Comput Optim Appl (2010) 45: 581–606
DOI 10.1007/s10589-008-9182-9

A one-parametric class of merit functions for the second-order cone complementarity problem

Jein-Shan Chen · Shaohua Pan

Received: 11 January 2007 / Revised: 21 April 2008 / Published online: 8 May 2008

Abstract We investigate a one-parametric class of merit functions for the second-order cone complementarity problem (SOCCP) which is closely related to the popular Fischer–Burmeister (FB) merit function and natural residual merit function. In fact, it will reduce to the FB merit function if the involved parameter τ equals 2, whereas as τ tends to zero, its limit will become a multiple of the natural residual merit function. In this paper, we show that this class of merit functions enjoys several favorable properties as the FB merit function holds, for example, the smoothness. These properties play an important role in the reformulation method of an unconstrained minimization or a nonsmooth system of equations for the SOCCP. Numerical results are reported for some convex second-order cone programs (SOCPs) by solving the unconstrained minimization reformulation of the KKT optimality conditions, which indicate that the FB merit function is not the best. For the sparse linear SOCPs, the merit function corresponding to $\tau = 2.5$ or 3 works better than the FB merit function, whereas for the dense convex SOCPs, the merit function with $\tau = 0.1$, 0.5 or 1.0 seems to have better numerical performance.

Keywords Second-order cone · Complementarity · Merit function · Jordan product

J.-S. Chen (✉)
Department of Mathematics, National Taiwan Normal University, 88 Section 4, Ting-Chou Road, 11677 Taipei, Taiwan
e-mail: jschen@math.ntnu.edu.tw

J.-S. Chen
Mathematics Division, National Center for Theoretical Sciences, Taipei Office, Taipei, Taiwain

S. Pan
School of Mathematical Sciences, South China University of Technology, Guangzhou 510640, China
e-mail: shhpan@scut.edu.cn

1 Introduction

We consider the conic complementarity problem of finding a vector $\zeta \in \mathbb{R}^n$ such that

$$F(\zeta) \in \mathcal{K}, \qquad G(\zeta) \in \mathcal{K}, \qquad \langle F(\zeta), G(\zeta)\rangle = 0, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, $F : \mathbb{R}^n \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathbb{R}^n$ are the mappings assumed to be continuously differentiable throughout this paper, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOCs). In other words,

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

where $N, n_1, \ldots, n_N \geq 1$, $n_1 + \cdots + n_N = n$, and

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

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of nonnegative reals $\mathbb{R}_+$. We will refer to (1)–(2) as the *second-order cone complementarity problem* (SOCCP).

An important special case of the SOCCP corresponds to $G(\zeta) = \zeta$ for all $\zeta \in \mathbb{R}^n$. Then (1) reduces to

$$F(\zeta) \in \mathcal{K}, \quad \zeta \in \mathcal{K}, \qquad \langle F(\zeta), \zeta\rangle = 0, \tag{3}$$

which is a natural extension of the nonlinear complementarity problem (NCP) [7, 8] with $\mathcal{K} = \mathbb{R}^n_+$, the nonnegative orthant cone of $\mathbb{R}^n$. Another important special case corresponds to the KKT optimality conditions of the convex second-order cone program (CSOCP):

$$\begin{array}{ll} \text{minimize } g(x) & \\ \text{subject to} \quad Ax = b, & x \in \mathcal{K}, \end{array} \tag{4}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a convex twice continuously differentiable function, $A \in \mathbb{R}^{m\times n}$ has full row rank and $b \in \mathbb{R}^m$. From [4], we know that the KKT conditions of (4), which are sufficient but not necessary for optimality, can be reformulated as (1) with

$$F(\zeta) := \bar{x} + \big(I - A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A\big)\zeta, \qquad G(\zeta) := \nabla g(F(\zeta)) - A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A\zeta, \tag{5}$$

where $\bar{x} \in \mathbb{R}^n$ is any point such that $A\bar{x} = b$. When g is linear, the CSOCP reduces to the linear SOCP which arises in numerous applications in engineering design, finance, robust optimization, and includes as special cases convex quadratically constrained quadratic programs and linear programs; see [1, 13] and references therein.

There have been various methods proposed for solving SOCPs and SOCCPs. They include the interior-point methods [2, 3, 15, 16, 19], the non-interior smoothing Newton methods [6, 9], and the smoothing-regularization method [11]. Recently, there was an alternative method [4] based on reformulating the SOCCP as an unconstrained minimization problem. In that approach, it aims to find a function $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ satisfying

$$\psi(x, y) = 0 \quad \Longleftrightarrow \quad x \in \mathcal{K},\ y \in \mathcal{K}, \quad \langle x, y\rangle = 0, \tag{6}$$

so that the SOCCP can be reformulated as an unconstrained minimization problem

$$\min_{\zeta \in \mathbb{R}^n} f(\zeta) := \psi(F(\zeta), G(\zeta)).$$

We call such ψ a *merit function* associated with the cone $\mathcal{K}$.

A popular choice of ψ is the Fischer–Burmeister (FB) merit function

$$\psi_{\text{FB}}(x, y) := \frac{1}{2}\|\phi_{\text{FB}}(x, y)\|^2, \tag{7}$$

where $\phi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the vector-valued FB function defined by

$$\phi_{\text{FB}}(x, y) := (x^2 + y^2)^{1/2} - (x + y), \tag{8}$$

with $x^2 = x \circ x$ denoting the Jordan product between x and itself, $x^{1/2}$ being a vector such that $(x^{1/2})^2 = x$, and $x + y$ meaning the usual componentwise addition of vectors. The function ψ_{FB} was studied in [4] and particularly shown to be continuously differentiable (smooth). Another popular choice of ψ is the natural residual merit function

$$\psi_{\text{NR}}(x, y) := \frac{1}{2}\|\phi_{\text{NR}}(x, y)\|^2,$$

where $\phi_{\text{NR}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the vector-valued natural residual function given by

$$\phi_{\text{NR}}(x, y) := x - (x - y)_+$$

with $(\cdot)_+$ meaning the projection in the Euclidean norm onto $\mathcal{K}$. The function ϕ_{NR} was studied in [9, 11] which is involved in smoothing methods for the SOCCP. Compared with the FB merit function ψ_{FB}, the function ψ_{NR} has a drawback, i.e., its non-differentiability.

In this paper, we will investigate the following one-parametric class of functions

$$\psi_\tau(x, y) := \frac{1}{2}\|\phi_\tau(x, y)\|^2, \tag{9}$$

where τ is a fixed parameter from $(0, 4)$ and $\phi_\tau : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is defined by

$$\phi_\tau(x, y) := \left[(x - y)^2 + \tau(x \circ y)\right]^{1/2} - (x + y). \tag{10}$$

Specifically, we prove that ψ_τ is a merit function associated with $\mathcal{K}$ which is continuously differentiable everywhere with computable gradient formulas (see Propositions 3.1–3.3), and hence the SOCCP can be reformulated as an unconstrained smooth minimization

$$\min_{\zeta \in \mathbb{R}^n} f_\tau(\zeta) := \psi_\tau(F(\zeta), G(\zeta)). \tag{11}$$

Also, we show that every stationary point of f_τ solves the SOCCP under the condition that ∇F and $-\nabla G$ are column monotone (see Proposition 4.1). Observe that ϕ_τ reduces to ϕ_{FB} when $\tau = 2$, whereas its limit as $\tau \to 0$ becomes a multiple of ϕ_{NR}.

Thus, this class of merit functions has a close relation to two of the most important merit functions so that a closer look and study for it is worthwhile. In addition, this study is motivated by the work [12] where ϕ_τ was used to develop a nonsmooth Newton method for the NCP. This paper is mainly concerned with the merit function approach based on the unconstrained minimization problem (11). Numerical results are also reported for some convex SOCPs, which indicate that ψ_τ can be an alternative for $\psi_{\rm FB}$ if a suitable τ is selected.

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors, and $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is identified with $\mathbb{R}^{n_1+\cdots+n_m}$. Thus, $(x_1, \ldots, x_m) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is viewed as a column vector in $\mathbb{R}^{n_1+\cdots+n_m}$. The notation I denotes an identity matrix of suitable dimension, and $\mathrm{int}(\mathcal{K}^n)$ denotes the interior of $\mathcal{K}^n$. For any differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) \in \mathbb{R}^{n\times m}$ denotes the transposed Jacobian of F at x. For a symmetric matrix A, we write $A \succeq O$ (respectively, $A \succ O$) to mean A is positive semidefinite (respectively, positive definite). For nonnegative α and β, we write $\alpha = O(\beta)$ to mean $\alpha \le C\beta$, with $C > 0$ independent of α and β. Without loss of generality, in the rest of this paper we assume that $\mathcal{K} = \mathcal{K}^n$ $(n > 1)$. All analysis can be carried over to the general case where $\mathcal{K}$ has the structure as (2). In addition, we always assume that τ satisfies $0 < \tau < 4$.

2 Preliminaries

It is known that $\mathcal{K}^n$ $(n > 1)$ is a closed convex self-dual cone with nonempty interior

$$\mathrm{int}(\mathcal{K}^n) := \left\{x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_1 > \|x_2\|\right\}.$$

For any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the *Jordan product* of x and y is defined by

$$x \circ y := (\langle x, y\rangle,\ y_1 x_2 + x_1 y_2). \tag{12}$$

The Jordan product, unlike scalar or matrix multiplication, is not associative, which is a main source on complication in the analysis of SOCCP. The identity element under this product is $e := (1, 0, \ldots, 0)^{\rm T} \in \mathbb{R}^n$. For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the *determinant* of x is defined by $\det(x) := x_1^2 - \|x_2\|^2$. If $\det(x) \neq 0$, then x is said to be *invertible*. If x is invertible, there exists a unique $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ satisfying $x \circ y = y \circ x = e$. We call this y the inverse of x and denote it by x^{-1}. For each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let

$$L_x := \begin{bmatrix} x_1 & x_2^{\rm T} \\ x_2 & x_1 I \end{bmatrix}. \tag{13}$$

It is easily verified that $L_x y = x \circ y$ and $L_{x+y} = L_x + L_y$ for any $x, y \in \mathbb{R}^n$, but generally $L_x^2 = L_x L_x \neq L_{x^2}$ and $L_x^{-1} \neq L_{x^{-1}}$. If L_x is invertible, then the inverse of L_x is given by

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^{\rm T} \\ -x_2 & \dfrac{\det(x)}{x_1} I + \dfrac{1}{x_1} x_2 x_2^{\rm T} \end{bmatrix}. \tag{14}$$

We next recall from [9] that each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ admits a spectral factorization, associated with $\mathcal{K}^n$, of the form

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

where $\lambda_1(x), \lambda_2(x)$ and $u_x^{(1)}, u_x^{(2)}$ are the spectral values and the associated spectral vectors of x given by

$$\lambda_i(x) = x_1 + (-1)^i \|x_2\|, \qquad u_x^{(i)} = \frac{1}{2}(1,\ (-1)^i \bar{x}_2) \quad \text{for } i = 1, 2,$$

with $\bar{x}_2 = \frac{x_2}{\|x_2\|}$ if $x_2 \neq 0$, and otherwise $\bar{x}_2$ being any vector in $\mathbb{R}^{n-1}$ such that $\|\bar{x}_2\| = 1$. If $x_2 \neq 0$, the factorization is unique. The spectral factorization of x has various interesting properties; see [9]. We list three properties that will be used later.

Property 2.1

(a) $x^2 = \lambda_1^2(x) \cdot u_x^{(1)} + \lambda_2^2(x) \cdot u_x^{(2)} \in \mathcal{K}^n$ for any $x \in \mathbb{R}^n$.
(b) If $x \in \mathcal{K}^n$, then $x^{1/2} = \sqrt{\lambda_1(x)} \cdot u_x^{(1)} + \sqrt{\lambda_2(x)} \cdot u_x^{(2)} \in \mathcal{K}^n$.
(c) $x \in \mathcal{K}^n \Longleftrightarrow \lambda_1(x) \geq 0 \Longleftrightarrow L_x \succeq O$, $x \in \mathrm{int}(\mathcal{K}^n) \Longleftrightarrow \lambda_1(x) > 0 \Longleftrightarrow L_x \succ O$.

3 Smoothness of the function ψ_τ

In this section we will show that ψ_τ defined by (9) is a smooth merit function. First, by Properties 2.1(a) and (b), ϕ_τ and ψ_τ are well-defined since for any $x, y \in \mathbb{R}^n$, we can verify that

$$\begin{aligned}(x-y)^2 + \tau(x \circ y) &= \left(x + \frac{\tau-2}{2}y\right)^2 + \frac{\tau(4-\tau)}{4}y^2 \\ &= \left(y + \frac{\tau-2}{2}x\right)^2 + \frac{\tau(4-\tau)}{4}x^2 \in \mathcal{K}^n. \end{aligned} \tag{15}$$

The following proposition shows that ψ_τ is indeed a merit function associated with $\mathcal{K}^n$.

Proposition 3.1 *Let ψ_τ and ϕ_τ be given as in* (9) *and* (10), *respectively. Then,*

$$\psi_\tau(x, y) = 0 \iff \phi_\tau(x, y) = 0 \iff x \in \mathcal{K}^n,\ y \in \mathcal{K}^n, \quad \langle x, y \rangle = 0.$$

Proof The first equivalence is clear by the definition of ψ_τ. We consider the second one.

"$\Leftarrow$". Since $x \in \mathcal{K}$, $y \in \mathcal{K}$ and $\langle x, y \rangle = 0$, we have $x \circ y = 0$. Substituting it into the expression of $\phi_\tau(x, y)$ then yields that $\phi_\tau(x, y) = (x^2 + y^2)^{1/2} - (x + y) = \phi_{\mathrm{FB}}(x, y)$. From Proposition 2.1 of [9], we immediately obtain $\phi_\tau(x, y) = 0$.

"$\Rightarrow$". Suppose that $\phi_\tau(x, y) = 0$. Then, $x + y = \left[(x-y)^2 + \tau(x \circ y)\right]^{1/2}$. Squaring both sides yields $x \circ y = 0$. This implies that $x + y = (x^2 + y^2)^{1/2}$, i.e.

$\phi_{\text{FB}}(x, y) = 0$. From Proposition 2.1 of [9], it then follows that $x \in \mathcal{K}^n$, $y \in \mathcal{K}^n$ and $\langle x, y \rangle = 0$. □

In what follows, we focus on the proof of the smoothness of ψ_τ. We first introduce some notation that will be used in the sequel. For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let

$$\begin{aligned} w &= (w_1, w_2) = w(x, y) := (x - y)^2 + \tau(x \circ y), \\ z &= (z_1, z_2) = z(x, y) := \left[(x - y)^2 + \tau(x \circ y)\right]^{1/2}. \end{aligned} \tag{16}$$

Then, $w \in \mathcal{K}^n$ and $z \in \mathcal{K}^n$. Moreover, by the definition of Jordan product,

$$\begin{aligned} w_1 &= w_1(x, y) = \|x\|^2 + \|y\|^2 + (\tau - 2)x^{\mathrm{T}} y, \\ w_2 &= w_2(x, y) = 2(x_1 x_2 + y_1 y_2) + (\tau - 2)(x_1 y_2 + y_1 x_2). \end{aligned} \tag{17}$$

Let $\lambda_1(w)$ and $\lambda_2(w)$ be the spectral values of w. By Property 2.1(b), we have that

$$\begin{aligned} z_1 &= z_1(x, y) = \frac{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}}{2}, \\ z_2 &= z_2(x, y) = \frac{\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}}{2} \bar{w}_2, \end{aligned} \tag{18}$$

where $\bar{w}_2 := \frac{w_2}{\|w_2\|}$ if $w_2 \neq 0$ and otherwise $\bar{w}_2$ is any vector in $\mathbb{R}^{n-1}$ satisfying $\|\bar{w}_2\| = 1$.

The following technical lemma describes the behavior of x, y when $w = (x - y)^2 + \tau(x \circ y)$ is on the boundary of $\mathcal{K}^n$. In fact, it may be viewed as an extension of [4, Lemma 3.2].

Lemma 3.1 *For any* $x = (x_1, x_2),\ y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, *if* $w \notin \operatorname{int}(\mathcal{K}^n)$, *then*

$$x_1^2 = \|x_2\|^2, \qquad y_1^2 = \|y_2\|^2, \qquad x_1 y_1 = x_2^{\mathrm{T}} y_2, \qquad x_1 y_2 = y_1 x_2; \tag{19}$$

$$\begin{aligned} x_1^2 + y_1^2 + (\tau - 2)x_1 y_1 &= \|x_1 x_2 + y_1 y_2 + (\tau - 2)x_1 y_2\| \\ &= \|x_2\|^2 + \|y_2\|^2 + (\tau - 2)x_2^{\mathrm{T}} y_2. \end{aligned} \tag{20}$$

If, in addition, $(x, y) \neq (0, 0)$, *then* $w_2 \neq 0$, *and furthermore,*

$$x_2^{\mathrm{T}} \frac{w_2}{\|w_2\|} = x_1, \qquad x_1 \frac{w_2}{\|w_2\|} = x_2, \qquad y_2^{\mathrm{T}} \frac{w_2}{\|w_2\|} = y_1, \qquad y_1 \frac{w_2}{\|w_2\|} = y_2. \tag{21}$$

Proof Since $w = (x - y)^2 + \tau(x \circ y) \notin \operatorname{int}(\mathcal{K}^n)$, using (15) and [4, Lemma 3.2] yields

$$\left(x_1 + \frac{\tau - 2}{2} y_1\right)^2 = \left\|x_2 + \frac{\tau - 2}{2} y_2\right\|^2, \qquad y_1^2 = \|y_2\|^2,$$

$$\left(x_1 + \frac{\tau - 2}{2} y_1\right) y_2 = \left(x_2 + \frac{\tau - 2}{2} y_2\right) y_1,$$

$$\left(x_1 + \frac{\tau-2}{2}y_1\right)y_1 = \left(x_2 + \frac{\tau-2}{2}y_2\right)^{\mathrm{T}} y_2;$$

$$\left(y_1 + \frac{\tau-2}{2}x_1\right)^2 = \left\|y_2 + \frac{\tau-2}{2}x_2\right\|^2, \qquad x_1^2 = \|x_2\|^2,$$

$$\left(y_1 + \frac{\tau-2}{2}x_1\right)x_2 = \left(y_2 + \frac{\tau-2}{2}x_2\right)x_1,$$

$$\left(y_1 + \frac{\tau-2}{2}x_1\right)x_1 = \left(y_2 + \frac{\tau-2}{2}x_2\right)^{\mathrm{T}} x_2.$$

From these equalities, we readily get the results in (19). Since $w \in \mathcal{K}^n$ but $w \notin \mathrm{int}(\mathcal{K}^n)$, we have $\|x\|^2 + \|y\|^2 + (\tau-2)x^{\mathrm{T}}y = \|2x_1x_2 + 2y_1y_2 + (\tau-2)(x_1y_2 + y_1x_2)\|$ by $\lambda_1(w) = 0$. Applying the relations in (19) then gives the equalities in (20). If, in addition, $(x, y) \neq (0, 0)$, then it is clear that $\|x_1x_2 + y_1y_2 + (\tau-2)x_1y_2\| = x_1^2 + y_1^2 + (\tau-2)x_1y_1 \neq 0$. To prove the equalities in (21), it suffices to verify that $x_2^{\mathrm{T}}\frac{w_2}{\|w_2\|} = x_1$ and $x_1\frac{w_2}{\|w_2\|} = x_2$ by the symmetry of x and y in w. The verifications are straightforward by (20) and $x_1y_2 = y_1x_2$. □

By Lemma 3.1, when $w \notin \mathrm{int}(\mathcal{K}^n)$, the spectral values of w are calculated as follows:

$$\lambda_1(w) = 0, \qquad \lambda_2(w) = 4\big(x_1^2 + y_1^2 + (\tau-2)x_1y_1\big). \tag{22}$$

If $(x, y) \neq (0, 0)$ also holds, then using (18), (20) and (22) yields that

$$z_1(x, y) = \sqrt{x_1^2 + y_1^2 + (\tau-2)x_1y_1}, \qquad z_2(x, y) = \frac{x_1x_2 + y_1y_2 + (\tau-2)x_1y_2}{\sqrt{x_1^2 + y_1^2 + (\tau-2)x_1y_1}}.$$

Thus, if $(x, y) \neq (0, 0)$ and $(x-y)^2 + \tau(x \circ y) \notin \mathrm{int}(\mathcal{K}^n)$, $\phi_\tau(x, y)$ can be rewritten as

$$\phi_\tau(x, y) = z(x, y) - (x+y) = \begin{pmatrix} \sqrt{x_1^2 + y_1^2 + (\tau-2)x_1y_1} - (x_1+y_1) \\ \frac{x_1x_2 + y_1y_2 + (\tau-2)x_1y_2}{\sqrt{x_1^2 + y_1^2 + (\tau-2)x_1y_1}} - (x_2+y_2) \end{pmatrix}. \tag{23}$$

This specific expression will be employed in the proof of the following main result.

Proposition 3.2 *The function ψ_τ given by (9) is differentiable at every $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. Moreover, $\nabla_x\psi_\tau(0, 0) = \nabla_y\psi_\tau(0, 0) = 0$; if $(x-y)^2 + \tau(x \circ y) \in \mathrm{int}(\mathcal{K}^n)$, then*

$$\begin{aligned} \nabla_x\psi_\tau(x, y) &= \big[L_{x+\frac{\tau-2}{2}y}L_z^{-1} - I\big]\phi_\tau(x, y), \\ \nabla_y\psi_\tau(x, y) &= \big[L_{y+\frac{\tau-2}{2}x}L_z^{-1} - I\big]\phi_\tau(x, y); \end{aligned} \tag{24}$$

if $(x,y)\neq(0,0)$ *and* $(x-y)^2+\tau(x\circ y)\notin \mathrm{int}(\mathcal{K}^n)$, *then* $x_1^2+y_1^2+(\tau-2)x_1y_1\neq 0$ *and*

$$\nabla_x\psi_\tau(x,y)=\left[\frac{x_1+\frac{\tau-2}{2}y_1}{\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1}}-1\right]\phi_\tau(x,y),$$
$$\nabla_y\psi_\tau(x,y)=\left[\frac{y_1+\frac{\tau-2}{2}x_1}{\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1}}-1\right]\phi_\tau(x,y). \tag{25}$$

Proof Case (1): $(x,y)=(0,0)$. For any $u=(u_1,u_2), v=(v_1,v_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$, let μ_1,μ_2 be the spectral values of $(u-v)^2+\tau(u\circ v)$ and $\xi^{(1)},\xi^{(2)}$ be the spectral vectors. Then,

$$\begin{aligned}2\left[\psi_\tau(u,v)-\psi_\tau(0,0)\right]&=\left\|\left[u^2+v^2+(\tau-2)(u\circ v)\right]^{1/2}-u-v\right\|^2\\&=\left\|\sqrt{\mu_1}\,\xi^{(1)}+\sqrt{\mu_2}\,\xi^{(2)}-u-v\right\|^2\\&\le\left(\sqrt{2\mu_2}+\|u\|+\|v\|\right)^2.\end{aligned}$$

In addition, from the definition of spectral value, it follows that

$$\begin{aligned}\mu_2&=\|u\|^2+\|v\|^2+(\tau-2)u^{\mathrm{T}}v+2(u_1u_2+v_1v_2)+(\tau-2)(u_1v_2+v_1u_2)\\&\le 2\|u\|^2+2\|v\|^2+3|\tau-2|\|u\|\|v\|\le 5(\|u\|^2+\|v\|^2).\end{aligned}$$

Now combining the last two equations, we have $\psi_\tau(u,v)-\psi_\tau(0,0)=O(\|u\|^2+\|v\|^2)$. This shows that ψ_τ is differentiable at $(0,0)$ with $\nabla_x\psi_\tau(0,0)=\nabla_y\psi_\tau(0,0)=0$.

Case (2): $(x-y)^2+\tau(x\circ y)\in\mathrm{int}(\mathcal{K}^n)$. By [5, Proposition 5], $z(x,y)$ defined by (18) is continuously differentiable at such (x,y), and consequently $\phi_\tau(x,y)$ is also continuously differentiable at such (x,y) since $\phi_\tau(x,y)=z(x,y)-(x+y)$. Notice that

$$z^2(x,y)=\left(x+\frac{\tau-2}{2}y\right)^2+\frac{\tau(4-\tau)}{4}y^2,$$

which leads to $\nabla_x z(x,y)L_z=L_{x+\frac{\tau-2}{2}y}$ by taking differentiation on both sides about x. Since $L_z\succ O$ by Property 2.1 (c), it follows that $\nabla_x z(x,y)=L_{x+\frac{\tau-2}{2}y}L_z^{-1}$. Consequently,

$$\nabla_x\phi_\tau(x,y)=\nabla_x z(x,y)-I=L_{x+\frac{\tau-2}{2}y}L_z^{-1}-I.$$

This together with $\nabla_x\psi_\tau(x,y)=\nabla_x\phi_\tau(x,y)\phi_\tau(x,y)$ proves the first formula of (24). For the symmetry of x and y in ψ_τ, the second formula also holds.

Case (3): $(x,y)\neq(0,0)$ and $(x-y)^2+\tau(x\circ y)\notin\mathrm{int}(\mathcal{K}^n)$. For any $x'=(x_1',x_2'),\ y'=(y_1',y_2')\in\mathbb{R}\times\mathbb{R}^{n-1}$, it is easy to verify that

$$2\psi_\tau(x',y')=\left\|\left[x'^2+y'^2+(\tau-2)(x'\circ y')\right]^{1/2}\right\|^2+\|x'+y'\|^2$$

$$\begin{aligned}
&\quad - 2\big\langle\big[x'^2 + y'^2 + (\tau - 2)(x' \circ y')\big]^{1/2}, x' + y'\big\rangle \\
&= \|x'\|^2 + \|y'\|^2 + (\tau - 2)\langle x', y'\rangle + \|x' + y'\|^2 \\
&\quad - 2\big\langle\big[x'^2 + y'^2 + (\tau - 2)(x' \circ y')\big]^{1/2}, x' + y'\big\rangle,
\end{aligned}$$

where the second equality uses the fact that $\|z\|^2 = \langle z^2, e\rangle$ for any $z \in \mathbb{R}^n$. Since $\|x'\|^2 + \|y'\|^2 + (\tau - 2)\langle x', y'\rangle + \|x' + y'\|^2$ is clearly differentiable in (x', y'), it suffices to show that $\langle [x'^2 + y'^2 + (\tau - 2)(x' \circ y')]^{1/2}, x' + y'\rangle$ is differentiable at $(x', y') = (x, y)$. By Lemma 3.1, $w_2 = w_2(x, y) \neq 0$, which implies $w_2' = w_2(x', y') = 2x_1'x_2' + 2y_1'y_2' + (\tau - 2)(x_1'y_2' + y_1'x_2') \neq 0$ for all $(x', y') \in \mathbb{R}^n \times \mathbb{R}^n$ sufficiently near to (x, y). Let μ_1, μ_2 be the spectral values of $x'^2 + y'^2 + (\tau - 2)(x' \circ y')$. Then we can compute that

$$\begin{aligned}
&2\big\langle\big[x'^2 + y'^2 + (\tau - 2)(x' \circ y')\big]^{1/2},\ x' + y'\big\rangle \\
&\quad = \sqrt{\mu_2}\left[x_1' + y_1' + \frac{[2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')]^{\mathrm{T}}(x_2' + y_2')}{\|2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')\|}\right] \\
&\qquad + \sqrt{\mu_1}\left[x_1' + y_1' - \frac{[2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')]^{\mathrm{T}}(x_2' + y_2')}{\|2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')\|}\right].
\end{aligned} \tag{26}$$

Since $\lambda_2(w) > 0$ and $w_2(x, y) \neq 0$, the first term on the right-hand side of (26) is differentiable at $(x', y') = (x, y)$. Now, we claim that the second term is $o(\|x' - x\| + \|y' - y\|)$, i.e., it is differentiable at (x, y) with zero gradient. To see this, notice that $w_2(x, y) \neq 0$, and hence $\mu_1 = \|x'\|^2 + \|y'\|^2 + (\tau - 2)\langle x', y'\rangle - \|2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')\|$, viewed as a function of (x', y'), is differentiable at $(x', y') = (x, y)$. Moreover, $\mu_1 = \lambda_1(w) = 0$ when $(x', y') = (x, y)$. Thus, the first-order Taylor's expansion of μ_1 at (x, y) yields

$$\mu_1 \ = \ O(\|x' - x\| + \|y' - y\|).$$

Also, since $w_2(x, y) \neq 0$, by the product and quotient rules for differentiation, the function

$$x_1' + y_1' - \frac{[2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')]^{\mathrm{T}}(x_2' + y_2')}{\|2(x_1'x_2' + y_1'y_2') + (\tau - 2)(x_1'y_2' + y_1'x_2')\|} \tag{27}$$

is differentiable at $(x', y') = (x, y)$, and it has value 0 at $(x', y') = (x, y)$ due to

$$\begin{aligned}
&x_1 + y_1 - \frac{[x_1x_2 + y_1y_2 + (\tau - 2)x_1y_2]^{\mathrm{T}}(x_2 + y_2)}{\|x_1x_2 + y_1y_2 + (\tau - 2)x_1y_2\|} \\
&\quad = x_1 - x_2^{\mathrm{T}}\frac{w_2}{\|w_2\|} + y_1 - y_2^{\mathrm{T}}\frac{w_2}{\|w_2\|} = 0.
\end{aligned}$$

Hence, the function in (27) is $O(\|x' - x\| + \|y' - y\|)$ in magnitude, which together with $\mu_1 = O(\|x' - x\| + \|y' - y\|)$ shows that the second term on the right-hand side of (26) is

$$O\big((\|x'-x\|+\|y'-y\|)^{3/2}\big)=o\big(\|x'-x\|+\|y'-y\|\big).$$

Thus, we have shown that ψ_τ is differentiable at (x,y). Moreover, we see that $2\nabla\psi_\tau(x,y)$ is the sum of the gradient of $\|x'\|^2+\|y'\|^2+(\tau-2)\langle x',y'\rangle+\|x'+y'\|^2$ and the gradient of the first term on the right-hand side of (26), evaluated at $(x',y')=(x,y)$.

The gradient of $\|x'\|^2+\|y'\|^2+(\tau-2)\langle x',y'\rangle+\|x'+y'\|^2$ with respect to x', evaluated at $(x',y')=(x,y)$, is $2x+(\tau-2)y+2(x+y)$. The derivative of the first term on the right-hand side of (26) with respect to x_1', evaluated at $(x',y')=(x,y)$, works out to be

$$\begin{aligned}
&\frac{1}{\sqrt{\lambda_2(w)}}\left[\left(x_1+\frac{\tau-2}{2}y_1\right)+\left(x_2+\frac{\tau-2}{2}y_2\right)^{\mathrm T}\frac{w_2}{\|w_2\|}\right]\\
&\quad\times\left(x_1+y_1+(x_2+y_2)^{\mathrm T}\frac{w_2}{\|w_2\|}\right)\\
&\quad+\sqrt{\lambda_2(w)}\left[1+\frac{(x_2+\frac{\tau-2}{2}y_2)^{\mathrm T}(x_2+y_2)}{\|x_1x_2+y_1y_2+(\tau-2)x_1y_2\|}\right.\\
&\quad\left.-\frac{w_2^{\mathrm T}(x_2+y_2)\cdot w_2^{\mathrm T}(x_2+\frac{\tau-2}{2}y_2)}{\|x_1x_2+y_1y_2+(\tau-2)x_1y_2\|\cdot\|w_2\|^2}\right]\\
&=\frac{2(x_1+\frac{\tau-2}{2}y_1)(x_1+y_1)}{\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1}}+2\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1},
\end{aligned}$$

where the equality follows from Lemma 3.1. Similarly, the gradient of the first term on the right of (26) with respect to x_2', evaluated at $(x',y')=(x,y)$, works out to be

$$\begin{aligned}
&\frac{1}{\sqrt{\lambda_2(w)}}\left[\left(x_2+\frac{\tau-2}{2}y_2\right)+\left(x_1+\frac{\tau-2}{2}y_1\right)\frac{w_2}{\|w_2\|}\right]\\
&\quad\times\left(x_1+y_1+(x_2+y_2)^{\mathrm T}\frac{w_2}{\|w_2\|}\right)\\
&\quad+\sqrt{\lambda_2(w)}\left[\frac{(2x_1+(\tau-2)y_1)x_2+\frac{\tau}{2}(x_1+y_1)y_2}{\|x_1x_2+y_1y_2+(\tau-2)x_1y_2\|}\right.\\
&\quad\left.\cdot-\frac{w_2^{\mathrm T}(x_2+y_2)\cdot(x_1+\frac{\tau-2}{2}y_1)w_2}{\|x_1x_2+y_1y_2+(\tau-2)x_1y_2\|\cdot\|w_2\|^2}\right]\\
&=2\frac{(2x_1+(\tau-2)y_1)x_2+\frac{\tau}{2}(x_1+y_1)y_2}{\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1}}.
\end{aligned}$$

Then, combining the last two gradient expressions yields that

$$
\begin{aligned}
&2\nabla_x\psi_\tau(x,y)\\
&\quad = 2x + (\tau-2)y + 2(x+y) - \begin{bmatrix} 2\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1} \\ 0 \end{bmatrix}\\
&\qquad - \frac{2}{\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1}}\begin{bmatrix} (x_1+\frac{\tau-2}{2}y_1)(x_1+y_1) \\ (2x_1+(\tau-2)y_1)x_2+\frac{\tau}{2}(x_1+y_1)y_2 \end{bmatrix}.
\end{aligned}
$$

Using the fact that $x_1y_2 = y_1x_2$ and noting that ϕ_τ can be simplified as the one in (23) under this case, we readily rewrite the above expression for $\nabla_x\psi_\tau(x,y)$ in the form of (25). By symmetry, $\nabla_y\psi_\tau(x,y)$ also holds as the form of (25). □

Proposition 3.2 shows that ψ_τ is differentiable with a computable gradient. To establish the continuity of the gradient of ψ_τ or the smoothness of ψ_τ, we need the following two crucial technical lemmas whose proofs are provided in the Appendix.

Lemma 3.2 *For any* $x=(x_1,x_2)$, $y=(y_1,y_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$, *if* $w_2\neq 0$, *then*

$$
\begin{aligned}
&\left[\left(x_1+\frac{\tau-2}{2}y_1\right)+(-1)^i\left(x_2+\frac{\tau-2}{2}y_2\right)^{\mathrm T}\frac{w_2}{\|w_2\|}\right]^2\\
&\quad\le\left\|\left(x_2+\frac{\tau-2}{2}y_2\right)+(-1)^i\left(x_1+\frac{\tau-2}{2}y_1\right)\frac{w_2}{\|w_2\|}\right\|^2\le\lambda_i(w)
\end{aligned}
$$

for $i=1,2$. *Furthermore, these relations also hold when interchanging* x *and* y.

Lemma 3.3 *For all* (x,y) *satisfying* $(x-y)^2+\tau(x\circ y)\in\operatorname{int}(\mathcal{K}^n)$, *we have that*

$$
\|L_{x+\frac{\tau-2}{2}y}L_z^{-1}\|_F\le C,\qquad \|L_{y+\frac{\tau-2}{2}x}L_z^{-1}\|_F\le C,\tag{28}
$$

where $C>0$ *is a constant independent of* x, y *and* τ, *and* $\|\cdot\|_F$ *denotes the Frobenius norm.*

Proposition 3.3 *The function* ψ_τ *defined by* (9) *is smooth everywhere on* $\mathbb{R}^n\times\mathbb{R}^n$.

Proof By Proposition 3.2 and the symmetry of x and y in $\nabla\psi_\tau$, it suffices to show that $\nabla_x\psi_\tau$ is continuous at every $(a,b)\in\mathbb{R}^n\times\mathbb{R}^n$. If $(a-b)^2+\tau(a\circ b)\in\operatorname{int}(\mathcal{K}^n)$, the conclusion has been shown in Proposition 3.2. We next consider the other two cases.

Case (1): $(a,b)=(0,0)$. By Proposition 3.2, we need to show that $\nabla_x\psi_\tau(x,y)\to 0$ as $(x,y)\to(0,0)$. If $(x-y)^2+\tau(x\circ y)\in\operatorname{int}(\mathcal{K}^n)$, then $\nabla_x\psi_\tau(x,y)$ is given by (24), whereas if $(x,y)\neq(0,0)$ and $(x-y)^2+\tau(x\circ y)\notin\operatorname{int}(\mathcal{K}^n)$, then $\nabla_x\psi_\tau(x,y)$ is given by (25). Notice that

$$
L_{x+\frac{\tau-2}{2}y}L_z^{-1}\quad\text{and}\quad\frac{x_1+\frac{\tau-2}{2}y_1}{\sqrt{x_1^2+y_1^2+(\tau-2)x_1y_1}}
$$

are bounded with bound independent of x, y and τ, using the continuity of $\phi_\tau(x, y)$ immediately yields the desired result.

Case (2): $(a, b) \neq (0, 0)$ and $(a - b)^2 + \tau(a \circ b) \notin \mathrm{int}(\mathcal{K}^n)$. We will show that $\nabla_x \psi_\tau(x, y) \to \nabla_x \psi_\tau(a, b)$ by the two subcases: (2a) $(x, y) \neq (0, 0)$ and $(x - y)^2 + \tau(x \circ y) \notin \mathrm{int}(\mathcal{K}^n)$ and (2b) $(x-y)^2+\tau(x \circ y) \in \mathrm{int}(\mathcal{K}^n)$. In subcase (2a), $\nabla_x \psi_\tau(x, y)$ is given by (25). Noting that the right-hand side of (25) is continuous at (a, b), the desired result follows.

Next, we prove that $\nabla_x \psi_\tau(x, y) \to \nabla_x \psi_\tau(a, b)$ in subcase (2b). From (24), we have that

$$\nabla_x \psi_\tau(x, y) = \left(x + \frac{\tau - 2}{2} y\right) - L_{x+\frac{\tau-2}{2}y} L_z^{-1}(x + y) - \phi_\tau(x, y). \tag{29}$$

On the other hand, since $(a, b) \neq (0, 0)$ and $(a - b)^2 + \tau(a \circ b) \notin \mathrm{int}(\mathcal{K}^n)$,

$$\|a\|^2 + \|b\|^2 + (\tau - 2)a^{\mathrm{T}} b = \|2(a_1 a_2 + b_1 b_2) + (\tau - 2)(a_1 b_2 + b_1 a_2)\| \neq 0, \tag{30}$$

and moreover from (20) it follows that

$$\begin{aligned}\|a\|^2 + \|b\|^2 + (\tau - 2)a^{\mathrm{T}} b &= 2\big(a_1^2 + b_1^2 + (\tau - 2)a_1 b_1\big)\\ &= 2\big(\|a_2\|^2 + \|b_2\|^2 + (\tau - 2)a_2^{\mathrm{T}} b_2\big)\\ &= 2\big\|(a_1 a_2 + b_1 b_2) + (\tau - 2)a_1 b_2\big\|.\end{aligned} \tag{31}$$

Using the equalities in (31), it is not hard to verify that

$$\frac{a_1 + \frac{\tau-2}{2} b_1}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}} \big((a - b)^2 + \tau(a \circ b)\big)^{1/2} = a + \frac{\tau - 2}{2} b.$$

This together with the expression of $\nabla_x \psi_\tau(a, b)$ given by (25) yields

$$\nabla_x \psi_\tau(a, b) = \left(a + \frac{\tau - 2}{2} b\right) - \frac{a_1 + \frac{\tau-2}{2} b_1}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}} (a + b) - \phi_\tau(a, b). \tag{32}$$

Comparing (29) with (32), we see that if we wish to prove $\nabla_x \psi_\tau(x, y) \to \nabla_x \psi_\tau(a, b)$ as $(x, y) \to (a, b)$, it suffices to show that

$$L_{x+\frac{\tau-2}{2}y} L_z^{-1}(x + y) \to \frac{a_1 + \frac{\tau-2}{2} b_1}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}} (a + b), \tag{33}$$

which is also equivalent to proving the following three relations

$$L_{x+\frac{\tau-2}{2}y} L_z^{-1} \left(x + \frac{\tau - 2}{2} y\right) \to \frac{a_1 + \frac{\tau-2}{2} b_1}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}} \left(a + \frac{\tau - 2}{2} b\right), \tag{34}$$

$$L_{y+\frac{\tau-2}{2}x}L_z^{-1}\left(y+\frac{\tau-2}{2}x\right) \to \frac{b_1+\frac{\tau-2}{2}a_1}{\sqrt{a_1^2+b_1^2+(\tau-2)a_1b_1}}\left(b+\frac{\tau-2}{2}a\right), \tag{35}$$

$$\frac{4-\tau}{2}L_{x-y}L_z^{-1}\left(y+\frac{\tau-2}{2}x\right) \to \frac{\frac{4-\tau}{2}(a_1-b_1)}{\sqrt{a_1^2+b_1^2+(\tau-2)a_1b_1}}\left(b+\frac{\tau-2}{2}a\right). \tag{36}$$

By the symmetry of x and y in (34) and (35), we only prove (34) and (36). Let

$$\begin{aligned}(\zeta_1,\zeta_2) &:= L_{x+\frac{\tau-2}{2}y}L_z^{-1}\left(x+\frac{\tau-2}{2}y\right),\\ (\xi_1,\xi_2) &:= L_{x-y}L_z^{-1}\left(y+\frac{\tau-2}{2}x\right).\end{aligned} \tag{37}$$

Then showing (34) and (36) reduces to proving the following relations hold as $(x,y)\to(a,b)$:

$$\begin{aligned}\zeta_1 &\to \frac{(a_1+\frac{\tau-2}{2}b_1)^2}{\sqrt{a_1^2+b_1^2+(\tau-2)a_1b_1}},\\ \zeta_2 &\to \frac{a_1+\frac{\tau-2}{2}b_1}{\sqrt{a_1^2+b_1^2+(\tau-2)a_1b_1}}\left(a_2+\frac{\tau-2}{2}b_2\right),\end{aligned} \tag{38}$$

$$\begin{aligned}\xi_1 &\to \frac{(a_1-b_1)(b_1+\frac{\tau-2}{2}a_1)}{\sqrt{a_1^2+b_1^2+(\tau-2)a_1b_1}},\\ \xi_2 &\to \frac{(a_1-b_1)}{\sqrt{a_1^2+b_1^2+(\tau-2)a_1b_1}}\left(b_2+\frac{\tau-2}{2}a_2\right).\end{aligned} \tag{39}$$

To verify (38), we take (x,y) sufficiently near to (a,b). By (30), we may assume that $w_2=w_2(x,y)\neq 0$. Let $s=(s_1,s_2)=x+\frac{\tau-2}{2}y$. Using (14) and (37), we can calculate that

$$\begin{aligned}\zeta_1 &= \frac{1}{\det(z)}\left(s_1^2z_1-2s_1s_2^{\mathrm T}z_2+\frac{\det(z)}{z_1}\|s_2\|^2+\frac{(z_2^{\mathrm T}s_2)^2}{z_1}\right),\\ &= \frac{\|s_2\|^2}{z_1}+\frac{(s_1z_1-s_2^{\mathrm T}z_2)^2}{z_1\det(z)},\end{aligned} \tag{40}$$

$$\begin{aligned}\zeta_2 &= \frac{1}{\det(z)}\left(s_1z_1s_2-z_2^{\mathrm T}s_2s_2-s_1^2z_2+\frac{s_1\det(z)}{z_1}s_2+\frac{s_1}{z_1}z_2^{\mathrm T}s_2z_2\right)\\ &= \frac{s_1}{z_1}s_2+\frac{(s_1z_1-s_2^{\mathrm T}z_2)}{\det(z)}\left(s_2-\frac{s_1}{z_1}z_2\right).\end{aligned} \tag{41}$$

Notice that, as $(x, y) \to (a, b)$, from equations (17) and (30)–(31) it follows that

$$\lambda_1(w) \to 0 \quad \text{and} \quad \lambda_2(w) \to 4\big(a_1^2 + b_1^2 + (\tau - 2)a_1 b_1\big). \tag{42}$$

In addition, by the proof of Lemma 3.1, we also have

$$\left(a_1 + \frac{\tau - 2}{2} b_1\right)^2 = \left\| a_2 + \frac{\tau - 2}{2} b_2 \right\|^2 \quad \text{and} \quad b_1^2 = \|b_2\|^2.$$

Thus, from the last two equations and the expression of z given by (18), we have

$$\frac{\|s_2\|^2}{z_1} = \frac{2\|s_2\|^2}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}} \to \frac{(a_1 + \frac{\tau-2}{2} b_1)^2}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}}. \tag{43}$$

On the other hand, for the second term in the right-hand side of (40), we can compute that

$$\begin{aligned} \frac{(s_1 z_1 - s_2^{\mathrm{T}} z_2)^2}{z_1 \det(z)} = \frac{1}{z_1 \sqrt{\lambda_2(w)}} \Bigg[& s_1^2 \sqrt{\lambda_1(w)} + s_1\big(\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}\big)\left(s_1 - \frac{s_2^{\mathrm{T}} w_2}{\|w_2\|}\right) \\ & + \frac{1}{4}\big(\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}\big)^2 \cdot \frac{1}{\sqrt{\lambda_1(w)}} \left(s_1 - \frac{s_2^{\mathrm{T}} w_2}{\|w_2\|}\right)^2 \Bigg]. \end{aligned} \tag{44}$$

Since $s_1^2\sqrt{\lambda_1(w)}$, $s_1 - \frac{s_2^{\mathrm{T}} w_2}{\|w_2\|} \to 0$ as $(x, y) \to (a, b)$ and $|s_1 - \frac{s_2^{\mathrm{T}} w_2}{\|w_2\|}| \le \sqrt{\lambda_1(w)}$ by Lemma 3.2, the right-hand side of (44) tends to 0 as $(x, y) \to (a, b)$. Combining with (43), we prove the first relation in (38). We next prove the second relation in (38). Note that ζ_2 is given by (41). From (42) and (21), it follows that, as $(x, y) \to (a, b)$,

$$\frac{s_1}{z_1} s_2 = \frac{2 s_1 s_2}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}} \to \frac{a_1 + \frac{\tau-2}{2} b_1}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}} \left(a_2 + \frac{\tau - 2}{2} b_2\right), \tag{45}$$

$$s_2 - \frac{s_1}{z_1} z_2 \to \left(a_2 + \frac{\tau - 2}{2} b_2\right) - \frac{(a_1^2 + b_1^2 + (\tau - 2)a_1 b_1)(a_2 + \frac{\tau-2}{2} b_2)}{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1} = 0. \tag{46}$$

In addition, by the expression of z, we can compute that

$$\frac{(s_1 z_1 - s_2^{\mathrm{T}} z_2)}{\det(z)} = \frac{s_1}{\sqrt{\lambda_2(w)}} + \frac{1 - \sqrt{\lambda_1(w)}/\sqrt{\lambda_2(w)}}{2\sqrt{\lambda_1(w)}} \left(s_1 - \frac{s_2^{\mathrm{T}} w_2}{\|w_2\|}\right). \tag{47}$$

By (42), the first term on the right-hand side of (47) tends to $\frac{a_1 + \frac{\tau-2}{2} b_1}{2\sqrt{a_1^2 + b_1^2 + (\tau-2)a_1 b_2}}$ as $(x, y) \to (a, b)$, while the second term is bounded since

$$|s_1 - \frac{s_2^{\mathrm{T}} w_2}{\|w_2\|}| \le \sqrt{\lambda_1(w)}$$

by Lemma 3.2. Combining with (45), (46) and (41) yields the second relation in (38). Hence, (34) holds.

Now, we focus on the proof of (40). Let $u = x - y$ and $v = y + \frac{\tau-2}{2}x$. From (14) and (37),

$$\xi_1 = \frac{1}{\det(z)}\left(u_1 z_1 v_1 - u_1 z_2^T v_2 - v_1 u_2^T z_2 + \frac{\det(z)}{z_1} u_2^T v_2 + \frac{u_2^T z_2 (z_2^T v_2)}{z_1}\right)$$
$$= \frac{u_2^T v_2}{z_1} + \frac{(u_1 z_1 - u_2^T z_2)(v_1 z_1 - v_2^T z_2)}{z_1 \det(z)}, \tag{48}$$
$$\xi_2 = \frac{1}{\det(z)}\left(z_1 v_1 u_2 - z_2^T v_2 u_2 - u_1 v_1 z_2 + \frac{u_1 \det(z)}{z_1} v_2 + \frac{u_1}{z_1} z_2^T v_2 z_2\right)$$
$$= \frac{u_1}{z_1} v_2 + \frac{(z_1 v_1 - z_2^T v_2)}{\det(z)}\left(u_2 - \frac{u_1}{z_1} z_2\right). \tag{49}$$

Since $(a-b)^2 + \tau(a \circ b) \notin \operatorname{int}(\mathcal{K}^n)$, we have $a_2^T b_2 = a_1 b_1$, $a_1^2 = \|a_2\|^2$ and $b_1^2 = \|b_2\|^2$ due to Lemma 3.1. This together with (42) implies that

$$\frac{u_2^T v_2}{z_1} = \frac{2(x_2 - y_2)^T v_2}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}} \to \frac{(a_1 - b_1)(b_1 + \frac{\tau-2}{2} a_1)}{\sqrt{a_1^2 + b_1^2 + (\tau-2) a_1 b_1}} \quad \text{as } (x, y) \to (a, b). \tag{50}$$

We next prove that the second term in the right-hand side of (48) tends to 0. By computing,

$$\frac{(u_1 z_1 - u_2^T z_2)(v_1 z_1 - v_2^T z_2)}{z_1 \det(z)}$$
$$= \frac{\sqrt{\lambda_1(w)\lambda_2(w)}}{z_1}\left[\frac{u_1}{\sqrt{\lambda_2(w)}} + \frac{\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}}{2\sqrt{\lambda_1(w)\lambda_2(w)}}\left(u_1 - \frac{u_2^T w_2}{\|w_2\|}\right)\right]$$
$$\times \left[\frac{v_1}{\sqrt{\lambda_2(w)}} + \frac{\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}}{2\sqrt{\lambda_1(w)\lambda_2(w)}}\left(v_1 - \frac{v_2^T w_2}{\|w_2\|}\right)\right].$$

When $(x, y) \to (a, b)$, we have $\frac{1}{z_1}\sqrt{\lambda_1(w)\lambda_2(w)} \to 0$. In addition, by Lemma 3.2,

$$\left|u_1 - \frac{u_2^T w_2}{\|w_2\|}\right| = \frac{2}{4-\tau}\left|\left[\left(x_1 + \frac{\tau-2}{2} y_1\right) - \left(x_2 + \frac{\tau-2}{2} y_2\right)^T \frac{w_2}{\|w_2\|}\right]\right.$$
$$\left. - \left[\left(y_1 + \frac{\tau-2}{2} x_1\right) - \left(y_2 + \frac{\tau-2}{2} x_2\right)^T \frac{w_2}{\|w_2\|}\right]\right| \le \frac{4\sqrt{\lambda_1(w)}}{4-\tau},$$
$$\left|v_1 - \frac{v_2^T w_2}{\|w_2\|}\right| = \left|\left(y_1 + \frac{\tau-2}{2} x_1\right) - \left(y_2 + \frac{\tau-2}{2} x_2\right)^T \frac{w_2}{\|w_2\|}\right| \le \sqrt{\lambda_1(w)}.$$

This means that

$$\frac{1}{\sqrt{\lambda_1(w)}}\left(u_1 - \frac{u_2^T w_2}{\|w_2\|}\right) \quad \text{and} \quad \frac{1}{\sqrt{\lambda_1(w)}}\left(v_1 - \frac{v_2^T w_2}{\|w_2\|}\right)$$

are uniformly bounded. Notice that

$$\frac{u_1}{\sqrt{\lambda_2(w)}}, \frac{v_1}{\sqrt{\lambda_2(w)}} \quad \text{and} \quad \frac{\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}}{\sqrt{\lambda_2(w)}}$$

are also uniformly bounded. Therefore,

$$\frac{(u_1 z_1 - u_2^{\mathrm{T}} z_2)(v_1 z_1 - v_2^{\mathrm{T}} z_2)}{z_1 \det(z)} \to 0 \quad \text{as } (x, y) \to (a, b).$$

Combining with (50), we prove the first relation in (40). It remains to show the second relation in (40). Note that ξ_2 is given by (49). When $(x, y) \to (a, b)$, from (42) and (21),

$$\frac{u_1}{z_1} v_2 \;\to\; \frac{(a_1 - b_1)}{\sqrt{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1}} \left(b_2 + \frac{\tau - 2}{2} a_2 \right), \tag{51}$$

$$u_2 - \frac{u_1}{z_1} z_2 \;\to\; (a_2 - b_2) - \frac{(a_1^2 + b_1^2 + (\tau - 2)a_1 b_1)\,(a_2 - b_2)}{a_1^2 + b_1^2 + (\tau - 2)a_1 b_1} = 0. \tag{52}$$

In addition, by the expression of z, we can compute that

$$\frac{(z_1 v_1 - z_2^{\mathrm{T}} v_2)}{\det(z)} = \frac{v_1}{\sqrt{\lambda_2(w)}} + \frac{1 - \sqrt{\lambda_1(w)}/\sqrt{\lambda_2(w)}}{2\sqrt{\lambda_1(w)}} \left(v_1 - \frac{v_2^{\mathrm{T}} w_2}{\|w_2\|} \right). \tag{53}$$

From (42), the first term on the right-hand side of (53) converges to $\frac{b_1 + \frac{\tau-2}{2} a_1}{2\sqrt{a_1^2 + b_1^2 + (\tau-2)a_1 b_2}}$ when $(x, y) \to (a, b)$, while the second term is bounded since $|v_1 - \frac{v_2^{\mathrm{T}} w_2}{\|w_2\|}| \le \sqrt{\lambda_1(w)}$ by Lemma 3.2. Combining with (51), (52) and (49), we obtain the second relation in (40) which implies (36) holds. Thus, the proof is completed. □

4 Unconstrained stationary points

In this section we consider the monotone SOCCP and show that every stationary point of unconstrained minimization (11) is a solution of the SOCCP. First, we prove the following important properties of $\nabla \psi_\tau$, which will reduce to the results of [4, Lemma 4.2] when $\tau = 2$.

Lemma 4.1 *For any $x = (x_1, x_2),\ y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have*

$$\langle x, \nabla_x \psi_\tau(x, y) \rangle + \langle y, \nabla_y \psi_\tau(x, y) \rangle = \|\phi_\tau(x, y)\|^2, \tag{54}$$

$$\langle \nabla_x \psi_\tau(x, y), \nabla_y \psi_\tau(x, y) \rangle \;\ge\; 0. \tag{55}$$

Furthermore, the equality in (55) holds if and only if $\phi_\tau(x, y) = 0$.

Proof When $(x,y)=(0,0)$, $\nabla_x\psi_\tau(x,y)=\nabla_y\psi_\tau(x,y)=0$ by Proposition 3.2, and the conclusion is clear. We next consider the other two cases.

Case (1): $(x-y)^2+\tau(x\circ y)\in\text{int}(\mathcal{K}^n)$. By Proposition 3.2, we can compute that

$$\begin{aligned}
&\langle x,\nabla_x\psi_\tau(x,y)\rangle+\langle y,\nabla_y\psi_\tau(x,y)\rangle\\
&\quad=\big\langle x,\big(L_{x+\frac{\tau-2}{2}y}L_z^{-1}-I\big)\phi_\tau\big\rangle+\big\langle y,\big(L_{y+\frac{\tau-2}{2}x}L_z^{-1}-I\big)\phi_\tau\big\rangle\\
&\quad=\big\langle\big(L_z^{-1}L_{x+\frac{\tau-2}{2}y}-I\big)x,\phi_\tau\big\rangle+\big\langle\big(L_z^{-1}L_{y+\frac{\tau-2}{2}x}-I\big)y,\phi_\tau\big\rangle\\
&\quad=\big\langle L_z^{-1}\big[(x^2+y^2)+(\tau-2)(x\circ y)\big]-(x+y),\phi_\tau\big\rangle\\
&\quad=\langle L_z^{-1}z^2-(x+y),\phi_\tau\rangle=\|\phi_\tau\|^2,
\end{aligned}$$

where, for simplicity, $\phi_\tau(x,y)$ is written as ϕ_τ. This proves (54). Notice that

$$\langle\nabla_x\psi_\tau(x,y),\nabla_y\psi_\tau(x,y)\rangle=\big\langle(L_{y+\frac{\tau-2}{2}x}-L_z)(L_{x+\frac{\tau-2}{2}y}-L_z)L_z^{-1}\phi_\tau,L_z^{-1}\phi_\tau\big\rangle.$$

Let S be the symmetric part of $(L_{y+\frac{\tau-2}{2}x}-L_z)(L_{x+\frac{\tau-2}{2}y}-L_z)$. Then,

$$\begin{aligned}
S&=\frac{1}{2}\big[\big(L_{y+\frac{\tau-2}{2}x}-L_z\big)\big(L_{x+\frac{\tau-2}{2}y}-L_z\big)+\big(L_{x+\frac{\tau-2}{2}y}-L_z\big)\big(L_{y+\frac{\tau-2}{2}x}-L_z\big)\big]\\
&=\frac{1}{2}\bigg[L_yL_x+\frac{\tau-2}{2}L_x^2-L_zL_x+\frac{\tau-2}{2}L_y^2+\frac{(\tau-2)^2}{4}L_xL_y-\frac{\tau-2}{2}L_zL_y\\
&\quad-L_yL_z-\frac{\tau-2}{2}L_xL_z+L_z^2+L_xL_y+\frac{\tau-2}{2}L_y^2-L_zL_y\\
&\quad+\frac{\tau-2}{2}L_x^2+\frac{(\tau-2)^2}{4}L_yL_x-\frac{\tau-2}{2}L_zL_x-L_xL_z-\frac{\tau-2}{2}L_yL_z+L_z^2\bigg]\\
&=\frac{\tau}{4}(L_z-L_x-L_y)^2+\frac{4-\tau}{4}\big(L_z^2-L_{\widetilde{x}}^2-L_{\widetilde{y}}^2\big),
\end{aligned}$$

where $\widetilde{x}:=x+\frac{\tau-2}{2}y$ and $\widetilde{y}:=\frac{1}{2}\sqrt{\tau(4-\tau)}\,y$. Noting that $z\in\mathcal{K}^n$ and $z^2=\widetilde{x}^2+\widetilde{y}^2$, we have $L_z^2-L_{\widetilde{x}}^2-L_{\widetilde{y}}^2\succeq O$ by Proposition 3.4 of [9]. Consequently,

$$\begin{aligned}
\langle\nabla_x\psi_\tau(x,y),\nabla_y\psi_\tau(x,y)\rangle&=\langle SL_z^{-1}\phi_\tau,L_z^{-1}\phi_\tau\rangle\\
&\geq\frac{\tau}{4}\big\langle(L_z-L_x-L_y)^2L_z^{-1}\phi_\tau,L_z^{-1}\phi_\tau\big\rangle\\
&=\frac{\tau}{4}\big\|L_{\phi_\tau}L_z^{-1}\phi_\tau\big\|^2,
\end{aligned}\tag{56}$$

where the equality is due to $L_z-L_x-L_y=L_{\phi_\tau}$. This implies (55). If the inequality in (55) holds with equality, then the above relation yields $\|L_{\phi_\tau}L_z^{-1}\phi_\tau\|^2=0$, which says

$$L_{\phi_\tau}L_z^{-1}\phi_\tau=\phi_\tau\circ(L_z^{-1}\phi_\tau)=0.$$

By the definition of Jordan product, $\langle \phi_\tau, L_z^{-1}\phi_\tau \rangle = 0$. This implies $\phi_\tau = 0$ since $L_z^{-1} \succ O$. Conversely, if $\phi_\tau = 0$, then it follows from (24) that $\langle \nabla_x \psi_\tau(x, y), \nabla_y \psi_\tau(x, y) \rangle = 0$.

Case (2): $(x, y) \neq (0, 0)$ and $(x - y)^2 + \tau(x \circ y) \notin \mathrm{int}(\mathcal{K}^n)$. By (25), we can compute

$$\begin{aligned}
&\langle x, \nabla_x \psi_\tau(x, y) \rangle + \langle y, \nabla_y \psi_\tau(x, y) \rangle \\
&= \left\langle \frac{x_1 x + y_1 y + \frac{\tau-2}{2}(y_1 x + x_1 y)}{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1} - (x + y),\ \phi_\tau(x, y) \right\rangle \\
&= \|\phi_\tau(x, y)\|^2,
\end{aligned}$$

where the last equality uses (23). This proves (54). Equation (55) holds since

$$\begin{aligned}
&\langle \nabla_x \psi_\tau(x, y), \nabla_y \psi_\tau(x, y) \rangle \\
&= \left[\frac{x_1 + \frac{\tau-2}{2} y_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1}} - 1 \right] \left[\frac{y_1 + \frac{\tau-2}{2} x_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1}} - 1 \right] \|\phi_\tau(x, y)\|^2 \\
&\geq 0,
\end{aligned}$$

where the inequality is due to

$$\frac{x_1 + \frac{\tau-2}{2} y_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1}} \leq 1 \quad \text{and} \quad \frac{y_1 + \frac{\tau-2}{2} x_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1}} \leq 1.$$

If (55) holds with equality, then either $\phi_\tau(x, y) = 0$ or

$$\frac{x_1 + \frac{\tau-2}{2} y_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1}} = 1 \quad \text{or} \quad \frac{y_1 + \frac{\tau-2}{2} x_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2)x_1 y_1}} = 1.$$

In the second case, we have $y_1 = 0$ and $x_1 \geq 0$, so that Lemma 3.1 yields $y_2 = 0$ and $x_1 = \|x_2\|$. In the third case, we have $x_1 = 0$ and $y_1 \geq 0$, so that Lemma 3.1 yields $x_2 = 0$ and $y_1 = \|y_2\|$. Thus, in the two cases, we have $\langle x, y \rangle = 0,\ x \in \mathcal{K}^n,\ y \in \mathcal{K}^n$. Consequently, $\phi_\tau(x, y) = 0$ by Proposition 3.1. Conversely, if $\phi_\tau = 0$, then from (24) it follows that $\langle \nabla_x \psi_\tau(x, y), \nabla_y \psi_\tau(x, y) \rangle = 0$. The proof is thus completed. □

Next, we are in a position to establish another main result of this paper, that is, each stationary point of f_τ solves the SOCCP under the condition that

$$\nabla F(\zeta) \ \text{and} \ -\nabla G(\zeta) \ \text{are column monotone for any} \ \zeta \in \mathbb{R}^n. \tag{57}$$

Proposition 4.1 *Let f_τ be given by* (11). *If F and G satisfies the condition* (57), *then for every $\zeta \in \mathbb{R}^n$, either $f_\tau(\zeta) = 0$ or $\nabla f_\tau(\zeta) \neq 0$. If $\nabla f_\tau(\zeta) \neq 0$ and $\nabla G(\zeta)$ is invertible, then $\langle d(\zeta), \nabla f_\tau(\zeta) \rangle < 0$, where $d(\zeta) := -(\nabla G(\zeta)^{-1})^{\mathrm{T}} \nabla_x \psi_\tau(F(\zeta), G(\zeta))$.*

Proof By Lemma 4.1, using the same arguments as those of [4, Proposition 3] except that $\psi_{\rm FB}$ and $f_{\rm FB}$ are replaced by ψ_τ and f_τ yields the desired result, and we omit it. □

From [7, p. 1014] or [14, p. 222], $A, B \in \mathbb{R}^{n\times n}$ are column monotone if

$$Au + Bv = 0 \quad \Longrightarrow \quad u^{\rm T}v \geq 0 \quad \text{for any } u, v \in \mathbb{R}^n.$$

By this, it is not hard to verify that, if $\nabla G(\zeta)$ is invertible, the condition (57) is equivalent to requiring that $\nabla G(\zeta)^{-1}\nabla F(\zeta) \succeq O$ for any $\zeta \in \mathbb{R}^n$. This implies that, for the SOCCP (3), the condition (57) is actually equivalent to F being monotone.

5 Numerical experiments

In this section, we use the merit function approach based on the unconstrained minimization (11) to solve convex SOCPs in the form of (4), and compare the numerical performance of the method corresponding to different $\tau \in (0, 4)$. Throughout the experiments, we adopted the Cholesky factorization of $AA^{\rm T}$ to evaluate F, G in (5), which was completed via the Matlab routine `chol`. For the vector $\bar{x}$ satisfying $A\bar{x} = b$, we computed it as a solution of $\min_x \|Ax - b\|$ using Matlab's least square solver.

All experiments were done at a PC of Pentium 4 with 2.8 GHz CPU and 512 MB memory, and the computer codes were all written in Matlab 6.5. In the experiments, we adopted the L-BFGS method, a limited-memory quasi-Newton method, with 5 limited-memory vector-updates to solve the minimization problem (11). For the scaling matrix $H^0 = \gamma I$ in the L-BFGS, we adopted $\gamma = p^{\rm T}q/q^{\rm T}q$ as recommended by [17, p. 226], where $p := \zeta - \zeta^{\rm old}$ and $q := \nabla f_\tau(\zeta) - \nabla f_\tau(\zeta^{\rm old})$. To ensure convergence, we reverted to the steepest descent direction $-\nabla f_\tau(\zeta)$ whenever the current direction d failed to satisfy $q^{\rm T}d \leq 10^{-5}\|q\| \cdot \|d\|$. In addition, we used the nonmonotone line search as described in [10] to seek a suitable steplength, i.e., we computed the smallest nonnegative integer l such that

$$f_\tau(\zeta^k + \rho^l d^k) \leq \mathcal{W}_k + \sigma\rho^l \nabla f_\tau(\zeta^k)^{\rm T} d^k$$

where d^k denotes the direction in the k-th iteration generated by L-BFGS, ρ and σ are given parameters in the interval (0, 1), and $\mathcal{W}_k$ is given by

$$\mathcal{W}_k = \max_{j=k-m_k,\ldots,k} f_\tau(\zeta^j)$$

and where, for a given nonnegative integer $\hat{m}$ and s, we set

$$m_k = \begin{cases} 0 & \text{if } k \leq s, \\ \min\left\{m_{k-1} + 1, \hat{m}\right\} & \text{otherwise.} \end{cases}$$

Throughout the experiments, we chose the following parameters for the algorithm:

$$\rho = 0.5, \qquad \sigma = 1.0e - 4, \qquad \hat{m} = 5 \quad \text{and} \quad s = 5.$$

The starting point was set to be $\zeta^{\text{init}} = 0$, and the algorithm was stopped whenever the number of function evaluations for f_τ is over 10000 or $\max\{f_\tau(\zeta), |\langle F(\zeta), G(\zeta)\rangle|\} \le 10^{-6}$.

The first group of test instances is the linear SOCP from the DIMACS library [18], in which A is sparse and $g(x) = c^{\mathrm{T}}x$. Numerical results are summarized in Table 1, where the first row lists the name of the problem and the dimension (m, n) of A, NF denotes the number of function evaluations for solving each test problem, Gap means the absolute complementarity gap, i.e., the value of the function $|\langle F(\zeta), G(\zeta)\rangle|$ at the final iteration, and Cpu records the CPU time in second to reach termination condition for each test problem. For these test problems, we see from Table 1 that the

Table 1 Numerical results with different τ for three DIMACS problems

τ	nb (123, 2383)			nb-L2 (123, 4195)			nb-L2-Bessel (123, 2641)		
	NF	Gap	Cpu	NF	Gap	Cpu	NF	Gap	Cpu
0.05	6153	7.53e-8	1708.9	723	7.47e-7	230.9	305	1.17e-7	75.0
0.1	4093	6.68e-7	931.9	700	5.00e-8	172.7	241	4.29e-7	58.9
0.5	3370	2.78e-7	763.7	422	2.78e-7	116.7	234	2.52e-7	53.5
1.0	–	7.83e-6	2365.0	475	6.32e-7	136.3	166	6.71e-7	43.9
1.5	3257	6.36e-8	734.5	605	9.78e-7	159.4	161	4.04e-7	42.3
2.0	3672	6.49e-7	788.3	839	4.40e-7	196.8	287	7.99e-7	59.4
2.5	1218	5.76e-7	263.7	597	6.32e-8	139.1	222	9.77e-7	53.8
3.0	2537	7.19e-7	533.5	634	6.27e-7	155.4	289	5.43e-7	64.1
3.5	7404	1.14e-7	1867.1	706	8.70e-7	184.0	239	1.21e-7	57.8
3.9	2606	6.23e-7	513.8	793	9.87e-7	201.1	245	1.55e-7	55.0

A hyphen means the number of function evaluations is over 10 000

Table 2 Numerical results with different τ for CSOCPs with dense A

τ	(500, 50, 2) (353, 904)			(500, 50, 10) (349, 900)			(800, 80, 8) (577, 1458)			(1000, 100, 5) (723, 1824)		
	NF	Gap	Cpu	NF	Gap	Cpu	NF	Gap	Cpu	NF	Gap	Cpu
0.05	1086	6.84e-7	58.31	617	9.39e-7	37.50	746	9.53e-7	98.17	1592	1.68e-7	320.4
0.1	848	9.91e-7	45.12	525	8.07e-7	30.92	629	2.17e-7	84.84	1353	9.09e-7	255.7
0.5	1046	6.87e-7	58.59	389	5.34e-7	23.41	916	1.92e-7	111.1	1529	6.05e-7	308.5
1.0	1015	4.41e-8	55.90	407	6.00e-7	24.57	703	2.41e-7	96.98	1514	3.30e-7	312.3
1.5	1487	3.75e-7	78.36	472	6.22e-8	28.23	834	8.62e-7	102.3	1498	2.78e-7	309.0
2.0	1169	9.53e-7	62.03	463	7.27e-7	28.07	877	3.78e-7	113.8	1777	3.16e-7	350.6
2.5	1468	3.07e-7	75.87	573	2.35e-7	34.56	713	1.11e-7	94.3	1403	4.02e-7	291.8
3.0	1532	9.30e-7	79.11	683	4.81e-7	39.93	1180	4.13e-7	144.9	2594	2.24e-7	525.5
3.5	1802	4.54e-7	85.65	980	9.95e-7	52.93	1847	3.65e-7	212.5	3587	7.02e-7	675.8
3.9	4236	9.97e-8	149.6	1891	7.98e-7	73.31	4515	5.72e-7	370.9	9281	8.58e-7	1244.9

merit function method has better numerical performance when $\tau \in [1.5, 3]$, whereas as $\tau \to 0$ or $\tau \to 4$, its numerical performance becomes worse, which is particularly remarkable for the more difficult problem "nb". Also, the method with $\tau = 1.5$, 2.5 or 3 works better than with $\tau = 2$.

The second group of test instances is the convex SOCP with dense A. To generate such test problems randomly, we consider the more realistic problem of minimizing a sum of the k largest Euclidean norms with a convex regularization term: $\min_{u\geq 0}\sum_{i=1}^{k}\|s_{[i]}\| + h(u)$, where $\|s_{[1]}\|,\ldots,\|s_{[r]}\|$ are the norms $\|s_1\|,\ldots,\|s_r\|$ sorted in nonincreasing order with $r \geq k$ and $s_i = b_i - A_i x$ for $i = 1,\ldots,r$ with $A_i \in \mathbb{R}^{m_i\times l}$ and $b_i \in \mathbb{R}^{m_i}$, and $h : \mathbb{R}^l \to \mathbb{R}$ is a twice continuously differentiable convex function. The problem can be converted to the CSOCP:

$$
\begin{array}{ll}
\min & (1-k/r)\displaystyle\sum_{i=1}^{r} v_i + \frac{k}{r}\sum_{i=1}^{r} w_i + h(u) \\
\text{s.t.} & A_i u + s_i = b_i, \quad i = 1,2,\ldots,r, \\
& (w_1 - v_1) - (w_2 - v_2) = 0, \\
& \qquad\qquad\vdots \\
& (w_1 - v_1) - (w_r - v_r) = 0, \\
& u \geq 0, \qquad v_i \geq 0, \qquad (w_i, s_i) \in \times\mathcal{K}^{m_i+1}, \quad i = 1,2,\ldots,r.
\end{array}
$$

In our tests, we set $h(u) := \frac{1}{3}\|u\|_3^3$ with $\|\cdot\|_3$ denoting the 3-norm, and generated each m_i randomly from $\{2,\ldots,10\}$ and each entry of A_i and b_i randomly by a uniform distribution from $[-1, 1]$ and $[-5, 5]$, respectively. Thus, if $d \geq m = m_1 + \cdots + m_r$, the constraint matrix is dense. The numerical results are reported in Table 2, in which the first row lists several groups of different (l, r, k), and the second row gives the dimension (m, n) of A.

For the dense test problems, we see from Table 2 that the merit function method with $\tau \in [0.1, 2.5]$ has better numerical performance, and as $\tau \to 4$ or $\tau \to 0$, its numerical performance also becomes much worse. Combining with the above experiment results, we may draw a conclusion that the function ψ_τ with $\tau = 2$ is not the best, and if the parameter τ is appropriately chosen, ψ_τ can be an alternative for the FB merit function. For example, for the sparse linear SOCPs, the parameter $\tau = 2.5$ or 3.0 is more satisfactory, whereas for the dense convex SOCPs, $\tau = 0.1$, 0.5 or 1.0 seems to be more favorable.

6 Conclusions

We have considered a one-parametric class of merit functions for the SOCCP. We showed that ψ_τ is continuously differentiable everywhere and shares the same favorable properties as the FB merit function ψ_{FB} which was recently studied in [4]. Although the proof techniques used in this paper may look similar to those used in [4], the algebraic analysis is indeed much harder and subtle since a generalization is considered. In addition, numerical results are reported for the merit function approach based on ψ_τ by solving the sparse and dense convex SOCPs, which indicate that ψ_τ can be an alternative for $\psi_{\rm FB}$ when the parameter τ is appropriately chosen, as well as provides some helpful advices on the choice of τ. Thus, it is worthwhile to study other algorithms based on the class of functions. Another direction is to study the convergence properties which are seldom analyzed for the SOCCP. We will leave them as future research topics.

Appendix: Proofs of Lemmas 3.2 and 3.3

Proof of Lemma 3.2 The first inequality can be proved by expanding the square on both sides and using the Cauchy-Schwartz inequality. It remains to show the second inequality. Since the left-hand side of the second inequality can be simplified as

$$\|x\|^2+\frac{(\tau-2)^2}{4}\|y\|^2+(\tau-2)x^{\rm T}y \\ +2(-1)^i\left(x_1+\frac{\tau-2}{2}y_1\right)\left(x_2+\frac{\tau-2}{2}y_2\right)^{\rm T}\frac{w_2}{\|w_2\|},$$

whereas the right-hand side equals to

$$\|x\|^2+\|y\|^2+(\tau-2)x^{\rm T}y+2(-1)^i+(-1)^i\|w_2\|,$$

we only need to prove the following inequality

$$(-1)^i\left[2\left(x_1+\frac{\tau-2}{2}y_1\right)\left(x_2+\frac{\tau-2}{2}y_2\right)^{\rm T}\frac{w_2}{\|w_2\|}-\|w_2\|\right]\le\frac{\tau(4-\tau)}{4}\|y\|^2.$$

Considering that $\frac{\tau(4-\tau)}{4}>0$ and $\|w_2\|>0$, the last inequality is actually equivalent to

$$\left|2\left(x_1+\frac{\tau-2}{2}y_1\right)\left(x_2+\frac{\tau-2}{2}y_2\right)^{\rm T}w_2-\|w_2\|^2\right|\le\frac{\tau(4-\tau)}{4}\|y\|^2\|w_2\|. \quad (58)$$

By using $w_2=2(x_1x_2+y_1y_2)+(\tau-2)(x_1y_2+y_1x_2)$, we can compute

$$2\big(x_1+((\tau-2)/2)y_1\big)\big(x_2+((\tau-2)/2)y_2\big)^{\rm T}w_2 \\ =\left[4x_1y_1+4(\tau-2)x_1^2+2(\tau-2)y_1^2+3(\tau-2)^2x_1y_1+\frac{(\tau-2)^3}{2}y_1^2\right]x_2^{\rm T}y_2 \\ +\left[2(\tau-2)x_1y_1+(\tau-2)^2x_1^2+(\tau-2)^2y_1^2+\frac{(\tau-2)^3}{2}x_1y_1\right]\|y_2\|^2 \\ +\left[4x_1^2+4(\tau-2)x_1y_1+(\tau-2)^2y_1^2\right]\|x_2\|^2$$

and

$$\|w_2\|^2=\left[8x_1y_1+2(\tau-2)^2x_1y_1+4(\tau-2)x_1^2+4(\tau-2)y_1^2\right]x_2^{\rm T}y_2$$

$$+\left[4y_1^2+(\tau-2)^2x_1^2+4(\tau-2)x_1y_1\right]\|y_2\|^2$$
$$+\left[4x_1^2+4(\tau-2)x_1y_1+(\tau-2)^2y_1^2\right]\|x_2\|^2.$$

Applying these two equalities, it then follows that

$$\begin{aligned}
&2\big(x_1+((\tau-2)/2)y_1\big)\big(x_2+((\tau-2)/2)y_2\big)^{\mathrm{T}}w_2-\|w_2\|^2\\
&=\Big[\big((\tau-2)^2-4\big)x_1y_1+\Big((\tau-2)^3/2-2(\tau-2)\Big)y_1^2\Big]x_2^{\mathrm{T}}y_2\\
&\quad+\Big[\big((\tau-2)^2-4\big)y_1^2+\big((\tau-2)^3/2-2(\tau-2)\big)x_1y_1\Big]\|y_2\|^2\\
&=(\tau^2-4\tau)\left[x_1y_1x_2^{\mathrm{T}}y_2+\frac{\tau-2}{2}y_1^2x_2^{\mathrm{T}}y_2+y_1^2\|y_2\|^2+\frac{\tau-2}{2}x_1y_1\|y_2\|^2\right].
\end{aligned}$$

From this, to show the inequality in (58), it suffices to prove that

$$\begin{aligned}
&\left|4x_1y_1x_2^{\mathrm{T}}y_2+2(\tau-2)x_1y_1\|y_2\|^2+4y_1^2\|y_2\|^2+2(\tau-2)y_1^2x_2^{\mathrm{T}}y_2\right|\\
&\quad\le\|y\|^2\|w_2\|.
\end{aligned}\tag{59}$$

Let L and R denote, respectively, the square of the left-hand side and the right-hand side of (59). We argue the assertion (59) by verifying that $R-L\ge 0$. Since

$$\begin{aligned}
L&=\big(2x_1+(\tau-2)y_1\big)^24y_1^2(x_2^{\mathrm{T}}y_2)^2+\big(2y_1+(\tau-2)x_1\big)^24y_1^2\|y_2\|^4\\
&\quad+8y_1^2\|y_2\|^2x_2^{\mathrm{T}}y_2\big(4x_1y_1+2(\tau-2)x_1^2+2(\tau-2)y_1^2+(\tau-2)^2x_1y_1\big),
\end{aligned}$$

and

$$\begin{aligned}
R&=\|y\|^4\big[(2x_1+(\tau-2)y_1)^2\|x_2\|^2+(2y_1+(\tau-2)x_1)^2\|y_2\|^2\big]\\
&\quad+\|y\|^4\big[8x_1y_1+2x_1y_1(\tau-2)^2+4(\tau-2)(x_1^2+y_1^2)\big]x_2^{\mathrm{T}}y_2.
\end{aligned}$$

Taking the difference between R and L leads to

$$\begin{aligned}
R-L&=\big(2x_1+(\tau-2)y_1\big)^2\big(\|y\|^4\|x_2\|^2-4y_1^2(x_2^{\mathrm{T}}y_2)^2\big)\\
&\quad+\big(2y_1+(\tau-2)x_1\big)^2\big(\|y\|^4\|y_2\|^2-4y_1^2\|y_2\|^4\big)\\
&\quad+8x_1y_1x_2^{\mathrm{T}}y_2\big(\|y\|^4-4y_1^2\|y_2\|^2\big)\\
&\quad+4(\tau-2)x_2^{\mathrm{T}}y_2x_1^2\big(\|y\|^4-4y_1^2\|y_2\|^2\big)\\
&\quad+4(\tau-2)x_2^{\mathrm{T}}y_2y_1^2\big(\|y\|^4-4y_1^2\|y_2\|^2\big)\\
&\quad+2(\tau-2)^2x_1y_1x_2^{\mathrm{T}}y_2\big(\|y\|^4-4y_1^2\|y_2\|^2\big)\\
&\ge\big(\|y\|^4-4y_1^2\|y_2\|^2\big)\Big[\big(2x_1+(\tau-2)y_1\big)^2\|x_2\|^2+\big(2y_1+(\tau-2)x_1\big)^2\|y_2\|^2\\
&\quad+8x_1y_1x_2^{\mathrm{T}}y_2+4(\tau-2)x_2^{\mathrm{T}}y_2x_1^2+4(\tau-2)x_2^{\mathrm{T}}y_2y_1^2+2(\tau-2)^2x_1y_1x_2^{\mathrm{T}}y_2\Big]\\
&=(y_1^2-\|y_2\|^2)^2\big\|2x_1x_2+(\tau-2)y_1x_2+2y_1y_2+(\tau-2)x_1y_2\big\|^2\ge 0.
\end{aligned}$$

By the symmetry of x and y, the above results also hold when interchanging x and y. □

Proof of Lemma 3.3 By the symmetry of x and y, it suffices to prove the first inequality in (28). Let $x=(x_1,x_2),\ y=(y_1,y_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$ with $(x-y)^2+\tau(x\circ y)\in\operatorname{int}(\mathcal{K}^n)$. We will proceed the proof by the following two cases: (1) $w_2=0$ and (2) $w_2\neq 0$.

Case (1): $w_2=0$. In this case, $z_2=0$ and $z_1=\sqrt{\|x\|^2+\|y\|^2+(\tau-2)x^{\mathrm T}y}>0$. Hence,

$$L_{x+\frac{\tau-2}{2}y}L_z^{-1}=\frac{1}{\sqrt{\|x\|^2+\|y\|^2+(\tau-2)x^{\mathrm T}y}}\begin{pmatrix} x_1+\frac{\tau-2}{2}y_1 & x_2^{\mathrm T}+\frac{\tau-2}{2}y_2^{\mathrm T}\\ x_2+\frac{\tau-2}{2}y_2 & (x_1+\frac{\tau-2}{2}y_1)I\end{pmatrix}.$$

Notice that $\|x\|^2+\|y\|^2+(\tau-2)x^{\mathrm T}y=\|x+\frac{\tau-2}{2}y\|^2+\frac{\tau(4-\tau)}{4}\|y\|^2$. Therefore,

$$\frac{|x_1+\frac{\tau-2}{2}y_1|}{\sqrt{\|x\|^2+\|y\|^2+(\tau-2)x^{\mathrm T}y}}\le 1\quad\text{and}\quad\frac{\|x_2+\frac{\tau-2}{2}y_2\|}{\sqrt{\|x\|^2+\|y\|^2+(\tau-2)x^{\mathrm T}y}}\le 1.$$

This shows that each element in $L_{x+\frac{\tau-2}{2}y}L_z^{-1}$ is bounded with the bound independent of x, y and τ, and consequently the first inequality in (28) holds under this case.

Case (2): $w_2\neq 0$. Now let λ_1 and λ_2 be the spectral values of w. By (18) and (14),

$$L_{x+\frac{\tau-2}{2}y}L_z^{-1}=\begin{pmatrix} bs_1+cs_2^{\mathrm T}\bar w_2 & cs_1\bar w_2^{\mathrm T}+as_2^{\mathrm T}+(b-a)s_2^{\mathrm T}\bar w_2\bar w_2^{\mathrm T}\\ bs_2+cs_1\bar w_2 & cs_2\bar w_2^{\mathrm T}+as_1I+(b-a)s_1\bar w_2\bar w_2^{\mathrm T}\end{pmatrix},$$

where $\bar w_2=\frac{w_2}{\|w_2\|}$, $s=(s_1,s_2)=x+\frac{\tau-2}{2}y$, and a, b and c are given by

$$a=\frac{2}{\sqrt{\lambda_2}+\sqrt{\lambda_1}},\qquad b=\frac{\sqrt{\lambda_2}+\sqrt{\lambda_1}}{2\sqrt{\lambda_2\lambda_1}},\qquad c=\frac{\sqrt{\lambda_1}-\sqrt{\lambda_2}}{2\sqrt{\lambda_2\lambda_1}}.$$

Using Lemma 3.2 and noting that $s_1=x_1+\frac{\tau-2}{2}y_1$ and $s_2=x_2+\frac{\tau-2}{2}y_2$, we have

$$\begin{aligned}
\left|bs_1+cs_2^{\mathrm T}\bar w_2\right| &\le \frac{1}{2\sqrt{\lambda_2}}\left|s_1+s_2^{\mathrm T}\bar w_2\right|+\frac{1}{2\sqrt{\lambda_1}}\left|s_1-s_2^{\mathrm T}\bar w_2\right|\le 1,\\
\|bs_2+cs_1\bar w_2\| &\le \frac{1}{2\sqrt{\lambda_2}}\|s_2+s_1\bar w_2\|+\frac{1}{2\sqrt{\lambda_1}}\|s_2-s_1\bar w_2\|\le 1,\\
\left\|cs_1\bar w_2^{\mathrm T}+bs_2^{\mathrm T}\bar w_2\bar w_2^{\mathrm T}\right\| &= \left\|\frac{1}{2\sqrt{\lambda_2}}\left(s_1+s_2^{\mathrm T}\bar w_2\right)\bar w_2^{\mathrm T}-\frac{1}{2\sqrt{\lambda_1}}\left(s_1-s_2^{\mathrm T}\bar w_2\right)\bar w_2^{\mathrm T}\right\|\\
&\le \frac{1}{2\sqrt{\lambda_2}}\left|s_1+s_2^{\mathrm T}\bar w_2\right|+\frac{1}{2\sqrt{\lambda_1}}\left|s_1-s_2^{\mathrm T}\bar w_2\right|\le 1,\\
\left\|as_2^{\mathrm T}-as_2^{\mathrm T}\bar w_2\bar w_2^{\mathrm T}\right\| &\le \left\|\frac{2s_2^{\mathrm T}}{\sqrt{\lambda_2}+\sqrt{\lambda_1}}\right\|\cdot\left\|I-\bar w_2\bar w_2^{\mathrm T}\right\|_F\le 2(n+1),
\end{aligned}$$

$$\left\| cs_2\bar{w}_2^T + bs_1\bar{w}_2\bar{w}_2^T \right\|_F = \left\| \frac{1}{2\sqrt{\lambda_2}}\left(s_2 + s_1\bar{w}_2\right)\bar{w}_2^T - \frac{1}{2\sqrt{\lambda_1}}\left(s_2 - s_1\bar{w}_2\right)\bar{w}_2^T \right\|_F$$

$$\leq \frac{1}{2\sqrt{\lambda_2}}\|s_2 + s_1\bar{w}_2\| + \frac{1}{2\sqrt{\lambda_1}}\|s_2 - s_1\bar{w}_2\| \leq 1,$$

$$\left\| as_1 I - as_1\bar{w}_2\bar{w}_2^T \right\|_F \leq \left\| \frac{2s_1}{\sqrt{\lambda_2} + \sqrt{\lambda_1}} \right\| \cdot \left\| I - \bar{w}_2\bar{w}_2^T \right\|_F \leq 2(n+1).$$

The above inequalities show that every entry of $L_{x+\frac{\tau-2}{2}y}L_z^{-1}$ is bounded with the bound independent of x, y and τ. Thus, the first inequality in (28) also holds under this case. □

Acknowledgement The work of J.-S. Chen is partially supported by National Science Council of Taiwan.

The authors would like to thank two anonymous referees for their helpful comments and suggestions on the revision of this paper.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**, 3–51 (2003)
2. Alizadeh, F., Schmieta, S.: Symmetric cones, potential reduction methods, and word-by-word extensions. In: Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.) Handbook of Semidefinite Programming, pp. 195–233. Kluwer Academic, Boston (2000)
3. Andersen, E.D., Roos, C., Terlaky, T.: On implementing a primal-dual interior-point method for conic quadratic optimization. Math. Program. **95**, 249–277 (2003)
4. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. Math. Program. **104**, 293–327 (2005)
5. Chen, J.-S., Chen, X., Tseng, P.: Analysis of nonsmooth vector-valued functions associated with second-order cone. Math. Program. **101**, 95–117 (2004)
6. Chen, X.-D., Sun, D., Sun, J.: Complementarity functions and numerical experiments for second-order cone complementarity problems. Comput. Optim. Appl. **25**, 39–56 (2003)
7. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vols. I and II. Springer, New York (2003)
8. Ferris, M.C., Mangasarian, O.L., Pang, J.-S. (eds.): Complementarity: Applications, Algorithms and Extensions. Kluwer Academic, Dordrecht (2001)
9. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2002)
10. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
11. Hayashi, S., Yamashita, N., Fukushima, M.: A combined smoothing and regularization method for monotone second-order cone complementarity problems. SIAM J. Optim. **15**, 593–615 (2005)
12. Kanzow, C., Kleinmichel, H.: A new class of semismooth Newton-type methods for nonlinear complementarity problems. Comput. Optim. Appl. **11**, 227–251 (1998)
13. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Application of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
14. Luo, Z.-Q., Tseng, P.: A new class of merit functions for the nonlinear complementarity problem. In: Ferris, M.C., Pang, J.-S. (eds.) Complementarity and Variational Problems: State of the Art, pp. 204–225. SIAM, Philadelphia (1997)
15. Mittelmann, H.D.: An independent benchmarking of SDP and SOCP solvers. Math. Program. **95**, 407–430 (2003)
16. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions. Math. Program. **88**, 61–83 (2000)

17. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, New York (1999)
18. Pataki, G., Schmieta, S.: The DIMACS library of semidefinite-quadratic-linear programs. Preliminary draft, Computational Optimization Research Center, Columbia University, New York. http://dimacs.rutgers.edu/Challenges
19. Tsuchiya, T.: A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim. Methods Softw. **11**, 141–182 (1999)