Nonlinear Analysis 73 (2010) 3083–3100

Contents lists available at ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

Interior proximal methods and central paths for convex second-order cone programming

Shaohua Pan [a], Jein-Shan Chen [b,∗,1]

[a] Department of Mathematics, South China University of Technology, Guangzhou 510640, China

[b] Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan

ARTICLE INFO

Article history:
Received 18 November 2008
Accepted 28 June 2010

Keywords:
Convex second-order cone optimization
Interior proximal methods
Proximal distances with respect to SOCs
Convergence
Central path

ABSTRACT

We make a unified analysis of interior proximal methods of solving convex second-order cone programming problems. These methods use a proximal distance with respect to second-order cones which can be produced with an appropriate closed proper univariate function in three ways. Under some mild conditions, the sequence generated is bounded with each limit point being a solution, and global rates of convergence estimates are obtained in terms of objective values. A class of regularized proximal distances is also constructed which can guarantee the global convergence of the sequence to an optimal solution. These results are illustrated with some examples. In addition, we also study the central paths associated with these distance-like functions, and for the linear SOCP we discuss their relations with the sequence generated by the interior proximal methods. From this, we obtain improved convergence results for the sequence for the interior proximal methods using a proximal distance continuous at the boundary of second-order cones.

1. Introduction

We consider the following convex second-order cone programming problem (CSOCP):

$$\begin{aligned}&\inf f(x)\\&\text{s.t. } Ax=b,\quad x\succeq_{\mathcal{K}}0,\end{aligned}\tag{1}$$

where $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is a closed proper convex function, A is an $m\times n$ matrix with full row rank m, b is a vector in $\mathbb{R}^m$, $x\succeq_{\mathcal{K}}0$ means $x\in\mathcal{K}$, and $\mathcal{K}$ is the Cartesian product of some second-order cones (SOCs), also called Lorentz cones [1]. In other words,

$$\mathcal{K}=\mathcal{K}^{n_1}\times\mathcal{K}^{n_2}\times\cdots\times\mathcal{K}^{n_r}\tag{2}$$

where $r,n_1,\ldots,n_r\geq 1$ with $n_1+\cdots+n_r=n$, and

$$\mathcal{K}^{n_i}:=\left\{(x_1,x_2)\in\mathbb{R}\times\mathbb{R}^{n_i-1}\mid x_1\geq\|x_2\|\right\}$$

with $\|\cdot\|$ being the Euclidean norm. When f reduces to a linear function, i.e. $f(x)=c^Tx$ for some $c\in\mathbb{R}^n$, (1) becomes the standard SOCP. Throughout this paper, we denote by X_* the optimal set of (1), and let $\mathcal{V}:=\{x\in\mathbb{R}^n\mid Ax=b\}$. The

∗ Corresponding author. Tel.: +886 2 29325417; fax: +886 2 29332342.
E-mail addresses: shhpan@scut.edu.cn (S. Pan), jschen@math.ntnu.edu.tw (J.-S. Chen).

1 Member of the Mathematics Division, National Center for Theoretical Sciences, Taipei Office.

0362-546X/$ – see front matter
doi:10.1016/j.na.2010.06.079

CSOCP, as an extension of the standard SOCP, has a wide range of applications from engineering, control, and finance to robust optimization and combinatorial optimization; see [2,3] and references therein.

There have proposed various methods for solving the CSOCP, which include the interior point methods [4–6], the smoothing Newton methods [7,8], the smoothing–regularization method [9], the semismooth Newton method [10], and the merit function method [11]. These methods are all developed by reformulating the KKT optimality conditions as a system of equations or an unconstrained minimization problem. This paper will focus on an iterative scheme which is proximal based and handles directly the CSOCP itself. Specifically, the proximal-type algorithm consists of generating a sequence $\{x^k\}$ via

$$x^k := \operatorname{argmin}\left\{\lambda_k f(x) + H(x, x^{k-1}) \mid x \in \mathcal{K} \cap \mathcal{V}\right\}, \quad k = 1, 2, \ldots \tag{3}$$

where $\{\lambda_k\}$ is a sequence of positive parameters, and $H: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proximal distance with respect to int $\mathcal{K}$ (see Definition 3.1) which plays the same role as the Euclidean distance $\|x-y\|^2$ in the classical proximal algorithms (see, e.g., [12,13]), but possesses certain more desirable properties for forcing the iterates to stay in $\mathcal{K} \cap \mathcal{V}$, thus eliminating the constraints automatically. As will be shown in Section 4, such proximal distances can be produced with an appropriate closed proper univariate function.

In this paper, under mild assumptions like those used in interior proximal methods for convex programs over nonnegative orthant cones (see, e.g., [14–20]), we show that the sequence $\{x^k\}$ is bounded with all limit points, being a solution of (1), and obtain global rates of convergence in terms of objective values. But, unlike for interior proximal methods for convex programs over nonnegative orthant cones, the global convergence of $\{x^k\}$ to an optimal solution can be guaranteed for the class of proximal distances $\mathcal{F}_1(\mathcal{K})$ or $\mathcal{F}_2(\mathcal{K})$ under a very restrictive assumption for X_* (see Theorem 3.2(a)), or for their subclasses $\widehat{\mathcal{F}}_1(\mathcal{K}^n)$ or $\widehat{\mathcal{F}}_2(\mathcal{K}^n)$ under mild assumptions for X_* (see Theorem 3.2(b)), or for the smallest subclass $\bar{\mathcal{F}}_2(\mathcal{K}^n)$. These results are illustrated with some examples.

Just like proximal point methods with generalized distances, the central paths derived from barrier functions have been the object of intensive study. Recently, the central paths for semidefinite programming were under active study (see, e.g., [21–24]). For example, da Cruz Neto et al. [21] established relations among the central paths in semidefinite programming, generalized proximal point methods, and Cauchy trajectories in Riemannian manifolds, extending the results of Iusem et al. [25] for monotone variational inequality problems. Motivated by this, we also investigate the properties of the central paths of (1) with respect to (w.r.t.) the distance-like functions used by interior proximal methods (see Propositions 5.2 and 5.3). For the linear SOCP, we discuss the relations between the central paths and the sequences generated by the interior proximal methods, and show that the sequence generated by interior proximal methods will converge under the usual assumptions if the proximal distance satisfies a certain continuity at the boundary of second-order cones (see Theorem 5.2).

Auslender and Teboulle [15] provided a unified technique for analyzing and designing interior proximal methods for convex and conic optimization. However, for the CSOCP, we notice that it seems hard to find a proximal distance example for the class $\mathcal{F}_+(\mathcal{K}^n)$ such that global convergence results similar to those for [15, Theorem 2.2] can apply for it. In this paper, we extend their unified analysis technique to interior proximal methods using a proximal distance which can be produced with an appropriate univariate function in three ways, and establish the global convergence results for the smallest class $\bar{\mathcal{F}}_2(\mathcal{K}^n)$, and the class $\widehat{\mathcal{F}}_2(\mathcal{K}^n)$ with some mild assumptions of X_*. The examples from the two classes of proximal distances are easy to find. In particular, for the linear SOCP, we obtain improved convergence results for these interior proximal methods, by exploring the relations between the sequence generated by the interior proximal methods and the central path associated with the corresponding proximal distances. In view of these contexts, this paper can be regarded as a refinement of [15] for the second-order cone optimization.

Throughout this paper, I denotes an identity matrix of suitable dimension and $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors. For any $x, y \in \mathbb{R}^n$, we write $x \succeq_{\mathcal{K}^n} y$ if $x - y \in \mathcal{K}^n$; and we write $x \succ_{\mathcal{K}^n} y$ if $x - y \in \operatorname{int} \mathcal{K}^n$. Given a matrix E, $\operatorname{Im}(E)$ means the subspace generated by the columns of E. A function is closed if and only if it is lower semicontinuous (lsc), and a function is proper if $f(x) < \infty$ for at least one $x \in \mathbb{R}^n$ and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$. For a lsc proper convex function $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, we denote its domain by $\operatorname{dom} f := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$ and the ϵ-subdifferential of f at $\bar{x}$ by $\partial_\epsilon f(\bar{x}) := \{w \in \mathbb{R}^n \mid f(x) \geq f(\bar{x}) + \langle w, x - \bar{x}\rangle - \epsilon,\ \forall x \in \mathbb{R}^n\}$. If f is differentiable at x, $\nabla f(x)$ means the gradient of f at x. For a differentiable h on $\mathbb{R}$, h' and h'' denote its first and second derivatives. For any closed set S, int S denotes the interior of S.

In the rest of this paper, we focus on the case where $\mathcal{K} = \mathcal{K}^n$, and all the analysis can be carried over to the case where $\mathcal{K}$ has the direct product structure as in (2). Unless otherwise stated, we make the following minimal assumption for the CSOCP (1):

(A1) $\operatorname{dom} f \cap (\mathcal{V} \cap \operatorname{int} \mathcal{K}^n) \neq \emptyset$ and $f_* := \inf\{f(x) \mid x \in \mathcal{V} \cap \mathcal{K}^n\} > -\infty$.

2. Preliminaries

This section recalls some preliminary results that will be used in the subsequent sections. For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, their Jordan product [1] is defined as

$$x \circ y := (\langle x, y\rangle,\ y_1 x_2 + x_1 y_2). \tag{4}$$

It is easy to verify that the identity element under the Jordan product is $e \equiv (1, 0, \ldots, 0)^T \in \mathbb{R}^n$, i.e., $e \circ x = x$ for all $x \in \mathbb{R}^n$. Note that the Jordan product is not associative, but it is power associated, i.e., $x \circ (x \circ x) = (x \circ x) \circ x$ for all $x \in \mathbb{R}^n$. Thus, we may without fear of ambiguity write x^m for the product of m copies of x and $x^{m+n} = x^m \circ x^n$ for all positive integers m and n. n We stipulate $x^0 = e$. For each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let

$$\det(x) := x_1^2 - \|x_2\|^2 \quad \text{and} \quad \mathrm{tr}(x) := 2x_1. \tag{5}$$

These are called the *determinant* and the *trace* of x, respectively. A vector x is said to be *invertible* if $\det(x) \neq 0$. If $x \in \mathbb{R}^n$ is invertible, there is a unique $y \in \mathbb{R}^n$ satisfying $x \circ y = y \circ x = e$. We call this y the inverse of x and denote it by x^{-1}.

We recall from [1] that each x admits a spectral factorization associated with $\mathcal{K}^n$:

$$x = \lambda_1(x)\, u_x^{(1)} + \lambda_2(x)\, u_x^{(2)}, \tag{6}$$

where $\lambda_i(x)$ and $u_x^{(i)}$ for $i = 1, 2$ are the spectral values of $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and the associated spectral vectors, defined by

$$\lambda_i(x) = x_1 + (-1)^i \|x_2\|, \qquad u_x^{(i)} = \frac{1}{2}\left(1,\ (-1)^i \bar{x}_2\right), \tag{7}$$

with $\bar{x}_2 = \frac{x_2}{\|x_2\|}$ if $x_2 \neq 0$, otherwise being any vector in $\mathbb{R}^{n-1}$ such that $\|\bar{x}_2\| = 1$. If $x_2 \neq 0$, then the factorization is unique. The following lemma is direct by formula (6).

Lemma 2.1. *For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the following results hold:*

(a) $\det(x) = \lambda_1(x)\lambda_2(x)$, $\mathrm{tr}(x) = \lambda_1(x) + \lambda_2(x)$ *and* $\|x\|^2 = \frac{1}{2}\left[(\lambda_1(x))^2 + (\lambda_2(x))^2\right]$.
(b) $x \in \mathcal{K}^n \iff \lambda_1(x) \geq 0$ *and* $x \in \mathrm{int}\,\mathcal{K}^n \iff \lambda_1(x) > 0$.
(c) $\lambda_1(x)\lambda_2(y) + \lambda_2(x)\lambda_1(y) \leq \mathrm{tr}(x \circ y) \leq \lambda_1(x)\lambda_1(y) + \lambda_2(x)\lambda_2(y)$.

With the spectral factorization above, one may define a vector-valued function using a univariate function. For any given $h\colon \mathrm{I}_{\mathbb{R}} \to \mathbb{R}$ with $\mathrm{I}_{\mathbb{R}} \subseteq \mathbb{R}$, define $h^{\mathrm{soc}}\colon S \to \mathbb{R}^n$ by

$$h^{\mathrm{soc}}(x) := h(\lambda_1(x)) \cdot u_x^{(1)} + h(\lambda_2(x)) \cdot u_x^{(2)}, \quad \forall x \in S. \tag{8}$$

The definition is unambiguous whether $x_2 \neq 0$ or $x_2 = 0$. For example, let $h(t) = t^{-1}$ for any $t > 0$; then using formulas (6) and (8) we can compute that

$$x^{-1} := h^{\mathrm{soc}}(x) = \frac{1}{x_1^2 - \|x_2\|^2}(x_1, -x_2) = \frac{\mathrm{tr}(x)e - x}{\det(x)} \quad \text{for } x \in \mathrm{int}\,\mathcal{K}^n. \tag{9}$$

Moreover, by Lemma 2.2 of [26], S is open whenever $\mathrm{I}_{\mathbb{R}}$ is open, and S is closed whenever $\mathrm{I}_{\mathbb{R}}$ is closed. The following lemma shows that some favorable properties of h can be transmitted to h^{soc}, whose proofs were given in Proposition 5.1 of [8] and Lemma 2.2 of [27].

Lemma 2.2. *Given $h\colon \mathrm{I}_{\mathbb{R}} \to \mathbb{R}$ with $\mathrm{I}_{\mathbb{R}} \subseteq \mathbb{R}$, let $h^{\mathrm{soc}}\colon S \to \mathbb{R}^n$ be the vector-valued function induced by h via (8), where $S \subseteq \mathbb{R}^n$. Then, the following results hold:*

(a) *If h is continuously differentiable on* $\mathrm{int}\,\mathrm{I}_{\mathbb{R}}$, *then h^{soc} is continuously differentiable on* $\mathrm{int}\,S$, *and for any $x \in \mathrm{int}\,S$ with $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$,*

$$\nabla h^{\mathrm{soc}}(x) = \begin{cases} h'(x_1)I & \text{if } x_2 = 0, \\ \begin{bmatrix} b & c\dfrac{x_2^T}{\|x_2\|} \\ c\dfrac{x_2}{\|x_2\|} & aI + (b-a)\dfrac{x_2 x_2^T}{\|x_2\|^2} \end{bmatrix} & \text{otherwise} \end{cases}$$

where $a = \frac{h(\lambda_2(x)) - h(\lambda_1(x))}{\lambda_2(x) - \lambda_1(x)}$, $b = \frac{h'(\lambda_2(x)) + h'(\lambda_1(x))}{2}$, $c = \frac{h'(\lambda_2(x)) - h'(\lambda_1(x))}{2}$.
(b) *If h is continuously differentiable on* $\mathrm{int}\,\mathrm{I}_{\mathbb{R}}$, *then* $\mathrm{tr}(h^{\mathrm{soc}}(x))$ *is continuously differentiable on* $\mathrm{int}\,S$ *with* $\nabla \mathrm{tr}(h^{\mathrm{soc}}(x)) = 2\nabla h^{\mathrm{soc}}(x)e = 2(h')^{\mathrm{soc}}(x)$.
(c) *If h is (strictly) convex on $\mathrm{I}_{\mathbb{R}}$, then* $\mathrm{tr}(h^{\mathrm{soc}}(x))$ *is (strictly) convex on S.*

Lemma 2.3. (a) *The real-valued function* $\ln(\det(x))$ *is strictly concave on* $\mathrm{int}\,\mathcal{K}^n$.
(b) *For any $x, y \in \mathrm{int}\,\mathcal{K}^n$ with $x \neq y$, it holds that*

$$\det(\alpha x + (1-\alpha)y) > (\det(x))^{\alpha}(\det(y))^{1-\alpha}, \quad \forall \alpha \in (0, 1).$$

Proof. Clearly, part (b) is a direct consequence of part (a). The proof of part (a) was given in [28, Prop. 2.4(a)] by computing the Hessian matrix of $\ln(\det(x))$. Here, we give a simpler proof. Let $\ln x$ be the vector-valued function induced by $\ln t$ via (8). From Lemma 2.1(a), $\ln(\det(x)) = \ln(\lambda_1(x)) + \ln(\lambda_2(x)) = \mathrm{tr}(\ln x)$ for any $x \in \mathrm{int}\,\mathcal{K}^n$. The result is then direct by Lemma 2.2(c) and the strict concavity of $\ln t$ $(t > 0)$. $\square$

To close this section, we review the definition of SOC-convexity and SOC-monotonicity. The two concepts, like matrix-convexity and the matrix-monotonicity in semidefinite programming, play an important role in the solution methods of SOCPs.

Definition 2.1 (*[28]*)**.** Given $h: \mathrm{I}_{\mathbb{R}} \to \mathbb{R}$ with $\mathrm{I}_{\mathbb{R}} \subseteq \mathbb{R}$. Let $h^{\mathrm{soc}}: S \to \mathbb{R}^n$ with $S \subseteq \mathbb{R}^n$ be the vector-valued function induced by h via formula (8). Then,

(a) h is said to be SOC-convex of order n on $\mathrm{I}_{\mathbb{R}}$ if for any $x, y \in S$ and $0 \le \beta \le 1$,

$$h^{\mathrm{soc}}(\beta x + (1-\beta)y) \preceq_{\mathcal{K}^n} \beta h^{\mathrm{soc}}(x) + (1-\beta)h^{\mathrm{soc}}(y). \tag{10}$$

(b) h is said to be SOC-monotone of order n on $\mathrm{I}_{\mathbb{R}}$ if for any $x, y \in S$,

$$x \succeq_{\mathcal{K}^n} y \Longrightarrow h^{\mathrm{soc}}(x) \succeq_{\mathcal{K}^n} h^{\mathrm{soc}}(y).$$

We say that h is SOC-convex (respectively, SOC-monotone) on $\mathrm{I}_{\mathbb{R}}$ if h is SOC-convex of all orders n (respectively, SOC-monotone of all orders n) on $\mathrm{I}_{\mathbb{R}}$. A function h is said to be SOC-concave on $\mathrm{I}_{\mathbb{R}}$ whenever $-h$ is SOC-convex on $\mathrm{I}_{\mathbb{R}}$. When h is continuous on $\mathrm{I}_{\mathbb{R}}$, the condition in (10) can be replaced by a more special condition:

$$h^{\mathrm{soc}}\left(\frac{x+y}{2}\right) \preceq_{\mathcal{K}^n} \frac{1}{2}(h^{\mathrm{soc}}(x) + h^{\mathrm{soc}}(y)). \tag{11}$$

Obviously, the set of SOC-monotone functions and the set of SOC-convex functions are both closed under positive linear combinations and under pointwise limits.

For the characterizations of SOC-convexity and SOC-monotonicity, the interested reader may refer to [28,29]. The following lemma collects some common SOC-concave functions whose proofs can be found in [27] or are direct by Lemma 3.2 of [27].

Lemma 2.4. (a) *For any fixed $u \in \mathbb{R}$, the function $h(t) = (t+u)^r$ with $r \in [0, 1]$ is SOC-concave and SOC-monotone on $[-u, +\infty)$.*

(b) *For any fixed $u \in \mathbb{R}$, the function $h(t) = -(t+u)^{-r}$ with $r \in [0, 1]$ is SOC-concave and SOC-monotone on $(-u, +\infty)$.*

(c) *For any fixed $\alpha \ge 0$, $\ln(\alpha + t)$ is SOC-concave and SOC-monotone on $[-a, +\infty)$.*

(d) *For any fixed $u \ge 0$, $\frac{t}{u+t}$ is SOC-concave and SOC-monotone on $(-u, +\infty)$.*

3. Interior proximal methods

First of all, we present the definition of a proximal distance w.r.t. the open cone $\operatorname{int}\mathcal{K}^n$.

Definition 3.1. An extended-valued function $H: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called a proximal distance with respect to $\operatorname{int}\mathcal{K}^n$ if it satisfies the following properties:

(P1) $\mathrm{dom}H(\cdot,\cdot) = \mathcal{C}_1 \times \mathcal{C}_2$ with $\operatorname{int}\mathcal{K}^n \times \operatorname{int}\mathcal{K}^n \subset \mathcal{C}_1 \times \mathcal{C}_2 \subseteq \mathcal{K}^n \times \mathcal{K}^n$.

(P2) For each given $y \in \operatorname{int}\mathcal{K}^n$, $H(\cdot, y)$ is continuous and strictly convex on $\mathcal{C}_1$, and it is continuously differentiable on $\operatorname{int}\mathcal{K}^n$ with $\mathrm{dom}\nabla_1 H(\cdot, y) = \operatorname{int}\mathcal{K}^n$.

(P3) $H(x, y) \ge 0$ for all $x, y \in \mathbb{R}^n$, and $H(y, y) = 0$ for all $y \in \operatorname{int}\mathcal{K}^n$.

(P4) For each fixed $y \in \mathcal{C}_2$, the sets $\{x \in \mathcal{C}_1 : H(x, y) \le \gamma\}$ are bounded for all $\gamma \in \mathbb{R}$.

Definition 3.1 has a little difference from Definition 2.1 of [15] for a proximal distance w.r.t. $\operatorname{int}\mathcal{K}^n$, since here $H(\cdot, y)$ is required to be strictly convex over $\mathcal{C}_1$ for any fixed $y \in \operatorname{int}\mathcal{K}^n$. We denote by $\mathcal{D}(\operatorname{int}\mathcal{K}^n)$ the family of functions H satisfying Definition 3.1. With a given $H \in \mathcal{D}(\operatorname{int}\mathcal{K}^n)$, we have the following basic iterative algorithm for (1).

Interior Proximal Algorithm (IPA). Given $H \in \mathcal{D}(\operatorname{int}\mathcal{K}^n)$ and $x^0 \in \mathcal{V} \cap \operatorname{int}\mathcal{K}^n$, for $k = 1, 2, \ldots$, with $\lambda_k > 0$ and $\epsilon_k \ge 0$, generate a sequence $\{x^k\} \subset \mathcal{V} \cap \operatorname{int}\mathcal{K}^n$ with $g^k \in \partial_{\epsilon_k} f(x^k)$ via the following iterative scheme:

$$x^k := \operatorname{argmin}\left\{\lambda_k f(x) + H(x, x^{k-1}) \mid x \in \mathcal{V}\right\} \tag{12}$$

such that

$$\lambda_k g^k + \nabla_1 H(x^k, x^{k-1}) = A^T u^k \quad \text{for some } u^k \in \mathbb{R}^m. \tag{13}$$

The following proposition implies that the IPA is well-defined, and moreover, from its proof we see that the iterative formula (12) is equivalent to the iterative scheme (3). When $\epsilon_k > 0$ for any $k \in \mathbb{N}$ (the set of natural numbers), the IPA can be viewed as an approximate interior proximal method, and it becomes exact if $\epsilon_k = 0$ for all $k \in \mathbb{N}$.

Proposition 3.1. *For any given $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$ and $y \in \text{int}\,\mathcal{K}^n$, consider the problem*

$$f_*(y,\tau) = \inf\{\tau f(x) + H(x,y) \mid x \in \mathcal{V}\} \quad \text{with } \tau > 0. \tag{14}$$

Then, for each $\epsilon \geq 0$, there exist $x(y,\tau) \in \mathcal{V} \cap \text{int}\,\mathcal{K}^n$ and $g \in \partial_\epsilon f(x(y,\tau))$ such that

$$\tau g + \nabla_1 H(x(y,\tau), y) = A^T u \tag{15}$$

for some $u \in \mathbb{R}^m$. Moreover, for such $x(y,\tau)$, we have

$$\tau f(x(y,\tau)) + H(x(y,\tau), y) \leq f_*(y,\tau) + \epsilon.$$

Proof. Set $F(x,\tau) := \tau f(x) + H(x,y) + \delta_{\mathcal{V}\cap\mathcal{K}^n}(x)$, where $\delta_{\mathcal{V}\cap\mathcal{K}^n}(x)$ is the indicator function defined on the set $\mathcal{V} \cap \mathcal{K}^n$. Since $\text{dom} H(\cdot,y) = \mathcal{C}_1 \subset \mathcal{K}^n$, it is clear that

$$f_*(y,\tau) = \inf\left\{F(x,\tau) \mid x \in \mathbb{R}^n\right\}. \tag{16}$$

Since $f_* > -\infty$, it is easy to verify that for any $\gamma \in \mathbb{R}$ the following relation holds:

$$\begin{aligned}\left\{x \in \mathbb{R}^n \mid F(x,\tau) \leq \gamma\right\} &\subset \left\{x \in \mathcal{V} \cap \mathcal{K}^n \mid H(x,y) \leq \gamma - \tau f_*\right\} \\ &\subset \left\{x \in \mathcal{C}_1 \mid H(x,y) \leq \gamma - \tau f_*\right\},\end{aligned}$$

which together with (P4) implies that $F(\cdot,\tau)$ has bounded level sets. In addition, by (P1)–(P3), $F(\cdot,\tau)$ is a closed proper and strictly convex function. Hence, the problem (16) has a unique solution, say $x(y,\tau)$. From the optimality conditions of (16), we get

$$0 \in \partial F(x(y,\tau)) = \tau \partial f(x(y,\tau)) + \nabla_1 H(x(y,\tau), y) + \partial \delta_{\mathcal{V}\cap\mathcal{K}^n}(x(y,\tau))$$

where the equality is due to Theorem 23.8 of [30] and $\text{dom} f \cap (\mathcal{V} \cap \text{int}\,\mathcal{K}^n) \neq \emptyset$. Notice that $\text{dom}\,\nabla_1 H(\cdot,y) = \text{int}\,\mathcal{K}^n$ and $\text{dom}\,\partial\delta_{\mathcal{V}\cap\mathcal{K}^n}(\cdot) = \mathcal{V} \cap \mathcal{K}^n$. Therefore, the last equation implies $x(y,\tau) \in \mathcal{V} \cap \text{int}\,\mathcal{K}^n$, and there exists $g \in \partial f(x(y,\tau))$ such that

$$-\tau g - \nabla_1 H(x(y,\tau), y) \in \partial \delta_{\mathcal{V}\cap\mathcal{K}^n}(x(y,\tau)).$$

On the other hand, by the definition of $\delta_{\mathcal{V}\cap\mathcal{K}^n}(\cdot)$, it is not hard to derive that

$$\partial \delta_{\mathcal{V}\cap\mathcal{K}^n}(x) = \text{Im}(A^T) \quad \forall x \in \mathcal{V} \cap \text{int}\,\mathcal{K}^n.$$

The last two equations imply that (15) holds for $\epsilon = 0$. When $\epsilon > 0$, (15) also holds for such $x(y,\tau)$ and g since $\partial f(x(y,\tau)) \subset \partial_\epsilon f(x(y,\tau))$. Finally, since for each $y \in \text{int}\,\mathcal{K}^n$ the function $H(\cdot,y)$ is strictly convex, and since $g \in \partial_\epsilon f(x(y,\tau))$, we have

$$\begin{aligned}\tau f(x) + H(x,y) &\geq \tau f(x(y,\tau)) + H(x(y,\tau), y) + \langle \tau g + \nabla_1 H(x(y,\tau), y), x - x(y,\tau)\rangle - \epsilon \\ &= \tau f(x(y,\tau)) + H(x(y,\tau), y) + \langle A^T u, x - x(y,\tau)\rangle - \epsilon \\ &= \tau f(x(y,\tau)) + H(x(y,\tau), y) - \epsilon \quad \text{for all } x \in \mathcal{V},\end{aligned}$$

where the first equality is from (15) and the last one is by $x, x(y,\tau) \in \mathcal{V}$. Thus, $f_*(y,\tau) = \inf\{\tau f(x) + H(x,y) \mid x \in \mathcal{V}\} \geq \tau f(x(y,\tau)) + H(x(y,\tau), y) - \epsilon$. □

In the rest of this section, we focus on the convergence behaviors of the IPA with H from several subclasses of $\mathcal{D}(\text{int}\,\mathcal{K}^n)$, which also satisfy one of the following properties.

(P5) For any $x, y \in \text{int}\,\mathcal{K}^n$ and $z \in \mathcal{C}_1$, $H(z,y) - H(z,x) \geq \langle \nabla_1 H(x,y), z - x\rangle$.
(P5′) For any $x, y \in \text{int}\,\mathcal{K}^n$ and $z \in \mathcal{C}_2$, $H(y,z) - H(x,z) \geq \langle \nabla_1 H(x,y), z - x\rangle$.
(P6) For each $x \in \mathcal{C}_1$, the level sets $\{y \in \mathcal{C}_2 : H(x,y) \leq \gamma\}$ are bounded for all $\gamma \in \mathbb{R}$.

Specifically, we denote as $\mathcal{F}_1(\text{int}\,\mathcal{K}^n)$ and $\mathcal{F}_2(\text{int}\,\mathcal{K}^n)$ the families of functions $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$ satisfying (P5) and (P5′), respectively. If $\mathcal{C}_1 = \mathcal{K}^n$, we denote as $\mathcal{F}_1(\mathcal{K}^n)$ the family of functions $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$ satisfying (P5) and (P6). If $\mathcal{C}_2 = \mathcal{K}^n$, we write $\mathcal{F}_2(\text{int}\,\mathcal{K}^n)$ as $\mathcal{F}(\mathcal{K}^n)$. It is easy to see that the class of proximal distance $\mathcal{F}(\text{int}\,\mathcal{K}^n)$ (respectively, $\mathcal{F}(\mathcal{K}^n)$) in [15] subsumes the (H,H) with $H \in \mathcal{F}_1(\text{int}\,\mathcal{K}^n)$ (respectively, $\mathcal{F}_1(\mathcal{K}^n)$), but it does not include any (H,H) with $H \in \mathcal{F}_2(\text{int}\,\mathcal{K}^n)$ (respectively, $\mathcal{F}_2(\mathcal{K}^n)$).

Theorem 3.1. *Let $\{x^k\}$ be the sequence generated by the IPA with $H \in \mathcal{F}_1(\text{int}\,\mathcal{K}^n)$ or $H \in \mathcal{F}_2(\text{int}\,\mathcal{K}^n)$. Set $\sigma_\nu = \sum_{k=1}^{\nu} \lambda_k$. Then, the following results hold:*

(a) *$f(x^\nu) - f(x) \leq \sigma_\nu^{-1} H(x, x^0) + \sigma_\nu^{-1} \sum_{k=1}^{\nu} \sigma_k \epsilon_k$ for any $x \in \mathcal{V} \cap \mathcal{C}_1$ if $H \in \mathcal{F}_1(\text{int}\,\mathcal{K}^n)$; $f(x^\nu) - f(x) \leq \sigma_\nu^{-1} H(x^0, x) + \sigma_\nu^{-1} \sum_{k=1}^{\nu} \sigma_k \epsilon_k$ for any $x \in \mathcal{V} \cap \mathcal{C}_2$ if $H \in \mathcal{F}_2(\text{int}\,\mathcal{K}^n)$.*
(b) *If $\sigma_\nu \to +\infty$ and $\epsilon_k \to 0$, then $\liminf_{\nu\to\infty} f(x^\nu) = f_*$.*
(c) *The sequence $\{f(x^k)\}$ converges to f_* whenever $\sum_{k=1}^{\infty} \epsilon_k < \infty$.*

(d) *If* $X_* \neq \emptyset$, *then* $\{x^k\}$ *is bounded with all limit points in* X_* *under* (d1) *or* (d2) *:*
(d1) X_* *is bounded and* $\sum_{k=1}^{\infty} \epsilon_k < \infty$*;*
(d2) $\sum_{k=1}^{\infty} \lambda_k \epsilon_k < \infty$ *and* $H \in \mathcal{F}_1(\mathcal{K}^n)$ *(or* $H \in \mathcal{F}_2(\mathcal{K}^n)$*).*

Proof. The proofs are similar to those of [15, Theorem 4.1]. For completeness, we here take $H \in \mathcal{F}_2(\operatorname{int} \mathcal{K}^n)$ for example to prove the results.
(a) Since $g^k \in \partial_{\epsilon_k} f(x^k)$, from the definition of the subdifferential, it follows that

$$f(x) \geq f(x^k) + \langle g^k, x - x^k \rangle - \epsilon_k \quad \forall x \in \mathbb{R}^n.$$

This, together with Eq. (13), implies that

$$\lambda_k (f(x^k) - f(x)) \leq \langle \nabla_1 H(x^k, x^{k-1}), x - x^k \rangle + \lambda_k \epsilon_k \quad \forall x \in \mathcal{V} \cap \mathcal{C}_2.$$

Using (P5′) with $x = x^k$, $y = x^{k-1}$ and $z = x \in \mathcal{V} \cap \mathcal{C}_2$, it then follows that

$$\lambda_k (f(x^k) - f(x)) \leq H(x^{k-1}, x) - H(x^k, x) + \lambda_k \epsilon_k \quad \forall x \in \mathcal{V} \cap \mathcal{C}_2. \tag{17}$$

Summing over $k = 1, 2, \ldots, \nu$ in this inequality yields that

$$-\sigma_\nu f(x) + \sum_{k=1}^{\nu} \lambda_k f(x^k) \leq H(x^0, x) - H(x^\nu, x) + \sum_{k=1}^{\nu} \lambda_k \epsilon_k. \tag{18}$$

On the other hand, setting $x = x^{k-1}$ in (17), we obtain

$$f(x^k) - f(x^{k-1}) \leq \lambda_k^{-1} \left[H(x^{k-1}, x^{k-1}) - H(x^k, x^{k-1}) \right] + \epsilon_k \leq \epsilon_k. \tag{19}$$

Multiplying the inequality by σ_{k-1} (with $\sigma_0 \equiv 0$) and summing over $k = 1, \ldots, \nu$, we get

$$\sum_{k=1}^{\nu} \sigma_{k-1} f(x^k) - \sum_{k=1}^{\nu} \sigma_{k-1} f(x^{k-1}) \leq \sum_{k=1}^{\nu} \sigma_{k-1} \epsilon_k.$$

Noting that $\sigma_k = \lambda_k + \sigma_{k-1}$ with $\sigma_0 \equiv 0$, the above inequality can reduce to

$$\sigma_\nu f(x^\nu) - \sum_{k=1}^{\nu} \lambda_k f(x^k) \leq \sum_{k=1}^{\nu} \sigma_{k-1} \epsilon_k. \tag{20}$$

Adding the inequalities (18) and (20) and recalling that $\sigma_k = \lambda_k + \sigma_{k-1}$, it follows that

$$f(x^\nu) - f(x) \leq \sigma_\nu^{-1} \left[H(x^0, x) - H(x^\nu, x) \right] + \sigma_\nu^{-1} \sum_{k=1}^{\nu} \sigma_k \epsilon_k \quad \forall x \in \mathcal{V} \cap \mathcal{C}_2,$$

which immediately implies the desired result due to the nonnegativity of $H(x^\nu, x)$.
(b) If $\sigma_\nu \to +\infty$ and $\epsilon_k \to 0$, then applying Lemma 2.2(ii) of [15] with $a_k = \epsilon_k$ and $b_\nu := \sigma_\nu^{-1} \sum_{k=1}^{\nu} \lambda_k \epsilon_k$ yields $\sigma_\nu^{-1} \sum_{k=1}^{\nu} \lambda_k \epsilon_k \to 0$. From part (a), it then follows that

$$\liminf_{\nu \to \infty} f(x^\nu) \leq \inf \left\{ f(x) \mid x \in \mathcal{V} \cap \operatorname{int} \mathcal{K}^n \right\}.$$

This together with $f(x^\nu) \geq \inf\{f(x) \mid x \in \mathcal{V} \cap \mathcal{K}^n\}$ implies that

$$\liminf_{\nu \to \infty} f(x^\nu) = \inf \left\{ f(x) \mid x \in \mathcal{V} \cap \operatorname{int} \mathcal{K}^n \right\} = f_*.$$

(c) From (19), $0 \leq f(x^k) - f_* \leq f(x^{k-1}) - f_* + \epsilon_k$. Using Lemma 2.1 of [15] with $\gamma_k \equiv 0$ and $v_k = f(x^k) - f_*$, we have that $\{f(x^k)\}$ converges to f_* whenever $\sum_{k=1}^{\infty} \epsilon_k < \infty$.
(d) If the condition (d1) holds, then the sets $\{x \in \mathcal{V} \cap \mathcal{K}^n \mid f(x) \leq \gamma\}$ are bounded for all $\gamma \in \mathbb{R}$, since f is closed proper convex and $X_* = \{x \in \mathcal{V} \cap \mathcal{K}^n \mid f(x) \leq f_*\}$. Note that (19) implies $\{x^k\} \subset \{x \in \mathcal{V} \cap \mathcal{K}^n \mid f(x) \leq f(x^0) + \sum_{j=1}^{k} \epsilon_j\}$. Combining with $\sum_{k=1}^{\infty} \epsilon_k < \infty$, clearly we have that $\{x^k\}$ is bounded. Since $\{f(x^k)\}$ converges to f_* and f is lsc, passing to the limit and recalling that $\{x^k\} \subset \mathcal{V} \cap \mathcal{K}^n$ yields that each limit point of $\{x^k\}$ is a solution of (1).

Suppose that the condition (d2) holds. If $H \in \mathcal{F}_2(\mathcal{K}^n)$, then inequality (17) holds for each $x \in \mathcal{V} \cap \mathcal{K}^n$, and particularly for $x_* \in X_*$. Consequently,

$$H(x^k, x_*) \leq H(x^{k-1}, x_*) + \lambda_k \epsilon_k \quad \forall x_* \in X_*. \tag{21}$$

Summing over $k = 1, 2, \ldots, \nu$ for the last inequality, we obtain

$$H(x^\nu, x_*) \leq H(x^0, x_*) + \sum_{k=1}^{\nu} \lambda_k \epsilon_k.$$

This, by (P4) and $\sum_{k=1}^{\infty}\lambda_k\epsilon_k < \infty$, implies that $\{x^k\}$ is bounded, and hence has an accumulation point. Without loss of generality, let $\hat{x} \in \mathcal{K}^n$ be an accumulation point of $\{x^k\}$. Then there exists a subsequence $\{x^{k_j}\}$ such that $x^{k_j} \to \hat{x}$ as $j \to +\infty$. From the lower semicontinuity of f and part (c), we get $f(\hat{x}) \le \lim_{j\to+\infty} f(x^{k_j}) = f_*$, which means that $\hat{x}$ is a solution of (1). If $H \in \mathcal{F}_1(\mathcal{K}^n)$, then the last inequality becomes

$$H(x_*, x^\nu) \le H(x_*, x^0) + \sum_{k=1}^{\nu}\lambda_k\epsilon_k.$$

By (P6) and $\sum_{k=1}^{\infty}\lambda_k\epsilon_k < \infty$, we also have that $\{x^k\}$ is bounded, and hence has an accumulation point. Using the same arguments as above, we get the desired result. □

An immediate by-product of the above analysis yields the following global rate of convergence estimate for the IPA with $H \in \mathcal{F}_1(\mathcal{K}^n)$ or $H \in \mathcal{F}_2(\mathcal{K}^n)$.

Corollary 3.1. *Let $\{x^k\}$ be the sequence given by the IPA with $H \in \mathcal{F}_1(\mathcal{K}^n)$ or $\mathcal{F}_2(\mathcal{K}^n)$. If $X_* \ne \emptyset$ and $\sum_{k=1}^{\infty}\epsilon_k < \infty$, then $f(x^\nu) - f_* = O(\sigma_\nu^{-1})$.*

Proof. The result is direct on setting $x = x^*$ for some $x^* \in X_*$ in the inequalities of Theorem 3.1(a), and noting that $0 < \frac{\sigma_k}{\sigma_\nu} \le 1$ for all $k = 1, 2, \ldots, \nu$. □

To establish the global convergence of $\{x^k\}$ to an optimal solution of (1), we need to make further assumptions on X_* or the proximal distances in $\mathcal{F}_1(\mathcal{K}^n)$ and $\mathcal{F}_2(\mathcal{K}^n)$. We denote as $\widehat{\mathcal{F}_1}(\mathcal{K}^n)$ the family of functions $H \in \mathcal{F}_1(\mathcal{K}^n)$ satisfying (P7)–(P8), as $\widehat{\mathcal{F}_2}(\mathcal{K}^n)$ the family of functions $H \in \mathcal{F}_2(\mathcal{K}^n)$ satisfying (P7′)–(P8′), and as $\bar{\mathcal{F}}(\mathcal{K}^n)$ the family of functions $H \in \mathcal{F}_2(\mathcal{K}^n)$ satisfying (P7′)–(P9′):

- (P7) For any $\{y^k\} \subseteq \operatorname{int}\mathcal{K}^n$ converging to $y^* \in \mathcal{K}^n$, we have $H(y^*, y^k) \to 0$.
- (P8) For any bounded sequence $\{y^k\} \subseteq \operatorname{int}\mathcal{K}^n$ and any $y^* \in \mathcal{K}^n$ with $H(y^*, y^k) \to 0$, it holds that $\lambda_i(y^k) \to \lambda_i(y^*)$ for $i = 1, 2$.
- (P7′) For any $\{y^k\} \subseteq \operatorname{int}\mathcal{K}^n$ converging to $y^* \in \mathcal{K}^n$, we have $H(y^k, y^*) \to 0$.
- (P8′) For any bounded sequence $\{y^k\} \subseteq \operatorname{int}\mathcal{K}^n$ and any $y^* \in \mathcal{K}^n$ with $H(y^k, y^*) \to 0$, it holds that $\lambda_i(y^k) \to \lambda_i(y^*)$ for $i = 1, 2$.
- (P9′) For any bounded sequence $\{y^k\} \subseteq \operatorname{int}\mathcal{K}^n$ and any $y^* \in \mathcal{K}^n$ with $H(y^k, y^*) \to 0$, it holds that $y^k \to y^*$.

It is easy to see that all previous subclasses of $\mathcal{D}(\operatorname{int}\mathcal{K}^n)$ have the following relations:

$$\widehat{\mathcal{F}_1}(\mathcal{K}^n) \subset \mathcal{F}_1(\mathcal{K}^n) \subset \mathcal{F}_1(\operatorname{int}\mathcal{K}^n), \qquad \bar{\mathcal{F}}_2(\mathcal{K}^n) \subset \widehat{\mathcal{F}_2}(\mathcal{K}^n) \subset \mathcal{F}_2(\mathcal{K}^n) \subset \mathcal{F}_2(\operatorname{int}\mathcal{K}^n).$$

Theorem 3.2. *Let $\{x^k\}$ be generated by the IPA with $H \in \mathcal{F}_1(\operatorname{int}\mathcal{K}^n)$ or $\mathcal{F}_2(\operatorname{int}\mathcal{K}^n)$. Suppose that X_* is nonempty, $\sum_{k=1}^{\infty}\lambda_k\epsilon_k < \infty$ and $\sum_{k=1}^{\infty}\epsilon_k < \infty$.*

(a) If X_* is a single point set, then $\{x^k\}$ converges to an optimal solution of (1).
(b) If X_* includes at least two elements and for any $x^* = (x_1^*, x_2^*), \bar{x}^* = (\bar{x}_1^*, \bar{x}_2^*) \in X_*$ with $x^* \ne \bar{x}^*$, it holds that $x_1^* \ne \bar{x}_1^*$ or $\|x_2^*\| \ne \|\bar{x}_2^*\|$, then $\{x^k\}$ converges to an optimal solution of (1) whenever $H \in \widehat{\mathcal{F}_1}(\mathcal{K}^n)$ (or $H \in \widehat{\mathcal{F}_2}(\mathcal{K}^n)$).
(c) If $H \in \bar{\mathcal{F}}_2(\mathcal{K}^n)$, then $\{x^k\}$ converges to an optimal solution of (1).

Proof. Part (a) is direct by Theorem 3.1(d1). We next consider part (b). Assume that $H \in \widehat{\mathcal{F}_2}(\mathcal{K}^n)$. Since $\sum_{k=1}^{\infty}\lambda_k\epsilon_k < \infty$, from (21) and Lemma 2.1 of [15], it follows that the sequence $\{H(x^k, x)\}$ is convergent for any $x \in X_*$. Let $\bar{x}$ be the limit of a subsequence $\{x^{k_l}\}$. By Theorem 3.1(d2), $\bar{x} \in X_*$. Consequently, $\{H(x^k, \bar{x})\}$ is convergent. By (P7′), $H(x^{k_l}, \bar{x}) \to 0$, and so $H(x^k, \bar{x}) \to 0$. Combining with (P8′), we have $\lambda_i(x^k) \to \lambda_i(\bar{x})$ for $i = 1, 2$, i.e.,

$$x_1^k - \|x_2^k\| \to \bar{x}_1 - \|\bar{x}_2\| \quad \text{and} \quad x_1^k + \|x_2^k\| \to \bar{x}_1 + \|\bar{x}_2\| \quad \text{as } k \to \infty.$$

This implies that $x_1^k \to \bar{x}_1$ and $\|x_2^k\| \to \|\bar{x}_2\|$. Combining this with the given assumption for X_*, we have that $x^k \to \bar{x}$. Suppose that $H \in \widehat{\mathcal{F}_1}(\mathcal{K}^n)$. The inequality (21) becomes

$$H(x_*, x^k) \le H(x_*, x^{k-1}) + \lambda_k\epsilon_k \quad \forall x_* \in X_*,$$

and using (P7)–(P8) and the same arguments as above then yields the result. Part (c) is direct by the arguments above and the property (P9′). □

When all points in the nonempty X_* lie on the boundary of $\mathcal{K}^n$, we must have $x_1^* \ne \bar{x}_1^*$ or $\|x_2^*\| \ne \|\bar{x}_2^*\|$ for any $x^* = (x_1^*, x_2^*), \bar{x}^* = (\bar{x}_1^*, \bar{x}_2^*) \in X_*$ with $x^* \ne \bar{x}^*$, and the assumption for X_* in (b) is automatically satisfied. Since the solutions of (1) are generally on the boundary of $\mathcal{K}^n$, the assumption for X_* in (b) is much weaker than the one in (a).

Up to now, we have studied two kinds of convergence results for the IPA using the class in which the proximal distance H lies. Theorem 3.1 and Corollary 3.1 show that the largest, and least demanding, classes $\mathcal{F}_1(\operatorname{int}\mathcal{K}^n)$ and $\mathcal{F}_2(\operatorname{int}\mathcal{K}^n)$ provide reasonable convergence properties for the IPA under minimal assumptions on the problem's data. This coincides with interior proximal methods for convex programming over nonnegative orthant cones; see [15]. The smallest subclass $\bar{\mathcal{F}}_2(\mathcal{K}^n)$ of $\mathcal{F}_2(\text{int}\,\mathcal{K}^n)$ guarantees that $\{x^k\}$ converges to an optimal solution provided that X_* is nonempty. The smaller class $\widehat{\mathcal{F}}_2(\mathcal{K}^n)$ may guarantee the global convergence of the sequence $\{x^k\}$ to an optimal solution under an additional assumption besides X_* being nonempty. Moreover, we illustrate in the next section that there are indeed examples for the class $\bar{\mathcal{F}}_2(\mathcal{K}^n)$. For the smallest subclass $\widehat{\mathcal{F}}_1(\mathcal{K}^n)$ of $\mathcal{F}_1(\text{int}\,\mathcal{K}^n)$, the analysis in the next section shows that it seems hard to find an example, although it guarantees the convergence of $\{x^k\}$ to an optimal solution by Theorem 3.2(b).

4. Proximal distances over SOCs

In this section, we provide three kinds of ways to construct a proximal distance w.r.t. $\text{int}\,\mathcal{K}^n$ and analyze their advantages and disadvantages. All of these ways exploit a lsc proper univariate function to produce such a proximal distance. In addition, with such a proximal distance and the Euclidean distance, we obtain the regularized ones.

The first way produces the proximal distances for the class $\mathcal{F}_1(\text{int}\,\mathcal{K}^n)$. This approach is based on the compound of a univariate function ϕ and the determinant function $\det(\cdot)$, where $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a lsc proper function satisfying the following conditions:

(B1) $\text{dom}\phi \subseteq [0, +\infty)$, $\text{int}(\text{dom}\phi) = (0, +\infty)$, and ϕ is continuous on its domain;
(B2) for any $t_1, t_2 \in \text{dom}\phi$, it holds that

$$\phi(t_1^r t_2^{1-r}) \le r\phi(t_1) + (1-r)\phi(t_2), \quad \forall r \in [0, 1]; \tag{22}$$

(B3) ϕ is continuously differentiable on $\text{int}(\text{dom}\phi)$ with $\text{dom}\phi' = (0, +\infty)$;
(B4) $\phi'(t) < 0$ for all $t \in (0, +\infty)$, $\lim_{t\to 0^+} \phi(t) = +\infty$, and $\lim_{t\to+\infty} t^{-1}\phi(t^2) \ge 0$.

With such a univariate ϕ, we define the function $H: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$H(x, y) := \begin{cases} \phi(\det(x)) - \phi(\det(y)) - \langle \nabla\phi(\det(y)), x - y\rangle & \forall x, y \in \text{int}(\mathcal{K}^n); \\ +\infty & \text{otherwise.} \end{cases} \tag{23}$$

By the conditions (B1)–(B4), we may prove that H has the following properties.

Proposition 4.1. *Let H be defined as in (23) with ϕ satisfying* (B1)–(B4). *Then,*

(a) *for any fixed $y \in \text{int}\,\mathcal{K}^n$, $H(\cdot, y)$ is strictly convex over $\text{int}\,\mathcal{K}^n$.*
(b) *For any fixed $y \in \text{int}\,\mathcal{K}^n$, $H(\cdot, y)$ is continuously differentiable on $\text{int}\,\mathcal{K}^n$ with*

$$\nabla_1 H(x, y) = 2\phi'(\det(x))\begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} - 2\phi'(\det(y))\begin{pmatrix} y_1 \\ -y_2 \end{pmatrix} \tag{24}$$

for all $x \in \text{int}\,\mathcal{K}^n$, where $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$.
(c) *$H(x, y) \ge 0$ for all $x, y \in \mathbb{R}^n$, and $H(y, y) = 0$ for all $y \in \text{int}\,\mathcal{K}^n$.*
(d) *For any $y \in \text{int}\,\mathcal{K}^n$, the sets $\{x \in \text{int}\,\mathcal{K}^n: H(x, y) \le \gamma\}$ are bounded for all $\gamma \in \mathbb{R}$.*
(e) *For any $x, y \in \text{int}\,\mathcal{K}^n$ and $z \in \text{int}\,\mathcal{K}^n$, the following three-point identity holds:*

$$H(z, y) = H(z, x) + H(x, y) + \langle \nabla_1 H(x, y), z - x\rangle. \tag{25}$$

Proof. (a) It suffices to prove that $\phi(\det(x))$ is strictly convex on $\text{int}\,\mathcal{K}^n$. By Lemma 2.3(b),

$$\det(\alpha x + (1-\alpha)z) > (\det(x))^\alpha (\det(z))^{1-\alpha} \quad \forall \alpha \in (0, 1)$$

for all $x, z \in \text{int}\,\mathcal{K}^n$ and $x \ne z$. Since $\phi'(t) < 0$ for all $t \in (0, +\infty)$, we have that ϕ is decreasing on $(0, +\infty)$. This, together with the condition (B2), yields that

$$\begin{aligned} \phi\,[\det(\alpha x + (1-\alpha)z)] &< \phi\left[(\det(x))^\alpha (\det(z))^{1-\alpha}\right] \\ &\le \alpha\phi[\det(x)] + (1-\alpha)\phi[\det(z)] \quad \forall \alpha \in (0, 1) \end{aligned}$$

for all $x, z \in \text{int}\,\mathcal{K}^n$ and $x \ne z$. This means that $\phi(\det(x))$ is strictly convex on $\text{int}\,\mathcal{K}^n$.
(b) Since $\det(x)$ is continuously differentiable on $\mathbb{R}^n$ and ϕ is continuously differentiable on $(0, +\infty)$, we have that $\phi(\det(x))$ is continuously differentiable on $\text{int}\,\mathcal{K}^n$. This means that for any fixed $y \in \text{int}\,\mathcal{K}^n$, $H(\cdot, y)$ is continuously differentiable on $\text{int}\,\mathcal{K}^n$. By a simple computation, we immediately obtain the formula in (24).
(c) Since $\phi(\det(x))$ is strictly convex and continuously differentiable on $\text{int}\,\mathcal{K}^n$, we have

$$\phi(\det(x)) > \phi(\det(y)) - \langle \nabla\phi(\det(y)), x - y\rangle \quad \forall x, y \in \text{int}\,\mathcal{K}^n \text{ with } x \ne y$$

for any $x, y \in \text{int}\,\mathcal{K}^n$ with $x \ne y$. This implies that $H(y, y) = 0$ for all $y \in \text{int}\,\mathcal{K}^n$. In addition, from the inequality and the continuity of ϕ on its domain, it follows that

$$\phi(\det(x)) \ge \phi(\det(y)) - \langle \nabla\phi(\det(y)), x - y\rangle$$

for any $x, y \in \text{int}\,\mathcal{K}^n$. By the definition of H, we have $H(x, y) \ge 0$ for all $x, y \in \mathbb{R}^n$.

(d) Let $\{x^k\} \subset \operatorname{int} \mathcal{K}^n$ be a sequence with $\|x^k\| \to \infty$. For any fixed $y = (y_1, y_2) \in \operatorname{int} \mathcal{K}^n$, we next prove that the sequence $\{H(x^k, y)\}$ is unbounded for three cases, and then the desired result follows. For convenience, we write $x^k = (x_1^k, x_2^k)$ for each k.

Case 1: the sequence $\{\det(x^k)\}$ has a zero limit point. Without loss of generality, we assume that $\det(x^k) \to 0$ as $k \to \infty$. Combining with $\lim_{t\to 0^+} \phi(t) = +\infty$, it readily follows that $\lim_{k\to\infty} \phi(\det(x^k)) \to +\infty$. In addition, for each k we have that

$$\begin{aligned}\langle \nabla\phi(\det(y)), x^k\rangle &= 2\phi'(\det(y))(x_1^k y_1 - (x_2^k)^T y_2)\\ &\le 2\phi'(\det(y))y_1(x_1^k - \|x_2^k\|) \le 0\end{aligned} \tag{26}$$

where the first inequality uses $\phi'(t) < 0$ for all $t > 0$, the Schwartz inequality, and $y \in \operatorname{int} \mathcal{K}^n$. Now from (23), it then follows that $\lim_{k\to\infty} H(x^k, y) = +\infty$.

Case 2: the sequence $\{\det(x^k)\}$ is unbounded. Noting that $\det(x^k) > 0$ for each k, we must have $\det(x^k) \to +\infty$ as $k \to \infty$. Since ϕ is decreasing on its domain, we have that

$$\frac{\phi(\det(x^k))}{\|x^k\|} = \frac{\sqrt{2}\phi(\lambda_1(x^k)\lambda_2(x^k))}{\sqrt{(\lambda_1(x^k))^2 + (\lambda_2(x^k))^2}} \ge \frac{\phi[(\lambda_2(x^k))^2]}{\lambda_2(x^k)}.$$

Note that $\lambda_2(x^k) \to \infty$ in this case, and from the last equation and (B4) it follows that

$$\lim_{k\to\infty} \frac{\phi(\det(x^k))}{\|x^k\|} \ge \lim_{k\to\infty} \frac{\phi[(\lambda_2(x^k))^2]}{\lambda_2(x^k)} \ge 0.$$

In addition, since $\left\{\frac{x^k}{\|x^k\|}\right\}$ is bounded, we without loss of generality assume that $\frac{x^k}{\|x^k\|} \to \hat{x} = (\hat{x}_1, \hat{x}_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$. Then, $\hat{x} \in \mathcal{K}^n$, $\|\hat{x}\| = 1$, and $\hat{x}_1 > 0$ (if not, $\hat{x} = 0$), and hence

$$\begin{aligned}\lim_{k\to\infty} \left\langle \nabla\phi(\det(y)), \frac{x^k}{\|x^k\|}\right\rangle &= \langle \nabla\phi(\det(y)), \hat{x}\rangle = 2\phi'(\det(y))(\hat{x}_1 y_1 - \hat{x}_2^T y_2)\\ &\le 2\phi'(\det(y))\hat{x}_1(y_1 - \|y_2\|) < 0.\end{aligned}$$

The two sides show that $\lim_{k\to\infty} \frac{H(x^k,y)}{\|x^k\|} > 0$, and consequently $\lim_{k\to\infty} H(x^k, y) = +\infty$.

Case 3: the sequence $\{\det(x^k)\}$ has some limit point ω with $0 < \omega < +\infty$. Without loss of generality, we assume that $\det(x^k) \to \omega$ as $k \to \infty$. Since $\{x^k\}$ is unbounded and $\{x^k\} \subset \operatorname{int} \mathcal{K}^n$, we must have $x_1^k \to +\infty$. In addition, by (26) and $\phi'(t) < 0$ for $t > 0$,

$$-\langle \nabla\phi(\det(y)), x^k\rangle \ge -2\phi'(\det(y))(x_1^k y_1 - \|x_2^k\|\,\|y_2\|) \ge -2\phi'(\det(y))x_1^k(y_1 - \|y_2\|).$$

This along with $y \in \operatorname{int} \mathcal{K}^n$ implies that $-\langle \nabla\phi(\det(y)), x^k\rangle \to +\infty$ as $k \to \infty$. Noting that $\phi(\det(x^k))$ is bounded, from (23) it follows that $\lim_{k\to\infty} H(x^k, y) \to +\infty$.

(e) For any $x, y \in \operatorname{int} \mathcal{K}^n$ and $z \in \operatorname{int} \mathcal{K}^n$, from the definition of H it follows that

$$\begin{aligned}H(z, y) - H(z, x) - H(x, y) &= \langle \nabla\phi(\det(x)) - \nabla\phi(\det(y)), z - x\rangle\\ &= \langle \nabla_1 H(x, y), z - x\rangle\end{aligned}$$

where the last equality is by part (b). The proof is thus completed. □

Proposition 4.1 shows that the function H defined by (23) with ϕ satisfying (B1)–(B4) is a proximal distance w.r.t. $\operatorname{int} \mathcal{K}^n$ and $\operatorname{dom} H = \operatorname{int} \mathcal{K}^n \times \operatorname{int} \mathcal{K}^n$. Also, $H \in \mathcal{F}_1(\operatorname{int} \mathcal{K}^n)$. The conditions (B1) and (B3)–(B4) are easy to check, whereas by Lemma 2.2 of [31] we have the following important characterizations for the condition (B2).

Lemma 4.1 (*[31, Lemma 2.2]*)**.** *A function ϕ: $(0, +\infty) \to \mathbb{R}$ satisfies (B2) if and only if one of the following conditions holds:*

(a) *The function $\phi(\exp(\cdot))$ is convex on $\mathbb{R}$.*
(b) $\phi(t_1 t_2) \le \frac{1}{2}\left(\phi(t_1^2) + \phi(t_2^2)\right)$ *for any $t_1, t_2 > 0$.*
(c) $\phi'(t) + t\phi''(t) \ge 0$ *if ϕ is twice differentiable.*

Example 4.1. Take $\phi(t) = -\ln t$ if $t > 0$, and otherwise $\phi(t) = +\infty$. It is easy to verify that ϕ satisfies (B1)–(B4). By formula (23), the induced proximal distance is

$$H(x, y) := \begin{cases} -\ln \dfrac{\det(x)}{\det(y)} + \dfrac{2x^T J_n y}{\det(y)} - 2 & \forall x, y \in \operatorname{int}(\mathcal{K}^n)\\ +\infty & \text{otherwise}\end{cases}$$

where J_n is a diagonal matrix with the first entry being 1 and the rest of the $(n-1)$ entries being -1. This is exactly the proximal distance given by [15]. Since $H \in \mathscr{F}_1(\mathrm{int}\,\mathcal{K}^n)$, we have the results of Theorem 3.1(a)–(d1) if the proximal distance is used for the IPA.

Example 4.2. Take $\phi(t) = t^{1-q}/(q-1)$ $(q > 1)$ if $t > 0$, and otherwise $\phi(t) = +\infty$. It is not hard to check that ϕ satisfies (B1)–(B4). By (23), we compute that

$$H(x, y) := \begin{cases} \dfrac{(\det(x))^{1-q} - (\det(y))^{1-q}}{q-1} + \dfrac{2x^T J_n y}{(\det(y))^q} - (\det(y))^{1-q} & \forall x, y \in \mathrm{int}(\mathcal{K}^n) \\ +\infty & \text{otherwise} \end{cases}$$

where J_n is the same diagonal matrix as in Example 4.1. Since $H \in \mathscr{F}(\mathrm{int}\,\mathcal{K}^n)$, when using the proximal distance for the IPA, the results of Theorem 3.1(a)–(d1) hold.

We should emphasize that using the first approach cannot produce the proximal distances of the class $\mathscr{F}_1(\mathcal{K}^n)$, and so $\widehat{\mathscr{F}}_1(\mathcal{K}^n)$, since the condition $\lim_{t\to 0^+} \phi(t) = +\infty$ is necessary for guaranteeing that H has the property (P4), but it implies that the domain of $H(\cdot, y)$ for any $y \in \mathrm{int}\,\mathcal{K}^n$ cannot be continuously extended to $\mathcal{K}^n$. Thus, when choosing such proximal distances for the IPA, we cannot apply Theorem 3.1(d2) and Theorem 3.2.

The other two ways are both based on the compound of the trace function $\mathrm{tr}(\cdot)$ and a vector-valued function induced by a univariate ϕ via (8). For convenience, in the sequel, for any lsc proper function $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, we write $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ as

$$d(s, t) := \begin{cases} \phi(s) - \phi(t) - \phi'(t)(s-t) & \text{if } s \in \mathrm{dom}\phi,\ t \in \mathrm{dom}\phi', \\ +\infty & \text{otherwise.} \end{cases} \tag{27}$$

The second approach also produces the proximal distances for the class $\mathscr{F}_1(\mathrm{int}\,\mathcal{K}^n)$, which requires $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ to be a lsc proper function satisfying the conditions:

(C1) $\mathrm{dom}\phi \subseteq [0, +\infty)$ and $\mathrm{int}(\mathrm{dom}\phi) = (0, +\infty)$;
(C2) ϕ is continuous and strictly convex on its domain;
(C3) ϕ is continuously differentiable on $\mathrm{int}(\mathrm{dom}\phi)$ with $\mathrm{dom}\phi' = (0, +\infty)$;
(C4) for any fixed $t > 0$, the sets $\{s \in \mathrm{dom}\phi \mid d(s, t) \le \gamma\}$ are bounded with all $\gamma \in \mathbb{R}$; for any fixed $s \in \mathrm{dom}\phi$, the sets $\{t > 0 \mid d(s, t) \le \gamma\}$ are bounded with all $\gamma \in \mathbb{R}$.

Let ϕ^{soc} be the vector-valued function induced by ϕ via (8) and write $\mathrm{dom}\phi^{\mathrm{soc}} = \mathcal{C}_1$. Clearly, $\mathcal{C}_1 \subseteq \mathcal{K}^n$ and $\mathrm{int}\,\mathcal{C}_1 = \mathrm{int}\,\mathcal{K}^n$. Define the function $H: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$H(x, y) := \begin{cases} \mathrm{tr}(\phi^{\mathrm{soc}}(x)) - \mathrm{tr}(\phi^{\mathrm{soc}}(y)) - \langle \nabla \mathrm{tr}(\phi^{\mathrm{soc}}(y)), x - y\rangle & \forall x \in \mathcal{C}_1, y \in \mathrm{int}\,\mathcal{K}^n, \\ +\infty & \text{otherwise.} \end{cases} \tag{28}$$

By Lemmas 2.1 and 2.2, the conditions (C1)–(C4), and arguments similar to those of [32, Prop. 3.1], it is not difficult to argue that H has the following favorable properties.

Proposition 4.2. *Let H be defined by (28) with ϕ satisfying (C1)–(C4). Then:*

(a) *For any fixed $y \in \mathrm{int}\,\mathcal{K}^n$, $H(\cdot, y)$ is continuous and strictly convex on $\mathcal{C}_1$.*
(b) *For any fixed $y \in \mathrm{int}\,\mathcal{K}^n$, $H(\cdot, y)$ is continuously differentiable on $\mathrm{int}\,\mathcal{K}^n$ with*

$$\nabla_1 H(x, y) = \nabla \mathrm{tr}(\phi^{\mathrm{soc}}(x)) - \nabla \mathrm{tr}(\phi^{\mathrm{soc}}(y)) = 2\left[(\phi')^{\mathrm{soc}}(x) - (\phi')^{\mathrm{soc}}(y)\right].$$

(c) *$H(x, y) \ge 0$ for all $x, y \in \mathbb{R}^n$, and $H(y, y) = 0$ for any $y \in \mathrm{int}\,\mathcal{K}^n$.*
(d) *$H(x, y) \ge \sum_{i=1}^{2} d(\lambda_i(x), \lambda_i(y)) \ge 0$ for any $x \in \mathcal{C}_1$ and $y \in \mathrm{int}\,\mathcal{K}^n$.*
(e) *For any fixed $y \in \mathrm{int}\,\mathcal{K}^n$, the sets $\{x \in \mathcal{C}_1 : H(x, y) \le \gamma\}$ are bounded for all $\gamma \in \mathbb{R}$; for any fixed $x \in \mathcal{C}_1$, the sets $\{y \in \mathrm{int}\,\mathcal{K}^n : H(x, y) \le \gamma\}$ are bounded for all $\gamma \in \mathbb{R}$.*
(f) *For any $x, y \in \mathrm{int}\,\mathcal{K}^n$ and $z \in \mathcal{C}_1$, the following three-point identity holds:*

$$H(z, y) = H(z, x) + H(x, y) + \langle \nabla_1 H(x, y), z - x\rangle.$$

Proposition 4.2 shows that the function H defined by (28) with ϕ satisfying (C1)–(C4) is a proximal distance w.r.t. $\mathrm{int}\,\mathcal{K}^n$ with $\mathrm{dom}\,H = \mathcal{C}_1 \times \mathrm{int}\,\mathcal{K}^n$, and furthermore, such proximal distances belong to the class $\mathscr{F}_1(\mathrm{int}\,\mathcal{K}^n)$. In particular, when $\mathrm{dom}\phi = [0, +\infty)$, they also belong to the class $\mathscr{F}_1(\mathcal{K}^n)$. We next present some specific examples.

Example 4.3. Take $\phi(t) = t\ln t - t$ if $t \ge 0$, and otherwise $\phi(t) = +\infty$, where we stipulate $0 \ln 0 = 0$. It is easy to verify that ϕ satisfies (C1)–(C4) with $\mathrm{dom}\phi = [0, +\infty)$. By formulas (8) and (28), we compute that H has the following expression:

$$H(x, y) = \begin{cases} \mathrm{tr}(x \circ \ln x - x \circ \ln y + y - x) & \forall x \in \mathcal{K}^n, y \in \mathrm{int}(\mathcal{K}^n), \\ +\infty & \text{otherwise.} \end{cases} \tag{29}$$

Example 4.4. Take $\phi(t) = t^p - t^q$ if $t \geq 0$, and otherwise $\phi(t) = +\infty$, where $p \geq 1$ and $0 < q < 1$. We can show that ϕ satisfies the conditions (C1)–(C4) with $\mathrm{dom}\phi = [0, +\infty)$. When $p = 1$ and $q = 1/2$, from formulas (8) and (28), we derive that

$$H(x,y) = \begin{cases} \mathrm{tr}\left[y^{\frac{1}{2}} - x^{\frac{1}{2}} + \dfrac{(\mathrm{tr}(y^{\frac{1}{2}})e - y^{\frac{1}{2}}) \circ (x-y)}{2\sqrt{\det(y)}} \right] & \forall x \in \mathcal{K}^n, y \in \mathrm{int}\,\mathcal{K}^n, \\ +\infty & \text{otherwise.} \end{cases}$$

Example 4.5. Take $\phi(t) = -t^q$ if $t \geq 0$, and otherwise $\phi(t) = +\infty$, where $0 < q < 1$. We can show that ϕ satisfies the conditions (C1)–(C4) with $\mathrm{dom}\phi = [0, +\infty)$. Now

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

Example 4.6. Take $\phi(t) = -\ln t + t - 1$ if $t > 0$, and otherwise $\phi(t) = +\infty$. It is easy to check that ϕ satisfies (C1)–(C4) with $\mathrm{dom}\phi = (0, +\infty)$. The induced proximal distance is

$$H(x,y) = \begin{cases} \mathrm{tr}(\ln y) - \mathrm{tr}(\ln x) + 2\langle y^{-1}, x\rangle - 2 & \forall x, y \in \mathrm{int}\,\mathcal{K}^n, \\ +\infty & \text{otherwise.} \end{cases}$$

By a simple computation, we have that the proximal distance is the same as the one given by Example 4.1, and the one induced by $\phi(t) = -\ln t$ $(t > 0)$ via formula (28).

Clearly, the proximal distances in Examples 4.3–4.5 belong to the class $\mathcal{F}_1(\mathcal{K}^n)$. Also, by Proposition 4.3, the proximal distances in Examples 4.3 and 4.4 also satisfy (P8) since the corresponding ϕ also satisfies the following condition (C5):

(C5) For any bounded sequence $\{a^k\} \subset \mathrm{int}(\mathrm{dom}\phi)$ and $a \in \mathrm{dom}\phi$ such that $\lim_{k\to\infty} d(a, a^k) = 0$, it holds that $a = \lim_{k\to\infty} a^k$, where d is defined as in (27).

Proposition 4.3. *Let H be defined as in (28) with ϕ satisfying* (C1)–(C5) *and* $\mathrm{dom}\phi = [0, +\infty)$. *Then, for any bounded sequence $\{y^k\} \subseteq \mathrm{int}\,\mathcal{K}^n$ and $y^* \in \mathcal{K}^n$ such that $H(y^*, y^k) \to 0$, we have $\lambda_i(y^k) \to \lambda_i(y^*)$ for $i = 1, 2$.*

Proof. From Proposition 4.2(d) and the nonnegativity of d, for each k we have

$$H(y^*, y^k) \geq d(\lambda_i(y^*), \lambda_i(y^k)) \geq 0, \quad i = 1, 2.$$

This, together with the given assumption $H(y^*, y^k) \to 0$, implies that

$$d(\lambda_i(y^*), \lambda_i(y^k)) \to 0, \quad i = 1, 2.$$

Notice that $\{\lambda_i(y^k)\} \subset \mathrm{int}(\mathrm{dom}\phi)$ and $\lambda_i(y^*) \in \mathcal{K}^n$ for $i = 1, 2$ by Lemma 2.1(b). From the condition (C5), we immediately obtain $\lambda_i(y^k) \to \lambda_i(y^*)$ for $i = 1, 2$. $\square$

Nevertheless, we should point out that the proximal distance H given by (28) with ϕ satisfying (C1)–(C4) and $\mathrm{dom}\phi = [0, +\infty)$ generally does not have the property (P7), even if ϕ satisfies the condition (C6). This fact will be illustrated by Example 4.7.

(C6) For any $\{a^k\} \subset (0, +\infty)$ converging to $a \in [0, +\infty)$, $\lim_{k\to\infty} d(a^*, a^k) \to 0$.

Example 4.7. Let H be the proximal distance induced by the entropy function ϕ in Example 4.3. It is easy to verify that ϕ satisfies the conditions (C1)–(C6). Here we shall present a sequence $\{y^k\} \subset \mathrm{int}(\mathcal{K}^3)$ which converges to $y^* \in \mathcal{K}^3$, but where $H(y^*, y^k) \to \infty$. Let

$$y^k = \begin{pmatrix} \sqrt{2(1+\mathrm{e}^{-k^3})} \\ \sqrt{1+k^{-1}-\mathrm{e}^{-k^3}} \\ \sqrt{1-k^{-1}+\mathrm{e}^{-k^3}} \end{pmatrix} \in \mathrm{int}(\mathcal{K}^3) \quad \text{and} \quad y^* = \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix} \in \mathcal{K}^3.$$

By the expression for $H(y^*, y^k)$, i.e., $H(y^*, y^k) = \mathrm{tr}(y^* \circ \ln y^*) - \mathrm{tr}(y^* \circ \ln y^k) + \mathrm{tr}(y^k - y^*)$, it suffices to prove that $\lim_{k\to\infty} -\mathrm{tr}(y^* \circ \ln y^k) = +\infty$ since $\lim_{k\to\infty} \mathrm{tr}(y^k - y^*) = 0$ and $\mathrm{tr}(y^* \circ \ln y^*) = \lambda_2(y^*)\ln(\lambda_2(y^*)) < +\infty$. By the definition of $\ln y^k$, we have

$$\mathrm{tr}(y^* \circ \ln y^k) = \ln(\lambda_1(y^k))\left(y_1^* - (y_2^*)^T \bar{y}_2^k\right) + \ln(\lambda_2(y^k))\left(y_1^* + (y_2^*)^T \bar{y}_2^k\right) \tag{30}$$

for $y^* = (y_1^*, y_2^*),\ y^k = (y_1^k, y_2^k) \in \mathbb{R} \times \mathbb{R}^2$ with $\bar{y}_2^k = y_2^k/\|y_2^k\|$. By computing,

$$\ln(\lambda_1(y^k)) = \ln\sqrt{2} - \ln\left(1 + \sqrt{1 + \mathrm{e}^{-k^3}}\right) - k^3,$$

$$y_1^* - (y_2^*)^T \bar{y}_2^k = \frac{1}{\|y_2^k\|}\left(\frac{-k^{-1} + \mathrm{e}^{-k^3}}{1 + \sqrt{1 + k^{-1} - \mathrm{e}^{-k^3}}} + \frac{k^{-1} - \mathrm{e}^{-k^3}}{1 + \sqrt{1 - k^{-1} + \mathrm{e}^{-k^3}}}\right).$$

The last two equalities imply that $\lim_{k\to\infty} \ln(\lambda_1(y^k))\left(y_1^* - (y_2^*)^T \bar{y}_2^k\right) = -\infty$. In addition, by noting that $y_2^k \neq 0$ for each k, we compute that

$$\lim_{k\to\infty} \ln(\lambda_2(y^k))\left(y_1^* - (y_2^*)^T \bar{y}_2^k\right) = \ln(\lambda_2(y^k))\left(y_1^* + (y_2^*)^T \frac{y_2^*}{\|y_2^*\|}\right) = \lambda_2(y^*)\ln(\lambda_2(y^*)).$$

From the last two equations, we immediately have $\lim_{k\to\infty} -\mathrm{tr}(y^* \circ \ln y^k) = +\infty$.

Thus, when the proximal distance in the IPA is chosen as the one given by (28) with ϕ satisfying (C1)–(C6) and $\mathrm{dom}\phi = [0, +\infty)$, Theorem 3.2(b) may not apply, i.e. the global convergence to an optimal solution may not be guaranteed. This is different from the case for interior proximal methods for convex programming over nonnegative orthant cones as we see by noting that ϕ is now a univariate Bregman function. Similarly, it seems hard to find examples for the class $\mathcal{F}_+(\mathcal{K}^n)$ in [15] such that Theorem 2.2 can apply, since it also requires (P7).

The third approach will produce the proximal distances for the class $\mathcal{F}_2(\mathrm{int}\ \mathcal{K}^n)$, which needs a lsc proper function $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ satisfying the following conditions:

(D1) ϕ is strictly convex and continuous on $\mathrm{dom}\phi$, and ϕ is continuously differentiable on a subset of $\mathrm{dom}\phi$, where $\mathrm{dom}\phi' \subseteq \mathrm{dom}\phi \subseteq [0, +\infty)$ and $\mathrm{int}(\mathrm{dom}\phi') = (0, +\infty)$;
(D2) ϕ is twice continuously differentiable on $\mathrm{int}(\mathrm{dom}\phi)$ and $\lim_{t\to 0^+} \phi''(t) = +\infty$;
(D3) $\phi'(t)t - \phi(t)$ is convex on $\mathrm{dom}\phi'$, and ϕ' is strictly concave on $\mathrm{dom}\phi'$;
(D4) ϕ' is SOC-concave on $\mathrm{dom}\phi'$.

With such a univariate ϕ, we define the proximal distance $H: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$H(x, y) := \begin{cases} \mathrm{tr}(\phi^{\mathrm{soc}}(y)) - \mathrm{tr}(\phi^{\mathrm{soc}}(x)) - \langle \nabla \mathrm{tr}(\phi^{\mathrm{soc}}(x)), y - x\rangle & \forall x \in \mathcal{C}_1, y \in \mathcal{C}_2 \\ +\infty & \text{otherwise} \end{cases} \tag{31}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are the domains of ϕ^{soc} and $(\phi')^{\mathrm{soc}}$, respectively. By the relation between $\mathrm{dom}\phi$ and $\mathrm{dom}\phi'$, obviously, $\mathcal{C}_2 \subseteq \mathcal{C}_1 \subseteq \mathcal{K}^n$ and $\mathrm{int}\ \mathcal{C}_1 = \mathrm{int}\ \mathcal{C}_2 = \mathrm{int}\ \mathcal{K}^n$.

Lemma 4.2. *Let $\phi: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a lsc proper function satisfying* (D1)–(D4). *Then:*

(a) $\mathrm{tr}\left[(\phi')^{\mathrm{soc}}(x) \circ x - \phi^{\mathrm{soc}}(x)\right]$ *is convex in $\mathcal{C}_1$ and continuously differentiable on* $\mathrm{int}\ \mathcal{C}_1$.
(b) *For any fixed $y \in \mathbb{R}^n$, $\langle (\phi')^{\mathrm{soc}}(x), y\rangle$ is continuously differentiable on* $\mathrm{int}\ \mathcal{C}_1$, *and moreover, it is strictly concave over $\mathcal{C}_1$ whenever $y \in \mathrm{int}\ \mathcal{K}^n$.*

Proof. (a) Let $\psi(t) := \phi'(t)t - \phi(t)$. Then, by (D2) and (D3), $\psi(t)$ is convex on $\mathrm{dom}\phi'$ and continuously differentiable on $\mathrm{int}(\mathrm{dom}\phi') = (0, +\infty)$. Since $\mathrm{tr}\left[(\phi')^{\mathrm{soc}}(x) \circ x - \phi^{\mathrm{soc}}(x)\right] = \mathrm{tr}[\psi^{\mathrm{soc}}(x)]$, using Lemma 2.2(b) and (c) immediately yields part (a).
(b) From (D2) and Lemma 2.2(a), $(\phi')^{\mathrm{soc}}(\cdot)$ is continuously differentiable on $\mathrm{int}\ \mathcal{C}_1$. This implies that $\langle y, (\phi')^{\mathrm{soc}}(x)\rangle$ for any fixed y is continuously differentiable on $\mathrm{int}\ \mathcal{C}_1$. We next show that it is also strictly concave in $\mathcal{C}_1$ whenever $y \in \mathrm{int}\ \mathcal{K}^n$. Note that $\mathrm{tr}[(\phi')^{\mathrm{soc}}(\cdot)]$ is strictly concave on $\mathcal{C}_1$ since ϕ' is strictly concave on $\mathrm{dom}\phi'$. Consequently,

$$\mathrm{tr}[(\phi')^{\mathrm{soc}}(\beta x + (1-\beta)z)] > \beta\mathrm{tr}[(\phi')^{\mathrm{soc}}(x)] + (1-\beta)\mathrm{tr}[(\phi')^{\mathrm{soc}}(z)] \quad \forall\, 0 < \beta < 1$$

for any $x, z \in \mathcal{C}_1$ and $x \neq z$. This implies that

$$(\phi')^{\mathrm{soc}}(\beta x + (1-\beta)z) - \beta(\phi')^{\mathrm{soc}}(x) - (1-\beta)(\phi')^{\mathrm{soc}}(z) \neq 0.$$

In addition, since ϕ' is SOC-concave on $\mathrm{dom}\phi'$, from Definition 2.1 it follows that

$$(\phi')^{\mathrm{soc}}[\beta x + (1-\beta)z] - \beta(\phi')^{\mathrm{soc}}(x) - (1-\beta)(\phi')^{\mathrm{soc}}(z) \succeq_{\mathcal{K}^n} 0.$$

Thus, for any fixed $y \in \mathrm{int}\ \mathcal{K}^n$, the last two equations imply that

$$\langle y, (\phi')^{\mathrm{soc}}[\beta x + (1-\beta)z] - \beta(\phi')^{\mathrm{soc}}(x) - (1-\beta)(\phi')^{\mathrm{soc}}(z)\rangle > 0.$$

This shows that $\langle y, (\phi')^{\mathrm{soc}}(x)\rangle$ for any fixed $y \in \mathrm{int}\ \mathcal{K}^n$ is strictly convex on $\mathcal{C}_1$. □

Using the conditions (D1)–(D4) and Lemma 4.2, and following the same arguments as for Propositions 4.1 and 4.2 of [27], we may prove the following proposition.

Proposition 4.4. *Let H be defined as in (31) with ϕ satisfying* (D1)–(D4). *Then:*

(a) $H(x,y)\geq 0$ *for any* $x,y\in\mathbb{R}^n$, *and* $H(y,y)=0$ *for any* $y\in\operatorname{int}\mathcal{K}^n$.
(b) *For any fixed* $y\in\mathcal{C}_2$, $H(\cdot,y)$ *is continuous in* $\mathcal{C}_1$, *and it is strictly convex on* $\mathcal{C}_1$ *whenever* $y\in\operatorname{int}\mathcal{K}^n$.
(c) *For any fixed* $y\in\mathcal{C}_2$, $H(\cdot,y)$ *is continuously differentiable on* $\operatorname{int}\mathcal{K}^n$ *with*

$$\nabla_1 H(x,y)=2\nabla(\phi')^{\mathrm{soc}}(x)(x-y). \tag{32}$$

Moreover, $\operatorname{dom}\nabla_1 H(\cdot,y)=\operatorname{int}\mathcal{K}^n$ *whenever* $y\in\operatorname{int}\mathcal{K}^n$.
(d) $H(x,y)\geq\sum_{i=1}^{2}d(\lambda_i(y),\lambda_i(x))\geq 0$ *for any* $x\in\mathcal{C}_1$ *and* $y\in\mathcal{C}_2$.
(e) *For any fixed* $y\in\mathcal{C}_2$, *the sets* $\{x\in\mathcal{C}_1\colon H(x,y)\leq\gamma\}$ *are bounded for all* $\gamma\in\mathbb{R}$.
(f) *For all* $x,y\in\operatorname{int}\mathcal{K}^n$ *and* $z\in\mathcal{C}_2$, $H(x,z)-H(y,z)\geq 2\langle\nabla_1 H(y,x),\ z-y\rangle$.

Proposition 4.4 demonstrates that the function H defined by (31) with ϕ satisfying (D1)–(D4) is a proximal distance w.r.t. the cone $\operatorname{int}\mathcal{K}^n$ and possesses the property (P5′), and therefore belongs to the class $\mathcal{F}_2(\operatorname{int}\mathcal{K}^n)$. If, in addition, $\operatorname{dom}\phi=[0,+\infty)$, then H belongs to the class $\mathcal{F}_2(\mathcal{K}^n)$. The conditions (D1)–(D3) are easy to check, and for the condition (D4), we can employ the characterizations in [28,29] to verify whether ϕ' is SOC-concave or not. Some examples are presented as follows.

Example 4.8. Let $\phi(t)=t\ln t-t+1$ if $t\geq 0$, and otherwise $\phi(t)=+\infty$. It is easy to verify that ϕ satisfies (D1)–(D3) with $\operatorname{dom}\phi=[0,+\infty)$ and $\operatorname{dom}\phi'=(0,+\infty)$. By Lemma 2.4(c), ϕ' is SOC-concave on $(0,+\infty)$. Using formulas (8) and (31), we have

$$H(x,y)=\begin{cases}\operatorname{tr}(y\circ\ln y-y\circ\ln x+x-y) & \forall x\in\operatorname{int}\mathcal{K}^n,\ y\in\mathcal{K}^n;\\ +\infty & \text{otherwise}.\end{cases} \tag{33}$$

Example 4.9. Take $\phi(t)=\frac{t^{q+1}}{q+1}$ if $t\geq 0$, and otherwise $\phi(t)=+\infty$, where $0<q<1$. It is easy to show that ϕ satisfies (D1)–(D3) with $\operatorname{dom}\phi=[0,+\infty)$ and $\operatorname{dom}\phi'=[0,+\infty)$. By Lemma 2.4(a), ϕ' is also SOC-concave on $[0,+\infty)$. By (8) and (31), we compute that

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

Example 4.10. Take $\phi(t)=(1+t)\ln(1+t)+\frac{t^{q+1}}{q+1}$ if $t\geq 0$, and otherwise $\phi(t)=+\infty$, where $0<q<1$. We can verify that ϕ satisfies (D1)–(D3) with $\operatorname{dom}\phi=\operatorname{dom}\phi'=[0,+\infty)$. From Lemma 2.4(a) and (c), ϕ' is also SOC-concave on $[0,+\infty)$. Using (8) and (31), it is not hard to compute that for any $x,y\in\mathcal{K}^n$,

$$H(x,y)=\operatorname{tr}\left[(e+y)\circ(\ln(e+y)-\ln(e+x))\right]-\operatorname{tr}(y-x)+\frac{1}{q+1}\operatorname{tr}(y^{q+1})+\frac{q}{q+1}\operatorname{tr}(x^{q+1})-\operatorname{tr}(x^q\circ y).$$

Note that the proximal distances in Examples 4.9 and 4.10 belong to the class $\mathcal{F}_2(\mathcal{K}^n)$. By Proposition 4.5, the ones in Examples 4.9 and 4.10 also belong to the class $\widehat{\mathcal{F}}_2(\mathcal{K}^n)$.

Proposition 4.5. *Let H be defined as in (31) with ϕ satisfying* (D1)–(D4). *Suppose that* $\operatorname{dom}\phi=\operatorname{dom}\phi'=[0,+\infty)$. *Then, H possesses the properties* (P7′) *and* (P8′).

Proof. By the given assumption, $\mathcal{C}_1=\mathcal{C}_2=\mathcal{K}^n$. From Proposition 4.4 (b), the function $H(\cdot,y^*)$ is continuous on $\mathcal{K}^n$. Consequently, $\lim_{k\to\infty}H(y^k,y^*)=H(y^*,y^*)=0$.

From Proposition 4.4(d), $H(y^k,y^*)\geq d(\lambda_i(y^*),\lambda_i(y^k))\geq 0$ for $i=1,2$. This together with the assumption $H(y^k,y^*)\to 0$ implies $d(\lambda_i(y^*),\lambda_i(y^k))\to 0$ for $i=1,2$. From this, we necessarily have $\lambda_i(y^k)\to\lambda_i(y^*)$ for $i=1,2$. Suppose not; then the bounded sequence $\{\lambda_i(y^k)\}$ must have another limit point $\nu_i^*\geq 0$ such that $\nu_i^*\neq\lambda_i(y^*)$. Without loss of generality, we assume that $\lim_{k\in K,k\to\infty}\lambda_i(y^k)=\nu_i^*$. Then, we have

$$d(\nu_i^*,\lambda_i(y^*))=\lim_{k\to\infty}d(\nu_i^*,\lambda_i(y^k))=\lim_{k\in K,k\to\infty}d(\nu_i^*,\lambda_i(y^k))=d(\nu_i^*,\nu_i^*)=0$$

where the first equality is due to the continuity of $d(s,\cdot)$ for any fixed $s\in[0,+\infty)$, and the second one is due to the convergence of $\{d(\nu_i^*,\lambda_i(y^k))\}$ implied by the first equality. This contradicts the fact that $d(\nu_i^*,\lambda_i(y^*))>0$ since $\nu_i^*\neq\lambda_i(y^*)$. $\square$

As illustrated by the following example, the proximal distance generated by (31) with ϕ satisfying (D1)–(D4) generally does not belong to the class $\bar{\mathcal{F}}_2(\mathcal{K}^n)$.

Example 4.11. Let H be the proximal distance in Example 4.8. Let

$$y^k = \begin{pmatrix} \sqrt{2} \\ (-1)^k \dfrac{k}{k+1} \\ (-1)^k \dfrac{k}{k+1} \end{pmatrix} \quad \text{for each } k \quad \text{and} \quad y^* = \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix}.$$

It is not hard to check that the sequence $\{y^k\} \subseteq \text{int}(\mathcal{K}^3)$ satisfies $H(y^k, y^*) \to 0$. Clearly, the sequence $y^k \nrightarrow y^*$ as $k \to \infty$, but $\lambda_1(y^k) \to \lambda_1(y^*) = 0$ and $\lambda_2(y^k) \to \lambda_2(y^*) = 2\sqrt{2}$.

Finally, let H_1 be a proximal distance produced via one of the approaches above, and define

$$H_\alpha(x, y) := H_1(x, y) + \frac{\alpha}{2}\|x - y\|^2 \tag{34}$$

where $\alpha > 0$ is a fixed parameter. Then, by Propositions 4.1, 4.2 and 4.4 and the identity

$$\|z - x\|^2 = \|z - y\|^2 + \|y - x\|^2 + 2\langle z - y, y - x\rangle, \quad \forall x, y, z \in \mathbb{R}^n,$$

it is easily shown that H_α is also a proximal distance w.r.t. int $\mathcal{K}^n$. In particular, when H_1 is given by (31) with ϕ satisfying (D1)–(D4) and domϕ = dom$\phi' = [0, +\infty)$ (for example the distances in Examples 4.9 and 4.10), the regularized proximal distance H_α satisfies (P7′) and (P9′), and hence $H_\alpha \in \bar{F}_2(\mathcal{K}^n)$. With such a regularized proximal distance, the sequence generated by the IPA converges to an optimal solution of (1) if $X_* \neq \emptyset$.

To sum up, we may construct a proximal distance w.r.t. the cone int $\mathcal{K}^n$ in three ways with an appropriate univariate function. The first approach in (23) can only produce a proximal distance belonging to $\mathcal{F}_1(\text{int}\,\mathcal{K}^n)$, the second approach in (28) produces a proximal distance of $\mathcal{F}_1(\mathcal{K}^n)$ if dom$\phi = [0, +\infty)$, whereas the third approach in (31) produces a proximal distance of the class $\widehat{\mathcal{F}_2}(\mathcal{K}^n)$ if domϕ = dom$\phi' = [0, +\infty)$. In particular, the regularized proximal distances H_α in (34) with H_1 given by (31) with domϕ = dom$\phi' = [0, +\infty)$ belong to the smallest class $\bar{F}_2(\mathcal{K}^n)$. With such regularized proximal distances, we have the convergence result of Theorem 3.2(c) for the general convex SOCP with $X_* \neq \emptyset$.

5. Central paths and interior proximal methods

In this section, for the linear SOCP, we will obtain some improved convergence results for the IPA by exploring the relations between the sequence generated by the IPA and the central path associated with the corresponding proximal distances.

Given a lsc proper strictly convex function Φ with dom$\Phi \subseteq \mathcal{K}^n$ and int(domΦ) = int $\mathcal{K}^n$, the *central path* of (1) associated with Φ is the set $\{x(\tau) : \tau > 0\}$ defined by

$$x(\tau) := \text{argmin}\{\tau f(x) + \Phi(x) \mid x \in \mathcal{V} \cap \mathcal{K}^n\} \quad \text{for } \tau > 0. \tag{35}$$

In what follows, we will focus on the central path of (1) w.r.t. a distance-like function $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$. From Definition 3.1 and Proposition 3.1, we immediately have the following result.

Proposition 5.1. *For any given $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$ and $\bar{x} \in \text{int}\,\mathcal{K}^n$, the central path $\{x(\tau) : \tau > 0\}$ associated with $H(\cdot, \bar{x})$ is well-defined and is in $\mathcal{V} \cap \text{int}\,\mathcal{K}^n$. For each $\tau > 0$, there exists $g_\tau \in \partial f(x(\tau))$ such that $\tau g_\tau + \nabla_1 H(x(\tau), \bar{x}) = A^T y(\tau)$ for some $y(\tau) \in \mathbb{R}^m$.*

We next study the favorable properties of the central path associated with $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$.

Proposition 5.2. *For any given $H \in \mathcal{D}(\text{int}\,\mathcal{K}^n)$ and $\bar{x} \in \text{int}\,\mathcal{K}^n$, let $\{x(\tau) : \tau > 0\}$ be the central path associated with $H(\cdot, \bar{x})$. Then, the following results hold:*

(a) *The function $H(x(\tau), \bar{x})$ is nondecreasing in τ.*
(b) *The set $\{x(\tau) : \hat{\tau} \le \tau \le \tilde{\tau}\}$ is bounded for any given $0 < \hat{\tau} < \tilde{\tau}$.*
(c) *$x(\tau)$ is continuous at any $\tau > 0$.*
(d) *The set $\{x(\tau) : \tau \ge \bar{\tau}\}$ is bounded for any $\bar{\tau} > 0$ if $X_* \neq \emptyset$ and dom$H(\cdot, \bar{x}) = \mathcal{K}^n$.*
(e) *All cluster points of $\{x(\tau) : \tau > 0\}$ are solutions of (1) if $X_* \neq \emptyset$.*

Proof. The proofs are similar to those of Propositions 3–5 of [25].
(a) Take $\tau_1, \tau_2 > 0$ and let $x^i = x(\tau_i)$ for $i = 1, 2$. Then, from Proposition 5.1, $x^1, x^2 \in \mathcal{V} \cap \text{int}\,\mathcal{K}^n$ and there exist $g^1 \in \partial f(x^1)$ and $g^2 \in \partial f(x^2)$ such that

$$\nabla_1 H(x^1, \bar{x}) = -\tau_1 g^1 + A^T y^1 \quad \text{and} \quad \nabla_1 H(x^2, \bar{x}) = -\tau_2 g^2 + A^T y^2 \tag{36}$$

for some $y^1, y^2 \in \mathbb{R}^m$. This together with the convexity of $H(\cdot, \bar{x})$ yields that

$$\tau_1^{-1}\left(H(x^1,\bar{x}) - H(x^2,\bar{x})\right) \le \tau_1^{-1}\langle \nabla_1 H(x^1,\bar{x}), x^1 - x^2\rangle = \langle g^1, x^2 - x^1\rangle,$$
$$\tau_2^{-1}\left(H(x^2,\bar{x}) - H(x^1,\bar{x})\right) \le \tau_2^{-1}\langle \nabla_1 H(x^2,\bar{x}), x^2 - x^1\rangle = \langle g^2, x^1 - x^2\rangle. \tag{37}$$

Adding the two inequalities and using the convexity of f, we obtain

$$\left(\tau_1^{-1} - \tau_2^{-1}\right)\left(H(x^1,\bar{x}) - H(x^2,\bar{x})\right) \le \langle g^1 - g^2, x^2 - x^1\rangle \le 0.$$

Thus, $H(x^1,\bar{x}) \le H(x^2,\bar{x})$ whenever $\tau_1 \le \tau_2$. In particular, from the last two equations,

$$\begin{aligned} 0 \le \tau_1^{-1}\left[H(x^1,\bar{x}) - H(x^2,\bar{x})\right] &\le \tau_1^{-1}\langle \nabla_1 H(x^1,\bar{x}), x^1 - x^2\rangle \le \langle g^2, x^2 - x^1\rangle \\ &\le \tau_2^{-1}\left[H(x^1,\bar{x}) - H(x^2,\bar{x})\right], \quad \forall \tau_1 \ge \tau_2 > 0. \end{aligned} \tag{38}$$

(b) By part (a), $H(x(\tau),\bar{x}) \le H(x(\tilde{\tau}),\bar{x})$ for any $\tau \le \tilde{\tau}$, which implies that

$$\{x(\tau) \colon \tau \le \tilde{\tau}\} \subseteq L_1 = \left\{x \in \operatorname{int}\mathcal{K}^n \mid H(x,\bar{x}) \le H(x(\tilde{\tau}),\bar{x})\right\}.$$

Noting that $\{x(\tau)\colon \hat{\tau} \le \tau \le \tilde{\tau}\} \subseteq \{x(\tau)\colon \tau \le \tilde{\tau}\} \subseteq L_1$, the desired result follows by (P4).
(c) Fix $\bar{\tau} > 0$. To prove that $x(\tau)$ is continuous at $\bar{\tau}$, it suffices to prove that $\lim_{k\to\infty} x(\tau_k) = x(\bar{\tau})$ for any sequence $\{\tau_k\}$ such that $\lim_{k\to\infty} \tau_k = \bar{\tau}$. Given such a sequence $\{\tau_k\}$, take $\hat{\tau}, \tilde{\tau}$ such that $\hat{\tau} > \bar{\tau} > \tilde{\tau}$. Then, $\{x(\tau)\colon \hat{\tau} \le \tau \le \tilde{\tau}\}$ is bounded by part (b), and $\tau_k \in (\hat{\tau}, \tilde{\tau})$ for sufficiently large k. Consequently, the sequence $\{x(\tau_k)\}$ is bounded. Let $\bar{y}$ be a cluster point of $\{x(\tau_k)\}$, and without loss of generality assume that $\lim_{k\to\infty} x(\tau_k) = \bar{y}$. Let $K_1 := \{k : \tau_k \le \bar{\tau}\}$ and take $k \in K_1$. Then, from (38) with $\tau_1 = \bar{\tau}$ and $\tau_2 = \tau_k$,

$$\begin{aligned} 0 \le \bar{\tau}^{-1}\left[H(x(\bar{\tau}),\bar{x}) - H(x(\tau_k),\bar{x})\right] &\le \bar{\tau}^{-1}\langle \nabla_1 H(x(\bar{\tau}),\bar{x}), x(\bar{\tau}) - x(\tau_k)\rangle \\ &\le \tau_k^{-1}\left[H(x(\bar{\tau}),\bar{x}) - H(x(\tau_k),\bar{x})\right]. \end{aligned}$$

If K_1 is infinite, taking the limit $k \to \infty$ with $k \in K_1$ in the last inequality and using the continuity of $H(\cdot,\bar{x})$ on $\operatorname{int}\mathcal{K}^n$ yields that

$$H(x(\bar{\tau}),\bar{x}) - H(\bar{y},\bar{x}) = \langle \nabla_1 H(x(\bar{\tau}),\bar{x}), x(\bar{\tau}) - \bar{y}\rangle.$$

This together with the strict convexity of $H(\cdot,\bar{x})$ implies $x(\bar{\tau}) = \bar{y}$. If K_1 is finite, then $K_2 := \{k : \tau_k \ge \bar{\tau}\}$ must be infinite. Using the same arguments, we also have $x(\bar{\tau}) = \bar{y}$.
(d) By (P3) and Proposition 5.1, there exists $g_\tau \in \partial f(x(\tau))$ such that for any $z \in \mathcal{V} \cap \mathcal{K}^n$,

$$H(x(\tau),\bar{x}) - H(z,\bar{x}) \le \tau^{-1}\langle \nabla_1 H(x(\tau),\bar{x}), x(\tau) - z\rangle = \langle g_\tau, z - x(\tau)\rangle. \tag{39}$$

In particular, taking $z = x^* \in X_*$ in the last equality and using the fact that

$$0 \ge f(x^*) - f(x(\tau)) \ge \langle g_\tau, x^* - x(\tau)\rangle,$$

we have $H(x(\tau),\bar{x}) - H(x^*,\bar{x}) \le 0$. Hence, $\{x(\tau)\colon \tau > \bar{\tau}\} \subset \{x \in \operatorname{int}\mathcal{K}^n \mid H(x,\bar{x}) \le H(x^*,\bar{x})\}$. By (P4), the latter is bounded, and the desired result then follows.
(e) Let $\hat{x}$ be a cluster point of $\{x(\tau)\}$ and $\{\tau_k\}$ be a sequence such that $\lim_{k\to\infty} \tau_k = +\infty$ and $\lim_{k\to\infty} x(\tau_k) = \hat{x}$. Write $x^k := x(\tau_k)$ and take $x^* \in X_*$ and $z \in \mathcal{V} \cap \operatorname{int}\mathcal{K}^n$. Then, for any $\epsilon > 0$, we have $x(\epsilon) := (1-\epsilon)x^* + \epsilon z \in \mathcal{V} \cap \operatorname{int}\mathcal{K}^n$. From the property (P3),

$$\langle \nabla_1 H(x(\epsilon),\bar{x}) - \nabla_1 H(x^k,\bar{x}), x^k - x(\epsilon)\rangle \le 0.$$

On the other hand, taking $z = x(\epsilon)$ in (39), we readily have

$$\tau_k^{-1}\langle \nabla_1 H(x^k,\bar{x}), x^k - x(\epsilon)\rangle = \langle g^k, x(\epsilon) - x^k\rangle$$

with $g^k \in \partial f(x^k)$. Combining the last two equations, we obtain

$$\tau_k^{-1}\langle \nabla_1 H(x(\epsilon),\bar{x}), x^k - x(\epsilon)\rangle \le \langle g^k, x(\epsilon) - x^k\rangle.$$

Since the subdifferential set $\partial f(x^k)$ for each k is compact and $g^k \in \partial f(x^k)$, the sequence $\{g^k\}$ is bounded. Taking the limit in the last inequality yields $0 \le \langle \hat{g}, x(\epsilon) - \hat{x}\rangle$, where $\hat{g}$ is a limit point of $\{g^k\}$, and by Theorem 24.4 of [30], $\hat{g} \in \partial f(\hat{x})$. Taking the limit $\epsilon \to 0$ in the inequality, we get $0 \le \langle \hat{g}, x^* - \hat{x}\rangle$. This implies that $f(\hat{x}) \le f(x^*)$ since $x^* \in X_*$ and $\hat{g} \in \partial f(\hat{x})$. Consequently, $\hat{x}$ is a solution of the CSOCP (1). □

In particular, from the following theorem, we also have that the central path is convergent if $H \in \mathcal{D}(\operatorname{int}\mathcal{K}^n)$ satisfies $\operatorname{dom}H(\cdot,\bar{x}) = \mathcal{K}^n$, where $\bar{x} \in \operatorname{int}\mathcal{K}^n$ is a given point. Notice that $H(\cdot,\bar{x})$ is continuous on $\operatorname{dom}H(\cdot,\bar{x})$ by (P2), and hence the assumption for H is equivalent to saying that $H(\cdot,\bar{x})$ is continuous at the boundary of the cone $\mathcal{K}^n$.

Theorem 5.1. *For any given $\bar{x} \in \operatorname{int} \mathcal{K}^n$ and $H \in \mathcal{D}(\operatorname{int} \mathcal{K}^n)$ with $\operatorname{dom} H(\cdot, \bar{x}) = \mathcal{K}^n$, let $\{x(\tau): \tau > 0\}$ be the central path associated with $H(\cdot, \bar{x})$. If X_* is nonempty, then $\lim_{\tau\to+\infty} x(\tau)$ exists and is the unique solution of $\min\{H(x, \bar{x}) \mid x \in X_*\}$.*

Proof. Let $\hat{x}$ be a cluster point of $\{x(\tau)\}$ and $\{\tau_k\}$ be such that $\lim_{k\to\infty} \tau_k = +\infty$ and $\lim_{k\to\infty} x(\tau_k) = \hat{x}$. Then, for any $x \in X_*$, using (38) with $x^1 = x(\tau_k)$ and $x^2 = x$, we get

$$[H(x(\tau_k), \bar{x}) - H(x, \bar{x})] \le \tau_k \langle g^k, x - x(\tau_k)\rangle \le \tau_k [f(x) - f(x(\tau_k))] \le 0$$

where the second inequality holds since $g^k \in \partial f(x(\tau_k))$, and the last one is due to $x \in X_*$. Taking the limit $k \to \infty$ in the last inequality and using the continuity of $H(\cdot, \bar{x})$, we have $H(\hat{x}, \bar{x}) \le H(x, \bar{x})$ for all $x \in X_*$. Since $\hat{x} \in X_*$ by Proposition 5.2(e), this shows that any cluster point of $\{x(\tau): \tau > 0\}$ is a solution of $\min\{H(x, \bar{x}) \mid x \in X_*\}$. By the uniqueness of the solution of $\min\{H(x, \bar{x}) \mid x \in X_*\}$, we have $\lim_{\tau\to+\infty} x(\tau) = x^*$. □

For the linear SOCP, we may establish the relations between the sequence generated by the IPA and the central path associated with the corresponding distance-like functions.

Proposition 5.3. *For the linear SOCP, let $\{x^k\}$ be the sequence generated by the IPA with $H \in \mathcal{D}(\operatorname{int} \mathcal{K}^n)$, $x^0 \in \mathcal{V} \cap \operatorname{int} \mathcal{K}^n$ and $\epsilon_k \equiv 0$, and $\{x(\tau): \tau > 0\}$ be the central path associated with $H(\cdot, x^0)$. Then, $x^k = x(\tau_k)$ for $k = 1, 2, \ldots$ under either of the following conditions:*

(a) *H is constructed via (23) or (28), and $\{\tau_k\}$ is given by $\tau_k = \sum_{j=0}^{k} \lambda_j$ for $k = 1, 2, \ldots$;*
(b) *H is constructed via (31), the mapping $\nabla(\phi')^{\text{soc}}(\cdot)$ defined on $\operatorname{int} \mathcal{K}^n$ maps any vector $\mathbb{R}^n$ into $\operatorname{Im} A^T$, and the sequence $\{\tau_k\}$ is given by $\tau_k = \lambda_k$ for $k = 1, 2, \ldots$.*

Moreover, for any positive increasing sequence $\{\tau_k\}$, there exists a positive sequence $\{\lambda_k\}$ with $\sum_{k=1}^{\infty} \lambda_k = +\infty$ such that the proximal sequence $\{x^k\}$ satisfies $x_k = x(\tau_k)$.

Proof. (a) Suppose that H is constructed via (23). From (13) and Proposition 4.1(b),

$$\lambda_j c + \nabla\phi(\det(x^j)) - \nabla\phi(\det(x^{j-1})) = A^T u^j \quad \text{for } j = 0, 1, 2, \ldots. \tag{40}$$

Summing the equality from $j = 0$ to k and taking $\tau_k = \sum_{j=0}^{k} \lambda_j$, $y^k = \sum_{j=0}^{k} u^j$, we get

$$\tau_k c + \nabla\phi(\det(x^k)) - \nabla\phi(\det(x^0)) = A^T y^k.$$

This means that x^k satisfies the optimal conditions of the problem

$$\min\left\{\tau_k f(x) + H(x, x^0) \mid x \in \mathcal{V} \cap \operatorname{int} \mathcal{K}^n\right\}, \tag{41}$$

and so $x^k = x(\tau_k)$. Now let $\{x(\tau): \tau > 0\}$ be the central path. Take a positive increasing sequence $\{\tau_k\}$ and let $x^k \equiv x(\tau_k)$. Then from Propositions 4.1 and 5.1(b), it follows that

$$\tau_k c + \nabla\phi(\det(x^k)) - \nabla\phi(\det(x^0)) = A^T y^k \quad \text{for some } y^k \in \mathbb{R}^m.$$

Setting $\lambda_k = \tau_k - \tau_{k-1}$ and $u^k = y^k - y^{k-1}$, from the last equality it follows that

$$\lambda_k c + \nabla\phi(\det(x^k)) - \nabla\phi(\det(x^{k-1})) = A^T u^k.$$

This shows that $\{x^k\}$ is the sequence generated by the IPA with $\epsilon_k \equiv 0$. If H is given by (28), using Proposition 4.2(b) and the same arguments, we also have that the result holds.
(b) For this case, by Proposition 4.4(c), the above (40) becomes

$$\lambda_j c + \nabla(\phi')^{\text{soc}}(x^j) \cdot (x^j - x^{j-1}) = A^T u^j \quad \text{for } j = 0, 1, 2, \ldots.$$

Since $\phi''(t) > 0$ for all $t \in (0, +\infty)$ by (D1) and (D2), from Proposition 5.2 of [8] it follows that $\nabla(\phi')^{\text{soc}}(x)$ is positive definite on $\operatorname{int} \mathcal{K}^n$. Thus, the last equality is equivalent to

$$\left[\nabla(\phi')^{\text{soc}}(x^j)\right]^{-1} \lambda_j c + (x^j - x^{j-1}) = \left[\nabla(\phi')^{\text{soc}}(x^j)\right]^{-1} A^T u^j \quad \text{for } j = 0, 1, 2, \ldots. \tag{42}$$

Summing the equality (42) from $j = 0$ to k and making a suitable arrangement, we get

$$\lambda_k c + \nabla(\phi')^{\text{soc}}(x^k)(x^k - x^0) = A^T u^k + \nabla(\phi')^{\text{soc}}(x^k) \sum_{j=0}^{k-1} \left[\nabla(\phi')^{\text{soc}}(x^j)\right]^{-1} (A^T u^j - \lambda_j c),$$

which, using the given assumptions and setting $\tau_k = \lambda_k$, reduces to

$$\tau_k c + \nabla(\phi')^{\text{soc}}(x^k)(x^k - x^0) = A^T \bar{y}^k \quad \text{for some } \bar{y}^k \in \mathbb{R}^m.$$

This means that x^k is the unique solution of (41), and hence $x^k = x(\tau_k)$ for any k. Let $\{x(\tau) : \tau > 0\}$ be the central path. Take a positive increasing sequence $\{\tau_k\}$ and define the sequence $x^k = x(\tau_k)$. Then, from Propositions 4.4 and 5.1(c),

$$\tau_k c + \nabla(\phi')^{\rm soc}(x^k)(x^k - x^0) = A^T y^k \quad \text{for some } y^k \in \mathbb{R}^m,$$

which, by the positive definiteness of $\nabla(\phi')^{\rm soc}(\cdot)$ on $\operatorname{int}\mathcal{K}^n$, implies that

$$[\nabla(\phi')^{\rm soc}(x^k)]^{-1}(\tau_k c - A^T y^k) + [\nabla(\phi')^{\rm soc}(x^{k-1})]^{-1}(\tau_{k-1} c - A^T y^{k-1}) + (x^k - x^{k-1}) = 0.$$

Consequently,

$$\tau_k c + \nabla(\phi')^{\rm soc}(x^k)(x^k - x^{k-1}) = \nabla(\phi')^{\rm soc}(x^k)[\nabla(\phi')^{\rm soc}(x^{k-1})]^{-1}(A^T y^{k-1} - \tau_{k-1} c).$$

Using the given assumptions and setting $\lambda_k = \tau_k$, we have

$$\lambda_k c + \nabla(\phi')^{\rm soc}(x^k)(x^k - x^{k-1}) = A^T u^k \quad \text{for some } u^k \in \mathbb{R}^m$$

for some $u^k \in \mathbb{R}^m$. This implies that $\{x^k\}$ is the sequence generated by the IPA and the sequence $\{\lambda_k\}$ satisfies $\sum_{k=1}^{\infty} \lambda_k = +\infty$ since $\{\tau_k\}$ is a positive increasing sequence. □

From Theorem 5.1 and Proposition 5.3, we readily have the following improved convergence results for the sequence generated by the IPA for the linear SOCP.

Theorem 5.2. *For the linear SOCP, let $\{x^k\}$ be the sequence generated by the IPA with $H \in \mathcal{D}(\operatorname{int}\mathcal{K}^n)$, $x^0 \in \mathcal{V} \cap \operatorname{int}\mathcal{K}^n$ and $\epsilon_k \equiv 0$. If one of the following conditions is satisfied:*

(a) *H is constructed via (28) with $\operatorname{dom}H(\cdot, x^0) = \mathcal{K}^n$ and $\sum_{k=0}^{\infty} \lambda_k = +\infty$;*
(b) *H is constructed via (31) with $\operatorname{dom}H(\cdot, x^0) = \mathcal{K}^n$, the mapping $\nabla(\phi')^{\rm soc}(\cdot)$ defined on $\operatorname{int}\mathcal{K}^n$ maps any vector in $\mathbb{R}^n$ into $\operatorname{Im}A^T$, and $\lim_{k\to\infty} \lambda_k = +\infty$;*

and $X_* \neq \emptyset$, then $\{x^k\}$ converges to the unique solution of $\min\{H(x, x^0) \mid x \in X_*\}$.

6. Conclusions

We have extended the unified analysis technique given in [15] for interior proximal methods for solving the convex SOCP and presented three simple and effective ways to construct a proximal distance w.r.t. the cone $\operatorname{int}\mathcal{K}^n$. The advantages and disadvantages of the corresponding proximal distances were analyzed and illustrated with some examples. In particular, a class of regularized proximal distances was constructed, for which the global convergence result of Theorem 3.2(c) can apply. However, for the class of proximal distances $\mathcal{F}_+(\mathcal{K}^n)$ in [15], as illustrated in Section 4, it seems hard to find examples such that global

convergence results similar to those for [15, Theorem 2.2] can apply for them.

In addition, we have also made investigations for the central paths of (1) associated with these proximal-like functions, and for the linear SOCP, established the relations between the central paths and the sequence generated by the interior proximal methods, from which we, in particular, obtain the global convergence of the sequence under the usual assumptions and the continuity of $H(\cdot, x^0)$ at the boundary of second-order cones.

Acknowledgements

The first author's work was supported by the National Young Natural Science Foundation (No. 10901058) and Guangdong Natural Science Foundation (No. 9251802902000001). The second author's work was partially supported by the National Science Council of Taiwan.

References

[1] J. Faraut, A. Korányi, Analysis on Symmetric Cones, in: Oxford Mathematical Monographs, Oxford University Press, New York, 1994.
[2] F. Alizadeh, D. Goldfarb, Second-order cone programming, Mathematical Programming 95 (2003) 3–51.
[3] M.S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret, Application of second-order cone programming, Linear Algebra and its Applications 284 (1998) 193–228.
[4] E.D. Andersen, C. Roos, T. Terlaky, On implementing a primal–dual interior-point method for conic quadratic optimization, Mathematical Programming 95 (2003) 249–277.
[5] R.D.C. Monteiro, T. Tsuchiya, Polynomial convergence of primal–dual algorithms for the second-order cone programs based on the MZ-family of directions, Mathematical Programming 88 (2000) 61–83.
[6] T. Tsuchiya, A convergence analysis of the scaling-invariant primal–dual path-following algorithms for second-order cone programming, Optimization Methods and Software 11 (1999) 141–182.
[7] X.-D. Chen, D. Sun, J. Sun, Complementarity functions and numerical experiments for second-order cone complementarity problems, Computational Optimization and Applications 25 (2003) 39–56.
[8] M. Fukushima, Z.-Q. Luo, P. Tseng, Smoothing functions for second-order cone complementarity problems, SIAM Journal on Optimization 12 (2002) 436–460.

[9] S. Hayashi, N. Yamashita, M. Fukushima, A combined smoothing and regularization method for monotone second-order cone complementarity problems, SIAM Journal on Optimization 15 (2005) 593–615.
[10] C. Kanzow, I. Ferenczi, M. Fukushima, On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity, SIAM Journal on Optimization 20 (2009) 297–320.
[11] J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Mathematical Programming 104 (2005) 293–327.
[12] B. Martinet, Perturbation des methodes *d*'optimisation, application, RAIRO Analyse Numerique 12 (1978) 153–171.
[13] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM Journal on Control and Optimization 14 (1976) 877–898.
[14] A. Auslender, M. Haddou, An interior–proximal method for convex linearly constrained problems and its extension to variational inequalities, Mathematical Programming 71 (1995) 77–100.
[15] A. Auslender, M. Teboulle, Interior gradient and proximal methods for convex and conic optimization, SIAM Journal on Optimization 16 (2006) 697–725.
[16] Y. Censor, S.A. Zenios, The proximal minimization algorithm with *D*-functions, Journal of Optimization Theory and Applications 73 (1992) 451–464.
[17] G. Chen, M. Teboulle, Convergence analysis of a proximal-like minimization algorithm using Bregman functions, SIAM Journal on Optimization 3 (1993) 538–543.
[18] J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, Mathematics of Operations Research 18 (1993) 202–226.
[19] A.N. Iusem, M. Teboulle, Convergence rate analysis of nonquadratic proximal and augmented Lagrangian methods for convex and linear programming, Mathematics of Operations Research 20 (1995) 657–677.
[20] M. Teboulle, Entropic proximal mappings with applications to nonlinear programming, Mathematics of Operations Research 17 (1992) 670–690.
[21] J.X. da Cruz Neto, O.P. Ferreira, P.R. Oliveira, R.C.M. Silva, Central paths in semidefinite programming, generalized proximal point method and Cauchy trajectories in Riemannian manifolds, Journal of Optimization Theory and Applications 139 (2008) 227–242.
[22] M. Halická, E. de Klerk, C. Roos, On the convergence of the central path in semidefinite optimization, SIAM Journal on Optimization 12 (2002) 1090–1099.
[23] M. Halická, E. de Klerk, C. Roos, Limiting behavior of the central path in semidefinite optimization, Optimization Methods and Software 20 (2005) 99–113.
[24] L.M. Graña Drummond, H.Y. Peterzil, The central path in smooth convex semidefinite programs, Optimization 51 (2002) 207–233.
[25] A.N. Iusem, B.F. Svaiter, J.X. da Cruz Neto, Central paths, generalized proximal point methods and Cauchy trajectories in Riemannian manifolds, SIAM Journal on Control and Optimization 2 (1999) 566–588.
[26] J.-S. Chen, T.-K. Liao, S.-H. Pan, An application of Schur complementarity theorem for SOC functions, Linear Algebra and its Applications (2009) (submitted for publication).
[27] S.-H. Pan, J.-S. Chen, A class of interior proximal-like algorithms for convex second-order cone programming, SIAM Journal on Optimization 19 (2008) 883–910.
[28] J.-S. Chen, The convex and monotone functions associated with second-order cone, Optimization 55 (2006) 363–385.
[29] J.-S. Chen, X. Chen, S.-H. Pan, J.-W. Zhang, The characterizations of SOC-monotone and SOC-convex function, Journal of Global Optimization 45 (2009) 259–279.
[30] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, New Jersey, 1970.
[31] J.-M. Peng, C. Roos, T. Terlaky, Self-regular functions and new search directions for linear and semi-definite optimization, Mathematical Programming 93 (2002) 129–171.
[32] S.-H. Pan, J.-S. Chen, Proximal-like algorithm using quasi *D*-function for convex second-order cone programming, Journal of Optimization Theory and Applications 138 (2008) 95–113.