Optimization
A Journal of Mathematical Programming and Operations Research

ISSN: 0233-1934 (Print) 1029-4945 (Online) Journal homepage: http://www.tandfonline.com/loi/gopt20

# Construction of proximal distances over symmetric cones

Julio López & Erik Alex Papa Quiroz

OPTIMIZATION, 2017
VOL. 66, NO. 8, 1301–1321
https://doi.org/10.1080/02331934.2016.1277998

Taylor & Francis
Taylor & Francis Group



# Construction of proximal distances over symmetric cones

Julio López[a] and Erik Alex Papa Quiroz[b]

[a]Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Santiago, Chile; [b]Facultad de Ciencias Matemáticas, Universidad Nacional Mayor de San Marcos, Lima, Perú

**ABSTRACT**
This paper is devoted to the study of proximal distances defined over symmetric cones, which include the non-negative orthant, the second-order cone and the cone of positive semi-definite symmetric matrices. Specifically, our first aim is to provide two ways to build them. For this, we consider two classes of real-valued functions satisfying some assumptions. Then, we show that its corresponding spectrally defined function defines a proximal distance. In addition, we present several examples and some properties of this distance. Taking into account these properties, we analyse the convergence of proximal-type algorithms for solving convex symmetric cone programming (SCP) problems, and we study the asymptotic behaviour of primal central paths associated with a proximal distance. Finally, for linear SCP problems, we provide a relationship between the proximal sequence and the primal central path.

**ARTICLE HISTORY**
Received 29 February 2016
Accepted 22 December 2016

**KEYWORDS**
Proximal distance; spectrally defined function; symmetric cone programming; proximal-type algorithms; primal central paths

## 1. Introduction

Since the introduction made by Martinet in 1970 [1], several proximal algorithms for solving constrained problems (for e.g. convex programming, variational inequalities, non-linear complementarity problem, etc.) have been studied. In particular, for convex programming problems over: the nonnegative orthant in $\mathbb{R}^n$ (see e.g. [2,3]), the Cartesian product of second-order cones (see e.g. [4–6]), and the positive semi-definite symmetric matrices (see e.g. [7]). These algorithms have the particularity of using a non-quadratic distance-like function to exploit the geometry of the constraints.

With the purpose of providing a unified technique for analysing and designing interior proximal methods with non-quadratic distance, Auslender and Teboulle [8] defined a class of proximal distances with respect to an open nonempty convex set of a Euclidean space, that includes the class of Bregman [2,3], $\varphi$-divergence [9,10], and second-order homogeneous distances [11]. The idea of using these classes of proximal distances has been carried out by several researchers [6,12–17]. For instance, Pan and Chen [6] defined a proximal distance with respect to second-order cones, made a unified analysis of interior proximal methods for solving convex second-order cone programming (SOCP) problems, and studied the central paths associated with these distance functions. In Auslender et al. [12] studied nonmonotone projection gradient schemes based on proximal distances for solving constrained nonconvex problems. Papa Quiroz and Oliveira [15] proposed proximal-type methods with proximal distance to solve minimization problems with quasiconvex objective functions on the nonnegative orthant. In [17], Sarmiento et al. studied an inexact proximal multiplier method using proximal distances for solving convex minimization problems with a separable structure.

An important and popular non-quadratic distance is the Kullback–Leibler (KL) entropy distance defined over the nonnegative orthant [18]. Using algebraic techniques of matrices and Euclidean

**CONTACT** Julio López julio.lopez@udp.cl

Jordan algebra (EJA) [19], the KL entropy distance has been extended and used in order to introduce and analyse an entropy-like proximal algorithm, and an exponential multiplier method for the problem of minimizing a closed proper convex function subject to positive semi-definite symmetric matrices [7] and symmetric cone constraints [20], respectively.

In addition, in the context of linear symmetric cone programming (SCP), the KL entropy distance has been used for the study of convergence of central paths [21].

Surprisingly, in recent years, it has been established that every symmetric cone can be casted as one of square elements of a suitable chosen EJA [19], and it serves as a unifying framework to which the important cases of cones used in optimization, such as the non-negative orthant in $\mathbb{R}^n$, the second-order cone (SOC) and the cone of positive semi-definite symmetric matrices belong.

Using the framework of EJAs, several interior point methods (IPMs) have been successfully extended to optimization over symmetric cones [22–30]. For instance, Faybusovich [22] extended IPMs from semidefinite programming to the linear SCP. Gu et al. [25] generalized Roos et al.'s algorithm for linear programming (LP) in [31] to linear SCP. Wang et al. [28] generalized the classical primal-dual logarithmic barrier method for LP to convex quadratic optimization over symmetric cones. Ramírez and Sossa [30] studied the asymptotic behaviour of central paths with respect to a broad class of barrier functions for convex SCP.

Recently, Alvarez and López [32] extended the interior proximal algorithm with variable metric proposed in [33] for SOCP to convex SCP, and used the technique of proximal bundle methods to develop an implementable version of the proposed algorithm.

In this paper, we first focus on studying proximal distances defined over symmetric cones. Specifically, we provide two ways to build these distances. The first approach is based on Bregman distance, while the second one is based on Distance-like, which is studied by Pan and Chen [4] for SOCs. For both approaches, we use the spectrally defined function of a real-valued function. To the best of our knowledge, the only reference that defines a proximal distance over symmetric cones, specifically the KL entropy distance, is the work proposed by Chen and Pan [20]. Taking into account the above construction, we proceed to study interior proximal-type algorithms and primal central paths associated with this proximal distance for convex SCP.

This paper is structured as follows; Section 2 reviews some basic notions and results on EJAs, presents some properties about spectrally defined functions over symmetric cones, and introduces the class of proximal distances defined on the symmetric cones. In Section 3, we provide two ways to construct proximal distances via the compute of the spectrally defined functions of two suitable classes of real-valued functions, and we discuss some properties of these distances. Several examples of these classes of proximal distances are presented in Section 4. In Section 5, first we study the convergence of the proximal-type algorithms associated with a proximal distance. Then, we describe some results of the primal central path and its connection with the proximal-type algorithm. Finally, concluding remarks are given in Section 6.

## 2. Preliminaries

In this section, we briefly describe some concepts, properties and results from EJAs and symmetric cones that are needed in this paper; for more details, see, e.g. Faraut and Korányi [19], Schmieta and Alizadeh [24]. Subsequently, we recall some results from spectrally defined functions over symmetric cones and present the classical notion of essential smoothness in our context. Finally, we present the definition of proximal distance with respect to the interior of a symmetric cone.

### 2.1. Euclidean Jordan algebras

A EJA is a triple $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle_{\mathbb{V}})$, shortly denoted by $\mathbb{V}$, where $(\mathbb{V}, \langle\cdot,\cdot\rangle_{\mathbb{V}})$ is a finite-dimensional space over the real field $\mathbb{R}$ equipped with an inner product $\langle\cdot,\cdot\rangle_{\mathbb{V}}$, and $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ is a bilinear mapping satisfying: (i) $x \circ y = y \circ x$; (ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, where $x^2 := x \circ x$; (iii) $\langle x \circ y, z\rangle_{\mathbb{V}} = \langle y, x \circ z\rangle_{\mathbb{V}}$, for all $x, y, z \in \mathbb{V}$.

Here, $x \circ y$ is called the *Jordan product* of $x$ and $y$. In addition, we assume that there exists a (unique) unitary element $e \in \mathbb{V}$ such that $x \circ e = x$ for all $x \in \mathbb{V}$. An EJA is said to be *simple* if it is not a direct sum of two EJAs.

The set $\mathcal{K} := \{x^2 : x \in \mathbb{V}\}$ is called the cone of squares of the EJA $\mathbb{V}$, which is a *symmetric cone* (see [19, Theorem III.2.1]).

The *rank* of $\mathbb{V}$ is defined as $r := \max\{\deg(x) : x \in \mathbb{V}\}$, where $\deg(x)$ denotes the degree of $x \in \mathbb{V}$ given by $\deg(x) = \min\{k > 0 : \{e, x, x^2, \dots, x^k\}$ is linearly dependent$\}$. From now on, we assume that $\mathbb{V}$ is an EJA with rank $r$.

An element $c \in \mathbb{V}$ is an *idempotent* iff $c^2 = c$. An idempotent $c$ is *primitive* iff it is nonzero and cannot be written as a sum of two nonzero idempotents. We say that a finite set $\{e_1, \dots, e_r\}$ of primitive idempotents in $\mathbb{V}$ is a *Jordan frame* iff $e_i \circ e_j = 0$, for all $i \neq j$, and $\sum_{i=1}^r e_i = e$.

The following theorem gives us a spectral decomposition for the elements in an EJA (see Theorem III.1.2 of [19]).

**Theorem 2.1 (Spectral decomposition theorem):** *For every $x \in \mathbb{V}$, there exists a Jordan frame $\{e_1(x), \dots, e_r(x)\}$ and real numbers $\lambda_1(x), \dots, \lambda_r(x)$, arranged in the decreasing order, such that $x = \sum_{i=1}^r \lambda_i(x) e_i(x)$.*

The numbers $\lambda_i(x)$ (counting multiplicities) are uniquely determined by $x$, and are called the *eigenvalues* of $x$. The *trace* of $x$, denoted as $\mathrm{tr}(x)$, is defined by $\mathrm{tr}(x) := \sum_{i=1}^r \lambda_i(x)$; whereas the *determinant* of $x$ is defined by $\det(x) := \prod_{i=1}^r \lambda_i(x)$.

It is easy to show that $x \in \mathcal{K}$ (resp. $\mathrm{int}(\mathcal{K})$) iff $\lambda_i(x) \geq 0$ (resp. $\lambda_i(x) > 0$), for all $i = 1, \dots, r$. Moreover, an element $x \in \mathbb{V}$ is *invertible*, if $\det(x) \neq 0$.

According to [22, Proposition III.1.5], the bilinear form $\langle x, y\rangle := \mathrm{tr}(x \circ y)$ is an inner product on $\mathbb{V}$ and it is called the *canonical (or trace) inner product on* $\mathbb{V}$. Furthermore, we can define the norm induced by the inner product $\langle x, y\rangle$ on $\mathbb{V}$, called *Frobenius norm*, by

$$\|x\|_F := \sqrt{\langle x, x\rangle} = \mathrm{tr}(x^2) = \left(\sum_{i=1}^r \lambda_i^2(x)\right)^{1/2}, \quad \forall x \in \mathbb{V}.$$

Then, in this inner product, the norm of any primitive idempotent is one. Hence, $\|e\|_F = \sqrt{r}$. Also, this norm satisfies the Cauchy–Schwartz inequality $|\langle x, y\rangle| \leq \|x\|_F \|y\|_F$. Finally, we recall the following technical results ([34, Theorem 23], [23, Theorem 5.13]), which will be used to investigate some properties of the proximal distance $H$ that will be defined in Section 2.3.

**Lemma 2.2 (Von Neumann inequality):** *For any $x, y \in \mathbb{V}$, one has that*

$$\langle x, y\rangle = \mathrm{tr}(x \circ y) \leq \sum_{i=1}^r \lambda_i(x)\lambda_i(y).$$

*The equality holds iff $x$ and $y$ have a similar joint decomposition, that is, if there exists a Jordan frame $\{e_1, \dots, e_r\}$ such that $x = \sum_{i=1}^r \lambda_i(x) e_i$ and $y = \sum_{i=1}^r \lambda_i(y) e_i$.*

**Lemma 2.3:** *Let $x \in \mathbb{V}$. Then, $x \in \mathcal{K}$ iff $\langle x, y\rangle \geq 0$ holds for all $y \in \mathcal{K}$. Moreover, $x \in \mathrm{int}(\mathcal{K})$ iff $\langle x, y\rangle > 0$ for all $y \in \mathcal{K} \setminus \{0\}$.*

## 2.2. Spectrally defined function over symmetric cones

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a scalar-valued function. Following [34,35], one can define a vector-valued function $\phi^{\mathrm{sc}} : \mathbb{V} \to \mathbb{V}$, called *Löwner's operator*, by

$$\phi^{\mathrm{sc}}(x) := \sum_{i=1}^{r} \phi(\lambda_i(x))e_i(x), \quad \text{if } \lambda_i(x) \in \mathrm{dom}(\phi), \tag{1}$$

where $x \in \mathbb{V}$ has the following spectral decomposition $x = \sum_{i=1}^{r} \lambda_i(x)e_i(x)$, with $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_r(x)$. Then, we can consider its corresponding *spectrally defined* function $\Phi : \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$ by

$$\Phi(x) := \mathrm{tr}(\phi^{\mathrm{sc}}(x)) = \sum_{i=1}^{r} \phi(\lambda_i(x)), \quad \text{if } \lambda_i(x) \in \mathrm{dom}(\phi), \tag{2}$$

and $+\infty$ otherwise. Note that if $\phi : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function on a subset of $\mathrm{dom}(\phi)$, we can define the vector-valued function $(\phi')^{\mathrm{sc}} : \mathbb{V} \to \mathbb{V}$ by

$$(\phi')^{\mathrm{sc}}(x) := \sum_{i=1}^{r} \phi'(\lambda_i(x))e_i(x), \quad \text{if } \lambda_i(x) \in \mathrm{dom}(\phi'). \tag{3}$$

The following results give us some known properties of functions $\Phi$ and $\phi^{\mathrm{sc}}$. Their proofs can be found in [34, Theorem 38, Corollary 39, and Theorem 41] and [35, Theorem 3.2].

**Lemma 2.4:** *If the function $\phi$ is convex (resp. strictly convex) and continuous on its domain, then its spectrally defined function $\Phi$ is convex (resp. strictly convex) and continuous on its domain.*

**Lemma 2.5:** *Let $\phi$ be a continuously differentiable function on a subset of dom($\phi$) and $x \in \mathbb{V}$ with its spectral decomposition $x = \sum_{i=1}^{r} \lambda_i(x)e_i(x)$. Then*

(a) *$\Phi$ is continuously differentiable on* $\mathrm{dom}((\phi')^{\mathrm{sc}})$ *and*

$$\nabla\Phi(x) = \sum_{i=1}^{r} \phi'(\lambda_i(x))e_i(x) = (\phi')^{\mathrm{sc}}(x), \quad \forall x \in \mathrm{dom}((\phi')^{\mathrm{sc}}). \tag{4}$$

(b) *$\phi^{\mathrm{sc}}$ is continuously differentiable on* $\mathrm{dom}((\phi')^{\mathrm{sc}})$, *and*

$$\nabla\phi^{\mathrm{sc}}(x)h = \sum_{i=1}^{r} \phi'(\lambda_i(x))\langle e_i(x), h\rangle e_i(x) + 4 \sum_{1\leq i<j\leq r} [\lambda_i(x), \lambda_j(x)]_{\phi} e_i(x) \circ (e_j(x) \circ h), \tag{5}$$

*$\forall h \in \mathbb{V}$, where*

$$[\lambda_i(x), \lambda_j(x)]_{\phi} = \frac{\phi(\lambda_i(x)) - \phi(\lambda_j(x))}{\lambda_i(x) - \lambda_j(x)}, \quad \forall i, j = 1, \ldots, r \text{ and } i \neq j.$$

*In particular,*

$$\nabla\phi^{\mathrm{sc}}(x)x = \sum_{i=1}^{r} \phi'(\lambda_i(x))\lambda_i(x)e_i(x) = (\phi')^{\mathrm{sc}}(x) \circ x, \ \forall x \in \mathrm{dom}((\phi')^{\mathrm{sc}}). \tag{6}$$

Finally, we present the notion of essential smoothness [36], which will be used in this work. Let us denote by $\Gamma_0(\mathbb{X}) = \{f : \mathbb{X} \to \mathbb{R} \cup \{+\infty\} : f \text{ is convex, lsc and proper}\}$, with $\mathbb{X}$ denoting a Euclidean space.

**Definition 1:** A function $f \in \Gamma_0(\mathbb{V})$ is essentially smooth iff it is differentiable on $\text{int}(\text{dom}(f)) \neq \emptyset$ and $\|\nabla f(x_k)\| \to +\infty$ as $k \to +\infty$, whenever $\{x_k\}_{k\in\mathbb{N}} \subseteq \text{int}(\text{dom}(f))$ converges to boundary point of domain of $f$.

### 2.3. Proximal distances over symmetric cones

**Definition 2:** An extented-valued function $H : \mathbb{V} \times \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$ is called a proximal distance with respect to $\text{int}(\mathcal{K})$ iff it satisfies the following properties:

(P1) $\text{dom}(H(\cdot,\cdot)) = \mathcal{C}_1 \times \mathcal{C}_2$ with $\text{int}(\mathcal{K}) \times \text{int}(\mathcal{K}) \subseteq \mathcal{C}_1 \times \mathcal{C}_2 \subseteq \mathcal{K} \times \mathcal{K}$.
(P2) $H(u,v) \geq 0$ for all $u, v \in \mathbb{V}$, and $H(v,v) = 0$ for all $v \in \text{int}(\mathcal{K})$.
(P3) For each $v \in \text{int}(\mathcal{K})$, $H(\cdot,v)$ is continuous and strictly convex on $\mathcal{C}_1$, and it is continuously differentiable on $\text{int}(\mathcal{K})$ with $\text{dom}(\nabla_1 H(\cdot,v)) = \text{int}(\mathcal{K})$, where $\nabla_1 H(\cdot,v)$ denotes the gradient of $H(\cdot,v)$ with respect to the first variable.
(P4) For each $\gamma \in \mathbb{R}$, the level set $L_H(v,\gamma) = \{u \in \mathcal{C}_1 : H(u,v) \leq \gamma\}$ is bounded for any $v \in \mathcal{C}_2$.

This definition was considered in [6] in the context of second-order cones,[1] and it has a little difference from the one introduced by Auslender and Teboulle (see [8, Definition 2.1]), since here $H(\cdot,v)$ is required to be strictly convex over $\mathcal{C}_1$, for any fixed $v \in \text{int}(\mathcal{K})$.

Let us denote by $\mathcal{H}$ the family of functions $H$ satisfying (P1)–(P4). Note that by (P1) the function $H(\cdot,\cdot)$ is proper. From (P2) it follows that $H(\cdot,v)$ achieves its global minimum value at $v$. This implies that $\nabla_1 H(v,v) = 0$, for all $v \in \text{int}(\mathcal{K})$. Moreover, for any $f \in \Gamma_0(\mathbb{V})$, $\mu > 0$, and $\tilde{x} \in \text{int}(\mathcal{K})$, properties (P3) and (P4) are used to guarantee that the following convex optimization problem

$$\min\{f(u) + \mu H(u,\tilde{x}) : u \in \mathcal{K}\}$$

admits a unique solution $z_\mu(\tilde{x}) \in \text{int}(\mathcal{K})$ (see, e.g. Proposition 5.1).

## 3. Constructing proximal distances over symmetric cones

In this section, we provide two ways to construct a function $H$ in terms of some scalar-valued function $\phi$, and give some conditions on $\phi$ in order to show that $H$ be a proximal distance. Additionally, we discuss some properties of this distance.

### 3.1. Bregman's pseudo-distance

For the first approach, let us consider $\phi \in \Gamma_0(\mathbb{R})$ with $\text{dom}(\phi) \subseteq \mathbb{R}_+$, and $\text{int.dom}(\phi)) = \mathbb{R}_{++}$, and suppose the following conditions:

(S.1) $\phi$ is continuous and strictly convex on its domain.
(S.2) $\phi$ is continuously differentiable on $\mathbb{R}_{++}$.
(S.3) For each $\gamma \in \mathbb{R}$, the level set $\{t \in \mathbb{R}_{++} : d_\phi(s,t) \leq \gamma\}$ is bounded for any $s \in \text{dom}(\phi)$.
(S.4) If $\{t_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_{++}$ is such that $\lim_{k\to+\infty} t_k = 0$, then $\lim_{k\to+\infty} \phi'(t_k)(s - t_k) = -\infty$, $\forall s \in \mathbb{R}_{++}$,

where $d_\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$d_\phi(s,t) := \begin{cases} \phi(s) - \phi(t) - \phi'(t)(s-t), & s \in \text{dom}(\phi),\ t \in \mathbb{R}_{++}, \\ +\infty, & \text{otherwise}. \end{cases} \tag{7}$$

**Remark 1:** Note that from (S.1) to (S.2) and [38, Theorem 3.7(iii)], one has that the level set $\{s \in \text{dom}(\phi) : d_\phi(s,t) \leq \gamma\}$ is bounded for any $t \in \mathbb{R}_{++}$ and $\gamma \in \mathbb{R}$.

We denote by $\Sigma(\phi)$ the class of functions $\phi$ satisfying conditions (S.1)–(S.4). Given a function $\phi \in \Sigma(\phi)$, we define the following function $H : \mathbb{V} \times \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$ by

$$H(x,y) := \begin{cases} \Phi(x) - \Phi(y) - \langle \nabla\Phi(y), x - y\rangle, & x \in \mathrm{dom}(\Phi),\ y \in \mathrm{int}(\mathcal{K}), \\ +\infty, & \text{otherwise}, \end{cases} \tag{8}$$

where the functions $\Phi$ and $\nabla\Phi$ are defined by (2) and (4), respectively.

Next, we study some important properties of the function $H$ defined in (8). The first result extends [20, Proposition 3.1 and 3.3] to our context.

**Proposition 3.1:** *Given $\phi \in \Sigma(\phi)$, let $H$ be the function defined by (8). Then, the following results hold:*

(a) *$H(x,y) \geq 0$ for any $(x,y) \in \mathrm{dom}(\Phi) \times \mathrm{int}(\mathcal{K})$, and $H(x,y) = 0$ iff $x = y$.*
(b) *$H(\cdot,\cdot)$ is continuous on $\mathrm{dom}(\Phi) \times \mathrm{int}(\mathcal{K})$ and, for any $y \in \mathrm{int}(\mathcal{K})$, the function $H(\cdot,y)$ is strictly convex on $\mathrm{dom}(\Phi)$.*
(c) *For any fixed $y \in \mathrm{int}(\mathcal{K})$, $H(\cdot,y)$ is continuously differentiable on $\mathrm{int}(\mathcal{K})$, with*

$$\nabla_1 H(x,y) = \nabla\Phi(x) - \nabla\Phi(y) = (\phi')^{\mathrm{sc}}(x) - (\phi')^{\mathrm{sc}}(y). \tag{9}$$

(d) *$H(x,y) \geq \sum_{i=1}^{r} d_\phi(\lambda_i(x), \lambda_i(y)) \geq 0$, for any $(x,y) \in \mathrm{dom}(\Phi) \times \mathrm{int}(\mathcal{K})$.*
(e) *For all $\gamma \geq 0$, the level sets $L_H(y,\gamma) = \{x \in \mathrm{dom}(\Phi) : H(x,y) \leq \gamma\}$ and $L_H(x,\gamma) = \{y \in \mathrm{int}(\mathcal{K}) : H(x,y) \leq \gamma\}$ are bounded, for any fixed $y \in \mathrm{int}(\mathcal{K})$ and $x \in \mathrm{dom}(\Phi)$, respectively.*
(f) *For any $x, y \in \mathrm{int}(\mathcal{K})$ and $z \in \mathrm{dom}(\Phi)$, the following three-points identity holds:*

$$\langle \nabla_1 H(x,y), z - x\rangle = H(z,y) - H(z,x) - H(x,y). \tag{10}$$

(g) *For all sequence $\{z_k\}_{k\in\mathbb{N}} \subseteq \mathrm{int}(\mathcal{K})$ such that $\lim_{k\to+\infty} z_k = \bar{z} \in \mathrm{bd}(\mathcal{K})$ (boundary of $\mathcal{K}$), one has that*

$$\lim_{k\to+\infty} \langle \nabla\Phi(z_k), x - z_k\rangle = -\infty, \quad \forall x \in \mathrm{int}(\mathcal{K}). \tag{11}$$

*As consequence we obtain, for any fixed $y \in \mathrm{int}(\mathcal{K})$, that*

$$\lim_{k\to+\infty} \langle \nabla_1 H(z_k, y)), x - z_k\rangle = -\infty, \quad \forall x \in \mathrm{int}(\mathcal{K}). \tag{12}$$

***Proof*:**

(a) This statement follows directly from the strictly convex of the function $\Phi$ (cf. Lemma 2.4), and the definition of $H$.
(b) Since $\phi^{\mathrm{sc}}(x)$, $(\phi')^{\mathrm{sc}}(y)$, $(\phi')^{\mathrm{sc}}(y) \circ (x - y)$ are continuous functions for any $x \in \mathrm{dom}(\phi^{\mathrm{sc}})$, and $y \in \mathrm{int}(\mathcal{K})$ (by Lemma 2.4), and the trace function is also continuous, it follows that the function $H$ is continuous on $\mathrm{dom}(\Phi) \times \mathrm{int}(\mathcal{K})$. On the other hand, by using (S.1) and Lemma 2.4 one has that the function $\Phi$ is strictly convex on $\mathrm{dom}(\Phi)$. Moreover, the linear expression $-\Phi(y) - \langle \nabla\Phi(y), . - y\rangle$ is clearly convex on $\mathcal{K}$, for any fixed $y \in \mathrm{int}(\mathcal{K})$. Hence, $H(\cdot,y)$ is strictly convex on $\mathrm{dom}(\Phi)$ for any fixed $y \in \mathrm{int}(\mathcal{K})$.
(c) By (S.2) and Lemma 2.5(a) it follows that $H(\cdot,y)$ is continuously differentiable on $\mathrm{int}(\mathcal{K})$. The equality (9) is obtained directly from (8) and Lemma 2.5(a).

(d) Let $x = \sum_{i=1}^r \lambda_i(x)e_i(x)$ and $y = \sum_{i=1}^r \lambda_i(y)e_i(y)$ be the spectral decompositions of $x$ and $y$, respectively. Using the definition of $H$, it follows that for any $(x, y) \in \mathrm{dom}(\Phi) \times \mathrm{int}(\mathcal{K})$,

$$\begin{aligned} H(x,y) &= \Phi(x) - \Phi(y) - \mathrm{tr}(\nabla\Phi(y) \circ x) + \mathrm{tr}(\nabla\Phi(y) \circ y) \\ &\geq \Phi(x) - \Phi(y) - \sum_{i=1}^r \phi'(\lambda_i(y))\lambda_i(x) + \mathrm{tr}(\nabla\Phi(y) \circ y) \\ &= \sum_{i=1}^r [\phi(\lambda_i(x)) - \phi(\lambda_i(y)) - \phi'(\lambda_i(y))\lambda_i(x) + \phi'(\lambda_i(y))\lambda_i(y)] \\ &= \sum_{i=1}^r d_\phi(\lambda_i(x), \lambda_i(y)), \end{aligned}$$

where the inequality follows from (4) and *Von Neumann inequality* (cf. Lemma 2.2), the second equality is due to (4), Lemma 2.2, and (2), and the last one follows from (7). The non-negativity of $d_\phi(s,t)$ is due to the strict convexity of $\phi$ on its domain.

(e) For any fixed $y \in \mathrm{int}(\mathcal{K})$ and $\gamma \geq 0$, from part (d) we have that $L_H(y,\gamma) \subseteq \{x \in \mathrm{dom}(\Phi) : \sum_{i=1}^r d_\phi(\lambda_i(x), \lambda_i(y)) \leq \gamma\}$. By Remark 1, it follows that the set in the right-hand side is bounded. Then, $L_H(y,\gamma)$ is bounded for all $\gamma \geq 0$. Similarly, for fixed $x \in \mathrm{dom}(\Phi)$, one has that $L_H(x,\gamma) \subseteq \{y \in \mathrm{int}(\mathcal{K}) : \sum_{i=1}^r d_\phi(\lambda_i(x), \lambda_i(y)) \leq \gamma\}$. From (S.3), we deduce that the sets $L_H(x,\gamma)$ are bounded for all $\gamma \geq 0$.

(f) This statement can be easily verified using the definition of $H$.

(g) Let us suppose that $z_k$, $\bar{z}$, and $x$ have the following spectral decompositions $z_k = \sum_{i=1}^r \lambda_i(z_k) e_i(z_k)$, $\bar{z} = \sum_{i=1}^r \lambda_i(\bar{z})e_i(\bar{z})$, and $x = \sum_{i=1}^r \lambda_i(x)e_i(x)$. By (4) and (3), one has that

$$\langle \nabla\Phi(z_k), z_k \rangle = \sum_{i=1}^r \phi'(\lambda_i(z_k))\lambda_i(z_k).$$

Then, using this equality and the *Von Neumann inequality* (cf. Lemma 2.2) we obtain that

$$\langle \nabla\Phi(z_k), x - z_k \rangle \leq \sum_{i=1}^r \phi'(\lambda_i(z_k))(\lambda_i(x) - \lambda_i(z_k)). \tag{13}$$

On the other hand, as $x \in \mathrm{int}(\mathcal{K})$, and $\bar{z} \in \mathrm{bd}(\mathcal{K})$, one has that $\lambda_i(x) > 0$, for all $i = 1, \ldots, r$, and that there exists an $l \in \{1, \ldots, r\}$ such that $\lambda_i(\bar{z}) = 0$, for all $i = l, \ldots, r$, and $\lambda_i(\bar{z}) > 0$, for all $i = 1, \ldots, l-1$. Then, from (S.4) and the continuity of the $\lambda_i$ values, we have that

$$\lim_{k\to+\infty} \phi'(\lambda_i(z_k))(\lambda_i(x) - \lambda_i(z_k)) = -\infty, \quad \forall i = l, \ldots, r. \tag{14}$$

Since the expression $\lim_{k\to+\infty} \phi'(\lambda_i(z_k))(\lambda_i(x) - \lambda_i(z_k))$ is finite, for all $i = 1, \ldots, l-1$, the relation (11) follows from (13) and (14).

Finally, the equality (12) is obtained taking into account that

$$\langle \nabla_1 H(z_k, y)), x - z_k \rangle = \langle \nabla\Phi(z_k), x - z_k \rangle - \langle \nabla\Phi(y), x - z_k \rangle, \quad \forall x, y \in \mathrm{int}(\mathcal{K}),$$

that the first term of the right-hand side of last equality goes to $-\infty$, and the second one converges as $k \to +\infty$. □

**Remark 2:** A similar result to Part (g) of Proposition 3.1 can be found in [30, Lemma 3.2].

**Remark 3:** By Proposition 3.1(g) and [36, Lemma 26.2] it follows that the function $H(\cdot,y)$ is essentially smooth (cf. Definition 1) for any fixed $y \in \operatorname{int}(\mathcal{K})$ (likewise the function $\Phi$). Then, by [36, Theorem 26.1] one has that $\partial_x H(x,y) = \emptyset$ for all $x \in \operatorname{bd}(\mathcal{K})$, and $\partial_x H(x,y) = \{\nabla_1 H(x,y)\}$ for all $x \in \operatorname{int}(\mathcal{K})$, where $\partial_x H(\cdot,y)$ denotes the subdifferential[2] of $H(\cdot,y)$ with respect to the first variable. In particular, $\operatorname{dom}(\partial_x H(\cdot,y)) = \operatorname{int}(\mathcal{K})$ for any fixed $y \in \operatorname{int}(\mathcal{K})$.

Proposition 3.1 and Remark 3 show that the function $H$ defined by (8), with $\phi \in \Sigma(\phi)$, is a proximal distance with respect to $\operatorname{int}(\mathcal{K})$, and whose domain is $\operatorname{dom}(H) = \operatorname{dom}(\Phi) \times \operatorname{int}(\mathcal{K})$.

The following result is quite useful when we analyze convergence of proximal algorithms associated with a proximal distance $H$.

**Proposition 3.2:** *Let $H$ be the function defined by (8) with $\phi \in \Sigma(\phi)$. Consider the following statements:*

(B.1) *For any $\{v^k\}_{k\in\mathbb{N}} \subseteq \operatorname{int}(\mathcal{K})$ converging to $y^* \in \mathcal{K}$, we have that $H(y^*, v^k) \to 0$.*

(B.2) *For any $\{v^k\}_{k\in\mathbb{N}} \subseteq \operatorname{int}(\mathcal{K})$ such that $v^k \to v^* \in \operatorname{dom}(\Phi)$, and $c \in \mathcal{K}$, we have that $H(c, v^k) \to H(c, v^*)$.*

*The following implication holds:* (B.1) $\Rightarrow$ (B.2).

***Proof*:** Let $\{v^k\}_{k\in\mathbb{N}} \subseteq \operatorname{int}(\mathcal{K})$ be a sequence converging to $v^* \in \operatorname{dom}(\Phi)$, and let $c \in \mathcal{K}$. Suppose that $v^* \in \operatorname{bd}(\mathcal{K})$, and define $u = \theta v^k + (1-\theta)v^*$ with $\theta \in (0,1)$. Clearly $u \in \operatorname{int}(\mathcal{K})$. Since $\nabla\Phi$ is monotone (cf. (S1) and Lemma 2.4), one has

$$\theta\langle \nabla\Phi(u) - \nabla\Phi(v^k), u - v^k\rangle \geq 0.$$

Then,

$$\theta\langle \nabla\Phi(u) - \nabla\Phi(v^k), u - v^*\rangle \geq \theta\langle \nabla\Phi(u) - \nabla\Phi(v^k), v^k - v^*\rangle = \langle \nabla\Phi(u) - \nabla\Phi(v^k), u - v^*\rangle,$$

whence

$$\langle \nabla_1 H(u, v^k), u - v^*\rangle = \langle \nabla\Phi(u) - \nabla\Phi(v^k), u - v^*\rangle \leq 0.$$

Now, by using (10) at the points $z = v^*$, $y = v^k$, $x = u$, the above inequality implies that

$$H(v^*, u) + H(u, v^k) \leq H(v^*, v^k).$$

On the other hand, as $\Phi$ is essentially smooth (cf. Remark 3) from [38, Theorem 3.8] it follows that $\lim_{k\to+\infty} H(u, v^k) = 0$. Then, letting $k \to +\infty$ in the above inequality and using the assumption (B.1), we obtain a contradiction. Hence, $v^* \in \operatorname{int}(\mathcal{K})$. The result follows by the continuity of function $H(\cdot,\cdot)$ on $\operatorname{dom}(\Phi) \times \operatorname{int}(\mathcal{K})$ (cf. Proposition 3.1(b)). □

**Remark 4:** We should point out that the proximal distance $H$, with $\phi \in \Sigma(\phi)$ and $\operatorname{dom}(\phi) = \mathbb{R}_+$, generally does not satisfy the condition (B.1) of Proposition 3.2 (see [6, Example 4.7] for $\mathcal{K} = \mathcal{L}^n_+$, and [7, Example 4.1] for $\mathcal{K} = S^n_+$), even if $\phi$ satisfies the following condition:

(S.5) For any $\{t_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_{++}$ such that $\lim_{k\to+\infty} t_k = t^* \in \mathbb{R}_+$, one has $\lim_{k\to+\infty} d_\phi(t^*, t_k) = 0$.

Note that, in the case $\mathcal{K} = \mathbb{R}^n_+$, the KL relative entropy distance satisfies (B.1) (see [10, Lemma 2.1]), and hence, it does the same (B.2).

### 3.2. Distance-like functions

In this approach, we assume that $\mathbb{V}$ is a simple EJA. Let $\psi \in \Gamma_0(\mathbb{R})$ with $\operatorname{dom}(\psi) = \mathbb{R}_+$, and assume the following conditions [4,6]:

(C.1) $\psi$ is continuous and strictly convex on $\mathbb{R}_+$, and it is continuously differentiable on a subset of $\operatorname{dom}(\psi)$, where $\operatorname{dom}(\psi') \subseteq \operatorname{dom}(\psi)$ and $\operatorname{int}(\operatorname{dom}(\psi')) = \mathbb{R}_{++}$.

(C.2) $\psi$ is twice continuously differentiable on $\mathbb{R}_{++}$ and $\lim_{t\to 0^+} \psi''(t) = +\infty$.
(C.3) $\xi(t) := \psi'(t)t - \psi(t)$ is convex on $\mathrm{dom}(\psi')$, and $\psi'$ is strictly concave on $\mathrm{dom}(\psi')$.
(C.4) $(\psi')^{\mathrm{sc}}$ defined in (3) is concave with respect to $\mathcal{K}$, that is, for all $x, y \in \mathrm{dom}((\psi')^{\mathrm{sc}})$, and $\beta \in [0,1]$, one has

$$(\psi')^{\mathrm{sc}}(\beta x + (1-\beta)y) - \beta(\psi')^{\mathrm{sc}}(x) - (1-\beta)(\psi')^{\mathrm{sc}}(y) \in \mathcal{K}.$$

**Remark 5:** The notion of concavity with respect to symmetric cone $\mathcal{K}$ (cf. (C.4)) was recently introduced in [39, Definition 1.2(b)] for a simple EJA. This notion extends [40, Definition 3.1(b)] given for second-order cones.

In the sequel, we denote by $\widehat{\Sigma}(\psi)$ the class of functions $\psi$ satisfying the assumptions (C.1)–(C.4). Given $\psi \in \widehat{\Sigma}(\psi)$, let $\Psi$ be its spectrally defined function (cf. (2)), and $\nabla\Psi$ be the gradient of that $\Psi$ defined in (4). With such function, we can define $\widehat{H} : \mathbb{V} \times \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$ by

$$\widehat{H}(x,y) := \begin{cases} \Psi(y) - \Psi(x) - \langle \nabla\Psi(x), y - x\rangle, & x \in \mathrm{dom}((\psi')^{\mathrm{sc}}),\ y \in \mathcal{K}, \\ +\infty, & \text{otherwise.} \end{cases} \tag{15}$$

The following result extends [6, Lemma 4.2] to our context.

**Lemma 3.3:** *Let $\psi \in \widehat{\Sigma}(\psi)$. Then*

(a) *The function $\Xi(x) := \langle \nabla\Psi(x), x\rangle - \Psi(x) = \mathrm{tr}(\xi^{\mathrm{sc}}(x))$, with $\xi^{\mathrm{sc}}(x) = (\psi')^{\mathrm{sc}}(x) \circ x - (\psi)^{\mathrm{sc}}(x)$, is convex in* $\mathrm{dom}((\psi')^{\mathrm{sc}})$ *and continuously differentiable on* $\mathrm{int}(\mathcal{K})$.
(b) *For any fixed $y \in \mathbb{V}$, the function $\rho_y(x) := \langle \nabla\Psi(x), y\rangle = \langle (\psi')^{\mathrm{sc}}(x), y\rangle$ is continuously differentiable on* $\mathrm{int}(\mathcal{K})$ *with $\nabla\rho_y(x) = \nabla(\psi')^{\mathrm{sc}}(x)y$, and moreover, it is strictly concave over* $\mathrm{dom}((\psi')^{\mathrm{sc}})$ *whenever $y \in \mathrm{int}(\mathcal{K})$.*

***Proof:***

(a) By (C.3), $\xi(t)$ is convex on $\mathrm{dom}(\psi')$. Then, from Lemma 2.4 one obtains that $\Xi(x)$ is convex in $\mathrm{dom}((\psi')^{\mathrm{sc}})$. For the other part, we note that $\psi$ and $\psi'$ are continuously differentiable on $\mathbb{R}_{++}$ by (C.2). Then, by using Lemma 2.5(b) we obtain that $(\psi)^{\mathrm{sc}}$ and $(\psi')^{\mathrm{sc}}$ are continuously differentiable on $\mathrm{int}(\mathcal{K})$. Hence, Lemma 2.5(a) implies the result.
(b) Clearly the function $\rho_y(x)$, for any fixed $y \in \mathbb{V}$, is continuously differentiable on $\mathrm{int}(\mathcal{K})$ by (C.2) and Lemma 2.5(b). Moreover, by applying the chain rule for inner product of two functions, one obtains that $\nabla\rho_y(x) = \nabla(\psi')^{\mathrm{sc}}(x)y$.

Now, we prove the second part. By (C.3) and Lemma 2.4, we have that the function $\mathrm{tr}((\psi')^{\mathrm{sc}}(x))$ is strictly concave on $\mathrm{dom}((\psi')^{\mathrm{sc}})$, that is, for any $x, z \in \mathrm{dom}((\psi')^{\mathrm{sc}})$ with $x \neq z$, and $\beta \in (0,1)$ one has that

$$\mathrm{tr}((\psi')^{\mathrm{sc}}(\beta x + (1-\beta)z)) > \beta\mathrm{tr}((\psi')^{\mathrm{sc}}(x)) + (1-\beta)\mathrm{tr}((\psi')^{\mathrm{sc}}(z)).$$

This implies that $(\psi')^{\mathrm{sc}}(\beta x + (1-\beta)z) - \beta(\psi')^{\mathrm{sc}}(x) - (1-\beta)(\psi')^{\mathrm{sc}}(z) \neq 0$. Then, from this relation, (C.4), and Lemma 2.3 we have that

$$\langle (\psi')^{\mathrm{sc}}(\beta x + (1-\beta)z) - \beta(\psi')^{\mathrm{sc}}(x) - (1-\beta)(\psi')^{\mathrm{sc}}(z), y\rangle > 0, \quad \forall y \in \mathrm{int}(\mathcal{K}).$$

This shows that the function $\rho_y(x)$ is strictly concave over $\mathrm{dom}((\psi')^{\mathrm{sc}})$, for any fixed $y \in \mathrm{int}(\mathcal{K})$. □

The next results present some properties of the function $\widehat{H}$. They are related to [4, Propositions 4.1 and 4.2] given for second-order cones.

**Proposition 3.4:** *Let $\psi \in \widehat{\Sigma}(\psi)$ and $\widehat{H}$ be defined by (15). Then,*

(a) *$\widehat{H}(x,y) \geq 0$ for any $(x,y) \in \mathrm{dom}((\psi')^{\mathrm{sc}}) \times \mathcal{K}$, and $\widehat{H}(x,y) = 0$ iff $x = y$.*

(b) $\widehat{H}(\cdot,\cdot)$ *is continuous on* $\mathrm{dom}((\psi')^{\mathrm{sc}}) \times \mathcal{K}$ *and, for any fixed* $y \in \mathrm{int}(\mathcal{K})$, *the function* $\widehat{H}(\cdot, y)$ *is strictly convex on* $\mathrm{dom}((\psi')^{\mathrm{sc}})$.

(c) *For any fixed* $y \in \mathcal{K}$, $\widehat{H}(\cdot, y)$ *is continuously differentiable on* $\mathrm{int}(\mathcal{K})$ *with*

$$\nabla_1 \widehat{H}(x, y) = \nabla(\psi')^{\mathrm{sc}}(x)(x - y). \tag{16}$$

(d) $\widehat{H}(x, y) \geq \sum_{i=1}^{r} d_\psi(\lambda_i(y), \lambda_i(x)) \geq 0$, *for any* $(x, y) \in \mathrm{dom}((\psi')^{\mathrm{sc}}) \times \mathcal{K}$, *where* $d_\psi$ *is defined in* (7).

(e) *For all* $\gamma \geq 0$, *the level sets* $L_{\widehat{H}}(y, \gamma) = \{x \in \mathrm{dom}((\psi')^{\mathrm{sc}}) : \widehat{H}(x, y) \leq \gamma\}$ *and* $L_{\widehat{H}}(x, \gamma) = \{y \in \mathcal{K} : \widehat{H}(x, y) \leq \gamma\}$ *are bounded, for any fixed* $y \in \mathcal{K}$ *and* $x \in \mathrm{dom}((\psi')^{\mathrm{sc}})$, *respectively.*

***Proof:***

(a) This statement follows directly from the strictly convex of the function $\Psi$ on $\mathrm{dom}((\psi')^{\mathrm{sc}})$ (cf. Lemma 2.4), and the definition of $\widehat{H}$.

(b) Since $\psi^{\mathrm{sc}}(y)$, $(\psi')^{\mathrm{sc}}(x)$, $(\psi')^{\mathrm{sc}}(x) \circ (y - x)$ are continuous functions for any $x \in \mathrm{dom}((\psi')^{\mathrm{sc}})$, and $y \in \mathcal{K}$ (by Lemma 2.4), and the trace function is also continuous, it follows that the function $\widehat{H}$ is continuous on $\mathrm{dom}((\psi')^{\mathrm{sc}}) \times \mathcal{K}$. On the other hand, from Lemma 3.3 it follows that $\langle \nabla\Psi(x), x\rangle - \Psi(x)$ is convex on $\mathrm{dom}((\psi')^{\mathrm{sc}})$, and that $\langle \nabla\Psi(x), y\rangle$ is strictly concave on $\mathrm{dom}((\psi')^{\mathrm{sc}})$, for any fixed $y \in \mathrm{int}(\mathcal{K})$. Hence, $\widehat{H}(\cdot, y)$ defined in (15) is strictly convex on $\mathrm{dom}((\psi')^{\mathrm{sc}})$, for any fixed $y \in \mathrm{int}(\mathcal{K})$.

(c) By (C.1) and Lemma 2.5(a) it follows that $\widehat{H}(\cdot, y)$ is continuously differentiable on $\mathrm{int}(\mathcal{K})$. The equality (16) is obtained directly from applying the chain rule for inner product in (15), and Lemma 3.3(b).

(d) Let $x = \sum_{i=1}^{r} \lambda_i(x) e_i(x)$ and $y = \sum_{i=1}^{r} \lambda_i(y) e_i(y)$ be the spectral decompositions of $x$ and $y$, respectively. Using the definition of $\widehat{H}$, it follows that for any $(x, y) \in \mathrm{dom}((\psi')^{\mathrm{sc}}) \times \mathcal{K}$,

$$\begin{aligned}
\widehat{H}(x, y) &= \Psi(y) - \Psi(x) - \mathrm{tr}(\nabla\Psi(x) \circ y) + \mathrm{tr}(\nabla\Psi(x) \circ x) \\
&\geq \Psi(y) - \Psi(x) - \sum_{i=1}^{r} \psi'(\lambda_i(x))\lambda_i(y) + \mathrm{tr}(\nabla\Psi(x) \circ x) \\
&= \sum_{i=1}^{r} [\psi(\lambda_i(y)) - \psi(\lambda_i(x)) - \psi'(\lambda_i(x))(\lambda_i(y) - \lambda_i(x))] \\
&= \sum_{i=1}^{r} d_\psi(\lambda_i(y), \lambda_i(x)),
\end{aligned}$$

where the inequality follows from (4) and *Von Neumann inequality* (cf. Lemma 2.2), the second equality is due to (2), (4) and Lemma 2.2, and the last one follows from (7). The non-negativity of $d_\psi(s, t)$ is due to the strict convexity of $\psi$ on its domain.

(e) For any fixed $y \in \mathcal{K}$ and $\gamma \geq 0$, from part (d) we have that $L_{\widehat{H}}(y, \gamma) \subseteq \{x \in \mathrm{dom}((\psi')^{\mathrm{sc}}) : \sum_{i=1}^{r} d_\psi(\lambda_i(y), \lambda_i(x)) \leq \gamma\}$. Since the set $\{t \in \mathrm{dom}(\psi') : d_\psi(s, t) \leq 0\}$ is equal to $\{s\}$ or $\emptyset$, one has that it is bounded, for any fixed $s \geq 0$. Then, from the convexity of $\xi(t) - \psi'(t)s$ (cf. (C.3)) and [36, Corollary 8.7.1] it follows that the level sets $\{t \in \mathrm{dom}(\psi') : d_\psi(s, t) \leq \gamma\}$, for any fixed $s \geq 0$, are bounded. Hence, $L_{\widehat{H}}(y, \gamma)$ is bounded for all $\gamma \geq 0$. Similarly, for fixed $x \in \mathrm{dom}((\psi')^{\mathrm{sc}})$, the sets $L_{\widehat{H}}(x, \gamma)$ are bounded for all $\gamma \geq 0$. □

**Proposition 3.5:** *Suppose that* $\psi \in \widehat{\Sigma}(\psi)$. *Then, for any fixed* $y \in \mathrm{int}(\mathcal{K})$, *the function* $\widehat{H}(\cdot, y)$, *defined by* (15) *is essentially smooth.*

***Proof:*** Let $\{x_k\}_{k\in\mathbb{N}} \subseteq \mathrm{int}(\mathcal{K})$, with $x_k \to x \in \mathrm{bd}(\mathcal{K})$. In order to show that $\widehat{H}(\cdot, y)$ is essentially smooth we prove that $\|\nabla_1 \widehat{H}(x_k, y))\|_F \to +\infty$, for any fixed $y \in \mathrm{int}(\mathcal{K})$. From (16) and (6), one has that

$$\nabla_1 \widehat{H}(x_k, y) = (\psi'')^{\rm sc}(x_k) \circ x_k - \nabla(\psi')^{\rm sc}(x_k)(y).$$

Moreover, if $x_k = \sum_{i=1}^r \lambda_i(x_k)e_i(x_k)$ is the spectral decomposition of $x_k$, the use of the relations (3) and (5) in the above equality implies that

$$\begin{aligned}\nabla_1 \widehat{H}(x_k, y) = &\sum_{i=1}^r \psi''(\lambda_i(x_k))(\lambda_i(x_k) - \langle e_i(x_k), y\rangle)e_i(x_k)\\ &+ 4 \sum_{1\le i<j\le r} [\lambda_i(x_k), \lambda_j(x_k)]_{\psi'} e_i(x_k) \circ (e_j(x_k) \circ y),\end{aligned}$$

where

$$[\lambda_i(x_k), \lambda_j(x_k)]_{\psi'} = \frac{\psi'(\lambda_i(x_k)) - \psi'(\lambda_j(x_k))}{\lambda_i(x_k) - \lambda_j(x_k)}, \quad \forall i, j = 1, \dots, r \text{ and } i \neq j.$$

Let $x = \sum_{i=1}^r \lambda_i(x)e_i(x)$ be the spectral decomposition of $x$. Since $x \in {\rm bd}(\mathcal{K})$, there exists an $l \in \{1, \dots, r\}$ such that $\lambda_i(x) = 0$, for $i = l, \dots, r$ and $\lambda_i(x) > 0$, for $i = 1, \dots, l-1$. Then, by using the Cauchy–Schwartz inequality it follows that

$$\|\nabla_1 \widehat{H}(x_k, y))\|_F \ge |\langle \nabla_1 \widehat{H}(x_k, y)), e_\ell(x_k)\rangle| = |\psi''(\lambda_\ell(x_k))||\lambda_\ell(x_k) - \langle e_\ell(x_k), y\rangle|, \ \forall \ell \ge l.$$

From Lemma 2.3 one has that $\langle e_\ell(x_k), y\rangle > 0$, $\forall k \in \mathbb{N}$. Then, as $\lim_{k\to+\infty} \lambda_\ell(x_k) = \lambda_\ell(x) = 0$, $\forall \ell \ge l$ (by continuity of the eigenvalues), the term $|\lambda_\ell(x_k) - \langle e_\ell(x_k), y\rangle|$ converges to a finite value as $k \to +\infty$. Since $\lim_{k\to+\infty} \psi''(\lambda_\ell(x_k)) = +\infty$ (cf. condition (C.2)), one has that $\|\nabla_1 \widehat{H}(x_k, y))\|_F \to +\infty$ as $k \to +\infty$. □

Propositions 3.4–3.5 show that the function $\widehat{H}$ defined by (15), with $\psi \in \widehat{\Sigma}(\psi)$, is a proximal distance with respect to ${\rm int}(\mathcal{K})$, and whose domain is ${\rm dom}(\widehat{H}) = {\rm dom}(\Psi) \times \mathcal{K}$.

The following result gives some properties about the distance $\widehat{H}$, which are important for proximal algorithm associated with this distance.

**Proposition 3.6:** *Let $\psi \in \widehat{\Sigma}(\psi)$ and $\widehat{H}$ be defined by (15). Then,*

(a) *For all $x, y \in {\rm int}(\mathcal{K})$ and $z \in \mathcal{K}$, we have that*

$$\langle \nabla_1 \widehat{H}(y, x), z - y\rangle \le \widehat{H}(x, z) - \widehat{H}(y, z). \tag{17}$$

(b) *If ${\rm dom}(\psi) = {\rm dom}(\psi') = \mathbb{R}_+$, then $\widehat{H}$ satisfies the following condition:*
  (D) *For any $\{v^k\}_{k\in\mathbb{N}} \subseteq {\rm int}(\mathcal{K})$ converging to $v^* \in \mathcal{K}$, we have that $\widehat{H}(v^k, v^*) \to 0$.*

***Proof:***

(a) From the definition of $\widehat{H}$, $\Xi$ and $\rho_z(\cdot)$ (defined in Lemma 3.3), we have for any $x, y \in {\rm int}(\mathcal{K})$ and $z \in \mathcal{K}$ that

$$\begin{aligned}\widehat{H}(x, z) - \widehat{H}(y, z) &= \Xi(x) - \Xi(y) + \rho_z(y) - \rho_z(x)\\ &\ge \langle \nabla\Xi(y), x - y\rangle + \langle \nabla\rho_z(y), y - x\rangle\\ &= \langle \nabla\Xi(y) - \nabla(\psi')^{\rm sc}(y)z, x - y\rangle,\end{aligned} \tag{18}$$

where the inequality follows from the convexity of $\Xi$ and the strictly concave of $\rho_z$ (see Lemma 3.3), and the last equality is due to Lemma 3.3(b). On the other hand, first, we note from Lemma 3.3(a) that $\Xi$ is the spectrally defined function of $\xi$, then by Lemma 2.5(a) and (3) we obtain that

$$\nabla\Xi(y) = (\xi')^{\rm sc}(y) = (\psi'')^{\rm sc}(y) \circ y.$$

Second, by Lemma 2.5(b) one has that $\nabla(\psi')^{sc}(y)y = (\psi'')^{sc}(y)\circ y$. Then, from both equalities we have that

$$\nabla(\psi')^{sc}(y)z = \nabla(\psi')^{sc}(y)(z-y) + \nabla(\psi')^{sc}(y)y = \nabla(\psi')^{sc}(y)(z-y) + \nabla\Xi(y).$$

Using this relation in (18) we obtain

$$\widehat{H}(x,z) - \widehat{H}(y,z) \geq \langle \nabla(\psi')^{sc}(y)(y-z), x-y\rangle = \langle \nabla(\psi')^{sc}(y)(y-x), z-y\rangle,$$

where the equality is due to symmetry of $\nabla(\psi')^{sc}(y)$. Finally, the result follows using (16).

(b) By assumption, we have that $\operatorname{dom}((\psi')^{sc}) = \mathcal{K}$. Then, $\widehat{H}(\cdot, v^*)$ is continuous on $\mathcal{K}$ by Proposition 3.4(b). Thus, the result follows.

□

## 4. Examples of proximal distances

In this section, we present several examples of proximal distances (cf. Definition 2) taking into account the two classes of real-valued functions considered in Section 3.

### 4.1. Examples of Bregman's pseudo-distance

In the first four examples, we give an explicit expression for the proximal distance $H$, while the last example presents a way to obtain functions $\phi$ satisfying conditions (S1)–(S4), and hence a proximal distance $H$ (this example is motivated by [41, Theorem 2]).

**Example 4.1 (Entropy-like proximal distance):** Let $\phi(t) = t\ln(t) - t$, if $t \geq 0$ (with the convention $0\ln(0) = 0$), and $\phi(t) = +\infty$, if $t < 0$. It is easy to check that $\phi \in \Sigma(\phi)$ with $\operatorname{dom}(\phi) = \mathbb{R}_+$. Let $x = \sum_{i=1}^r \lambda_i(x)e_i(x)$ be the spectral decomposition of $x$. Then,

$$\ln(x) = \sum_{i=1}^r \ln(\lambda_i(x))e_i(x), \quad \forall x \in \operatorname{int}(\mathcal{K}).$$

Thus, the spectrally defined function associated to $\phi$ is $\Phi(x) = \operatorname{tr}(x\circ\ln(x) - x)$, if $x \in \mathcal{K}$. By Lemma 2.5(a), we have that $\nabla\Phi(x) = \ln(x)$, for $x \in \operatorname{int}(\mathcal{K})$. Then, the function $H : \mathbb{V}\times\mathbb{V} \to \mathbb{R}\cup\{+\infty\}$ is given by

$$H(x,y) = \begin{cases} \operatorname{tr}(x\circ\ln(x) - x\circ\ln(y) + y - x), & \forall x \in \mathcal{K}, y \in \operatorname{int}(\mathcal{K}),\\ +\infty, & \text{otherwise.}\end{cases}$$

It follows from Proposition 3.1 and Remark 3 (see also [20, Propositions 3.1 and 4.1]) that $H$ is a proximal distance with $\mathcal{C}_1 = \mathcal{K}$ and $\mathcal{C}_2 = \operatorname{int}(\mathcal{K})$.

(a) Note that when $\mathbb{V} = \mathbb{R}^n$, $\mathcal{K} = \mathbb{R}^n_+$, the function $H$ has the form

$$H(x,y) = \begin{cases} \sum_{i=1}^n (x_i\ln(x_i/y_i) + y_i - x_i), & \forall x \in \mathbb{R}^n_+, y \in \mathbb{R}^n_{++},\\ +\infty, & \text{otherwise,}\end{cases}$$

which is the so-called KL relative entropy distance [18].

(b) For $\mathbb{V} = \mathbb{R}^n$, $\mathcal{K} = \mathcal{L}^n_+$, the function $H$ has the form

$$H(x,y) = \begin{cases} \operatorname{tr}(x\circ\ln(x) - x\circ\ln(y) + y - x), & \forall x \in \mathcal{L}^n_+, y \in \mathcal{L}^n_{++},\\ +\infty, & \text{otherwise.}\end{cases}$$

(c) Let $\mathcal{S}^n$ be the set of all $n \times n$ real symmetric matrices, and $\mathcal{K} = \mathcal{S}^n_+$ be the cone of $n \times n$ symmetric positive semi-definite matrices. In this case, the function $H$ has the form

$$H(X,Y) = \begin{cases} \operatorname{tr}(X \ln (X - Y) + Y - X), & \forall X \in \mathcal{S}^n_+, Y \in \mathcal{S}^n_{++}, \\ +\infty, & \text{otherwise.} \end{cases}$$

**Example 4.2:** For $p \in (0,1)$, let us consider the family of functions $\phi(t) = (pt - t^p)/(1-p)$ if $t \geq 0$, and $\phi(t) = +\infty$ if, $t < 0$. It follows from [9, Example 3.1(3)] and [8, Example 3.1] that $\phi \in \Sigma(\phi)$ with $\operatorname{dom}(\phi) = \mathbb{R}_+$. Its spectrally defined function associated is $\Phi(x) = \frac{1}{1-p}\operatorname{tr}(px - x^p)$, for $x \in \mathcal{K}$. By Lemma 2.5(a), we have that $\nabla\Phi(x) = \frac{p}{1-p}(e - x^{p-1})$, for $x \in \operatorname{int}(\mathcal{K})$. Thus, we obtain that

$$H(x,y) = \begin{cases} \frac{1}{1-p}\operatorname{tr}((1-p)y^p + (py^{p-1} - x^{p-1}) \circ x), & \forall x \in \mathcal{K}, y \in \operatorname{int}(\mathcal{K}), \\ +\infty, & \text{otherwise.} \end{cases}$$

From Proposition 3.1 and Remark 3, we obtain that $H$ is a proximal distance with $\mathcal{C}_1 = \mathcal{K}$ and $\mathcal{C}_2 = \operatorname{int}(\mathcal{K})$.

Note that when $\mathbb{V} = \mathbb{R}^n$ and $\mathcal{K} = \mathbb{R}^n_+$, the proximal distance $H$ takes the form

$$H(x,y) = \begin{cases} \sum_{i=1}^n (y_i^{p-1}(y_i + qx_i) - (1+q)x_i^p), & \forall x \in \mathbb{R}^n_+, y \in \mathbb{R}^n_{++}, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $q = \frac{p}{1-p}$.

**Example 4.3:** Take $\phi(t) = pt^q - qt^p - (p-q)$ if $t \geq 0$, and $\phi(t) = +\infty$, if $t < 0$, where $0 < p < 1$, $q \geq 1$ (see [42, Example 2.1]). It is easy to show that $\phi \in \Sigma(\phi)$ with $\operatorname{dom}(\phi) = \mathbb{R}_+$, and that its spectrally defined function associated is given by $\Phi(x) = \operatorname{tr}(px^q - qx^p) - (p-q)r$, for $x \in \mathcal{K}$. By Lemma 2.5(a) one has that $\nabla\Phi(x) = pq(x^{q-1} - x^{p-1})$, for $x \in \operatorname{int}(\mathcal{K})$. Then, the function $H : \mathbb{V} \times \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$ takes the form

$$H(x,y) = \begin{cases} \operatorname{tr}(p(x^q + (q-1)y^q) - q(x^p + (p-1)y^p) & \\ -pqx \circ (y^{q-1} - y^{p-1})), & \forall x \in \mathcal{K}, y \in \operatorname{int}(\mathcal{K}), \\ +\infty, & \text{otherwise.} \end{cases}$$

This function is a proximal distance with $\mathcal{C}_1 = \mathcal{K}$ and $\mathcal{C}_2 = \operatorname{int}(\mathcal{K})$ (cf. Proposition 3.1 and Remark 3). Note that when $p = \frac{1}{2}$ and $q = 1$, the proximal distance takes the form

$$H(x,y) = \begin{cases} \operatorname{tr}(\frac{1}{2}y^{1/2} - x^{1/2} + \frac{1}{2}x \circ y^{-\frac{1}{2}}), & \forall x \in \mathcal{K}, y \in \operatorname{int}(\mathcal{K}), \\ +\infty, & \text{otherwise.} \end{cases}$$

**Example 4.4 (Log-barrier proximal distance):** Let $\phi(t) = -\ln(t)$, for $t > 0$. It is easy to verify that $\phi \in \Sigma(\phi)$ with $\operatorname{dom}(\phi) = \mathbb{R}_{++}$ (see for instance [8, Example 3.1]), and that its spectrally defined function associated is $\Phi(x) = -\operatorname{tr}(\ln(x)) = -\ln(\det(x))$, for $x \in \operatorname{int}(\mathcal{K})$. By Lemma 2.5 it follows that $\nabla\Phi(x) = -x^{-1}$, for $x \in \operatorname{int}(\mathcal{K})$. Then, one has that

$$H(x,y) = \begin{cases} \operatorname{tr}(\ln(y) - \ln(x) + y^{-1} \circ x) - r, & \forall x, y \in \operatorname{int}(\mathcal{K}), \\ +\infty, & \text{otherwise.} \end{cases}$$

This function extends the log-barrier proximal distance given in [7,8,43] to our context. Moreover, using Proposition 3.1 and Remark 3, one has that $H$ is a proximal distance with $\mathcal{C}_1 = \mathcal{C}_2 = \operatorname{int}(\mathcal{K})$.

**Example 4.5:** Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a continuous and strictly convex function with $\mathrm{dom}(\varphi) = \mathbb{R}_+$. Assume that $\varphi$ is continuously differentiable on $\mathbb{R}_{++}$, and that it satisfies the following condition:

$$\lim_{t \downarrow 0} t\varphi'(t) = 0, \quad \lim_{t \downarrow 0} \varphi'(t) = -\infty. \tag{19}$$

Then, the function $\phi$ defined by $\phi(t) = \varphi(t) + |t|^p$, with $p > 1$ fixed belongs to $\Sigma(\phi)$. Indeed, $\phi$ satisfies conditions (S.1), (S.2), and (S.4). Condition (S.3) follows from [44, Proposition 2]. Therefore, under the above assumptions, if $\Phi$ denotes its corresponding spectrally defined function, the function $H$ defined by (8) is a proximal distance with $\mathcal{C}_1 = \mathcal{K}$ and $\mathcal{C}_2 = \mathrm{int}(\mathcal{K})$.

**Remark 6:** Let $\phi : \mathbb{R} \to \mathbb{R}$, and let us consider the following assumption (see [44])

**(K)** For each $s \in \mathbb{R}_+$, there exist constants $\alpha(s) > 0$, $c(s)$ such that

$$d_\phi(s,t) + c(s) \geq \alpha(s)|s - t|, \quad \forall t \in \mathbb{R}_{++}.$$

If $\phi$ satisfies the conditions of Example 4.5 and the condition **(K)** holds, one has that $\phi \in \Sigma(\phi)$. Some examples of $\phi$ for which the condition **(K)** holds are [44]:

- $\phi_1(t) = t\ln(e^t - 1)$,
- $\phi_2(t) = |t|^p$, with $p > 1$,
- $\phi_3(t) = t\ln(t) - t$.

### *4.2. Examples of distance-like functions*

**Example 4.6:** Let $\psi(t) = \frac{t^{q+1}}{q+1}$, if $t \geq 0$, and $\psi(t) = +\infty$, if $t < 0$, where $q \in (0,1)$. It is easy to show that $\psi$ satisfies (C.1)–(C.3) with $\mathrm{dom}(\psi) = \mathrm{dom}(\psi') = \mathbb{R}_+$, where $\psi'(t) = t^q$. Moreover, since $\Psi(x) = \frac{1}{q+1}\mathrm{tr}(x^{q+1})$ and $\nabla\Psi(x) = (\psi')^{\mathrm{sc}}(x) = x^q$, for $x \in \mathcal{K}$, we can use the same arguments of [40, Proposition 3.7] to prove that (C.4) holds (see also [39, Example 3.9(ii)]). Then, the function $\widehat{H}$ given by

$$\widehat{H}(x,y) = \begin{cases} \frac{1}{q+1}\mathrm{tr}(y^{q+1}) + \frac{q}{q+1}\mathrm{tr}(x^{q+1}) - \mathrm{tr}(x^q \circ y), & \forall x,y \in \mathcal{K}, \\ +\infty, & \text{otherwise}, \end{cases}$$

is a proximal distance with $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{K}$.

**Example 4.7:** Let $\psi(t) = t\ln(t) - t + 1$, if $t \geq 0$, and $\psi(t) = +\infty$, if $t < 0$. It is easy to verify that $\psi$ satisfies conditions (C.1)–(C.3) with $\mathrm{dom}(\psi) = \mathbb{R}_+$, and $\mathrm{dom}(\psi') = \mathbb{R}_{++}$, where $\psi'(t) = \ln(t)$. Moreover, by [39, Example 3.9(i)] condition (C.4) also holds. Since $\Psi(x) = \mathrm{tr}(x \circ \ln(x) - x + e)$, for $x \in \mathcal{K}$, and $\nabla\Psi(x) = \ln(x)$, for $x \in \mathrm{int}(\mathcal{K})$, the function $\widehat{H}$ given by

$$\widehat{H}(x,y) = \begin{cases} \mathrm{tr}(y \circ \ln(y) - y \circ \ln(x) + x - y), & \forall x \in \mathrm{int}(\mathcal{K}),\ y \in \mathcal{K}, \\ +\infty, & \text{otherwise}, \end{cases}$$

is a proximal distance with $\mathcal{C}_1 = \mathrm{int}(\mathcal{K})$, and $\mathcal{C}_2 = \mathcal{K}$.

**Example 4.8:** Let $\psi(t) = t\ln(t) + (t+1)\ln(t+1) + t^{q+1}$, if $t \geq 0$, and $\psi(t) = +\infty$, if $t < 0$, where $q \in (0,1)$. Clearly, $\psi$ satisfies conditions (C.1)–(C.3) with $\mathrm{dom}(\psi) = \mathbb{R}_+$, and $\mathrm{dom}(\psi') = \mathbb{R}_{++}$, where $\psi'(t) = \ln(t) + \ln(t+1) + 2 + (q+1)t^q$. Moreover, from [39, Example 3.9] it follows that $\psi$ satisfies condition (C.4). Since $\Psi(x) = \mathrm{tr}(x \circ \ln(x) + (x+e) \circ \ln(x+e) + x^{q+1})$, for $x \in \mathcal{K}$, and $\nabla\Psi(x) = \ln(x) + \ln(x+e) + 2e + (q+1)x^q$, for $x \in \mathrm{int}(\mathcal{K})$, the function $\widehat{H}$ given by

$$\widehat{H}(x,y) = \begin{cases} \mathrm{tr}(y \circ (\ln(y) - \ln(x)) + (y+e) \circ \ln(y+e) & \\ -(y+e) \circ \ln(x+e) + qx^{q+1} + y^{q+1} & \\ -2(y-x) - (1+q)x^q \circ y), & \forall x \in \mathrm{int}(\mathcal{K}),\ y \in \mathcal{K}, \\ +\infty, & \text{otherwise}, \end{cases}$$

is a proximal distance with $\mathcal{C}_1 = \mathrm{int}(\mathcal{K})$, and $\mathcal{C}_2 = \mathcal{K}$.

**Example 4.9:** Let $\psi_1(t) = \psi(t) + \frac{\nu}{2}t^2$ with $\psi \in \widehat{\Sigma}(\psi)$ and $\nu > 0$. Clearly, $\psi_1$ satisfies (C.1)–(C.4) with $\mathrm{dom}(\psi_1) = \mathrm{dom}(\psi)$, and $\mathrm{dom}(\psi_1') = \mathrm{dom}(\psi')$. Since $\Psi_1(x) = \Psi(x) + \frac{\nu}{2}\mathrm{tr}(x^2)$, for $x \in \mathrm{dom}(\Psi)$, and $\nabla\Psi_1(x) = \nabla\Psi(x) + \nu x$, for $x \in \mathrm{dom}((\psi')^{\mathrm{sc}})$, the function $\widehat{H}_\nu$ is defined by

$$\widehat{H}_\nu(x,y) = \widehat{H}(x,y) + \frac{\nu}{2}\|x-y\|^2, \tag{20}$$

with $\widehat{H}$ given by (15) is a proximal distance with $\mathcal{C}_1 = \mathrm{dom}((\psi')^{\mathrm{sc}})$, and $\mathcal{C}_2 = \mathrm{dom}(\Psi)$.

**Remark 7:** Note that the functions $\psi$ considered in Examples 4.6–4.9 satisfy condition (C.4), for any simple EJA $\mathbb{V}$ except for the algebra of all $3 \times 3$ octonion Hermitian matrices (denoted by $\mathbb{O}^3$). Two reasons why condition (C.4) fails for $\mathbb{O}^3$ are explained in [39]. For instance, one of them is that it seems impossible to embed $\mathbb{O}^3$ into some $\mathcal{S}^n$. Without considering this fact, the relation between concavity with respect to $\mathcal{K}$ and matrix concavity cannot be used. The last fact is employed in the proof of [39, Theorem 3.5]. Hence, Examples 4.6–4.9 are valid for any simple EJA except for $\mathbb{O}^3$.

## 5. Study of proximal-type algorithms and primal central paths in SCP problem

We consider the following convex SCP problem

$$f_* = \min\{f(x) : x \in \mathcal{V} \cap \mathcal{K}\}, \tag{SCP}$$

where $f \in \Gamma_0(\mathbb{V})$, $\mathcal{V} = \{x \in \mathbb{V} : \mathcal{A}x = b\}$ with $\mathcal{A} : \mathbb{V} \to \mathbb{R}^m$ a surjective linear map, and $b \in \mathbb{R}^m$. Let us denote by $\mathcal{A}^*$ the adjoint of the linear map $\mathcal{A}$, by $\mathcal{F} = \mathcal{V} \cap \mathcal{K}$, $\mathcal{F}^0 = \mathcal{V} \cap \mathrm{int}(\mathcal{K})$, and $\mathcal{F}_* = \{x \in \mathcal{F} : f(x) = f_*\}$ the feasible, the interior feasible set, and the optimal set of problem (SCP), respectively, and by $\mathrm{dom}(f) = \{x \in \mathbb{V} : f(x) < +\infty\}$ the domain of $f$.

In this section, first we study the convergence of proximal-type algorithms for solving problem (SCP). Then, we define the primal central path for (SCP) and study its asymptotic behaviour.

From now on, we suppose the following assumptions on the problem (SCP):

(A1) $f_* > -\infty$;
(A2) $\mathrm{dom}(f) \cap \mathcal{F}^0 \neq \emptyset$.

### 5.1. Proximal-type algorithms

We consider the following interior proximal algorithm with proximal distance (IPAPD) for solving problem (SCP): **Algorithm IPAPD:** Given $H \in \mathcal{H}$, $x^0 \in \mathcal{F}^0$, and scalars $\gamma_k > 0$, $\varepsilon_k \geq 0$, for $k = 1, 2, \dots$, we generate iteratively the sequence $\{x^k\} \subset \mathcal{F}^0$ such that

$$g^k \in \partial_{\varepsilon_k} f(x^k), \tag{21}$$
$$\gamma_k g^k + \nabla_1 H(x^k, x^{k-1}) = \mathcal{A}^* s^k, \tag{22}$$

for some sequence $\{s^k\}$ in $\mathbb{R}^m$.

The following result shows that the algorithm IPAPD is well-defined. Its proof follows with the same arguments as the ones given in the proof of [6, Proposition 3.1] (see also [8, Proposition 2.1]).

**Proposition 5.1:** *Let $H \in \mathcal{H}$. Then, for any $x^{k-1} \in \mathcal{F}^0$, $\gamma_k > 0$, and $\varepsilon_k \geq 0$ $\forall k \geq 0$, there exists a unique point $x^k \in \mathcal{F}^0$ satisfying (21) and (22), for some $s^k \in \mathbb{R}^m$. Moreover, for such $x^k$ one has*

$$\gamma_k f(x^k) + H(x^k, x^{k-1}) \leq f_*(x^{k-1}, \gamma_k) + \varepsilon,$$

*where $f_*(x^{k-1}, \gamma_k) = \min\{\gamma_k f(x) + H(x, x^{k-1}) : x \in \mathcal{V}\}$.*

Let us denote by $\mathcal{H}_1$ (resp. $\mathcal{H}_2$) the family of functions $H$ defined by (8) (resp. (15)) with $\phi \in \Sigma(\phi)$ (resp. $\psi \in \widehat{\Sigma}(\psi)$). Clearly, Proposition 5.1 holds when $H \in \mathcal{H}_1$ or $H \in \mathcal{H}_2$.

Next, we establish the convergence of the algorithm IPAPD under limit points. The proof of this result repeats word by word the arguments in [8, Theorem 2.1].

**Theorem 5.2:** *Let $\{x^k\}$ be the sequence generated by the algorithm IPAPD with $H \in \mathcal{H}_1$ or $H \in \mathcal{H}_2$, and let $\sigma_n = \sum_{k=1}^n \gamma_k$. Then, the following results hold:*

(a) $f(x^n) - f(x) \le \sigma_n^{-1}(H(x, x^0) + \sum_{k=1}^n \sigma_k \varepsilon_k)$, *for any* $x \in \mathcal{V} \cap \operatorname{dom}(\Phi)$, *if* $H \in \mathcal{H}_1$;
(b) $f(x^n) - f(x) \le \sigma_n^{-1}(H(x^0, x) + \sum_{k=1}^n \sigma_k \varepsilon_k)$, *for any* $x \in \mathcal{V} \cap \operatorname{dom}((\psi')^{\mathrm{soc}})$, *if* $H \in \mathcal{H}_2$;
(c) *If* $\sigma_n \to +\infty$ *and* $\varepsilon_k \to 0$, *then* $\liminf_{n\to+\infty} f(x^n) = f_*$;
(d) *The sequence* $\{f(x^k)\}$ *converges to* $f_*$ *whenever* $\sum_{k=1}^\infty \varepsilon_k < +\infty$;
(e) *Suppose that* $\mathcal{F}_* \ne \emptyset$ *and consider the following two cases:*
  (i) $\mathcal{F}_*$ *is bounded, and* $\sum_{k=1}^\infty \varepsilon_k < +\infty$;
  (ii) $\sum_{k=1}^\infty \varepsilon_k \gamma_k < +\infty$, *and* $\operatorname{dom}(\phi) = \mathbb{R}_+$ *(or* $\operatorname{dom}((\psi')) = \mathbb{R}_+$*).*
  *Then, under either* (i) *or* (ii) *the sequence* $\{x^k\}$ *is bounded, and each limit point belongs to* $\mathcal{F}_*$.

**Remark 8:** Note that Theorem 5.2 holds for all Examples given in previous section. In particular, for $\phi$ given in Example 4.1, we recover [20, Proposition 4.3].

As consequence of Theorem 5.2, the following result indicates an estimate for the global rate of convergence for the algorithm IPAPD.

**Corollary 5.3:** *Let $\{x^k\}$ be the sequence generated by the algorithm IPAPD with $H \in \mathcal{H}_1$ or $H \in \mathcal{H}_2$. If $\mathcal{F}_* \ne \emptyset$, $\operatorname{dom}(\phi) = \mathbb{R}_+$ (or $\operatorname{dom}((\psi')) = \mathbb{R}_+$), and $\sum_{k=1}^\infty \varepsilon_k < +\infty$, then we have the estimate $f(x^n) - f_* = O(\sigma_n^{-1})$.*

***Proof*:** Under the given hypothesis, Parts (a) and (b) of Theorem 5.2 hold, for any $x = x^* \in \mathcal{F}_*$. Since $0 < \sigma_k \le \sigma_n$, $\sum_{k=1}^\infty \varepsilon_k < +\infty$, the result follows. □

To establish the global convergence of $\{x^k\}$ to an optimal solution of problem (SCP), we need some conditions on the proximal distance $H \in \mathcal{H}$ (see, e.g. [8]):

(E.1) For any sequence bounded $\{v^k\}_{k\in\mathbb{N}} \subseteq \operatorname{int}(\mathcal{K})$ and any $v^* \in \mathcal{K}$ with $H(v^*, v^k) \to 0$, one has $v^k \to v^*$.
(E.2) For any sequence bounded $\{v^k\}_{k\in\mathbb{N}} \subseteq \operatorname{int}(\mathcal{K})$ and any $v^* \in \mathcal{K}$ with $H(v^k, v^*) \to 0$, one has $v^k \to v^*$.

**Theorem 5.4:** *Let $\{x^k\}$ be the sequence generated by the algorithm IPAPD with $H \in \mathcal{H}_1$ or $H \in \mathcal{H}_2$, and let $\sigma_n = \sum_{k=1}^n \gamma_k$. Suppose that $\mathcal{F}_* \ne \emptyset$, $\sigma_n \to +\infty$, $\sum_{k=1}^\infty \varepsilon_k < +\infty$, and $\sum_{k=1}^\infty \varepsilon_k \gamma_k < +\infty$. If*

(a) (B.1), (E.1) *hold, and* $\operatorname{dom}(\phi) = \mathbb{R}_+$, *when* $H \in \mathcal{H}_1$ , *or*
(b) (E.2) *holds and* $\operatorname{dom}(\psi) = \operatorname{dom}((\psi')) = \mathbb{R}_+$, *when* $H \in \mathcal{H}_2$,

*then the sequence $\{x^k\}$ converges to an optimal solution of problem* (SCP).

***Proof*:** The proof of case (a) is similar to that in [8, Theorem 2.2]. For completeness, we only consider the case (b). From (21)–(22) it follows that

$$\gamma_k(f(x^k) - f(x)) \le \langle \nabla_1 H(x^k, x^{k-1}), x - x^k \rangle + \gamma_k \varepsilon_k, \quad \forall x \in \mathcal{F}.$$

Using (17) with $y = x^k$, $x = x^{k-1}$ and $z = x$, the above inequality implies that

$$\gamma_k(f(x^k) - f(x)) \le H(x^{k-1}, x) - H(x^k, x) + \gamma_k \varepsilon_k, \quad \forall x \in \mathcal{F}. \tag{23}$$

Taking $x = x^* \in \mathcal{F}_*$ in (23), we get

$$H(x^k, x^*) \le H(x^{k-1}, x^*) + \gamma_k \varepsilon_k. \tag{24}$$

Since $\sum_{k=1}^{\infty} \varepsilon_k \gamma_k < +\infty$, it follows from [8, Lemma 2.1] that the sequence $\{H(x^k, x^*)\}$ converges to some $l(x^*) \in \mathbb{R}_+$. On the other hand, since the sequence $\{x^k\}$ is bounded (cf. Theorem 5.2(e)), there exists a subsequence $\{x^{k_j}\}$ converging to some optimal solution $x_\infty$ (cf. Theorem 5.2(e)). Then, using Proposition 3.6(b), we have $H(x^{k_j}, x_\infty) \to 0$, which implies that $H(x^k, x_\infty) \to 0$. Finally, by (E.2) it follows that the sequence $\{x^k\}$ converges to $x_\infty$. □

**Remark 9:** Note that it is slightly difficult to build proximal distances in our context, such that (B.1) holds (see Remark 4). However, it is not difficult to build proximal distances such that either (E.1) or (E.2) holds. For instance, the regularized proximal distance

$$H_\nu(x, y) = H(x, y) + \frac{\nu}{2}\|x - y\|^2, \quad \nu > 0,$$

with $H \in \mathcal{H}_1$ and $\mathrm{dom}(\phi) = \mathbb{R}_+$, (resp. $H \in \mathcal{H}_2$ and $\mathrm{dom}(\psi) = \mathbb{R}_+$) satisfies (E.1) (resp. (E.2)).

### 5.2. Primal central paths with proximal distance

Let $H$ be a proximal distance and $\bar{x} \in \mathrm{int}(\mathcal{K})$. The *primal central path* for problem (SCP) with respect to $H(\cdot, \bar{x})$ is the set $\{x(\mu) : \mu > 0\}$, where $x(\mu)$ is solution of problem

$$\min\{f(x) + \mu H(x, \bar{x}) : x \in \mathcal{F}\}, \quad \mu > 0. \tag{25}$$

**Remark 10:** Note that when $f$ is linear and $H$ is the Entropy-like proximal distance (cf. Example 4.1), the primal central path was studied in [21]. For $f$ convex and continuously differentiable, in [30] the authors studied this central path for a class spectral barrier function instead of proximal distance $H$.

The next result establishes that the primal central path is well defined. Its proof follows immediately from Definition 2 and [8, Proposition 2.1].

**Proposition 5.5:** *For any $H \in \mathcal{H}$ and $\bar{x} \in \mathrm{int}(\mathcal{K})$, the primal central path $\{x(\mu) : \mu > 0\}$ with respect to $H(\cdot, \bar{x})$ is well defined, belong to $\mathcal{F}^0$, and for each $\mu > 0$, it is the unique solution of*

$$g_\mu + \mu \nabla_1 H(x(\mu), \bar{x}) = \mathcal{A}^* y(\mu), \tag{26}$$

*for some $g_\mu \in \partial f(x(\mu))$ and $y(\mu) \in \mathbb{R}^m$.*

**Remark 11:** Since $\mathcal{A}$ is surjective we have the relationship $y(\mu) = (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}(g_\mu - \mu \nabla_1 H(x(\mu), \bar{x}))$.

The following result is related to the boundedness of the primal central path (for small valued of $\mu$). Its proof is similar to the arguments given in Propositions 3–5 in [45].

**Proposition 5.6:** *For any $H \in \mathcal{H}$ and $\bar{x} \in \mathrm{int}(\mathcal{K})$, the following assertions hold true:*

(a) *The function $H(x(\mu), \bar{x})$ is non-increasing in $\mu$.*
(b) *The function $f(x(\mu))$ is non-decreasing in $\mu$.*
(c) *Suppose that $\mathcal{F}_* \neq \emptyset$. If*
  (i) *$\mathcal{F}_*$ is bounded or*
  (ii) $\mathrm{dom}(H(\cdot, \bar{x})) = \mathcal{K}$,
  *then the set $\{x(\mu) : 0 < \mu < \bar{\mu}\}$ is bounded for any $\bar{\mu} > 0$, and all their limit points are optimal solutions of* (SCP).

The next result shows the convergence of the primal central path, and provides a characterization of its limit points when $\mathrm{dom}(H(\cdot, \bar{x})) = \mathcal{K}$. It slightly generalizes [45, Proposition 6] to our context (see also e.g. [30, Theorem 5.1]).

**Theorem 5.7:** *Suppose that $\mathcal{F}_* \neq \emptyset$. Then, for any $\bar{x} \in \mathrm{int}(\mathcal{K})$ and $H \in \mathcal{H}$ with $\mathrm{dom}(H(\cdot, \bar{x})) = \mathcal{K}$, the primal central path w.r.t. $H(\cdot, \bar{x})$ converges, when $\mu \to 0$, toward the unique optimal solution of* $\min\{H(x, \bar{x}) : x \in \mathcal{F}_*\}$.

In the following result, we present a natural extension to linear SCP of [45, Theorem 3], which provides a connection between the proximal sequence and the primal central path.

**Theorem 5.8:** *Suppose that $f(x) = c^\top x$ with $c \in \mathbb{V}$. Let $\{x^k\}$ be the sequence generated by the algorithm IPAPD with $H \in \mathcal{H}$, $x^0 \in \mathcal{F}^0$ and $\varepsilon_k \equiv 0$, and $\{x(\mu) : \mu > 0\}$ be the primal central path w.r.t. $H(\cdot, x^0)$. If*

(a) *$H \in \mathcal{H}_1$ and $\{\mu_k\}$ is defined as $\mu_k = (\sum_{j=1}^k \gamma_j)^{-1}$, for $k = 1, 2, \ldots$, or*

(b) *$H \in \mathcal{H}_2$, $\nabla(\psi')^{\mathrm{sc}}(x)(\mathbb{V}) = \mathrm{Im}(\mathcal{A}^*)$ for any $x \in \mathrm{int}(\mathcal{K})$, and $\{\mu_k\}$ is defined as $\mu_k = \gamma_k^{-1}$, for $k = 1, 2, \ldots$,*

*then $x^k = x(\mu_k)$ for $k = 1, 2, \ldots$. Moreover, for each positive decreasing sequence $\{\mu_k\}$, there exists a positive sequence $\{\gamma_k\}$ with $\sum_{k=1}^\infty \gamma_k = +\infty$ such that the proximal sequence $\{x^k\}$, with regularization parameter $\gamma_k$, satisfies $x(\mu_k) = x^k$.*

***Proof*:** The proof of case (a) is similar to that in [45, Theorem 3]. For completeness, we only consider the case (b).

From (22) and Proposition 3.4(c), we have that the sequence $\{x^j\}$ satisfies

$$\gamma_j c + \nabla(\psi')^{\mathrm{sc}}(x^j)(x^j - x^{j-1}) = \mathcal{A}^* s^j, \tag{27}$$

for some $s^j \in \mathbb{R}^m$. Since $\psi''(t) > 0$, for all $t \in \mathbb{R}_{++}$ (cf. (C.1)–(C.2)), from [46, Theorem 3.2(b)] it follows that $\nabla(\psi')^{\mathrm{sc}}(x)$ is positive definite on $\mathrm{int}(\mathcal{K})$. Then, (27) can be written as

$$[\nabla(\psi')^{\mathrm{sc}}(x^j)]^{-1}\gamma_j c + (x^j - x^{j-1}) = [\nabla(\psi')^{\mathrm{sc}}(x^j)]^{-1}\mathcal{A}^* s^j.$$

Summing this equality from $j = 1$ to $k$ and making a suitable arrangement, we obtain

$$\gamma_k c + \nabla(\psi')^{\mathrm{sc}}(x^k)(x^k - x^0) = \mathcal{A}^* s^k + \nabla(\psi')^{\mathrm{sc}}(x^k)\sum_{j=1}^{k-1}[\nabla(\psi')^{\mathrm{sc}}(x^j)]^{-1}(\mathcal{A}^* s^j - \gamma_j c).$$

Taking $\mu_k = \gamma_k^{-1}$ and using the fact that $\nabla(\psi')^{\mathrm{sc}}(x^j)(\mathbb{V}) = \mathrm{Im}(\mathcal{A}^*)$, the above equality it reduced to

$$\mathcal{A}^* y^k = c + \mu_k \nabla(\psi')^{\mathrm{sc}}(x^k)(x^k - x^0) = c + \mu_k \nabla_1 H(x^k, x^0),$$

for some $y^k \in \mathbb{R}^m$. Thus, from this equality and (26) we deduce that $x^k = x(\mu_k)$. Now, let $\{x(\mu) : \mu > 0\}$ be the primal central path w.r.t $H(\cdot, x^0)$ and let $\{y(\mu) : \mu > 0\}$ be given in Remark 11. Take a positive decreasing sequence $\{\mu_k\}$ and define the sequences $x^k = x(\mu_k)$, $y^k = y(\mu_k)$. It follows from (26), Proposition 3.4(c), and positive definiteness of $\nabla(\psi')^{\mathrm{sc}}(\,\cdot\,)$ on $\mathrm{int}(\mathcal{K})$ that

$$[\nabla(\psi')^{\mathrm{sc}}(x^k)]^{-1} c\mu_k^{-1} + (x^k - x^0) = [\nabla(\psi')^{\mathrm{sc}}(x^k)]^{-1}\mathcal{A}^*(y^k \mu_k^{-1}).$$

From this equality, we have that

$$[\nabla(\psi')^{\mathrm{sc}}(x^k)]^{-1}(c\mu_k^{-1} - \mathcal{A}^*(y^k\mu_k^{-1})) + (x^k - x^{k-1}) = [\nabla(\psi')^{\mathrm{sc}}(x^{k-1})]^{-1}(c\mu_{k-1}^{-1} - \mathcal{A}^*(y^{k-1}\mu_{k-1}^{-1})),$$

whence

$$\begin{aligned} c\mu_k^{-1} + \nabla(\psi')^{\mathrm{sc}}(x^k)(x^k - x^{k-1}) = {}& \mathcal{A}^*(y^k\mu_k^{-1}) \\ & + \nabla(\psi')^{\mathrm{sc}}(x^k)[\nabla(\psi')^{\mathrm{sc}}(x^{k-1})]^{-1}(c\mu_{k-1}^{-1} - \mathcal{A}^*(y^{k-1}\mu_{k-1}^{-1})). \end{aligned}$$

Letting $\gamma_k = \mu_k^{-1}$ and using the fact that $\nabla(\psi')^{\mathrm{sc}}(x^j)(\mathbb{V}) = \mathrm{Im}(\mathcal{A}^*)$, from the last equality we obtain that

$$\mathcal{A}^* s^k = c\gamma_k + \nabla(\psi')^{\mathrm{sc}}(x^k)(x^k - x^{k-1}) = c\gamma_k + \nabla_1 H(x^k, x^{k-1}), \tag{28}$$

for some $s^k \in \mathbb{R}^m$. Since $\{\gamma_k\}$ is a positive increasing sequence, it satisfies $\sum_{k=1}^{\infty} \gamma_k = +\infty$. Then, from (28), we deduce that $\{x^k\}$ is the proximal sequence generated by the algorithm IPAPD with $H \in \mathcal{H}_2$. □

## 6. Conclusions

In this paper, we have provided two ways to construct a proximal distance with respect to the interior of the symmetric cone $\mathcal{K}$ of the EJA. This distance has been generated by a real-valued function, and under some mild assumptions, we have showed that it is a proximal one. In addition, several examples and properties of this distance have been presented. As application of this proximal distance, we have studied the convergence of proximal-type algorithms for solving convex SCP problems. Moreover, we have analysed the asymptotic behaviour of the primal central path in this context. Finally, for linear SCP, we have established the relations between the primal central paths and the sequence generated by the proximal algorithm IPAPD.

## Notes

1. Recall that a second-order cone [37] is defined by the set $\mathcal{L}_+^n := \{x = (x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} : \|\bar{x}\| \le x_1\}$.
2. Recall that the $\varepsilon$-subdifferential of a $f \in \Gamma_0(\mathbb{V})$ at $x$ is defined by $\partial_\varepsilon f(x) = \{g \in \mathbb{V} : f(x) + \langle g, z - x\rangle - \varepsilon \le f(z),\ \forall z \in \mathbb{V}\}$, for some $\varepsilon \ge 0$, and the subdifferential by $\partial f = \partial_0 f$.

## Acknowledgements

The authors are grateful to the anonymous referees for their careful reading and helpful suggestions that improved the paper greatly.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This research was supported by CONICYT-Chile, via FONDECYT [project 1160894] (Julio López); Postdoctoral Scholarship CAPES-FAPERJ Edital [PAPD-2011] (Erik Alex Papa).

## References

[1] Martinet B. Régularisation d'inquations variationelles par approximations successives. Reveu d' Automatique Informatique et Reserche Opérationelle. 1970;4(R3):154–158.
[2] Censor Y, Zenios S. The proximal minimization algorithm with d-functions. J Optim Theory Appl. 1992;73:451–464.
[3] Chen G, Teboulle M. Convergence analysis of a proximal-like minimization algorithm using bregman's function. SIAM J Optim. 1993;3(3):538–543.
[4] Pan S, Chen J. A class of interior proximal-like algorithms for convex second-order cone programming. SIAM J Optim. 2008;19(2):883–910.
[5] Pan S, Chen J. Proximal-like algorithm using the quasi D-function for convex second-order cone programming. J Optim Theory Appl. 2008;138:95–113.
[6] Pan S, Chen J. Interior proximal methods and central paths for convex second-order cone programming. Nonlinear Anal: Theory Methods Appl. 2010;73(9):3083–3100.
[7] Doljansky M, Teboulle M. An interior proximal algorithm and the exponential multiplier method for semidefinite programming. SIAM J Optim. 1998;9(1):1–13.

[8] Auslender A, Teboulle M. Interior gradient and proximal methods for convex and conic optimization. SIAM J Optim. 2006;16(3):697–725.
[9] Teboulle M. Entropic proximal mappings with applications to nonlinear programming. Math Oper Res. 1992;17(3):670–690.
[10] Teboulle M. Convergence of proximal-like algorithms. SIAM J Optim. 1997;7(4):1069–1083.
[11] Auslender A, Teboulle M, Ben-Tiba S. Interior proximal and multiplier methods based on second order homogeneous functionals. Math Oper Res. 1999;24(3):645–668.
[12] Auslender A, Silva P, Teboulle M. Nonmonotone projected gradient methods based on barrier and euclidean distances. Comput Optim Appl. 2007;38:305–327.
[13] Auslender A, Teboulle M. Projected subgradient methods with non-euclidean distances for non-differentiable convex minimization and variational inequalities. Math Program. 2009;120(1):27–48.
[14] Villacorta K, Oliveira P. An interior proximal method in vector optimization. Eur J Oper Res. 2011;214(3):485–492.
[15] Papa Quiroz E, Oliveira P. An extension of proximal methods for quasiconvex minimization on the nonnegative orthant. Eur J Oper Res. 2012;216:26–32.
[16] Papa Quiroz E, Mallma Ramirez L, Oliveira P. An inexact proximal method for quasiconvex minimization. Eur J Oper Res. 2015;246:721–729.
[17] Sarmiento O. Papa Quiroz E, Oliveira P. A proximal multiplier method for separable convex minimization. Optimization. 2016;65(2):501–537.
[18] Liese F, Vajda I. Convex statistical distances. Leipzig: Teubner; 1987.
[19] Faraut J, Korányi A. Analysis on symmetric cones. New York (NY): The Clarendon Press, Oxford University Press, Oxford Science Publications; 1994. (Oxford mathematical monographs).
[20] Chen J, Pan S. An entropy-like proximal algorithm and the exponential multiplier method for convex symmetric cone programming. Comput Optim Appl. 2010;47(3):477–499.
[21] Yu Z, Zhu Y, Cao Q. On the convergence of central path and generalized proximal point method for symmetric cone linear programming. Appl Math Inf Sci. 2013;7:2327–2333.
[22] Faybusovich L. Euclidean jordan algebras and interior-point algorithms. Positivity. 1997;1(4):331–357.
[23] Faybusovich F. Linear systems in jordan algebras and primal-dual interior-point algorithms. J Comput Appl Math. 1997;86(1):149–175.
[24] Schmieta S, Alizadeh F. Extension of primal-dual interior point algorithms to symmetric cones. Math Program. 2003;96(3):409–438.
[25] Gu G, Zangiabadi M, Roos C. Full nesterov-todd step infeasible interior-point method for symmetric optimization. Eur J Oper Res. 2011;214(3):473–484.
[26] Wang G, Bai Y. A new full nesterov-todd step primal-dual path- following interior-point algorithm for symmetric optimization. J Optim Theory Appl. 2012;154(3):966–985.
[27] Valkonen T. Extension of primal-dual interior point methods to diff-convex problems on symmetric cones. Optimization. 2013;62(3):345–377.
[28] Wang G, Yu C, Teo K. A new full nesterov-todd step feasible interior-point method for convex quadratic symmetric cone optimization. Appl Math Comput. 2013;221:329–343.
[29] Wang G, Zhang Z, Zhu D. On extending primal-dual interior-point method for linear optimization to convex quadratic symmetric cone optimization. Numer Funct Anal Optim. 2013;34(5):576–603.
[30] Ramírez H, Sossa D. On the central paths in symmetric cone programming. J Optim Theory Appl. in press 2016.
[31] Roos C, Terlaky T, Vial J. Theory and algorithms for linear optimization: an interior-point approach. New York (NY): Wiley; 1997. (Wiley-Interscience series in discrete mathematics and optimization).
[32] Alvarez F, López J. Interior proximal bundle algorithm with variable metric for nonsmooth convex symmetric cone programming. Optimization. 2016;65(9):1757–1779.
[33] Alvarez F, López J, Ramírez CH. Interior proximal algorithm with variable metric for second-order cone programming: applications to structural optimization and support vector machines. Optim Methods Softw. 2010;25(6):859–881.
[34] Baes M. Convexity and differentiability properties of spectral functions and spectral mappings on euclidean jordan algebras. Linear Algebra Appl. 2007;422(2–3):664–700.
[35] Sun D, Sun J. Löwner's operator and spectral functions in euclidean jordan algebras. Math Oper Res. 2008;33(2):421–445.
[36] Rockafellar R. Convex analysis. Princeton (NJ): Princeton University Press; 1970. (Princeton mathematical series).
[37] Alizadeh F, Goldfarb D. Second-order cone programming. Math. Program. 2003;95:3–51.
[38] Bauschke H, Borwein J. Legendre functions and the method of random bregman projections. J Convex Anal. 1997;4(1):27–68.
[39] Chang Y, Chen J, Pan S. Symmetric cone monotone functions and symmetric cone convex functions. J Nonlinear Convex Anal. 2016;17(3):499–512.
[40] Chen J. The convex and monotone functions associated with second-order cone. Optimization. 2006;55(4):363–385.

[41] Kaplan A, Tichatschke R. Interior proximal method for variational inequalities on non-polyhedral sets. Discussiones Mathematicae: Differ Inclusions Control Optim. 2007;27(1):71–93.
[42] Iusem A, Svaiter B, Teboulle M. Entropy-like proximal methods in convex programming. Math Oper Res. 1994;19(4):790–814.
[43] Csiszar I. An axiomatic approach to inference for linear inverse problems. Ann. Stat. 1991;19(4):2032–2066.
[44] Kaplan A, Tichatschke R. On inexact generalized proximal methods with a weakened error tolerance criterion. Optimization. 2004;53(1):3–17.
[45] Iusem A, Svaiter B, da Cruz Neto J. Central paths, generalized proximal point methods and cauchy trajectories in riemannian manifolds. SIAM J Control Optim. 1999;37(2):566–588.
[46] Chang Y, Chen J. Convexity of symmetric cone trace functions in euclidean jordan algebras. J Nonlinear Convex Anal. 2013;14(1):53–61.