



A Suitable Conjugacy for the l_0 Pseudonorm

Jean-Philippe Chancelier, Michel de Lara

► To cite this version:

Jean-Philippe Chancelier, Michel de Lara. A Suitable Conjugacy for the l_0 Pseudonorm. 2019. hal-02013977

HAL Id: hal-02013977

<https://hal.science/hal-02013977>

Preprint submitted on 11 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Suitable Conjugacy for the l_0 Pseudonorm

Jean-Philippe Chancelier and Michel De Lara,
CERMICS, École des Ponts ParisTech

February 11, 2019

Abstract

The so-called l_0 pseudonorm on $\mathbb{R}^d$ counts the number of nonzero components of a vector. It is well-known that the l_0 pseudonorm is not convex, as its Fenchel biconjugate is zero. In this paper, we introduce a suitable conjugacy, induced by a novel coupling, Caprac, having the property of being constant along primal rays, like the l_0 pseudonorm. The Caprac coupling belongs to the class of one-sided linear couplings, that we introduce. We show that they induce conjugacies that share nice properties with the classic Fenchel conjugacy. For the Caprac conjugacy, induced by the coupling Caprac, we prove that the l_0 pseudonorm is equal to its biconjugate: hence, the l_0 pseudonorm is Caprac-convex in the sense of generalized convexity. As a corollary, we show that the l_0 pseudonorm coincides, on the sphere, with a convex lsc function. We also provide expressions for conjugates in terms of two families of dual norms, the 2 - k -symmetric gauge norms and the k -support norms.

Key words: l_0 pseudonorm, coupling, Fenchel-Moreau conjugacy, 2 - k -symmetric gauge norms, k -support norms.

1 Introduction

The *counting function*, also called *cardinality function* or *l_0 pseudonorm*, counts the number of nonzero components of a vector in $\mathbb{R}^d$. It is related to the rank function defined over matrices [5]. It is well-known that the l_0 pseudonorm is lower semi continuous but is not convex. This can be deduced from the computation of its Fenchel biconjugate, which is zero.

In this paper, we display a suitable conjugacy for which we show that the l_0 pseudonorm is “convex” in the sense of generalized convexity (equal to its biconjugate). As a corollary, we also show that the l_0 pseudonorm coincides, on the sphere, with a convex lsc function.

The paper is organized as follows. In Sect. 2, we provide background on Fenchel-Moreau conjugacies, then introduce a novel class of *one-sided linear couplings*, which includes the *constant along primal rays coupling* $\dot{\circ}$ (*Caprac*). We show that one-sided linear couplings induce conjugacies that share nice properties with the classic Fenchel conjugacy, by giving

expressions for conjugate and biconjugate functions. We elucidate the structure of Caprac-convex functions. Then, in Sect. 3, we relate the Caprac conjugate and biconjugate of the l_0 pseudonorm, the characteristic functions of its level sets and the symmetric gauge norms. In particular, we show that the l_0 pseudonorm is Caprac biconjugate (a Caprac-convex function), from which we deduce that it coincides, on the sphere, with a convex lsc function. The Appendix A gathers background on J. J. Moreau lower and upper additions, properties of 2- k -symmetric gauge norms, and properties of the l_0 pseudonorm level sets.

2 The constant along primal rays coupling (Caprac)

After having recalled background on Fenchel-Moreau conjugacies in §2.1, we introduce *one-sided linear couplings* in §2.2, and finally the *constant along primal rays coupling* $\dot{c}$ (Caprac) in §2.3.

2.1 Background on Fenchel-Moreau conjugacies

We review general concepts and notations, then we focus on the special case of the Fenchel conjugacy. We denote $\overline{\mathbb{R}} = [-\infty, +\infty]$. Background on J. J. Moreau lower and upper additions can be found in §A.1.

The general case

Let be given two sets $\mathbb{X}$ (“primal”), $\mathbb{Y}$ (“dual”), together with a *coupling* function

$$c : \mathbb{X} \times \mathbb{Y} \rightarrow \overline{\mathbb{R}}. \quad (1)$$

With any coupling, we associate *conjugacies* from $\overline{\mathbb{R}}^{\mathbb{X}}$ to $\overline{\mathbb{R}}^{\mathbb{Y}}$ and from $\overline{\mathbb{R}}^{\mathbb{Y}}$ to $\overline{\mathbb{R}}^{\mathbb{X}}$ as follows.

Definition 1 *The c -Fenchel-Moreau conjugate of a function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, with respect to the coupling c , is the function $f^c : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$ defined by*

$$f^c(y) = \sup_{x \in \mathbb{X}} \left(c(x, y) \dot{+} (-f(x)) \right), \quad \forall y \in \mathbb{Y}. \quad (2)$$

With the coupling c , we associate the reverse coupling c' defined by

$$c' : \mathbb{Y} \times \mathbb{X} \rightarrow \overline{\mathbb{R}}, \quad c'(y, x) = c(x, y), \quad \forall (y, x) \in \mathbb{Y} \times \mathbb{X}. \quad (3)$$

The c' -Fenchel-Moreau conjugate of a function $g : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$, with respect to the coupling c' , is the function $g^{c'} : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ defined by

$$g^{c'}(x) = \sup_{y \in \mathbb{Y}} \left(c(x, y) \dot{+} (-g(y)) \right), \quad \forall x \in \mathbb{X}. \quad (4)$$

The c -Fenchel-Moreau biconjugate of a function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, with respect to the coupling c , is the function $f^{cc'} : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ defined by

$$f^{cc'}(x) = (f^c)^{c'}(x) = \sup_{y \in \mathbb{Y}} \left(c(x, y) \dot{+} (-f^c(y)) \right), \quad \forall x \in \mathbb{X}. \quad (5)$$

For any coupling c ,

- the biconjugate of a function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ satisfies

$$f^{cc'}(x) \leq f(x), \quad \forall x \in \mathbb{X}, \quad (6a)$$

- for any couple of functions $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ and $h : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, we have the inequality

$$\sup_{y \in \mathbb{Y}} \left((-f^c(y)) \dot{+} (-h^{-c}(y)) \right) \leq \inf_{x \in \mathbb{X}} \left(f(x) \dot{+} h(x) \right), \quad (6b)$$

where the $(-c)$ -Fenchel-Moreau conjugate is given by

$$h^{-c}(y) = \sup_{x \in \mathbb{X}} \left((-c(x, y)) \dot{+} (-h(x)) \right), \quad \forall y \in \mathbb{Y}, \quad (6c)$$

- for any function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ and subset $X \subset \mathbb{X}$, we have the inequality

$$\sup_{y \in \mathbb{Y}} \left((-f^c(y)) \dot{+} (-\delta_X^{-c}(y)) \right) \leq \inf_{x \in \mathbb{X}} \left(f(x) \dot{+} \delta_X(x) \right) = \inf_{x \in X} f(x). \quad (6d)$$

The Fenchel conjugacy

When the sets $\mathbb{X}$ and $\mathbb{Y}$ are vector spaces equipped with a bilinear form $\langle \cdot, \cdot \rangle$, the corresponding conjugacy is the classical *Fenchel conjugacy*. For any functions $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ and $g : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$, we denote

$$f^*(y) = \sup_{x \in \mathbb{X}} \left(\langle x, y \rangle \dot{+} (-f(x)) \right), \quad \forall y \in \mathbb{Y}, \quad (7a)$$

$$g^*(x) = \sup_{y \in \mathbb{Y}} \left(\langle x, y \rangle \dot{+} (-g(y)) \right), \quad \forall x \in \mathbb{X} \quad (7b)$$

$$f^{**}(x) = \sup_{y \in \mathbb{Y}} \left(\langle x, y \rangle \dot{+} (-f^*(y)) \right), \quad \forall x \in \mathbb{X}. \quad (7c)$$

Due to the presence of the coupling $(-c)$ in the Inequality (6b), we also introduce¹

$$f^{(-*)}(y) = \sup_{x \in \mathbb{X}} \left(-\langle x, y \rangle \dot{+} (-f(x)) \right) = f^*(-y), \quad \forall y \in \mathbb{Y}, \quad (8a)$$

$$g^{(-*)'}(x) = \sup_{y \in \mathbb{Y}} \left(-\langle x, y \rangle \dot{+} (-g(y)) \right) = g^*(-x), \quad \forall x \in \mathbb{X} \quad (8b)$$

$$f^{(-*)^{(-*)'}}(x) = \sup_{y \in \mathbb{Y}} \left(-\langle x, y \rangle \dot{+} (-f^{(-*)}(y)) \right) = f^{**}(x), \quad \forall x \in \mathbb{X}. \quad (8c)$$

When the two vector spaces $\mathbb{X}$ and $\mathbb{Y}$ are *paired* in the sense of convex analysis², Fenchel conjugates are convex *lower semi continuous (lsc)* functions, and their opposites are concave *upper semi continuous (usc)* functions.

¹In convex analysis, one does not use the notations below, but rather uses $f^\vee(x) = f(-x)$, for all $x \in \mathbb{X}$, and $g^\vee(y) = g(-y)$, for all $y \in \mathbb{Y}$. The connection between both notations is given by $f^{(-*)} = (f^\vee)^* = (f^*)^\vee$.

²That is, $\mathbb{X}$ and $\mathbb{Y}$ are equipped with a bilinear form $\langle \cdot, \cdot \rangle$, and locally convex topologies that are compatible in the sense that the continuous linear forms on $\mathbb{X}$ are the functions $x \in \mathbb{X} \mapsto \langle x, y \rangle$, for all $y \in \mathbb{Y}$, and that the continuous linear forms on $\mathbb{Y}$ are the functions $y \in \mathbb{Y} \mapsto \langle x, y \rangle$, for all $x \in \mathbb{X}$.

2.2 One-sided linear couplings

Let $\mathbb{W}$ and $\mathbb{X}$ be any two sets and $\theta : \mathbb{W} \rightarrow \mathbb{X}$ be a mapping. We recall the definition [2, p. 214] of the *infimal postcomposition* $(\theta \triangleright h) : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ of a function $h : \mathbb{W} \rightarrow \overline{\mathbb{R}}$:

$$(\theta \triangleright h)(x) = \inf \{h(w) \mid w \in \mathbb{W}, \theta(w) = x\}, \quad \forall x \in \mathbb{X}, \quad (9)$$

with the convention that $\inf \emptyset = +\infty$ (and with the consequence that $\theta : \mathbb{W} \rightarrow \mathbb{X}$ need not be defined on all $\mathbb{W}$, but only on $\text{dom} h$). The infimal postcomposition has the following *invariance property*

$$h = f \circ \theta \text{ where } f : \mathbb{X} \rightarrow \overline{\mathbb{R}} \Rightarrow \theta \triangleright h = f \dot{+} \delta_{\theta(\mathbb{W})}, \quad (10)$$

where δ_Z denotes the *characteristic function* of a set Z :

$$\delta_Z(z) = \begin{cases} 0 & \text{if } z \in Z, \\ +\infty & \text{if } z \notin Z. \end{cases} \quad (11)$$

Definition 2 Let $\mathbb{X}$ and $\mathbb{Y}$ be two vector spaces equipped with a bilinear form $\langle \cdot, \cdot \rangle$. Let $\mathbb{W}$ be a set and $\theta : \mathbb{W} \rightarrow \mathbb{X}$ a mapping. We define the *one-sided linear coupling* c_θ between $\mathbb{W}$ and $\mathbb{Y}$ by

$$c_\theta : \mathbb{W} \times \mathbb{Y} \rightarrow \overline{\mathbb{R}}, \quad c_\theta(w, y) = \langle \theta(w), y \rangle, \quad \forall w \in \mathbb{W}, \quad \forall y \in \mathbb{Y}. \quad (12)$$

Here are expressions for the conjugates and biconjugates of a function.

Proposition 3 For any function $g : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$, the c'_θ -Fenchel-Moreau conjugate is given by

$$g^{c'_\theta} = g^* \circ \theta. \quad (13)$$

For any function $h : \mathbb{W} \rightarrow \overline{\mathbb{R}}$, the c_θ -Fenchel-Moreau conjugate is given by

$$h^{c_\theta} = (\theta \triangleright h)^*, \quad (14)$$

and the c_θ -Fenchel-Moreau biconjugate is given by

$$h^{c_\theta c_\theta'} = (h^{c_\theta})^* \circ \theta = (\theta \triangleright h)^{\star\star'} \circ \theta. \quad (15)$$

For any subset $W \subset \mathbb{W}$, the $(-c_\theta)$ -Fenchel-Moreau conjugate of the characteristic function of W is given by

$$\delta_W^{-c_\theta} = \sigma_{-\theta(W)}. \quad (16)$$

We recall that, in convex analysis, $\sigma_X : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$ denotes the *support function* of a subset $X \subset \mathbb{X}$:

$$\sigma_X(y) = \sup_{x \in X} \langle x, y \rangle, \quad \forall y \in \mathbb{Y}. \quad (17)$$

Proof. We prove (13). Letting $w \in \mathbb{W}$, we have that

$$\begin{aligned} (g)^{c'_\theta}(w) &= \sup_{y \in \mathbb{Y}} \left(\langle \theta(w), y \rangle \dot{+} (-g(y)) \right) \quad (\text{by the conjugate formula (2) and the coupling (12)}) \\ &= g^*(\theta(w)) . \quad (\text{by the expression (7a) of the Fenchel conjugate}) \end{aligned}$$

We prove (14). Letting $y \in \mathbb{Y}$, we have that

$$\begin{aligned} h^{c_\theta}(y) &= \sup_{w \in \mathbb{X}} \left(\langle \theta(w), y \rangle \dot{+} (-h(w)) \right) \quad (\text{by the conjugate formula (2) and the coupling (12)}) \\ &= \sup_{x \in \mathbb{X}} \sup_{w \in \mathbb{X}, \theta(w)=x} \left(\langle \theta(w), y \rangle \dot{+} (-h(w)) \right) \\ &= \sup_{x \in \mathbb{X}} \left(\langle x, y \rangle \dot{+} \sup_{w \in \mathbb{X}, \theta(w)=x} (-h(w)) \right) \quad (\text{by (36e)}) \\ &= \sup_{x \in \mathbb{X}} \left(\langle x, y \rangle \dot{+} \left(- \inf_{w \in \mathbb{X}, \theta(w)=x} h(w) \right) \right) \\ &= \sup_{x \in \mathbb{X}} \left(\langle x, y \rangle \dot{+} \left(-(\theta \triangleright h)(x) \right) \right) \quad (\text{by the infimal postcomposition expression (9)}) \\ &= (\theta \triangleright h)^*(y) \quad (\text{by the expression (7a) of the Fenchel conjugate}) \end{aligned}$$

We prove (15). Letting $x \in \mathbb{X}$, we have that

$$\begin{aligned} h^{c_\theta c_{\theta'}}(x) &= (h^{c_\theta})^{c'_{\theta'}}(x) \quad (\text{by the definition (5) of the biconjugate}) \\ &= ((\theta \triangleright h)^*)^{c'_{\theta'}}(x) \quad (\text{by (14)}) \\ &= \sup_{y \in \mathbb{Y}} \left(\langle \theta(x), y \rangle \dot{+} (-((\theta \triangleright h)^*(y))) \right) \\ &\quad (\text{by the conjugate formula (2) and the coupling (12)}) \\ &= (\theta \triangleright h)^{\star\star'}(\theta(x)) \quad (\text{by the expression (7a) of the Fenchel conjugate}) \end{aligned}$$

We prove (16):

$$\begin{aligned} \delta_W^{-c_\theta} &= \delta_W^{c_{(-\theta)}} \quad (\text{because } -c_\theta = c_{(-\theta)} \text{ by (12)}) \\ &= ((-\theta) \triangleright \delta_W)^* \quad (\text{by (14)}) \\ &= \delta_{-\theta(W)}^* \quad (\text{because } \theta \triangleright \delta_W = \delta_{\theta(W)} \text{ by (9)}) \\ &= \sigma_{-\theta(W)} . \quad (\text{as is well-known in convex analysis}) \end{aligned}$$

This ends the proof. $\square$

2.3 Constant along primal rays coupling (Caprac)

Now, we introduce a novel coupling, which is a special case of one-sided linear couplings.

Definition 4 Let $\mathbb{X}$ and $\mathbb{Y}$ be two vector spaces equipped with a bilinear form $\langle \cdot, \cdot \rangle$, and suppose that $\mathbb{X}$ is equipped with a norm $||| \cdot |||$. We define the Caprac coupling $\dot{\varsigma}$ between $\mathbb{X}$ and $\mathbb{Y}$ by

$$\forall y \in \mathbb{Y}, \quad \begin{cases} \dot{\varsigma}(x, y) = \frac{\langle x, y \rangle}{|||x|||}, & \forall x \in \mathbb{X} \setminus \{0\} \\ \dot{\varsigma}(0, y) = 0. \end{cases} \quad (18)$$

We stress the point that, in (18), the bilinear form term $\langle x, y \rangle$ and the norm term $|||x|||$ need not be related. Indeed, the bilinear form $\langle \cdot, \cdot \rangle$ is not necessarily a scalar product and the norm $||| \cdot |||$ is not necessarily induced by this latter.

The Caprac coupling has the property of being *constant along primal rays*, hence the acronym Caprac. We introduce the unit sphere $S_{|||\cdot|||}$ of the normed space $(\mathbb{X}, ||| \cdot |||)$, and the primal *normalization mapping* n

$$S_{|||\cdot|||} = \{x \in \mathbb{X} \mid |||x||| = 1\} \quad \text{and} \quad n : \mathbb{X} \rightarrow S_{|||\cdot|||} \cup \{0\}, \quad n(x) = \begin{cases} \frac{x}{|||x|||} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases} \quad (19)$$

We immediatly obtain that, for all subset $D \subset \mathbb{R}^d$ that contains zero ($0 \in D$):

$$n^{-1}(D) = n^{-1}((\{0\} \cup S_{|||\cdot|||}) \cap D) = \{0\} \cup n^{-1}(S_{|||\cdot|||} \cap D). \quad (20)$$

With these notations, the Caprac coupling (18) is a special case of one-sided linear coupling c_n , as in (12) with $\theta = n$, the Fenchel coupling after primal normalization:

$$\dot{\varsigma}(x, y) = c_n(x, y) = \langle n(x), y \rangle, \quad \forall x \in \mathbb{X}, \quad \forall y \in \mathbb{Y}. \quad (21)$$

Here are expressions for the Caprac-conjugates and biconjugates of a function. The following Proposition simply is Proposition 3 in the case where the mapping θ is the normalization mapping n in (19).

Proposition 5 Let $\mathbb{X}$ and $\mathbb{Y}$ be two vector spaces equipped with a bilinear form $\langle \cdot, \cdot \rangle$, and suppose that $\mathbb{X}$ is equipped with a norm $||| \cdot |||$.

For any function $g : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$, the $\dot{\varsigma}'$ -Fenchel-Moreau conjugate is given by

$$g^{\dot{\varsigma}'} = g^* \circ n. \quad (22)$$

For any function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, the $\dot{\varsigma}$ -Fenchel-Moreau conjugate is given by

$$f^{\dot{\varsigma}} = (n \triangleright f)^*, \quad (23)$$

where the infimal postcomposition (9) has the expression

$$(n \triangleright f)(x) = \inf \{f(x') \mid n(x') = x\} = \begin{cases} \inf_{\lambda > 0} f(\lambda x) & \text{if } x \in S_{|||\cdot|||} \cup \{0\} \\ +\infty & \text{if } x \notin S_{|||\cdot|||} \cup \{0\} \end{cases} \quad (24)$$

and the $\dot{\varsigma}$ -Fenchel-Moreau biconjugate is given by

$$f^{\dot{\varsigma}\dot{\varsigma}'} = (f^{\dot{\varsigma}})^* \circ n = (n \triangleright f)^{\star\star'} \circ n. \quad (25)$$

We recall that so-called Caprac $\dot{\varsigma}$ -convex functions are all functions $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ of the form $(g)^{\dot{\varsigma}'}$, for any $g \in \overline{\mathbb{R}}^{\mathbb{Y}}$, or, equivalently, all functions of the form $f^{\dot{\varsigma}\dot{\varsigma}'}$, for any $f \in \overline{\mathbb{R}}^{\mathbb{X}}$, or, equivalently, all functions that are equal to their $\dot{\varsigma}$ -biconjugate ($f^{\dot{\varsigma}\dot{\varsigma}'} = f$) [10, 9, 6]. From the expressions (22), (23) and (25), we easily deduce the following result.

Corollary 6 *When $\mathbb{X}$ and $\mathbb{Y}$ are two paired vector spaces, and $\mathbb{X}$ is equipped with a norm $|||\cdot|||$, the $\dot{\varsigma}$ -Fenchel-Moreau conjugate $f^{\dot{\varsigma}}$ is a convex lower semi continuous (lsc) function on $\mathbb{Y}$. In addition, using (22), a function is $\dot{\varsigma}$ -convex if and only if it is the composition of a convex lower semi continuous function on $\mathbb{X}$ with the normalization mapping (19).*

3 Caprac conjugates and biconjugates related to the l_0 pseudonorm

In this Section, we work on the Euclidian space $\mathbb{R}^d$ (with $d \in \mathbb{N}^*$), equipped with the scalar product $\langle \cdot, \cdot \rangle$ and with the Euclidian norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. In particular, we consider the *Euclidian unit sphere*

$$S = \{x \in \mathbb{X} \mid \|x\| = 1\}, \quad (26)$$

and the (*Euclidian*) *coupling Caprac* $\dot{\varsigma}$ between $\mathbb{R}^d$ and $\mathbb{R}^d$ by

$$\forall y \in \mathbb{R}^d, \quad \begin{cases} \dot{\varsigma}(x, y) &= \frac{\langle x, y \rangle}{\|x\|}, \quad \forall x \in \mathbb{R}^d \setminus \{0\}, \\ \dot{\varsigma}(0, y) &= 0. \end{cases} \quad (27)$$

The so-called l_0 *pseudonorm* is the function $\ell_0 : \mathbb{R}^d \rightarrow \{0, 1, \dots, d\}$ defined, for any $x \in \mathbb{R}^d$, by

$$\ell_0(x) = |x|_0 = \text{number of nonzero components of } x. \quad (28)$$

The l_0 pseudonorm displays the invariance property

$$\ell_0 \circ n = \ell_0 \quad (29)$$

with respect to the normalization mapping (19). This property will be instrumental to show that the l_0 pseudonorm is a Caprac $\dot{\varsigma}$ -convex function. For this purpose, we will start by introducing two dual norms.

For any $x \in \mathbb{R}^d$ and $K \subset \{1, \dots, d\}$, we denote by $x_K \in \mathbb{R}^d$ the vector which coincides with x , except for the components outside of K that vanish: x_K is the orthogonal projection of x onto the subspace $\mathbb{R}^K \times \{0\}^{-K} \subset \mathbb{R}^d$. Here, following notation from Game Theory, we have denoted by $-K$ the complementary subset of K in $\{1, \dots, d\}$: $K \cup (-K) = \{1, \dots, d\}$ and $K \cap (-K) = \emptyset$. In what follows, $|K|$ denotes the cardinal of the set K and the notation $\sup_{|K| \leq k}$ is a shorthand for $\sup_{K \subset \{1, \dots, d\}, |K| \leq k}$ (the same holds for $\sup_{|K|=k}$).

Definition 7 Let $x \in \mathbb{R}^d$. For $k \in \{1, \dots, d\}$, we denote by $\|x\|_{(k)}^{\text{sgn}}$ the maximum of $\|x_K\|$ over all subsets $K \subset \{1, \dots, d\}$ with cardinal (less than) k :

$$\|x\|_{(k)}^{\text{sgn}} = \sup_{|K| \leq k} \|x_K\| = \sup_{|K|=k} \|x_K\|. \quad (30)$$

Thus defined, $\|\cdot\|_{(k)}^{\text{sgn}}$ is a norm, the 2- k -symmetric gauge norm, or Ky Fan vector norm. Its dual norm (see Definition 10) is called k -support norm [1], denoted by $\|\cdot\|_{(k)}^{\text{sn}}$:

$$\|\cdot\|_{(k)}^{\text{sn}} = \left(\|\cdot\|_{(k)}^{\text{sgn}} \right)_\star. \quad (31)$$

The property that $\sup_{|K| \leq k} \|x_K\| = \sup_{|K|=k} \|x_K\|$ in (30) comes from the easy observation that $K \subset K' \Rightarrow \|x_K\| \leq \|x_{K'}\|$.

The l_0 pseudonorm is used in exact sparse optimization problems of the form $\inf_{|x|_0 \leq k} f(x)$. This is why we introduce the *level sets*

$$\ell_0^{\leq k} = \{x \in \mathbb{R}^d \mid \ell_0(x) \leq k\}, \quad \forall k \in \{0, 1, \dots, d\}, \quad (32a)$$

and the *level curves*

$$\ell_0^=k = \{x \in \mathbb{R}^d \mid \ell_0(x) = k\}, \quad \forall k \in \{0, 1, \dots, d\}. \quad (32b)$$

The l_0 pseudonorm in (28), the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets and the symmetric gauge norms in (30) are related by the following conjugate formulas. The proof relies on results gathered in the Appendix A.

Theorem 8 Let $\dot{\cdot}$ be the Euclidian coupling Caprac (27). Let $k \in \{0, 1, \dots, d\}$. We have that:

$$\delta_{\ell_0^{\leq k}}^{-\dot{\cdot}} = \delta_{\ell_0^{\leq k}}^{\dot{\cdot}} = \|\cdot\|_{(k)}^{\text{sgn}}, \quad (33a)$$

$$\delta_{\ell_0^{\leq k}}^{\dot{\cdot}\dot{\cdot}'} = \delta_{\ell_0^{\leq k}}, \quad (33b)$$

$$\ell_0^{\dot{\cdot}} = \sup_{l=0,1,\dots,d} \left[\|\cdot\|_{(l)}^{\text{sgn}} - l \right], \quad (33c)$$

$$\ell_0^{\dot{\cdot}\dot{\cdot}'} = \ell_0, \quad (33d)$$

with the convention, in (33a) and in (33c), that $\|\cdot\|_{(0)}^{\text{sgn}} = 0$.

Proof. We will use the framework and results of Sect. 2 with $\mathbb{X} = \mathbb{Y} = \mathbb{R}^d$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and with the Euclidian norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$.

- We prove (33a):

$$\begin{aligned}
\delta_{\ell_0^{\leq k}}^{-\dot{\zeta}} &= \sigma_{-n(\ell_0^{\leq k})} && \text{(by (16))} \\
&= \sigma_{n(\ell_0^{\leq k})} && \text{(by symmetry of the set } \ell_0^{\leq k} \text{ and of the mapping } n) \\
&= \delta_{\ell_0^{\leq k}}^{\dot{\zeta}} && \text{(by (16))} \\
&= \sigma_{n(\ell_0^{\leq k})} && \text{(by (16))} \\
&= \sigma_{(S \cap \ell_0^{\leq k}) \cup \{0\}} && \text{(by the expression (19) of the normalization mapping } n) \\
&= \sup \{ \sigma_{\ell_0^{\leq k} \cap S}, 0 \} && \text{(as is well-known in convex analysis)} \\
&= \sup \{ \sigma_{\bigcup_{|K| \leq k} S_K}, 0 \} && \text{(as } \ell_0^{\leq k} \cap S = \bigcup_{|K| \leq k} S_K \text{ by (54a))} \\
&= \sup \left\{ \sup_{|K| \leq k} \sigma_{S_K}, 0 \right\} && \text{(as is well-known in convex analysis)} \\
&= \sup \{ \| \cdot \|_{(k)}^{\text{sgn}}, 0 \} && \text{(as } \sup_{|K| \leq k} \sigma_{S_K} = \| \cdot \|_{(k)}^{\text{sgn}} \text{ by (46))} \\
&= \| \cdot \|_{(k)}^{\text{sgn}} .
\end{aligned}$$

- We prove (33b):

$$\begin{aligned}
\delta_{\ell_0^{\leq k}}^{\dot{\zeta}\dot{\zeta}'} &= (\delta_{\ell_0^{\leq k}}^{\dot{\zeta}})^* \circ n && \text{(by the formula (25) for the biconjugate)} \\
&= (\| \cdot \|_{(k)}^{\text{sgn}})^* \circ n && \text{(by (33a))} \\
&= (\sigma_{B_{(k)}^{\text{sn}}})^* \circ n && \text{(by (41), that expresses a norm as a support function)} \\
&= \delta_{B_{(k)}^{\text{sn}}} \circ n && \text{(as } (\sigma_{B_{(k)}^{\text{sn}}})^* = \delta_{B_{(k)}^{\text{sn}}} \text{ since } B_{(k)}^{\text{sn}} \text{ is closed convex)} \\
&= \delta_{n^{-1}(B_{(k)}^{\text{sn}})} && \text{(by the definition (11) of a characteristic function)} \\
&= \delta_{\{0\} \cup n^{-1}(S \cap B_{(k)}^{\text{sn}})} && \text{(by (20) since } 0 \in B_{(k)}^{\text{sn}}) \\
&= \delta_{\{0\} \cup n^{-1}(S \cap \ell_0^{\leq k})} && \text{(as } S \cap B_{(k)}^{\text{sn}} = S \cap \ell_0^{\leq k} \text{ by (54b))} \\
&= \delta_{n^{-1}(\ell_0^{\leq k})} && \text{(by (20) since } 0 \in \ell_0^{\leq k}) \\
&= \delta_{\ell_0^{\leq k}} . && \text{(as } \ell_0 \circ n = \ell_0 \text{ by (29))}
\end{aligned}$$

- We prove (33c):

$$\begin{aligned}
\ell_0^\zeta &= \left(\inf_{l=0,1,\dots,d} \{ \delta_{\ell_0^\zeta}^\zeta + l \} \right)^\zeta && \text{(since } \ell_0 = \inf_{l=0,1,\dots,d} \{ \delta_{\ell_0^\zeta}^\zeta + l \} \text{ by using the level curves (32b))} \\
&= \sup_{l=0,1,\dots,d} \{ \delta_{\ell_0^\zeta}^\zeta + (-l) \} && \text{(as conjugacies, being dualities, turn infima into suprema)} \\
&= \sup_{l=0,1,\dots,d} \{ \sigma_{n(\ell_0^\zeta)}^\zeta + (-l) \} && \text{(as } \delta_{\ell_0^\zeta}^\zeta = \sigma_{n(\ell_0^\zeta)}^\zeta \text{ by (16))} \\
&= \sup \left\{ 0, \sup_{l=1,\dots,d} \{ \sigma_{S \cap \ell_0^\zeta}^\zeta + (-l) \} \right\} && \text{(as } n(\ell_0^\zeta) = S \cap \ell_0^\zeta \text{ when } l \geq 1 \text{ by (19))} \\
&= \sup \left\{ 0, \sup_{l=1,\dots,d} \{ \sigma_{\overline{S \cap \ell_0^\zeta}}^\zeta + (-l) \} \right\} && \text{(as } \sigma_X = \sigma_{\overline{X}} \text{ for any } X \subset \mathbb{X}) \\
&= \sup \left\{ 0, \sup_{l=1,\dots,d} \{ \sigma_{S \cap \ell_0^{\leq l}}^\zeta + (-l) \} \right\} && \text{(as } \overline{S \cap \ell_0^\zeta} = S \cap \ell_0^{\leq l} \text{ by (54c))} \\
&= \sup \left\{ 0, \sup_{l=1,\dots,d} \{ \sigma_{\cup_{|K| \leq k} S_K}^\zeta + (-l) \} \right\} && \text{(as } S \cap \ell_0^{\leq l} = \cup_{|K| \leq k} S_K \text{ by (54a))} \\
&= \sup \left\{ 0, \sup_{l=1,\dots,d} \left[\|y\|_{(l)}^{\text{sgn}} - l \right] \right\} && \text{(as } \sup_{|K| \leq k} \sigma_{S_K}^\zeta = \|\cdot\|_{(k)}^{\text{sgn}} \text{ by (46))} \\
&= \sup_{l=0,1,\dots,d} \left[\|y\|_{(l)}^{\text{sgn}} - l \right]. && \text{(with the convention that } \|\cdot\|_{(0)}^{\text{sgn}} = 0)
\end{aligned}$$

- We prove (33d).

It is easy to check that $\ell_0^{\zeta\zeta'}(0) = 0 = \ell_0(0)$. Therefore, let $x \in \mathbb{R}^d \setminus \{0\}$ be given and assume that $\ell_0(x) = l \in \{1, \dots, d\}$. We consider the mapping $\phi :]0, +\infty[\rightarrow \mathbb{R}$ defined by

$$\phi(\lambda) = \frac{\langle x, \lambda x \rangle}{\|x\|} + \left(- \sup \left\{ 0, \sup_{j=1,\dots,d} \left[\|\lambda x\|_{(j)}^{\text{sgn}} - j \right] \right\} \right), \quad \forall \lambda > 0, \quad (34)$$

and we will show that $\lim_{\lambda \rightarrow +\infty} \phi(\lambda) = l$. We have

$$\begin{aligned}
\phi(\lambda) &= \lambda \|x\| + \left(- \sup \left\{ 0, \sup_{j=1,\dots,d} \left[\|\lambda x\|_{(j)}^{\text{sgn}} - j \right] \right\} \right) && \text{(by definition (34) of } \phi) \\
&= \lambda \|x\|_{(l)}^{\text{sgn}} + \inf \left\{ 0, - \sup_{j=1,\dots,d} \left[\lambda \|x\|_{(j)}^{\text{sgn}} - j \right] \right\} && \text{(as } \|x\| = \|x\|_{(l)}^{\text{sgn}} \text{ when } \ell_0(x) = l \text{ by (51))} \\
&= \inf \left\{ \lambda \|x\|_{(l)}^{\text{sgn}}, \lambda \|x\|_{(l)}^{\text{sgn}} + \inf_{j=1,\dots,d} \left(- \left[\lambda \|x\|_{(j)}^{\text{sgn}} - j \right] \right) \right\} \\
&= \inf \left\{ \lambda \|x\|_{(l)}^{\text{sgn}}, \inf_{j=1,\dots,d} \left(\lambda (\|x\|_{(l)}^{\text{sgn}} - \|x\|_{(j)}^{\text{sgn}}) + j \right) \right\} \\
&= \inf \left\{ \lambda \|x\|_{(l)}^{\text{sgn}}, \inf_{j=1,\dots,l} \left(\lambda (\|x\|_{(l)}^{\text{sgn}} - \|x\|_{(j)}^{\text{sgn}}) + j \right) \right\} && \text{(as } \|x\|_{(j)}^{\text{sgn}} = \|x\|_{(l)}^{\text{sgn}} \text{ for } j \geq l \text{ by (51))} \\
&= \inf \left\{ \lambda \|x\|_{(l)}^{\text{sgn}}, \inf_{j=1,\dots,l-1} \left(\lambda (\|x\|_{(l)}^{\text{sgn}} - \|x\|_{(j)}^{\text{sgn}}) + j \right), l \right\}.
\end{aligned}$$

Let us show that the two first terms in the infimum go to $+\infty$ when $\lambda \rightarrow +\infty$. The first term goes to $+\infty$ because $\|x\|_{(l)}^{\text{sgn}} = \|x\| > 0$ by assumption ($x \neq 0$). The second term also goes to

$+\infty$ because $\ell_0(x) = l$, so that $\|x\| = \|x\|_{(l)}^{\text{sgn}} > \|x\|_{(j)}^{\text{sgn}}$ for $j = 1, \dots, l-1$ by (51). Therefore, $\lim_{\lambda \rightarrow +\infty} \phi(\lambda) = \inf\{+\infty, +\infty, l\} = l$. This concludes the proof since

$$\begin{aligned}
l &= \lim_{\lambda \rightarrow +\infty} \phi(\lambda) \leq \sup_{y \in \mathbb{R}^d} \left(\frac{\langle x, y \rangle}{\|x\|} \dot{+} \left(- \sup \left\{ 0, \sup_{j=1, \dots, d} \left[\|y\|_{(j)}^{\text{sgn}} - j \right] \right\} \right) \right) \quad (\text{by definition (34) of } \phi) \\
&= \sup_{y \in \mathbb{R}^d} \left(\frac{\langle x, y \rangle}{\|x\|} \dot{+} \left(- \sup_{j=0, 1, \dots, d} \left[\|y\|_{(j)}^{\text{sgn}} - j \right] \right) \right) \\
&\quad (\text{by the convention that } \|\cdot\|_{(0)}^{\text{sgn}} = 0) \\
&= \sup_{y \in \mathbb{R}^d} \left(\frac{\langle x, y \rangle}{\|x\|} \dot{+} \left(- \ell_0^\zeta(y) \right) \right) \quad (\text{by the formula (33c) for } \ell_0^\zeta) \\
&= \ell_0^{\zeta\zeta'}(x) \quad (\text{by the biconjugate formula (5)}) \\
&\leq \ell_0(x) \quad (\text{by (6a) giving } \ell_0^{\zeta\zeta'} \leq \ell_0) \\
&= l. \quad (\text{by assumption})
\end{aligned}$$

Therefore, we have obtained $l = \ell_0^{\zeta\zeta'}(x) = \ell_0(x)$.

This ends the proof. $\square$

Corollary 9 *The ℓ_0 pseudonorm ℓ_0 coincides, on the sphere S , with a convex lsc function defined on the whole space $\mathbb{R}^d$:*

$$\ell_0(x) = \left(\sup_{l=0, 1, \dots, d} \left[\|\cdot\|_{(l)}^{\text{sgn}} - l \right] \right)^*(x), \quad \forall x \in S. \quad (35)$$

Proof. For $x \in S$, we have

$$\begin{aligned}
\ell_0(x) &= \ell_0^{\zeta\zeta'}(x) \quad (\text{by (33d)}) \\
&= \sup_{y \in \mathbb{R}^d} \left(\zeta(x, y) \dot{+} \left(- \ell_0^\zeta(y) \right) \right) \quad (\text{by the biconjugate formula (5)}) \\
&= \sup_{y \in \mathbb{R}^d} \left(\langle x, y \rangle \dot{+} \left(- \ell_0^\zeta(y) \right) \right) \quad (\text{by (18) with } \|x\| = 1 \text{ since } x \in S) \\
&= \sup_{y \in \mathbb{R}^d} \left(\langle x, y \rangle \dot{+} \left(- \left(\sup_{l=0, 1, \dots, d} \left[\|y\|_{(l)}^{\text{sgn}} - l \right] \right) \right) \right) \quad (\text{by (33c)}) \\
&= \left(\sup_{l=0, 1, \dots, d} \left[\|y\|_{(l)}^{\text{sgn}} - l \right] \right)^*(x). \quad (\text{by the expression (7a) of the Fenchel conjugate})
\end{aligned}$$

This ends the proof. $\square$

4 Conclusion

In this paper, we have introduced a novel class of one-sided linear couplings, and have shown that they induce conjugacies that share nice properties with the classic Fenchel conjugacy.

Among them, we have distinguished a novel coupling, Caprac, having the property of being constant along primal rays, like the l_0 pseudonorm. For the Caprac conjugacy, induced by the coupling Caprac, we have proved that the l_0 pseudonorm is equal to its biconjugate: hence, the l_0 pseudonorm is Caprac-convex in the sense of generalized convexity. We have also provided expressions for conjugates in terms of two families of dual norms, the 2- k -symmetric gauge norms and the k -support norms.

In a companion paper [3], we apply our results to so-called sparse optimization, that is, problems where one looks for solution that have few nonzero components. We provide a systematic way to obtain convex minimization programs (over unit balls of some norms) that are lower bounds for the original exact sparse optimization problem.

Acknowledgements. We want to thank Juan Enrique Martínez Legaz and Jean-Baptiste Hiriart-Urruty for discussions on first versions of this work.

A Appendix

A.1 Background on J. J. Moreau lower and upper additions

When we manipulate functions with values in $\overline{\mathbb{R}} = [-\infty, +\infty]$, we adopt the following Moreau *lower addition* or *upper addition*, depending on whether we deal with sup or inf operations. We follow [7]. In the sequel, u, v and w are any elements of $\overline{\mathbb{R}}$.

Moreau lower addition

The Moreau *lower addition* extends the usual addition with

$$(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = -\infty . \quad (36a)$$

With the *lower addition*, $(\overline{\mathbb{R}}, \dot{+})$ is a convex cone, with $\dot{+}$ commutative and associative. The lower addition displays the following properties:

$$u \leq u' , \quad v \leq v' \Rightarrow u \dot{+} v \leq u' \dot{+} v' , \quad (36b)$$

$$(-u) \dot{+} (-v) \leq -(u \dot{+} v) , \quad (36c)$$

$$(-u) \dot{+} u \leq 0 , \quad (36d)$$

$$\sup_{a \in \mathbb{A}} f(a) \dot{+} \sup_{b \in \mathbb{B}} g(b) = \sup_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{+} g(b)) , \quad (36e)$$

$$\inf_{a \in \mathbb{A}} f(a) \dot{+} \inf_{b \in \mathbb{B}} g(b) \leq \inf_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{+} g(b)) , \quad (36f)$$

$$t < +\infty \Rightarrow \inf_{a \in \mathbb{A}} f(a) \dot{+} t = \inf_{a \in \mathbb{A}} (f(a) \dot{+} t) . \quad (36g)$$

Moreau upper addition

The Moreau *upper addition* extends the usual addition with

$$(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = +\infty . \quad (37a)$$

With the *upper addition*, $(\overline{\mathbb{R}}, \dot{+})$ is a convex cone, with $\dot{+}$ commutative and associative. The upper addition displays the following properties:

$$u \leq u' , \quad v \leq v' \Rightarrow u \dot{+} v \leq u' \dot{+} v' , \quad (37b)$$

$$(-u) \dot{+} (-v) \geq -(u \dot{+} v) , \quad (37c)$$

$$(-u) \dot{+} u \geq 0 , \quad (37d)$$

$$\inf_{a \in \mathbb{A}} f(a) \dot{+} \inf_{b \in \mathbb{B}} g(b) = \inf_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{+} g(b)) , \quad (37e)$$

$$\sup_{a \in \mathbb{A}} f(a) \dot{+} \sup_{b \in \mathbb{B}} g(b) \geq \sup_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{+} g(b)) , \quad (37f)$$

$$-\infty < t \Rightarrow \sup_{a \in \mathbb{A}} f(a) \dot{+} t = \sup_{a \in \mathbb{A}} (f(a) \dot{+} t) . \quad (37g)$$

Joint properties of the Moreau lower and upper addition

We obviously have that

$$u \dot{+} v \leq u \dot{+} v . \quad (38a)$$

The Moreau lower and upper additions are related by

$$-(u \dot{+} v) = (-u) \dot{+} (-v) , \quad -(u \dot{-} v) = (-u) \dot{+} (-v) . \quad (38b)$$

They satisfy the inequality

$$(u \dot{+} v) \dot{+} w \leq u \dot{+} (v \dot{+} w) . \quad (38c)$$

with

$$(u \dot{+} v) \dot{+} w < u \dot{+} (v \dot{+} w) \iff \begin{cases} u = +\infty \text{ and } w = -\infty , \\ \text{or} \\ u = -\infty \text{ and } w = +\infty \text{ and } -\infty < v < +\infty . \end{cases} \quad (38d)$$

Finally, we have that

$$u \dot{+} (-v) \leq 0 \iff u \leq v \iff 0 \leq v \dot{+} (-u) , \quad (38e)$$

$$u \dot{+} (-v) \leq w \iff u \leq v \dot{+} w \iff u \dot{+} (-w) \leq v , \quad (38f)$$

$$w \leq v \dot{+} (-u) \iff u \dot{+} w \leq v \iff u \leq v \dot{+} (-w) . \quad (38g)$$

A.2 Properties of 2- k -symmetric gauge norms

Before studying properties of 2- k -symmetric gauge norms, we recall the notion of dual norm.

Let $||| \cdot |||$ be a norm on $\mathbb{R}^d$, with unit ball denoted by

$$B_{||| \cdot |||} = \{x \in \mathbb{R}^d \mid |||x||| \leq 1\} . \quad (39)$$

Definition 10 *The following expression*

$$|||y|||_\star = \sup_{|||x||| \leq 1} \langle x, y \rangle, \quad \forall y \in \mathbb{Y} \quad (40)$$

defines a norm on $\mathbb{Y}$, called the dual norm $||| \cdot |||_\star$.

We have

$$||| \cdot |||_\star = \sigma_{B_{||| \cdot |||}} \quad \text{and} \quad ||| \cdot ||| = \sigma_{B_{||| \cdot |||_\star}}, \quad (41)$$

where $B_{||| \cdot |||_\star}$ is the unit ball of the dual norm:

$$B_{||| \cdot |||_\star} = \{y \in \mathbb{Y} \mid |||y|||_\star \leq 1\}. \quad (42)$$

For all $K \subset \{1, \dots, d\}$, we introduce *degenerate* unit “spheres” and “balls” of $\mathbb{R}^d$ by

$$\mathbb{S}_K = \{x \in \mathbb{R}^d \mid \|x_K\| = 1\}, \quad (43a)$$

$$\mathbb{B}_K = \{x \in \mathbb{R}^d \mid \|x_K\| \leq 1\}, \quad (43b)$$

$$S_K = \{x \in \mathbb{R}^d \mid x_{-K} = 0 \text{ and } \|x_K\| = 1\}, \quad (43c)$$

$$B_K = \{x \in \mathbb{R}^d \mid x_{-K} = 0 \text{ and } \|x_K\| \leq 1\}, \quad (43d)$$

where x_K has been defined right before Definition 7.

In what follows, the notation $\bigcup_{|K| \leq k}$ is a shorthand for $\bigcup_{K \subset \{1, \dots, d\}, |K| \leq k}$, $\bigcap_{|K| \leq k}$ for $\bigcap_{K \subset \{1, \dots, d\}, |K| \leq k}$, and $\sup_{|K| \leq k}$ for $\sup_{K \subset \{1, \dots, d\}, |K| \leq k}$. The same holds true for $\bigcup_{|K|=k}$, $\bigcap_{|K|=k}$ and $\sup_{|K|=k}$.

Proposition 11 *Let $k \in \{1, \dots, d\}$.*

- *The following inequalities hold true*

$$\sup_{j=1, \dots, d} |x_j| = \|x\|_\infty = \|x\|_{(1)}^{\text{sgn}} \leq \dots \leq \|x\|_{(k)}^{\text{sgn}} \leq \|x\|_{(k+1)}^{\text{sgn}} \leq \dots \leq \|x\|_{(n)}^{\text{sgn}} = \|x\|. \quad (44)$$

- *The two “spheres” in (43a) and (43c) are related by*

$$S_K = S \cap \mathbb{S}_K, \quad \forall K \subset \{1, \dots, d\}. \quad (45)$$

- *The 2- k -symmetric gauge norm $\|\cdot\|_{(k)}^{\text{sgn}}$ in Definition 7 satisfies*

$$\|\cdot\|_{(k)}^{\text{sgn}} = \sigma_{\bigcup_{|K| \leq k} B_K} = \sup_{|K| \leq k} \sigma_{B_K} = \sup_{|K| \leq k} \sigma_{S_K} = \sigma_{\bigcup_{|K| \leq k} S_K}, \quad (46)$$

where $|K| \leq k$ can be replaced by $|K| = k$ everywhere.

- The unit sphere $S_{(k)}^{\text{sgn}}$ and ball $B_{(k)}^{\text{sgn}}$ of $\mathbb{R}^d$ for the 2- k -symmetric gauge norm $\|\cdot\|_{(k)}^{\text{sgn}}$ in Definition 7 satisfy

$$B_{(k)}^{\text{sgn}} = \left\{ x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\text{sgn}} \leq 1 \right\} = \bigcap_{|K| \leq k} \mathbb{B}_K, \quad (47a)$$

$$S_{(k)}^{\text{sgn}} = \left\{ x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\text{sgn}} = 1 \right\} = B_{(k)}^{\text{sgn}} \cap \left(\bigcup_{|K| \leq k} \mathbb{S}_K \right), \quad (47b)$$

where $|K| \leq k$ can be replaced by $|K| = k$ everywhere.

- The unit ball $B_{(k)}^{\text{sn}}$ of the k -support norm $\|\cdot\|_{(k)}^{\text{sn}}$ in Definition 7 satisfies

$$B_{(k)}^{\text{sn}} = \left\{ x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\text{sn}} \leq 1 \right\} = \overline{\text{co}}\left(\bigcup_{|K| \leq k} B_K \right) = \overline{\text{co}}\left(\bigcup_{|K| \leq k} S_K \right), \quad (48)$$

where $|K| \leq k$ can be replaced by $|K| = k$ everywhere.

Proof.

- The Inequalities and Equalities (44) easily derive from the very definition (30) of the 2- k -symmetric gauge norm $\|\cdot\|_{(k)}^{\text{sgn}}$.
- We prove Equation (45). Recall that, following notation from Game Theory, we denote by $-K$ the complementary subset of K in $\{1, \dots, d\}$: $K \cup (-K) = \{1, \dots, d\}$ and $K \cap (-K) = \emptyset$. Then, we have that $x = x_K + x_{-K}$, for any $x \in \mathbb{R}^d$, and the decomposition is orthogonal, leading to

$$(\forall x \in \mathbb{R}^d) \quad x = x_K + x_{-K}, \quad x_K \perp x_{-K} \quad \text{and} \quad \|x\|^2 = \|x_K\|^2 + \|x_{-K}\|^2. \quad (49)$$

For $K \subset \{1, \dots, d\}$, we have that

$$\begin{aligned} x \in S \quad \text{and} \quad x \in S_K &\iff 1 = \|x\|^2 \quad \text{and} \quad 1 = \|x_K\|^2 && \text{(by (43a))} \\ &\iff 1 = \|x\|^2 = \|x_K\|^2 + \|x_{-K}\|^2 \quad \text{and} \quad 1 = \|x_K\|^2 && \text{(by (49))} \\ &\iff \|x_{-K}\| = 0 \quad \text{and} \quad 1 = \|x_K\| && \text{(by (49))} \\ &\iff x \in S_K. && \text{(by (43c))} \end{aligned}$$

- We prove Equation (46). For this purpose, we first establish that

$$\sigma_{B_K}(y) = \|y_K\|, \quad \forall y \in \mathbb{R}^d. \quad (50)$$

Indeed, for $y \in \mathbb{R}^d$, we have

$$\begin{aligned} \sigma_{B_K}(y) &= \sup_{x \in B_K} \langle x, y \rangle && \text{(by definition (17) of a support function)} \\ &= \sup_{x \in B_K} \langle x_K + x_{-K}, y_K + y_{-K} \rangle && \text{(by the decomposition (49))} \\ &= \sup_{x \in B_K} (\langle x_K, y_K \rangle + \langle x_{-K}, y_{-K} \rangle) && \text{(because } x_K \perp y_{-K} \text{ and } x_{-K} \perp y_K \text{ by (49))} \\ &= \sup \left\{ \langle x_K, y_K \rangle + \langle x_{-K}, y_{-K} \rangle, x_{-K} = 0 \quad \text{and} \quad \|x_K\| \leq 1 \right\} && \text{(by definition (43d) of } B_K) \\ &= \sup \left\{ \langle x_K, y_K \rangle, \|x_K\| \leq 1 \right\} \\ &= \|y_K\| \end{aligned}$$

as is well-known for the Euclidian norm $\|\cdot\|$, when restricted to the subspace $\{x \in \mathbb{R}^d \mid x_{-K} = 0\}$ (because it is equal to its dual norm). Then, for all $y \in \mathbb{R}^d$, we have that

$$\begin{aligned}\sigma_{\cup_{|K| \leq k} B_K}(y) &= \sup_{|K| \leq k} \sigma_{B_K}(y) && \text{(as is well-known in convex analysis)} \\ &= \sup_{|K| \leq k} \|y_K\| && \text{(by (50))} \\ &= \|y\|_{(k)}^{\text{sgn}}. && \text{(by definition (30) of } \|\cdot\|_{(k)}^{\text{sgn}}\text{)}\end{aligned}$$

Now, by (43c) and (43d), it is straightforward that $\overline{\text{co}}(S_K) = B_K$ and we deduce that

$$\|\cdot\|_{(k)}^{\text{sgn}} = \sigma_{\cup_{|K| \leq k} B_K} = \sup_{|K| \leq k} \sigma_{B_K} = \sup_{|K| \leq k} \sigma_{\overline{\text{co}}(S_K)} = \sup_{|K| \leq k} \sigma_{S_K} = \sigma_{\cup_{|K| \leq k} S_K},$$

giving Equation (46).

If we take over the proof using the property that $\sup_{|K| \leq k} \|y_K\| = \sup_{|K|=k} \|y_K\|$ in (30), we deduce that $|K| \leq k$ can be replaced by $|K| = k$ everywhere.

- We prove Equation (47a):

$$\begin{aligned}B_{(k)}^{\text{sgn}} &= \left\{x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\text{sgn}} \leq 1\right\} && \text{(by definition of the ball } B_{(k)}^{\text{sgn}}\text{)} \\ &= \left\{x \in \mathbb{R}^d \mid \sup_{|K| \leq k} \|x_K\| \leq 1\right\} && \text{(by definition (30) of } \|\cdot\|_{(k)}^{\text{sgn}}\text{)} \\ &= \bigcap_{|K| \leq k} \left\{x \in \mathbb{R}^d \mid \|x_K\| \leq 1\right\} \\ &= \bigcap_{|K| \leq k} \mathbb{B}_K. && \text{(by definition (43b) of } \mathbb{B}_K\text{)}\end{aligned}$$

If we take over the proof using the property that $\sup_{|K| \leq k} \|y_K\| = \sup_{|K|=k} \|y_K\|$ in (30), we deduce that $|K| \leq k$ can be replaced by $|K| = k$ everywhere.

- We prove Equation (47b):

$$\begin{aligned}S_{(k)}^{\text{sgn}} &= \left\{x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\text{sgn}} = 1\right\} && \text{(by definition of the sphere } S_{(k)}^{\text{sgn}}\text{)} \\ &= \left\{x \in \mathbb{R}^d \mid \sup_{|K| \leq k} \|x_K\| = 1\right\} && \text{(by definition (30) of } \|\cdot\|_{(k)}^{\text{sgn}}\text{)} \\ &= \left\{x \in \mathbb{R}^d \mid \sup_{|K| \leq k} \|x_K\| \leq 1\right\} \\ &\quad \cap \left\{x \in \mathbb{R}^d \mid \exists K \subset \{1, \dots, d\}, \ |K| \leq k, \ \|x_K\| = 1\right\} \\ &= B_{(k)}^{\text{sgn}} \cap \left(\bigcup_{|K| \leq k} \left\{x \in \mathbb{R}^d \mid \|x_K\| = 1\right\}\right) && \text{(by definition of the ball } B_{(k)}^{\text{sgn}}\text{)} \\ &= B_{(k)}^{\text{sgn}} \cap \left(\bigcup_{|K| \leq k} \mathbb{S}_K\right). && \text{(by definition (43a) of } \mathbb{S}_K\text{)}\end{aligned}$$

If we take over the proof using the property that $\sup_{|K| \leq k} \|y_K\| = \sup_{|K|=k} \|y_K\|$ in (30), we deduce that $|K| \leq k$ can be replaced by $|K| = k$ everywhere.

• We prove Equation (48). On the one hand, by the first relation in (41), we have that $\|\cdot\|_{(k)}^{\text{sgn}} = \sigma_{B_{(k)}^{\text{sn}}}$. On the other hand, by (46), we have that $\|\cdot\|_{(k)}^{\text{sgn}} = \sigma_{\cup_{|K| \leq k} B_K} = \sigma_{\cup_{|K| \leq k} S_K}$. Then, as is well-known in convex analysis, we deduce that $\overline{\text{co}}(B_{(k)}^{\text{sn}}) = \overline{\text{co}}(\cup_{|K| \leq k} B_K) = \overline{\text{co}}(\cup_{|K| \leq k} S_K)$. As the unit ball $B_{(k)}^{\text{sn}}$ is closed and convex, we immediately obtain (48).

If we take over the proof using the property that $\sigma_{\cup_{|K| \leq k} B_K} = \sigma_{\cup_{|K| \leq k} S_K} = \sigma_{\cup_{|K|=k} B_K} = \sigma_{\cup_{|K|=k} S_K}$ in (46), we deduce that $|K| \leq k$ can be replaced by $|K| = k$ everywhere. $\square$

A.3 Properties of the level sets of the l_0 pseudonorm

A connection between the l_0 pseudonorm in (28) and the 2- k -symmetric gauge norm $\|\cdot\|_{(k)}^{\text{sgn}}$ in (30) is given by the (easily proved) following Proposition.

Proposition 12 *Let $k \in \{0, 1, \dots, d\}$. For any $x \in \mathbb{R}^d$, we have*

$$\ell_0(x) = k \iff 0 = \|x\|_{(0)}^{\text{sgn}} \leq \dots \leq \|x\|_{(k-1)}^{\text{sgn}} < \|x\|_{(k)}^{\text{sgn}} = \dots = \|x\|_{(n)}^{\text{sgn}} = \|x\|, \quad (51)$$

from which we deduce the formula

$$\ell_0(x) = \min \left\{ j \in \{0, 1, \dots, d\} \mid \|x\|_{(j)}^{\text{sgn}} = \|x\| \right\}, \quad (52)$$

with the convention that $\|\cdot\|_{(0)}^{\text{sgn}} = 0$.

We prove the following Proposition about the level sets of the l_0 pseudonorm.

Proposition 13 *Let $k \in \{0, 1, \dots, d\}$. The level set $\ell_0^{\leq k}$ in (32a) of the l_0 pseudonorm in (28) satisfies*

$$(\forall x \in \mathbb{R}^d) \quad x \in \ell_0^{\leq k} \iff \ell_0(x) \leq k \iff \|x\|_{(k)}^{\text{sgn}} = \|x\|, \quad (53a)$$

$$(\forall x \in \mathbb{R}^d) \quad x \in \ell_0^{\leq k} \setminus \{0\} \iff 0 < \ell_0(x) \leq k \iff x \neq 0 \text{ and } \frac{x}{\|x\|} \in S \cap S_{(k)}^{\text{sgn}}, \quad (53b)$$

and its intersection with the sphere S has the three following expressions

$$S \cap \ell_0^{\leq k} = \bigcup_{|K| \leq k} S_K = \bigcup_{|K|=k} S_K, \quad (54a)$$

$$S \cap \ell_0^{\leq k} = S \cap B_{(k)}^{\text{sn}}, \quad (54b)$$

$$S \cap \ell_0^{\leq k} = \overline{S \cap \ell_0^{\leq k}}. \quad (54c)$$

Proof.

- The Equivalence (53a) easily follows from (51).
- We prove the Equivalence (53b). Indeed, using Equation (53a) we have that, for $x \in \mathbb{R}^d \setminus \{0\}$:

$$\ell_0(x) \leq k \iff \|x\|_{(k)}^{\text{sgn}} = \|x\| \iff \left\| \frac{x}{\|x\|} \right\|_{(k)}^{\text{sgn}} = 1 \iff \frac{x}{\|x\|} \in S_{(k)}^{\text{sgn}} \iff \frac{x}{\|x\|} \in S \cap S_{(k)}^{\text{sgn}}.$$

- We prove Equation (54a):

$$\begin{aligned} S \cap \ell_0^{\leq k} &= \left\{ x \in \mathbb{R}^d \mid \|x\| = 1 \text{ and } \ell_0(x) \leq k \right\} && \text{(by definitions (26) of } S \text{ and (32b) of } \ell_0^{\leq k}) \\ &= \left\{ x \in \mathbb{R}^d \mid \|x\| = 1 \text{ and } \|x\|_{(k)}^{\text{sgn}} = \|x\| \right\} && \text{(by (53a))} \\ &= \left\{ x \in \mathbb{R}^d \mid \|x\| = 1 \text{ and } \|x\|_{(k)}^{\text{sgn}} = 1 \right\} \\ &= S \cap S_{(k)}^{\text{sgn}} && \text{(by definitions (26) and (47b) of the spheres } S \text{ and } S_{(k)}^{\text{sgn}}) \\ &= S \cap B_{(k)}^{\text{sgn}} \cap \left(\bigcup_{|K| \leq k} \mathbb{S}_K \right) && \text{(by property (47b) of the sphere } S_{(k)}^{\text{sgn}}) \\ &= S \cap \left(\bigcup_{|K| \leq k} \mathbb{S}_K \right) && \text{(as, by (44), we have that } S \subset B_{(k)}^{\text{sgn}}) \\ &= \bigcup_{|K| \leq k} (S \cap \mathbb{S}_K) \\ &= \bigcup_{|K| \leq k} S_K. && \text{(as } S \cap \mathbb{S}_K = S_K \text{ by (45))} \end{aligned}$$

If we take over the proof where we use $S_{(k)}^{\text{sgn}} = B_{(k)}^{\text{sgn}} \cap \left(\bigcup_{|K|=k} \mathbb{S}_K \right)$ in (47b), we obtain that $S \cap \ell_0^{\leq k} = \bigcup_{|K|=k} S_K$.

- We prove Equation (54b). First, we observe that the level set $\ell_0^{\leq k}$ is closed because, by (53a), it can be expressed as $\ell_0^{\leq k} = \left\{ x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\text{sgn}} = \|x\| \right\}$. This also follows from the well-known property that the ℓ_0 pseudonorm ℓ_0 is lower semi continuous. Second, we have

$$\begin{aligned} S \cap \ell_0^{\leq k} &= S \cap \overline{\text{co}}(S \cap \ell_0^{\leq k}) && \text{(by Lemma 14 since } S \cap \ell_0^{\leq k} \subset S \text{ and is closed)} \\ &= S \cap \overline{\text{co}}\left(\bigcup_{|K| \leq k} S_K \right) && \text{(as } S \cap \ell_0^{\leq k} = \bigcup_{|K| \leq k} S_K \text{ by (54a))} \\ &= S \cap B_{(k)}^{\text{sn}}. && \text{(as } \overline{\text{co}}\left(\bigcup_{|K| \leq k} S_K \right) = B_{(k)}^{\text{sn}} \text{ by (48))} \end{aligned}$$

- We prove Equation (54c). For this purpose, we first establish the (known) fact that $\overline{\ell_0^{\leq k}} = \ell_0^{\leq k}$. The inclusion $\overline{\ell_0^{\leq k}} \subset \ell_0^{\leq k}$ is easy. Indeed, as we have seen that $\ell_0^{\leq k}$ is closed, we have $\ell_0^{\leq k} \subset \overline{\ell_0^{\leq k}} \Rightarrow \overline{\ell_0^{\leq k}} \subset \ell_0^{\leq k} = \ell_0^{\leq k}$. There remains to prove the reverse inclusion $\ell_0^{\leq k} \subset \overline{\ell_0^{\leq k}}$. For this purpose, we consider $x \in \ell_0^{\leq k}$. If $x \in \ell_0^{\leq k}$, obviously $x \in \overline{\ell_0^{\leq k}}$. Therefore, we suppose that $\ell_0(x) = l < k$. By

definition of $\ell_0(x)$, there exists $L \subset \{1, \dots, d\}$ such that $|L| = l < k$ and $x = x_L$. For $\epsilon > 0$, define x^ϵ as coinciding with x except for $k - l$ indices outside L for which the components are $\epsilon > 0$. By construction $\ell_0(x^\epsilon) = k$ and $x^\epsilon \rightarrow x$ when $\epsilon \rightarrow 0$. This proves that $\ell_0^{\leq k} \subset \overline{\ell_0^k}$.

Second, we prove that $S \cap \ell_0^{\leq k} = \overline{S \cap \ell_0^k}$. The inclusion $\overline{S \cap \ell_0^k} \subset S \cap \ell_0^{\leq k}$, is easy. Indeed, $\overline{\ell_0^k} = \ell_0^{\leq k} \Rightarrow \overline{S \cap \ell_0^k} \subset \overline{S \cap \ell_0^{\leq k}} = S \cap \ell_0^{\leq k}$. To prove the reverse inclusion $S \cap \ell_0^{\leq k} \subset \overline{S \cap \ell_0^k}$, we consider $x \in S \cap \ell_0^{\leq k}$. As we have just seen that $\ell_0^{\leq k} = \overline{\ell_0^k}$, we deduce that $x \in \overline{\ell_0^k}$. Therefore, there exists a sequence $\{z_n\}_{n \in \mathbb{N}}$ in ℓ_0^k such that $z_n \rightarrow x$ when $n \rightarrow +\infty$. Since $x \in S$, we can always suppose that $z_n \neq 0$, for all $n \in \mathbb{N}$. Therefore $z_n/\|z_n\|$ is well defined and, when $n \rightarrow +\infty$, we have $z_n/\|z_n\| \rightarrow x/\|x\| = x$ since $x \in S = \{x \in \mathbb{X} \mid \|x\| = 1\}$. Now, on the one hand, $z_n/\|z_n\| \in \ell_0^k$, for all $n \in \mathbb{N}$, and, on the other hand, $z_n/\|z_n\| \in S$. As a consequence $z_n/\|z_n\| \in S \cap \ell_0^k$, and we conclude that $x \in \overline{S \cap \ell_0^k}$. Thus, we have proved that $S \cap \ell_0^{\leq k} \subset \overline{S \cap \ell_0^k}$.

This ends the proof. $\square$

Lemma 14 *If A is a subset of the Euclidian sphere S of $\mathbb{R}^d$, then $A = \text{co}(A) \cap S$. If A is a closed subset of the Euclidian sphere S of $\mathbb{R}^d$, then $A = \overline{\text{co}}(A) \cap S$.*

Proof. We first prove that $A = \text{co}(A) \cap S$ when $A \subset S$. Since $A \subset \text{co}(A)$ and $A \subset S$, we immediately get that $A \subset \text{co}(A) \cap S$. To prove the reverse inclusion, we first start by proving that $\text{co}(A) \cap S \subset \text{extr}(\text{co}(A))$, the set of extreme points of $\text{co}(A)$.

The proof is by contradiction. Suppose indeed that there exists $x \in \text{co}(A) \cap S$ and $x \notin \text{extr}(\text{co}(A))$. Then, we could find $y \in \text{co}(A)$ and $z \in \text{co}(A)$, distinct from x , and such that $x = \lambda y + (1 - \lambda)z$ for some $\lambda \in (0, 1)$. Notice that necessarily $y \neq z$ (because, else, we would have $x = y = z$ which would contradict $y \neq x$ and $z \neq x$). By assumption $A \subset S$, we deduce that $\text{co}(A) \subset B = \{x \in \mathbb{X} \mid \|x\| \leq 1\}$, the unit ball, and therefore that $\|y\| \leq 1$ and $\|z\| \leq 1$. If y or z were not in S — that is, if either $\|y\| < 1$ or $\|z\| < 1$ — then we would obtain that $\|x\| \leq \lambda\|y\| + (1 - \lambda)\|z\| < 1$ since $\lambda \in (0, 1)$; we would thus arrive at a contradiction since x could not be in S . Thus, both y and z must be in S , and we have a contradiction since no $x \in S$, the Euclidian sphere, can be obtained as a convex combination of $y \in S$ and $z \in S$, with $y \neq z$.

Hence, we have proved by contradiction that $\text{co}(A) \cap S \subset \text{extr}(\text{co}(A))$. We can conclude using the fact that $\text{extr}(\text{co}(A)) \subset A$ (see [4, Exercice 6.4]).

Now, we consider the case where the subset A of the Euclidian sphere S is closed. Using the first part of the proof we have that $A = \text{co}(A) \cap S$. Now, A is closed by assumption and bounded since $A \subset S$. Thus, A is compact and in a finite dimensional space we have that $\text{co}(A)$ is compact [8, Th. 17.2], thus closed. We conclude that $A = \text{co}(A) \cap S = \overline{\text{co}(A)} \cap S = \overline{\text{co}}(A) \cap S$, where the last equality comes from [2, Prop. 3.46]. $\square$

References

- [1] A. Argyriou, R. Foygel, and N. Srebro. Sparse prediction with the k -support norm. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, pages 1457–1465, USA, 2012. Curran Associates Inc.

- [2] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer-Verlag, second edition, 2017.
- [3] J.-P. Chancelier and M. De Lara. Lower bound convex programs for exact sparse optimization, 2019. preprint.
- [4] J. Hiriart-Urruty. *Optimisation et analyse convexe*. Presses Universitaires de France, 1998.
- [5] J.-B. Hiriart-Urruty and H. Le. A variational approach of the rank function. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, 21(2):207–240, 2013.
- [6] J. E. Martínez-Legaz. Generalized convex duality and its economic applications. In S. S. Hadjisavvas N., Komlósi S., editor, *Handbook of Generalized Convexity and Generalized Monotonicity. Nonconvex Optimization and Its Applications*, volume 76, pages 237–292. Springer-Verlag, 2005.
- [7] J. J. Moreau. Inf-convolution, sous-additivité, convexité des fonctions numériques. *J. Math. Pures Appl. (9)*, 49:109–154, 1970.
- [8] T. R. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N.J., 1970.
- [9] A. Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.
- [10] I. Singer. *Abstract convex analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons, Inc., New York, 1997.