Math. Meth. Oper. Res. (2006) 64: 495–519
DOI 10.1007/s00186-006-0098-9

ORIGINAL ARTICLE

Jein-Shan Chen

Two classes of merit functions for the second-order cone complementarity problem

Received: 4 June 2005 / Accepted: 11 January 2006 / Published online: 7 November 2006

Abstract Recently Tseng (Math Program 83:159–185, 1998) extended a class of merit functions, proposed by Luo and Tseng (*A new class of merit functions for the nonlinear complementarity problem*, in Complementarity and Variational Problems: State of the Art, pp. 204–225, 1997), for the nonlinear complementarity problem (NCP) to the semidefinite complementarity problem (SDCP) and showed several related properties. In this paper, we extend this class of merit functions to the second-order cone complementarity problem (SOCCP) and show analogous properties as in NCP and SDCP cases. In addition, we study another class of merit functions which are based on a slight modification of the aforementioned class of merit functions. Both classes of merit functions provide an error bound for the SOCCP and have bounded level sets.

Keywords Error bound · Jordan product · Level set · Merit function · Second-order cone · Spectral factorization

AMS subject classifications 26B05 · 90C33

1 Introduction

We consider the following conic complementarity problem of finding $x, y \in \mathbb{R}^n$ and $\zeta \in \mathbb{R}^n$ satisfying

$$\langle x, y\rangle = 0, \quad x \in \mathcal{K}, \quad y \in \mathcal{K}, \tag{1}$$

$$x = F(\zeta), \quad y = G(\zeta), \tag{2}$$

Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

J.-S. Chen
Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan
E-mail: jschen@math.ntnu.edu.tw

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 smooth (i.e., continuously differentiable) mappings, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOC), also called Lorentz cones (Faraut and Korányi 1994). In other words,

$$\mathcal{K}=\mathcal{K}^{n_1}\times\cdots\times\mathcal{K}^{n_N}, \tag{3}$$

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

$$\mathcal{K}^{n_i}:=\{(x_1,x_2)\in\mathbb{R}\times\mathbb{R}^{n_i-1}\mid \|x_2\|\leq x_1\}, \tag{4}$$

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of nonnegative reals $\mathbb{R}_+$. A special case of (3) is $\mathcal{K}=\mathbb{R}^n_+$, the nonnegative orthant in $\mathbb{R}^n$, which corresponds to $N=n$ and $n_1=\cdots=n_N=1$. We will refer to (1), (2), (3) as the *second-order cone complementarity problem* (SOCCP).

An important special case of SOCCP corresponds to $G(\zeta)=\zeta$ for all $\zeta\in\mathbb{R}^n$. Then (1) and (2) reduce to

$$\langle F(\zeta),\zeta\rangle=0,\quad F(\zeta)\in\mathcal{K},\ \ \zeta\in\mathcal{K}, \tag{5}$$

which is a natural extension of the nonlinear complementarity problem (NCP) where $\mathcal{K}=\mathbb{R}^n_+$. Another important special case of SOCCP corresponds to the Karush–Kuhn–Tucker (KKT) optimality conditions for the second-order cone program (SOCP) (see Chen and Tseng 2005 for details):

$$\begin{aligned}&\text{minimize } c^{\mathrm{T}}x\\&\text{subject to } Ax=b,\ \ x\in\mathcal{K},\end{aligned} \tag{6}$$

where $A\in\mathbb{R}^{m\times n}$ has full row rank, $b\in\mathbb{R}^m$ and $c\in\mathbb{R}^n$.

For simplicity, we will focus on $\mathcal{K}=\mathcal{K}^n$ throughout the whole paper. All the analysis can be carried over to the general case where $\mathcal{K}$ has the direct product structure as (3). It is known that $\mathcal{K}^n$ is a closed convex cone with interior given by

$$\operatorname{int}(\mathcal{K}^n)=\{(x_1,x_2)\in\mathbb{R}\times\mathbb{R}^{n-1}\mid \|x_2\|<x_1\}.$$

For any x,y in $\mathbb{R}^n$, we write $x\succeq_{\mathcal{K}^n}y$ if $x-y\in\mathcal{K}^n$; and write $x\succ_{\mathcal{K}^n}y$ if $x-y\in\operatorname{int}(\mathcal{K}^n)$. In other words, we have $x\succeq_{\mathcal{K}^n}0$ if and only if $x\in\mathcal{K}^n$ and $x\succ_{\mathcal{K}^n}0$ if and only if $x\in\operatorname{int}(\mathcal{K}^n)$. The relation $\succeq_{\mathcal{K}^n}$ is a partial ordering, i.e., it is anti-symmetric, transitive, and reflexive. Nonetheless, it is not a total ordering in $\mathcal{K}^n$.

There have been various methods proposed for solving SOCP and SOCCP. They include interior-point methods (Alizadeh and Schmieta 2000; Andersen et al. 2003; Lobo et al. 1998; Mittelmann 2003; Monteiro and Tsuchiya 2000; Schmieta and Alizadeh 2001; Tsuchiya 1999), non-interior smoothing Newton methods (Chen et al. 2003; Fukushima et al. 2002; Hayashi et al. 2002), and smoothing–regularization methods (Hayashi et al. 2005). Recently, the author and his co-author studied an alternative approach based on reformulating SOCP and SOCCP as an unconstrained smooth minimization problem (Chen and Tseng 2005). In that approach, it aimed to find a smooth function $\psi:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}_+$ such that

$$\psi(x,y)=0\iff x\in\mathcal{K}^n,\quad y\in\mathcal{K}^n,\ \ \langle x,y\rangle=0. \tag{7}$$

Then SOCCP can be expressed as an unconstrained smooth (global) minimization problem:

$$\min_{\zeta \in \mathbb{R}^n} \; f(\zeta) := \psi(F(\zeta), G(\zeta)). \tag{8}$$

We call such a f a *merit function* for the SOCCP.

A popular choice of ψ is the squared norm of Fischer–Burmeister function, i.e., $\psi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ associated with second-order cone given by

$$\psi_{\text{FB}}(x, y) \;=\; \frac{1}{2}\|\phi_{\text{FB}}(x, y)\|^2, \tag{9}$$

where $\phi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the well-known Fischer–Burmeister function (Fischer 1992, 1997) defined by

$$\phi_{\text{FB}}(x, y) = (x^2 + y^2)^{1/2} - x - y. \tag{10}$$

More specifically, for any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their *Jordan product* associated with $\mathcal{K}^n$ as

$$x \circ y \;:=\; (\langle x, y\rangle,\; y_1 x_2 + x_1 y_2). \tag{11}$$

The Jordan product $\circ$, 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 $e := (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbb{R}^n$. We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in \mathbb{R}^n$. Moreover, if $x \in \mathcal{K}^n$, then there exists a unique vector in $\mathcal{K}^n$, denoted by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Thus, ϕ_{FB} defined as (10) is well-defined for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ and maps $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. It was shown by Fukushima et al. (2002) that $\phi_{\text{FB}}(x, y) = 0$ if and only if (x, y) satisfies (1). Therefore, ψ_{FB} defined as (9) induces a merit function for the SOCCP.

In this paper, we study two classes of merit functions for the SOCCP. The first class is

$$f_{\text{LT}}(\zeta) := \psi_0(\langle F(\zeta), G(\zeta)\rangle) + \psi(F(\zeta), G(\zeta)), \tag{12}$$

where $\psi_0 : \mathbb{R} \to \mathbb{R}_+$ satisfies

$$\psi_0(t) = 0 \;\; \forall t \le 0 \text{ and } \psi_0'(t) > 0 \;\; \forall t > 0, \tag{13}$$

and $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ satisfies

$$\psi(x, y) = 0, \;\; \langle x, y\rangle \le 0 \iff (x, y) \in \mathcal{K}^n \times \mathcal{K}^n, \;\; \langle x, y\rangle = 0. \tag{14}$$

The function f_{LT} was proposed by Luo and Tseng (1997) for NCP case and was extended to the SDCP case by Tseng (1998). We explore the extension to the SOCCP as will be seen in Sects. 3 and 4. In addition, we make a slight modification of f_{LT} which forms another class of merit function as below.

$$\widehat{f_{\text{LT}}}(\zeta) := \psi_0^*(F(\zeta) \circ G(\zeta)) + \psi(F(\zeta), G(\zeta)), \tag{15}$$

where $\psi_0^* : \mathbb{R}^n \to \mathbb{R}_+$ is given as

$$\psi_0^*(w) = \frac{1}{2}\|(w)_+\|^2. \tag{16}$$

and $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ satisfies (14). We notice that ψ_0^* possesses the following property:

$$\psi_0^*(w) = 0 \iff w \preceq_{\mathcal{K}^n} 0, \tag{17}$$

which is a similar feature to (13) in some sense. Examples of ψ_0 and ψ will be given in Sect. 3. The second class of merit functions for SDCP case was recently studied (Goes and Oliveira 2002) and a variant of $\widehat{f_{\rm LT}}$ was also studied by the author (Chen 2006).

We will show that both $f_{\rm LT}$ and $\widehat{f_{\rm LT}}$ provide global error bound (Propositions 4.1 and 4.2), which plays an important role in analyzing the convergence rate of some iterative methods for solving the SOCCP, if F and G are jointly strongly monotone. We will also prove that if F and G are jointly monotone and a strictly feasible solution exists then both $f_{\rm LT}$ and $\widehat{f_{\rm LT}}$ have bounded level sets (Propositions 4.3 and 4.4) which will ensure that the sequence generated by a descent algorithm has at least an accumulation point. All these properties will make it possible to construct a descent algorithm for solving the equivalent unconstrained reformulation of the SOCCP. In contrast, the merit function induced by $\psi_{\rm FB}$ lacks these properties. In addition, we will show that both $f_{\rm LT}$ and $\widehat{f_{\rm LT}}$ are differentiable and their gradients have computable formulas. All the aforementioned features are significant reasons for choosing and studying these new merit functions.

Finally, we point out that SOCCP can be reduced to an SDCP by observing that, for any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have $x \in \mathcal{K}^n$ if and only if

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

is positive semidefinite (also see Fukushima et al. 2002, p. 437 and Sim and Zhao 2005). However, this reduction increases the problem dimension from n to $n(n+1)/2$ and it is not known whether this increase can be mitigated by exploiting the special "arrow" structure of L_x.

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors and $^{\rm T}$ denotes transpose. For any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x. For any differentiable mapping $F = (F_1, ..., F_m)^{\rm T} : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x) \cdots \nabla F_m(x)]$ is a $n \times m$ matrix which denotes the transpose Jacobian of F at x. For any symmetric matrices $A, B \in \mathbb{R}^{n\times n}$, we write $A \succeq B$ (respectively, $A \succ B$) to mean $A - B$ is positive semidefinite (respectively, positive definite). For nonnegative scalars α and β, we write $\alpha = O(\beta)$ to mean $\alpha \le C\beta$, with C independent of α and β. For any $x \in \mathbb{R}^n$, $(x)_+$ is used to denote the orthogonal projection of x onto $\mathcal{K}^n$, whereas $(x)_-$ means the orthogonal projection of x onto $-\mathcal{K}^n$. Also we denote $\mathcal{C}^* := \{y \mid \langle x, y\rangle \ge 0 \ \forall x \in K\}$ the dual cone of $\mathcal{C}$, given any closed convex cone $\mathcal{C}$.

2 Preliminaries

In this section, we review some background materials and preliminary results obtained by the author and his co-author in (Chen and Tseng 2005) that will be used later. We begin with the determinant and trace of x. For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, its *determinant* and *trace* are defined by

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

In general, $\det(x \circ y) \neq \det(x)\det(y)$ unless $x_2 = y_2$. Besides, we observe that $\mathrm{tr}(x \circ y) = 2\langle x, y\rangle$. We next recall from Fukushima et al. (2002) 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

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

where λ_1, λ_2 and $u^{(1)}, u^{(2)}$ are the spectral values and the associated spectral vectors of x given by

$$\begin{aligned} \lambda_i &= x_1 + (-1)^i \|x_2\|, \\ u^{(i)} &= \begin{cases} \frac{1}{2}\left(1,\ (-1)^i \dfrac{x_2}{\|x_2\|}\right) & \text{if } x_2 \neq 0; \\ \frac{1}{2}\left(1,\ (-1)^i w_2\right) & \text{if } x_2 = 0, \end{cases} \end{aligned}$$

for $i = 1, 2$, with w_2 being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w_2\| = 1$. If $x_2 \neq 0$, the factorization is unique.

The above spectral factorization of x, as well as x^2 and $x^{1/2}$ and the matrix L_x, have various interesting properties; see Fukushima et al. (2002). We list four properties that we will use in the subsequent sections.

Property 2.1 *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, with spectral values λ_1, λ_2 and spectral vectors $u^{(1)}, u^{(2)}$, the following results hold.*

(a) *$tr(x) = \lambda_1 + \lambda_2$ and $det(x) = \lambda_1\lambda_2$.*
(b) *If $x \in \mathcal{K}^n$, then $0 \le \lambda_1 \le \lambda_2$ and $x^{1/2} = \sqrt{\lambda_1}\, u^{(1)} + \sqrt{\lambda_2}\, u^{(2)}$.*
(c) *If $x \in \mathrm{int}(\mathcal{K}^n)$, then $0 < \lambda_1 \le \lambda_2$, and L_x is invertible with*

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

(d) *$x \circ y = L_x y$ for all $y \in \mathbb{R}^n$, and $L_x \succ 0$ if and only if $x \in \mathrm{int}(\mathcal{K}^n)$.*

In the following, we present some preliminary properties about ϕ_{FB} and ψ_{FB} given as (10) and (9), respectively, which are crucial to proving the results in Sects. 3 and 4. We only indicate their sources and omit the proofs since they can be found in Chen and Tseng (2005) and Fukushima et al. (2002).

Lemma 2.1 ([Fukushima et al. (2002), Proposition 2.1]) *Let* $\phi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *be given by* (10). *Then*

$$\begin{aligned}\phi_{\text{FB}}(x, y) = 0 &\iff x, y \in \mathcal{K}^n,\ x \circ y = 0,\\ &\iff x, y \in \mathcal{K}^n,\ \langle x, y\rangle = 0.\end{aligned}$$

Lemma 2.2 ([Chen and Tseng (2005), Lemma 3.2]) *For any* $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ *with* $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, *we have*

$$\begin{aligned}x_1^2 &= \|x_2\|^2,\\ y_1^2 &= \|y_2\|^2,\\ x_1 y_1 &= x_2^T y_2,\\ x_1 y_2 &= y_1 x_2.\end{aligned}$$

Lemma 2.3 ([Chen and Tseng (2005), Proposition 3.1, 3.2]) *Let* ϕ_{FB}, ψ_{FB} *be given as* (10) *and* (9), *respectively. Then,* ψ_{FB} *has the following properties.*

(a) $\psi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ *satisfies* (7).
(b) ψ_{FB} *is continuously differentiable at every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. *Moreover,* $\nabla_x \psi_{\text{FB}}(0, 0) = \nabla_y \psi_{\text{BF}}(0, 0) = 0$. *If* $(x, y) \neq (0, 0)$ *and* $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, *then*

$$\begin{aligned}\nabla_x \psi_{\text{FB}}(x, y) &= \left(L_x L^{-1}_{(x^2+y^2)^{1/2}} - I\right)\phi_{\text{FB}}(x, y),\\ \nabla_y \psi_{\text{FB}}(x, y) &= \left(L_y L^{-1}_{(x^2+y^2)^{1/2}} - I\right)\phi_{\text{FB}}(x, y).\end{aligned} \tag{18}$$

If $(x, y) \neq (0, 0)$ *and* $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, *then* $x_1^2 + y_1^2 \neq 0$ *and*

$$\nabla_x \psi_{\text{FB}}(x, y) = \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1\right)\phi_{\text{FB}}(x, y), \tag{19}$$

$$\nabla_y \psi_{\text{FB}}(x, y) = \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1\right)\phi_{\text{FB}}(x, y). \tag{20}$$

Lemma 2.4 ([Chen and Tseng (2005), Lemma 5.1]) *Let* $\mathcal{C}$ *be any closed convex cone in* $\mathbb{R}^n$. *For each* $x \in \mathbb{R}^n$, *let* $x_{\mathcal{C}}^+$ *and* $x_{\mathcal{C}}^-$ *denote the nearest-point* (*in the Euclidean norm*) *projection of* x *onto* $\mathcal{C}$ *and* $-\mathcal{C}^*$, *respectively. Then, the following results hold.*

(a) *For any* $x \in \mathbb{R}^n$, *we have* $x = x_{\mathcal{C}}^+ + x_{\mathcal{C}}^-$ *and* $\|x\|^2 = \|x_{\mathcal{C}}^+\|^2 + \|x_{\mathcal{C}}^-\|^2$.
(b) *For any* $x \in \mathbb{R}^n$ *and* $y \in \mathcal{C}$, *we have* $\langle x, y\rangle \leq \langle x_{\mathcal{C}}^+, y\rangle$.
(c) *If* $\mathcal{C}$ *is self-dual, then for any* $x \in \mathbb{R}^n$ *and* $y \in \mathcal{C}$, *we have* $\left\|(x + y)_{\mathcal{C}}^+\right\| \geq \left\|x_{\mathcal{C}}^+\right\|$.

Proof In fact, part (a) and (b) are classical results of Korányi (1984). □

Lemma 2.5 ([Chen and Tseng 2005, Lemma 5.2]) *Let* ϕ_{FB}, ψ_{FB} *be given by* (10) *and* (9), *respectively. For any* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, *we have*

$$4\psi_{\text{FB}}(x, y) \geq 2\left\|\phi_{\text{FB}}(x, y)_+\right\|^2 \geq \left\|(-x)_+\right\|^2 + \left\|(-y)_+\right\|^2.$$

To close this section, we recall some definitions that will be used for analysis in subsequent sections. We say that F and G are *jointly monotone* if

$$\langle F(\zeta) - F(\xi), G(\zeta) - G(\xi) \rangle \geq 0 \quad \forall \zeta,\ \xi \in \mathbb{R}^n.$$

Similarly, F and G are *jointly strongly monotone* if there exists $\rho > 0$ such that

$$\langle F(\zeta) - F(\xi), G(\zeta) - G(\xi) \rangle \geq \rho \|\zeta - \xi\|^2 \quad \forall \zeta,\ \xi \in \mathbb{R}^n.$$

In the case where $G(\zeta) = \zeta$ for all $\zeta \in \mathbb{R}^n$, the above notions are equivalent to the well-known notions of F being, respectively, monotone and strongly monotone (Facchinei and Pang 2003, Sect. 2.3).

3 Two classes of merit functions

In this section, we study two classes of merit functions for the SOCCP. We are motivated by a class of merit functions proposed by Luo and Tseng (1997) for the NCP case originally and was already extended to the SDCP by Tseng (1998). We introduce them as below. Let $f_{\rm LT}$ be given as (12), i.e.,

$$f_{\rm LT}(\zeta) := \psi_0(\langle F(\zeta), G(\zeta) \rangle) + \psi(F(\zeta), G(\zeta)),$$

where ψ_0 satisfies (13) and ψ satisfies (14). We notice that ψ_0 is differentiable and strictly increasing on $[0, \infty)$. An example of ψ_0 is $\psi_0(t) = \frac{1}{4}(\max\{0, t\})^4$. Let Ψ_+ (we adopt the notation used as in Tseng 1998) denote the collection of $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ satisfying (14) that are differentiable and satisfy the following conditions:

$$\begin{cases} \langle \nabla_x \psi(x, y), \nabla_y \psi(x, y) \rangle \geq 0, & \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n. \\ \langle x, \nabla_x \psi(x, y) \rangle + \langle y, \nabla_y \psi(x, y) \rangle \geq 0 & \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n. \end{cases} \tag{21}$$

We will give an example of ψ belonging to Ψ_+ in Proposition 3.1. Before that, we need couple technical lemmas which will be used for proving Propositions 3.1 and 3.2.

Lemma 3.1 (a) *For any* $x \in \mathbb{R}^n$, $\langle x, (x)_- \rangle = \|(x)_-\|^2$ *and* $\langle x, (x)_+ \rangle = \|(x)_+\|^2$.
(b) *For any* $x \in \mathbb{R}^n$ *and* $y \in \mathbb{R}^n$, *we have*

$$x \in \mathcal{K}^n \iff \langle x, y \rangle \geq 0 \quad \forall y \in \mathcal{K}^n. \tag{22}$$

Proof (a) By definition of trace, we know that $\mathrm{tr}(x \circ y) = 2\langle x, y \rangle$. Thus,

$$\begin{aligned} \langle x, (x)_- \rangle &= \frac{1}{2}\mathrm{tr}\Big(x \circ (x)_-\Big) \\ &= \frac{1}{2}\mathrm{tr}\Big([(x)_+ + (x)_-] \circ (x)_-\Big) \\ &= \frac{1}{2}\mathrm{tr}\Big((x)_-^2\Big) \\ &= \|(x)_-\|^2, \end{aligned}$$

where the last inequality is from definition of trace again. Similar arguments applied for $\langle x, (x)_+ \rangle = \|(x)_+\|^2$.

(b) Since $\mathcal{K}^n$ is self-dual, that is $\mathcal{K}^n = (\mathcal{K}^n)^*$. Hence, the desired result follows. □

Lemma 3.2 [Fukushima et al. 2002, Proposition 3.4] *For any* $x, y \in \mathbb{R}^n$ *and* $w \in \mathcal{K}^n$, *we have*

$$w^2 \succeq x^2 + y^2 \Longrightarrow L_w^2 \succeq L_x^2 + L_y^2,$$
$$w^2 \succeq x^2 \Longrightarrow w \succeq x.$$

Proposition 3.1 *Let* $\psi_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ *be given by*

$$\psi_1(x, y) := \frac{1}{2}\left(\|(-x)_+\|^2 + \|(-y)_+\|^2\right). \tag{23}$$

Then, the following results hold.

(a) ψ_1 *satisfies* (14).
(b) ψ_1 *is convex and differentiable at every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ *with* $\nabla_x \psi_1(x, y) = (x)_-$ *and* $\nabla_y \psi_1(x, y) = (y)_-$.
(c) *For every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, *we have*

$$\langle \nabla_x \psi_1(x, y), \nabla_y \psi_1(x, y)\rangle \geq 0.$$

(d) *For every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, *we have*

$$\langle x, \nabla_x \psi_1(x, y)\rangle + \langle y, \nabla_y \psi_1(x, y)\rangle = \|(x)_-\|^2 + \|(y)_-\|^2.$$

(e) ψ_1 *belongs to* Ψ_+.

Proof (a) Suppose $\psi_1(x, y) = 0$ and $\langle x, y\rangle \leq 0$. Then by definition of ψ_1 as (23), we have $(-x)_+ = 0$, $(-y)_+ = 0$ which implies $x \in \mathcal{K}^n, y \in \mathcal{K}^n$. Since $\mathcal{K}^n$ is self-dual, $x, y \in \mathcal{K}^n$ leads to $\langle x, y\rangle \geq 0$ by (22). This together with $\langle x, y\rangle \leq 0$ yields $\langle x, y\rangle = 0$. The other direction is clear from the above arguments. Hence, we proved that ψ_1 satisfies (14).

(b) For any $x \in \mathbb{R}^n$, we have the decomposition $x = (x)_+ + (x)_- = (x)_+ - (-x)_+$. Hence,

$$\frac{1}{2}\|(-x)_+\|^2 = \frac{1}{2}\|(x)_+ - x\|^2 = \min_{w \in \mathcal{K}^n} \frac{1}{2}\|w - x\|^2,$$

which is convex and differentiable in x (see Rockafellar 1970; page 255). Moreover, the chain rule gives

$$\nabla_x \left[\frac{1}{2}\|(-x)_+\|^2\right] = -(-x)_+ = (x)_-.$$

Similar formula holds for y. Thus, ψ_1 is convex and differentiable at every $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ with $\nabla_x \psi_1(x, y) = -(-x)_+ = (x)_-$ and $\nabla_y \psi_1(x, y) = -(-y)_+ = (y)_-$.

(c) From part(b), we have

$$\langle \nabla_x \psi_1(x, y), \nabla_y \psi_1(x, y)\rangle = \langle (x)_-, (y)_-\rangle = \langle (-x)_+, (-y)_+\rangle \geq 0,$$

where the inequality is true by (22).

(d) By applying Lemma 3.1(a), we obtain

$$\langle x, \nabla_x \psi_1(x, y)\rangle = \langle x, (x)_-\rangle = \|(x)_-\|^2.$$

Similarly, $\langle y, \nabla_x \psi_1(x, y)\rangle = \|(y)_-\|^2$ and hence the desired result holds.

(e) This is an immediate consequence of (a) through (d). □

Next, we consider a further restriction on ψ. Let Ψ_{++} denote the collection of $\psi \in \Psi_+$ satisfying the following conditions:

$$\psi(x, y) = 0 \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \quad \text{whenever } \langle \nabla_x \psi(x, y), \nabla_y \psi(x, y)\rangle = 0. \tag{24}$$

We notice that the ψ_1 defined as (23) in Proposition 3.1 does not belong to Ψ_{++}. An example of such ψ belonging to Ψ_{++} is given in Proposition 3.2.

Proposition 3.2 *Let* $\psi_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ *be given by*

$$\psi_2(x, y) := \frac{1}{2}\|\phi_{\text{FB}}(x, y)_+\|^2, \tag{25}$$

where ϕ_{FB} *is defined as* (10). *Then, the following results hold.*

(a) ψ_2 *satisfies* (14).

(b) ψ_2 *is differentiable at every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ *Moreover,* $\nabla_x \psi_2(0, 0) = \nabla_y \psi_2(0, 0) = 0$. *If* $(x, y) \neq (0, 0)$ *and* $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$*, then*

$$\begin{aligned} \nabla_x \psi_2(x, y) &= \left(L_x L^{-1}_{(x^2+y^2)^{1/2}} - I\right)\phi_{\text{FB}}(x, y)_+, \\ \nabla_y \psi_2(x, y) &= \left(L_y L^{-1}_{(x^2+y^2)^{1/2}} - I\right)\phi_{\text{FB}}(x, y)_+. \end{aligned} \tag{26}$$

If $(x, y) \neq (0, 0)$ *and* $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$*, then* $x_1^2 + y_1^2 \neq 0$ *and*

$$\begin{aligned} \nabla_x \psi_2(x, y) &= \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1\right)\phi_{\text{FB}}(x, y)_+, \\ \nabla_y \psi_2(x, y) &= \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1\right)\phi_{\text{FB}}(x, y)_+. \end{aligned} \tag{27}$$

(c) *For every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$*, we have*

$$\langle \nabla_x \psi_2(x, y), \nabla_y \psi_2(x, y)\rangle \geq 0,$$

and the equality holds whenever $\psi_2(x, y) = 0$.

(d) *For every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$*, we have*

$$\langle x, \nabla_x \psi_2(x, y)\rangle + \langle y, \nabla_y \psi_2(x, y)\rangle = \|\phi_{\text{FB}}(x, y)_+\|^2.$$

(e) ψ_2 *belongs to* Ψ_{++}.

Proof (a) Suppose $\psi_2(x, y) = 0$ and $\langle x, y\rangle \le 0$. Let $z := -\phi_{\rm FB}(x, y)$. Then $(-z)_+ = \phi_{\rm FB}(x, y)_+ = 0$ which says $z \in \mathcal{K}^n$. Since $x + y = (x^2 + y^2)^{1/2} + z$, squaring both sides and simplifying yield

$$2(x \circ y) = 2\left((x^2 + y^2)^{1/2} \circ z\right) + z^2.$$

Now, taking trace of both sides and using the fact $\mathrm{tr}(x \circ y) = 2\langle x, y\rangle$, we obtain

$$4\langle x, y\rangle = 4\langle (x^2 + y^2)^{1/2}, z\rangle + 2\|z\|^2. \tag{28}$$

Since $(x^2 + y^2)^{1/2} \in \mathcal{K}^n$ and $z \in \mathcal{K}^n$, then we know $\langle (x^2 + y^2)^{1/2}, z\rangle \ge 0$ by Lemma 3.1(b). Thus, the right hand-side of (28) is nonnegative, which togethers with $\langle x, y\rangle \le 0$ implies $\langle x, y\rangle = 0$. Therefore, with this, the equation (28) says $z = 0$ which is equivalent to $\phi_{\rm FB}(x, y) = 0$. Then by Lemma 2.1, we have $x, y \in \mathcal{K}^n$. Conversely, if $x, y \in \mathcal{K}^n$ and $\langle x, y\rangle = 0$, then again Lemma 2.1 yields $\phi_{\rm FB}(x, y) = 0$. Thus, $\psi_2(x, y) = 0$ and $\langle x, y\rangle \le 0$.

(b) For the proof of part(b), we need to discuss three cases.

Case 1: If $(x, y) = (0, 0)$, then for any $h, k \in \mathbb{R}^n$, let $\mu_1 \le \mu_2$ be the spectral values and let $v^{(1)}, v^{(2)}$ be the corresponding spectral vectors of $h^2 + k^2$. Hence, by Property 2.1(b),

$$\begin{aligned}\|(h^2 + k^2)^{1/2} - h - k\| &= \|\sqrt{\mu_1}v^{(1)} + \sqrt{\mu_2}v^{(2)} - h - k\| \\ &\le \sqrt{\mu_1}\|v^{(1)}\| + \sqrt{\mu_2}\|v^{(2)}\| + \|h\| + \|k\| \\ &= (\sqrt{\mu_1} + \sqrt{\mu_2})/\sqrt{2} + \|h\| + \|k\|.\end{aligned}$$

Also

$$\begin{aligned}\mu_1 \le \mu_2 &= \|h\|^2 + \|k\|^2 + 2\|h_1h_2 + k_1k_2\| \\ &\le \|h\|^2 + \|k\|^2 + 2|h_1|\|h_2\| + 2|k_1|\|k_2\| \\ &\le 2\|h\|^2 + 2\|k\|^2.\end{aligned}$$

Combining the above two inequalities yields

$$\begin{aligned}\psi_2(h, k) - \psi_2(0, 0) &= \frac{1}{2}\|\phi_{\rm FB}(h, k)_+\|^2 \\ &\le \|\phi_{\rm FB}(h, k)\|^2 \\ &= \|(h^2 + k^2)^{1/2} - h - k\|^2 \\ &\le \left((\sqrt{\mu_1} + \sqrt{\mu_2})/\sqrt{2} + \|h\| + \|k\|\right)^2 \\ &\le \left(2\sqrt{2\|h\|^2 + 2\|k\|^2}/\sqrt{2} + \|h\| + \|k\|\right)^2 \\ &= O(\|h\|^2 + \|k\|^2),\end{aligned}$$

where the first inequality is from Lemma 2.5. This shows that ψ_2 is differentiable at $(0, 0)$ with

$$\nabla_x\psi_2(0, 0) = \nabla_y\psi_2(0, 0) = 0.$$

Case 2: If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, let z be factored as $z = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}$ for any $z \in \mathbb{R}^n$. Now, let $g : \mathbb{R}^n \to \mathbb{R}^n$ be defined as

$$g(z) := \frac{1}{2}((z)_+)^2 = \hat{g}(\lambda_1)u^{(1)} + \hat{g}(\lambda_2)u^{(2)},$$

where $\hat{g} : \mathbb{R} \to \mathbb{R}$ is given by $\hat{g}(\lambda) := \frac{1}{2}(\max(0, \lambda))^2$. From the continuous differentiability of $\hat{g}$ and Proposition 5.2 of Chen et al. (2004), the vector-valued function g is also continuously differentiable. Hence, the first component $g_1(z) = \frac{1}{2}\|(z)_+\|^2$ of $g(z)$ is continuously differentiable as well. By an easy computation, we have $\nabla g_1(z) = (z)_+$. Since $\psi_2(x, y) = g_1(\phi_{\text{FB}}(x, y))$ and ϕ_{FB} is differentiable at $(x, y) \neq (0, 0)$ with $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$ (see Fukushima et al. 2002, Corrollary 5.2). Hence, the chain rule yields

$$\nabla_x \psi_2(x, y) = \nabla_x \phi_{\text{FB}}(x, y) \nabla g_1(\phi_{\text{FB}}(x, y)) = \left(L_x L^{-1}_{(x^2+y^2)^{1/2}} - I \right) \phi_{\text{FB}}(x, y)_+,$$

$$\nabla_y \psi_2(x, y) = \nabla_y \phi_{\text{FB}}(x, y) \nabla g_1(\phi_{\text{FB}}(x, y)) = \left(L_y L^{-1}_{(x^2+y^2)^{1/2}} - I \right) \phi_{\text{FB}}(x, y)_+.$$

Case 3: If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, by direct computation, we know $\|x\|^2 + \|y\|^2 = 2\|x_1 x_2 + y_1 y_2\|$ under this case. Since $(x, y) \neq (0, 0)$, this also implies $x_1 x_2 + y_1 y_2 \neq 0$. We notice that we can not apply the chain rule as in case 2 since ϕ_{FB} is no longer differentiable at such (x, y) of case 3. By the spectral factorization, we observe that

$$\begin{aligned} \phi_{\text{FB}}(x, y)_+ = \phi_{\text{FB}}(x, y) &\Longleftrightarrow \phi_{\text{FB}}(x, y) \in \mathcal{K}^n \\ \phi_{\text{FB}}(x, y)_+ = 0 &\Longleftrightarrow \phi_{\text{FB}}(x, y) \in -\mathcal{K}^n \\ \phi_{\text{FB}}(x, y)_+ = \lambda_2 u^{(2)} &\Longleftrightarrow \phi_{\text{FB}}(x, y) \notin \mathcal{K}^n \cup -\mathcal{K}^n, \end{aligned} \tag{29}$$

where λ_2 is the bigger spectral value of $\phi_{\text{FB}}(x, y)$ and $u^{(2)}$ is the corresponding spectral vector. Indeed, by applying Lemma 2.2, under this case, we have (as in Chen and Tseng 2005, Eq. (26))

$$\phi_{\text{FB}}(x, y) = \left(\sqrt{x_1^2 + y_1^2} - (x_1 + y_1), \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2}} - (x_2 + y_2) \right). \tag{30}$$

Therefore, λ_2 and $u^{(2)}$ are given as below:

$$\lambda_2 = \sqrt{x_1^2 + y_1^2} - (x_1 + y_1) + \|w_2\|, \tag{31}$$

$$u^{(2)} = \frac{1}{2}\left(1, \frac{w_2}{\|w_2\|}\right), \tag{32}$$

where $w_2 = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2}} - (x_2 + y_2)$. To prove the differentiability of ψ_2 under this case, we shall discuss the following three subcases according to the above observation (29).

(i) If $\phi_{\rm FB}(x,y) \notin \mathcal{K}^n \cup -\mathcal{K}^n$ then $\phi_{\rm FB}(x,y)_+ = \lambda_2 u^{(2)}$ where λ_2 and $u^{(2)}$ are given as in (31). From the fact that $\|u^{(2)}\| = \frac{1}{\sqrt{2}}$, we obtain

$$\begin{aligned}\psi_2(x,y) &= \frac{1}{2}\|\phi_{\rm FB}(x,y)_+\|^2 = \frac{1}{4}\lambda_2^2 \\ &= \frac{1}{4}\Bigg[\left(\sqrt{x_1^2+y_1^2} - (x_1+y_1)\right)^2 \\ &\quad + 2\left(\sqrt{x_1^2+y_1^2} - (x_1+y_1)\right)\cdot \|w_2\| + \|w_2\|^2\Bigg].\end{aligned}$$

Since $(x,y) \neq (0,0)$ in this case, ψ_2 is differentiable clearly. Moreover, using the product rule and chain rule for differentiation, the derivative of ψ_2 with respect to x_1 works out to be

$$\begin{aligned}\frac{\partial}{\partial x_1}\psi_2(x,y) &= \frac{1}{4}\Bigg[2\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)\right)\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right) \\ &\quad + 2\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\|w_2\| \\ &\quad + 2\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)\right)\cdot\frac{w_2^{\rm T}\nabla_{x_1}w_2}{\|w_2\|} + 2w_2^{\rm T}\nabla_{x_1}w_2\Bigg] \\ &= \frac{1}{2}\Bigg[\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)+\|w_2\|\right)\Bigg].\end{aligned}$$

The last equality of the above expression is true because of

$$\begin{aligned}\nabla_{x_1}w_2 &= \frac{x_2\cdot\sqrt{x_1^2+y_1^2} - (x_1x_2+y_1y_2)\cdot\frac{x_1}{\sqrt{x_1^2+y_1^2}}}{(x_1^2+y_1^2)} \\ &= \frac{\frac{1}{\sqrt{x_1^2+y_1^2}}\left[x_2(x_1^2+y_1^2)-(x_1^2x_2+x_1y_1y_2)\right]}{(x_1^2+y_1^2)} \\ &= \frac{x_1^2x_2+y_1^2x_2-x_1^2x_2-x_1y_1y_2}{\left(\sqrt{x_1^2+y_1^2}\right)^3} \\ &= 0,\end{aligned}$$

where the last equality holds by Lemma 2.2. Similarly, the gradient of ψ_2 with respect to x_2 works out to be

$$\begin{aligned}\nabla_{x_2}\psi_2(x,y) &= \frac{1}{4}\left[2\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)\right)\frac{\nabla_{x_2}w_2\cdot w_2}{\|w_2\|}+2\nabla_{x_2}w_2\cdot w_2\right]\\ &= \frac{1}{2}\left[\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)\right)\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\frac{w_2}{\|w_2\|}\right.\\ &\qquad\left.+\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)w_2\right]\\ &= \frac{1}{2}\left[\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)+\|w_2\|\right)\frac{w_2}{\|w_2\|}\right].\end{aligned}$$

Then, we can rewrite $\nabla_x\psi_2(x,y)$ as

$$\begin{aligned}\nabla_x\psi_2(x,y) &= \begin{bmatrix}\frac{\partial}{\partial x_1}\psi_2(x,y)\\ \nabla_{x_2}\psi_2(x,y)\end{bmatrix}\\ &:= \begin{bmatrix}\Xi_1\\ \Xi_2\end{bmatrix}\\ &= \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\lambda_2 u^{(2)}\\ &= \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\phi_{\text{FB}}(x,y)_+, \qquad (33)\end{aligned}$$

where

$$\begin{aligned}\Xi_1 &:= \frac{1}{2}\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)+\|w_2\|\right)\in\mathbb{R}\\ \Xi_2 &:= \frac{1}{2}\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\left(\sqrt{x_1^2+y_1^2}-(x_1+y_1)+\|w_2\|\right)\frac{w_2}{\|w_2\|}\in\mathbb{R}^{n-1}.\end{aligned}$$

(ii) If $\phi_{\text{FB}}(x,y)\in\mathcal{K}^n$ then $\phi_{\text{FB}}(x,y)_+=\phi_{\text{FB}}(x,y)$ and hence $\psi_2(x,y)=\frac{1}{2}\|\phi_{\text{FB}}(x,y)_+\|^2=\frac{1}{2}\|\phi_{\text{FB}}(x,y)\|^2$. Thus, by Chen and Tseng (2005, Prop. 3.1(b)), we know that the gradient of ψ_2 under this subcase is as below:

$$\nabla_x \psi_2(x,y) = \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}} - 1\right)\phi_{\rm FB}(x,y) = \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}} - 1\right)\phi_{\rm FB}(x,y)_+$$

$$\nabla_y \psi_2(x,y) = \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}} - 1\right)\phi_{\rm FB}(x,y) = \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}} - 1\right)\phi_{\rm FB}(x,y)_+. \quad (34)$$

If there is (x', y') such that $\phi_{\rm FB}(x', y') \notin \mathcal{K}^n \cup -\mathcal{K}^n$ and $\phi_{\rm FB}(x', y') \to \phi_{\rm FB}(x, y) \in \mathcal{K}^n$ (the neighborhood of point belonging to this subcase). From (33) and (34), it can be seen that

$$\nabla_x \psi_2(x', y') \to \nabla_x \psi_2(x, y), \quad \nabla_y \psi_2(x', y') \to \nabla_y \psi_2(x, y).$$

Thus, ψ_2 is differentiable under this subcase.
(iii) If $\phi_{\rm FB}(x, y) \in -\mathcal{K}^n$ then $\phi_{\rm FB}(x, y)_+ = 0$. Thus, $\psi_2(x, y) = \frac{1}{2}\|\phi_{\rm FB}(x, y)_+\|^2 = 0$ and it is clear that its gradient under this subcase is

$$\nabla_x \psi_2(x,y) = 0 = \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}} - 1\right)\phi_{\rm FB}(x,y)_+,$$

$$\nabla_y \psi_2(x,y) = 0 = \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}} - 1\right)\phi_{\rm FB}(x,y)_+. \quad (35)$$

Again, if there is (x', y') such that $\phi_{\rm FB}(x', y') \notin \mathcal{K}^n \cup -\mathcal{K}^n$ and $\phi_{\rm FB}(x', y') \to \phi_{\rm FB}(x, y) \in -\mathcal{K}^n$ (the neighborhood of point belonging to this subcase). From (33) and (35), it can be seen that

$$\nabla_x \psi_2(x', y') \to 0 = \nabla_x \psi_2(x, y), \quad \nabla_y \psi_2(x', y') \to 0 = \nabla_y \psi_2(x, y).$$

Thus, ψ_2 is differentiable under this subcase.
From the above, we complete the proof of this case and therefore the proof for part(b) is done.
(c) We wish to show that $\langle \nabla_x \psi_2(x, y), \nabla_y \psi_2(x, y)\rangle \geq 0$ and the equality holds if and only if $\psi_2(x, y) = 0$. We follow the three cases as above.
Case 1: If $(x, y) = (0, 0)$, by part (b), we know $\nabla_x \psi_2(x, y) = \nabla_y \psi_2(x, y) = 0$. Therefore, the desired equality holds.
Case 2: If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, by part (b), we have

$$\begin{aligned}\langle \nabla_x \psi_2(x,y),\ \nabla_y \psi_2(x,y)\rangle &= \langle (L_x L_z^{-1} - I)(\phi_{\rm FB})_+,\ (L_y L_z^{-1} - I)(\phi_{\rm FB})_+\rangle \\ &= \langle (L_x - L_z)L_z^{-1}(\phi_{\rm FB})_+,\ (L_y - L_z)L_z^{-1}(\phi_{\rm FB})_+\rangle \\ &= \langle (L_y - L_z)(L_x - L_z)L_z^{-1}(\phi_{\rm FB})_+,\ L_z^{-1}(\phi_{\rm FB})_+\rangle.\end{aligned} \quad (36)$$

Let S be the symmetric part of $(L_y - L_z)(L_x - L_z)$. Then

$$\begin{aligned} S &= \frac{1}{2}\Big((L_y - L_z)(L_x - L_z) + (L_x - L_z)(L_y - L_z)\Big) \\ &= \frac{1}{2}\Big(L_x L_y + L_y L_x - L_z(L_x + L_y) - (L_x + L_y)L_z + 2L_z^2\Big) \\ &= \frac{1}{2}(L_z - L_x - L_y)^2 + \frac{1}{2}(L_z^2 - L_x^2 - L_y^2). \end{aligned}$$

Since $z \in \mathcal{K}^n$ and $z^2 = x^2 + y^2$, Lemma 3.2 implies $L_z^2 - L_x^2 - L_y^2 \succeq O$. Then (36) yields

$$\begin{aligned} &\langle \nabla_x \psi_2(x, y),\ \nabla_y \psi_2(x, y)\rangle \\ &\quad = \langle S L_z^{-1}(\phi_{\text{FB}})_+, L_z^{-1}(\phi_{\text{FB}})_+\rangle \\ &\quad = \frac{1}{2}\langle (L_z - L_x - L_y)^2 L_z^{-1}(\phi_{\text{FB}})_+, L_z^{-1}(\phi_{\text{FB}})_+\rangle \\ &\qquad + \frac{1}{2}\langle (L_z^2 - L_x^2 - L_y^2) L_z^{-1}(\phi_{\text{FB}})_+, L_z^{-1}(\phi_{\text{FB}})_+\rangle \\ &\quad \geq \frac{1}{2}\langle (L_z - L_x - L_y)^2 L_z^{-1}(\phi_{\text{FB}})_+, L_z^{-1}(\phi_{\text{FB}})_+\rangle \\ &\quad = \frac{1}{2}\|L_{\phi_{\text{FB}}} L_z^{-1}(\phi_{\text{FB}})_+\|^2, \end{aligned}$$

where the last equality uses $L_z - L_x - L_y = L_{z-x-y} = L_{\phi_{\text{FB}}}$. If the equality holds, then the above relation yields $\|L_{\phi_{\text{FB}}} L_z^{-1}(\phi_{\text{FB}})_+\|^2 = 0$ and, by Property 2.1(d),

$$L_{\phi_{\text{FB}}} L_z^{-1}(\phi_{\text{FB}})_+ = \phi_{\text{FB}} \circ (L_z^{-1}(\phi_{\text{FB}})_+) = L_z^{-1}(\phi_{\text{FB}})_+ \circ \phi_{\text{FB}} = 0.$$

Since $z = (x^2 + y^2)^{1/2} \in \text{int}(\mathcal{K}^n)$ so that $L_z^{-1} \succ O$ (see Property 2.1(d)), multiplying L_z^{-1} both sides gives $\phi_{\text{FB}} \circ (\phi_{\text{FB}})_+ = 0$. From definition of Jordan product (11) and Lemma 3.1(a), it implies $(\phi_{\text{FB}})_+ = 0$; and hence $\psi_2 = 0$. Conversely, if $(\phi_{\text{FB}})_+ = 0$, then it is clear that $\langle \nabla_x \psi_2(x, y), \nabla_y \psi_2(x, y)\rangle = 0$.
Case 3: If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, by part (b), we have

$$\begin{aligned} &\langle \nabla_x \psi_2(x, y),\ \nabla_y \psi_2(x, y)\rangle \\ &= \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1\right)\left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1\right)\|\phi_{\text{FB}}(x, y)_+\|^2 \ \geq\ 0. \end{aligned}$$

If the equality holds, then either $\phi_{\text{FB}}(x, y)_+ = 0$ or $\frac{x_1}{\sqrt{x_1^2+y_1^2}} = 1$ or $\frac{y_1}{\sqrt{x_1^2+y_1^2}} = 1$. In the second case, we have $y_1 = 0$ and $x_1 \geq 0$, so that Lemma 2.2 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 2.2 yields $x_2 = 0$ and $y_1 = \|y_2\|$. Thus, in these two cases, we have $x \circ y = 0,\ x \in \mathcal{K}^n,\ y \in \mathcal{K}^n$. Then, by (14), $\psi_2(x, y) = 0$.

(d) Again, we need to discuss the three cases as below.

Case 1: If $(x, y) = (0, 0)$, by part (b), we know $\nabla_x \psi_2(x, y) = \nabla_y \psi_2(x, y) = 0$. Therefore, the desired equality holds.
Case 2: If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$, by part (b), we have

$$\nabla_x \psi_2(x, y) = \left(L_x L_z^{-1} - I \right) \phi_{\mathrm{FB}}(x, y)_+,$$
$$\nabla_y \psi_2(x, y) = \left(L_y L_z^{-1} - I \right) \phi_{\mathrm{FB}}(x, y)_+,$$

where we let $z := (x^2 + y^2)^{1/2}$. For simplicity, we will write $\phi(x, y)_+$ as ϕ_+. Thus,

$$\begin{aligned}
&\langle x, \nabla_x \psi_2(x, y) \rangle + \langle y, \nabla_y \psi_2(x, y) \rangle \\
&\quad = \langle x, (L_x L_z^{-1} - I)(\phi_{\mathrm{FB}})_+ \rangle + \langle y, (L_y L_z^{-1} - I)(\phi_{\mathrm{FB}})_+ \rangle \\
&\quad = \langle (L_z^{-1} L_x - I)x, (\phi_{\mathrm{FB}})_+ \rangle + \langle (L_z^{-1} L_y - I)y, (\phi_{\mathrm{FB}})_+ \rangle \\
&\quad = \langle L_z^{-1} L_x x + L_z^{-1} L_y y - x - y, (\phi_{\mathrm{FB}})_+ \rangle \\
&\quad = \langle L_z^{-1}(x^2 + y^2) - x - y, (\phi_{\mathrm{FB}})_+ \rangle \\
&\quad = \langle L_z^{-1} z^2 - x - y, (\phi_{\mathrm{FB}})_+ \rangle \\
&\quad = \langle z - x - y, (\phi_{\mathrm{FB}})_+ \rangle \\
&\quad = \| (\phi_{\mathrm{FB}})_+ \|^2,
\end{aligned}$$

where the next-to-last equality follows from $L_z z = z^2$, so that $L_z^{-1} z^2 = z$ and the last equality is from Lemma 3.1(a).
Case 3: If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \notin \mathrm{int}(\mathcal{K}^n)$, by part(b), we have

$$\nabla_x \psi_2(x, y) = \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\mathrm{FB}}(x, y)_+,$$
$$\nabla_y \psi_2(x, y) = \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\mathrm{FB}}(x, y)_+.$$

Thus,

$$\begin{aligned}
&\langle x, \nabla_x \psi_2(x, y) \rangle + \langle y, \nabla_y \psi_2(x, y) \rangle \\
&= \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \langle x, (\phi_{\mathrm{FB}})_+ \rangle + \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \langle y, (\phi_{\mathrm{FB}})_+ \rangle \\
&= \left\langle \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) x + \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) y, (\phi_{\mathrm{FB}})_+ \right\rangle \\
&= \left\langle \frac{x_1 x + y_1 y}{\sqrt{x_1^2 + y_1^2}} - x - y, (\phi_{\mathrm{FB}})_+ \right\rangle
\end{aligned}$$

$$= \langle \phi_{\rm FB}, (\phi_{\rm FB})_+ \rangle$$
$$= \|(\phi_{\rm FB})_+\|^2,$$

where the next-to-last equality uses (30) and the last equality is from Lemma 3.1(a) again.

(e) This is an immediate consequence of (a) through (d). □

We notice that (26) can be rewritten as

$$\nabla_x \psi_2(x, y) = L_z^{-1}\Big[[z - x - y]_+\Big] \circ (x - z),$$
$$\nabla_y \psi_2(x, y) = L_z^{-1}\Big[[z - x - y]_+\Big] \circ (y - z),$$

where $z = (x^2 + y^2)^{1/2}$. This is a similar form as in Tseng (1998, Lemma 7.2). Nonetheless, (27) can not be rewritten as the above form since L_z^{-1} does not exist whenever $x^2 + y^2$ is on the boundary of $\mathcal{K}^n$. The next proposition is a result which is an extension of (Tseng 1998, Proposition 7.1) for SDCP to the case of SOCCP. Though the ideas for arguments are similar, we present the proof for completion.

Proposition 3.3 *Let* $f_{\rm LT} : \mathbb{R}^n \to \mathbb{R}_+$ *be given as* (12) *with* ψ_0 *satisfying* (13) *and* ψ *satisfying* (14). *Then, the following results hold.*

(a) *For all* $\zeta \in \mathbb{R}^n$, *we have* $f_{\rm LT}(\zeta) \geq 0$ *and* $f_{\rm LT}(\zeta) = 0$ *if and only if* ζ *solves the SOCCP.*

(b) *If* ψ_0, ψ *and* F, G *are differentiable, then so is* $f_{\rm LT}$ *and*

$$\nabla f_{\rm LT}(\zeta) = \psi_0'(\langle F(\zeta), G(\zeta)\rangle)\Big[\nabla F(\zeta)G(\zeta) + \nabla G(\zeta)F(\zeta)\Big]$$
$$+\nabla F(\zeta)\nabla_x\psi(F(\zeta), G(\zeta))$$
$$+\nabla G(\zeta)\nabla_y\psi(F(\zeta), G(\zeta)).$$

(c) *Assume* F, G *are differentiable on* $\mathbb{R}^n$ *and* ψ *belongs to* Ψ_+ (*respectively,* Ψ_{++}). *Then, for every* $\zeta \in \mathbb{R}^n$ *where* $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ *is positive definite* (*respectively, positive semi-definite*), *either* (i) $f_{\rm LT}(\zeta) = 0$ *or* (ii) $\nabla f_{\rm LT}(\zeta) \neq 0$ *with* $\langle d(\zeta), \nabla f_{\rm LT}(\zeta)\rangle < 0$, *where*

$$d(\zeta) := -(\nabla G(\zeta)^{-1})^{\rm T}\Big[\psi_0'(\langle F(\zeta), G(\zeta)\rangle)G(\zeta) + \nabla_x\psi(F(\zeta), G(\zeta))\Big].$$

Proof (a) This consequence follows from (12) and (13), (14).

(b) By direct computation and chain rule, the result follows.

(c) First, we consider the case of $\psi \in \Psi_{++}$ and fix $\zeta \in \mathbb{R}^n$ where $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ is positive semi-definite. Let $\alpha := \psi_0'(\langle F(\zeta), G(\zeta)\rangle)$ and drop the argument "(ζ)" for simplicity. Then

$$\langle d, \nabla f_{\rm LT}\rangle$$
$$= \langle -(\nabla G^{-1})^{\rm T}(\alpha G + \nabla_x\psi(F, G)), \nabla F(\alpha G + \nabla_x\psi(F, G))$$
$$+\nabla G(\alpha F + \nabla_y\psi(F, G))\rangle$$

$$
\begin{aligned}
&= -\langle \alpha G + \nabla_x \psi(F,G), \nabla G^{-1} \nabla F(\alpha G + \nabla_x \psi(F,G)) \rangle \\
&\quad - \langle \alpha G + \nabla_x \psi(F,G), \alpha F + \nabla_y \psi(F,G) \rangle \\
&\leq -\langle \alpha G + \nabla_x \psi(F,G), \alpha F + \nabla_y \psi(F,G) \rangle \\
&= -\alpha^2 \langle F, G \rangle - \alpha \big(\langle F, \nabla_x \psi(F,G) \rangle + \langle G, \nabla_y \psi(F,G) \rangle \big) \\
&\quad - \langle \nabla_x \psi(F,G), \nabla_y \psi(F,G) \rangle \\
&= -\alpha^2 \langle F, G \rangle - \langle \nabla_x \psi(F,G), \nabla_y \psi(F,G) \rangle,
\end{aligned}
$$

where the first inequality holds since $\nabla G^{-1} \nabla F$ is positive semi-definite and the inequality follows from $\alpha \geq 0$ and equation (21). Now, we observe that $t\psi_0'(t) > 0$ if and only if $t > 0$ since ψ_0 is strictly increasing on $[0, \infty)$. Therefore, the first term on the right-hand side is non-positive and equals zero if $\langle F, G \rangle \leq 0$. In addition, by equations (21) and (24), the second term on the right-hand side is non-positive and equals zero only if $\psi(F,G) = 0$. Thus, we have $\langle d(\zeta), \nabla f_{\mathrm{LT}}(\zeta) \rangle \leq 0$ and the equality holds only when $\langle F(\zeta), G(\zeta) \rangle \leq 0$ and $\psi(F(\zeta), G(\zeta)) = 0$, in which equation (14) implies ζ satisfies (1)–(2), i.e., $f_{\mathrm{LT}}(\zeta) = 0$.

Similar arguments can be applied for the case of $\psi \in \Psi_+$ and $\nabla G(\zeta)^{-1} \nabla F(\zeta)$ being positive definite. □

Next, we further consider another class of merit functions by modifying f_{LT} a bit where ψ_0 is replaced by $\psi_0^* : \mathbb{R}^n \to \mathbb{R}_+$ given as (16), i.e., $\psi_0^*(w) = \frac{1}{2}\|(w)_+\|^2$. It is known that the function ψ_0^* given in (16) is continuously differentiable (see Rockafellar 1970, p. 255) with $\nabla \psi_0^*(w) = [w]_+$ (by the chain rule). In other words, we will study $\widehat{f_{\mathrm{LT}}} : \mathbb{R}^n \to \mathbb{R}_+$ defined as (15), (16):

$$
\widehat{f_{\mathrm{LT}}}(\zeta) := \psi_0^*(F(\zeta) \circ G(\zeta)) + \psi(F(\zeta), G(\zeta)),
$$

where ψ_0^* is given as (16) and ψ satisfies (14). By imitating the steps for proving Proposition 3.3 and using Lemma 3.3 as below, we obtain Proposition 3.4 which is a result analogous to Proposition 3.3. We omit its proof.

Lemma 3.3 *The function* $\psi_0^*(x \circ y) := \frac{1}{2}\|(x \circ y)_+\|^2$ *is differentiable for all* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. *Moreover,*

$$
\begin{aligned}
\nabla_x \psi_0^*(x \circ y) &= L_y \cdot (x \circ y)_+ \\
\nabla_y \psi_0^*(x \circ y) &= L_x \cdot (x \circ y)_+
\end{aligned}
$$

Proof This is result of Chen (2006, Lemma 3.1). □

Proposition 3.4 *Let* $\widehat{f_{\mathrm{LT}}} : \mathbb{R}^n \to \mathbb{R}_+$ *be given as* (15), (16). *Then, the following results hold.*

(a) *For all* $x \in \mathbb{R}^n$, *we have* $\widehat{f_{\mathrm{LT}}}(\zeta) \geq 0$ *and* $\widehat{f_{\mathrm{LT}}}(\zeta) = 0$ *if and only if* ζ *solves the SOCCP.*
(b) *If* ψ_0^*, ψ *and* F, G *are differentiable, then so is* $\widehat{f_{\mathrm{LT}}}$ *and*

$$
\begin{aligned}
\nabla \widehat{f_{\mathrm{LT}}}(\zeta) = &\Big[\nabla F(\zeta) L_{G(\zeta)} + \nabla G(\zeta) L_{F(\zeta)} \Big] (F(\zeta) \circ G(\zeta))_+ \\
&+ \nabla F(\zeta) \nabla_x \psi(F(\zeta), G(\zeta)) \\
&+ \nabla G(\zeta) \nabla_y \psi(F(\zeta), G(\zeta)).
\end{aligned}
$$

We originally thought there should have parallel results to Proposition 3.3(c) for $\widehat{f_{\rm LT}}$ and whose proofs are also similar. In other words, we wish to have the following:
Assume F, G are differentiable on $\mathbb{R}^n$ and ψ belongs to Ψ_+ (respectively, Ψ_{++}). Then, for every $\zeta \in \mathbb{R}^n$ where $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ is positive definite (respectively, positive semi-definite), either (i) $\widehat{f_{\rm LT}}(\zeta) = 0$ or (ii) $\nabla \widehat{f_{\rm LT}}(\zeta) \neq 0$ with $\langle d(\zeta), \nabla \widehat{f_{\rm LT}}(\zeta)\rangle < 0$, where

$$d(\zeta) := -(\nabla G(\zeta)^{-1})^{\rm T}\Big[L_{G(\zeta)} \cdot (F(\zeta) \circ G(\zeta))_+ + \nabla_x \psi(F(\zeta), G(\zeta))\Big].$$

However, we are not able to complete the arguments even though ψ_0^* is in relation to ψ_0 in certain sense. We thank a referee for pointing this out. We suspect that there needs more subtle properties of ψ_0^* to finish it.

4 Error bound and bounded level sets

The error bound is an important concept that indicates how close an arbitrary point is to the solution set of SOCCP. Thus, an error bound may be used to provide stopping criterion for an iterative method. As below, we establish propositions about the error bound properties of $f_{\rm LT}$, $\widehat{f_{\rm LT}}$ given as (12) and (15). We need some technical lemmas as below to prove the error bound properties.

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

$$\langle x, y\rangle \le \sqrt{2}\|(x \circ y)_+\|.$$

Proof See Chen (2006, Lemma 4.1). □

Lemma 4.2 *Let* ψ_1, ψ_2 *be given as* (23) *and* (25)*, respectively. Then,* ψ_1 *and* ψ_2 *satisfy the following inequality.*

$$\psi_i(x, y) \ge \alpha\Big(\|(-x)_+\|^2 + \|(-y)_+\|^2\Big) \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \tag{37}$$

for some positive constant α *and* $i = 1, 2$.

Proof For ψ_1, it is clear by definition (23) where $\alpha = \frac{1}{2}$. For ψ_2, the inequality is still true, where $\alpha = \frac{1}{4}$, due to Lemma 2.5. □

Lemma 4.3 *Let* ψ_0^* *be given as* (16)*. Then,* ψ_0^* *satisfies*

$$\psi_0^*(w) \ge \beta\|(w)_+\|^2 \quad \forall w \in \mathbb{R}^n, \tag{38}$$

for some positive constant β.

Proof It is clear by definition of ψ_0^* given as (16) where $\beta = \frac{1}{2}$. □

Proposition 4.1 *Let f_{LT} be given by* (12)–(14) *with ψ satisfying* (37). *Suppose that F and G are jointly strongly monotone mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ and SOCCP has a solution ζ^*. Then, there exists a scalar $\tau > 0$ such that*

$$\tau\|\zeta - \zeta^*\|^2 \leq \max\{0, \langle F(\zeta), G(\zeta)\rangle\} + \|(-F(\zeta))_+\| + \|(-G(\zeta))_+\| \quad \forall \zeta \in \mathbb{R}^n. \tag{39}$$

Moreover,

$$\tau\|\zeta - \zeta^*\|^2 \leq \psi_0^{-1}(f_{\mathrm{LT}}(\zeta)) + \frac{\sqrt{2}}{\sqrt{\alpha}} f_{\mathrm{LT}}(\zeta)^{1/2} \quad \forall \zeta \in \mathbb{R}^n, \tag{40}$$

where α is a positive constant .

Proof Since F and G are jointly strongly monotone, there exists a scalar $\rho > 0$ such that, for any $\zeta \in \mathbb{R}^n$,

$$\begin{aligned}
&\rho\|\zeta - \zeta^*\|^2 \\
&\quad \leq \langle F(\zeta) - F(\zeta^*),\ G(\zeta) - G(\zeta^*)\rangle \\
&\quad = \langle F(\zeta), G(\zeta)\rangle + \langle -F(\zeta), G(\zeta^*)\rangle + \langle F(\zeta^*), -G(\zeta)\rangle \\
&\quad \leq \max\{0, \langle F(\zeta), G(\zeta)\rangle\} + \langle (-F(\zeta))_+, G(\zeta^*)\rangle + \langle F(\zeta^*), (-G(\zeta))_+\rangle \\
&\quad \leq \max\{0, \langle F(\zeta), G(\zeta)\rangle\} + \|(-F(\zeta))_+\|\ \|G(\zeta^*)\| + \|F(\zeta^*)\|\ \|(-G(\zeta))_+\| \\
&\quad \leq \max\left\{1, \|F(\zeta^*)\|, \|G(\zeta^*)\|\right\} \\
&\qquad \times \left(\max\{0, \langle F(\zeta), G(\zeta)\rangle\} + \|(-F(\zeta))_+\| + \|(-G(\zeta))_+\|\right),
\end{aligned}$$

where the second inequality uses Lemma 2.4(b). Setting $\tau := \frac{\rho}{\max\{1, \|F(\zeta^*)\|, \|G(\zeta^*)\|\}}$ yields (39).

Notice that ψ_0^{-1} is well-defined by (13), and by using that ψ_0 is strictly increasing on $[0, \infty)$, we thus have

$$\max\{0, \langle F(\zeta), G(\zeta)\rangle\} \leq \psi_0^{-1}\left(f_{\mathrm{LT}}(\zeta)\right).$$

In addition, it is clear that

$$\psi(F(\zeta), G(\zeta)) \leq f_{\mathrm{LT}}(\zeta).$$

Now using Lemma 4.2 and the above inequality, we obtain

$$\begin{aligned}
\|(-F(\zeta))_+\| + \|(-G(\zeta))_+\| &\leq \sqrt{2}\left(\|(-F(\zeta))_+\|^2 + \|(-G(\zeta))_+\|^2\right)^{1/2} \\
&\leq \frac{\sqrt{2}}{\sqrt{\alpha}} \psi(F(\zeta), G(\zeta))^{1/2} \\
&\leq \frac{\sqrt{2}}{\sqrt{\alpha}} f_{\mathrm{LT}}(\zeta)^{1/2}.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\max\{0, \langle F(\zeta), G(\zeta)\rangle\} + \|(-F(\zeta))_+\| + \|(-G(\zeta))_+\| \\
&\quad \leq \psi_0^{-1}\left(f_{\mathrm{LT}}(\zeta)\right) + \frac{\sqrt{2}}{\sqrt{\alpha}} f_{\mathrm{LT}}(\zeta)^{1/2}.
\end{aligned}$$

This together with (39) yields (40). □

Proposition 4.2 *Let $\widehat{f_{\rm LT}}$ be given by* (15), (16) *with ψ satisfying* (37). *Suppose that F and G are jointly strongly monotone mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ and the SOCCP has a solution ζ^*. Then, there exists a scalar $\tau > 0$ such that*

$$\tau\|\zeta - \zeta^*\|^2 \le \|(F(\zeta)\circ G(\zeta))_+\| + \|(-F(\zeta))_+\| + \|(-G(\zeta))_+\| \quad \forall \zeta \in \mathbb{R}^n. \tag{41}$$

Moreover,

$$\tau\|\zeta - \zeta^*\|^2 \le \left(\frac{1}{\sqrt{\beta}} + \frac{\sqrt{2}}{\sqrt{\alpha}}\right)\widehat{f_{\rm LT}}(\zeta)^{1/2} \quad \forall \zeta \in \mathbb{R}^n, \tag{42}$$

where α and β are positive constants.

Proof Since F and G are jointly strongly monotone, there exists a scalar $\rho > 0$ such that, for any $\zeta \in \mathbb{R}^n$,

$$\begin{aligned}
&\rho\|\zeta - \zeta^*\|^2 \\
&\le \langle F(\zeta) - F(\zeta^*),\ G(\zeta) - G(\zeta^*)\rangle \\
&= \langle F(\zeta), G(\zeta)\rangle + \langle -F(\zeta), G(\zeta^*)\rangle + \langle F(\zeta^*), -G(\zeta)\rangle \\
&\le \langle F(\zeta), G(\zeta)\rangle + \langle (-F(\zeta))_+, G(\zeta^*)\rangle + \langle F(\zeta^*), (-G(\zeta))_+\rangle \\
&\le \langle F(\zeta), G(\zeta)\rangle + \|(-F(\zeta))_+\|\,\|G(\zeta^*)\| + \|F(\zeta^*)\|\,\|(-G(\zeta))_+\| \\
&\le \sqrt{2}\|(F(\zeta)\circ G(\zeta))_+\| + \|(-F(\zeta))_+\|\,\|G(\zeta^*)\| + \|F(\zeta^*)\|\,\|(-G(\zeta))_+\| \\
&\le \max\{\sqrt{2}, \|F(\zeta^*)\|, \|G(\zeta^*)\|\} \\
&\quad \times \left(\|(F(\zeta)\circ G(\zeta))_+\| + \|(-F(\zeta))_+\| + \|(-G(\zeta))_+\|\right),
\end{aligned}$$

where the second inequality uses Lemma 2.4(b) while the fourth inequality is from Lemma 4.1. Then, setting $\tau := \dfrac{\rho}{\max\{\sqrt{2}, \|F(\zeta^*)\|, \|G(\zeta^*)\|\}}$ yields (41).

Moreover, by Lemma 4.3, we have

$$\|(F(\zeta)\circ G(\zeta))_+\| \le \frac{1}{\sqrt{\beta}}\,\psi_0^*(F(\zeta)\circ G(\zeta))^{1/2} \le \frac{1}{\sqrt{\beta}}\,\widehat{f_{\rm LT}}(\zeta)^{1/2},$$

and (as in Proposition 4.1)

$$\begin{aligned}
\|(-F(\zeta))_+\| + \|(-G(\zeta))_+\| &\le \sqrt{2}\left(\|(-F(\zeta))_+\|^2 + \|(-G(\zeta))_+\|^2\right)^{1/2} \\
&\le \frac{\sqrt{2}}{\sqrt{\alpha}}\,\psi(F(\zeta), G(\zeta))^{1/2} \\
&\le \frac{\sqrt{2}}{\sqrt{\alpha}}\,\widehat{f_{\rm LT}}(\zeta)^{1/2},
\end{aligned}$$

where the second inequality is true by Lemma 4.2. Thus,

$$\|(F(\zeta)\circ G(\zeta))_+\| + \|(-F(\zeta))_+\| + \|(-G(\zeta))_+\| \le \left(\frac{1}{\sqrt{\beta}} + \frac{\sqrt{2}}{\sqrt{\alpha}}\right)\widehat{f_{\rm LT}}(\zeta)^{1/2}.$$

This together with (41) yields (42). □

Now, we give conditions under which $f_{\rm LT}$, $\widehat{f_{\rm LT}}$ has bounded level sets in Propositions 4.3 and 4.4, respectively. We need the next lemma which is key to proving the properties of bounded level sets.

Lemma 4.4 *Let ψ_1, ψ_2 be given by (23) and (25), respectively. For any $\{(x^k, y^k)\}_{k=1}^{\infty} \subseteq \mathbb{R}^n \times \mathbb{R}^n$, let $\lambda_1^k \le \lambda_2^k$ and $\mu_1^k \le \mu_2^k$ denote the spectral values of x^k and y^k, respectively. Then, the following results hold.*

(a) *If $\lambda_1^k \to -\infty$ or $\mu_1^k \to -\infty$, then $\psi_i(x^k, y^k) \to \infty$, for $i = 1, 2$.*
(b) *Suppose that $\{\lambda_1^k\}$ and $\{\mu_1^k\}$ are bounded below. If $\lambda_2^k \to \infty$ or $\mu_2^k \to \infty$, then $\langle x, x^k\rangle + \langle y, y^k\rangle \to \infty$ for any $x, y \in \mathrm{int}(\mathcal{K}^n)$.*

Proof (a) For ψ_1, the proof follows by the fact that

$$2\|(-x^k)_+\|^2 = \sum_{i=1}^{2}\left(\max\{0, -\lambda_i^k\}\right)^2$$

and similarly for $\|(-y^k)_+\|^2$; see Fukushima et al. (2002), Property 2.2 and Proposition 3.3.
For ψ_2, using the same fact plus Lemma 2.5 leads to the desired result.

(b) Fix any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}\times\mathbb{R}^{n-1}$ with $\|x_2\| < x_1$, $\|y_2\| < y_1$. Using the spectral decomposition

$$x^k = \left(\frac{\lambda_1^k + \lambda_2^k}{2}, \frac{\lambda_2^k - \lambda_1^k}{2} w_2^k\right) \quad \text{with } \|w_2^k\| = 1,$$

we have

$$\begin{aligned}\langle x, x^k\rangle &= \left(\frac{\lambda_1^k + \lambda_2^k}{2}\right)x_1 + \left(\frac{\lambda_2^k - \lambda_1^k}{2}\right)x_2^{\mathrm{T}} w_2^k \\ &= \frac{\lambda_1^k}{2}\left(x_1 - x_2^{\mathrm{T}} w_2^k\right) + \frac{\lambda_2^k}{2}\left(x_1 + x_2^{\mathrm{T}} w_2^k\right).\end{aligned} \tag{43}$$

Since $\|w_2^k\| = 1$, we have $x_1 - x_2^{\mathrm{T}} w_2^k \ge x_1 - \|x_2\| > 0$ and $x_1 + x_2^{\mathrm{T}} w_2^k \ge x_1 - \|x_2\| > 0$. Since $\{\lambda_1^k\}$ is bounded below, the first term on the right-hand side of (43) is bounded below. If $\{\lambda_2^k\} \to \infty$, then the second term on the right-hand side of (43) tends to infinity. Hence, $\langle x, x^k\rangle \to \infty$. A similar argument shows that $\langle y, y^k\rangle$ is bounded below. Thus, $\langle x, x^k\rangle + \langle y, y^k\rangle \to \infty$. If $\{\mu_2^k\} \to \infty$, the argument is symmetric to the one above. □

Proposition 4.3 *Let f_{LT} be given as* (12)–(14) *with ψ satisfying the condition of Lemma* 4.4(a). *Suppose that F, G are differentiable, jointly monotone mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying*

$$\lim_{\|\zeta\|\to\infty}\left(\|F(\zeta)\| + \|G(\zeta)\|\right) = \infty. \tag{44}$$

Suppose also that SOCCP is strictly feasible, i.e., there exists $\bar{\zeta} \in \mathbb{R}^n$ such that $F(\bar{\zeta}), G(\bar{\zeta}) \in \mathrm{int}(\mathcal{K}^n)$. Then, the level set

$$\mathcal{L}(\gamma) := \{\zeta \in \mathbb{R}^n \mid f_{\mathrm{LT}}(\zeta) \le \gamma\}$$

is bounded for all $\gamma \ge 0$.

Proof Suppose there exists an unbounded sequence $\{\zeta^k\} \subseteq \mathcal{L}(\gamma)$ for some $\gamma \geq 0$. It can be seen that the sequence of the smaller spectral values of $\{F(\zeta^k)\}$ and $\{G(\zeta^k)\}$ are bounded below. In fact, if not, it follows from Lemma 4.4(a) that $\psi(F(\zeta^k), G(\zeta^k)) \to \infty$. Thus, we have $f_{\text{LT}}(\zeta^k) \to \infty$, which contradicts $\{\zeta^k\} \subseteq \mathcal{L}(\gamma)$. Therefore, the unboundedness of $\{\zeta^k\}$ and (44) yield that the sequence of the bigger spectral values of $\{F(\zeta^k)\}$ or $\{G(\zeta^k)\}$ tends to infinity. Since F, G are jointly monotone, we have

$$\langle F(\zeta^k) - F(\bar{\zeta}), \quad G(\zeta^k) - G(\bar{\zeta})\rangle \geq 0,$$

which is equivalent to

$$\langle F(\zeta^k), G(\bar{\zeta})\rangle + \langle F(\bar{\zeta}), G(\zeta^k)\rangle \leq \langle F(\zeta^k), G(\zeta^k)\rangle + \langle F(\bar{\zeta}), G(\bar{\zeta})\rangle. \quad (45)$$

Then, by Lemma 4.4(b) and $F(\bar{\zeta}), G(\bar{\zeta}) \in \text{int}(\mathcal{K}^n)$, we obtain $\langle F(\zeta^k), G(\bar{\zeta})\rangle + \langle F(\bar{\zeta}), G(\zeta^k)\rangle \to \infty$, which together with (45) lead to $\langle F(\zeta^k), G(\zeta^k)\rangle \to \infty$. Thus, $f_{\text{LT}}(\zeta^k) \to \infty$. But, this contradicts $\{\zeta^k\} \subseteq \mathcal{L}(\gamma)$. Hence, we proved that $\mathcal{L}(\gamma)$ is bounded. □

Proposition 4.4 *Let* $\widehat{f_{\text{LT}}}$ *be given as* (15)-(16) *with* ψ *satisfying the condition of Lemma* 4.4(a). *Suppose that* F, G *are differentiable, jointly monotone mappings from* $\mathbb{R}^n$ *to* $\mathbb{R}^n$ *satisfying*

$$\lim_{\|\zeta\|\to\infty} \Big(\|F(\zeta)\| + \|G(\zeta)\|\Big) = \infty.$$

Suppose also that the SOCCP is strictly feasible, i.e., *there exists* $\bar{\zeta} \in \mathbb{R}^n$ *such that* $F(\bar{\zeta}), G(\bar{\zeta}) \in \text{int}(\mathcal{K}^n)$. *Then, the level set*

$$\mathcal{L}(\gamma) := \{\zeta \in \mathbb{R}^n \mid \widehat{f_{\text{LT}}}(\zeta) \leq \gamma\}$$

is bounded for all $\gamma \geq 0$.

Proof The arguments are similar to those in Proposition 4.3, so we omit the proof. □

5 Final remarks

In this paper, we have studied two classes of merit functions for the second-order cone complementarity problem. The first class is motivated by a class of merit functions for NCP Luo and Tseng (1997) and SDCP Tseng (1998), while the second class is based on a slight modification of the first one. We have also presented examples of merit functions which belong to the two classes we studied. Moreover, we have shown conditions under which the merit functions have properties of error bounds and bounded level sets. The related topics for future study are about the descent methods including numerical examples for solving the unconstrained minimization via these merit functions. On the other hand, recently there have been definitions of P-properties for nonlinear transformations on Euclidean Jordan Algebras (see Tao and Gowda 2004 for details), which are related to SOCCP due to the Jordan Algebra. In particular, there have been some special implications as below:

$$\text{strongly monotone} \Longrightarrow \text{uniform Jordan } P\text{-property} \Longrightarrow \\ \text{uniform } P\text{-property} \Longrightarrow P\text{-property}.$$

In a recent paper (Liu et al. 2005) where a symmetric cone complementarity problem (SCCP) is considered, it indicates that the uniform Jordan P-property is sufficient to guarantee the boundedness of the level sets of some merit functions which is a weaker assumption than that used in this paper. We suspect that similar conditions will hold for the merit functions studied in this paper.

Acknowledgment. The author thanks for the referees for their careful reading of the paper and helpful suggestions.

References

Alizadeh F, Schmieta S (2000) Symmetric cones, potential reduction methods, and word-by-word extensions. In: Wolkowicz H, Saigal R, Vandenberghe L (eds) Handbook of semidefinite programming. Kluwer, Boston, pp 195–233

Andersen ED, Roos C, Terlaky T (2003) On implementing a primal-dual interior-point method for conic quadratic optimization. Math Program Ser B 95:249–277

Chen J-S (2006) A new merit function and its related properties for the second-order cone complementarity problem. Pacific J Optim 2:167–179

Chen J-S, Tseng P (2005) An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. Math Program 104:293–327

Chen J-S, Chen X, Tseng P (2004) Analysis of nonsmooth vector-valued functions associated with second-order cone. Math Program 101:95–117

Chen X-D, Sun D, Sun J (2003) Complementarity functions and numerical experiments for second-order cone complementarity problems. Comput Optim Appl 25:39–56

Facchinei F, Pang J-S (2003) Finite-dimensional variational inequalities and complementarity problems. Vol I, II. Springer, Berlin Heidelberg New York

Faraut U, Korányi A (1994) Analysis on symmetric cones. Oxford Mathematical Monographs. Oxford University Press, New York

Fischer A (1992) A special Newton-type optimization methods. Optimization 24:269–284

Fischer A (1997) Solution of the monotone complementarity problem with locally Lipschitzian functions. Math Program 76:513–532

Fukushima M, Luo Z-Q, Tseng P (2002) Smoothing functions for second-order cone complementarity problems. SIAM J Optim 12:436–460

Goes RMB, Oliveira PR (2002) A new class of merit functions for the semidefinite complementarity problem. Annals of the Brazilian Workshop on Continuous Optimization, pp 1–18

Hayashi S, Yamashita N, Fukushima M (2002) On the coerciveness of merit functions for the second-order cone complementarity problem. Technical Report, Department of Applied Mathematics and Physics, Kyoto University

Hayashi S, Yamashita N, Fukushima M (2005) A combined smoothing and regularization method for monotone second-order cone complementarity problems. SIAM J Optim 15:593–615

Korányi A (1984) Monotone functions on formally real Jordan algebras. Mathematische Annalen 269:73–76

Liu Y-J, Zhang Z-W, Wang Y-H (2005) Some properties of a class of merit functions for symmetric cone complementarity problems. Asia-Pacific J Oper Res (to appear)

Lobo MS, Vandenberghe L, Boyd S, Lebret H (1998) Application of second-order cone programming. Linear Algebra Appl 284:193–228

Luo Z-Q, Tseng P (1997) A new class of merit functions for the nonlinear complementarity problem. In: Ferris MC, Pang J-S (eds) Complementarity and variational problems: state of the art. SIAM, Philadelphia, pp 204–225

Mittelmann HD (2003) An independent benchmarking of SDP and SOCP solvers. Math Program 95:407–430

Monteiro RDC, Tsuchiya T (2000) Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions. Math Program 88:61–83

Rockafellar RT (1970) Convex analysis. Princeton Mathematical Series, Princeton
Schmieta S, Alizadeh F (2001) Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. Math Oper Res 26:543–564
Sim C-K, Zhao G (2005) A note on treating second order cone problem as a special case of semidefinite problem. Math Program 102:609–613
Tao J, Gowda MS (2004) Some P-properties for the nonlinear transformations on Euclidean Jordan Algebra. Technical Report, Department of Mathematics and Statistics, University of Maryland
Tseng P (1998) Merit function for semidefinite complementarity problems. Math Program 83:159–185
Tsuchiya T (1999) A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim Meth Softw 11:141–182