any set $C \subseteq \mathcal{E}$ we use $\mathcal{L}_C$ to denote a subspace such that $\operatorname{aff}(C) = \mathcal{L}_C + a$ with an arbitrary $a \in \operatorname{aff}(C)$.

Lemma 2.6 (Support function of the affine hull)**.** *Given any set $C \subseteq \mathcal{E}$, write $\operatorname{aff}(C) = \mathcal{L}_C + a$, where $\mathcal{L}_C$ is a subspace and $a \in \operatorname{aff}(C)$ is arbitrary. Then, $\sigma_{\operatorname{aff}(C)} = \langle \cdot,\, a\rangle + \delta_{\mathcal{L}_C^\perp}$. Moreover, $(\operatorname{aff}(C))^\circ = \mathcal{L}_C^\perp \cap \{a\}^\circ$.*

The *infimal convolution* of two functions $f, g : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$(f \,\square\, g)(x) := \inf_{x_1}\{f(x_1) + g(x - x_1)\}.$$

The infimal convolution of f and g is said to be *exact* at x if the infimum is attained at x, i.e., there exists x_1 such that $(f \,\square\, g)(x) = f(x_1) + g(x - x_1)$. We use the notation

$$f \boxdot g$$

to indicate that the infimal convolution of f and g is exact at all point in $\operatorname{dom}(f \,\square\, g)$. By [1, Proposition 12.8], when the infimal convolution is exact, we have $\operatorname{epi}(f \boxdot g) = \operatorname{epi} f + \operatorname{epi} g$.

For two closed convex sets $C_1, C_2 \subseteq \mathcal{E}$, we have from [35, Corollary 16.4.1] that $\sigma_{C_1\cap C_2} = \operatorname{cl}(\sigma_{C_1} \,\square\, \sigma_{C_2})$. Moreover, if $\operatorname{ri}(C_1) \cap \operatorname{ri}(C_2) \neq \varnothing$, then $\sigma_{C_1\cap C_2} = \sigma_{C_1} \boxdot \sigma_{C_2}$. Infimal convolution of closed proper convex functions is commutative and associative, i.e., for closed proper convex functions f, g and h, we have $f \,\square\, g = g \,\square\, f$ and $(f \,\square\, g) \,\square\, h = f \,\square\, (g \,\square\, h)$; see, for example, [35, Page 34] and [10, IV (2.3.3), (2.3.4)]. For multiple convex proper functions $f_1, \ldots, f_m$ we have

$$(f_1 \,\square \cdots \square\, f_m)(x) = \inf\{f_1(x_1) + \cdots f_m(x_m) \mid x_1 + \cdots + x_m = x\}.$$

Separating and supporting hyperplanes A hyperplane H is said to *separate* two nonempty sets $C_1, C_2 \subseteq \mathcal{E}$ if C_1 and C_2 are contained in distinct closed half-spaces associated with H. If C_1 and C_2 are not *both* contained in H itself, then we say H separates C_1 and C_2 *properly*.

For nonempty convex sets $C_1, C_2 \subseteq \mathcal{E}$, by [35, Theorem 11.3], $\operatorname{ri}(C_1) \cap \operatorname{ri}(C_2) = \varnothing$ holds if and only if there exists a hyperplane $H := \{z \mid \langle s,\, z\rangle = \theta\}$ (with $s \neq 0$) separating C_1 and C_2 properly. Overall, by [35, Theorems 11.1 and 11.3] we have

$$\begin{aligned} &\operatorname{ri}(C_1) \cap \operatorname{ri}(C_2) = \varnothing \iff \exists H \text{ separating } C_1, C_2 \text{ properly} \\ \iff &\exists(s,\theta) \text{ s.t. } \inf_{x\in C_1}\langle s,\, x\rangle \geqslant \theta \geqslant \sup_{x\in C_2}\langle s,\, x\rangle \text{ and } \sup_{x\in C_1}\langle s,\, x\rangle > \theta > \inf_{x\in C_2}\langle s,\, x\rangle. \end{aligned} \tag{2.8}$$

The following lemma shows how inequalities related to the proper separation of two nonempty convex sets C_1 and C_2 can be simplified when the sets intersect.

Lemma 2.7 (Separation and intersection)**.** *Given two nonempty and convex sets $C_1, C_2 \subseteq \mathcal{E}$. Let (s,θ) be such that $\inf_{x\in C_2}\langle s,\, x\rangle \geqslant \theta \geqslant \sup_{x\in C_1}\langle s,\, x\rangle$. If $C_1 \cap C_2 \neq \varnothing$, then we have*

$$\theta = \sup_{x\in C_1}\langle s,\, x\rangle = \sup_{x\in C_1\cap C_2}\langle s,\, x\rangle = \inf_{x\in C_1\cap C_2}\langle s,\, x\rangle = \inf_{x\in C_2}\langle s,\, x\rangle.$$

It follows that $s \in \operatorname{dom}\sigma_{C_1}$, $s \in -\operatorname{dom}\sigma_{C_2}$, $s \in \operatorname{dom}\sigma_{C_1\cap C_2}$ and $C_1 \cap C_2 \subseteq H := \{z \mid \langle s,\, z\rangle = \theta\}$.

For supporting hyperplanes, we further have the following property.

Lemma 2.8 (Intersection of supporting hyperplanes)**.** *Let C be a convex set and $s_1, s_2 \in \operatorname{dom} \sigma_C$. Let $H_1 := \{z \mid \langle s_1, z\rangle = \sigma_C(s_1)\}$, $H_2 := \{z \mid \langle s_2, z\rangle = \sigma_C(s_2)\}$, $H := \{z \mid \langle s_1 + s_2, z\rangle = \sigma_C(s_1) + \sigma_C(s_2)\}$ and $\widetilde{H} := \{z \mid \langle s_1 + s_2, z\rangle = \sigma_C(s_1 + s_2)\}$. Then we have $C \cap H_1 \cap H_2 = C \cap H$. If further $C \cap H_1 \cap H_2 \neq \varnothing$, we have $C \cap H_1 \cap H_2 = C \cap \widetilde{H}$.*

Strong duality We define *strong duality* between a pair of primal and dual problems as follows.

Definition 2.9 (Strong duality)**.** *Let (P) and (D) be pairs of primal and dual problems (e.g., (P) being (1.1) and (D) being (1.2)), where (P) is feasible. We say that* (P) *and* (D) *satisfy* strong duality *if their optimal values coincide and (D) is attained.*

Some remarks are in order. First, we do not define precisely what it means for an optimization problem to be the "dual" of some other problem. That said, the only cases where we will use Definition 2.9 is when (D) is the Fenchel dual of either (P) or some reformulation of (P). Also, when we say that "(D) is attained" we mean that there exists an optimal solution for the problem (D) that achieves the common optimal value. This includes the case where the common optimal value is $-\infty$, as in this case (D) is "attained" at any point that is not in the feasible region of (D). This usage is consistent with the conventions in [35, Section 31].

Our definition of strong duality is not symmetric with respect to (P) and (D). Furthermore, it may happen that (P) is not attained even when there is strong duality between (P) and (D). While it could make sense to define separately *primal strong duality* and *dual strong duality* (e.g., Definition 2.9 but with the roles of (P) and (D) exchanged), as we will only be concerned with strong duality in the sense of Definition 2.9, we will simply use the term *strong duality* throughout.

Slater's condition We say that (1.1) satisfies *Slater's condition* if $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(\operatorname{dom} g) \neq \varnothing$ holds. Under Slater's condition the problem (1.1) and its Fenchel dual (1.2) satisfy strong duality in the sense of Definition 2.9 by [35, Theorem 31.1].

3 Facial Reduction for Convex Sets

In this section, we develop a new facial reduction algorithm for convex sets to regularize (1.1). More generally, we consider the following feasibility problem

$$\text{find } x \in C_1 \cap C_2, \tag{3.1}$$

where $C_1, C_2 \subseteq \mathcal{E}$ are nonempty convex sets that need not to be closed.

The goal of our facial reduction algorithm (Algorithm 1) is to either detect infeasibility of (3.1) (i.e., $C_1 \cap C_2 = \varnothing$), or find two unique faces $F^1_{\min} \trianglelefteq C_1$ and $F^2_{\min} \trianglelefteq C_2$ satisfying

$$F^1_{\min} \trianglelefteq C_1, \quad F^2_{\min} \trianglelefteq C_2, \quad F^1_{\min} \cap F^2_{\min} = C_1 \cap C_2, \quad \operatorname{ri}(F^1_{\min}) \cap \operatorname{ri}(F^2_{\min}) \neq \varnothing. \tag{3.2}$$

These two unique faces always exist if (3.1) is feasible by applying Proposition 2.3 to C_1, C_2.

This can be applied to (1.1) as follows. If $C_1 = \operatorname{dom} f$ and $C_2 = \operatorname{dom} g$ hold[2], then (3.2) implies that (1.1) is equivalent to

$$\min_x f_{|F^1_{\min}}(x) + g_{|F^1_{\min}}(x). \tag{3.3}$$

However, since $F^1_{\min}$ and $F^2_{\min}$ are faces, we have $F^1_{\min} = C_1 \cap \operatorname{aff}(F^1_{\min})$ and $F^2_{\min} = C_2 \cap \operatorname{aff}(F^2_{\min})$. That is, $f_{|F^1_{\min}} = f_{|\operatorname{aff}(F^1_{\min})}$ and $g_{|F^2_{\min}} = f_{|\operatorname{aff}(F^2_{\min})}$, which leads to

$$(f_{|F^1_{\min}})^* = \operatorname{cl}(f^* \,\square\, \sigma_{\operatorname{aff}(F^1_{\min})}), \qquad (g_{|F^2_{\min}})^* = \operatorname{cl}(g^* \,\square\, \sigma_{\operatorname{aff}(F^2_{\min})}), \tag{3.4}$$

by [35, Theorem 16.4]. Therefore, we also have that (1.1) is equivalent to

$$\inf_x f_{|\operatorname{aff}(F^1_{\min})}(x) + g_{|\operatorname{aff}(F^2_{\min})}(x). \tag{3.5}$$

More precisely, the equivalence between these three problems is as follows. The problems (1.1), (3.3) and (3.5) share the same feasible set, optimal solution set and optimal value. However a distinct advantage of (3.3) and (3.5) over (1.1) is that both of them satisfy Slater's condition (see (3.2)) and therefore their Fenchel duals

$$\sup_\lambda -(f_{|F^1_{\min}})^*(-\lambda) - (g_{|F^2_{\min}})^*(\lambda), \tag{3.6}$$

and

$$\sup_\lambda - \operatorname{cl}(f^* \,\square\, \sigma_{\operatorname{aff}(F^1_{\min})})(-\lambda) - \operatorname{cl}(g^* \,\square\, \sigma_{\operatorname{aff}(F^2_{\min})})(\lambda), \tag{3.7}$$

satisfy strong duality in the sense that (3.5) and (3.7) share the same optimal value and the latter is attained. The same is true for (3.3) and (3.6). We summarize the discussion so far as follows.

Theorem 3.1. *Suppose that (1.1) is feasible (i.e., $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$). Then, the primal dual pairs formed by (1.1), (3.6) and (1.1), (3.7) satisfy strong duality in the sense of Definition 2.9.*

In what follows we say that (3.3) and (3.5) are *regularized primal* problems of (1.1). Similarly, we say that (3.6) and (3.7) are *regularized Fenchel dual* problems of (1.1).

3.1 Characterizations of proper separation

The facial reduction algorithm we will introduce shortly hinges on finding proper separating hyperplanes for two convex sets S_1 and S_2 when their relative interiors do not intersect. Before we proceed, we take a brief look at this. We start with the following lemma.

Lemma 3.2. *Let $S_1, S_2 \subseteq \mathcal{E}$ be nonempty convex sets. Suppose that the hyperplane $H \coloneqq \{z \mid \langle s,\, z\rangle = \theta\}$ separates S_1 and S_2. Then for $i \in \{1, 2\}$,*

$$S_i \subseteq H \Leftrightarrow \langle s,\, e\rangle = \theta, \forall s \in \operatorname{ri} S_i \Leftrightarrow \exists\, e \in \operatorname{ri}(S_i) \text{ such that } \langle s,\, e\rangle = \theta$$

[2]Here we note that even if f and g are assumed to be closed, their domains may not be closed.

Proof. All the "$\Rightarrow$" implications are straightforward, so we only show the "$\Leftarrow$" implications.

Suppose that $i \in \{1, 2\}$, and $e \in \mathrm{ri}\,(S_i)$ is such that

$$\langle s,\, e\rangle = \theta \tag{3.8}$$

and let $x \in S_i$ be arbitrary. Then, there exists some $\alpha > 1$ such that $\alpha e + (1-\alpha)x \in S_i$, e.g., see [35, Theorem 6.4]. Since H separates S_1 and S_2, without loss of generality, we assume that

$$\langle s,\, x\rangle \leqslant \theta, \quad \forall\, x \in S_i. \tag{3.9}$$

Now, using (3.8), (3.9) and $\alpha e + (1-\alpha)x \in S_i$ from the convexity of S_i, we have

$$\theta \geqslant \langle s,\, \alpha e + (1-\alpha)x\rangle = \alpha\langle s,\, e\rangle + (1-\alpha)\langle s,\, x\rangle = \alpha\theta + (1-\alpha)\langle s,\, x\rangle,$$

which together with $\alpha > 1$ implies that $\langle s,\, x\rangle \geqslant \theta$. This further together with (3.9) implies that $x \in H$ and completes the proof. $\square$

Using this lemma, we characterize proper separation as follows.

Theorem 3.3 (Proper separation equivalence)**.** *Let $S_1, S_2 \subseteq \mathcal{E}$ be nonempty convex sets and let $e_1 \in \mathrm{ri}\,(S_1)$ and $e_2 \in \mathrm{ri}\,(S_2)$.*

(i) S_1 and S_2 can be properly separated if and only if there exists $s \in (S_2 - S_1)^*$ with $\langle s,\, e_2 - e_1\rangle > 0$.

(ii) Let $s \in (S_2 - S_1)^*$ and $\theta \in \mathbb{R}$. Then, $H \coloneqq \{z \mid \langle s,\, z\rangle = \theta\}$ separates S_1 and S_2 if and only if θ satisfies $\inf\limits_{y \in S_2}\langle s,\, y\rangle \geqslant \theta \geqslant \sup\limits_{y \in S_1}\langle s,\, y\rangle$.

Proof. (*i*) First, we prove "if" part. Suppose that there exists $s \in (S_2 - S_1)^*$ with $\langle s,\, e_2 - e_1\rangle > 0$. We then have

$$\langle s,\, y - x\rangle \geqslant 0, \quad \forall\, x \in S_1,\, y \in S_2,$$

which further implies that

$$\inf_{y \in S_2}\langle s,\, y\rangle \geqslant \sup_{x \in S_1}\langle s,\, x\rangle.$$

This implies that there exists a hyperplane, say, $H \coloneqq \{z \mid \langle s,\, z\rangle = \theta\}$, which separates S_1 and S_2. If $S_1 \subseteq H$, we then see from $e_1 \in \mathrm{ri}\,(S_1)$ that $\langle s,\, e_1\rangle = \theta$. This together with $\langle s,\, e_2 - e_1\rangle > 0$ implies that $\langle s,\, e_2\rangle > \theta$. Using Lemma 3.2, we have $S_2 \nsubseteq H$. Consequently, S_1 and S_2 are not both contained in H, which proves that H separates S_1 and S_2 properly. This completes the proof of "if" part.

Now we prove "only if" part. Suppose that there exists a hyperplane $H \coloneqq \{z \mid \langle s,\, z\rangle = \theta\}$ which separates S_1 and S_2 properly. We then have that at least one of S_1 and S_2 is not contained in H, and without loss of generality (otherwise, we replace s and θ by $-s$ and $-\theta$, respectively),

$$\langle s,\, x\rangle \leqslant \theta \leqslant \langle s,\, y\rangle, \quad \forall\, x \in S_1,\, y \in S_2. \tag{3.10}$$

We see from (3.10) that

$$\langle s,\, y - x\rangle \geqslant 0, \quad \forall\, x \in S_1,\, y \in S_2,$$

which implies $s \in (S_2 - S_1)^*$. Specifically, it holds that $\langle s, e_2 - e_1 \rangle \geqslant 0$. Now it remains to show that $\langle s, e_2 - e_1 \rangle > 0$. Suppose that $\langle s, e_2 - e_1 \rangle = 0$. This together with (3.10) further gives

$$\langle s, e_2 \rangle = \langle s, e_1 \rangle \leqslant \theta \leqslant \langle s, e_2 \rangle,$$

which implies that $\langle s, e_1 \rangle = \langle s, e_2 \rangle = \theta$. This together with Lemma 3.2 implies that $S_1 \subseteq H$ and $S_2 \subseteq H$, which causes a contradiction. Consequently, we have $\langle s, e_2 - e_1 \rangle > 0$.

(ii) We first show "if" part. Suppose that H separates S_1 and S_2, then by definition we know that S_1 and S_2 are contained in distinct closed half-spaces associated with H. Since $s \in (S_2 - S_1)^*$, for any $y_1 \in S_1$ and $y_2 \in S_2$, we have $\langle s, y_2 - y_1 \rangle \geqslant 0$, which implies $\langle s, y_2 \rangle \geqslant \langle s, y_2 \rangle$. Then, we know from the separation that S_1 and S_2 are contained in the half-spaces $H_- := \{z \mid \langle s, z \rangle \leqslant \theta\}$ and $H_+ := \{z \mid \langle s, z \rangle \geqslant \theta\}$, respectively. This observation immediately leads to the desired result.

We then prove "only if" part. Suppose that $\inf_{y \in S_2} \langle s, y \rangle \geqslant \theta \geqslant \sup_{y \in S_1} \langle s, y \rangle$. Then we know S_1 and S_2 are contained in the half-spaces $H_- := \{z \mid \langle s, z \rangle \leqslant \theta\}$ and $H_+ := \{z \mid \langle s, z \rangle \geqslant \theta\}$, respectively, which by definition implies that S_1 and S_1 are separated by H. □

3.2 A facial reduction algorithm

We now move on to the development of our facial reduction algorithm for convex sets. Our algorithm is summarized as follows, where we regard as *one facial reduction step* the computation of H_k as in **Step 1**.

Algorithm 1 Facial reduction algorithm

Step 0. Set $k \leftarrow 0$, $F_k^1 \leftarrow C_1$ and $F_k^2 \leftarrow C_2$.

Step 1. If $\mathrm{ri}\,(F_k^1) \cap \mathrm{ri}\,(F_k^2) = \varnothing$ holds, choose $s_k \in (F_k^2 - F_k^1)^*$, $\theta_k \in \mathbb{R}$ such that the hyperplane $H_k := \{z \mid \langle s_k, z \rangle = \theta_k\}$ separates F_k^1 and F_k^2 properly (Theorem 3.3); otherwise, **stop** and return F_k^1 and F_k^2.

Step 2. If $\inf_{y \in F_k^2} \langle s_k, y \rangle > \sup_{y \in F_k^1} \langle s_k, y \rangle$ holds, **stop** and declare that (3.1) is infeasible.

Step 3. Set $F_{k+1}^1 \leftarrow F_k^1 \cap H_k$, $F_{k+1}^2 \leftarrow F_k^2 \cap H_k$.

Step 4. If $F_{k+1}^1 = \varnothing$ or $F_{k+1}^2 = \varnothing$ holds, **stop** and declare that (3.1) is infeasible. Otherwise, set $k \leftarrow k + 1$ and go to **Step 1**.

Lemma 3.4 (Properties of Algorithm 1)**.** *The following items hold.*

(i) If k is such that $F_k^1 \cap F_k^2 \neq \varnothing$, then we have

$$\theta_k = \sup_{y \in F_k^1} \langle s_k, y \rangle = \sup_{y \in F_k^1 \cap F_k^2} \langle s_k, y \rangle = \inf_{y \in F_k^1 \cap F_k^2} \langle s_k, y \rangle = \inf_{y \in F_k^2} \langle s_k, y \rangle.$$

It follows that $s_k \in \mathrm{dom}\, \sigma_{F_k^1}$, $s_k \in -\mathrm{dom}\, \sigma_{F_k^2}$, $s_k \in \mathrm{dom}\, \sigma_{F_k^1 \cap F_k^2}$ and $F_k^1 \cap F_k^2 \subseteq H_k$.

(ii) For all k, $(s_i, \theta_i)_{i=0}^k$ are linearly independent.

Proof. Items (*i*) follows from Lemma 2.7 and $F_k^1 \cap F_k^2 \neq \emptyset$.

(*ii*) We show the linear independence of $(s_i, \theta_i)_{i=0}^k$ by induction on k. Assume that $(s_i, \theta_i)_{i=0}^{k-1}$ are linearly independent, we shall show that $(s_i, \theta_i)_{i=0}^k$ are also linearly independent. Suppose that $(s_i, \theta_i)_{i=0}^k$ are linearly dependent, then there exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}$, not all zero, such that

$$s_k = \sum_{i=0}^{k-1} \alpha_i s_i, \qquad \theta_k = \sum_{i=0}^{k-1} \alpha_i \theta_i.$$

Therefore, we have $H_k = \{z \mid \langle \sum_{i=0}^{k-1} \alpha_i s_i,\, z \rangle = \sum_{i=0}^{k-1} \alpha_i \theta_i\}$. We conclude that

$$F_k^1 = C_1 \cap H_0 \cap H_1 \cap \ldots \cap H_{k-1} = F_k^1 \cap H_k, \quad F_k^2 = C_2 \cap H_0 \cap H_1 \cap \ldots \cap H_{k-1} = F_k^2 \cap H_k,$$

which follows that both F_k^1 and F_k^2 are in H_k, violating the assumption that H_k separates F_k^1 and F_k^2 *properly*. Therefore, $(s_i, \theta_i)_{i=0}^k$ must be linearly independent. □

Given a convex set C, let ℓ_C be the length of the *longest* strictly ascending chain of nonempty faces in C. Denote

$$\ell := \min\{\ell_{C_1} + \ell_{C_2} - 2, \dim(\mathcal{E}) + 1\}\}. \tag{3.11}$$

Then, Algorithm 1 terminates after at most ℓ facial reduction steps, as shown in the next theorem.

Theorem 3.5 (Finite termination)**.** Algorithm 1 *terminates after at most ℓ facial reduction steps. Furthermore, the following holds.*

*(i) $C_1 \cap C_2 = \emptyset$ holds if and only if Algorithm 1 stops at either **Step 2** or **Step 4***

(ii) Let $\widehat{\ell}$ be the number of facial reduction steps. If $C_1 \cap C_2 \neq \emptyset$, then the faces $F_{\widehat{\ell}}^1$ and $F_{\widehat{\ell}}^2$ satisfy (3.2).

Proof. The basic idea for checking the finite termination of Algorithm 1 is that whenever **Step 4** is reached and none of the stopping criteria are satisfied, then either F_{k+1}^1 is a proper face of F_k^1 or F_{k+1}^2 is a proper faces of F_k^2 or both.

In more details, let us examine one facial reduction step of Algorithm 1. At **Step 1**, if $\mathrm{ri}\,(F_k^1) \cap \mathrm{ri}\,(F_k^2) = \emptyset$, then by (2.8) there exists a hyperplane H_k that separates F_k^1 and F_k^2 properly. Moreover, if $F_k^1 \cap F_k^2 \neq \emptyset$, we have

$$\inf_{y \in F_k^2} \langle s_k,\, y \rangle = \sup_{y \in F_k^1} \langle s_k,\, y \rangle,$$

by Lemma 3.4(*i*)

Therefore, if $\inf_{y \in F_k^2} \langle s_k,\, y \rangle > \sup_{x \in F_k^1} \langle s_k,\, x \rangle$ holds in **Step 2**, then we have $F_k^1 \cap F_k^2 = \emptyset$.

If we proceed to **Step 3** without stopping, we have $\inf_{y \in F_k^2} \langle s_k,\, y \rangle = \sup_{x \in F_k^1} \langle s_k,\, x \rangle$.

Next, we observe that $F_k^1 \cap F_k^2 \subseteq H_k$ by looking at two cases. If $F_k^1 \cap F_k^2 = \emptyset$, then we naturally have $F_k^1 \cap F_k^2 \subseteq H_k$. If $F_k^1 \cap F_k^2 \neq \emptyset$, then by Lemma 3.4, we have $F_k^1 \cap F_k^2 \subseteq H_k$. If either $F_{k+1}^1 = \emptyset$ or $F_{k+1}^2 = \emptyset$ holds at **Step 4**, it is because $F_k^1 \cap F_k^2 = \emptyset$. Otherwise, H_k

defines a supporting hyperplane to both F_k^1 and F_k^2 that exposes the nonempty faces F_{k+1}^1 and F_{k+1}^2, respectively. By induction, we have

$$F_{k+1}^1 = F_k^1 \cap H_k \trianglelefteq F_k^1 \trianglelefteq \ldots \trianglelefteq C_1, \quad F_{k+1}^2 = F_k^2 \cap H_k \trianglelefteq F_k^2 \trianglelefteq \ldots \trianglelefteq C_2.$$

and

$$F_{k+1}^1 \cap F_{k+1}^2 = F_k^1 \cap F_k^2 \cap H_k = F_k^1 \cap F_k^2 = \cdots = C_1 \cap C_2. \tag{3.12}$$

Since F_k^1 and F_k^2 are not simultaneously contained in H_k, we have

$$\dim(F_{k+1}^1) = \dim(F_k^1 \cap H_k) < \dim(F_k^1) \quad \text{or} \quad \dim(F_{k+1}^2) = \dim(F_k^2 \cap H_k) < \dim(F_k^2), \tag{3.13}$$

which corresponds to

$$F_{k+1}^1 \trianglelefteq\!\!\!\!\!\neq F_k^1 \quad \text{or} \quad F_{k+1}^2 \trianglelefteq\!\!\!\!\!\neq F_k^2. \tag{3.14}$$

Given that C_1 and C_2 are contained in the finite dimensional space $\mathcal{E}$, we conclude from (3.13), (3.14) and **Step 1** that Algorithm 1 will terminate after finitely many facial reduction steps. Denote by $\widehat{\ell}$ the number of facial reduction steps.

By Lemma 3.4(*ii*), $(s_i, \theta_i)_{i=0}^{\widehat{\ell}-1}$ are linearly independent in $\mathcal{E} \times \mathbb{R}$, so $\widehat{\ell} \leqslant \dim(E) + 1$. Meanwhile, (3.14) and the properness of the separation **Step 1**, have $\widehat{\ell} \leqslant \ell_{C_1} + \ell_{C_2} - 2$. Then, Algorithm 1 will terminate after at most ℓ (cf. (3.11)) facial reduction steps.

To summarize, if we stop at **Step 1** with $k = \widehat{\ell}$, then F_k^1, F_k^2 satisfy (3.2) by (3.12); if we stop at either **Step 2** or **Step 4** it is because (3.1) is infeasible.

Conversely, we have

- If (3.1) is infeasible, then $\operatorname{ri}(F_k^1) \cap \operatorname{ri}(F_k^2) = \varnothing$ will always hold at **Step 1**. By finite termination, we will eventually stop at either **Step 2** or **Step 4**.
- If (3.1) is feasible, (3.12) tells us that whenever smaller faces are identified, we have $F_{k+1}^1 \cap F_{k+1}^2 = C_1 \cap C_2 \neq \varnothing$, so the algorithm never stops at **Step 2** or **Step 4**. Again, by finite termination, eventually we will stop at **Step 1** and F_k^1, F_k^2 satisfy (3.2) by (3.12).

□

In Algorithm 1, at each k, there may be several choices of H_k that properly separate F_k^1 and F_k^2. This motivates the following definition.

Definition 3.6. *Let $C_1, C_2 \subseteq \mathcal{E}$ be convex sets. We call the* minimal *number of facial reduction steps for Algorithm 1 to stop the* joint singularity degree of C_1 and C_2, *denoted by $\mathcal{SD}(C_1, C_2)$.*

When the context is clear we will omit C_1 and C_2 and simply write $\mathcal{SD}$. Then, Lemma 5.1(*i*) tells us that $\mathcal{SD} \leqslant \ell$ holds. When $C_1 \cap C_2 \neq \varnothing$, since Algorithm 1 must stop at **Step 1**, the joint singularity degree C_1 and C_2 corresponds to the minimal number of reducing steps to reach the minimal faces as in (3.2).

4 Nice convex sets and vertical niceness

In this section, we generalize the notion of *niceness* from closed convex cones to general convex sets and proper convex functions. The motivation for doing so comes from the Fenchel dual (3.7) of the regularized problem (3.5) and (3.4): we wish to simplify the expressions for the conjugate functions $(f_{|F^1_{\min}})^*$ and $(g_{|F^2_{\min}})^*$.

Before we proceed, we need a discussion on lifting a closed convex set to a convex cone.

4.1 Interlude on conic liftings of convex sets

In this subsection, we discuss some properties of faces and the *conic lifting* $\overline{\mathrm{cone}}(C \times \{-1\})$ of a convex set $C \subseteq \mathcal{E}$. We note that all results in this subsection are still valid if $\overline{\mathrm{cone}}(C \times \{-1\})$ is replaced with $\overline{\mathrm{cone}}(C \times 1)$, $\overline{\mathrm{cone}}(-1 \times C)$ or $\overline{\mathrm{cone}}(1 \times C)$, with the caveat that some formulae may need to be adjusted slightly.

When C is compact the situation is relatively straightforward as $\overline{\mathrm{cone}}(C \times \{-1\}) = \mathrm{cone}(C \times \{-1\})$ and it is well-known that there exists a bijective correspondence between the faces of C and the faces of the cone $\mathrm{cone}(C \times \{-1\})$, e.g., see [37, Section 3]. Without compacts there could be certain subtleties, which we examine in this interlude.

The following lemma shows that the recession cone of a face of a convex set C is itself a face of the recession cone of C, and describes the relationship between their corresponding exposing vectors, when such vectors exist.

Lemma 4.1 (Recession cones of faces)**.** *Let $C \subseteq \mathcal{E}$ be a closed convex set and $F \trianglelefteq C$. Then, $\mathrm{rec}(F) \trianglelefteq \mathrm{rec}(C)$ holds. Furthermore, if F is exposed by $H := \{z \mid \langle z,\, s\rangle = \theta\}$ with $s \in \mathrm{dom}\,\sigma_C$, then $\mathrm{rec}(F)$ is exposed by $H_0 := \{z \mid \langle z,\, s\rangle = 0\}$.*

Proof. By [35, Theorem 8.3], we have $\mathrm{rec}(F) \subseteq \mathrm{rec}(C)$. For arbitrary $d_1, d_2 \in \mathrm{rec}(C)$, we have, for any $x \in C$ and $t \geqslant 0$, that $x + td_1 \in C$ and $x + td_2 \in C$. Suppose that for some $\lambda \in (0, 1)$, it holds that $\lambda d_1 + (1-\lambda)d_2 \in \mathrm{rec}(F)$. Then, by definition, for any $x \in F$ and $t \geqslant 0$, we have

$$x + t(\lambda d_1 + (1-\lambda)d_2) = x + \lambda t d_1 + (1-\lambda)td_2 \in F.$$

That is,

$$\lambda(x + td_1) + (1-\lambda)(x + td_2) \in F.$$

Since $F \trianglelefteq C$, and given that $x+td_1 \in C$ and $x+td_2 \in C$, we conclude that $x+td_1 \in F$ and $x+td_2 \in F$ for any $t \geqslant 0$. This demonstrates that $d_1, d_2 \in \mathrm{rec}(F)$, thus proving that $\mathrm{rec}(F) \trianglelefteq \mathrm{rec}(C)$.

If F is exposed by $H := \{z \mid \langle z,\, s\rangle = \theta\}$ with $s \in \mathrm{dom}\,\sigma_C$, we have

$$\langle s,\, x\rangle < \theta \quad \forall\, x \in C \backslash F; \qquad \langle s,\, x\rangle = \theta \quad \forall\, x \in F. \tag{4.1}$$

For any $d \in \mathrm{rec}(F)$, we know from the definition of recession cone that $x + td \in F$ for any $x \in F$ and $t > 0$. Then, by (4.1), we have $\langle s,\, x + td\rangle = \theta$, which implies $\langle s,\, d\rangle = 0$ since $\langle s,\, x\rangle = \theta$. Similar, for any $d \in \mathrm{rec}(C)\backslash\mathrm{rec}(F)$, we have $x + td \in C\backslash F$ for some $x \in F$ and some $t > 0$. Then $\langle s,\, x + td\rangle < \theta$, implying $\langle s,\, d\rangle < 0$. Summarizing, we have

$$\langle s,\, d\rangle < 0 \quad \forall\, d \in \mathrm{rec}(C)\backslash\,\mathrm{rec}(F); \qquad \langle s,\, d\rangle = 0 \quad \forall\, x \in \mathrm{rec}(F).$$

This shows that $\mathrm{rec}(F)$ is exposed by $H_0 := \{z \mid \langle z,\, s\rangle = 0\}$. $\square$

Next we prove some properties of conic liftings of closed convex sets, which we will repeatedly use in the sequel.

Proposition 4.2 (Properties of conic liftings)**.** *Let $C \subseteq \mathcal{E}$ be a nonempty convex set. It holds that*

$$\overline{\mathrm{cone}}(C \times \{-1\})^\circ = \mathrm{epi}\,\sigma_C, \tag{4.2}$$
$$\overline{\mathrm{cone}}(C \times \{-1\})^\perp = \mathrm{gra}\,\sigma_{\mathrm{aff}(C)}. \tag{4.3}$$

Furthermore, if C is closed, we have

- *(i) A face $F \trianglelefteq C$ is exposed by $H := \{x \mid \langle x,\, s\rangle = \theta\}$ with $s \in \mathrm{dom}\,\sigma_C$ and $\theta = \sigma_C(s)$ if and only if $\overline{\mathrm{cone}}(F \times \{-1\}) \trianglelefteq \overline{\mathrm{cone}}(C \times \{-1\})$ is exposed by $\{(s, \theta)\}^\perp$ with $-(s, \theta) \in \overline{\mathrm{cone}}(C \times \{-1\})^*$.*
- *(ii) $F \trianglelefteq C$ if and only if $\overline{\mathrm{cone}}(F \times \{-1\}) \trianglelefteq \overline{\mathrm{cone}}(C \times \{-1\})$.*

Proof. We first show (4.2). Let $(z, \lambda) \in \mathrm{cone}(C \times \{-1\})^\circ$. For any $x \in C$ and $\mu > 0$, we have $(\mu x,\, -\mu) \in \mathrm{cone}(C \times \{-1\})$ and therefore

$$0 \geqslant \langle (z,\, \lambda),\, (\mu x,\, -\mu)\rangle = \mu\langle x,\, z\rangle - \lambda\mu.$$

This together with $\mu > 0$ implies that $\langle x,\, z\rangle \leqslant \lambda$ for any $x \in C$. Then $\sigma_C(z) \leqslant \lambda$ and hence $(z,\, \lambda) \in \mathrm{epi}\,\sigma_C$. Since $(z,\, \lambda)$ is arbitrary, this shows that $\overline{\mathrm{cone}}(C \times \{-1\})^\circ = \mathrm{cone}(C \times \{-1\})^\circ \subseteq \mathrm{epi}\,\sigma_C$.

Conversely, let $(z,\, \lambda) \in \mathrm{epi}\,\sigma_C$. Then, it holds that $\langle x,\, z\rangle \leqslant \lambda$ for any $x \in C$. For any $(y,\, -\mu) \in \mathrm{cone}(C \times \{-1\})$ with $\mu > 0$, there exists $x \in C$ such that $(y, -\mu) = \mu(x, -1)$. Then,

$$\langle (z,\, \lambda),\, (y,\, -\mu)\rangle = \langle z,\, y\rangle - \lambda\mu = \mu\langle z,\, x\rangle - \lambda\mu \leqslant \lambda\mu - \lambda\mu = 0.$$

That is, $(z,\, \lambda) \in \mathrm{cone}(C \times \{-1\})^\circ$. Since $(z,\, \lambda)$ is arbitrary, we conclude that $\mathrm{epi}\,\sigma_C \subseteq \mathrm{cone}(C \times \{-1\})^\circ = \overline{\mathrm{cone}}(C \times \{-1\})^\circ$. Therefore, we obtain (4.2).

We next show (4.3). Let $(z,\, \lambda) \in \mathrm{cone}(C \times \{-1\})^\perp$. For any $x \in C$ and $\mu > 0$, we have $(\mu x,\, -\mu) \in \mathrm{cone}(C \times \{-1\})$. Then,

$$0 = \langle (z,\, \lambda),\, (\mu x,\, -\mu)\rangle = \mu\langle x,\, z\rangle - \lambda\mu.$$

This together with $\mu > 0$ implies that $\langle x,\, z\rangle = \lambda$ for any $x \in C$. Now, for any $a \in \mathrm{aff}(C)$, there exist $k \in \mathbb{N}$, $\{\alpha_i\}_{i=1}^k$ and $\{x_i\}_{i=1}^k \subseteq C$ such that $\sum_{i=1}^k \alpha_i = 1$ and $a = \sum_{i=1}^k \alpha_i x_i$. Thus,

$$\langle a,\, z\rangle = \sum_{i=1}^k \langle \alpha_i x_i,\, z\rangle = \sum_{i=1}^k \alpha_i \lambda = \lambda.$$

This together with the arbitrariness of a implies that $(z,\, \lambda) \in \mathrm{gra}\,\sigma_{\mathrm{aff}(C)}$ and $\overline{\mathrm{cone}}(C \times \{-1\})^\perp = \mathrm{cone}(C \times \{-1\})^\perp \subseteq \mathrm{gra}\,\sigma_{\mathrm{aff}(C)}$.

Conversely, let $(z, \lambda) \in \mathrm{gra}\,\sigma_{\mathrm{aff}(C)}$, so that $\sigma_{\mathrm{aff}(C)}(z) = \lambda$ holds. Let $x \in C$. Since $x \in \mathrm{aff}(C)$, Lemma 2.6 implies that

$$\sigma_{\mathrm{aff}(C)}(z) = \langle x,\, z\rangle = \lambda.$$

Therefore, for any $\mu > 0$, we have

$$0 = \mu\langle x,\, z\rangle - \lambda\mu = \langle (z, \lambda),\, (\mu x, -\mu)\rangle.$$

Since $(\mu x, -\mu) \in \operatorname{cone}(C \times \{-1\})$ and x and μ are arbitrary, we obtain $(z, \lambda) \in \operatorname{cone}(C \times \{-1\})^{\perp}$. Hence, $\operatorname{gra}\sigma_{\operatorname{aff}(C)} \subseteq \operatorname{cone}(C \times \{-1\})^{\perp} = \overline{\operatorname{cone}}(C \times \{-1\})^{\perp}$.

Next we assume that C is closed and show the remaining statements.

(i) This proof is inspired by [37, Proposition 3.2].

($\Longrightarrow$) We first note that

$$\overline{\operatorname{cone}}(C \times \{-1\}) = \operatorname{cone}(C \times \{-1\}) \cup (\operatorname{rec}(C) \times \{0\})$$

holds by [35, Theorem 8.2]. Then, the basic idea is to show that $\operatorname{cone}(F \times \{-1\})$ is a face of $\operatorname{cone}(C \times \{-1\})$ exposed by $\{(s, \theta)\}^{\perp}$, and $(\operatorname{rec}(F) \times \{0\})$ is a face of $\operatorname{rec}(C) \times \{0\}$ also exposed by $\{(s, \theta)\}^{\perp}$, then

$$\begin{aligned}
&\overline{\operatorname{cone}}(F \times \{-1\}) = \operatorname{cone}(F \times \{-1\}) \cup (\operatorname{rec}(F) \times \{0\}) \\
&= \left(\operatorname{cone}(C \times \{-1\}) \cap \{(s, \theta)\}^{\perp}\right) \cup \left((\operatorname{rec}(C) \times \{0\}) \cap \{(s, \theta)\}^{\perp}\right) \\
&= (\operatorname{cone}(C \times \{-1\}) \cup (\operatorname{rec}(C) \times \{0\})) \cap \{(s, \theta)\}^{\perp} = \overline{\operatorname{cone}}(C \times \{-1\}) \cap \{(s, \theta)\}^{\perp},
\end{aligned}$$

which shows that $\overline{\operatorname{cone}}(F \times \{-1\})$ is exposed by $\{(s, \theta)\}^{\perp}$. From Lemma 4.1 we know $\operatorname{rec}(F)$ is exposed by $\{s\}^{\perp}$, which implies that $\operatorname{rec}(F) \times \{0\}$ is exposed by $\{(s, \theta)\}^{\perp}$.

Suppose that $F \trianglelefteq C$ is exposed by $H \coloneqq \{x \mid \langle x,\, s\rangle = \theta\}$ with $s \in \operatorname{dom}\sigma_C$ and $\theta = \sigma_C(s)$. Then,

$$\langle s,\, x\rangle < \theta \quad \forall\, x \in C \backslash F; \qquad \langle s,\, x\rangle = \theta \quad \forall\, x \in F.$$

For any $(\alpha x, -\alpha) \in \operatorname{cone}(F \times \{-1\})$ with $x \in F$ and $\alpha > 0$, we have

$$\langle \alpha x,\, s\rangle - \alpha\theta = \alpha(\langle x,\, s\rangle - \theta) = 0.$$

Similarly, for $(\alpha x, -\alpha) \in \operatorname{cone}(C \times \{-1\}) \backslash \operatorname{cone}(F \times \{-1\})$ with $\alpha > 0$, it holds that $x \in C \backslash F$, then

$$\langle \alpha x,\, s\rangle - \alpha\theta = \alpha(\langle x,\, s\rangle - \theta) < 0.$$

Therefore, $\{(s, \theta)\}^{\perp}$ exposes $\operatorname{cone}(F \times \{-1\}) \trianglelefteq \operatorname{cone}(C \times \{-1\})$. The fact that $-(s, \theta) \in \overline{\operatorname{cone}}(C \times \{-1\})^*$ comes from (4.2).

($\Longleftarrow$) Suppose that $\{(s, \theta)\}^{\perp}$ exposes $\overline{\operatorname{cone}}(F \times \{-1\})$. One has

$$\begin{aligned}
&\langle s,\, \overline{x}\rangle - \alpha\theta < 0 \quad \forall\, (\overline{x}, -\alpha) \in \overline{\operatorname{cone}}(C \times \{-1\}) \backslash \overline{\operatorname{cone}}(F \times \{-1\}); \\
&\langle s,\, \overline{x}\rangle - \alpha\theta = 0 \quad \forall\, (\overline{x}, -\alpha) \in \overline{\operatorname{cone}}(F \times \{-1\}).
\end{aligned}$$

Then fix $\alpha = 1$, we know from $F \times \{-1\} = \overline{\operatorname{cone}}(F \times \{-1\}) \cap (\mathcal{E} \times \{-1\})$ that $(x, -1) \in F \times \{-1\}$. We thus have

$$\langle s,\, x\rangle - \theta < 0 \;\; \forall\, (x, -1) \in C \backslash F \times \{-1\}; \qquad \langle s,\, x\rangle - \theta = 0 \;\; \forall\, (x, -1) \in F \times \{-1\}.$$

This immediately shows that $F \trianglelefteq C$ is exposed by H with $s \in \operatorname{dom} \sigma_C$ and $\theta = \sigma_C(s)$. Then $F \trianglelefteq C$ is also exposed by H.

(ii) ($\Longrightarrow$) Suppose that $F \trianglelefteq C$ and let $z_1, z_2 \in \overline{\text{cone}}(C \times \{-1\})$ be such that $z_1 + z_2 \in \overline{\text{cone}}(F \times \{-1\})$. We need to show that z_1 and z_2 are both in $\overline{\text{cone}}(F \times \{-1\})$.

By [35, Theorem 8.2], we have $\overline{\text{cone}}(C \times \{-1\}) = \text{cone}(C \times \{-1\}) \cup (\text{rec}(C) \times \{0\})$. Therefore, without loss of generality, there are three possibilities for z_1, z_2:

(I) both z_1 and z_2 are in $\overline{\text{cone}}(C \times \{-1\}) \backslash (\text{rec}(C) \times \{0\})$;

(II) both z_1 and z_2 are in $\text{rec}(C) \times \{0\}$;

(III) $z_1 \in \overline{\text{cone}}(C \times \{-1\}) \backslash (\text{rec}(C) \times \{0\})$ and $z_2 \in \text{rec}(C) \times \{0\}$.

(I) We have $z_1 = (\alpha x_1, -\alpha), z_2 = (\beta x_2, -\beta)$ for some $\alpha > 0, \beta > 0$, $x_1, x_2 \in C$. From the assumption that $z_1 + z_2 \in \overline{\text{cone}}(F \times \{-1\})$, we know that there exist $\gamma > 0$ and $y \in F$ such that

$$z_1 + z_2 = (\alpha x_1 + \beta x_2, -\alpha - \beta) = (\gamma y, -\gamma).$$

Since $-\gamma = -\alpha - \beta < 0$, it holds that $(\gamma y, -\gamma) \in \overline{\text{cone}}(F \times \{-1\}) \backslash (\text{rec}(F) \times \{0\})$. Then we have,

$$\frac{\alpha}{\gamma} x_1 + \frac{\beta}{\gamma} x_2 = y \in F.$$

As α, β, γ are all positive, left-hand size of the above display is indeed a convex combination of x_1 and x_2. Recalling that $F \trianglelefteq C$, then we have $x_1, x_2 \in F$. Thus, we conclude that $z_1, z_2 \in \overline{\text{cone}}(F \times \{-1\}) \backslash (\text{rec}(F) \times \{0\})$.

(II) In this case, since $z_1 + z_2 \in \overline{\text{cone}}(F \times \{-1\}) = \text{cone}(F \times \{-1\}) \cup (\text{rec}(F) \times \{0\})$, we have $z_1 + z_2 \in \text{rec}(F) \times \{0\}$, since the last coordinates of z_1, z_2 are both zero. Because $\text{rec}(F) \trianglelefteq \text{rec}(C)$ by Lemma 4.1, then one can readily see that $z_1, z_2 \in \overline{\text{cone}}(F \times \{-1\})$.

(III) We have $z_1 = (\alpha x_1, -\alpha), z_2 = (\beta x_2, 0)$ for some $\alpha > 0$, $\beta \geqslant 0$, $x_1 \in C$, $x_2 \in \text{rec}(C)$. This leads to $z_1 + z_2 = (\alpha x_1 + \beta x_2, -\alpha) \in \overline{\text{cone}}(F \times \{-1\})$. Notice that since $\alpha > 0$, then $z_1 + z_2$ cannot be in $\text{rec}(F) \times \{0\}$.

If $\beta = 0$, i.e., $z_2 = 0$, then $z_1 + z_2 \in \overline{\text{cone}}(F \times \{-1\}) \backslash (\text{rec}(F) \times \{0\})$ immediately implies that $z_1, z_2 \in \overline{\text{cone}}(F \times \{-1\})$. Suppose that $\beta > 0$. Then, $z_1 + z_2 = (\alpha(x_1 + \frac{\beta}{\alpha} x_2), -\alpha) \in \overline{\text{cone}}(F \times \{-1\}) \backslash (\text{rec}(F) \times \{0\})$ implies that $x_1 + \frac{\beta}{\alpha} x_2 \in F$. Next, let $\eta \in (0,1)$ be arbitrary, we then have

$$x_1 + \frac{\beta}{\alpha} x_2 = (1 - \eta) x_1 + \eta (x_1 + \frac{\beta}{\eta \alpha} x_2) \in F.$$

Since $\frac{\beta}{\eta\alpha} > 0$ and $x_2 \in \text{rec}(C)$, it holds that $x_1 + \frac{\beta}{\eta\alpha} x_2 \in C$. This and the facts that $F \trianglelefteq C$ and $x_1 \in C$ imply $x_1 \in F$ and $x_1 + \frac{\beta}{\eta\alpha} x_2 \in F$. It then follows from the arbitrariness of $\eta \in (0,1)$ that $x_2 \in \text{rec}(F)$. Therefore, $z_1 \in \overline{\text{cone}}(F \times \{-1\}) \backslash (\text{rec}(F) \times \{0\})$, $z_2 \in \text{rec}(F) \times \{0\} \subseteq \overline{\text{cone}}(F \times \{-1\})$.

Combining the cases (I), (II), (III) above we conclude that $\overline{\text{cone}}(F \times \{-1\}) \trianglelefteq \overline{\text{cone}}(C \times \{-1\})$.

($\Longleftarrow$) We first note that $C \times \{-1\} = \overline{\text{cone}}(C \times \{-1\}) \cap (\mathcal{E} \times \{-1\})$ and that $F \times \{-1\} \trianglelefteq C \times \{-1\}$ if and only if $F \trianglelefteq C$. From [40, Theorem 7.13(1)], if $\overline{\text{cone}}(F \times \{-1\}) \trianglelefteq \overline{\text{cone}}(C \times -1)$ for $F \times \{-1\} \subseteq \mathcal{E} \times \{-1\}$, then $F \times \{-1\} = \overline{\text{cone}}(F \times \{-1\}) \cap (\mathcal{E} \times \{-1\})$ is a face of $C \times \{-1\}$ because $\mathcal{E} \times \{-1\}$ is the only face of $\mathcal{E} \times \{-1\}$. Therefore, $F \trianglelefteq C$.

$\square$

4.2 Nice convex sets

We recall that a face F of a convex cone $\mathcal{K} \subseteq \mathcal{E}$ is said to be *nice* if $F^* = \mathcal{K}^* + F^\perp$. If all faces of $\mathcal{K}$ are nice then $\mathcal{K}$ itself is said to be *nice*. The extension of the notion of niceness that we propose is based on the following observation. For any $F \trianglelefteq \mathcal{K}$, we have $F = \mathcal{K} \cap \operatorname{aff}(F)$ and then

$$\sigma_F = \sigma_{\mathcal{K}\cap\operatorname{aff}(F)} = (\delta_{\mathcal{K}} + \delta_{\operatorname{aff}(F)})^* = \operatorname{cl}(\sigma_{\mathcal{K}} \,\square\, \sigma_{\operatorname{aff}(F)}),$$

where the last equality follows from [35, Theorem 16.4]. It follows that $\operatorname{dom}\sigma_F = F^\circ = \operatorname{cl}(\operatorname{dom}\sigma_{\mathcal{K}} + \operatorname{dom}\sigma_{\operatorname{aff}(F)}) = \operatorname{cl}(\mathcal{K}^\circ + F^\perp)$, i.e., $F^* = \operatorname{cl}(\mathcal{K}^* + F^\perp)$. In general, we cannot drop the closure, so the assumption that F is nice amounts to requiring the closedness of $\mathcal{K}^* + F^\perp$.

Now, if C is a nonempty closed convex set and $F \trianglelefteq C$ is a face, we also have $F = C \cap \operatorname{aff} F$ which leads to $\sigma_F = \operatorname{cl}(\sigma_C \,\square\, \sigma_{\operatorname{aff}(F)})$ by [35, Theorem 16.4]. This suggests the following general definition of nice faces using the infimal convolution.

Definition 4.3 (Niceness of convex sets)**.** *A face F of a convex set $C \subseteq \mathcal{E}$ is said to be* nice *if $\sigma_F = \sigma_C \,\boxdot\, \sigma_{\operatorname{aff}(F)}$ holds. A convex set $C \subseteq \mathcal{E}$ is* nice *if all of its faces are nice.*

That is, in order for a face $F \trianglelefteq C$ to be nice two conditions must be met. First, the infimal convolution $\sigma_C \,\square\, \sigma_{\operatorname{aff}(F)}$ must be a closed function. Second, the infimum in the definition of $(\sigma_C \,\square\, \sigma_{\operatorname{aff}(F)})(s)$ must be attained for every $s \in \operatorname{dom}\sigma_F$.

Next we examine some basic properties of nice convex sets.

Proposition 4.4. *The following items hold.*

(i) Let C be a nice convex set and let $F \trianglelefteq C$ be a face, then F is a nice convex set.

(ii) Let $C_1, C_2 \subseteq \mathcal{E}$ be two nice convex sets with $C_1 \cap C_2 \neq \emptyset$, then $C_1 \cap C_2$ is nice.

Proof. (i) Let $\widehat{F}$ be a face of F. We need to show that $\sigma_{\widehat{F}} = \sigma_F \,\boxdot\, \sigma_{\operatorname{aff}\widehat{F}}$. We have

$$\sigma_F \,\square\, \sigma_{\operatorname{aff}\widehat{F}} \overset{\text{(a)}}{=} (\sigma_C \,\square\, \sigma_{\operatorname{aff} F}) \,\square\, \sigma_{\operatorname{aff}\widehat{F}} \overset{\text{(b)}}{=} \sigma_C \,\square\, (\sigma_{\operatorname{aff} F} \,\square\, \sigma_{\operatorname{aff}\widehat{F}}) \overset{\text{(c)}}{=} \sigma_C \,\square\, \sigma_{\operatorname{aff}\widehat{F}},$$

where (a) follows from the niceness of C; (b) follows from the associativity of the infimal convolution and (c) follows from $\operatorname{aff}\widehat{F} \subseteq \operatorname{aff} F$, so $\delta_{\operatorname{aff}\widehat{F}} = \delta_{\operatorname{aff}\widehat{F}} + \delta_{\operatorname{aff} F}$ which leads to $\sigma_{\operatorname{aff}\widehat{F}} = \operatorname{cl}(\sigma_{\operatorname{aff} F} \,\square\, \sigma_{\operatorname{aff}\widehat{F}}) = \sigma_{\operatorname{aff} F} \,\boxdot\, \sigma_{\operatorname{aff}\widehat{F}}$, where the infimal convolution is exact because $\operatorname{aff} F$ and $\operatorname{aff}\widehat{F}$ are relatively open sets that intersect.

We conclude that $\sigma_F \,\square\, \sigma_{\operatorname{aff}\widehat{F}} = \sigma_C \,\square\, \sigma_{\operatorname{aff}\widehat{F}}$. But $\widehat{F}$ is also a face of the nice set C, so $\sigma_C \,\square\, \sigma_{\operatorname{aff}\widehat{F}} = \sigma_C \,\boxdot\, \sigma_{\operatorname{aff}\widehat{F}}$ is a closed function. Therefore, $\sigma_F \,\square\, \sigma_{\operatorname{aff}\widehat{F}}$ is closed as well and we have

$$\sigma_{\widehat{F}} = \sigma_F \,\square\, \sigma_{\operatorname{aff}\widehat{F}} = \sigma_C \,\boxdot\, \sigma_{\operatorname{aff}\widehat{F}}. \tag{4.4}$$

It remains to verify that the infimal convolution $\sigma_F \,\square\, \sigma_{\operatorname{aff}\widehat{F}}$ is exact so let $s \in \operatorname{dom}\sigma_{\widehat{F}}$ and consider

$$(\sigma_F \,\square\, \sigma_{\operatorname{aff}\widehat{F}})(s) = \inf_{s_1} \sigma_F(s_1) + \sigma_{\operatorname{aff}\widehat{F}}(s - s_1). \tag{4.5}$$

From (4.4) there exists $s^* \in \mathcal{E}$ such that

$$(\sigma_F \,\square\, \sigma_{\operatorname{aff} \hat{F}})(s) = \sigma_C(s^*) + \sigma_{\operatorname{aff} \hat{F}}(s - s^*) \geqslant \sigma_F(s^*) + \sigma_{\operatorname{aff} \hat{F}}(s - s^*),$$

where the inequality holds because $F \subseteq C$ so $\sigma_F \leqslant \sigma_C$. This shows that the infimum in (4.5) is attained at s^*.

(*ii*) Without loss of generality, we can assume that $\operatorname{ri}(C_1) \cap \operatorname{ri}(C_2) \neq \varnothing$, otherwise by applying Proposition 2.3 with the trivial face $C_1 \cap C_2$, we can always find $F_1 \trianglelefteq C_1$ and $F_2 \trianglelefteq C_2$ such that $F_1 \cap F_2 = C_1 \cap C_2$ and $\operatorname{ri}(F_1) \cap \operatorname{ri}(F_2) \neq \varnothing$. Furthermore, by the previous item, the faces F_1 and F_2 are nice convex sets as well.

Let $C := C_1 \cap C_2$. By Proposition 2.3, for any $F \trianglelefteq C$, there exist $F_1 \trianglelefteq C_1$ and $F_2 \trianglelefteq C_2$ such that $F = F_1 \cap F_2$ and $\operatorname{ri}(F) = \operatorname{ri}(F_1) \cap \operatorname{ri}(F_2) \neq \varnothing$. Now, we have

$$\begin{aligned}&\sigma_F \overset{(a)}{=} \sigma_{F_1 \cap F_2} \overset{(b)}{=} \sigma_{F_1} \boxdot \sigma_{F_2} \overset{(c)}{=} \sigma_{C_1} \boxdot \sigma_{\operatorname{aff}(F_1)} \boxdot \sigma_{C_2} \boxdot \sigma_{\operatorname{aff}(F_2)} \overset{(d)}{=} \sigma_{C_1} \boxdot \sigma_{C_2} \boxdot \sigma_{\operatorname{aff}(F_1)} \boxdot \sigma_{\operatorname{aff}(F_2)}\\ &\overset{(e)}{=} \sigma_{C_1 \cap C_2} \boxdot \sigma_{\operatorname{aff}(F_1) \cap \operatorname{aff}(F_2)} \overset{(f)}{=} \sigma_C \boxdot \sigma_{\operatorname{aff}(F_1 \cap F_2)} \overset{(g)}{=} \sigma_C \boxdot \sigma_{\operatorname{aff}(F)},\end{aligned}$$

where (a) and (b) come from $F = F_1 \cap F_2$ and $\operatorname{ri}(F_1) \cap \operatorname{ri}(F_2) \neq \varnothing$; (c) is true thanks to the niceness of C_1 and C_2; (d) holds because of the commutativity and the associativity of infimal convolution; (e) is true since $\operatorname{ri}(C_1) \cap \operatorname{ri}(C_2) \neq \varnothing$ and $\operatorname{ri}(\operatorname{aff}(F_1)) \cap \operatorname{ri}(\operatorname{aff}(F_2)) = \operatorname{aff}(F_1) \cap \operatorname{aff}(F_2) \neq \varnothing$; (f) comes from $C = C_1 \cap C_2$ and $\operatorname{aff}(F_1) \cap \operatorname{aff}(F_2) = \operatorname{aff}(F_1 \cap F_2)$ by Lemma 2.1; (g) holds by $F = F_1 \cap F_2$. Consequently, we have $\sigma_F = \sigma_C \boxdot \sigma_{\operatorname{aff}(F)}$ for any $F \trianglelefteq C$. Then $C = C_1 \cap C_2$ is nice. □

The next theorem states the connections between the niceness of a closed convex set and its conic lifting.

Theorem 4.5 (Niceness of conic lifting)**.** *For any closed convex set $C \subseteq \mathcal{E}$, the following hold:*

(i) If $\overline{\operatorname{cone}}(C \times \{-1\})$ is nice, then C is nice.

(ii) If C is compact, then C is nice if and only if $\overline{\operatorname{cone}}(C \times \{-1\})$ is nice.

Proof. (*i*) Thanks to Proposition 4.4, the observation $C \times \{-1\} = \overline{\operatorname{cone}}(C \times \{-1\}) \cap (\mathcal{E} \times \{-1\})$ and the fact that any affine subspace is nice, we deduce that $C \times \{-1\}$ is nice. The niceness of C follows from this and that $C \times \{-1\}$ is nice if and only if C is nice.

(*ii*) Using (i), C is nice if $\overline{\operatorname{cone}}(C \times \{-1\})$ is nice, so we only need to show the converse. We first recall that $\overline{\operatorname{cone}}(C \times \{-1\}) = \operatorname{cone}(C \times \{-1\}) \cup (\operatorname{rec}(C) \times \{0\})$ from [35, Theorem 8.2]. If C is compact, then $\operatorname{rec}(C) = \{0\}$ and hence $\overline{\operatorname{cone}}(C \times \{-1\}) = \operatorname{cone}(C \times \{-1\})$. Then by Proposition 4.2(ii), except for the trivial face $\{0\}$, all faces of $\overline{\operatorname{cone}}(C \times \{-1\})$ admit the form $\overline{\operatorname{cone}}(F \times \{-1\})$ with some $F \trianglelefteq C$.

If C is nice, then for any $F \trianglelefteq C$, we have

$$\overline{\operatorname{cone}}(F \times \{-1\})^\circ \overset{(a)}{=} \operatorname{epi} \sigma_F \overset{(b)}{=} \operatorname{epi} \sigma_C + \operatorname{epi} \sigma_{\operatorname{aff}(F)} \overset{(c)}{=} \overline{\operatorname{cone}}(C \times \{-1\})^\circ + \overline{\operatorname{cone}}(F \times \{-1\})^\perp,$$

where (a) comes from (4.2); (b) holds due to the exactness of the infimal convolution from the niceness of C; (c) is true thanks to (4.2) and (4.3). This shows that $\overline{\operatorname{cone}}(F \times \{-1\})$ is a nice face of $\overline{\operatorname{cone}}(C \times \{-1\})$. Since $\{0\}$ is always nice, then $\overline{\operatorname{cone}}(C \times \{-1\})$ is nice since all of its faces are nice.

Consequently, we know that C is nice if and only if $\overline{\operatorname{cone}}(C \times \{-1\})$ is nice. □

Pataki proved that nice closed convex cones are facially exposed [29]. Combining Theorem 4.5 with Proposition 4.2, we conclude that a nice compact convex set must be facially exposed. Since we do not know if the compactness assumption in item (*ii*) can be relaxed, the discussion so far does not allow us to conclude that nice closed convex sets are facially exposed. However, later in Corollary 6.6, as a consequence of a more general discussion on extended duality, we will indeed show that nice closed convex sets are facially exposed.

4.3 Vertically nice functions

For convex functions, we also consider a notion of niceness, which as expected, is defined in terms of its epigraph. However, requiring the niceness of the whole epigraph is unnecessary for our purposes.

Definition 4.6 (Vertical faces and niceness)**.** *Let $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function and let $\widehat{F} \trianglelefteq \operatorname{epi} f$. $\widehat{F}$ is said to be a* vertical face *if $(x, \mu) \in \widehat{F}$ implies that $(x, \bar{\mu}) \in \widehat{F}$ for all $\bar{\mu} \geqslant \mu$. We say that f is* vertically nice *if every vertical face of $\operatorname{epi} f$ is nice.*

When applying the facial reduction algorithm to regularize (1.1), we obtain the minimal face of $\operatorname{dom} f$ containing the feasible region $\operatorname{dom} f \cap \operatorname{dom} g$. Denoting this minimal face by F, it is natural to consider the restriction $f_{|F}$ and its epigraph $\operatorname{epi} f_{|F}$, see (3.7). As we will show shortly in Lemma 4.7, $\operatorname{epi} f_{|F}$ is a vertical face of $\operatorname{epi} f$. In this sense, assuming vertical niceness of f amounts to requiring that all possible faces of interest (i.e., vertical faces) of $\operatorname{epi} f$ possess good properties, especially the ability to simplify $(f_{|F})^*$ as we will see in Theorem 4.8.

Lemma 4.7 (Vertical faces and faces of domain)**.** *Let $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. The following statements hold.*

(i) Suppose $F \subseteq \operatorname{dom} f$ is a nonempty convex set. Then, $\operatorname{epi} f_{|F}$ is a vertical face of $\operatorname{epi} f$ if and only if $F \trianglelefteq \operatorname{dom} f$.

(ii) If $\widehat{F} \trianglelefteq \operatorname{epi} f$ is a vertical face, then $F := \{x \mid \exists \mu \text{ s.t. } (x, \mu) \in \widehat{F}\}$ is a face of $\operatorname{dom} f$ satisfying $\widehat{F} = \operatorname{epi} f_{|F}$.

Proof. (*i*) Suppose that $\operatorname{epi} f_{|F}$ is a vertical face of $\operatorname{epi} f$ and let $x_1, x_2 \in \operatorname{dom} f$, $\alpha \in (0, 1)$ be such that $z := \alpha x_1 + (1 - \alpha)x_2 \in F$. Then, we have $(z, f(z)) \in \operatorname{epi} f_{|F}$. Since f is convex we have

$$f(z) \leqslant \alpha f(x_1) + (1 - \alpha)f(x_2).$$

By the definition of $\operatorname{epi} f_{|F}$ this leads to $(z, \alpha f(x_1) + (1 - \alpha)f(x_2)) \in \operatorname{epi} f_{|F}$. That is, $\alpha(x_1, f(x_1)) + (1 - \alpha)(x_2, f(x_2)) \in \operatorname{epi} f_{|F}$, which implies, in particular, that $x_1, x_2 \in F$, since $\operatorname{epi} f_{|F}$ is a face. Then, $F \trianglelefteq \operatorname{dom} f$.

Conversely, suppose that $F \trianglelefteq \operatorname{dom} f$. Let (x_1, μ_1), $(x_2, \mu_2) \in \operatorname{epi} f$, $\alpha \in (0, 1)$ be such that $\alpha(x_1, \mu_1) + (1 - \alpha)(x_2, \mu_2) \in \operatorname{epi} f_{|F}$. In particular, this implies $x_1, x_2 \in \operatorname{dom} f$, $\alpha x_1 + (1 - \alpha)x_2 \in F$. Therefore, $x_1, x_2 \in F$ and $(x_1, \mu_1), (x_2, \mu_2)$ belong to $\operatorname{epi} f_{|F}$, by definition. This shows that $\operatorname{epi} f_{|F}$ is a face of $\operatorname{epi} f$, which must be vertical by definition.

(*ii*) First we check that $\widehat{F} = \operatorname{epi} f_{|F}$ holds. Let $(x, \mu) \in \widehat{F}$, then $x \in F$ and $f(x) \leqslant \mu$ holds, so $(x, \mu) \in \operatorname{epi} f_{|F}$.

Conversely, suppose that $(x,\mu) \in \operatorname{epi} f_{|F}$, i.e., $f(x) \leqslant \mu$ and $x \in F$. Then, there exists μ_1 such that $(x,\mu_1) \in \widehat{F}$ and $f(x) \leqslant \mu_1$. If $\mu \geqslant \mu_1$ holds, then $(x,\mu) \in \widehat{F}$ because $\widehat{F}$ is assumed to be vertical. So suppose that $\mu < \mu_1$ holds. Let $\mu_2 := \mu_1 + 1$ so that $(x,\mu_2) \in \operatorname{epi} f$ and we have $\mu < \mu_1 < \mu_2$.

Letting $\alpha := \frac{\mu_1-\mu}{\mu_2-\mu}$, we have

$$\mu_1 = \mu + \alpha(\mu_2 - \mu),$$

which implies that $(x,\mu_1) = (1-\alpha)(x,\mu) + \alpha(x,\mu_2)$. Because (x,μ) and (x,μ_2) both belong to $\operatorname{epi} f$, $\alpha \in (0,1)$ and $\widehat{F}$ is a face of $\operatorname{epi} f$ containing (x,μ_1), we conclude that (x,μ) belongs to $\widehat{F}$.

Having checked that $\widehat{F} = \operatorname{epi} f_{|F}$ holds, this item then follows from item (*i*). $\square$

Theorem 4.8 (Characterization of niceness of a vertical face)**.** *Let $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function and $F \trianglelefteq \operatorname{dom} f$. Then $\widehat{F} := \operatorname{epi} f_{|F}$ is nice if and only if F is nice and $(f_{|F})^* = f^* \boxdot \sigma_{\operatorname{aff}(F)}$.*

Proof. We start by setting up some formulae. Since $\operatorname{aff}(\widehat{F}) = \operatorname{aff}(F) \times \mathbb{R}$, then for any $(x,\mu) \in \mathcal{E} \times \mathbb{R}$,

$$\sigma_{\operatorname{aff}(\hat{F})}(x,\mu) = \begin{cases} +\infty & \text{if } \mu \neq 0, \\ \sigma_{\operatorname{aff}(\hat{F})}(x,0) = \sigma_{\operatorname{aff}(F)}(x) & \text{if } \mu = 0. \end{cases} \tag{4.6}$$

Moreover, by [11, X Proposition 1.2.1], for any $(s,\tau) \in \mathcal{E} \times \mathbb{R}$, we have

$$\sigma_{\hat{F}}(s,\tau) = \begin{cases} |\tau|(f_{|F})^*\left(\frac{s}{|\tau|}\right) & \text{if } \tau < 0, \\ \sigma_{\hat{F}}(s,0) = \sigma_F(s) & \text{if } \tau = 0, \\ +\infty & \text{if } \tau > 0. \end{cases} \tag{4.7}$$

For any $s \in \mathcal{E}$, it holds that

$$\begin{aligned} &(\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\hat{F})})(s,0) = \inf_{(x,\mu)} \left\{\sigma_{\operatorname{epi} f}(s-x,0-\mu) + \sigma_{\operatorname{aff}(\hat{F})}(x,\mu)\right\} \\ &\overset{\text{(a)}}{=} \inf_x \left\{\sigma_{\operatorname{epi} f}(s-x,0) + \sigma_{\operatorname{aff}(F)}(x)\right\} \overset{\text{(b)}}{=} \inf_x \left\{\sigma_{\operatorname{dom} f}(s-x) + \sigma_{\operatorname{aff}(F)}(x)\right\} \\ &= \left(\sigma_{\operatorname{dom} f} \square \sigma_{\operatorname{aff}(F)}\right)(s), \end{aligned} \tag{4.8}$$

where (a) holds by (4.6); (b) comes from (4.7) with F replaced with $\operatorname{dom} f$ and $\tau = 0$. This shows that $\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\hat{F})}$ is exact at $(s,0)$ if and only if $\sigma_{\operatorname{dom} f} \square \sigma_{\operatorname{aff}(F)}$ is exact at s.[3]

Meanwhile, for any $s \in \mathcal{E}$,

$$\begin{aligned} &(\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\hat{F})})(s,-1) = \inf_{(x,\mu)} \left\{\sigma_{\operatorname{epi} f}(s-x,-1-\mu) + \sigma_{\operatorname{aff}(\hat{F})}(x,\mu)\right\} \\ &\overset{\text{(a)}}{=} \inf_x \left\{\sigma_{\operatorname{epi} f}(s-x,-1) + \sigma_{\operatorname{aff}(F)}(x)\right\} \overset{\text{(b)}}{=} \inf_x \left\{f^*(s-x) + \sigma_{\operatorname{aff}(F)}(x)\right\} \\ &= (f^* \square \sigma_{\operatorname{aff}(F)})(s), \end{aligned} \tag{4.9}$$

[3]This is not only because the infima in (4.8) all coincide; it is also because $\sigma_{\operatorname{epi} f}(s-x,-\mu) + \sigma_{\operatorname{aff}(\hat{F})}(x,\mu) = \sigma_{\operatorname{dom} f}(s-x) + \sigma_{\operatorname{aff}(F)}(x)$ holds for all x and μ. A similar comment applies to (4.9).

where (a) holds by (4.6); (b) comes from (4.7) with F replaced with $\operatorname{dom} f$ and $\tau = -1$. Similarly, this shows that $\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})}$ is exact at $(s, -1)$ if and only if $f^* \square \sigma_{\operatorname{aff}(F)}$ is exact at s.

($\Longrightarrow$) We first check that F is a nice face of $\operatorname{dom} f$. For any $s \in \mathcal{E}$, we have,

$$\sigma_F(s) \overset{\text{(a)}}{=} \sigma_{\widehat{F}}(s, 0) \overset{\text{(b)}}{=} (\sigma_{\operatorname{epi} f} \boxdot \sigma_{\operatorname{aff}(\widehat{F})})(s, 0) \overset{\text{(c)}}{=} \left(\sigma_{\operatorname{dom} f} \square \sigma_{\operatorname{aff}(F)}\right)(s),$$

where (a) holds by (4.7); (b) is true thanks to the niceness of $\widehat{F}$, i.e., $\sigma_{\widehat{F}} = \sigma_{\operatorname{epi} f} \boxdot \sigma_{\operatorname{aff}(\widehat{F})}$; (c) comes from (4.8). Since $\sigma_{\operatorname{epi} f} \boxdot \sigma_{\operatorname{aff}(\widehat{F})}$ is exact at $(s, 0)$, we conclude that $\sigma_{\operatorname{dom} f} \square \sigma_{\operatorname{aff}(F)}$ is exact at s. As s is arbitrary, this shows that F is nice.

Similarly, for any $s \in \mathcal{E}$,

$$(f_{|F})^*(s) \overset{\text{(a)}}{=} \sigma_{\widehat{F}}(s, -1) \overset{\text{(b)}}{=} (\sigma_{\operatorname{epi} f} \boxdot \sigma_{\operatorname{aff}(\widehat{F})})(s, -1) \overset{\text{(c)}}{=} (f^* \square \sigma_{\operatorname{aff}(F)})(s),$$

where (a) comes from (4.7) with $\tau = -1$; (b) is true thanks to the niceness of $\widehat{F}$; (c) holds by (4.9). Since $\sigma_{\operatorname{epi} f} \boxdot \sigma_{\operatorname{aff}(\widehat{F})}$ is exact at $(s, -1)$, we conclude that $f^* \square \sigma_{\operatorname{aff}(F)}$ is exact at s. Hence, $(f_{|F})^* = f^* \boxdot \sigma_{\operatorname{aff}(F)}$, since s is arbitrary.

($\Longleftarrow$) Let $(s, \tau) \in \mathcal{E} \times \mathbb{R}$ and we will consider three cases. For $\tau = 0$, we have

$$\sigma_{\widehat{F}}(s, 0) \overset{\text{(a)}}{=} \sigma_F(s) \overset{\text{(b)}}{=} (\sigma_{\operatorname{dom} f} \boxdot \sigma_{\operatorname{aff}(F)})(s) \overset{\text{(c)}}{=} (\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})})(s, 0),$$

where (a) holds by (4.7); (b) comes from the niceness of F; (c) is true because of (4.8). Since $\sigma_{\operatorname{dom} f} \boxdot \sigma_{\operatorname{aff}(F)}$ is exact at s, we conclude that $\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})}$ is also exact at $(s, 0)$.

For $\tau < 0$, since support functions are positively homogeneous so that $\sigma_{\widehat{F}}(u, \tau) = |\tau| \sigma_{\widehat{F}}(\frac{u}{|\tau|}, -1)$ holds for every $u \in \mathcal{E}$, it is enough to consider the case $\tau = -1$. It holds that

$$\begin{aligned}
\sigma_{\widehat{F}}(s, -1) &\overset{\text{(a)}}{=} (f_{|F})^*(s) \overset{\text{(b)}}{=} \left(f^* \boxdot \sigma_{\operatorname{aff}(F)}\right)(s) = \inf_x \left\{f^*(s - x) + \sigma_{\operatorname{aff}(F)}(x)\right\} \\
&\overset{\text{(c)}}{=} \inf_x \left\{\sigma_{\operatorname{epi} f}(s - x, -1) + \sigma_{\operatorname{aff}(F)}(x)\right\} \\
&\overset{\text{(d)}}{=} \inf_{(x, \mu)} \left\{\sigma_{\operatorname{epi} f}(s - x, -1 - \mu) + \sigma_{\operatorname{aff}(\widehat{F})}(x, \mu)\right\} = (\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})})(s, -1),
\end{aligned}$$

where (a) holds by (4.7); (b) comes from the assumption that $(f_{|F})^* = f^* \boxdot \sigma_{\operatorname{aff}(F)}$; (c) is true thanks to (4.7) with F replaced with $\operatorname{dom} f$; (d) holds because of (4.6). Since $f^* \boxdot \sigma_{\operatorname{aff}(F)}$ is exact at s, we conclude that $\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})}$ is exact at $(s, -1)$ as well.

Finally, for $\tau > 0$ it can be readily shown that $(\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})})(s, \tau) = +\infty$. Indeed, we have $(\sigma_{\operatorname{epi} f} \square \sigma_{\operatorname{aff}(\widehat{F})})(s, \tau) = \inf_x \left\{\sigma_{\operatorname{epi} f}(s - x, \tau) + \sigma_{\operatorname{aff}(F)}(x)\right\} = +\infty$ because $\sigma_{\operatorname{epi} f}(s - x, \tau) = +\infty$ for $\tau > 0$.

Overall, we conclude that $\sigma_{\widehat{F}} = \sigma_{\operatorname{epi} f} \boxdot \sigma_{\operatorname{aff}(\widehat{F})}$, i.e., $\widehat{F}$ is a nice face of $\operatorname{epi} f$. $\square$

Combining Theorem 4.8 with Lemma 4.7, we have the following corollary.

Corollary 4.9 (Characterization of vertical niceness of functions)**.** *Let $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Then f is vertically nice if and only if* $\operatorname{dom} f$ *is nice and for any* $F \trianglelefteq \operatorname{dom} f$*,* $(f_{|F})^* = f^* \boxdot \sigma_{\operatorname{aff}(F)}$.

The next proposition states that the sum of two convex vertically nice functions is also vertically nice if their domains have relative interior intersection.

Proposition 4.10 (Sum of vertically nice functions)**.** *Let* $f_1, f_2 : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ *be two vertically nice functions. If* $\operatorname{ri}(\operatorname{dom} f_1) \cap \operatorname{ri}(\operatorname{dom} f_2) \neq \emptyset$*, then* $f := f_1 + f_2$ *is also vertically nice.*

Proof. Since $\operatorname{ri}(\operatorname{dom} f_1) \cap \operatorname{ri}(\operatorname{dom} f_2) \neq \emptyset$, we have that f is a proper convex function. By Corollary 4.9, to show that f is also vertically nice, it suffices to show that $\operatorname{dom} f = \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ is nice and that for any $F \trianglelefteq \operatorname{dom} f$, $(f_{|F})^* = f^* \boxdot \sigma_{\operatorname{aff}(F)}$.

Since both f_1 and f_2 are vertically nice, Corollary 4.9 implies that both $\operatorname{dom} f_1$ and $\operatorname{dom} f_2$ are nice, then $\operatorname{dom} f = \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ is also nice by Proposition 4.4. Moreover, for any $F_1 \trianglelefteq \operatorname{dom} f_1$ and $F_2 \trianglelefteq \operatorname{dom} f_2$, we have

$$(f_{1|F_1})^* = f_1^* \boxdot \sigma_{\operatorname{aff}(F_1)} \quad \text{and} \quad (f_{2|F_2})^* = f_2^* \boxdot \sigma_{\operatorname{aff}(F_2)}. \tag{4.10}$$

Let $F \trianglelefteq \operatorname{dom} f = \operatorname{dom} f_1 \cap \operatorname{dom} f_2$. By Proposition 2.3, there exist $F_1 \trianglelefteq \operatorname{dom} f_1$ and $F_2 \trianglelefteq \operatorname{dom} f_2$ such that $F = F_1 \cap F_2$ and $\operatorname{ri}(F_1) \cap \operatorname{ri}(F_2) \neq \emptyset$. Therefore,

$$\begin{aligned}(f_{|F})^* &= (f_{|F_1 \cap F_2})^* = (f_{1|F_1} + f_{2|F_2})^* \overset{\text{(a)}}{=} (f_{1|F_1})^* \boxdot (f_{2|F_2})^* \overset{\text{(b)}}{=} f_1^* \boxdot \sigma_{\operatorname{aff}(F_1)} \boxdot f_2^* \boxdot \sigma_{\operatorname{aff}(F_2)} \\ &= f_1^* \boxdot f_2^* \boxdot \sigma_{\operatorname{aff}(F_1)} \boxdot \sigma_{\operatorname{aff}(F_2)} \overset{\text{(c)}}{=} (f_1 + f_2)^* \boxdot \sigma_{\operatorname{aff}(F_1)} \boxdot \sigma_{\operatorname{aff}(F_2)} \overset{\text{(d)}}{=} f^* \boxdot \sigma_{\operatorname{aff}(F)},\end{aligned}$$

where (a) follows from [35, Theorem 16.4] and the condition $\operatorname{ri}(F_1) \cap \operatorname{ri}(F_2) \neq \emptyset$; (b) comes from (4.10); (c) is true because $f_1^* \boxdot f_2^* = (f_1 + f_2)^*$ by [35, Theorem 16.4] and $\operatorname{ri}(\operatorname{dom} f_1) \cap \operatorname{ri}(\operatorname{dom} f_2) \neq \emptyset$; (d) holds due to the observation that $\sigma_{\operatorname{aff}(F_1)} \boxdot \sigma_{\operatorname{aff}(F_2)} = \sigma_{\operatorname{aff}(F_1) \cap \operatorname{aff}(F_2)} = \sigma_{\operatorname{aff}(F_1 \cap F_2)} = \sigma_{\operatorname{aff}(F)}$ since $\operatorname{ri}(F_1) \cap \operatorname{ri}(F_2) \neq \emptyset$. This shows that $f = f_1 + f_2$ is vertically nice. ☐

Making use of the results so far, we can construct several examples of vertically nice functions.

Corollary 4.11 (Examples of vertically nice functions)**.** *Let* $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ *be a proper convex function and* C *be a convex set. It holds that*

(i) If $\operatorname{dom} f$ *is an affine subspace, then* f *is vertically nice.*

(ii) If C *is nice, then* δ_C *is vertically nice.*

(iii) If f *is vertically nice,* C *is nice and* $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(C) \neq \emptyset$*, then* $f + \delta_C$ *is vertically nice.*

Proof. (*i*) If $\operatorname{dom} f$ is affine, its only face is $F = \operatorname{dom} f$. As $f = f + \delta_F$ we have $f^* = f^* \boxdot \sigma_F$ either by direct verification or by observing that $(\operatorname{ri} \operatorname{dom} f) \cap \operatorname{ri} F \neq \emptyset$ holds and invoking [35, Theorems 12.2 and 16.4].

(*ii*) We have $\operatorname{epi} \delta_C = C \times \mathbb{R}_+$. By Lemma 4.7, any vertical face $\widehat{F}$ of $\operatorname{epi} \delta_C$ takes the form $\widehat{F} := \operatorname{epi} \delta_F = F \times \mathbb{R}_+$ for some $F \trianglelefteq C$. In view of the fact that niceness is preserved by direct products (which follows directly from the definition) and the niceness of F by Proposition 4.4(i), as well as the niceness of $\mathbb{R}_+$, we know that $\widehat{F}$ is a nice face of $\operatorname{epi} \delta_C$, which follows that δ_C is vertically nice.

(*iii*) It follows directly from Proposition 4.10 and item (*ii*). ☐

Corollary 4.11 tells us that $f := \langle c, \cdot\rangle + \delta_{\mathcal{K}}$, with $c \in \mathcal{E}$ and $\mathcal{K} \subseteq \mathcal{E}$ being a nice closed convex cone, is vertically nice. Therefore, a conic linear optimization problem with a nice cone is a particular case of minimizing a vertically nice function over a nice set.

Before we move on, we discuss an example of a vertically nice function whose epigraph is not a nice convex set.

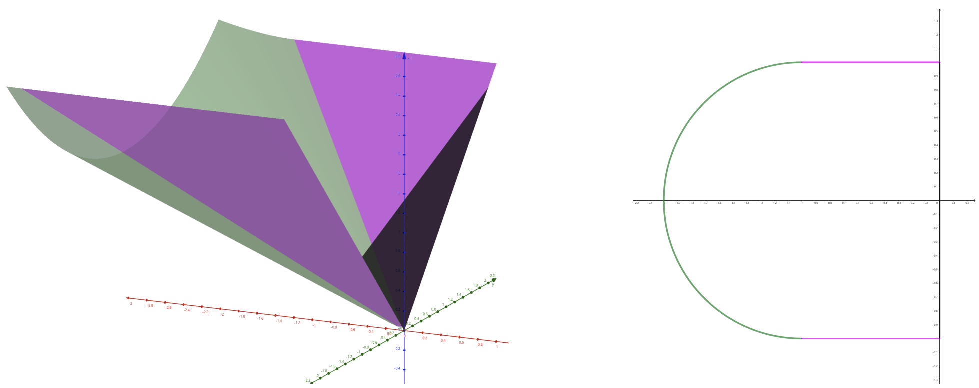
Example 4.12. *Consider the closed proper convex function defined as*

$$f(x,y) := \begin{cases} |y| & \text{if } x = 0, \\ \min\left\{\frac{x^2+y^2}{-2x}, \max\{|x|,|y|\}\right\} & \text{if } x < 0. \end{cases}$$

The graph of f is shown in Figure 1(a). We note that f is positively homogeneous, so $\operatorname{epi} f$ is a closed convex cone. Indeed, $\operatorname{epi} f = \operatorname{cone}(C \times \{1\})$, where

$$C := \{(x,y) \mid (x+1)^2 + y^2 \leqslant 1, x \leqslant 0\} \cup \{(x,y) \mid x \leqslant 0, \max\{|x|,|y|\} \leqslant 1\}$$

is shown in Figure 1(b). Moreover, one can readily see that the two faces at the junction of the purple and green surfaces in Figure 1(a), i.e., $F_1 := \{(x,y,z) \mid -x = y = z\}$ and $F_2 := \{(x,y,z) \mid -x = -y = z\}$, are not exposed. Therefore, by [29, Theorem 3], $\operatorname{epi} f$ is not nice since it is not facially exposed. However, using Corollary 4.11, we can conclude f is vertically nice by noting that



(a) Graph of $f(x,y)$.

(b) The closed convex set C such that $\operatorname{epi} f = \operatorname{cone}(C \times \{1\})$.

Figure 1: Graph of $f(x,y)$

$f = g + \delta_{\mathbb{R}_- \times \mathbb{R}}$ *where*

$$g(x,y) := \begin{cases} |y| & \text{if } x \geqslant 0, \\ \min\left\{\frac{x^2+y^2}{-2x}, \max\{|x|,|y|\}\right\} & \text{if } x < 0 \end{cases}$$

is vertically nice (because $\operatorname{dom} g = \mathbb{R}^2$) and $\mathbb{R}_- \times \mathbb{R}$ is nice.

4.4 The regularized Fenchel dual revisited

We return to the problem in (1.1). As long as $(\operatorname{dom} f) \cap (\operatorname{dom} g)$ is nonempty, if we have access to $F^1_{\min}$ and $F^2_{\min}$ (computed, say, through Algorithm 1) as in (3.2) we can form the regularized Fenchel dual (3.7) and, by Theorem 3.1, the pair (1.1), (3.7) satisfies strong duality as in Definition 2.9.

If we further assume that f and g are vertically nice, then (3.7) can be equivalently rewritten as

$$\sup_{\lambda} -f^* \boxdot \sigma_{\operatorname{aff}(F^1_{\min})}(-\lambda) - g^* \boxdot \sigma_{\operatorname{aff}(F^2_{\min})}(\lambda). \tag{4.11}$$

"Opening up" the infimal convolution leads to the following theorem.

Theorem 4.13. *Suppose that f, g are vertically nice proper closed convex functions satisfying $(\operatorname{dom} f) \cap (\operatorname{dom} g) \neq \emptyset$. Let $F^1_{\min}$ and $F^2_{\min}$ be as in (3.2) and consider the following problem*

$$\sup_{\lambda,y,z} - f^*(-y-\lambda) - \sigma_{\operatorname{aff}(F^1_{\min})}(y) - g^*(-z+\lambda) - \sigma_{\operatorname{aff}(F^2_{\min})}(z). \tag{4.12}$$

The pair (1.1) and (4.12) satisfy strong duality.

Proof. From the definition of exact infimal convolution, we have

$$-f^* \boxdot \sigma_{\operatorname{aff}(F^1_{\min})}(-\lambda) = \sup_{y} -f^*(-\lambda - y) - \sigma_{\operatorname{aff}(F^1_{\min})}(y),$$
$$-g^* \boxdot \sigma_{\operatorname{aff}(F^2_{\min})}(\lambda) = \sup_{z} -g^*(\lambda - z) - \sigma_{\operatorname{aff}(F^2_{\min})}(z),$$

where exactness ensures that for each λ both suprema are attained at some y and z, whenever the left-hand-sides are finite. Therefore, not only the optimal values of (4.11) and (4.12) coincide, but (4.12) must be attained as well. As Theorem 3.1 implies that (1.1) and (4.11) satisfy strong duality, the same is true of the pair (1.1) and (4.12). ☐

The support function of an affine set is just a linear function restricted to some subspace (see Lemma 2.6). Thus, in the setting of Theorem 4.13, the difference between (4.12) and (1.2) is that the domains of f^* and g^* are enlarged along subspaces and linear functions are added to f^* and g^*. This is a comparatively mild regularization, provided that we have $F^1_{\min}$ and $F^2_{\min}$ at hand.

5 Extended Duals

The disadvantage of the discussion in Sections 3 and 4 is that the minimal faces as in (3.2) need to be *explicitly determined*. Inspired by Ramana's dual [33] and analogous developments in conic linear programming [30, 18, 19], we leverage the facial reduction algorithm in Algorithm 1 to construct an *extended dual* for (1.1) in terms of f, g, $C_1 = \operatorname{dom} f$, $C_2 = \operatorname{dom} g$ and convex objects associated to them. This extended dual, although quite involved, will not require that the minimal face be explicitly determined.

In view of (3.7), what we need are appropriate expressions for the support functions of $\operatorname{aff}(F^1_{\min})$ and $\operatorname{aff}(F^2_{\min})$. Unfortunately, we will only be able to do so in Section 6 where we assume vertical niceness and the closedness of the domains of f and g. In the general case, expressions for $\sigma_{\mathrm{aff}(F^1_{\min})}$ and $\sigma_{\mathrm{aff}(F^2_{\min})}$ seem hard to be obtained directly, so in this section we will settle for the next best option which is to get expressions for the support functions of $F^1_{\min}$ and $F^2_{\min}$. Recalling (3.6), we have

$$(f_{|F^1_{\min}})^* = \mathrm{cl}(f^* \,\square\, \sigma_{F^1_{\min}}), \qquad (g_{|g^1_{\min}})^* = \mathrm{cl}(g^* \,\square\, \sigma_{F^2_{\min}}),$$

so obtaining expressions for $\sigma_{F^1_{\min}}$ and $\sigma_{F^2_{\min}}$ will also be enough for the purpose of constructing an extended dual.

Thus the task before us now is the computation of $\sigma_{F^1_{\min}}$ and $\sigma_{F^2_{\min}}$ without having explicit access to neither $F^1_{\min}$ nor $F^2_{\min}$. In the case of CLPs, the discussion in [30, 18, 19] suggests that a path towards obtaining an extended dual for CLPs is to "encode" the facial reduction algorithm as conic linear constraints.

Here, given arbitrary nonempty convex sets $C_1, C_2 \subseteq \mathcal{E}$, we will follow a similar idea and encode Algorithm 1 via one large set of convex constraints involving C_1 and C_2. Then, based on that, we will determine formulae for the support functions of $F^1_{\min}$ and $F^2_{\min}$, where $F^1_{\min}$ and $F^2_{\min}$ are as in (3.2).

For this section we assume that $C_1 \cap C_2 \neq \emptyset$, so that $F^1_{\min}$ and $F^2_{\min}$ (cf. (3.2)) exist. We first consider a simplified version of Algorithm 1 where we no longer require the separation to be proper and allow for trivial iterations where $(s_k, \theta_k) = (0, 0)$ may hold. However, since $C_1 \cap C_2 \neq \emptyset$ holds by assumption, we may assume that θ_k is as in Lemma 3.4(*i*).

Algorithm 2 A conceptual facial reduction algorithm for when $C_1 \cap C_2 \neq \emptyset$

Step 0. Set $F^1_0 \leftarrow C_1$ $F^2_0 \leftarrow C_2$ and $k \leftarrow 0$.

Step 1. Let (s_k, θ_k) be such that $s_k \in (F^2_k - F^1_k)^*$ and $\inf_{y\in F^2_k}\langle s_k,\, y\rangle = \theta_k = \sup_{y\in F^1_k}\langle s_k,\, y\rangle$. Set $H_k \leftarrow \{z \mid \langle s_k,\, z\rangle = \theta_k\}$.

Step 2. Set $F^1_{k+1} \leftarrow F^1_k \cap H_k$, $F^2_{k+1} \leftarrow F^2_k \cap H_k$, $k \leftarrow k+1$ and go to **Step 1**.

If at a given k the hyperplane H_k properly separates F^1_k and F^2_k we will say that this iteration is *reducing*. In view of Lemma 3.4(*i*), we also have the following relations.

$$\begin{cases} \theta_k = \sup\limits_{y\in F^1_k}\langle s_k,\, y\rangle = \sup\limits_{y\in F^1_k\cap F^2_k}\langle s_k,\, y\rangle = \inf\limits_{y\in F^1_k\cap F^2_k}\langle s_k,\, y\rangle = \inf\limits_{y\in F^2_k}\langle s_k,\, y\rangle, \\ s_k \in \mathrm{dom}\,\sigma_{F^1_k}, \quad s_k \in -\mathrm{dom}\,\sigma_{F^2_k}, \quad s_k \in \mathrm{dom}\,\sigma_{F^1_k\cap F^2_k} \quad \text{and} \quad F^1_k \cap F^2_k \subseteq H_k. \end{cases} \tag{5.1}$$

Next, we fix some $i \geqslant 0$ and consider the following so-called *extended system*:

$$\begin{cases} F_0^1 = C_1, F_0^2 = C_2, \\ s_k \in (F_k^2 - F_k^1)^*, \ \sup_{y \in F_k^1} \langle s_k, y \rangle = \theta_k = \inf_{y \in F_k^2} \langle s_k, y \rangle, k = 0, \ldots, i, \\ \text{where} \\ F_k^1 = C_1 \cap H_0 \cap \ldots \cap H_{k-1}, k = 1, 2, \ldots, i, \\ F_k^2 = C_2 \cap H_0 \cap \ldots \cap H_{k-1}, k = 1, 2, \ldots, i, \\ H_k = \{z \mid \langle s_k, z \rangle = \theta_k\}, k = 0, 1, \ldots, i. \end{cases} \tag{EXT$_i$}$$

We denote the feasible set of (EXT$_i$) by Feas(EXT$_i$), i.e., the set of all possible sequences $(s_k, \theta_k)_{k=0}^i$ that satisfy (EXT$_i$). This is the same as the set of all possible $(s_k, \theta_k)_{k=0}^i$ that could appear during the first $i+1$ iterations of a run of Algorithm 2. We now set some nomenclature.

- Given $(s_k, \theta_k)_{k=0}^i$ satisfying (EXT$_i$), we say that the faces $\mathcal{F} := (F_k^1, F_k^2)_{k=0}^i$ as in (EXT$_i$) are *generated by* $(s_k, \theta_k)_{k=0}^i$.
- Given a sequence of faces $\mathcal{F} := (F_k^1, F_k^2)_{k=0}^i$ where we have $F_k^j \trianglelefteq C_j$ for all $k \in \{0, \ldots, i+1\}$ and $j \in \{1, 2\}$, we say that $\mathcal{F}$ is *generated by* (EXT$_i$) if $\mathcal{F}$ is generated by some $(s_k, \theta_k)_{k=0}^i$ satisfying (EXT$_i$).

As a consequence of Theorem 3.5, we have the following lemma, where ℓ is as in (3.11) and $\mathcal{SD}$ is joint singularity degree of C_1, C_2 (see Definition 3.6).

Lemma 5.1. *The following hold about Algorithm 2.*

(i) For any $i \geqslant \mathcal{SD}$, there exists $(s_k, \theta_k)_{k=0}^i \in$ Feas(EXT$_i$) such that the generated faces $(F_k^1, F_k^2)_{k=0}^i$ satisfy $F_i^1 = F_{\min}^1$ and $F_i^2 = F_{\min}^2$.

(ii) For all $k \geqslant 0$, $F_{\min}^1 \trianglelefteq F_k^1$, $F_{\min}^2 \trianglelefteq F_k^2$, and $F_k^1 \cap F_k^2 = C_1 \cap C_2 \neq \varnothing$.

Before we proceed we introduce some notation. For every i, let $\mathfrak{F}_i$ be the set of all possible sequences of faces generated by (EXT$_i$), i.e.,

$$\mathfrak{F}_i := \left\{\mathcal{F} := (F_k^1, F_k^2)_{k=0}^i \mid \mathcal{F} \text{ is generated by (EXT}_i)\right\}. \tag{5.2}$$

For $\mathcal{F} \in \mathfrak{F}_i$, we denote the set of all possible $(s_k, \theta_k)_{k=0}^i \in$ Feas(EXT$_i$) that generates $\mathcal{F}$ by

$$\mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_i) := \left\{(s_k, \theta_k)_{k=0}^i \in \mathrm{Feas}(\mathrm{EXT}_i) \mid (s_k, \theta_k)_{k=0}^i \text{ generates } \mathcal{F}\right\}. \tag{5.3}$$

For every i we have

$$\begin{gathered} \bigcup_{\mathcal{F} \in \mathfrak{F}_i} \mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_i) = \mathrm{Feas}(\mathrm{EXT}_i); \\ \mathrm{Feas}_{\mathcal{F}_1}(\mathrm{EXT}_i) \cap \mathrm{Feas}_{\mathcal{F}_2}(\mathrm{EXT}_i) = \varnothing \text{ for any } \mathcal{F}_1, \mathcal{F}_2 \in \mathfrak{F}_i \text{ with } \mathcal{F}_1 \neq \mathcal{F}_2. \end{gathered} \tag{5.4}$$

For any $\mathcal{F} \in \mathfrak{F}_i$, we denote by $\overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_i)$ the component-wise rescaling of $\mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_i)$ by

$$\overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_i) := \{(\lambda_k s_k, \lambda_k \theta_k)_{k=0}^i \mid (s_k, \theta_k)_{k=0}^i \in \mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_i), \lambda_k \in \mathbb{R} \ \forall\, k\}. \tag{5.5}$$

Analogously, we define

$$\overline{\text{Feas}}(\text{EXT}_i) := \{(\lambda_k s_k, \lambda_k \theta_k)_{k=0}^i \mid (s_k, \theta_k)_{k=0}^i \in \text{Feas}(\text{EXT}_i), \lambda_k \in \mathbb{R}\ \forall\, k\}. \tag{5.6}$$

Then, it can be readily observed from (5.4) that

$$\overline{\text{Feas}}(\text{EXT}_i) = \bigcup_{\mathcal{F}\in\mathfrak{F}_i} \overline{\text{Feas}}_{\mathcal{F}}(\text{EXT}_i). \tag{5.7}$$

Let

$$A_i := \left\{ \sum_{k=0}^{i-1} r_k \mid (r_k, \rho_k)_{k=0}^i \in \overline{\text{Feas}}(\text{EXT}_i) \right\}. \tag{5.8}$$

Here, by convention we let $\sum_{k=0}^{-1} r_k = 0$. Using (5.7), we have

$$\text{span}\,(A_i) = \left\{ \sum_{k=0}^{i-1} r_k \mid (r_k, \rho_k)_{k=0}^i \in \text{span}\,(\text{Feas}(\text{EXT}_i)) \right\}. \tag{5.9}$$

The following lemma will aid in our construction of the extended dual.

Lemma 5.2. *For $j \in \{1, 2\}$ and any $i \geqslant \mathcal{SD}$, we have $\sigma_{F^j_{\min}} = \min\limits_{\mathcal{F}\in\mathfrak{F}_i} \sigma_{F^j_i}$ and $\sigma_{\text{aff}(F^j_{\min})} = \min\limits_{\mathcal{F}\in\mathfrak{F}_i} \sigma_{\text{aff}(F^j_i)}$.*

Proof. It suffices to show the statement holds for $i = \mathcal{SD}$. Let $\mathcal{F} \in \mathfrak{F}_{\mathcal{SD}}$ and $F^j_{\mathcal{SD}}$ in $\mathcal{F}$. By Lemma 5.1(*ii*), we know that

$$\sigma_{F^j_{\min}} \leqslant \sigma_{F^j_{\mathcal{SD}}} \text{ and } \sigma_{\text{aff}(F^j_{\min})} \leqslant \sigma_{\text{aff}(F^j_{\mathcal{SD}})}.$$

Hence, $\sigma_{F^j_{\min}} \leqslant \min\limits_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \sigma_{F^j_{\mathcal{SD}}}$ and $\sigma_{\text{aff}(F^j_{\min})} \leqslant \min\limits_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \sigma_{\text{aff}(F^j_{\mathcal{SD}})}$. Conversely, thanks to Lemma 5.1(*i*), there exists $\mathcal{F} \in \mathfrak{F}_{\mathcal{SD}}$ such that $F^j_{\mathcal{SD}} = F^j_{\min}$ for $j \in \{1, 2\}$. Then

$$\sigma_{F^j_{\min}} \geqslant \min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \sigma_{F^j_{\mathcal{SD}}} \text{ and } \sigma_{\text{aff}(F^j_{\min})} \geqslant \min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \sigma_{\text{aff}(F^j_{\mathcal{SD}})}.$$

This implies that $\sigma_{F^j_{\min}} = \min\limits_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \sigma_{F^j_{\mathcal{SD}}}$ and $\sigma_{\text{aff}(F^j_{\min})} = \min\limits_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \sigma_{\text{aff}(F^j_{\mathcal{SD}})}$ for $j \in \{1, 2\}$. □

We can now deduce expressions for $\sigma_{F^1_{\min}}$ and $\sigma_{F^2_{\min}}$ using Feas(EXT$_i$) with $i \geqslant \mathcal{SD}$.

Theorem 5.3. *Let $C_1, C_2 \subseteq \mathcal{E}$ be nonempty convex sets satisfying $C_1 \cap C_2 \neq \varnothing$ and let $F^1_{\min}$ and $F^2_{\min}$ be as in (3.2). For $j \in \{1, 2\}$ and $i \geqslant \mathcal{SD}$,*

$$\sigma_{F^j_{\min}}(x) = \text{cl} \inf_{\substack{(r_k,\rho_k)_{k=0}^i \in \\ \text{span}\,(\text{Feas}(\text{EXT}_i))}} \left\{ \sigma_{C_j}\Big(x - \sum_{k=0}^{i-1} r_k\Big) + \sum_{k=0}^{i-1} \rho_k \right\}. \tag{5.10}$$

Proof. It suffices to show the result for $i = \mathcal{SD}$. For any $\mathcal{F} \in \mathfrak{F}_{\mathcal{SD}}$, we define for $j \in \{1, 2\}$,

$$g^j_{\mathcal{F}}(x) := \operatorname{cl} \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}} \in \\ \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\},$$

where $\overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$ is defined in (5.5). We also define for $j \in \{1, 2\}$,

$$g^j(x) := \operatorname{cl} \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}} \in \\ \overline{\mathrm{Feas}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\}, \tag{5.11}$$

where $\overline{\mathrm{Feas}}(\mathrm{EXT}_{\mathcal{SD}})$ is defined in (5.6); and

$$p^j(x) := \operatorname{cl} \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}} \in \\ \mathrm{span}\,(\mathrm{Feas}(\mathrm{EXT}_{\mathcal{SD}}))}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\}.$$

Then we shall show (5.10) by showing for $j \in \{1, 2\}$,

$$\sigma_{F^j_{\min}}(x) \overset{\text{(a)}}{=} \min_{\mathcal{F} \in \mathfrak{F}_{\mathcal{SD}}} \sigma_{F^j_{\mathcal{SD}}}(x) \overset{\text{(b)}}{=} \min_{\mathcal{F} \in \mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}}(x) \overset{\text{(c)}}{=} g^j(x) \overset{\text{(d)}}{=} p^j(x),$$

where (a) follows from Lemma 5.2. In what follows, we will prove each of the equalities (b), (c) and (d).

$\boxed{\text{(b)}}$ It suffices to show that for any $\mathcal{F} \in \mathfrak{F}_{\mathcal{SD}}$, we have $\sigma_{F^j_{\mathcal{SD}}} = g^j_{\mathcal{F}}$.

Let $\mathcal{F} = (F^1_k, F^2_k)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}$ and let $(s_k, \theta_k)_{k=0}^{\mathcal{SD}} \in \mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$. For each k we define

$$H_k := \{z \mid \langle s_k, z \rangle = \theta_k\}.$$

By Lemma 2.6 we have that

$$\operatorname{dom} \sigma_{H_k} = \mathrm{span}\,\{s_k\}, \qquad \sigma_{H_k}(\lambda s_k) = \lambda \theta_k \tag{5.12}$$

hold for every $\lambda \in \mathbb{R}$.

By (EXT_i), we have $F^j_{\mathcal{SD}} = C_j \cap H_0 \cap H_1 \cap \ldots \cap H_{\mathcal{SD}-1}$. Therefore,

$$\begin{aligned} \sigma_{F^j_{\mathcal{SD}}}(x) &= \operatorname{cl} \left(\sigma_{C_j} \,\square\, \sigma_{H_0} \,\square\, \ldots \,\square\, \sigma_{H_{\mathcal{SD}-1}} \right)(x) \\ &= \operatorname{cl} \inf_{(x_k)_{k=0}^{\mathcal{SD}}} \left\{ \sigma_{C_j}(x_{\mathcal{SD}}) + \sum_{k=0}^{\mathcal{SD}-1} \sigma_{H_k}(x_k) \mid \sum_{k=0}^{\mathcal{SD}} x_k = x \right\} \\ &= \operatorname{cl} \inf_{(\lambda_k)_{k=0}^{\mathcal{SD}-1} \subseteq \mathbb{R}^{\mathcal{SD}}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} \lambda_k s_k) + \sum_{k=0}^{\mathcal{SD}-1} \lambda_k \theta_k \right\}, \end{aligned} \tag{5.13}$$

where the last equality comes from (5.12).

Next, let

$$L^j(x) := \inf_{(\lambda_k)_{k=0}^{\mathcal{SD}-1}\subseteq \mathbb{R}^{\mathcal{SD}}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} \lambda_k s_k) + \sum_{k=0}^{\mathcal{SD}-1} \lambda_k \theta_k \right\}$$
$$h^j_{\mathcal{F}}(x) := \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}}\in \\ \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\}. \tag{5.14}$$

We will show that

$$L^j(x) \geqslant h^j_{\mathcal{F}}(x) \geqslant \mathrm{cl}\, L^j(x). \tag{5.15}$$

We have $L^j(x) \geqslant h^j_{\mathcal{F}}(x)$ because $(s_k, \theta_k)_{k=0}^{\mathcal{SD}} \in \mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$ implies that $(\lambda_k s_k, \lambda_k\theta_k)_{k=0}^{\mathcal{SD}} \in \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$ for any $(\lambda_k)_{k=0}^{\mathcal{SD}-1} \subseteq \mathbb{R}^{\mathcal{SD}}$.

Let $(r_k, \rho_k)_{k=0}^{\mathcal{SD}} \in \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$. Then, $(r_k, \rho_k) = (\widehat{\lambda}_k \widehat{s}_k, \widehat{\lambda}_k \widehat{\theta}_k)$ for some $(\widehat{s}_k, \widehat{\theta}_k)_{k=0}^{\mathcal{SD}} \in \mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$ and $(\widehat{\lambda}_k)_{k=0}^{\mathcal{SD}-1} \subseteq \mathbb{R}^{\mathcal{SD}}$. Now, (5.13) is valid for any element of $\mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$, so it is valid, in particular, for $(\widehat{s}_k, \widehat{\theta}_k)_{k=0}^{\mathcal{SD}}$ as well. Therefore,

$$\begin{aligned}
\sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k &= \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} \widehat{\lambda}_k \widehat{s}_k) + \sum_{k=0}^{\mathcal{SD}-1} \widehat{\lambda}_k \widehat{\theta}_k \\
&\geqslant \inf_{(\lambda_k)_{k=0}^{\mathcal{SD}-1}\subseteq \mathbb{R}^{\mathcal{SD}}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} \lambda_k \widehat{s}_k) + \sum_{k=0}^{\mathcal{SD}-1} \lambda_k \widehat{\theta}_k \right\} \\
&\geqslant \mathrm{cl} \inf_{(\lambda_k)_{k=0}^{\mathcal{SD}-1}\subseteq \mathbb{R}^{\mathcal{SD}}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} \lambda_k \widehat{s}_k) + \sum_{k=0}^{\mathcal{SD}-1} \lambda_k \widehat{\theta}_k \right\} \\
&= \sigma_{F^j_{\mathcal{SD}}}(x) = \mathrm{cl}\, L^j(x),
\end{aligned}$$

where the second inequality and the last equality both follow from (5.13) and the fact that a convex function is lower bounded by its closure. Therefore, $h^j_{\mathcal{F}}(x) \geqslant \mathrm{cl}\, L^j(x)$ holds.

Then, taking the closure at both inequalities in (5.15) and recalling that $\sigma_{F^j_{\mathcal{SD}}}(x) = \mathrm{cl}\, L^j(x)$ allow us to conclude that

$$\sigma_{F^j_{\mathcal{SD}}}(x) = \mathrm{cl} \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}}\in \\ \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(x - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\} = g^j_{\mathcal{F}}(x).$$

$\boxed{\text{(c)}}$ We have by (a) and (b) that

$$\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \mathrm{cl}\, h^j_{\mathcal{F}} = \min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}} = \sigma_{F^j_{\min}} \tag{5.16}$$

holds. This implies that the infimums

$$\inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \mathrm{cl}\, h^j_{\mathcal{F}} \qquad \text{and} \qquad \inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}}$$

are attained as $\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \mathrm{cl}\, h^j_{\mathcal{F}}$ and $\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}}$, respectively. Moreover, since $\sigma_{F^j_{\min}}$ is a closed function, we know that $\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \mathrm{cl}\, h^j_{\mathcal{F}}$ is also a closed function, then

$$\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \mathrm{cl}\, h^j_{\mathcal{F}} = \mathrm{cl} \min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \mathrm{cl}\, h^j_{\mathcal{F}}. \tag{5.17}$$

Therefore, we claim the closure of $h^j_{\mathcal{F}}$ inside the infimum can be taken away as following:

$$\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \operatorname{cl} h^j_{\mathcal{F}} \overset{\text{(a)}}{=} \operatorname{cl} \min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \operatorname{cl} h^j_{\mathcal{F}} \overset{\text{(b)}}{=} \operatorname{cl} \inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \operatorname{cl} h^j_{\mathcal{F}} \overset{\text{(c)}}{=} \operatorname{cl} \inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} h^j_{\mathcal{F}}, \tag{5.18}$$

where (a) comes from (5.17); (b) holds since $\inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \operatorname{cl} h^j_{\mathcal{F}}$ is attained; (c) is true thanks to Lemma 2.5(*i*). Regarding $\inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} h^j_{\mathcal{F}}$, it is not clear whether the infimum is attained so we use "inf" instead of "min". This leads to

$$\min_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}} = \inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}} = \operatorname{cl} \inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}}\in \\ \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(\cdot - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\}, \tag{5.19}$$

where the first equality comes from the fact that $\inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} g^j_{\mathcal{F}}$ is attained; the second equality comes from (5.16), (5.18) and the definition of $h^j_{\mathcal{F}}$ in (5.14).

Next, by (5.19) and the definition of g^j in (5.11), to show (c), it remains to show that

$$\inf_{\mathcal{F}\in\mathfrak{F}_{\mathcal{SD}}} \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}}\in \\ \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(\cdot - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\}$$
$$= \inf_{\substack{(r_k,\rho_k)_{k=0}^{\mathcal{SD}}\in \\ \overline{\mathrm{Feas}}(\mathrm{EXT}_{\mathcal{SD}})}} \left\{ \sigma_{C_j}(\cdot - \sum_{k=0}^{\mathcal{SD}-1} r_k) + \sum_{k=0}^{\mathcal{SD}-1} \rho_k \right\}.$$

This follows directly by using Lemma 2.5(*ii*) and (5.7).

$\boxed{\text{(d)}}$ On one hand, since $F^j_{\min} = C_j \cap \operatorname{aff}(F^j_{\min})$, we have $\sigma_{F^j_{\min}} = \operatorname{cl}(\sigma_{C_j} \Box \sigma_{\operatorname{aff}(F^j_{\min})})$. Let $a \in C_1 \cap C_2$ be arbitrary. From Lemma 2.6, we have $\sigma_{\operatorname{aff} F^j_{\min}} = \langle a, \cdot\rangle + \delta_{\mathcal{L}^\perp_{F^j_{\min}}}$. Then, we have

$$\sigma_{F^j_{\min}}(x) = \operatorname{cl} \inf_{y\in\mathcal{L}^\perp_{F^j_{\min}}} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\}. \tag{5.20}$$

On the other hand, let $(r_k,\rho_k)_{k=0}^{\mathcal{SD}} \in \overline{\mathrm{Feas}}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$. Then, $(r_k,\rho_k) = (\lambda_k s_k, \lambda_k\theta_k)$ for some $(s_k,\theta_k)_{k=0}^{\mathcal{SD}} \in \mathrm{Feas}_{\mathcal{F}}(\mathrm{EXT}_{\mathcal{SD}})$ and $(\lambda_k)_{k=0}^{\mathcal{SD}-1} \subseteq \mathbb{R}^{\mathcal{SD}}$ Now, defining hyperplanes H_k's as in (5.12), we have $a \in C_1 \cap C_2 \subseteq H_k$. That is, $\langle a, s_k\rangle = \theta_k$ holds and we have

$$\sum_{k=0}^{\mathcal{SD}-1} \rho_k = \langle a, \sum_{k=0}^{\mathcal{SD}-1} r_k\rangle.$$

Then, by (a), (b), (c), the definition of g^j in (5.11) and the definition of $A_{\mathcal{SD}}$ in (5.8), we have

$$\sigma_{F^j_{\min}}(x) = \operatorname{cl} \inf_{y\in A_{\mathcal{SD}}} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\}. \tag{5.21}$$

From (5.20) and (5.21), we have

$$\operatorname{cl} \inf_{y\in\mathcal{L}^\perp_{F^j_{\min}}} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\} = \operatorname{cl} \inf_{y\in A_{\mathcal{SD}}} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\}. \tag{5.22}$$

For any $(s_k, \theta_k)_{k=0}^{\mathcal{SD}} \in \mathrm{Feas}(\mathrm{EXT}_{\mathcal{SD}})$, we know that $F_{\min}^j \subseteq H_k := \{z \mid \langle s_k, z\rangle = \theta_k\}$ for all k. Then it holds by the definition of $\mathcal{L}_{F_{\min}^j}$ that $\mathcal{L}_{F_{\min}^j} \subseteq \{s_k\}^\perp$, which follows that $\mathcal{L}_{F_{\min}^j}^\perp \supseteq \mathrm{span}\,\{s_k\}$, i.e., for any $\lambda_k \in \mathbb{R}$,

$$\lambda_k s_k \in \mathcal{L}_{F_{\min}^j}^\perp.$$

Since $\mathcal{L}_{F_{\min}^j}^\perp$ is a subspace, for any $(\lambda_k)_{k=0}^{\mathcal{SD}-1} \subseteq \mathbb{R}^{\mathcal{SD}}$, $\sum_{k=0}^{\mathcal{SD}-1} \lambda_k s_k \in \mathcal{L}_{F_{\min}^j}^\perp$. Therefore, we have $A_{\mathcal{SD}} \subseteq \mathcal{L}_{F_{\min}^j}^\perp$ and hence $A_{\mathcal{SD}} \subseteq \mathrm{span}\,(A_{\mathcal{SD}}) \subseteq \mathcal{L}_{F_{\min}^j}^\perp$. Therefore, we have

$$\begin{aligned}
\mathrm{cl} \inf_{y \in \mathcal{L}_{F_{\min}^j}^\perp} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\} &\leqslant \mathrm{cl} \inf_{y \in \mathrm{span}\,(A_{\mathcal{SD}})} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\} \\
&\leqslant \mathrm{cl} \inf_{y \in A_{\mathcal{SD}}} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\}.
\end{aligned}$$

By (5.22), we know all inequalities in the above display becomes equalities, which implies

$$g^j(x) = \mathrm{cl} \inf_{y \in \mathrm{span}\,(A_{\mathcal{SD}})} \left\{\sigma_{C_j}(x-y) + \langle a, y\rangle\right\}.$$

This and (5.9) show that (d) holds. $\square$

Now, we return to the setting of (1.1). Let $C_1 = \mathrm{dom}\, f$, $C_2 = \mathrm{dom}\, g$ and we assume that (1.1) is feasible so that $C_1 \cap C_2 \neq \varnothing$. Let φ_j denote the right-hand-side of (5.10) for $j \in \{1, 2\}$ and $i = \mathcal{SD}$. Using Theorem 5.3, we then obtain the following extended Fenchel dual.

Theorem 5.4 (Extended Fenchel dual)**.** *Consider the following problem:*

$$\max_{\lambda} -\mathrm{cl}(f^* \,\square\, \varphi_1)(-\lambda) - \mathrm{cl}(g^* \,\square\, \varphi_2)(\lambda). \tag{5.23}$$

If (1.1) *is feasible, then* (1.1) *and* (5.23) *satisfy strong duality.*

Proof. By Theorem 5.3, $\sigma_{F_{\min}^j} = \varphi_j$ holds for $j \in \{1, 2\}$. Then, the result follows from Theorem 3.1. $\square$

Through Theorem 5.3 we obtain a dual for (5.4) that always affords strong duality in the sense of Definition 2.9. Furthermore, it does not require explicit expressions for $F_{\min}^1$ an $F_{\min}^2$. However, saying that (5.23) is *somewhat involved* is a massive understatement.

Given that (1.1) is fairly general, unfortunately, there seems to be little room for obtaining more concrete expressions. As such, at this level of generality, Theorem 5.3 may be one of the closest analogues to results such as Theorem 2 of [19], which similarly provides an extended dual for a general conic linear program by encoding the facial reduction process through the so-called *facial reduction cone*.

In the next section we will see that we can further simplify (5.23) under additional assumptions.

6 Extended Fenchel duality with niceness

In this section, our goal is to further simplify the expression in Theorem 5.3 under the assumption that C_1 and C_2 are *closed* and *nice*. We start with the following lemma.

Lemma 6.1. *Suppose $C_1, C_2 \subseteq \mathcal{E}$ are nonempty convex sets and $F_1 \trianglelefteq C_1$ and $F_2 \trianglelefteq C_2$ are nice faces. Then, $s \in (F_2 - F_1)^*$ if and only if there exist $u_1 \in \operatorname{dom} \sigma_{C_1}$, $v_1 \in \operatorname{dom} \sigma_{\operatorname{aff}(F_1)}$, $u_2 \in \operatorname{dom} \sigma_{C_2}$ and $v_2 \in \operatorname{dom} \sigma_{\operatorname{aff}(F_2)}$ such that*

$$\begin{cases} \sigma_{F_1}(s) = \sigma_{C_1}(u_1) + \sigma_{\operatorname{aff}(F_1)}(v_1), \quad \sigma_{F_2}(-s) = \sigma_{C_2}(u_2) + \sigma_{\operatorname{aff}(F_2)}(v_2), \\ s = u_1 + v_1 = -u_2 - v_2, \\ \sigma_{C_1}(u_1) + \sigma_{C_2}(u_2) + \sigma_{\operatorname{aff}(F_1)}(v_1) + \sigma_{\operatorname{aff}(F_2)}(v_2) \leqslant 0. \end{cases} \tag{6.1}$$

If further $F_1 \cap F_2 \neq \varnothing$, then the equivalence holds with the last inequality of (6.1) being an equality.

Proof. We see from the first part in Lemma 2.4 that

$$s \in (F_2 - F_1)^* \iff \sigma_{F_1}(s) + \sigma_{F_2}(-s) \leqslant 0. \tag{6.2}$$

Suppose that $s \in (F_2 - F_1)^*$. The niceness of C_1 and C_2 leads to

$$\sigma_{F_1}(s) = \sigma_{C_1} \boxdot \sigma_{\operatorname{aff}(F_1)}(s), \quad \sigma_{F_2}(-s) = \sigma_{C_2} \boxdot \sigma_{\operatorname{aff}(F_2)}(-s).$$

As the infimal convolutions are exact, there exist $u_1 \in \operatorname{dom} \sigma_{C_1}$, $v_1 \in \operatorname{dom} \sigma_{\operatorname{aff}(F_1)}$, $u_2 \in \operatorname{dom} \sigma_{C_2}$ and $v_2 \in \operatorname{dom} \sigma_{\operatorname{aff}(F_2)}$ such that

$$\begin{aligned} s = u_1 + v_1, &\quad \sigma_{F_1}(s) = \sigma_{C_1}(u_1) + \sigma_{\operatorname{aff}(F_1)}(v_1), \\ -s = u_2 + v_2, &\quad \sigma_{F_2}(-s) = \sigma_{C_2}(u_2) + \sigma_{\operatorname{aff}(F_2)}(v_2). \end{aligned} \tag{6.3}$$

Therefore, (6.1) is a consequence of (6.2) and (6.3).

Conversely, if (6.1) holds, we obtain from the first and the last equations in (6.1) that

$$\sigma_{F_1}(s) + \sigma_{F_2}(-s) \leqslant 0.$$

This together with (6.2) proves that $s \in (F_2 - F_1)^*$.

If $F_1 \cap F_2 \neq \varnothing$ holds, then the second part in Lemma 2.4 implies that that inequality in (6.2) holds as an equality. This implies that the inequality in (6.1) holds as an equality as well. ☐

In what follows, we recall that $\operatorname{epi} \sigma_C$ is a closed convex cone and the tangent space at a given point is defined as in (2.4).

Lemma 6.2. *Let $C \subseteq \mathcal{E}$ be a closed convex set, $s \in \operatorname{dom} \sigma_C$, and $H \coloneqq \{x \mid \langle x, s \rangle = \sigma_C(s)\}$. A face $F \trianglelefteq C$ is exposed by H if and only if*

$$\operatorname{gra} \sigma_{\operatorname{aff}(F)} = \tan\left((s, \sigma_C(s)), \operatorname{epi} \sigma_C\right).$$

Proof. ($\Longrightarrow$) From Proposition 4.2(*i*), $\overline{\text{cone}}(F \times \{-1\})$ is exposed by $\{(s, \sigma_C(s))\}^\perp$, i.e.,

$$\overline{\text{cone}}(F \times \{-1\}) = \overline{\text{cone}}(C \times \{-1\}) \cap \{(s, \sigma_C(s))\}^\perp .$$

Using (2.6), we have

$$\overline{\text{cone}}(F \times \{-1\}) = \text{minFace}\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right)^\triangle .$$

Taking the orthogonal complement at both sides, we have

$$\begin{aligned}\overline{\text{cone}}(F \times \{-1\})^\perp &= \text{minFace}\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right)^{\triangle\perp} \\ &= \tan\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right),\end{aligned} \tag{6.4}$$

where the last equality comes from (2.5). Then,

$$\begin{aligned}\text{gra}\,\sigma_{\text{aff}(F)} &\overset{\text{(a)}}{=} \overline{\text{cone}}(F \times \{-1\})^\perp \overset{\text{(b)}}{=} \tan\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right) \\ &= \tan\left(\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^\circ\right) \overset{\text{(c)}}{=} \tan\left(\left(s, \sigma_C(s)\right), \text{epi}\,\sigma_C\right),\end{aligned}$$

where (a) comes from (4.3); (b) holds thanks to (6.4); (c) is true by (4.2).

($\Longleftarrow$) We have

$$\begin{aligned}\overline{\text{cone}}(F \times \{-1\})^\perp &\overset{\text{(a)}}{=} \text{gra}\,\sigma_{\text{aff}(F)} \overset{\text{(b)}}{=} \tan\left(\left(s, \sigma_C(s)\right), \text{epi}\,\sigma_C\right) \\ &\overset{\text{(c)}}{=} \tan\left(\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^\circ\right) = \tan\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right),\end{aligned}$$

where (a) comes from (4.3); (b) holds as assumed; (c) is true by (4.2).

Therefore, it holds that

$$\begin{aligned}\overline{\text{cone}}(F \times \{-1\})^\perp &= \tan\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right) \\ &= \text{minFace}\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right)^{\triangle\perp},\end{aligned}$$

where the last equality comes from (2.5). Taking the orthogonal complement at both sides, we have

$$\text{span}\left(\overline{\text{cone}}(F \times \{-1\})\right) = \text{span}\left(\text{minFace}\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right)^\triangle\right).$$

That is, the face $\overline{\text{cone}}(F \times \{-1\})$ and the face exposed by $-(s, \sigma_C(s))$ share the same linear span. Therefore, they must coincide by (2.2), i.e.,

$$\overline{\text{cone}}(F \times \{-1\}) = \text{minFace}\left(-\left(s, \sigma_C(s)\right), \overline{\text{cone}}(C \times \{-1\})^*\right)^\triangle .$$

Then, Proposition 4.2(*i*) implies that F is exposed by H. □

We now re-examine the extended system in (EXT_i) and explain the intuition behind what we will do next.

- By Lemma 6.1, each s_k can be expressed in two different ways: as the sum of elements of $\text{dom}\,\sigma_{C_1}$ and $\text{dom}\,\sigma_{\text{aff}(F_k^1)}$; or the sum of the negative of elements of $\text{dom}\,\sigma_{C_2}$ and $\text{dom}\,\sigma_{\text{aff}(F_k^2)}$.

- The pair (s_k, θ_k) defines the hyperplane $H_k := \{z \mid \langle z,\, s_k\rangle = \theta_k\}$. Now, H_k exposes F^1_{k+1} as a face of F^1_k but it may fail to expose F^1_{k+1} as a face of C_1 because H_k may not be a supporting hyperplane of C_1.

  However, s_k does have a component u^1_k that belongs to $\operatorname{dom}\sigma_{C_1}$ (by Lemma 6.1). It turns out that if we sum all the u_j's until $j = k$, we can construct a hyperplane that directly exposes F_{k+1} as face of C_1, see (6.5) below. The key for that is ensuring that the tangent space condition in Lemma 6.2 is met. A similar discussion holds for the faces of C_2. We note that the importance of a tangent space constraint has been previously identified by Pataki in [30] in the context of conic linear programs.

  By exposing the faces F^1_k, F^2_k directly as faces of C_1 and C_2 and using the tangent space condition in Lemma 6.2, we no longer need to keep track of expressions such as $(F^2_k - F^1_k)^*$ in (EXT_i). In terms of computation, this suggests that, under niceness and closedness, the only requirement for obtaining an extended dual is knowing how to compute the tangent spaces at points of the epigraphs of the support functions of C_1 and C_2. This is a relatively more doable task than computing the objects involved in (5.10).

Based on these ideas, we consider the following system.

$$\begin{cases} (u^1_0, v^1_0, \zeta^1_0) \in \operatorname{dom}\sigma_{C_1} \times \operatorname{gra}\sigma_{\operatorname{aff}(C_1)}, \quad (u^2_0, v^2_0, \zeta^2_0) \in \operatorname{dom}\sigma_{C_2} \times \operatorname{gra}\sigma_{\operatorname{aff}(C_2)}, \\ \left(u^j_k, v^j_k, \zeta^j_k\right) \in \operatorname{dom}\sigma_{C_j} \times \tan\left(\left(\sum_{l=0}^{k-1} u^j_l, \sigma_{C_j}\big(\sum_{l=0}^{k-1} u^j_l\big)\right), \operatorname{epi}\sigma_{C_j}\right), j \in \{1,2\}, k = 1, \dots, i, \\ u^1_k + v^1_k = -u^2_k - v^2_k, \qquad k = 0, 1, \dots, i, \\ \sigma_{C_1}(u^1_k) + \sigma_{C_2}(u^2_k) + \zeta^1_k + \zeta^2_k \leqslant 0, \qquad k = 0, 1, \dots, i. \end{cases} \tag{EXT_i-nice}$$

Next, we set the notation. The feasible set of (EXT_i-nice) is denoted by Feas(EXT_i-nice), i.e., it is the set of all possible sequences

$$(u^1_k, v^1_k, \zeta^1_k, u^2_k, v^2_k, \zeta^2_k)^i_{k=0}$$

that are feasible for (EXT_i-nice).

For any $(u^1_k, v^1_k, \zeta^1_k, u^2_k, v^2_k, \zeta^2_k)^i_{k=0} \in$ Feas(EXT_i-nice), for $k \in \{1, 2, \dots, i\}$, we let

$$\widetilde{H}^j_{k-1} := \left\{ x \mid \langle x,\, \sum_{l=0}^{k-1} u^j_l\rangle = \sigma_{C_j}\big(\sum_{l=0}^{k-1} u^j_l\big) \right\} \tag{6.5}$$

and let G^j_k be the face of C_j exposed by $\widetilde{H}^j_{k-1}$, i.e.,

$$G^j_k := C_j \cap \widetilde{H}^j_{k-1}. \tag{6.6}$$

We also let $G^j_0 := C_j$. With these G^j_k's, we define the pairs of faces $\mathcal{G} := (G^1_k, G^2_k)^i_{k=0}$ by the pairs of faces generated by $(u^1_k, v^1_k, \zeta^1_k, u^2_k, v^2_k, \zeta^2_k)^i_{k=0} \in$ Feas(EXT_i-nice). When the context is clear, we also say $\mathcal{G}$ *is generated by* (EXT_i-nice).

Similar to what we have done in Section 5, for $\mathcal{G}$ we use $\mathrm{Feas}_{\mathcal{G}}$(EXT$_i$-nice) to represent the set of all possible $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i$ that generate $\mathcal{G}$. In general, it could happen that $G_k^j = \varnothing$ for some k and j, however, as we shall show in the Proposition 6.4, this does not happen here thanks to the assumption that $C_1 \cap C_2 \neq \varnothing$. Specifically, in Proposition 6.4, we will show that every $\mathcal{G}$ generated by (EXT$_i$-nice) is also generated by (EXT$_i$), and vice-versa. First, we consider the following lemma.

Lemma 6.3. *Let $C \subseteq \mathcal{E}$ be a convex set and let $G \trianglelefteq C$. Suppose that $s, u, v \in \mathcal{E}$ and $H \subseteq \mathcal{E}$ are as follows.*

- $s \in \operatorname{dom} \sigma_G$.
- $H := \{z \mid \langle z, s\rangle = \sigma_G(s)\}$ *and* $G \cap H \neq \varnothing$.
- $\sigma_G(s) = \sigma_C(u) + \sigma_{\mathrm{aff}(G)}(v)$ *and* $s = u + v$.

Then,

$$G \cap H = G \cap \{z \mid \langle u, z\rangle = \sigma_C(u)\}.$$

Proof. Since $v \in \operatorname{dom} \sigma_{\mathrm{aff}(G)}$ holds, Lemma 2.6 implies that for every $x \in G$ we have

$$\sigma_{\mathrm{aff}(G)}(v) = \langle v, x\rangle. \tag{6.7}$$

Now let $x \in G \cap H$. Then by definition of H,

$$\sigma_G(s) = \langle s, x\rangle = \langle u + v, x\rangle = \langle u, x\rangle + \sigma_{\mathrm{aff}(G)}(v).$$

From $\sigma_G(s) = \sigma_C(u) + \sigma_{\mathrm{aff}(G)}(v)$ and (6.7) we conclude that

$$\langle u, x\rangle = \sigma_C(u), \qquad \forall x \in G \cap H.$$

This shows the "$\subseteq$" inclusion. Conversely, suppose that $x \in G$ satisfies $\langle u, x\rangle = \sigma_C(u)$. Then, (6.7) (which holds for every element in G) implies that $\langle v, x\rangle = \sigma_{\mathrm{aff}(G)}(v)$ so that

$$\langle s, x\rangle = \langle u + v, x\rangle = \sigma_C(u) + \sigma_{\mathrm{aff}(G)}(v) = \sigma_G(s).$$

This shows that the "$\supseteq$" inclusion holds. ☐

Proposition 6.4 (Equivalent faces)**.** *Let $C_1, C_2 \subseteq \mathcal{E}$ be two nonempty closed convex nice sets with $C_1 \cap C_2 \neq \varnothing$.*

(i) Let $\mathcal{G} := (G_k^1, G_k^2)_{k=0}^i$ be generated by $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i \in \mathrm{Feas}$(EXT$_i$-nice). For $k \in \{0, 1, \ldots, i\}$, let

$$s_k := u_k^1 + v_k^1, \quad \theta_k := \sigma_{C_1}(u_k^1) + \zeta_k^1. \tag{6.8}$$

Then, $(s_k, \theta_k)_{k=0}^i \in \mathrm{Feas}$(EXT$_i$) holds and the generated faces $\mathcal{F} := (F_k^1, F_k^2)_{k=0}^i$ satisfy $\mathcal{F} = \mathcal{G}$.

(ii) *Let* $\mathcal{F} := (F_k^1, F_k^2)_{k=0}^i$ *be generated by* $(s_k, \theta_k)_{k=0}^i \in \mathrm{Feas}(\mathrm{EXT}_i)$. *For* $k \in \{0, 1, \ldots, i\}$, *let* $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)$ *be such that* $u_k^1 \in \operatorname{dom} \sigma_{C_1}$, $v_k^1 \in \operatorname{dom} \sigma_{\mathrm{aff}(F_k^1)}$, $u_k^2 \in \operatorname{dom} \sigma_{C_2}$, $v_k^2 \in \operatorname{dom} \sigma_{\mathrm{aff}(F_k^2)}$ *and satisfy*[4]

$$\begin{cases} \sigma_{F_k^1}(s_k) = \sigma_{C_1}(u_k^1) + \sigma_{\mathrm{aff}(F_k^1)}(v_k^1), \quad \sigma_{F_k^2}(-s_k) = \sigma_{C_2}(u_k^2) + \sigma_{\mathrm{aff}(F_k^2)}(v_k^2), \\ s_k = u_k^1 + v_k^1 = -u_k^2 - v_k^2, \\ \zeta_k^1 := \sigma_{\mathrm{aff}(F_k^1)}(v_k^1), \quad \zeta_k^2 := \sigma_{\mathrm{aff}(F_k^2)}(v_k^2) \\ \sigma_{C_1}(u_k^1) + \sigma_{C_2}(u_k^2) + \zeta_k^1 + \zeta_k^2 \leqslant 0. \end{cases} \tag{6.9}$$

Then, $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i \in \mathrm{Feas}(\mathrm{EXT}_i\text{-nice})$ *holds and the generated faces* $\mathcal{G} := (G_k^1, G_k^2)_{k=0}^i$ *satisfy* $\mathcal{G} = \mathcal{F}$.

(iii) *Let* $\mathcal{G} := (G_k^1, G_k^2)_{k=0}^i$ *be generated by* (EXT_i-nice). *Then, for every* $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i \in \mathrm{Feas}_{\mathcal{G}}(\mathrm{EXT}_i\text{-nice})$ *we have*

$$\operatorname{gra} \sigma_{\mathrm{aff}(G_k^j)} = \tan\left(\left(\sum_{l=0}^{k-1} u_l^j, \sigma_{C_j}\left(\sum_{l=0}^{k-1} u_l^j\right)\right), \operatorname{epi} \sigma_{C_j}\right), \tag{6.10}$$

for $j \in \{1, 2\}$ *and* $k \in \{1, \ldots, i\}$.

Proof. ***Proof of (i).*** We proceed by induction on i. First, let $i = 0$. Then, we have

$$(u_0^1, v_0^1, \zeta_0^1, u_0^2, v_0^2, \zeta_0^2) \in \mathrm{Feas}(\mathrm{EXT}_0\text{-nice})$$

and we need to show the following.

(a) Defining s_0, θ_0 as in (6.8), we have $(s_0, \theta_0) \in \mathrm{Feas}(\mathrm{EXT}_0)$.

(b) The faces $G_0^1 \trianglelefteq C_1, G_0^2 \trianglelefteq C_2$ generated by $(u_0^1, v_0^1, \zeta_0^1, u_0^2, v_0^2, \zeta_0^2)$ (see (6.6)) coincide with the faces $F_0^1 \trianglelefteq C_1, F_0^2 \trianglelefteq C_2$ generated by (s_0, θ_0) (see (EXT_i)).

We start by showing (a). From (EXT_i-nice), we have $\zeta_0^j = \sigma_{\mathrm{aff}(C_j)}(v_0^j)$, $j \in \{1, 2\}$. This together with the last inequality in (EXT_i-nice) gives

$$\sigma_{C_1}(u_0^1) + \sigma_{\mathrm{aff}(C_1)}(v_0^1) + \sigma_{C_2}(u_0^2) + \sigma_{\mathrm{aff}(C_2)}(v_0^2) \leqslant 0.$$

The niceness of C_1 and C_2 implies that

$$\begin{aligned} \sigma_{C_1}(s_0) = \sigma_{C_1}(u_0^1 + v_0^1) = \sigma_{C_1} \boxdot \sigma_{\mathrm{aff}(C_1)}(u_0^1 + v_0^1) \leqslant \sigma_{C_1}(u_0^1) + \sigma_{\mathrm{aff}(C_1)}(v_0^1), \\ \sigma_{C_2}(-s_0) = \sigma_{C_2}(u_0^2 + v_0^2) = \sigma_{C_2} \boxdot \sigma_{\mathrm{aff}(C_2)}(u_0^2 + v_0^2) \leqslant \sigma_{C_2}(u_0^2) + \sigma_{\mathrm{aff}(C_2)}(v_0^2). \end{aligned} \tag{6.11}$$

Summing both sides of the equations above, we obtain

$$\sigma_{C_1}(s_0) + \sigma_{C_2}(-s_0) \leqslant \sigma_{C_1}(u_0^1) + \sigma_{\mathrm{aff}(C_1)}(v_0^1) + \sigma_{C_2}(u_0^2) + \sigma_{\mathrm{aff}(C_2)}(v_0^2) \leqslant 0. \tag{6.12}$$

[4]Note that $s_k \in (F_k^2 - F_k^1)^*$, $F_k^1 \trianglelefteq C_1$ and $F_k^2 \trianglelefteq C_2$ are nice faces since C_1, C_2 are nice. According to Lemma 6.1, we see that such $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)$ always exists.

Applying the first part in Lemma 2.4 to (6.12), we have $s_0 \in (C_2 - C_1)^*$. Furthermore, recall that $C_1 \cap C_2 \neq \varnothing$. We then see from $s_0 \in (C_2 - C_1)^*$ and the second part in Lemma 2.4 that $\sigma_{C_1}(s_0) + \sigma_{C_2}(-s_0) = 0$. This together with (6.12) further implies that the inequalities in (6.11) hold as equalities, i.e.,

$$\sigma_{C_1}(s_0) = \sigma_{C_1}(u_0^1) + \sigma_{\mathrm{aff}(C_1)}(v_0^1), \quad \sigma_{C_2}(-s_0) = \sigma_{C_2}(u_0^2) + \sigma_{\mathrm{aff}(C_2)}(v_0^2).$$

Combining this with $\theta_0 = \sigma_{C_1}(u_0^1) + \zeta_0^1$ and $\zeta_0^j = \sigma_{\mathrm{aff}(C_j)}(v_0^j)$, $j \in \{1, 2\}$, we obtain

$$\sup_{x \in C_1} \langle s_0, x \rangle = \sigma_{C_1}(s_0) = \theta_0, \quad \inf_{x \in C_2} \langle s_0, x \rangle = -\sigma_{C_2}(-s_0) = \sigma_{C_1}(s_0) = \theta_0.$$

This together with $s_0 \in (C_2 - C_1)^*$ shows that $(s_0, \theta_0) \in \mathrm{Feas}(\mathrm{EXT}_0)$. That is, item (a) holds. Furthermore, as $F_0^j = C_j = G_0^j$ holds, we also obtain (b).

Next, suppose that the conclusion holds for some $i \geqslant 0$. We will prove that the conclusion holds for $i + 1$, namely, if $\mathcal{G} = (G_k^1, G_k^2)_{k=0}^{i+1}$ is generated by $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^{i+1} \in \mathrm{Feas}(\mathrm{EXT}_{i+1}\text{-nice})$ and $(s_k, \theta_k)_{k=0}^{i+1}$ are defined by

$$s_k := u_k^1 + v_k^1, \quad \theta_k := \sigma_{C_1}(u_k^1) + \zeta_k^1, \quad k \in \{0, 1, \ldots, i+1\},$$

then we must have $(s_k, \theta_k)_{k=0}^{i+1} \in \mathrm{Feas}(\mathrm{EXT}_{i+1})$ and $\mathcal{G} = \mathcal{F}$ with $\mathcal{F} = (F_k^1, F_k^2)_{k=0}^{i+1}$ being generated by $(s_k, \theta_k)_{k=0}^{i+1}$.

By assumption, $\mathcal{G} = (G_k^1, G_k^2)_{k=0}^{i}$ is generated by $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^{i} \in \mathrm{Feas}(\mathrm{EXT}_i\text{-nice})$. By the inductive hypothesis, we have $(s_k, \theta_k)_{k=0}^{i} \in \mathrm{Feas}(\mathrm{EXT}_i)$ and

$$\begin{aligned} (F_i^1, F_i^2) &= (G_i^1, G_i^2), \\ s_i &\in (F_i^2 - F_i^1)^*, \\ \sup_{x \in F_i^1} \langle s_i, x \rangle = \theta_i &= \inf_{x \in F_i^2} \langle s_i, x \rangle. \end{aligned} \tag{6.13}$$

To complete the proof, it then remains to show that

- $s_{i+1} \in (F_{i+1}^2 - F_{i+1}^1)^*$, $\sup_{x \in F_{i+1}^1} \langle s_{i+1}, x \rangle = \theta_{i+1} = \inf_{x \in F_{i+1}^2} \langle s_{i+1}, x \rangle$, and
- $(F_{i+1}^1, F_{i+1}^2) = (G_{i+1}^1, G_{i+1}^2)$.

We first show $F_{i+1}^j = G_{i+1}^j$. Before that, we recall the hyperplanes defined by the (s_k, θ_k) and the $\{u_l^j\}$ as in (EXT_i) and (6.5), respectively. By definition, for $k \in \{0, 1, \ldots, i+1\}$,

$$H_k := \{z \mid \langle s_k, z \rangle = \theta_k\}, \quad \widetilde{H}_k^j := \{z \mid \langle \sum_{l=0}^{k} u_l^j, z \rangle = \sigma_{C_j}(\sum_{l=0}^{k} u_l^j)\}.$$

Next, we will prove the following equality

$$G_i^j \cap H_i = G_i^j \cap \{x \mid \langle u_i^j, x \rangle = \sigma_{C_j}(u_i^j)\}. \tag{6.14}$$

First, by (EXT$_i$-nice) we have for $j \in \{1,2\}$

$$(v_i^j, \zeta_i^j) \in \tan\left(\Big(\sum_{l=0}^{i-1} u_l^j, \sigma_{C_j}\big(\sum_{l=0}^{i-1} u_l^j\big)\Big), \operatorname{epi}\sigma_{C_j}\right).$$

In view of Lemma 6.2 and the fact that $\widetilde{H}_{i-1}^j$ exposes G_i^j as face of C_j (see (6.6)), we obtain for $j \in \{1,2\}$,

$$(v_i^j, \zeta_i^j) \in \operatorname{gra}\sigma_{\operatorname{aff}(G_i^j)}.$$

Furthermore, since $(G_i^1, G_i^2) = (F_i^1, F_i^2)$ by (6.13), we have for $j \in \{1,2\}$

$$\sigma_{\operatorname{aff}(F_i^j)}(v_i^j) = \zeta_i^j. \tag{6.15}$$

We will now prove (6.14) for $j = 1$. By (6.13), $\theta_i = \sigma_{F_i^1}(s_i)$ holds and, by (6.8) and (6.15), we have

$$\sigma_{F_i^1}(s_i) = \sigma_{C_1}(u_i^1) + \zeta_i^1 = \sigma_{C_1}(u_i^1) + \sigma_{\operatorname{aff}(F_i^1)}(v_i^1).$$

As $G_i^1 = F_i^1$ holds by (6.13), we can invoke Lemma 6.3 to conclude that (6.14) holds for $j = 1$. For $j = 2$, we have $-\theta_i = \sigma_{F_i^2}(-s_i)$ by (6.13). Since $-u_i^2 - v_i^2 = u_i^1 + v_i^1$ (see (EXT$_i$-nice)), we also have $-s_i = u_i^2 + v_i^2$ by (6.8). Then, the niceness of F_i^2 leads to

$$\sigma_{F_i^2}(-s_i) = \sigma_{F_i^2}(u_i^2 + v_i^2) = \sigma_{C_2} \boxdot \sigma_{\operatorname{aff}(F_i^2)}(u_i^2 + v_i^2) \leqslant \sigma_{C_2}(u_i^2) + \sigma_{\operatorname{aff}(F_i^2)}(v_i^2). \tag{6.16}$$

Since $\sigma_{F_i^1}(s_i) = \sigma_{C_1}(u_i^1) + \sigma_{\operatorname{aff}(F_i^1)}(v_i^1)$ and $\sigma_{F_i^1}(s_i) + \sigma_{F_i^2}(-s_i) = 0$ both hold, we have

$$0 = \sigma_{F_i^1}(s_i) + \sigma_{F_i^2}(-s_i) \leqslant \sigma_{C_1}(u_i^1) + \sigma_{\operatorname{aff}(F_i^1)}(v_i^1) + \sigma_{C_2}(u_i^2) + \sigma_{\operatorname{aff}(F_i^2)}(v_i^2) \leqslant 0,$$

where the last inequality follows from (EXT$_i$-nice). This implies that (6.16) must be an equality as well and we have $\sigma_{F_i^2}(-s_2) = \sigma_{C_2}(u_i^2) + \sigma_{\operatorname{aff}(F_i^2)}(v_i^2)$. Then, finally, we apply Lemma 6.3 to $H_i = \{z \mid \langle -s_i, z\rangle = -\theta_i\}$, $F_i^2 = G_i^2$, $-s_i$, u_i^2 and v_i^2 to conclude that (6.14) holds when $j = 2$.

We then see from (EXT$_i$) and (6.13) that for $j \in \{1,2\}$

$$\begin{aligned}
F_{i+1}^j &= F_i^j \cap H_i = G_i^j \cap H_i \overset{\text{(a)}}{=} G_i^j \cap \{z \mid \langle u_i^j, z\rangle = \sigma_{C_j}(u_i^j)\} \\
&\overset{\text{(b)}}{=} C_j \cap \widetilde{H}_{i-1}^j \cap \{z \mid \langle u_i^j, z\rangle = \sigma_{C_j}(u_i^j)\} \\
&\overset{\text{(c)}}{=} C_j \cap \left\{ z \mid \langle \sum_{l=0}^{i} u_l^j, z\rangle = \sigma_{C_j}\big(\sum_{l=0}^{i} u_l^j\big)\right\} = C_j \cap \widetilde{H}_i^j = G_{i+1}^j,
\end{aligned} \tag{6.17}$$

where (a) follows from (6.14), (b) comes from the definition of G_i^j in (6.6) and (c) follows by applying Lemma 2.8.

Next, we prove $\sup_{x \in F_{i+1}^1} \langle s_{i+1}, x\rangle = \theta_{i+1} = \inf_{x \in F_{i+1}^2} \langle s_{i+1}, x\rangle$. By the definition of (EXT$_{i+1}$-nice), we have

$$u_{i+1}^j \in \operatorname{dom}\sigma_{C_j}, \quad \text{and} \quad (v_{i+1}^j, \zeta_{i+1}^j) \in \tan\left(\Big(\sum_{l=0}^{i} u_l^j, \sigma_{C_j}\big(\sum_{l=0}^{i} u_l^j\big)\Big), \operatorname{epi}\sigma_{C_j}\right).$$

In view of Lemma 6.2 and the fact that $\widetilde{H}_i^j$ exposes G_{i+1}^j as face of C_j (see (6.6)), and noting that $\sum_{l=0}^i u_l^j \in \operatorname{dom} \sigma_{C_j}$, we obtain for $j \in \{1, 2\}$,

$$(v_{i+1}^j, \zeta_{i+1}^j) \in \operatorname{gra} \sigma_{\operatorname{aff}(G_{i+1}^j)}. \tag{6.18}$$

This together with the last inequality in (EXT$_i$-nice) gives

$$\sigma_{C_1}(u_{i+1}^1) + \sigma_{\operatorname{aff}(G_{i+1}^1)}(v_{i+1}^1) + \sigma_{C_2}(u_{i+1}^2) + \sigma_{\operatorname{aff}(G_{i+1}^2)}(v_{i+1}^2) \leqslant 0.$$

The niceness of C_1 and C_2 implies that

$$\sigma_{G_{i+1}^1}(s_{i+1}) = \sigma_{G_{i+1}^1}(u_{i+1}^1 + v_{i+1}^1) = \sigma_{C_1} \boxdot \sigma_{\operatorname{aff}(G_{i+1}^1)}(u_{i+1}^1 + v_{i+1}^1) \leqslant \sigma_{C_1}(u_{i+1}^1) + \sigma_{\operatorname{aff}(G_{i+1}^1)}(v_{i+1}^1),$$
$$\sigma_{G_{i+1}^2}(-s_{i+1}) = \sigma_{G_{i+1}^2}(u_{i+1}^2 + v_{i+1}^2) = \sigma_{C_2} \boxdot \sigma_{\operatorname{aff}(G_{i+1}^2)}(u_{i+1}^2 + v_{i+1}^2) \leqslant \sigma_{C_2}(u_{i+1}^2) + \sigma_{\operatorname{aff}(G_{i+1}^2)}(v_{i+1}^2).$$

Summing both sides of the above equations, we further obtain

$$\sigma_{G_{i+1}^1}(s_{i+1}) + \sigma_{G_{i+1}^2}(-s_{i+1}) \leqslant \sigma_{C_1}(u_{i+1}^1) + \sigma_{\operatorname{aff}(G_{i+1}^1)}(v_{i+1}^1) + \sigma_{C_2}(u_{i+1}^2) + \sigma_{\operatorname{aff}(G_{i+1}^2)}(v_{i+1}^2) \leqslant 0. \tag{6.19}$$

Applying the first part in Lemma 2.4 to (6.19), we have $s_{i+1} \in (G_{i+1}^2 - G_{i+1}^1)^* = (F_{i+1}^2 - F_{i+1}^1)^*$.

Furthermore, recall from Lemma 5.1 that $G_{i+1}^1 \cap G_{i+1}^2 = F_{i+1} \cap F_{i+1}^2 = C_1 \cap C_2 \neq \varnothing$. We then see from $s_{i+1} \in (F_{i+1}^2 - F_{i+1}^1)^*$ and the second part in Lemma 2.4 that $\sigma_{G_{i+1}^1}(s_{i+1}) + \sigma_{G_{i+1}^2}(-s_{i+1}) = 0$. This together with (6.19) implies that

$$\sigma_{G_{i+1}^1}(s_{i+1}) = \sigma_{C_1}(u_{i+1}^1) + \sigma_{\operatorname{aff}(G_{i+1}^1)}(v_{i+1}^1), \quad \sigma_{G_{i+1}^2}(-s_{i+1}) = \sigma_{C_2}(u_{i+1}^2) + \sigma_{\operatorname{aff}(G_{i+1}^2)}(v_{i+1}^2).$$

Combining this with $\theta_{i+1} := \sigma_{C_1}(u_{i+1}^1) + \zeta_{i+1}^1$ and (6.18), we obtain

$$\sup_{x \in F_{i+1}^1} \langle s_{i+1}, x \rangle = \sup_{x \in G_{i+1}^1} \langle s_{i+1}, x \rangle = \sigma_{G_{i+1}^1}(s_{i+1}) = \theta_{i+1},$$
$$\inf_{x \in F_{i+1}^2} \langle s_{i+1}, x \rangle = \inf_{x \in G_{i+1}^2} \langle s_{i+1}, x \rangle = -\sigma_{G_{i+1}^2}(-s_{i+1}) = \theta_{i+1}.$$

This completes the proof of item *(i)*.

Proof of (ii). Here, the situation is reversed. We are given $\mathcal{F} := (F_k^1, F_k^2)_{k=0}^i$ generated by $(s_k, \theta_k)_{k=0}^i \in \operatorname{Feas}(\text{EXT}_i)$ and need to show that if $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i$ is constructed as in the statement, it will belong to $\operatorname{Feas}(\text{EXT}_i\text{-nice})$ and the generated faces $\mathcal{G}$ will satisfy $\mathcal{G} = \mathcal{F}$.

We proceed by induction on i. First, let $i = 0$. We need to prove that

- $(u_0^1, v_0^1, \zeta_0^1, u_0^2, v_0^2, \zeta_0^2) \in \operatorname{Feas}(\text{EXT}_0\text{-nice})$ and
- $(F_0^1, F_0^2) = (G_0^1, G_0^2)$.

By assumption $(u_0^1, v_0^1, \zeta_0^1, u_0^2, v_0^2, \zeta_0^2)$ satisfies (6.9). Comparing (6.9) with (EXT$_i$-nice), we see that all constraints are satisfied[5]. Also, by definition, $(F_0^1, F_0^2) = (G_0^1, G_0^2) = (C_1, C_2)$. This proves that the statement holds for $i = 0$.

Next, suppose that the conclusion holds for some $i \geqslant 0$. We will prove that the conclusion holds for $i + 1$. That is,

[5] For $i = 0$, the tangent space constraint in (EXT$_i$-nice) is absent.

- $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^{i+1} \in \operatorname{Feas}(\mathrm{EXT}_{i+1}\text{-nice})$ holds and
- the generated faces $\mathcal{G} := (G_k^1, G_k^2)_{k=0}^{i+1}$ satisfy $\mathcal{G} = \mathcal{F} := (F_k^1, F_k^2)_{k=0}^{i+1}$.

By the induction hypothesis, we already have

$$\begin{aligned} (u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^{i} &\in \operatorname{Feas}(\mathrm{EXT}_i\text{-nice}) \\ (G_k^1, G_k^2) &= (F_k^1, F_k^2), \qquad k \in \{0, 1, \ldots, i\}. \end{aligned} \tag{6.20}$$

First we will check that $(F_{i+1}^1, F_{i+1}^2) = (G_{i+1}^1, G_{i+1}^2)$ holds.

The first step is to prove that

$$G_i^j \cap H_i = G_i^j \cap \{z \mid \langle u_i^j, z\rangle = \sigma_{C_j}(u_i^j)\}, \tag{6.21}$$

where H_i as in (EXT_i). By the induction hypothesis $F_i^j = G_i^j$ holds and we have $F_i^1 \cap F_i^2 = C_1 \cap C_2 \neq \varnothing$ by Lemma 5.1(*ii*). Let us examine the case $j = 1$ first. By definition, we have $H_i := \{z \mid \langle s_i, z\rangle = \theta_i\}$ and θ_i satisfies $\sigma_{F_i^1}(s_i) = \theta_i$ by (EXT_i). By the assumption that (6.9) holds, we have

$$\begin{aligned} \sigma_{F_i^1}(s_i) &= \sigma_{C_1}(u_i^1) + \sigma_{\operatorname{aff}(F_i^1)}(v_i^1) = \theta_i, \\ s_i &= u_i^1 + v_i^1. \end{aligned}$$

Then, by Lemma 6.3 and the fact that F_i^1 coincides with G_i^1 we obtain (6.21) for $j = 1$. For $j = 2$, the argument is analogous as follows. From (EXT_i) we have $\sigma_{F_i^2}(-s_i) = -\theta_i$, so that $H_i = \{z \mid \langle -s_i, z\rangle = \sigma_{F_i^2}(-s_i)\}$ holds. By the assumption that (6.9) holds we have $\sigma_{F_i^2}(-s_i) = \sigma_{C_2}(u_i^2) + \sigma_{\operatorname{aff}(F_i^2)}(v_i^2)$. Then, applying Lemma 6.3 to H_i, $F_i^2 = G_i^2$, $-s_i$, u_i^2 and v_i^2 implies that (6.21) holds when $j = 2$.

We have thus established that (6.21) holds for $j \in \{1, 2\}$. Having (6.21), using the same argument as in (6.17), we conclude that $F_{i+1}^j = G_{i+1}^j$ holds.

The final step is to show that $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^{i+1} \in \operatorname{Feas}(\mathrm{EXT}_{i+1}\text{-nice})$. By the induction hypothesis (see (6.20)), all that remains is to check that

$$\begin{aligned} &\left(u_{i+1}^j, v_{i+1}^j, \zeta_{i+1}^j\right) \in \operatorname{dom} \sigma_{C_j} \times \tan\left(\Big(\sum_{l=0}^{i} u_l^j, \sigma_{C_j}\big(\sum_{l=0}^{i} u_l^j\big)\Big), \operatorname{epi} \sigma_{C_j}\right), j \in \{1, 2\}, \\ &u_{i+1}^1 + v_{i+1}^1 = -u_{i+1}^2 - v_{i+1}^2, \\ &\sigma_{C_1}(u_{i+1}^1) + \sigma_{C_2}(u_{i+1}^2) + \zeta_{i+1}^1 + \zeta_{i+1}^2 \leqslant 0. \end{aligned} \tag{6.22}$$

Now, in view of the assumption that (6.9) holds, we only need to check that the tangent space constraint is satisfied.

By assumption, $u_{i+1}^j \in \operatorname{dom} \sigma_{C_j}$. We have also established that $(F_k^1, F_k^2) = (G_k^1, G_k^2)$ for $k \in \{0, 1, \ldots, i+1\}$. In particular, $F_i^j = G_i^j$ and the definition $\zeta_{i+1}^j := \sigma_{\operatorname{aff}(F_{i+1}^j)}(v_{i+1}^j)$ imply that $(v_{i+1}^j, \zeta_{i+1}^j) \in \operatorname{gra} \sigma_{\operatorname{aff}(G_{i+1}^j)}$.

Moreover, in view of Lemma 6.2 and the fact that $\widetilde{H}_i^j$ exposes G_{i+1}^j as face of C_j (see (6.6)), we have $\operatorname{gra} \sigma_{\operatorname{aff}(G_{i+1}^j)} = \tan\left(\left(\sum_{l=0}^{i} u_l^j, \sigma_{C_j}\left(\sum_{l=0}^{i} u_l^j\right)\right), \operatorname{epi} \sigma_{C_j}\right)$ and thus (6.22) holds. This completes the proof of item (ii).

Proof of (iii). By item (i) and the discussion in Section 5, we know $G_k^j \neq \varnothing$ for all $k \in \{0, 1, \ldots\}$ and $j \in \{1, 2\}$. Moreover, in view of (6.6), each face G_k^j is exposed by $\widetilde{H}_{k-1}^j$. Then, applying Lemma 6.2 we obtain

$$\operatorname{gra} \sigma_{\operatorname{aff}(G_k^j)} = \tan\left(\left(\sum_{l=0}^{k-1} u_l^j, \sigma_{C_j}(\sum_{l=0}^{k-1} u_l^j)\right), \operatorname{epi} \sigma_{C_j}\right),$$

which yields item (iii). ☐

Making use of the niceness and closedness of C_1 and C_2, we can finally obtain simplified expressions for $\operatorname{gra} \sigma_{\operatorname{aff}(F_{\min}^1)}$ and $\operatorname{gra} \sigma_{\operatorname{aff}(F_{\min}^2)}$. If we further have vertical niceness of f and g, the extended dual for (1.1) can be simplified to only involve f^*, g^* and Feas(EXT$_i$-nice) for any $i \geqslant \mathcal{SD}$. The following theorem summarizes these simplifications. In order to simplify the notation, for $(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i \in$ Feas(EXT$_i$-nice) we define for $j \in \{1, 2\}$:

$$\tan(u_i^j) := \tan\left(\Big(\sum_{k=0}^{i-1} u_k^j, \sigma_{C_j}\big(\sum_{k=0}^{i-1} u_k^j\big)\Big), \operatorname{epi} \sigma_{C_j}\right). \tag{6.23}$$

Theorem 6.5. *Suppose $C_1, C_2 \subseteq \mathcal{E}$ are closed nice convex sets with $C_1 \cap C_2 \neq \varnothing$. Let $F_{\min}^1$ and $F_{\min}^2$ be defined as in (3.2), then the following hold.*

(i) *For $i \geqslant \mathcal{SD}$,*

$$\operatorname{gra} \sigma_{\operatorname{aff}(F_{\min}^1)} \times \operatorname{gra} \sigma_{\operatorname{aff}(F_{\min}^2)} = \bigcup_{\substack{(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i \\ \in \operatorname{Feas}(\mathrm{EXT}_i\text{-nice})}} \tan(u_i^1) \times \tan(u_i^2), \tag{6.24}$$

where $\tan(u_i^j)$ is as in (6.23).

(ii) *Suppose that f, g in (1.1) are closed proper convex and vertically nice functions with $C_1 = \operatorname{dom} f$ and $C_2 = \operatorname{dom} g$ being closed. Then, for any $i \geqslant \mathcal{SD}$, the following is an extended dual for (1.1)*

$$\begin{aligned}
&\max -f^*(-\lambda - y) - \xi - g^*(\lambda - z) - \eta \\
&\text{s.t. } (y, \xi) \in \tan\left(\Big(\sum_{k=0}^{i-1} u_k^1, \sigma_{C_1}\big(\sum_{k=0}^{i-1} u_k^1\big)\Big), \operatorname{epi} \sigma_{C_1}\right), \\
&\qquad (z, \eta) \in \tan\left(\Big(\sum_{k=0}^{i-1} u_k^2, \sigma_{C_2}\big(\sum_{k=0}^{i-1} u_k^2\big)\Big), \operatorname{epi} \sigma_{C_2}\right), \\
&\qquad (u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^i \in \operatorname{Feas}(\mathrm{EXT}_i\text{-nice}).
\end{aligned} \tag{6.25}$$

That is, if (1.1) is feasible, then Definition 2.9 is satisfied for the pair (1.1) and (6.25).

Proof. We first prove item (i). It suffices to show (6.24) holds for $i = \mathcal{SD}$. Fix $j \in \{1, 2\}$ and $a \in F_{\min}^1 \cap F_{\min}^2$.

Recall that $\mathfrak{F}_{\mathcal{SD}}$ denotes all sequences of faces generated by (EXT$_{\mathcal{SD}}$), see (5.2). Let $(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}$. By Lemma 5.1(*ii*), all the faces F_k^j contain the minimal face $F_{\min}^j$. Therefore,

$$\sigma_{\mathrm{aff}(F_{\min}^j)} \leqslant \sigma_{\mathrm{aff}(F_{\mathcal{SD}}^j)},$$

which implies that the domain of $\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^j)}$ is contained in the domain of $\sigma_{\mathrm{aff}(F_{\min}^j)}$. This can be refined as follows. Since $a \in F_{\min}^j \subseteq F_{\mathcal{SD}}^j$ holds, from Lemma 2.6, we have

$$\sigma_{\mathrm{aff}(F_{\min}^j)}(s) = \langle s, a\rangle + \delta_{(\mathcal{L}_{F_{\min}^j})^\perp}(s) \leqslant \sigma_{\mathrm{aff}(F_{\mathcal{SD}}^j)}(s) = \langle s, a\rangle + \delta_{(\mathcal{L}_{F_{\mathcal{SD}}^j})^\perp}(s),$$

where $\mathcal{L}_{F_{\mathcal{SD}}^j}$ is the subspace defined in Lemma 2.6 (i.e., the subspace parallel to $\mathrm{aff}(F_{\mathcal{SD}}^j)$). This means that $\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^j)}(s)$ either coincides with $\sigma_{\mathrm{aff}(F_{\min}^j)}(s)$ or is $+\infty$. Therefore, the graph of $\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^j)}$ is entirely contained in the graph of $\sigma_{\mathrm{aff}(F_{\min}^j)}$.

This leads to the following observation:

$$\operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^2)} \supseteq \bigcup_{(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}} \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^2)}.$$

Using Lemma 5.1(*i*), there exists $\overline{\mathcal{F}} = (\overline{F}_k^1, \overline{F}_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}$ such that $\overline{F}_{\mathcal{SD}}^j = F_{\min}^j$. This implies the containment

$$\operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^2)} \subseteq \bigcup_{(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}} \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^2)}.$$

Overall, we have

$$\operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^2)} = \bigcup_{(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}} \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^2)}.$$

Now, items (*i*) and (*ii*) of Proposition 6.4 imply that every sequence of faces $(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}$ (i.e., generated by (EXT$_{\mathcal{SD}}$)) is also generated by (EXT$_{\mathcal{SD}}$-nice) and vice-versa. Therefore

$$\begin{aligned}
\operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\min}^2)} &= \bigcup_{(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \in \mathfrak{F}_{\mathcal{SD}}} \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^2)} \\
&= \bigcup_{\substack{(F_k^1, F_k^2)_{k=0}^{\mathcal{SD}} \text{ generated by} \\ (\mathrm{EXT}_{\mathcal{SD}}\text{-nice})}} \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^1)} \times \operatorname{gra}\sigma_{\mathrm{aff}(F_{\mathcal{SD}}^2)} \\
&= \bigcup_{\substack{(u_k^1, v_k^1, \zeta_k^1, u_k^2, v_k^2, \zeta_k^2)_{k=0}^{\mathcal{SD}} \\ \in \mathrm{Feas}(\mathrm{EXT}_{\mathcal{SD}}\text{-nice})}} \tan(u_i^1) \times \tan(u_i^2),
\end{aligned}$$

where the third equality comes from item (*iii*) of Proposition 6.4.

We next show item (*ii*). Suppose that (1.1) is feasible. Since f and g are vertically nice, the pair (1.1) and (4.12) satisfy strong duality by Theorem 4.13. We consider two cases.

First, suppose that the common optimal value of (1.1) and (4.12) is finite and let y^*, λ^*, z^* be an optimal solution to (4.12). Let $\xi^* := \sigma_{\mathrm{aff}(F^1_{\min})}(y^*)$, $\eta^* := \sigma_{\mathrm{aff}(F^2_{\min})}(z^*)$. Then,

$$(y^*, \xi^*, z^*, \eta^*) \in \operatorname{gra} \sigma_{\mathrm{aff}(F^1_{\min})} \times \operatorname{gra} \sigma_{\mathrm{aff}(F^2_{\min})}.$$

By item (i), there exists $(u^1_k, v^1_k, \zeta^1_k, u^2_k, v^2_k, \zeta^2_k)^i_{k=0} \in \mathrm{Feas}$(EXT$_i$-nice), such that

$$(y^*, \xi^*) \in \tan(u^1_i), \qquad (z^*, \eta^*) \in \tan(u^2_i),$$

where $\tan(u^j_i)$ is as in (6.23). This shows that (6.25) has a feasible solution that attains the optimal value of (1.1).

It only remains to check that no feasible solution of (6.25) has a larger objective function value than the optimal value of (1.1). This is also a consequence of item (i): if λ, y, ξ, z, η satisfy the constraints in (6.25), then (y, ξ, z, η) must be an element of $\operatorname{gra} \sigma_{\mathrm{aff}(F^1_{\min})} \times \operatorname{gra} \sigma_{\mathrm{aff}(F^2_{\min})}$, which implies that λ, y, z is a feasible solution to (4.12) affording the same objective function value. Therefore, the corresponding objective function value cannot be larger than the optimal value of (4.12).

Finally, suppose that the common optimal value of (1.1) and (4.12) is $+\infty$. We need to show that the optimal value of (6.25) is $+\infty$ too. In order to derive a contradiction, suppose that (6.25) has a solution with finite objective function value. In particular, the λ, y, ξ, z, η components of such a solution satisfy the constraints in (6.25). Then, by item (i), λ, y, z would be a solution to (4.12) with finite objective function value, which is a contradiction. $\square$

Before we end this section, we present a brief detour on the relationship between niceness and facial exposedness. It is known that a nice closed convex cone must be facially exposed [29]. Here, we also show a corollary of Proposition 6.4 that a nice closed convex set must be facially exposed.

Corollary 6.6. *A nice closed convex set is facially exposed.*

Proof. Let $C \subseteq \mathcal{E}$ be a nice closed convex set, F be an arbitrary face of C, and $a \in \operatorname{ri} F$.

Then, F is the minimal face of C containing a by (2.3). Let $C_1 := C$ and $C_2 := \{a\}$. Then, $F^1_{\min}$ and $F^2_{\min}$ as in (3.2) must be such that $F^1_{\min} = F$ and $F^2_{\min} = \{a\}$.

Applying Lemma 5.1 to C_1 and C_2, we conclude that, for some $i \geqslant 0$, there exists a sequence of faces $\mathcal{F} = (F^1_k, F^2_k)^i_{k=0}$ generated by some $(s_k, \theta_k)^i_{k=0} \in$ (EXT$_i$) such that

$$F^1_i = F \quad \text{and} \quad F^2_i = \{a\}.$$

By Proposition 6.4(ii), there exists $\mathcal{G} = (G^1_k, G^2_k)^i_{k=0}$ generated by some $(u^1_k, v^1_k, \zeta^1_k, u^2_k, v^2_k, \zeta^2_k)^i_{k=0} \in \mathrm{Feas}$(EXT$_i$-nice) such that $\mathcal{F} = \mathcal{G}$. From the definitions of $\widetilde{H}^j_k$ and G^j_k in (6.5) and (6.6), respectively, and noting that $\sum_{l=0}^{i-1} u^1_l \in \operatorname{dom} \sigma_C$, we know that $F = F^1_i = G^1_i$ is exposed as face of C by the hyperplane $\widetilde{H}^1_{i-1}$ which satisfies

$$\widetilde{H}^1_{i-1} = \left\{ z \mid \langle z, \sum_{l=0}^{i-1} u^j_l \rangle = \sigma_C\big(\sum_{l=0}^{k-1} u^j_l\big) \right\}.$$

Since F is arbitrary, we conclude that C is facially exposed. $\square$

7 An Example

In this section, we give an example inspired by [33, Example 4] to illustrate Algorithm 1 and the extended dual discussed in Sections 5 and 6. In what follows, we let $\mathcal{S}^n$ denote the space of $n \times n$ real symmetric matrices and $\mathcal{S}^n_+$ the positive semidefinite cone in $\mathcal{S}^n$. Let $f : \mathcal{S}^3_+ \to \mathbb{R} \cup \{+\infty\}$ be such that

$$f(X) = \begin{cases} \frac{X_{11}^2}{2} & \text{if } X \in \mathcal{S}^3_+; \\ +\infty & \text{otherwise.} \end{cases}$$

and let $g := \delta_S$ with $S := \left\{X \in \mathcal{S}^3 \mid X_{22} \leqslant 0, X_{11} + 2X_{23} = 1\right\}$. As in (1.1) we consider the problem of minimizing $f + g$ which can be written as follows.

$$\begin{aligned} \min_{X \in \mathbb{R}^{3\times 3}} \quad & f(X) = \frac{X_{11}^2}{2} \\ \text{s.t.} \quad & X \in \operatorname{dom} f \cap S = \mathcal{S}^3_+ \cap \left\{X \in \mathcal{S}^3 \mid X_{22} \leqslant 0, X_{11} + 2X_{23} = 1\right\}. \end{aligned} \tag{7.1}$$

Suppose that X is feasible for (7.1). The constraint $X_{22} \leqslant 0$ implies $X_{22} = 0$ and $X_{12} = X_{21} = X_{23} = X_{32} = 0$. Therefore, $X_{11} = 1$. It follows that

$$\operatorname{dom} f \cap S = \left\{ X \mid X = \begin{pmatrix} 1 & 0 & \beta \\ 0 & 0 & 0 \\ \beta & 0 & \alpha \end{pmatrix} \in \mathcal{S}^3_+, \text{ with } \alpha \geqslant 0, \beta \in \mathbb{R} \right\}. \tag{7.2}$$

Hence the optimal value of (7.1) is $\nu_p^* = 0.5$. Moreover, (7.1) does not satisfy Slater's condition since for any $X \in \operatorname{ri}(\operatorname{dom} f)$, $X_{22} > 0$ must hold so $X \notin S$, implying $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(S) = \emptyset$.

Our first task is to write down the Fenchel dual of (7.1) as in (1.2). The function f is the sum of the function $X \mapsto \frac{X_{11}^2}{2}$ with $\delta_{\mathcal{S}^3_+}$. As the domain of the former is the whole space, the conjugate of the sum is the exact infimal convolution of the conjugates, e.g., see [33, Theorem 16.4]. Therefore

$$\begin{aligned} f^*(\Lambda) &= \min_{Y,Z} \left\{ \frac{Y_{11}^2}{2} + \delta_{-\mathcal{S}^3_+}(Z) \mid Y + Z = \Lambda, Y_{ij} = 0 \ \forall (i,j) \neq (1,1) \right\} \\ &= \min_{z} \left\{ \frac{(\Lambda_{11} - z)^2}{2} \mid \begin{pmatrix} z & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{pmatrix} \in -\mathcal{S}^3_+ \right\}. \end{aligned}$$

The set S is polyhedral and is the intersection of $P_1 := \{X \in \mathcal{S}^3 \mid X_{22} \leqslant 0\}$ and $P_2 := \{X \in \mathcal{S}^3 \mid X_{11} + 2X_{23} = 1\}$. Therefore, the support function of S is the infimal convolution between the indicator function of the polar cone of P_1, i.e., $P_1^\circ = \{X \in \mathcal{S}^3 \mid X_{22} \geqslant 0, X_{ij} = 0, (i,j) \neq (2,2)\}$, and the support function of P_2, which satisfies

$$\sigma_{P_2}(Z) = \begin{cases} Z_{11} & \text{if } Z \in \left\{ \begin{pmatrix} \beta & 0 & 0 \\ 0 & 0 & \beta \\ 0 & \beta & 0 \end{pmatrix} \mid \beta \in \mathbb{R} \right\}; \\ +\infty & \text{otherwise.} \end{cases}$$

Overall, we have

$$\sigma_S(Z) = \begin{cases} Z_{11} & \text{if } Z \in \left\{ \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix} \mid \beta \in \mathbb{R}, \alpha \geqslant 0 \right\}; \\ +\infty & \text{otherwise.} \end{cases} \tag{7.3}$$

Making use of auxiliary variables, the Fenchel dual problem of (7.1) can be written as

$$\begin{aligned} \max_{\Lambda \in \mathcal{S}^3, z, \alpha, \beta \in \mathbb{R}} \quad & -\frac{(-\Lambda_{11} - z)^2}{2} - \Lambda_{11} \\ \text{s.t.} \quad & \begin{pmatrix} z & 0 & 0 \\ 0 & -\alpha & -\beta \\ 0 & -\beta & 0 \end{pmatrix} \in -\mathcal{S}^3_+, \qquad \Lambda = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix}, \beta \in \mathbb{R}, \alpha \geqslant 0. \end{aligned} \tag{7.4}$$

Suppose that $\Lambda, z, \alpha, \beta$ is feasible for (7.4). From the first constraint in (7.4) we have $\beta = 0$. From the second constraint, we have $\beta = \Lambda_{11}$ and, therefore, $\Lambda_{11} = 0$ holds. So the maximal value achievable by $-\frac{(-\Lambda_{11}-z)^2}{2} - \Lambda_{11} = -z^2/2$ is zero, which can be obtained by letting $\Lambda = 0$, $\alpha = \beta = z = 0$. Noting that $v_d^* = 0 \neq 0.5 = v_p^*$, there is a positive duality gap, making the problem difficult to solve.

Indeed, we tried to solve (7.1) with the modelling tools JuMP [24], CVX [6] and different choices of solvers such as SDPT3 [43, 44], SeDuMi [41], Hypatia [5] and Clarabel [9]. In all cases, the solvers fail to compute the correct optimal value of (7.1).

Applying Facial Reduction Let $C_1 := \operatorname{dom} f = \mathcal{S}^3_+$ and $C_2 := S$. We apply Algorithm 1 to C_1 and C_2 in hopes of closing the duality gap. The first step is obtaining S_0 that separates properly C_1 and C_2. For that, we will compute $(C_2 - C_1)^*$. Here we are using S_0 instead of s_0 to emphasize that S_0 is matrix. We have

$$C_2^* = \left\{ \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix} \mid \alpha \leqslant 0, \beta \geqslant 0 \right\}.$$

Since $C_1^\circ = -\mathcal{S}^3_+$, we have from Lemma 2.2 that

$$(C_2 - C_1)^* = C_2^* \cap C_1^\circ = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \alpha \leqslant 0 \right\}. \tag{7.5}$$

We can select

$$S_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in C_2^* \cap C_1^\circ = (C_2 - C_1)^*, \quad \theta_0 = \sigma_{C_1}(S_0) = 0, \tag{7.6}$$

and

$$H_0 = \{S_0\}^\perp = \left\{ X \in \mathcal{S}^3 \mid X_{22} = 0 \right\}.$$

Then,

$$F_1^1 = C_1 \cap H_0 = \left\{ \begin{pmatrix} \alpha & 0 & \beta \\ 0 & 0 & 0 \\ \beta & 0 & \gamma \end{pmatrix} \in \mathcal{S}^3_+ \mid \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \in \mathcal{S}^2_+ \right\},$$

$$F_1^2 = C_2 \cap H_0 = \left\{ X \in \mathcal{S}^3 \mid X_{22} = 0, X_{11} + 2X_{23} = 1 \right\}.$$

We can now readily check that

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \operatorname{ri}(F_1^1) \cap \operatorname{ri}(F_1^2),$$

which implies that $\operatorname{ri}(F_1^1) \cap \operatorname{ri}(F_1^2) \neq \varnothing$. Moreover, recalling (7.2), we have $F_1^1 \cap F_1^2 = C_1 \cap C_2$. The facial reduction algorithm therefore terminates and produces

$$F_{\min}^1 = F_1^1, \quad F_{\min}^2 = F_1^2.$$

We have

$$\begin{aligned}
(F_{\min}^1)^\circ &= \left\{ \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} \in \mathcal{S}^3 \mid \begin{pmatrix} X_{11} & X_{13} \\ X_{31} & X_{33} \end{pmatrix} \in -\mathcal{S}_+^2 \right\}, \\
\sigma_{F_{\min}^1} &= \delta_{(F_{\min}^1)^\circ}, \\
\sigma_{\operatorname{aff}(F_{\min}^1)} &= \delta_{(F_{\min}^1)^\perp}.
\end{aligned} \tag{7.7}$$

Similar to (7.3), we obtain

$$\sigma_{\operatorname{aff}(F_{\min}^2)}(Z) = \sigma_{F_{\min}^2}(Z) = \begin{cases} Z_{11} & \text{if } Z \in \left\{ \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix} \mid \beta \in \mathbb{R}, \alpha \in \mathbb{R} \right\}, \\ +\infty & \text{otherwise}, \end{cases} \tag{7.8}$$

where the first equality holds because $F_{\min}^2$ is an affine subspace. Let $\widehat{f} : \mathcal{S}^3 \to \mathbb{R}$ be the function mapping X to $X_{11}^2/2$. Then $f = \widehat{f} + \delta_{\mathcal{S}_+^3}$ is vertically nice by Corollary 4.11. Therefore,

$$(f_{|F_{\min}^1})^* = f^* \boxdot \sigma_{\operatorname{aff}(F_{\min}^1)} = (\widehat{f} + \delta_{\mathcal{S}_+^3})^* \boxdot \sigma_{\operatorname{aff}(F_{\min}^1)} = (\widehat{f}^* \boxdot \delta_{-\mathcal{S}_+^3}) \boxdot \sigma_{\operatorname{aff}(F_{\min}^1)} = \widehat{f}^* \boxdot (\delta_{-\mathcal{S}_+^3} \boxdot \sigma_{\operatorname{aff}(F_{\min}^1)}),$$

where the third equality comes from [35, Theorem 16.4]; the last equality follows by the associativity of the infimal convolution. Since $F_{\min}^1 \trianglelefteq \mathcal{S}_+^3$, $\sigma_{\operatorname{aff}(F_{\min}^1)} = \delta_{(F_{\min}^1)^\perp}$ and $\mathcal{S}_+^3$ is a nice cone, we have $(F_{\min}^1)^\circ = -\mathcal{S}_+^3 + (F_{\min}^1)^\perp$ and $\delta_{-\mathcal{S}_+^3} \boxdot \sigma_{\operatorname{aff}(F_{\min}^1)} = \delta_{-\mathcal{S}_+^3 + (F_{\min}^1)^\perp} = \delta_{(F_{\min}^1)^\circ}$. Therefore,

$$(f_{|F_{\min}^1})^*(\Lambda) = \widehat{f}^* \boxdot \delta_{(F_{\min}^1)^\circ}(\Lambda) = \min_z \left\{ \frac{(\Lambda_{11} - z)^2}{2} \mid \begin{pmatrix} z & \Lambda_{13} \\ \Lambda_{31} & \Lambda_{33} \end{pmatrix} \in -\mathcal{S}_+^2 \right\}. \tag{7.9}$$

Using (7.8) and (7.9), we consider the problem of maximizing $-(f_{|F_{\min}^1})^*(-\Lambda) - (g_{|F_{\min}^2})^*(\Lambda)$, which, after introducing auxiliary variables, becomes

$$\begin{aligned}
\max_{\Lambda \in \mathcal{S}^3, z, \alpha, \beta \in \mathbb{R}} \quad & -\frac{(-\Lambda_{11} - z)^2}{2} - \Lambda_{11} \\
\text{s.t.} \quad & \begin{pmatrix} z & 0 \\ 0 & 0 \end{pmatrix} \in -\mathcal{S}_+^2, \qquad \Lambda = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix}, \beta \in \mathbb{R}, \alpha \in \mathbb{R}.
\end{aligned} \tag{7.10}$$

Theorem 3.1 asserts that (7.1) and (7.10) satisfy strong duality. Indeed, letting

$$\Lambda = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

and $z = 0$, we obtain a feasible solution (7.10) whose corresponding objection function value is $0.5 = \nu_p^*$. We have thus closed the duality gap.

Formulae for $\sigma_{F^1_{\min}}$ and $\sigma_{F^2_{\min}}$ The discussion so far required the computation of $F^1_{\min}$ and $F^2_{\min}$. Next, we will pretend that we do not know the pair $F^1_{\min}$, $F^2_{\min}$ and, using Theorem 5.3, we will compute $\sigma_{F^1_{\min}}$ and $\sigma_{F^2_{\min}}$ without explicitly determining $F^1_{\min}$ and $F^2_{\min}$ beforehand. Of course, there is no magic here, so we must assume that we know f and g "well enough", including its conjugate functions, faces of the domain and so on.

In principle, as we do not know $F^1_{\min}$ and $F^2_{\min}$, we also do not know that the joint singularity degree is $\mathcal{SD} = 1$ (i.e., facial reduction stops in one step). However, we have access to the upper bound ℓ in (3.11), which only depends on our knowledge of the faces of $C_1 = \operatorname{dom} f$ and $C_2 = \operatorname{dom} g$. As $\ell_{C_1} = 4$, $\ell_{C_2} = 2$, $\dim(\mathcal{E}) = 6$, we know that

$$\ell = \min\{4 + 2 - 2, 6 + 1\} = 4.$$

This tells us that invoking Theorem 5.3 with $i = 4$ is enough.

First, we determine Feas(EXT$_1$). Then, for any $(\widehat{s}_k, \widehat{\theta}_k)_{k=0}^1 \in$ Feas(EXT$_1$), for any $k = 0, 1$, $(\widehat{s}_k, \widehat{\theta}_k)$ takes three possibilities:

(I) $(\widehat{s}_k, \widehat{\theta}_k) \in (C_2 - C_1)^*$ that generates C_1 and C_2, i.e., a non-FRA reducing step. By the formula of $(C_2 - C_1)^*$ in (7.5), we must have $(\widehat{s}_k, \widehat{\theta}_k) = (0, 0)$.

(II) $(\widehat{s}_k, \widehat{\theta}_k) \in (C_2 - C_1)^*$ that generates F^1_1 and F^1_2, which can only appear up to once. From (7.6) we know that

$$\widehat{s}_k = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ with } \alpha < 0, \quad \widehat{\theta}_k = 0.$$

(III) $(\widehat{s}_k, \widehat{\theta}_k) \in (F^2_1 - F^1_1)^*$. Using Lemma 2.2, we have

$$\begin{aligned} &(F^2_1 - F^1_1)^* = (F^2_1)^* \cap (F^1_1)^\circ \\ &= \left\{ \left(\begin{smallmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{smallmatrix} \right) \mid \beta \geqslant 0, \alpha \in \mathbb{R} \right\} \bigcap \left\{ \left(\begin{smallmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{smallmatrix} \right) \in \mathcal{S}^3 \mid \left(\begin{smallmatrix} X_{11} & X_{13} \\ X_{31} & X_{33} \end{smallmatrix} \right) \in -\mathcal{S}^2_+ \right\}. \end{aligned}$$

Therefore,

$$\widehat{s}_k \in (F^2_1 - F^1_1)^* = \left\{ \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{smallmatrix} \right) \mid \alpha \in \mathbb{R} \right\}, \quad \widehat{\theta}_k = 0.$$

Collecting the above three possibilities together, we conclude that

$$\mathrm{Feas}(\mathrm{EXT}_1) = \{((0,0),(0,0))\} \bigcup \left\{((0,0),(s_1,\theta_1)) \mid s_1 = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right) \text{ with } \alpha < 0, \theta_1 = 0\right\} \bigcup$$
$$\left\{(s_k,\theta_k)_{k=0}^{1} \mid s_0 = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right) \text{ with } \alpha < 0, \theta_0 = 0; s_1 = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right) \text{ with } \alpha \in \mathbb{R}, \theta_1 = 0\right\}.$$

Before we go through the trouble of computing $\mathrm{Feas}(\mathrm{EXT}_2)$, we consider the expression in (5.10). For every $j \in \{1,2\}$ and every $i \geqslant \mathcal{SD}$, in order to compute the right-hand-side of (5.10), we need to consider sums of the form $\sum_{k=0}^{i-1} r_k$ and $\sum_{k=0}^{i-1} \rho_k$ with $(r_k,\rho_k)_{k=0}^{i}$ in $\mathrm{span}\,(\mathrm{Feas}(\mathrm{EXT}_i))$. We will save the reader the trouble of checking that increasing i do not change the possible values that the sums $\sum_{k=0}^{i-1} r_k$ and $\sum_{k=0}^{i-1} \rho_k$ can take. This makes sense in view of our forbidden knowledge that $\mathcal{SD}$ is in fact one.

So for $i = 1$, the possible values of the sums $\sum_{k=0}^{i-1} r_k = r_0$ and $\sum_{k=0}^{i-1} \rho_k = \rho_0$ is as follows:

$$\left\{\left(\sum_{k=0}^{1-1} r_k, \sum_{k=0}^{1-1} \rho_k\right) \mid (r_k,\rho_k)_{k=0}^{1} \in \mathrm{span}\,(\mathrm{Feas}(\mathrm{EXT}_1))\right\} = \left\{\left(\begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{pmatrix}, 0\right) \mid \alpha \in \mathbb{R}\right\}. \quad (7.11)$$

Next, for $j = 1$ and $i = 1$, we consider the expression in (5.10) in view of (7.11). First, we analyze the term inside the "cl", i.e., just the infimum part in (5.10). Now, σ_{C_1} is the indicator function of $-\mathcal{S}^3_+$, so, since the "sum of the ρ_k's" part in (7.11) is zero, the infimum in (5.10) is 0 or $+\infty$. The infimum will be 0 if and only if there exists $\alpha \in \mathbb{R}$ such that

$$X - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{dom}\,\sigma_{C_1} = -\mathcal{S}^3_+.$$

Put otherwise, the infimum is 0 if and only if X belongs to $-\mathcal{S}^3_+ + \mathrm{span}\,(\{E\})$, where $E = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. However, $-\mathcal{S}^3_+ + \mathrm{span}\,(\{E\})$ is the cone of feasible directions of $-\mathcal{S}^3_+$ at E, see (2.4). The closure of $\mathrm{dir}(E, -\mathcal{S}^3_+)$ is the tangent cone of $-\mathcal{S}^3_+$ at E. Therefore, (5.10) corresponds to the assertion that $\sigma_{F^1_{\min}}$ satisfies

$$\sigma_{F^1_{\min}} = \mathrm{cl}\,\delta_{\mathrm{dir}(E,-\mathcal{S}^3_+)} = \delta_{\mathrm{tanCone}(E,-\mathcal{S}^3_+)}.$$

Also, in view of (2.5) and [30, (28.4)], we have

$$\mathrm{tanCone}(E, -\mathcal{S}^3_+) = \left\{\begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} \in \mathcal{S}^3 \mid \begin{pmatrix} X_{11} & X_{13} \\ X_{31} & X_{33} \end{pmatrix} \in -\mathcal{S}^2_+\right\}.$$

This shows that the expression for $\sigma_{F^1_{\min}}$ obtained through Theorem 5.3 is, as expected, the indicator function of $F^\circ_{\min}$ as computed in (7.7). This also shows that the closure operation in Theorem 5.3 cannot be omitted in general.

As for $\sigma_{F^2_{\min}}$, using the same techniques and recalling the formula of σ_{C_2} in (7.3), we can readily verify the formula of $\sigma_{F^2_{\min}}$ as in (7.8).

Utilizing niceness In this example, f and g are vertically nice, C_1 and C_2 are both nice, so the techniques of Section 6 apply. Here we play the same game as before and pretend that we do not know $F^1_{\min}$ and $F^2_{\min}$. On the other hand, we assume we know deeply the structure of the individual sets C_1, C_2. Then, the goal is to compute the graph of $\sigma_{F^j_{\min}}$ for $j \in \{1, 2\}$ using Theorem 6.5. This, by its turn, requires that we determine Feas(EXT$_i$-nice) for some $i \geqslant \mathcal{SD}$. As before, in view of (3.11), it is enough to take $i = 4$.

First we compute Feas(EXT$_1$-nice). This requires understanding the tangent spaces of the epigraphs of σ_{C_1} and σ_{C_2}. Since $\sigma_{C_1} = \delta_{C_1^\circ}$, we have $\operatorname{epi}\sigma_{C_1} = C_1^\circ \times \mathbb{R}_+ = -\mathcal{S}^3_+ \times \mathbb{R}_+$. Then, for any $X \in \operatorname{dom}\sigma_{C_1}$,

$$\begin{aligned} &\tan\left((X, \sigma_{C_1}(X)), \operatorname{epi}\sigma_{C_1}\right) = \tan\left((X, 0), C_1^\circ \times \mathbb{R}_+\right) \\ &= \tan\left(X, C_1^\circ\right) \times \tan\left(0, \mathbb{R}_+\right) = \tan\left(X, -\mathcal{S}^3_+\right) \times \{0\}. \end{aligned} \tag{7.12}$$

The tangent space of $-\mathcal{S}^3_+$ at some $X \in -\mathcal{S}^3_+$ is well-understood. It can be computed, for example, using (2.5), see also [30, Proposition 28.1] or the discussion in [28, Section 2.5]. Meanwhile, recalling σ_{C_2} in (7.3), for $X = \begin{pmatrix} \beta_0 & 0 & 0 \\ 0 & \alpha_0 & \beta_0 \\ 0 & \beta_0 & 0 \end{pmatrix} \in \operatorname{dom}\sigma_{C_2}$ with $\alpha_0 \geqslant 0$, we have

$$\tan\left((X, \sigma_{C_2}(X)), \operatorname{epi}\sigma_{C_2}\right) = \begin{cases} \left\{ \left(\begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix}, \ \beta \right) \;\middle|\; \alpha, \beta \in \mathbb{R} \right\}, & \text{if } \alpha_0 > 0, \\ \left\{ \left(\begin{pmatrix} \beta & 0 & 0 \\ 0 & 0 & \beta \\ 0 & \beta & 0 \end{pmatrix}, \ \beta \right) \;\middle|\; \beta \in \mathbb{R} \right\}, & \text{if } \alpha_0 = 0. \end{cases} \tag{7.13}$$

Next, let $(u^1_k, v^1_k, \zeta^1_k, u^2_k, v^2_k, \zeta^2_k)_{k=0}^1 \in$ Feas(EXT$_1$-nice). For $k = 0$ we have the following constraints:

$$\begin{aligned} &(u^1_0, v^1_0, \zeta^1_0) \in \operatorname{dom}\sigma_{C_1} \times \operatorname{gra}\sigma_{\operatorname{aff}(C_1)}, \qquad (u^2_0, v^2_0, \zeta^2_0) \in \operatorname{dom}\sigma_{C_2} \times \operatorname{gra}\sigma_{\operatorname{aff}(C_2)}, \\ &u^1_0 + v^1_0 = -u^2_0 - v^2_0 \quad \text{ and } \quad \sigma_{C_1}(u^1_0) + \sigma_{C_2}(u^2_0) + \zeta^1_0 + \zeta^2_0 \leqslant 0. \end{aligned} \tag{7.14}$$

Since $\operatorname{aff}(C_1) = \mathcal{S}^3$ and hence $\operatorname{gra}\sigma_{\operatorname{aff}(C_1)} = \{(0, 0)\}$, we know $v^1_0 = 0$ and $\zeta^1_0 = 0$. Noticing that $\operatorname{aff}(C_2) = \{X \in \mathcal{S}^3 \mid X_{11} + 2X_{23} = 1\}$, it holds that

$$\sigma_{\operatorname{aff}(C_2)}(Z) = \begin{cases} Z_{11} & \text{if } Z \in \left\{ \begin{pmatrix} \beta & 0 & 0 \\ 0 & 0 & \beta \\ 0 & \beta & 0 \end{pmatrix} \mid \beta \in \mathbb{R} \right\}; \\ +\infty & \text{otherwise.} \end{cases}$$

Recall from (7.3) that $\operatorname{dom}\sigma_{C_2} = \left\{ \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix} \mid \beta \in \mathbb{R}, \alpha \geqslant 0 \right\}$. Then, there exist $\beta_1, \beta_2 \in \mathbb{R}$ and $\alpha_0 \geqslant 0$ such that

$$u^2_0 = \begin{pmatrix} \beta_1 & 0 & 0 \\ 0 & \alpha_0 & \beta_1 \\ 0 & \beta_1 & 0 \end{pmatrix}, \quad v^2_0 = \begin{pmatrix} \beta_2 & 0 & 0 \\ 0 & 0 & \beta_2 \\ 0 & \beta_2 & 0 \end{pmatrix}.$$

Then,

$$u^2_0 + v^2_0 = \begin{pmatrix} \beta_1 + \beta_2 & 0 & 0 \\ 0 & \alpha_0 & \beta_1 + \beta_2 \\ 0 & \beta_1 + \beta_2 & 0 \end{pmatrix} = -u^1_0 \in -\operatorname{dom}\sigma_{C_1} = \mathcal{S}^3_+, \quad \sigma_{C_2}(u^2_0) + \sigma_{\operatorname{aff}(C_2)}(v^2_0) = \beta_1 + \beta_2.$$

As $-u^1_0 = u^2_0 + v^2_0$ must be positive semidefinite, $\beta_1 + \beta_2 = 0$ holds. This implies that

$$\sigma_{C_2}(u^2_0) + \zeta^2_0 = \sigma_{C_2}(u^2_0) + \sigma_{\operatorname{aff}(C_2)}(v^2_0) = 0.$$

Therefore, any $(u_0^1, v_0^1, \zeta_0^1, u_0^2, v_0^2, \zeta_0^2)$ satisfying (7.14) must be of the form

$$u_0^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\alpha_0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad v_0^1 = 0, \quad u_0^2 + v_0^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha_0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \zeta_0^1 = 0, \quad \zeta_0^2 = \beta_2, \tag{7.15}$$

with $\alpha_0 \geqslant 0$.

In view of (7.12) and (7.13), if $\alpha_0 > 0$ we have

$$\begin{aligned} \tan\left((u_0^1, \sigma_{C_1}(u_0^1)), \operatorname{epi} \sigma_{C_1}\right) &= \tan(u_0^1, -\mathcal{S}^3_+) \times \{0\} = \left(\left\{ \begin{pmatrix} 0 & v^1_{12,1} & 0 \\ v^1_{21,1} & v^1_{22,1} & v^1_{23,1} \\ 0 & v^1_{32,1} & 0 \end{pmatrix} \in \mathcal{S}^3 \right\}\right) \times \{0\}, \\ \tan\left((u_0^2, \sigma_{C_2}(u_0^2)), \operatorname{epi} \sigma_{C_2}\right) &= \left\{ \left(\begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & 0 \end{pmatrix}, \ \beta \right) \;\middle|\; \alpha, \beta \in \mathbb{R} \right\}. \end{aligned} \tag{7.16}$$

If $\alpha_0 = 0$, the tangent space $\tan\left((u_0^1, \sigma_{C_1}(u_0^1)), \operatorname{epi} \sigma_{C_1}\right)$ only contains the zero element. And, $\tan\left((u_0^2, \sigma_{C_2}(u_0^2)), \operatorname{epi} \sigma_{C_2}\right)$ becomes as in the second case of (7.13).

Here we will also omit the computations showing that if we increase i the right-hand-sides in (7.16) do not change. This is consistent with the fact that $\mathcal{SD} = 1$. And, indeed, the right-hand-sides in (7.16) correspond to the graphs of $\sigma_{\operatorname{aff}(F^1_{\min})}$ and $\sigma_{\operatorname{aff}(F^2_{\min})}$, respectively. This follows from (7.7) and (7.8), respectively. Plugging f^*, Feas(EXT$_i$-nice) in (7.15), and the right-hand-sides of (7.16) into (6.25) gives an extended dual that is equivalent to (7.10).

8 Concluding Remarks

The goal of this work was to design a facial reduction algorithm appropriate for handling the problem (1.1) in order to regularize it so that Slater's condition is satisfied. More genrally, given two nonempty convex sets C_1 and C_2, our proposed Algorithm 1 is able to either find faces $F^1_{\min}, F^2_{\min}$ as in (3.2), or to verify that their intersection is empty. We can then apply this to (1.1) by setting $C_1 = \operatorname{dom} f$ and $C_2 = \operatorname{dom} g$.

Restricting f and g to faces of their respective domains leads to the problem of computing the conjugates $f^*_{|F^1_{\min}}$ and $g^*_{|F^2_{\min}}$. By its turn, this provides a natural motivation for the notion of nice convex sets and the class of vertically nice functions considered in Section 4. We then leverage these tools to obtain extended duals that, while fairly technical, to do not require that $F^1_{\min}$ and $F^2_{\min}$ be explicitly determined beforehand. This is first done in as much generality as possible (Theorem 5.4) and, then, under closedness and niceness assumptions (Theorem 6.5).

In conic linear programming there are many applications of facial reduction and its natural to wonder which ones could be extended to general convex programs in a natural way. For example, in [23, 16, 17, 15, 14] the authors used a framework based on the facial reduction algorithm and the so-called *facial residual functions* to deduce error bounds for conic linear programmings. The idea of using facial reduction algorithm to deduce error bounds dates back to Sturm's pioneering work [42]. This inspires one possible direction of future research: can we use the facial reduction algorithm proposed in this paper to deduce error bounds for general convex feasibility problems as in (3.1)?

Finally, we remark that while this work was in its finishing stages, a paper by Scott [39] was released that has a very similar motivation as ours: extending facial reduction ideas from the conic setting to the general convex optimization case. Along the way, Scott proves several interesting results such as one that is analogous to Theorem 3.1, e.g., see [39, Corollary 5.5]. That said, apart from the philosophical similarities, we believe the technical development of both papers is fairly distinct. In particular, the bilateral facial reduction algorithm presented in [39] seems quite different from Algorithm 1 described here. Furthermore, as far as we could see, the discussions and results described in Sections 4, 5 and 6 do not seem to have a counterpart in [39].

References

[1] Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[2] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM, Philadelphia, January 2001.

[3] Jon M. Borwein and Henry Wolkowicz. Facial reduction for a cone-convex programming problem. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 30(3):369–380, February 1981.

[4] Jonathan M. Borwein and Henry Wolkowicz. Regularizing the abstract convex program. *Journal of Mathematical Analysis and Applications*, 83(2):495 – 530, 1981.

[5] Chris Coey, Lea Kapelevich, and Juan Pablo Vielma. Solving natural conic formulations with Hypatia.jl. *INFORMS Journal on Computing*, 34:2686–2699, 2022.

[6] Inc. CVX Research. CVX: Matlab software for disciplined convex programming, version 2.0. https://cvxr.com/cvx, August 2012.

[7] Lester E. Dubins. On extreme points of convex sets. *Journal of Mathematical Analysis and Applications*, 5(2):237–244, 1962.

[8] Katsuki Fujisawa, Kazuhide Nakata, Makoto Yamashita, and Mituhiro Fukuda. Sdpa project: Solving large-scale semidefinite programs. *Journal of the Operations Research Society of Japan*, 50(4):278–298, dec 2007.

[9] Paul J. Goulart and Yuwen Chen. Clarabel: An interior-point solver for conic programs with quadratic objectives, 2024.

[10] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms I: Fundamentals*. Springer Berlin Heidelberg, 1993.

[11] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods*. Springer, Berlin Heidelberg, 1993.

[12] Mehdi Karimi and Levent Tunçel. Domain-driven solver (DDS) version 2.1: a MATLAB-based software package for convex optimization problems in domain-driven form. *Mathematical Programming Computation*, 16(1):37–92, 2024.

[13] Igor Klep and Markus Schweighofer. An exact duality theory for semidefinite programming based on sums of squares. *Mathematics of Operations Research*, 38(3):569–590, August 2013.

[14] Ying Lin, Scott B. Lindstrom, Bruno F. Lourenço, and Ting Kei Pong. Tight error bounds for log-determinant cones without constraint qualifications. *Journal of Optimization Theory and Applications*, 205(3), 2025.

[15] Ying Lin, Scott B. Lindstrom, Bruno F. Lourenço, and Ting Kei Pong. Generalized power cones: Optimal error bounds and automorphisms. *SIAM Journal on Optimization*, 34(2):1316–1340, 2024.

[16] Scott B. Lindstrom, Bruno F. Lourenço, and Ting Kei Pong. Error bounds, facial residual functions and applications to the exponential cone. *Mathematical Programming*, 200:229–278, 2023.

[17] Scott B. Lindstrom, Bruno F. Lourenço, and Ting Kei Pong. Optimal error bounds in the absence of constraint qualifications with applications to p-cones and beyond. *Mathematics of Operations Research*, 2024.

[18] Minghui Liu and Gábor Pataki. Exact duality in semidefinite programming based on elementary reformulations. *SIAM Journal on Optimization*, 25(3):1441–1454, January 2015.

[19] Minghui Liu and Gábor Pataki. Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming. *Mathematical Programming*, 167(2):435–480, 2018.

[20] Bruno F. Lourenço, Masakazu Muramatsu, and Takashi Tsuchiya. Facial reduction and partial polyhedrality. *SIAM Journal on Optimization*, 28(3):2304–2326, 2018.

[21] Bruno F. Lourenço and Gábor Pataki. A simplified treatment of Ramana's exact dual for semidefinite programming. *Optimization Letters*, 17(2):219–243, October 2022.

[22] Bruno F. Lourenço, Vera Roshchina, and James Saunderson. Amenable cones are particularly nice. *SIAM Journal on Optimization*, 32(3):2347–2375, 2022.

[23] Bruno F. Lourenço. Amenable cones: error bounds without constraint qualifications. *Mathematical Programming*, 186:1–48, 2021.

[24] Miles Lubin, Oscar Dowson, Joaquim Dias Garcia, Joey Huchette, Benoît Legat, and Juan Pablo Vielma. JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, 2023.

[25] MOSEK ApS. *MOSEK Modeling Cookbook Release 3.3.0*, 2025.

[26] Dávid Papp and Sercan Yıldız. Alfonso: Matlab package for nonsymmetric conic optimization. *INFORMS Journal on Computing*, 34(1):11–19, 2022.

[27] Gábor Pataki. The geometry of semidefinite programming. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenberghe, editors, *Handbook of semidefinite programming: theory, algorithms, and applications*. Kluwer Academic Publishers, 2000.

[28] Gábor Pataki. On the closedness of the linear image of a closed convex cone. *Mathematics of Operations Research*, 32(2):395–412, 2007.

[29] Gábor Pataki. On the connection of facially exposed and nice cones. *Journal of Mathematical Analysis and Applications*, 400(1):211–221, 2013.

[30] Gábor Pataki. Strong duality in conic linear programming: Facial reduction and extended duals. In *Computational and Analytical Mathematics*, volume 50, pages 613–634. Springer, New York, 2013.

[31] Gábor Pataki. A combinatorial approach to Ramana's exact dual for semidefinite programming, 2025.

[32] Frank Permenter and Pablo Parrilo. Partial facial reduction: simplified, equivalent sdps via approximations of the psd cone. *Mathematical Programming*, 171(1–2):1–54, June 2017.

[33] Motakuri V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77(1):129–162, April 1997.

[34] Motakuri V. Ramana, Levent Tunçel, and Henry Wolkowicz. Strong duality for semidefinite programming. *SIAM Journal on Optimization*, 7(3):641–662, August 1997.

[35] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[36] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis.* Springer Berlin Heidelberg, 2009.

[37] Vera Roshchina. Facially exposed cones are not always nice. *SIAM Journal on Optimization*, 24(1):257–268, 2014.

[38] Vera Roshchina and Levent Tunçel. Facially dual complete (nice) cones and lexicographic tangents. *SIAM Journal on Optimization*, 29(3):2363–2387, 2019.

[39] Matthew S. Scott. Qualification-free convex analysis via the joint supporting subspace, 2025.

[40] Valeriu Soltan. *Lectures on Convex Sets.* World Scientific, 2015.

[41] Jos F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.

[42] Jos F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, 2000.

[43] K. C. Toh, M. J. Todd, and R. H. Tütüncü. SDPT3 — a matlab software package for semidefinite programming, version 1.3. *Optimization Methods and Software*, 11(1–4):545–581, January 1999.

[44] R. H. Tütüncü, K. C. Toh, and M. J. Todd. Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, 95(2):189–217, 2003.

[45] Hayato Waki and Masakazu Muramatsu. Facial reduction algorithms for conic optimization problems. *Journal of Optimization Theory and Applications*, 158(1):188–215, 2013.

[46] Hayato Waki, Maho Nakata, and Masakazu Muramatsu. Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization. *Computational Optimization and Applications*, 53(3):823–844, September 2011.

[47] Yuzixuan (Melody) Zhu, Gábor Pataki, and Quoc Tran-Dinh. Sieve-SDP: a simple facial reduction algorithm to preprocess semidefinite programs. *Mathematical Programming Computation*, 11(3):503–586, Sep 2019.

A Proofs of Section 2

Proof of Lemma 2.1. (*i*) Let $s \in \mathrm{ri}\,(C)$ and denote the right-hand side set of (2.1) by G. For any $x \in \mathrm{aff}(C)$, by the definition of relative interior, there must exists $t \in (0,1)$ such that $tx + (1-t)s = s + t(x-s) \in C$. Then letting $z = x - s$, we have $x = s + z$, which means $x \in G$ and so $\mathrm{aff}(C) \subseteq G$.

Conversely, suppose that there exist $t \in (0,1)$ and z such that $s + tz \in C$. We then have $tz \in C - s \subseteq \mathrm{span}\,(C-s)$, and hence $z \in \mathrm{span}\,(C-s)$. It follows that $s+z \in s+\mathrm{span}\,(C-s) = \mathrm{aff}(C)$ and so $G \subseteq \mathrm{aff}(C)$.

(*ii*) For any $x \in \mathrm{aff}(C_1 \cap C_2)$, there exists $k \in \mathbb{N}$, $v_i \in C_1 \cap C_2$ for $i = 1, 2, \ldots, k$, and $\{\alpha_i\}_{i=1}^k$ with $\sum_{i=1}^k \alpha_i = 1$ such that $x = \sum_{i=1}^k \alpha_i v_i$. Since $v_i \in C_1 \cap C_2$ for all i, we have $x \in \mathrm{aff}(C_1) \cap \mathrm{aff}(C_2)$ as well.

For the converse, we start by observing that $\mathrm{ri}\,(C_1 \cap C_2) = \mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2)$ holds, since $\mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2) \neq \varnothing$, e.g., see [35, Theorem 6.5]. Let $s \in \mathrm{ri}\,(C_1 \cap C_2) = \mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2)$ be arbitrary. For any $x \in \mathrm{aff}(C_1) \cap \mathrm{aff}(C_2)$, applying (2.1) for C_1 and C_2, respectively, there exist $t_1, t_2 \in (0,1)$ and z such that $x = s + z$, $s + t_1 z \in C_1$ and $s + t_2 z \in C_2$. By the convexity of C_1 and C_2, it holds that $s + \min\{t_1, t_2\} z \in C_1 \cap C_2$. Applying (2.1) for $C_1 \cap C_2$, we see that $x = s + z \in \mathrm{aff}(C_1 \cap C_2)$.

(*iii*) For any $x \in \mathrm{aff}(C_1 - C_2)$, there exists $k \in \mathbb{N}$, $v_i \in C_1$, $u_i \in C_2$ for $i = 1, 2, \ldots, k$ and $\{\alpha_i\}_{i=1}^k$ such that $x = \sum_{i=1}^k \alpha_i (v_i - u_i)$. Then $x = \sum_{i=1}^k \alpha_i v_i - \sum_{i=1}^k \alpha_i u_i \in \mathrm{aff}(C_1) - \mathrm{aff}(C_2)$.

For the converse, as in the previous item, we have $\mathrm{ri}\,(C_1 \cap C_2) = \mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2)$, since $\mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2) \neq \varnothing$. Moreover, from [35, Corollary 6.6.2], it holds that $\mathrm{ri}\,(C_1 - C_2) = \mathrm{ri}\,(C_1) - \mathrm{ri}\,(C_2)$ and hence $0 \in \mathrm{ri}\,(C_1 - C_2)$. Let $s \in \mathrm{ri}\,(C_1 \cap C_2) = \mathrm{ri}\,(C_1) \cap \mathrm{ri}\,(C_2)$ be arbitrary. For any $x \in \mathrm{aff}(C_1) - \mathrm{aff}(C_2)$, applying (2.1) for C_1 and C_2, respectively, there exist $t_1, t_2 \in (0,1)$ and z_1, z_2 such that $x = s + z_1 - s - z_2 = z_1 - z_2$ with $s + t_1 z_1 \in C_1$ and $s + t_2 z_2 \in C_2$. By the convexity of C_1 and C_2, it holds that $s + \min\{t_1, t_2\} z_1 \in C_1$ and $s + \min\{t_1, t_2\} z_2 \in C_2$. Then $s + \min\{t_1, t_2\} z_1 - s - \min\{t_1, t_2\} z_2 = 0 + \min\{t_1, t_2\}(z_1 - z_2) = 0 + \min\{t_1, t_2\} x \in C_1 - C_2$. Since $0 \in \mathrm{ri}\,(C_1 - C_2)$, then by (2.1), we obtain $x \in \mathrm{aff}(C_1 - C_2)$. $\square$

Proof of Lemma 2.2. For any $y \in C_2^* \cap C_1^\circ$, it holds that $\langle x_1, y \rangle \leqslant 0$ and $\langle x_2, y \rangle \geqslant 0$ for any $x_1 \in C_1$ and $x_2 \in C_2$, which follows that $\langle x_2 - x_1, y \rangle \geqslant 0$ for any $x_1 \in C_1$ and $x_2 \in C_2$ and hence $y \in (C_2 - C_1)^*$.

Conversely, since $0 \in C_1$, we have $C_2 \subseteq C_2 - C_1$ and hence $(C_2 - C_1)^* \subseteq C_2^*$. Moreover, for any $c \in C_1 \cap C_2$, since C_1 is a closed convex cone, $x + c \in C_1$ for any $x \in C_1$. Then $-x = c - (x + c) \in C_2 - C_1$, i.e., $-C_1 \subseteq C_2 - C_1$, which implies that $(C_2 - C_1)^* \subseteq C_1^\circ$. Therefore, we have $(C_2 - C_1)^* \subseteq C_2^* \cap C_1^\circ$ and consequently $(C_2 - C_1)^* = C_2^* \cap C_1^\circ$. $\square$

Proof of Lemma 2.5. (*i*) Since $\operatorname{cl} f_i \leqslant f_i$ for any $i \in \mathcal{I}$, we have $\inf_{i\in\mathcal{I}} \operatorname{cl} f_i \leqslant \inf_{i\in\mathcal{I}} f_i$. Taking the closure on both sides and noting that the closure operation is order-preserving, we have

$$\operatorname{cl} \inf_{i\in\mathcal{I}} \operatorname{cl} f_i \leqslant \operatorname{cl} \inf_{i\in\mathcal{I}} f_i.$$

Conversely, for any $x \in \mathcal{E}$, we have

$$\operatorname{cl} \inf_{i\in\mathcal{I}} f_i(x) \overset{\text{(a)}}{=} \liminf_{y\to x} \inf_{i\in\mathcal{I}} f_i(y) \overset{\text{(b)}}{\leqslant} \liminf_{y\to x} f_j(y) \text{ for all } j \in \mathcal{I}, \tag{A.1}$$

where (a) comes from the definition of the closure of a function; (b) holds since the limit inferior is order-preserving and $\inf_{i\in\mathcal{I}} f_i \leqslant f_j$ for all $j \in \mathcal{I}$. Furthermore, for any $j \in \mathcal{I}$, $\operatorname{cl} f_j(x) = \liminf_{y\to x} f_j(y)$, which together with (A.1) imply that $\operatorname{cl}\inf_{i\in\mathcal{I}} f_i(x) \leqslant \operatorname{cl} f_j(x)$ for all $j \in \mathcal{I}$. Therefore, we conclude that

$$\operatorname{cl} \inf_{i\in\mathcal{I}} f_i(x) \leqslant \inf_{i\in\mathcal{I}} \operatorname{cl} f_i(x).$$

Since $\operatorname{cl}\inf_{i\in\mathcal{I}} f_i(x)$ is lower-semicontinuous (by definition) and the closure operation is order-preserving, we obtain

$$\operatorname{cl} \inf_{i\in\mathcal{I}} f_i(x) \leqslant \operatorname{cl} \inf_{i\in\mathcal{I}} \operatorname{cl} f_i(x).$$

(*ii*) For any $i \in \mathcal{I}$, we have $C_i \subseteq C$, which implies that

$$\inf_{x\in C_i} f(x) \geqslant \inf_{x\in C} f(x),$$

and hence

$$\inf_{i\in\mathcal{I}} \inf_{x\in C_i} f(x) \geqslant \inf_{x\in C} f(x).$$

Since f is proper, its epigraph is nonempty, so $\inf_{x\in C} f(x) < +\infty$. Next, we divide in two cases. First suppose that $\inf_{x\in C} f(x)$ is finite. Then, for any $\epsilon > 0$, there exists $\overline{x} \in C$ such that

$$f(\overline{x}) < \inf_{x\in C} f(x) + \epsilon.$$

Since $\bigcup_{i\in\mathcal{I}} C_i = C$, then $\overline{x} \in C_j$ for some $j \in \mathcal{I}$. Therefore,

$$\inf_{x\in C_j} f(x) \leqslant f(\overline{x}).$$

It follows that

$$\inf_{i\in\mathcal{I}} \inf_{x\in C_i} f(x) \leqslant \inf_{x\in C_j} f(x) \leqslant f(\overline{x}) < \inf_{x\in C} f(x) + \epsilon.$$

As this holds for any $\epsilon > 0$, we must have

$$\inf_{i\in\mathcal{I}} \inf_{x\in C_i} f(x) \leqslant \inf_{x\in C} f(x).$$

The second case is when $\inf_{x\in C} f(x) = -\infty$. Then, we let $L < 0$ be arbitrary and take $\bar{x}$ such that $f(\bar{x}) < L$ holds. Following a similar sequence of arguments, we conclude that

$$\inf_{i\in\mathcal{I}} \inf_{x\in C_i} f(x) \leqslant \inf_{x\in C_j} f(x) \leqslant f(\overline{x}) < L$$

Therefore, as L is arbitrary, $\inf_{i\in\mathcal{I}} \inf_{x\in C_i} f(x) \leqslant \inf_{x\in C} f(x)$ holds in this case as well. $\square$

Proof of Lemma 2.6. Notice that

$$\sigma_{\mathrm{aff}(C)}(x) = \sigma_{\mathcal{L}_C+a}(x) = \sup_{y\in\mathcal{L}_C} \langle x,\, y+a\rangle = \sup_{y\in\mathcal{L}_C} \langle x,\, y\rangle + \langle x,\, a\rangle = \sigma_{\mathcal{L}_C}(x) + \langle x,\, a\rangle.$$

Because $\mathcal{L}_C$ is a subspace, then $\operatorname{dom}\sigma_{\mathcal{L}_C} = \mathcal{L}_C^{\perp}$ and $\sigma_{\mathcal{L}_C}(x) = 0$ for any $x \in \mathcal{L}_C^{\perp}$. Therefore, $\sigma_{\mathrm{aff}(C)} = \langle\cdot,\, a\rangle + \delta_{\mathcal{L}_C^{\perp}}$. Using the formula of $\sigma_{\mathrm{aff}(C)}$, we can obtain that

$$x \in (\mathrm{aff}(C))^\circ \;\Leftrightarrow\; \sigma_{\mathrm{aff}(C)}(x) \leqslant 0 \;\Leftrightarrow\; x \in \mathcal{L}_C^{\perp} \text{ and } \langle x,\, a\rangle \leqslant 0 \;\Leftrightarrow\; x \in \mathcal{L}_C^{\perp} \cap \{a\}^\circ.$$

Then it holds that $(\mathrm{aff}(C))^\circ = \mathcal{L}_C^{\perp} \cap \{a\}^\circ$. $\square$

Proof of Lemma 2.7. Notice that

$$\sup_{x\in C_1} \langle s,\, x\rangle \geqslant \sup_{x\in C_1\cap C_2} \langle s,\, x\rangle \geqslant \inf_{x\in C_1\cap C_2} \langle s,\, x\rangle \geqslant \inf_{x\in C_2} \langle s,\, x\rangle,$$

where the first and last inequalities come from the facts that $\varnothing \neq C_1 \cap C_2 \subseteq C_1$ and $\varnothing \neq C_1 \cap C_2 \subseteq C_2$. This and the assumption $\inf_{x\in C_1}\langle s,\, x\rangle \geqslant \theta \geqslant \sup_{x\in C_2}\langle s,\, x\rangle$ imply that all inequalities hold as equalities, leading to the desired statement. Since $C_1 \cap C_2 \neq \varnothing$, this also implies that $s \in \operatorname{dom}\sigma_{C_1}$, $s \in -\operatorname{dom}\sigma_{C_2}$, and $s \in \operatorname{dom}\sigma_{C_1\cap C_2}$. Moreover, $\theta = \sup_{x\in C_1\cap C_2}\langle s,\, x\rangle = \inf_{x\in C_1\cap C_2}\langle s,\, x\rangle$ implies that $\langle s,\, x\rangle = \theta$ for all $x \in C_1 \cap C_2$, hence we have $C_1 \cap C_2 \subseteq H := \{z \mid \langle s,\, z\rangle = \theta\}$. $\square$

Proof of Lemma 2.8. We first prove $C \cap H_1 \cap H_2 = C \cap H$.

($\subseteq$) If $C \cap H_1 \cap H_2 = \varnothing$, then $C \cap H_1 \cap H_2 \subseteq C \cap H$. Next we assume $C \cap H_1 \cap H_2 \neq \varnothing$, then we can pick $z \in C \cap H_1 \cap H_2$. By definition, $z \in C$, $\langle s_1,\, z\rangle = \sigma_C(s_1)$ and $\langle s_2,\, z\rangle = \sigma_C(s_2)$. Then, we have $\langle s_1 + s_2,\, z\rangle = \langle s_1,\, z\rangle + \langle s_2,\, z\rangle = \sigma_C(s_1) + \sigma_C(s_2)$. Thus, $z \in H$, and since $z \in C$, then $z \in C \cap H$ and hence $C \cap H_1 \cap H_2 \subseteq C \cap H$.

($\supseteq$) Conversely, similar to previous, we can assume $C \cap H \neq \varnothing$, then we pick $z \in C \cap H$. By definition, $z \in C$ and $\langle s_1 + s_2,\, z\rangle = \sigma_C(s_1) + \sigma_C(s_2)$. By the definition of support functions, we have $\langle s_1,\, z\rangle \leqslant \sigma_C(s_1)$ and $\langle s_2,\, z\rangle \leqslant \sigma_C(s_2)$. This gives $\langle s_1,\, z\rangle + \langle s_2,\, z\rangle \leqslant \sigma_C(s_1) + \sigma_C(s_2)$. From $\langle s_1 + s_2,\, z\rangle = \sigma_C(s_1) + \sigma_C(s_2)$, equality must hold in both support function inequalities, so $\langle s_1,\, z\rangle = \sigma_C(s_1)$ and $\langle s_2,\, z\rangle = \sigma_C(s_2)$. Thus, $z \in H_1 \cap H_2$. Since $z \in C$, it follows that $C \cap H \subseteq C \cap H_1 \cap H_2$.

Next, we prove $C \cap H_1 \cap H_2 = C \cap \widetilde{H}$. Since $C \cap H_1 \cap H_2 = C \cap H$, we only need to show that $C \cap H = C \cap \widetilde{H}$. Note that $C \cap H = C \cap H_1 \cap H_2 \neq \varnothing$. Pick any $z \in C \cap H$. We then have $z \in H$ and see from the subadditivity of a support function that

$$\langle s_1 + s_2,\, z\rangle = \sigma_C(s_1) + \sigma_C(s_2) \geqslant \sigma_C(s_1 + s_2).$$

This together with $\langle s_1 + s_2, z\rangle \leqslant \sigma_C(s_1 + s_2)$ yields that $z \in \widetilde{H}$ and therefore $C \cap H \subseteq C \cap \widetilde{H}$.

Since $C \cap H \neq \varnothing$, we see from $C \cap H \subseteq C \cap \widetilde{H}$ that there exists some $c \in C \cap H \subseteq C \cap \widetilde{H}$. Consequently, $c \in H \cap \widetilde{H}$ and therefore

$$\sigma_C(s_1 + s_2) = \langle s_1 + s_2, c\rangle = \sigma_C(s_1) + \sigma_C(s_2),$$

which implies that $H = \widetilde{H}$ and thus $C \cap H = C \cap \widetilde{H}$. □