# A semismooth Newton method for SOCCPs based on a one-parametric class of SOC complementarity functions

Shaohua Pan [1]
School of Mathematical Sciences
South China University of Technology
Guangzhou 510640, China

Jein-Shan Chen [2]
Department of Mathematics
National Taiwan Normal University
Taipei 11677, Taiwan

March 26, 2007
(revised on July 30, 2007)
(2nd revised on October 28, 2007)

**Abstract.** In this paper, we present a detailed investigation for the properties of a one-parametric class of SOC complementarity functions, which include the globally Lipschitz continuity, strong semismoothness, and the characterization of the B-subdifferential at a general point. Moreover, for the merit functions induced by them for the second-order cone complementarity problem (SOCCP), we provide a condition for each stationary point being a solution of the SOCCP and establish the boundedness of their level sets, by exploiting Cartesian $P$-properties. We also propose a semismooth Newton method based on the reformulation of the nonsmooth system of equations involving the class of SOC complementarity functions. The global and superlinear convergence results are obtained, and among others, the superlinear convergence is established under strict complementarity. Preliminary numerical results are reported for DIMACS second-order cone programs, which confirm the favorable theoretical properties of the method.

**Key words.** Second-order cone, complementarity, semismooth, B-subdifferential, Newton's method.

[1]The author's work is partially supported by the Doctoral Starting-up Foundation (B13B6050640) of GuangDong Province. E-mail:shhpan@scut.edu.cn.

[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. E-mail: jschen@math.ntnu.edu.tw.

# 1 Introduction

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

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

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $F$ and $G$ are the mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ which are assumed to be continuously differentiable, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOCs), also called Lorentz cones [8]. In other words,

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

where $m, n_1, \ldots, n_m \geq 1$, $n_1 + n_2 + \cdots + n_m = n$, and

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

with $\|\cdot\|$ 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)*. In addition, we write $F = (F_1, \ldots, F_m)$ and $G = (G_1, \ldots, G_m)$ with $F_i, G_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$.

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

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

which is a natural extension of the nonlinear complementarity problem (NCP) where $\mathcal{K} = \mathcal{K}^1 \times \cdots \times \mathcal{K}^1$. Another important special case corresponds to the Karush-Kuhn-Tucker (KKT) conditions of the convex second-order cone program (SOCP):

$$\begin{array}{ll} \min & g(x) \\ \text{s.t.} & Ax = b, \quad x \in \mathcal{K}, \end{array} \tag{4}$$

where $A \in \mathbb{R}^{m \times n}$ has full row rank, $b \in \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}$ is a twice continuously differentiable convex function. From [7], the KKT conditions for (4), which are sufficient but not necessary for optimality, can be written in the form of (1) with

$$F(\zeta) := d + (I - A^T(AA^T)^{-1}A)\zeta, \quad G(\zeta) := \nabla g(F(\zeta)) - A^T(AA^T)^{-1}A\zeta, \tag{5}$$

where $d \in \mathbb{R}^n$ is any vector satisfying $Ax = b$. For large problems with a sparse $A$, (5) has an advantage that the main cost of evaluating the Jacobian $\nabla F$ and $\nabla G$ lies in inverting $AA^T$, which can be done efficiently via sparse Cholesky factorization.

There have been various methods proposed for solving SOCPs and SOCCPs. They include interior-point methods [1, 2, 17, 18, 24], non-interior smoothing Newton methods [4, 9], the smoothing-regularization method [13], the merit function method [7] and the semismooth Newton method [15]. Among others, the last four kinds of methods are all based on an SOC complementarity function or a smooth merit function induced by it. Given a mapping $\phi : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}^l$ $(l \geq 1)$, we call $\phi$ an *SOC complementarity function* associated with the cone $\mathcal{K}^l$ if for any $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$,

$$\phi(x,y) = 0 \iff x \in \mathcal{K}^l, \quad y \in \mathcal{K}^l, \quad \langle x, y\rangle = 0. \tag{6}$$

Clearly, when $l = 1$, an SOC complementarity function reduces to an NCP function, which plays an important role in the solution of NCPs; see [22] and references therein.

A popular choice of $\phi$ is the Fischer-Burmeister (FB) function [10, 11], defined by

$$\phi_{\rm FB}(x,y) \ := \ (x^2 + y^2)^{1/2} - (x+y), \tag{7}$$

where $x^2$ means $x \circ x$ with "$\circ$" denoting the Jordan product, and $x+y$ denotes the usual componentwise addition of vectors. More specifically, for any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, we define their *Jordan product* associated with $\mathcal{K}^l$ as

$$x \circ y \ := \ (\langle x, y\rangle, y_1 x_2 + x_1 y_2). \tag{8}$$

The Jordan product, unlike scalar or matrix multiplication, is not associative, which is the main source on complication in the analysis of SOCCPs. The identity element under this product is $e := (1, 0, \cdots, 0)^T \in \mathbb{R}^l$. It is known that $x^2 \in \mathcal{K}^l$ for all $x \in \mathbb{R}^l$. Moreover, if $x \in \mathcal{K}^l$, then there exists a unique vector in $\mathcal{K}^l$, denoted by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Thus, $\phi_{\rm FB}$ in (7) is well-defined for all $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$. The function $\phi_{\rm FB}$ was proved in [9] to satisfy the equivalence (6), and its squared norm

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

has been shown to be continuously differentiable everywhere by Chen and Tseng [7]. Another popular choice of $\phi$ is the residual function $\phi_{\rm NR} : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}^l$ given by

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

where $[\,\cdot\,]_+$ means the minimum Euclidean distance projection onto $\mathcal{K}^l$. The function was studied in [9, 13] which is involved in smoothing methods for the SOCCP, and recently it was used to develop a semismooth Newton method for nonlinear SOCPs by Kanzow and Fukushima [15]. The function $\phi_{\rm NR}$ also induces a merit function

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

but, compared to $\psi_{\rm FB}$, it has a remarkable drawback, i.e. the non-differentiability.

In this paper, we consider a one-parametric class of vector-valued functions

$$\phi_\tau(x,y) := \left[(x-y)^2 + \tau(x \circ y)\right]^{1/2} - (x+y) \tag{9}$$

with $\tau$ being an arbitrary fixed parameter from $(0, 4)$. The class of functions is a natural extension of the family of NCP functions proposed by Kanzow and Kleinmichel [14], and has been shown to satisfy the characterization (6) in [6]. It is not hard to see that as $\tau = 2$, $\phi_\tau$ reduces to the FB function $\phi_{\rm FB}$ in (7) while it becomes a multiple of the natural residual function $\phi_{\rm NR}$ as $\tau \to 0^+$. With the class of SOC complementarity functions, the SOCCP can be reformulated as a nonsmooth system of equations

$$\Phi_\tau(\zeta) := \begin{pmatrix} \phi_\tau(F_1(\zeta), G_1(\zeta)) \\ \vdots \\ \phi_\tau(F_i(\zeta), G_i(\zeta)) \\ \vdots \\ \phi_\tau(F_m(\zeta), G_m(\zeta)) \end{pmatrix} = 0, \tag{10}$$

which induces a natural merit function $\Psi_\tau : \mathbb{R}^n \to \mathbb{R}_+$ given by

$$\Psi_\tau(\zeta) = \frac{1}{2}\|\Phi_\tau(\zeta)\|^2 = \sum_{i=1}^{m} \psi_\tau(F_i(\zeta), G_i(\zeta), \tag{11}$$

with

$$\psi_\tau(x, y) = \frac{1}{2}\|\phi_\tau(x, y)\|^2. \tag{12}$$

In [6], we studied the continuous differentiability of $\psi_\tau$ and proved that each stationary point of $\Psi_\tau$ is a solution of the SOCCP if $\nabla F$ and $-\nabla G$ are column monotone. This paper focuses on other properties of $\phi_\tau$, including the globally Lipschitz continuity, the strong semismoothness, and the characterization of the B-subdifferential. Particularly, we provide a weaker condition than [6] for each stationary point of $\Psi_\tau$ to be a solution of the SOCCP and establish the boundedness of the level sets of $\Psi_\tau$, by using Cartesian $P$-properties. We also propose a semismooth Newton method based on (10), and obtain the corresponding global and the superlinear convergence results. Among others, the superlinear convergence is established under strict complementarity.

Throughout this paper, $I$ represents an identity matrix of suitable dimension, and $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is identified with $\mathbb{R}^{n_1 + \cdots + n_m}$. For a differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x)$ denotes the transpose of the Jacobian $F'(x)$. For a symmetric matrix $A \in \mathbb{R}^{n \times n}$, we write $A \succeq O$ (respectively, $A \succ O$) to mean $A$ is positive semidefinite (respectively, positive definite). Given a finite number of square matrices $Q_1, \ldots, Q_m$, we denote the block diagonal matrix with these matrices as block diagonals by $\mathrm{diag}(Q_1, \ldots, Q_m)$ or by $\mathrm{diag}(Q_i,\ i = 1, \ldots, m)$. If $\mathcal{J}$ and $\mathcal{B}$ are index sets such that $\mathcal{J}, \mathcal{B} \subseteq \{1, 2, \ldots, m\}$, we denote $P_{\mathcal{J}\mathcal{B}}$ by the block matrix consisting of the submatrices $P_{jk} \in \mathbb{R}^{n_j \times n_k}$ of $P$ with $j \in \mathcal{J}, k \in \mathcal{B}$, and by $x_{\mathcal{B}}$ a vector consisting of subvectors $x_i \in \mathbb{R}^{n_i}$ with $i \in \mathcal{B}$.

## 2 Preliminaries

This section recalls some background materials and preliminary results that will be used in the subsequent sections. We begin with the interior and the boundary of $\mathcal{K}^l$ ($l \geq 1$). It is known that $\mathcal{K}^l$ is a closed convex self-dual cone with nonempty interior given by

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

and the boundary given by

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

For each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, the *determinant* and the *trace* of $x$ 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 $x_2 = \alpha y_2$ for some $\alpha \in \mathbb{R}$. A vector $x \in \mathbb{R}^l$ is said to be *invertible* if $\det(x) \neq 0$, and its inverse is denoted by $x^{-1}$. Given a vector $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, we often use the following symmetry matrix

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

which can be viewed as a linear mapping from $\mathbb{R}^l$ to $\mathbb{R}^l$. It is easy to verify $L_x y = x \circ y$ and $L_{x+y} = L_x + L_y$ for any $x, y \in \mathbb{R}^l$. Furthermore, $x \in \mathcal{K}^l$ if and only if $L_x \succeq O$, and $x \in \mathrm{int}(\mathcal{K}^l)$ if and only if $L_x \succ O$. If $x \in \mathrm{int}(\mathcal{K}^l)$, then $L_x$ is invertible with

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

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

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

where $\lambda_i(x)$ and $u_x^{(i)}$ for $i = 1, 2$ are the spectral values and the associated spectral vectors of $x$, respectively, given by

$$\lambda_i(x) = x_1 + (-1)^i \|x_2\|, \quad u_x^{(i)} = \frac{1}{2}\left(1, (-1)^i \bar{x}_2\right) \tag{15}$$

with $\bar{x}_2 = x_2/\|x_2\|$ if $x_2 \neq 0$, and otherwise $\bar{x}_2$ being any vector in $\mathbb{R}^{l-1}$ such that $\|\bar{x}_2\| = 1$. If $x_2 \neq 0$, then the factorization is unique. The spectral decomposition of $x, x^2$ and $x^{1/2}$ has some basic properties as below, whose proofs can be found in [9].

**Property 2.1** *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ with the spectral values $\lambda_1(x), \lambda_2(x)$ and spectral vectors $u_x^{(1)}, u_x^{(2)}$ given as above, the following results hold:*

**(a)** $x \in \mathcal{K}^l$ *if and only if* $\lambda_1(x) \geq 0$, *and* $x \in \mathrm{int}(\mathcal{K}^l)$ *if and only if* $\lambda_1(x) > 0$.

**(b)** $x^2 = \lambda_1^2(x) u_x^{(1)} + \lambda_2^2(x) u_x^{(2)} \in \mathcal{K}^l$;

**(c)** $x^{1/2} = \sqrt{\lambda_1(x)}\ u_x^{(1)} + \sqrt{\lambda_2(x)}\ u_x^{(2)} \in \mathcal{K}^l$ *if* $x \in \mathcal{K}^l$.

**(d)** $\det(x) = \lambda_1(x)\lambda_2(x),\ \mathrm{tr}(x) = \lambda_1(x) + \lambda_2(x)$ *and* $\|x\|^2 = [\lambda_1^2(x) + \lambda_2^2(x)]/2$.

For the sake of notation, throughout the rest of this paper, we always write

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

for any $x = (x_1, x_2),\ y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, and let $\bar{w}_2 = w_2/\|w_2\|$ if $w_2 \neq 0$, and otherwise $\bar{w}_2$ be any vector in $\mathbb{R}^{l-1}$ satisfying $\|\bar{w}_2\| = 1$. We have

$$w_1 = \|x\|^2 + \|y\|^2 + (\tau - 2)x^T y, \quad w_2 = 2(x_1 x_2 + y_1 y_2) + (\tau - 2)(x_1 y_2 + y_1 x_2).$$

Moreover, $w \in \mathcal{K}^l$ and $z \in \mathcal{K}^l$ hold by noting that

$$\begin{aligned} w = x^2 + y^2 + (\tau - 2)(x \circ y) &= \left(x + \frac{\tau - 2}{2} y\right)^2 + \frac{\tau(4 - \tau)}{4} y^2 \\ &= \left(y + \frac{\tau - 2}{2} x\right)^2 + \frac{\tau(4 - \tau)}{4} x^2. \end{aligned} \tag{17}$$

In addition, using Property 2.1 (b) and (c), it is not hard to compute that

$$z = \left( \frac{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}}{2}, \frac{\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}}{2} \bar{w}_2 \right) \in \mathcal{K}^l. \tag{18}$$

The following lemma characterizes the set of points where $z(x, y)$ is (continuously) differentiable. Since the proof is direct by the arguments in Case (2) of [6, Proposition 3.2] and formulas (18) and (14), we here omit it.

**Lemma 2.1** *The function $z(x, y)$ defined by (16) is (continuously) differentiable at a point $(x, y)$ if and only if $(x - y)^2 + \tau(x \circ y) \in \mathrm{int}(\mathcal{K}^l)$, and furthermore,*

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

*where*

$$L_z^{-1} = \begin{cases} \begin{pmatrix} b & c\bar{w}_2^T \\ c\bar{w}_2 & aI + (b - a)\bar{w}_2 \bar{w}_2^T \end{pmatrix} & \text{if } w_2 \neq 0; \\ \left(1/\sqrt{w_1}\right) I & \text{if } w_2 = 0, \end{cases} \tag{19}$$

*with* $a = \frac{2}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}}$, $b = \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2(w)}} + \frac{1}{\sqrt{\lambda_1(w)}}\right)$ *and* $c = \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2(w)}} - \frac{1}{\sqrt{\lambda_1(w)}}\right)$.

**Lemma 2.2** *[6, Lemma 3.1] For any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, let $w = (w_1, w_2)$ be given as in (16). If $(x-y)^2 + \tau(x \circ y) \notin \mathrm{int}(\mathcal{K}^l)$, then*

$$\begin{aligned} x_1^2 \;=\; \|x_2\|^2,\ y_1^2 \;=\; \|y_2\|^2,\ \ x_1 y_1 \;=\; x_2^T y_2,\ x_1 y_2 \;=\; y_1 x_2; \qquad (20) \\ x_1^2 + y_1^2 + (\tau - 2)x_1 y_1 \;=\; \|x_1 x_2 + y_1 y_2 + (\tau-2)x_1 y_2\| \\ =\; \|x_2\|^2 + \|y_2\|^2 + (\tau-2)x_2^T y_2. \qquad (21) \end{aligned}$$

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

$$x_2^T \bar{w}_2 = x_1, \quad x_1 \bar{w}_2 = x_2, \quad y_2^T \bar{w}_2 = y_1, \quad y_1 \bar{w}_2 = y_2. \qquad (22)$$

**Lemma 2.3** *[6, Lemma 3.2] For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, let $w = (w_1, w_2)$ be defined as in (16). If $w_2 \neq 0$, then for $i = 1, 2$,*

$$\begin{aligned} &\left[\left(x_1 + \frac{\tau-2}{2}y_1\right) + (-1)^i \left(x_2 + \frac{\tau-2}{2}y_2\right)^T \bar{w}_2\right]^2 \\ &\quad \leq \left\|\left(x_2 + \frac{\tau-2}{2}y_2\right) + (-1)^i \left(x_1 + \frac{\tau-2}{2}y_1\right)\bar{w}_2\right\|^2 \leq \lambda_i(w). \end{aligned}$$

*Furthermore, these relations also hold when interchanging $x$ and $y$.*

To close this section, we recall some concepts that will be used in the sequel. Given a mapping $H : \mathbb{R}^n \to \mathbb{R}^m$, if $H$ is locally Lipschitz continuous, the following set

$$\partial_B H(z) := \left\{V \in \mathbb{R}^{m\times n} |\ \exists \{z^k\} \subseteq D_H : z^k \to z, H'(z^k) \to V\right\}$$

is nonempty and is called the B-subdifferential of $H$ at $z$, where $D_H \subseteq \mathbb{R}^n$ denotes the set of points at which $H$ is differentiable. The convex hull $\partial H(z) := \mathrm{conv}\partial_B H(z)$ is the generalized Jacobian of $H$ at $z$ in the sense of Clarke [5]. For the concepts of (strongly) semismooth functions, please refer to [20, 21] for details. We next present definitions of Cartesian $P$-properties for a matrix $M \in \mathbb{R}^{n\times n}$, which are in fact special cases of those introduced by Chen and Qi [3] for a linear transformation.

**Definition 2.1** *A matrix $M \in \mathbb{R}^{n\times n}$ is said to have*

**(a)** *the Cartesian $P$-property if for any $0 \neq x = (x_1, \ldots, x_m) \in \mathbb{R}^n$ with $x_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \ldots, m\}$ such that $\langle x_\nu, (Mx)_\nu \rangle > 0$;*

**(b)** *the Cartesian $P_0$-property if for any $0 \neq x = (x_1, \ldots, x_m) \in \mathbb{R}^n$ with $x_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \ldots, m\}$ such that $x_\nu \neq 0$ and $\langle x_\nu, (Mx)_\nu \rangle \geq 0$.*

Some nonlinear generalizations of these concepts in the setting of $\mathcal{K}$ are defined as follows.

**Definition 2.2** *Given a mapping $F = (F_1, \ldots, F_m)$ with $F_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$, $F$ is said to*

**(a)** *have the uniform Cartesian $P$-property if for any $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m) \in \mathbb{R}^n$, there is an index $\nu \in \{1, 2, \ldots, m\}$ and a constant $\rho > 0$ such that*

$$\langle x_\nu - y_\nu, F_\nu(x) - F_\nu(y)\rangle \geq \rho\|x - y\|^2;$$

**(b)** *have the Cartesian $P_0$-property if for any $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m) \in \mathbb{R}^n$ and $x \neq y$, there exists an index $\nu \in \{1, 2, \ldots, m\}$ such that*

$$x_\nu \neq y_\nu \quad \text{and} \quad \langle x_\nu - y_\nu, F_\nu(x) - F_\nu(y)\rangle \geq 0.$$

# 3 Properties of the functions $\phi_\tau$ and $\Phi_\tau$

First, we study the favorable properties of $\phi_\tau$, including the globally Lipschitz continuity, the strong semismoothness and the characterization of the B-subdifferential at any point.

**Proposition 3.1** *The function $\phi_\tau$ defined as in (9) has the following properties.*

**(a)** *$\phi_\tau$ is (continuously) differentiable at $(x, y)$ if and only if $w(x, y) \in \mathrm{int}(\mathcal{K}^l)$. Also,*

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

**(b)** *$\phi_\tau$ is globally Lipschitz continuous with the Lipschitz constant independent of $\tau$.*

**(c)** *$\phi_\tau$ is strongly semismooth at any $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$.*

**(d)** *The squared norm of $\phi_\tau$, i.e. $\psi_\tau$, is continuously differentiable everywhere.*

**Proof.** (a) The proof directly follows from Lemma 2.1 and the following fact that

$$\phi_\tau(x, y) = z(x, y) - (x + y). \tag{23}$$

(b) It suffices to prove that $z(x, y)$ is globally Lipschitz continuous by (23). Let

$$\hat{z} \ = \ \hat{z}(x, y, \epsilon) \ := \ \left[(x - y)^2 + \tau(x \circ y) + \epsilon e\right]^{1/2} \tag{24}$$

for any $\epsilon > 0$ and $x = (x_1, x_2), \ y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$. Then, applying Lemma A.1 in the appendix and the Mean-Value Theorem, we have

$$\begin{aligned}
\Big\|z(x, y) - z(a, b)\Big\| &= \Big\|\lim_{\epsilon\to 0^+} \hat{z}(x, y, \epsilon) - \lim_{\epsilon\to 0^+} \hat{z}(a, b, \epsilon)\Big\| \\
&\leq \lim_{\epsilon\to 0^+} \|\hat{z}(x, y, \epsilon) - \hat{z}(a, y, \epsilon) + \hat{z}(a, y, \epsilon) - \hat{z}(a, b, \epsilon)\| \\
&\leq \lim_{\epsilon\to 0^+} \left\|\int_0^1 \nabla_x\hat{z}(a + t(x - a), y, \epsilon)(x - a)dt\right\| \\
&\quad + \lim_{\epsilon\to 0^+} \left\|\int_0^1 \nabla_y\hat{z}(a, b + t(y - b), \epsilon)(y - b)dt\right\| \\
&\leq \sqrt{2}C\|(x, y) - (a, b)\|
\end{aligned}$$

for any $(x, y), (a, b) \in \mathrm{IR}^l \times \mathrm{IR}^l$, where $C > 0$ is a constant independent of $\tau$.

(c) From the definition of $\phi_\tau$ and $\phi_{\mathrm{FB}}$, it is not hard to check that

$$\phi_\tau(x, y) = \phi_{\mathrm{FB}}\left(x + \frac{\tau-2}{2}y, \frac{\sqrt{\tau(4-\tau)}}{2}y\right) + \frac{1}{2}\left(\tau - 4 + \sqrt{\tau(4-\tau)}\right)y.$$

Notice that $\phi_{\mathrm{FB}}$ is strongly semismooth by [23, Corollary 3.3], and the functions $x+\frac{\tau-2}{2}y$, $\frac{1}{2}\sqrt{\tau(4-\tau)}y$ and $\frac{1}{2}(\tau-4+\sqrt{\tau(4-\tau)})y$ are also strongly semismooth. Therefore, $\phi_\tau$ is a strongly semismooth function since by [11, Theorem 19] the composition of strongly semismooth functions is strongly semismooth.

(d) The proof can be found in Proposition 3.3 of [6]. $\square$

Proposition 3.1 (c) indicates that, when a smoothing or nonsmooth Newton method is used to solve system (10), a fast convergence rate (at least superlinear) may be expected. To develop a semismooth Newton method for the SOCCP, we need to characterize the B-subdifferential $\partial_B\phi_\tau(x, y)$ at a general point $(x, y)$. The discussion of B-subdifferential for $\phi_{\mathrm{FB}}$ was given in [19], and we here generalize it to $\phi_\tau$ for any $\tau \in (0, 4)$. The detailed derivation process is included in the appendix for completeness.

**Proposition 3.2** *Given a general point $(x, y) \in \mathrm{IR} \times \mathrm{IR}^{l-1}$, each element in $\partial_B\phi_\tau(x, y)$ is of the form $V = [V_x - I \;\; V_y - I]$ with $V_x$ and $V_y$ having the following representation:*

**(a)** *If $(x-y)^2 + \tau(x \circ y) \in \mathrm{int}(\mathcal{K}^l)$, then $V_x = L_z^{-1}L_{x+\frac{\tau-2}{2}y}$ and $V_y = L_z^{-1}L_{y+\frac{\tau-2}{2}x}$.*

**(b)** *If $(x-y)^2 + \tau(x \circ y) \in \mathrm{bd}(\mathcal{K}^l)$ and $(x, y) \neq (0, 0)$, then*

$$\begin{aligned} V_x &\in \left\{\frac{1}{2\sqrt{2w_1}}\begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T\end{pmatrix}\left(L_x + \frac{\tau-2}{2}L_y\right) + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2\end{pmatrix}u^T\right\} \\ V_y &\in \left\{\frac{1}{2\sqrt{2w_1}}\begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T\end{pmatrix}\left(L_y + \frac{\tau-2}{2}L_x\right) + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2\end{pmatrix}v^T\right\} \end{aligned} \quad (25)$$

*for some $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathrm{IR} \times \mathrm{IR}^{l-1}$ satisfying $|u_1| \le \|u_2\| \le 1$ and $|v_1| \le \|v_2\| \le 1$, where $\bar{w}_2 = \frac{w_2}{\|w_2\|}$.*

**(c)** *If $(x, y) = (0, 0)$, then $V_x \in \{L_{\hat{u}}\}, V_y \in \{L_{\hat{v}}\}$ for some $\hat{u} = (\hat{u}_1, \hat{u}_2), \hat{v} = (\hat{v}_1, \hat{v}_2) \in \mathrm{IR} \times \mathrm{IR}^{l-1}$ satisfying $\|\hat{u}\|, \|\hat{v}\| \le 1$ and $\hat{u}_1\hat{v}_2 + \hat{v}_1\hat{u}_2 = 0$, or*

$$\begin{aligned} V_x &\in \left\{\frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2\end{pmatrix}\xi^T + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2\end{pmatrix}u^T + 2\begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2\bar{w}_2^T)s_2 & (I - \bar{w}_2\bar{w}_2^T)s_1\end{pmatrix}\right\} \\ V_y &\in \left\{\frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2\end{pmatrix}\eta^T + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2\end{pmatrix}v^T + 2\begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2\bar{w}_2^T)\omega_2 & (I - \bar{w}_2\bar{w}_2^T)\omega_1\end{pmatrix}\right\} \end{aligned} \quad (26)$$

*for some* $u=(u_1,u_2), v=(v_1,v_2), \xi=(\xi_1,\xi_2), \eta=(\eta_1,\eta_2) \in \mathbb{R}\times\mathbb{R}^{l-1}$ *satisfying* $|u_1| \le \|u_2\| \le 1$, $|v_1| \le \|v_2\| \le 1$, $|\xi_1| \le \|\xi_2\| \le 1$ *and* $|\eta_1| \le \|\eta_2\| \le 1$, $\bar{w}_2 \in \mathbb{R}^{l-1}$ *satisfying* $\|\bar{w}_2\| = 1$, *and* $s=(s_1,s_2)$, $\omega=(\omega_1,\omega_2) \in \mathbb{R}\times\mathbb{R}^{l-1}$ *such that* $\|s\|^2 + \|\omega\|^2 \le 1$.

In what follows, we investigate the properties of the operator $\Phi_\tau : \mathbb{R}^n \to \mathbb{R}^n$ given by (10). We start with the semismoothness of $\Phi_\tau$. Since $\Phi_\tau$ is (strongly) semismooth if and only if all component functions are (strongly) semismooth, and since the composite of (strongly) semismooth functions is (strongly) semismooth by [11, Theorem 19], we obtain the following conclusion as an immediate consequence of Proposition 3.1 (c).

**Proposition 3.3** *The operator* $\Phi_\tau : \mathbb{R}^n \to \mathbb{R}^n$ *given by (10) is semismooth. Moreover, it is strongly semismooth if* $F'$ *and* $G'$ *are locally Lipschitz continuous.*

To characterize the B-subdifferential of $\Phi_\tau$, we write $F_i(\zeta) = (F_{i1}(\zeta), F_{i2}(\zeta))$ and $G_i(\zeta) = (G_{i1}(\zeta), G_{i2}(\zeta))$, and denote $w_i$ and $z_i$ for $i = 1, 2, \ldots, m$ by

$$w_i = (w_{i1}(\zeta), w_{i2}(\zeta)) = w(F_i(\zeta), G_i(\zeta)), \quad z_i = (z_{i1}(\zeta), z_{i2}(\zeta)) = z(F_i(\zeta), G_i(\zeta)). \qquad (27)$$

For convenience, we sometimes suppress in $F_i(\zeta)$ and $G_i(\zeta)$ the dependence on $\zeta$.

**Proposition 3.4** *Let* $\Phi_\tau : \mathbb{R}^n \to \mathbb{R}^n$ *be defined as in (10). Then, for any* $\zeta \in \mathbb{R}^n$,

$$\partial_B \Phi_\tau(\zeta)^T \subseteq \nabla F(\zeta)\,(A(\zeta) - I) + \nabla G(\zeta)\,(B(\zeta) - I), \qquad (28)$$

*where* $A(\zeta)$ *and* $B(\zeta)$ *are possibly multivalued* $n \times n$ *block diagonal matrices whose* $i$*th blocks* $A_i(\zeta)$ *and* $B_i(\zeta)$ *for* $i = 1, 2, \ldots, m$ *have the following representation.*

**(a)** *If* $(F_i(\zeta) - G_i(\zeta))^2 + \tau\,(F_i(\zeta) \circ G_i(\zeta)) \in \mathrm{int}(\mathcal{K}^{n_i})$, *then*

$$A_i(\zeta) = L_{F_i + \frac{\tau-2}{2}G_i} L_{z_i}^{-1} \quad \text{and} \quad B_i(\zeta) = L_{G_i + \frac{\tau-2}{2}F_i} L_{z_i}^{-1}.$$

**(b)** *If* $(F_i(\zeta), G_i(\zeta)) \neq (0,0)$ *and* $(F_i(\zeta) - G_i(\zeta))^2 + \tau\,(F_i(\zeta) \circ G_i(\zeta)) \in \mathrm{bd}(\mathcal{K}^{n_i})$, *then*

$$A_i(\zeta) \in \left\{ \frac{1}{2\sqrt{2w_{i1}}} \left( L_{F_i} + \frac{\tau-2}{2} L_{G_i} \right) \begin{pmatrix} 1 & \bar{w}_{i2}^T \\ \bar{w}_{i2} & 4I - 3\bar{w}_{i2}\bar{w}_{i2}^T \end{pmatrix} + \frac{1}{2} u_i (1, -\bar{w}_{i2}^T) \right\}$$

$$B_i(\zeta) \in \left\{ \frac{1}{2\sqrt{2w_{i1}}} \left( L_{G_i} + \frac{\tau-2}{2} L_{F_i} \right) \begin{pmatrix} 1 & \bar{w}_{i2}^T \\ \bar{w}_{i2} & 4I - 3\bar{w}_{i2}\bar{w}_{i2}^T \end{pmatrix} + \frac{1}{2} v_i (1, -\bar{w}_{i2}^T) \right\}$$

*for some* $u_i = (u_{i1}, u_{i2}), v_i = (v_{i1}, v_{i2}) \in \mathbb{R}\times\mathbb{R}^{n_i-1}$ *satisfying* $|u_{i1}| \le \|u_{i2}\| \le 1$ *and* $|v_{i1}| \le \|v_{i2}\| \le 1$, *where* $\bar{w}_{i2} = \frac{w_{i2}}{\|w_{i2}\|}$.

**(c)** *If* $(F_i(\zeta), G_i(\zeta)) = (0,0)$, *then*

$$A_i(\zeta) \in \left\{L_{\hat{u}_i}\right\} \cup \left\{\frac{1}{2}\xi_i\left(1, \bar{w}_{i2}^T\right) + \frac{1}{2}u_i\left(1, -\bar{w}_{i2}^T\right) + \begin{pmatrix} 0 & 2s_{i2}^T(I - \bar{w}_{i2}\bar{w}_{i2}^T) \\ 0 & 2s_{i1}(I - \bar{w}_{i2}\bar{w}_{i2}^T) \end{pmatrix}\right\}$$
$$B_i(\zeta) \in \left\{L_{\hat{v}_i}\right\} \cup \left\{\frac{1}{2}\eta_i\left(1, \bar{w}_{i2}^T\right) + \frac{1}{2}v_i\left(1, -\bar{w}_{i2}^T\right) + \begin{pmatrix} 0 & 2\omega_{i2}^T(I - \bar{w}_{i2}\bar{w}_{i2}^T) \\ 0 & 2\omega_{i1}(I - \bar{w}_{i2}\bar{w}_{i2}^T) \end{pmatrix}\right\}$$

*for some* $\hat{u}_i = (\hat{u}_{i1}, \hat{u}_{i2}), \hat{v}_i = (\hat{v}_{i1}, \hat{v}_{i2}) \in \mathbb{R} \times \mathbb{R}^{n_i-1}$ *satisfying* $\|\hat{u}_i\|, \|\hat{v}_i\| \leq 1$ *and* $\hat{u}_{i1}\hat{v}_{i2} + \hat{v}_{i1}\hat{u}_{i2} = 0$, *some* $u_i = (u_{i1}, u_{i2}), v_i = (v_{i1}, v_{i2}), \xi_i = (\xi_{i1}, \xi_{i2}), \eta_i = (\eta_{i1}, \eta_{i2}) \in \mathbb{R} \times \mathbb{R}^{n_i-1}$ *with* $|u_{i1}| \leq \|u_{i2}\| \leq 1, |v_{i1}| \leq \|v_{i2}\| \leq 1, |\xi_{i1}| \leq \|\xi_{i2}\| \leq 1$ *and* $|\eta_{i1}| \leq \|\eta_{i2}\| \leq 1$, $\bar{w}_{i2} \in \mathbb{R}^{n_i-1}$ *satisfying* $\|\bar{w}_{i2}\| = 1$, *and* $s_i = (s_{i1}, s_{i2}), \omega_i = (\omega_{i1}, \omega_{i2}) \in \mathbb{R} \times \mathbb{R}^{n_i-1}$ *such that* $\|s_i\|^2 + \|\omega_i\|^2 \leq 1$.

**Proof.** Let $\Phi_{\tau,i}(\zeta)$ denote the $i$th subvector of $\Phi_\tau$, i.e. $\Phi_{\tau,i}(\zeta) = \phi_\tau(F_i(\zeta), G_i(\zeta))$ for all $i = 1, 2, \ldots, m$. From Proposition 2.6.2 of [5], it follows that

$$\partial_B \Phi_\tau(\zeta)^T \ \subseteq\ \partial_B \Phi_{\tau,1}(\zeta)^T \times \partial_B \Phi_{\tau,2}(\zeta)^T \times \cdots \times \partial_B \Phi_{\tau,m}(\zeta)^T, \tag{29}$$

where the latter denotes the set of all matrices whose $(n_{i-1}+1)$ to $n_i$th columns with $n_0 = 0$ belong to $\partial_B \Phi_{\tau,i}(\zeta)^T$. Using the definition of B-subdifferential and the continuous differentiability of $F$ and $G$, it is not difficult to verify that

$$\partial_B \Phi_{\tau,i}(\zeta)^T = [\nabla F_i(\zeta) \ \ \nabla G_i(\zeta)]\partial_B \phi_\tau(F_i(\zeta), G_i(\zeta))^T, \quad i = 1, \ldots, m. \tag{30}$$

Using Proposition 3.2 and the last two equations, we readily get the desired result. □

**Lemma 3.1** *For any* $\zeta \in \mathbb{R}^n$, *let* $A(\zeta)$ *and* $B(\zeta)$ *be the multivalued block diagonal matrices given as in Proposition 3.4. Then, for any* $i \in \{1, 2, \ldots, m\}$,

$$\langle (A_i(\zeta) - I)\Phi_{\tau,i}(\zeta),\ (B_i(\zeta) - I)\Phi_{\tau,i}(\zeta)\rangle \geq 0,$$

*and the equality holds if and only if* $\Phi_{\tau,i}(\zeta) = 0$. *Particularly, for the index* $i$ *such that* $(F_i(\zeta) - G_i(\zeta))^2 + \tau(F_i(\zeta) \circ G_i(\zeta) \in \mathrm{int}(\mathcal{K}^{n_i})$, *we have*

$$\langle (A_i(\zeta) - I)v_i,\ (B_i(\zeta) - I)v_i\rangle \geq 0, \quad \text{for any } v_i \in \mathbb{R}^{n_i}.$$

**Proof.** From Theorem 2.6.6 of [5] and Proposition 3.1 (d), we have

$$\nabla \psi_\tau(x, y) \ = \ \partial_B \phi_\tau(x, y)^T \phi_\tau(x, y).$$

Consequently, for any $i = 1, 2, \ldots, m$, it follows that

$$\nabla \psi_\tau(F_i(\zeta), G_i(\zeta)) \ = \ \partial_B \phi_\tau(F_i(\zeta), G_i(\zeta))^T \phi_\tau(F_i(\zeta), G_i(\zeta)).$$

In addition, from Propositions 3.2 and 3.4, it is not hard to see that

$$[A_i(\zeta)^T - I \quad B_i(\zeta)^T - I] \in \partial_B \phi_\tau(F_i(\zeta), G_i(\zeta)).$$

Combining with the last two equations yields that for any $i = 1, 2, \ldots, m$,

$$\begin{aligned} \nabla_x \psi_\tau(F_i(\zeta), G_i(\zeta)) &= (A_i(\zeta) - I)\Phi_{\tau,i}(\zeta) \\ \nabla_y \psi_\tau(F_i(\zeta), G_i(\zeta)) &= (B_i(\zeta) - I)\Phi_{\tau,i}(\zeta). \end{aligned} \tag{31}$$

Consequently, the first part of the conclusions is direct by Proposition 4.1 of [6]. Notice that for any $i$ such that $(F_i(\zeta) - G_i(\zeta))^2 + \tau(F_i(\zeta) \circ G_i(\zeta) \in \mathrm{int}(\mathcal{K}^{n_i})$ and any $\upsilon_i \in \mathrm{I\!R}^{n_i}$,

$$\begin{aligned} & \langle (A_i(\zeta) - I)\upsilon_i,\ (B_i(\zeta) - I)\upsilon_i \rangle \\ = & \left\langle \left(L_{F_i + \frac{\tau-2}{2}G_i} - L_{z_i}\right) L_{z_i}^{-1}\upsilon_i,\ \left(L_{G_i + \frac{\tau-2}{2}F_i} - L_{z_i}\right) L_{z_i}^{-1}\upsilon_i \right\rangle \\ = & \left\langle \left(L_{G_i + \frac{\tau-2}{2}F_i} - L_{z_i}\right)\left(L_{F_i + \frac{\tau-2}{2}G_i} - L_{z_i}\right) L_{z_i}^{-1}\upsilon_i,\ L_{z_i}^{-1}\upsilon_i \right\rangle. \end{aligned} \tag{32}$$

Therefore, using the same argument as Case (2) of [6, Proposition 4.1], we can obtain the second part of the conclusions. □

# 4 Nonsingularity conditions

In this section, we show that all elements of the B-subdifferential $\partial_B \Phi_\tau(\zeta)$ at a solution $\zeta^*$ of the SOCCP are nonsingular if $\zeta^*$ satisfies *strict complementarity*, i.e.,

$$F_i(\zeta^*) + G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i}) \quad \text{for all } i = 1, 2, \ldots, m. \tag{33}$$

First, we give a technical lemma which states that the multivalued matrix $(A_i(\zeta^*) - I) + (B_i(\zeta^*) - I)$ are nonsingular if the $i$th block component satisfies strict complementarity.

**Lemma 4.1** *Let $\zeta^*$ be a solution of the SOCCP, and $A(\zeta^*)$ and $B(\zeta^*)$ be the multivalued block diagonal matrices characterized by Proposition 3.4. Then, for any $i \in \{1, 2, \ldots, m\}$ such that $F_i(\zeta^*) + G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$, we have that $\Phi_{\tau,i}(\zeta)$ is continuously differentiable at $\zeta^*$ and $(A_i(\zeta^*) - I) + (B_i(\zeta^*) - I)$ is nonsingular.*

**Proof.** Since $\zeta^*$ is a solution of the SOCCP, we have for all $i = 1, 2, \ldots, m$

$$F_i(\zeta^*) \in \mathcal{K}^{n_i},\ G_i(\zeta^*) \in \mathcal{K}^{n_i},\ \langle F_i(\zeta^*), G_i(\zeta^*) \rangle = 0.$$

It is not hard to verify that $F_i(\zeta^*) + G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$ if and only if one of the three cases shown as below holds.

Case (1). $F_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$ and $G_i(\zeta^*) = 0$. Under this case,

$$w_i(\zeta^*) = (F_i(\zeta^*) - G_i(\zeta^*))^2 + \tau(F_i(\zeta^*) \circ G_i(\zeta^*)) = F_i(\zeta^*)^2 \in \mathrm{int}(\mathcal{K}^{n_i}).$$

By Proposition 3.1 (a), $\Phi_{\tau,i}(\zeta)$ is continuously differentiable at $\zeta^*$. Since $z_i(\zeta^*) = w_i(\zeta^*)^{1/2} = F_i(\zeta^*)$, from Proposition 3.4 (a) it follows that

$$A_i(\zeta^*) = I \ \text{ and } \ B_i(\zeta^*) = \frac{\tau - 2}{2} I,$$

which implies that $(A_i(\zeta^*) - I) + (B_i(\zeta^*) - I)$ is nonsingular since $0 < \tau < 4$.

Case (2). $F_i(\zeta^*) = 0$ and $G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$. Now, $w_i(\zeta^*) = G_i(\zeta^*)^2 \in \mathrm{int}(\mathcal{K}^{n_i})$. So, $\Phi_{\tau,i}(\zeta)$ is continuously differentiable at $\zeta^*$ by Proposition 3.1 (a). Since

$$z_i(\zeta^*) = w_i(\zeta^*)^{1/2} = G_i(\zeta^*),$$

using Proposition 3.4 (a) yields that $A_i(\zeta^*) = \frac{\tau-2}{2} I$ and $B_i(\zeta^*) = I$, which immediately implies that $(A_i(\zeta^*) - I) + (B_i(\zeta^*) - I)$ is nonsingular.

Case (3). $F_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i})$ and $G_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i})$. We claim that $w_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$ for this case. If not, then $w_i(\zeta^*) \in \mathrm{bd}(\mathcal{K}^{n_i})$. From (20) in Lemma 2.2, it follows that

$$F_{i1}(\zeta^*) G_{i1}(\zeta^*) = F_{i2}(\zeta^*)^T G_{i2}(\zeta^*). \tag{34}$$

Since $F_{i1}(\zeta^*) = \|F_{i2}(\zeta^*)\| \neq 0$ and $G_{i1}(\zeta^*) = \|G_{i2}(\zeta^*)\| \neq 0$, we have

$$\|F_{i2}(\zeta^*)\| \cdot \|G_{i2}(\zeta^*)\| = F_{i2}(\zeta^*)^T G_{i2}(\zeta^*),$$

which implies that $F_{i2}(\zeta^*) = \alpha G_{i2}(\zeta^*)$ for some constant $\alpha > 0$. Substituting it into (34) yields that $F_{i1}(\zeta^*) = \alpha G_{i1}(\zeta^*)$, and consequently, $F_i(\zeta^*) = \alpha G_i(\zeta^*)$. Noting that $\langle F_i(\zeta^*), G_i(\zeta^*) \rangle = 0$, we then obtain $F_i(\zeta^*) = G_i(\zeta^*) = 0$. This clearly contradicts the assumption that $F_i(\zeta^*) \neq 0$ and $G_i(\zeta^*) \neq 0$. Hence, $w_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$.

From the expression of $A_i(\zeta)$ and $B_i(\zeta)$ given by Proposition 3.4 (a),

$$(A_i(\zeta^*) - I) + (B_i(\zeta^*) - I) \ = \ -L_{2z_i(\zeta^*) - \frac{\tau}{2}(F_i(\zeta^*) + G_i(\zeta^*))} L^{-1}_{z_i(\zeta^*)}.$$

Therefore, to establish the nonsingularity of $(A_i(\zeta^*) - I) + (B_i(\zeta^*) - I)$, it suffices to prove that the matrix $L_{2z_i(\zeta^*) - \frac{\tau}{2}(F_i(\zeta^*) + G_i(\zeta^*))}$ is nonsingular. Since

$$\begin{aligned}\Big(2z_i(\zeta^*)\Big)^2 \ &= \ 2\left[\Big(F_i(\zeta^*) + \frac{\tau - 2}{2} G_i(\zeta^*)\Big)^2 + \frac{\tau(4-\tau)}{4} G_i(\zeta^*)^2\right] \\ &\quad +2\left[\Big(G_i(\zeta^*) + \frac{\tau - 2}{2} F_i(\zeta^*)\Big)^2 + \frac{\tau(4-\tau)}{4} F_i(\zeta^*)^2\right],\end{aligned}$$

it follows that

$$\begin{aligned}\Big(2z_i(\zeta^*)\Big)^2 - \frac{\tau^2}{4}\Big(F_i(\zeta^*)+G_i(\zeta^*)\Big)^2 &= \frac{\tau(4-\tau)}{2}\Big[G_i(\zeta^*)^2+F_i(\zeta^*)^2\Big] \\ &\quad + \frac{(4-\tau)^2}{4}\Big(F_i(\zeta^*)-G_i(\zeta^*)\Big)^2. \qquad (35)\end{aligned}$$

Notice that $w_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$ implies that $(F_i(\zeta^*) - G_i(\zeta^*))^2 \in \mathrm{int}(\mathcal{K}^{n_i})$ since $F_i(\zeta^*) \circ G_i(\zeta^*) = 0$, and hence from the equality (35) we immediately obtain that

$$\Big(2z_i(\zeta^*)\Big)^2 - \frac{\tau^2}{4}\Big(F_i(\zeta^*)+G_i(\zeta^*)\Big)^2 \in \mathrm{int}(\mathcal{K}^{n_i}).$$

Since $z_i(\zeta^*) = w_i(\zeta^*)^{1/2} \in \mathrm{int}(\mathcal{K}^{n_i})$, using Proposition 3.4 of [9] yields that

$$2z_i(\zeta^*) - \frac{\tau}{2}(F_i(\zeta^*)+G_i(\zeta^*)) \in \mathrm{int}(\mathcal{K}^{n_i}).$$

This means that $L_{2z_i(\zeta^*)-\frac{\tau}{2}(F_i(\zeta^*)+G_i(\zeta^*))} \succ O$, and consequently it is nonsingular. $\square$

Given a solution $\zeta^*$ of the SOCCP, we know from [1] that, if $\zeta^*$ is a strict complementarity one, i.e. satisfies the conditions in (33), the following index sets

$$\begin{aligned}\mathcal{I} &:= \Big\{i \in \{1,2,\ldots,m\} \mid F_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i}), G_i(\zeta^*) = 0\Big\}, \\ \mathcal{B} &:= \Big\{i \in \{1,2,\ldots,m\} \mid F_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i}), G_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i})\Big\}, \qquad (36) \\ \mathcal{J} &:= \Big\{i \in \{1,2,\ldots,m\} \mid F_i(\zeta^*) = 0, G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})\Big\}\end{aligned}$$

form a partition of $\{1,\ldots,m\}$, where $\mathrm{bd}^+(\mathcal{K}^{n_i}) = \mathrm{bd}(\mathcal{K}^{n_i}) \setminus \{0\}$. Thus, by supposing that $\nabla G(\zeta^*)$ is invertible and rearranging the matrices appropriately,

$$P(\zeta^*) = \nabla G(\zeta^*)^{-1}\nabla F(\zeta^*) = \begin{pmatrix} P(\zeta^*)_{\mathcal{I}\mathcal{I}} & P(\zeta^*)_{\mathcal{I}\mathcal{B}} & P(\zeta^*)_{\mathcal{I}\mathcal{J}} \\ P(\zeta^*)_{\mathcal{B}\mathcal{I}} & P(\zeta^*)_{\mathcal{B}\mathcal{B}} & P(\zeta^*)_{\mathcal{B}\mathcal{J}} \\ P(\zeta^*)_{\mathcal{J}\mathcal{I}} & P(\zeta^*)_{\mathcal{J}\mathcal{B}} & P(\zeta^*)_{\mathcal{J}\mathcal{J}} \end{pmatrix}.$$

Now we are in a position to establish the nonsingularity of all elements in $\partial_B \Phi_\tau(\zeta^*)$.

**Theorem 4.1** *Let $\zeta^*$ be a strict complementarity solution of the SOCCP. Suppose that $\nabla G(\zeta^*)$ is invertible and let $P(\zeta^*) = \nabla G(\zeta^*)^{-1}\nabla F(\zeta^*)$. If $P(\zeta^*)_{\mathcal{J}\mathcal{J}}$ is nonsingular and its Schur-complement, denoted by $\widehat{P}(\zeta^*)_{\mathcal{J}\mathcal{J}}$, in the matrix*

$$\begin{pmatrix} P(\zeta^*)_{\mathcal{B}\mathcal{B}} & P(\zeta^*)_{\mathcal{B}\mathcal{J}} \\ P(\zeta^*)_{\mathcal{J}\mathcal{B}} & P(\zeta^*)_{\mathcal{J}\mathcal{J}} \end{pmatrix}$$

*has the Cartesian P-property, then all $W \in \partial_B \Phi_\tau(\zeta^*)$ are nonsingular.*

**Proof.** By Proposition 3.4 and the invertibility of $\nabla G(\zeta^*)$, it suffices to show that any matrix $C$ belonging to $\nabla G(\zeta^*)^{-1}\nabla F(\zeta^*)(A(\zeta^*)-I)+(B(\zeta^*)-I)$ is invertible. Since $\zeta^*$ is a strict complementarity solution, it follows from Lemma 4.1 that the matrix $C$ can be written in the following partitioned form

$$C=\begin{pmatrix} \frac{\tau-4}{2}I_{\mathcal{I}} & P_{\mathcal{IB}}(A_{\mathcal{B}}-I_{\mathcal{B}}) & \frac{\tau-4}{2}P_{\mathcal{IJ}} \\ 0_{\mathcal{BI}} & P_{\mathcal{BB}}(A_{\mathcal{B}}-I_{\mathcal{B}})+(B_{\mathcal{B}}-I_{\mathcal{B}}) & \frac{\tau-4}{2}P_{\mathcal{BJ}} \\ 0_{\mathcal{JI}} & P_{\mathcal{JB}}(A_{\mathcal{B}}-I_{\mathcal{B}}) & \frac{\tau-4}{2}P_{\mathcal{JJ}} \end{pmatrix},$$

where $I_{\mathcal{I}}=\mathrm{diag}(I_i,\ i\in\mathcal{I})$ with $I_i$ being an $n_i\times n_i$ identity matrix, $A_{\mathcal{B}}=\mathrm{diag}(A_i,\ i\in\mathcal{B})$ and $B_{\mathcal{B}}=\mathrm{diag}(B_i,\ i\in\mathcal{B})$. For the sake of notation, we here omit the notation $\zeta^*$ in the functions. It is not hard to see that these $C$ are nonsingular if and only if

$$C_r=\begin{pmatrix} P_{\mathcal{BB}}(A_{\mathcal{B}}-I_{\mathcal{BB}})+(B_{\mathcal{B}}-I_{\mathcal{BB}}) & \frac{\tau-4}{2}P_{\mathcal{BJ}} \\ P_{\mathcal{JB}}(A_{\mathcal{B}}-I_{\mathcal{BB}}) & \frac{\tau-4}{2}P_{\mathcal{JJ}} \end{pmatrix}$$

is nonsingular. Showing that $C_r$ is nonsingular is equivalent to showing that the system

$$-C_r\begin{pmatrix} y_{\mathcal{B}} \\ y_{\mathcal{J}} \end{pmatrix}=0$$

for any $y=(y_{\mathcal{B}};y_{\mathcal{J}})$ has only the zero solution. This system can be rewritten as

$$\begin{cases} \frac{4-\tau}{2}P_{\mathcal{JJ}}y_{\mathcal{J}}+P_{\mathcal{JB}}(I_{\mathcal{BB}}-A_{\mathcal{B}})y_{\mathcal{B}}=0, \\ \\ \frac{4-\tau}{2}P_{\mathcal{BJ}}y_{\mathcal{J}}+P_{\mathcal{BB}}(I_{\mathcal{BB}}-A_{\mathcal{B}})y_{\mathcal{B}}=-(I_{\mathcal{BB}}-B_{\mathcal{B}})y_{\mathcal{B}}. \end{cases}$$

Recall that $P_{\mathcal{JJ}}$ is nonsingular, and we obtain from the last system that

$$\begin{cases} y_{\mathcal{J}}=-\frac{2}{4-\tau}P_{\mathcal{JJ}}^{-1}P_{\mathcal{JB}}(I_{\mathcal{BB}}-A_{\mathcal{B}})y_{\mathcal{B}}, \\ \\ (P_{\mathcal{BB}}-P_{\mathcal{BJ}}P_{\mathcal{JJ}}^{-1}P_{\mathcal{JB}})(I_{\mathcal{BB}}-A_{\mathcal{B}})y_{\mathcal{B}}=-(I_{\mathcal{BB}}-B_{\mathcal{B}})y_{\mathcal{B}}. \end{cases} \tag{37}$$

Thus, by Lemma 3.1 and Lemma 4.1, using the same arguments as Theorem 4.1 of [19] yields the desired result. □

Observe that, when $n_1=\cdots=n_m=1$, the assumption for $\widehat{P}_{\mathcal{JJ}}$ is actually equivalent to requiring that $\widehat{P}_{\mathcal{JJ}}$ is a $P$-matrix, which is common in the solution of NCPs. Now, we are not clear whether the result of Theorem 4.1 holds when removing the strict complementarity. We will leave it as a future research topic.

From Theorem 4.1 and [21, Lemma 2.6], we readily obtain the following result.

**Corollary 4.1** *Suppose that $\zeta^*$ is a strict complementarity solution of the SOCCP and the mapping $F$ and $G$ at the $\zeta^*$ satisfy the conditions of Theorem 4.1. Then, there exist a neighborhood $\mathcal{N}(\zeta^*)$ of $\zeta^*$ and a constant $C > 0$ such that for any $\zeta \in \mathcal{N}(\zeta^*)$ and any $W \in \partial_B \Phi_\tau(\zeta)$, $W$ is nonsingular and satisfies $\|W^{-1}\| \le C$.*

# 5 Stationary point condition and bounded level sets

In general a stationary point of a function is not a solution of the underlying problem. In [6], we showed that, when $\nabla F$ and $-\nabla G$ are column monotone, every stationary point of the smooth merit function $\Psi_\tau(\zeta)$ is a solution of the SOCCP. In this section, we provide a different stationary point condition by the Cartesian $P_0$-property of a matrix, which, as shown later, is weaker than that of [6] when $\nabla G$ is invertible. We also establish the boundedness of the level sets of $\Psi_\tau$ for the SOCCP (3) under the condition that $F$ has the uniform Cartesian $P$-property.

To present the first result of this section, we need the following technical lemma.

**Lemma 5.1** *Let $\psi_\tau : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}_+$ be given by (12). Then, for any $x, y \in \mathbb{R}^l$,*

$$\phi_\tau(x,y) \neq 0 \iff \nabla_x \psi_\tau(x,y) \neq 0,\ \nabla_y \psi_\tau(x,y) \neq 0.$$

**Proof.** From Proposition 3.2 of [6], the sufficiency is obvious. Suppose that $\phi_\tau(x,y) \neq 0$. If either $\nabla_x \psi_\tau(x,y) = 0$ or $\nabla_y \psi_\tau(x,y) = 0$, then $\langle \nabla_x \psi_\tau(x,y), \nabla_y \psi_\tau(x,y) \rangle = 0$. From Proposition 4.1 of [6], it follows that $\phi_\tau(x,y) = 0$. This gives a contradiction. □

**Proposition 5.1** *Let $\Psi_\tau : \mathbb{R}^n \to \mathbb{R}_+$ be given as (11). Suppose $\nabla G$ is invertible and $\nabla G(\zeta)^{-1} \nabla F(\zeta)$ at any $\zeta \in \mathbb{R}^n$ has the Cartesian $P_0$-property. Then, every stationary point of $\Psi_\tau$ is a solution of the SOCCP.*

**Proof.** Let $\zeta$ be an arbitrary stationary point of $\Psi_\tau(\zeta)$. Since $\Psi_\tau$ is continuously differentiable by Proposition 3.1 (d) and $\Phi_\tau$ is locally Lipschitz continuous, applying Theorem 2.6.6 of Clarke [5] then gives that for any $V \in \partial \Phi_\tau(\zeta)^T$

$$0 = \nabla \Psi_\tau(\zeta) = V \Phi_\tau(\zeta).$$

Let $V$ be an element of $\partial_B \Phi_\tau(\zeta)^T (\subseteq \partial \Phi_\tau(\zeta)^T)$. Then from (29) it follows that there exist matrices $V_i \in \partial_B \Phi_{\tau,i}(\zeta)^T$ such that

$$V \;=\; V_1 \times V_2 \times \cdots \times V_m.$$

In addition, for each $V_i \in \mathbb{R}^{n\times n_i}$, by Proposition 3.2 there exist matrices $A_i(\zeta) \in \mathbb{R}^{n_i\times n_i}$ and $B_i(\zeta) \in \mathbb{R}^{n_i\times n_i}$, as characterized by Proposition 3.4, such that

$$V_i = \nabla F_i(\zeta)(A_i(\zeta) - I) + \nabla G_i(\zeta)(B_i(\zeta) - I), \quad i = 1, 2, \ldots, m.$$

Let $A(\zeta) = \mathrm{diag}(A_1(\zeta), \ldots, A_m(\zeta))$ and $B(\zeta) = \mathrm{diag}(B_1(\zeta), \ldots, B_m(\zeta))$. Combining the last three equations, it then follows that

$$[\nabla F(\zeta)(A(\zeta) - I) + \nabla G(\zeta)(B(\zeta) - I)]\, \Phi_\tau(\zeta) = 0,$$

which, by the invertibility of $\nabla G(\zeta)$, is equivalent to

$$\left[\nabla G(\zeta)^{-1}\nabla F(\zeta)(A(\zeta) - I) + (B(\zeta) - I)\right] \Phi_\tau(\zeta) = 0. \tag{38}$$

Suppose that $\Phi_\tau(\zeta) \neq 0$. Then, there necessarily exists an index $\nu \in \{1, 2, \ldots, m\}$ such that $\Phi_{\tau,\nu}(\zeta) = \phi_\tau(F_\nu(\zeta), G_\nu(\zeta)) \neq 0$. Using Lemma 5.1 and equation (31) then yields

$$(A_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta) \neq 0 \;\text{ and }\; (B_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta) \neq 0. \tag{39}$$

In addition, from (38) it follows that

$$\left[\nabla G(\zeta)^{-1}\nabla F(\zeta)(A(\zeta) - I)\Phi_\tau(\zeta)\right]_\nu + (B_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta) = 0.$$

Making the inner product with $(A_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta)$ on both sides, we obtain

$$\begin{aligned}&\Big\langle (A_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta),\; \left[\nabla G(\zeta)^{-1}\nabla F(\zeta)(A(\zeta) - I)\Phi_\tau(\zeta)\right]_\nu \Big\rangle \\ &+ \Big\langle (A_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta),\; (B_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta) \Big\rangle \;=\; 0.\end{aligned}$$

Notice that the first term of the left hand side is nonnegative by (39) and the assumption that $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ has the Cartesian $P_0$-property at any $\zeta \in \mathbb{R}^n$, and the second term is positive by Lemma 3.1 since $\Phi_{\tau,\nu}(\zeta) \neq 0$. This leads to a contradiction. $\Box$

**Remark 5.1** **(i)** *It is easy to verify that $\nabla G(\zeta)^{-1}\nabla F(\zeta) \succeq O$ implies the Cartesian $P_0$-property of $\nabla G(\zeta)^{-1}\nabla F(\zeta)$. While, by [6], the column monotonicity of $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ is now equivalent to $\nabla G(\zeta)^{-1}\nabla F(\zeta) \succeq O$. This means that the condition in Proposition 5.1 is weaker than the one used by Proposition 4.2 of [6].*

**(ii)** *For the SOCCP (3), the condition of Proposition 5.1 is equivalent to requiring that $F$ has the Cartesian $P_0$-property. If $n_1 = n_2 = \cdots = n_m = 1$, this reduces to the common condition in the NCPs that $F$ is a $P_0$-function.*

**Lemma 5.2** *Let $\psi_\tau$ be given by (12). Then, for any $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$, we have*

$$4\psi_\tau(x, y) \geq 2\|[\phi_\tau(x, y)]_+\|^2 \geq \frac{(4-\tau)^2}{4}\left[\|(-x)_+\|^2 + \|(-y)_+\|^2\right]$$

**Proof.** Note that $z(x,y) - (x + \frac{\tau-2}{2}y) \in \mathcal{K}^l$ and $z(x,y) - (y + \frac{\tau-2}{2}x) \in \mathcal{K}^l$. Following the same proof line as Lemma 8 of [7] immediately yields the desired result. $\square$

**Lemma 5.3** *Let $\psi_\tau$ be defined as in (12). For any sequence $\{(x^k, y^k)\} \subseteq \mathbb{R}^l \times \mathbb{R}^l$, let $\lambda_1^k \leq \lambda_2^k$ and $\mu_1^k \leq \mu_2^k$ denote the spectral values of $x^k$ and $y^k$, respectively.*

**(a)** *If $\lambda_1^k \to -\infty$ or $\mu_1^k \to -\infty$, then $\psi_\tau(x^k, y^k) \to +\infty$.*

**(b)** *If $\{\lambda_1^k\}$ and $\{\mu_1^k\}$ are bounded below, but $\lambda_2^k \to +\infty$, $\mu_2^k \to +\infty$, and $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \nrightarrow 0$, then $\psi_\tau(x^k, y^k) \to +\infty$.*

**Proof.** Part (a) is direct by Lemma 5.2 and the following fact that

$$\left\|(-x^k)_+\right\|^2 = \frac{1}{2}\sum_{i=1}^{2}\left(\min\{0, \lambda_i^k\}\right)^2, \quad \left\|(-y^k)_+\right\|^2 = \frac{1}{2}\sum_{i=1}^{2}\left(\min\{0, \mu_i^k\}\right)^2.$$

We next prove part (b) by contradiction. Suppose that $\{\psi_\tau(x^k, y^k)\}$ is bounded. Since

$$x^k + y^k = z^k - \phi_\tau(x^k, y^k) \quad \forall k,$$

where $z^k = z(x^k, y^k)$ with $z(x,y)$ defined as in (16). Squaring the two sides of the last equality then yields that

$$(4-\tau)x^k \circ y^k = -2z^k \circ \phi_\tau(x^k, y^k) + (\phi_\tau(x^k, y^k))^2. \tag{40}$$

Noting that, for each $k$,

$$0 \leq \frac{z_1^k}{\|x^k\|\|y^k\|} \leq \frac{\sqrt{2w_1^k}}{\|x^k\|\|y^k\|} = \sqrt{\frac{\|x^k\|^2 + \|y^k\|^2 + (\tau-2)(x^k)^T y^k}{\|x^k\|^2\|y^k\|^2}},$$

we can verify that $\lim_{k\to+\infty} \dfrac{z_1^k}{\|x^k\|\|y^k\|} = 0$. Combining with $\dfrac{z^k}{\|x^k\|\|y^k\|} \in \mathcal{K}^l$ yields

$$\lim_{k\to+\infty} \frac{z^k}{\|x^k\|\|y^k\|} = 0.$$

Using equation (40) and the boundedness of $\{\phi_\tau(x^k, y^k)\}$, it then follows that

$$\lim_{k\to+\infty} \frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} = 0,$$

which clearly contradicts the given assumption. The proof is complete. $\square$

Now using Lemma 5.3 and the same arguments as Proposition 5.2 of [19], we can establish the boundedness of the level sets of $\Psi_\tau(\zeta)$ for the SOCCP (3) under the assumption that $F$ has the uniform Cartesian $P$-property and satisfies the following condition:

**Condition A.** For any sequence $\{\zeta^k\} \subseteq \mathbb{R}^n$ such that $\|\zeta^k\| \to +\infty$, if there exists $i \in \{1, \ldots, m\}$ such that $\lambda_1(\zeta_i^k), \lambda_1(F_i(\zeta^k)) > -\infty$ and $\lambda_2(\zeta_i^k), \lambda_2(F_i(\zeta^k)) \to +\infty$, then

$$\limsup_{k \to +\infty} \left\langle \frac{\zeta_i^k}{\|\zeta_i^k\|}, \frac{F_i(\zeta^k)}{\|F_i(\zeta^k)\|} \right\rangle > 0.$$

Consequently, we extend the coerciveness of the FB merit function to the function $\Psi_\tau$.

**Proposition 5.2** *For the SOCCP (3), if $F : \mathbb{R}^n \to \mathbb{R}^n$ has the uniform Cartesian $P$-property and satisfies Condition A, then the merit function $\Psi_\tau$ has bounded level sets.*

# 6 Algorithm and numerical results

The previous discussions show that the SOC complementarity function $\phi_\tau$ possesses all nice features of the FB SOC complementarity function. In this section, we test the numerical performance of the class of SOC functions by using the semismooth Newton method proposed by De Luca, Facchinei and Kanzow [16], which is described as follows.

**Algorithm 6.1:**

**Step 0.** Given a $\tau \in (0, 4)$ and a starting point $\zeta^0 \in \mathbb{R}^n$, and choose $\gamma > 0$, $p > 2$, $\rho \in (0, 1)$, $\sigma \in (0, 1/2)$, and $\varepsilon > 0$. Set $k := 0$.

**Step 1.** If $\|\nabla \Psi_\tau(\zeta^k)\| \leq \varepsilon$, then stop.

**Step 2.** Select an element $W_k \in \partial_B \Phi_\tau(\zeta^k)$. Find a solution $d^k \in \mathbb{R}^n$ of the linear system

$$W_k d = -\Phi_\tau(\zeta^k). \tag{41}$$

If the system is not solvable or if the descent condition

$$\nabla \Psi_\tau(\zeta^k)^T d^k \leq -\gamma \|d^k\|^p$$

is not satisfied, set $d^k := -\nabla \Psi_\tau(\zeta^k)$.

**Step 3.** Let $m_k$ be the smallest nonnegative integer $m$ such that

$$\Psi_\tau(\zeta^k + \rho^m d^k) \leq \Psi_\tau(\zeta^k) + \sigma \rho^m \nabla \Psi_\tau(\zeta^k)^T d^k, \tag{42}$$

and set $\zeta^{k+1} := \zeta^k + \rho^{m_k} d^k$, $k := k + 1$, and go to Step 1.

The global and local convergence properties of Algorithm 6.1 are summarized in the following theorem, in which we implicitly assume that the termination parameter $\varepsilon$ equals to 0, i.e. the algorithm generates an infinite sequence.

**Theorem 6.1** *Suppose that $\{\zeta^k\}$ is a sequence generated by Algorithm 6.1. Then,*

**(a)** *each accumulation point of $\{\zeta^k\}$ is a stationary point of the merit function $\Psi_\tau$.*

**(b)** *If $\zeta^*$ is an isolated accumulation point of $\{\zeta^k\}$, then the entire sequence $\{\zeta^k\}$ converges to $\zeta^*$.*

**(c)** *If $\zeta^*$ is an accumulation point such that $\zeta^*$ is a strict complementarity solution and $F(\zeta)$ and $G(\zeta)$ at $\zeta^*$ satisfy the conditions of Theorem 4.1. Then,*

- **(i)** *the search direction $d^k$ is eventually given by the solution of (41);*
- **(ii)** *the sequence $\{\zeta^k\}$ converges to $\zeta^*$ Q-superlinearly;*
- **(iii)** *if, in addition, $F'$ and $G'$ are Lipschitz continuous at $\zeta^*$, then the rate of convergence is Q-quadratic.*

**Proof.** Since the proofs are similar to that of [14, Theorem 4.2] or [16, Theorem 3.1] by the results obtained in Section 3–5, we here omit them. $\square$

Note that Theorem 6.1 (a) and (b) only gives global convergence results to stationary points of the merit function $\Psi_\tau$ whereas we are much concerned with finding a global minimizer of $\Psi_\tau$ and consequently a solution of the SOCCP. Fortunately, Proposition 5.1 provides a rather weak condition to guarantee such a stationary point is a solution of the SOCCP. The existence of an accumulation point and thus of a stationary point of $\Psi_\tau$ is guaranteed by Proposition 5.2. From Definition 2.2, we see that the assumption from Proposition 5.2 may be satisfied by some monotone SOCCPs, and our numerical experiences also verify this fact.

In what follows, we report the computational experience with solving some linear SOCPs, which correspond to the SOCP (4) with $g(x) = c^T x$, by Algorithm 6.1. From the introduction, the class of problems can be reformulated as the SOCCP with $F(\zeta)$ and $G(\zeta)$ given as in (5). The test instances are taken from the DIMACS Implementation Challenge library and described in Table 1 in which, the notation [4 $\times$ 1; 1 $\times$ 123; 838 $\times$ 3] in the column of structure of SOCs means that $\mathcal{K}$ consists of the product of four $\mathcal{K}^1$, one $\mathcal{K}^{123}$, and 838 $\mathcal{K}^3$, and $m \times n$ specifies the size of the matrix $A$.

All experiments were done at a PC with 2.8GHz CPU and 512MB memory. The computer codes were all written in Matlab 6.5. During the experiments, we replaced

Table 1: Set of test problems

| No. | Problem Names | n | m | # of nonzero elts of matrix $A$ | structure of SOCs |
|---|---|---|---|---|---|
| 1 | nb | 2383 | 123 | 192439 | [4 × 1; 793 × 3] |
| 2 | nb-L1 | 3176 | 915 | 193104 | [797 × 1; 793 × 3] |
| 3 | nb-L2-bessel | 2641 | 123 | 209924 | [4 × 1; 1 × 123; 838 × 3] |

the standard Armijo linesearch rule in Algorithm 6.1 with a nonmonotone linesearch as described in [12]. The motivation of adopting this variant is to circumvent very small stepsizes which will lead to the difficulty in the solution of SOCCPs. In addition, the nonmonotone linesearch was proved in [12] to have better numerical performance for the unconstrained minimization of smooth functions. Specifically, we computed the smallest nonnegative integer $m$ such that

$$\Psi_\tau(\zeta^k + \rho^m d^k) \leq \mathcal{W}_k + \sigma\rho^m \nabla\Psi_\tau(\zeta^k)^T d^k,$$

where

$$\mathcal{W}_k := \max\left\{\Psi_\tau(\zeta^j) \mid j = k - m_k, \ldots, k\right\},$$

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, the following parameters were used in the algorithm:

$$\gamma = 10^{-8}, \quad p = 2.1, \quad \rho = 0.5, \quad \sigma = 10^{-4}, \quad \hat{m} = 5 \quad \text{and} \quad s = 5.$$

The starting point was chosen to be $\zeta^0 = 0$. The Algorithm was terminated whenever one of the following conditions is satisfied

$$\max\left\{|F(\zeta^k)^T G(\zeta^k)|, \Psi_\tau(\zeta^k)\right\} \leq 10^{-5}, \quad k > 200, \quad \alpha_k := \rho^{m_k} < 10^{-15}. \tag{43}$$

The term $|F(\zeta^k)^T G(\zeta^k)|$ in the first condition aims to obtain a solution with a favorable dual gap. In addition, it also helps to stop the algorithm when the decrease of $\Psi_\tau(\zeta)$ has little advantage in reducing the dual gap.

Numerical results are summarized in Table 2, where **NF** and $k$ denote the number of function evaluations and iterations for solving each test problem, **Obj.** means the objective value of the test problems at the final iteration, and **Time** denotes the CPU

Table 2: Numerical results of Algorithm 6.1 for linear SOCPs with a different $\tau$

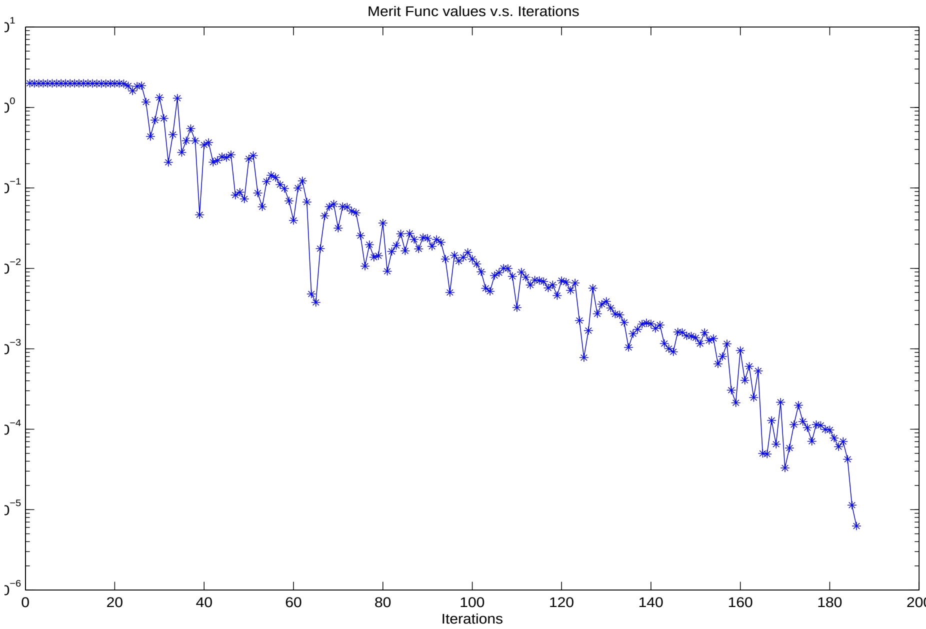
| No. | $\tau$ | **Obj.** | **NF** | $k$ | **Time** | $\tau$ | **Obj.** | **NF** | $k$ | **Time** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.5 | –0.0507101 | 177 | 59 | 644.1 | 1.5 | –0.0507184 | 75 | 28 | 303.2 |
| | 2.0 | –0.0507130 | 85 | 29 | 313.8 | 2.5 | –0.0507088 | 66 | 32 | 342.2 |
| | 3.0 | –0.0507256 | 74 | 29 | 311.2 | 3.5 | –0.0507091 | 63 | 38 | 406.0 |
| 2 | 0.5 | – | – | $> 200$ | – | 1.5 | –13.0122435 | 144 | 87 | 1587.4 |
| | 2.0 | –13.0120761 | 219 | 112 | 2047.2 | 2.5 | –13.0121923 | 227 | 112 | 2149.3 |
| | 3.0 | –13.0121999 | 393 | 197 | 3762.1 | 3.5 | – | – | $> 200$ | – |
| 3 | 0.5 | –0.1025695 | 35 | 18 | 235.3 | 1.5 | –0.1025728 | 23 | 10 | 128.6 |
| | 2.0 | –0.1025766 | 15 | 9 | 113.7 | 2.5 | –0.1025706 | 17 | 10 | 125.6 |
| | 3.0 | –0.1025695 | 21 | 14 | 181.4 | 3.5 | –0.1025695 | 39 | 29 | 364.4 |

time in second that the iterates satisfy the termination condition.

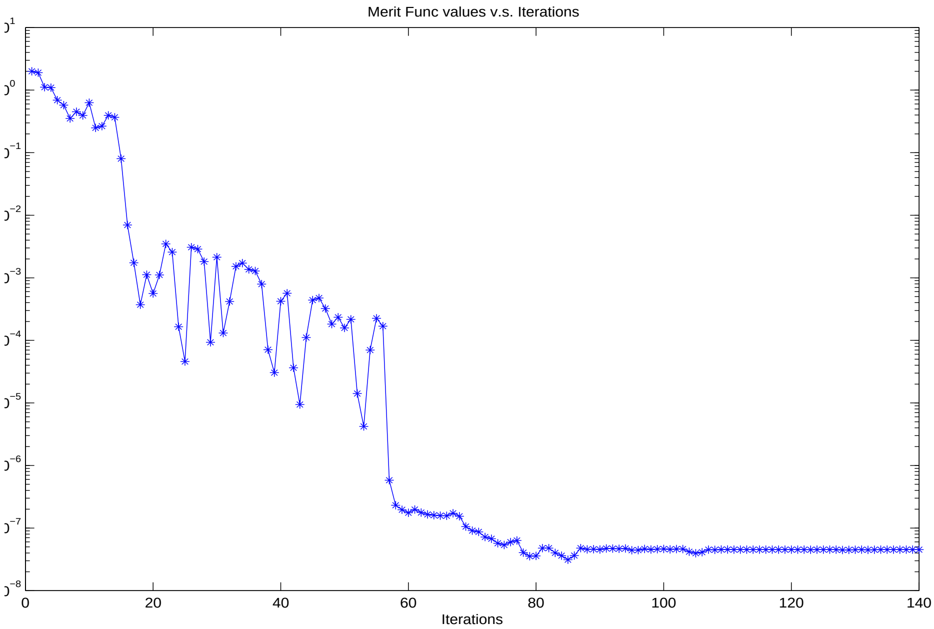
From Table 2, we see that the semismooth Newton method proposed can solve all test problems with $\tau \in [1.5, 3]$ and has better numerical performance with $\tau \in [1.5, 2.5]$ for all test problems. When $\tau$ tends to 0 or 4, the number of iteration has a remarkable increase. For problem "nb-L1", Algorithm 6.1 requires much more iterations. After a check, the solution of this problem does not satisfy strict complementarity, and now we are not clear whether this takes charge in much more iterations. We also observe that the parameter $\tau$ close to 4 often gives a better global convergence, whereas the parameter $\tau$ close to 0 leads to a fast local convergence. Figure 1 below displays the convergence of $\Psi_\tau$ for problem "nb" with $\tau = 0.1$ and $\tau = 3.9$, respectively. The performance of $\Psi_\tau$ coincides with the case described by [14] for the NCPs, which is very important for the use of the class of SOC complementarity functions. Based on this feature of $\phi_\tau$, we may adopt a dynamic choice of $\tau$ in the algorithm by following a line similar to [14].

# 7 Conclusions

In this paper, we continued to investigate the properties of the one-parametric class of SOC complementarity functions $\phi_\tau$, which includes the FB SOC complementarity




(a) $\tau = 0.1$




(b) $\tau = 3.9$

Figure 1: The convergence of Algorithm 6.1 with different $\tau$ for ‘nb’.

function and the natural residual SOC complementarity function as a special case. We showed that $\phi_\tau$ is globally Lipschitz continuous and strongly semismooth and characterized its B-subdifferential at any point. Furthermore, for the induced merit function $\Psi_\tau$, we provided a weaker condition than [6] to guarantee every stationary point to be a solution of the SOCCP, and proved that it has bounded level sets for the SOCCP (3) if the mapping has the uniform Cartesian $P$-property and satisfies Condition A. Combining with the results of [6], we thus extended most of favorable properties of the class of complementarity functions for the NCP to the setting of the SOCCP.

A semismooth Newton method is also proposed by the nonsmooth reformulation (10) involving the class of SOC complementarity functions. The superlinear convergence of the algorithm is established by requiring the solution to be strict complementarity. The condition is stronger than the counterpart in the NCPs, and we will consider to weaken this condition in the future research work.

**Acknowledgements** The authors would like to thank the two anonymous referees for their helpful comments which improved the presentation of this paper.

# References

[1] F. Alizadeh and D. Goldfarb (2003), *Second-order cone programming*, Mathematical Programming, vol. 95, pp. 3–51.

[2] E. D. Andersen, C. Roos, and T. Terlaky (2003), *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming Ser. B, vol. 95, pp. 249–277.

[3] X. Chen and H. Qi (2006), *Cartesian P-proeprty and its applications to the semidefinite linear complementarity problem*, Mathematical Programming, vol. 106, pp. 177–201.

[4] X.-D. Chen, D. Sun, and J. Sun (2003), *Complementarity functions and numerical experiments for second-order cone complementarity problems*, Computational Optimization and Applications, vol. 25, pp. 39–56.

[5] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).

[6] J.-S. Chen and S.-H. Pan (2007), *A one-parametric class of merit functions for the second-order cone complementarity problem*, Submitted to Computational Optimization and Applications.

[7] J.-S. Chen and P. Tseng (2005), *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Mathematical Programming, vol. 104, pp. 293–327.

[8] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[9] M. Fukushima, Z.-Q. Luo, and P. Tseng (2002), *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimization, vol. 12, pp. 436–460.

[10] A. Fischer (1992), *A special Newton-type optimization methods*, Optimization, vol. 24, pp. 269-284.

[11] A. Fischer (1997), *Solution of the monotone complementarity problem with locally Lipschitzian functions*, Mathematical Programming, vol. 76, pp. 513-532.

[12] L. Grippo, F. Lampariello and S. Lucidi (1986), *A nonmonotone line search technique for Newton's method*, SIAM Journal on Numerical Analysis, 1986, vol. 23, pp. 707–716.

[13] S. Hayashi, N. Yamashita, and M. Fukushima (2005), *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM Journal of Optimization, vol. 15, pp. 593–615.

[14] C. Kanzow and H. Kleinmichel (1998), *A new class of semismooth Newton-type methods for nonlinear complementarity problems*, Computational Optimization and Applications, vol. 11, pp. 227–251.

[15] C. Kanzow and M. Fukushima (2006), *Semismooth metods for linear and nonlinear second-order cone programs*, Technical Report, Department of Applied Mathematics and Physics, Kyoto University.

[16] T. De. Luca, F. Facchinei and C. Kanzow (1996), *A semismooth equation approach to the solution of nonlinear complementarity problems*, Mathematical Programming, vol. 75, pp. 407–439.

[17] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret (1998), *Application of second-order cone programming*, Linear Algebra and its Applications, vol. 284, pp. 193–228.

[18] R. D. C. Monteiro and T. Tsuchiya (2000) *Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions*, Mathematical Programming, vol. 88, pp. 61–83.

[19] S.-H. Pan and J.-S. Chen (2006), *A damped Gauss-Newton method for the second-order cone complementarity problem*, Accepted by Applied Mathematics and Optimization.

[20] L. Qi and J. Sun (1993) *A nonsmooth version of Newton's method*, Mathematical Programming, vol. 58, pp. 353–367.

[21] L. Qi (1993) *Convergence analysis of some algorithms for solving nonsmooth equations*, Mathematics of Operations Research, vol. 18, pp. 227–244.

[22] D. Sun and L.-Q. Qi, *On NCP-functions* (1999), Computational Optimization and Applications, vol. 13, pp. 201-220.

[23] D. Sun and J. Sun (2005), *Strong semismoothness of the Fischer-Burmeister SDC and SOC complmentarity functions*, Mathematical Programming, vol. 103, pp. 575–581.

[24] T. Tsuchiya (1999), *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, vol. 11, pp. 141–182.

## Appendix

**Lemma A.1** *The function $\hat{z}(x,y,\epsilon)$ defined by (24) for any $\epsilon > 0$ is continuously differentiable everywhere, and there exists a scalar $C > 0$ such that*

$$\|\nabla_x \hat{z}(x,y,\epsilon)\|_F \le C, \quad \|\nabla_y \hat{z}(x,y,\epsilon)\|_F \le C \tag{44}$$

*for all $(x,y) \in \mathbb{R}^l \times \mathbb{R}^l$, where $\|A\|_F$ denotes the Frobenius norm of the matrix $A$.*

**Proof.** Since $(x-y)^2 + \tau(x \circ y) + \epsilon e \in \mathrm{int}(\mathcal{K}^l)$ for any $(x,y) \in \mathbb{R}^l \times \mathbb{R}^l$ and $\epsilon > 0$, by Lemma 2.1 the function $\hat{z}(x,y,\epsilon)$ is continuously differentiable everywhere and

$$\nabla_x \hat{z}(x,y,\epsilon) = \left(L_x + \frac{\tau-2}{2}L_y\right)L_{\hat{z}}^{-1}, \quad \nabla_y \hat{z}(x,y,\epsilon) = \left(L_y + \frac{\tau-2}{2}L_x\right)L_{\hat{z}}^{-1}. \tag{45}$$

We next prove the bound in (44) by the two cases: $w_2 \neq 0$ and $w_2 = 0$. Let

$$\hat{w} = (\hat{w}_1, \hat{w}_2) = \hat{w}(x,y,\epsilon) \ := \ (x-y)^2 + \tau(x \circ y) + \epsilon e.$$

Case (1). $w_2 \neq 0$. Then, $\hat{w}_2 \neq 0$ since $\hat{w}_2 = w_2$. Let $g = (g_1, g_2) := x + \frac{\tau-2}{2}y$. By (45) and the formula of $L_{\hat{z}}^{-1}$ given by (19), we can compute that

$$\nabla_x \hat{z}(x,y,\epsilon) = \begin{pmatrix} \hat{b}g_1 + \hat{c}g_2^T\bar{w}_2 & \hat{c}g_1\bar{w}_2 + \hat{a}g_2^T + (\hat{b}-\hat{a})g_2^T\bar{w}_2\bar{w}_2^T \\ \hat{b}g_2 + \hat{c}g_1\bar{w}_2 & \hat{c}g_2\bar{w}_2^T + \hat{a}g_1 I + (\hat{b}-\hat{a})g_1\bar{w}_2\bar{w}_2 \end{pmatrix},$$

where $\hat{a}, \hat{b}$ and $\hat{c}$ are defined as in Lemma 2.1 with $w = \hat{w}$. Notice that

$$g_1 = x_1 + \frac{\tau-2}{2}y_1,\ g_2 = x_2 + \frac{\tau-2}{2}y_2;\ \ \lambda_1(\hat{w}) = \lambda_1(w) + \epsilon,\ \lambda_2(\hat{w}) = \lambda_2(w) + \epsilon.$$

Using the expression of $\hat{a}, \hat{b}$ and $\hat{c}$ and the result of Lemma 2.3 then yields that

$$\begin{aligned}
\left|\hat{b}g_1 + \hat{c}g_2^T\bar{w}_2\right| &\le \frac{1}{2\sqrt{\lambda_2(w)}}\left|g_1 + g_2^T\bar{w}_2\right| + \frac{1}{2\sqrt{\lambda_1(w)}}\left|g_1 - g_2^T\bar{w}_2\right| \ \le\ 1,\\
\left\|\hat{c}g_1\bar{w}_2^T + \hat{b}g_2^T\bar{w}_2\bar{w}_2^T\right\| &\le \frac{1}{2\sqrt{\lambda_2(w)}}\left|g_1 + g_2^T\bar{w}_2\right| + \frac{1}{2\sqrt{\lambda_1(w)}}\left|g_1 - g_2^T\bar{w}_2\right| \ \le\ 1,\\
\left\|\hat{a}g_2^T - \hat{a}g_2^T\bar{w}_2\bar{w}_2^T\right\| &\le \frac{\|2g_2\|}{\sqrt{\|x\|^2+\|y\|^2+(\tau-2)x^Ty}}(1+\|\bar{w}_2\|) \ \le\ 4,\\
\left\|\hat{b}g_2 + \hat{c}g_1\bar{w}_2\right\| &\le \frac{1}{2\sqrt{\lambda_2(w)}}\|g_2 + g_1\bar{w}_2\| + \frac{1}{2\sqrt{\lambda_1(w)}}\|g_2 - g_1\bar{w}_2\| \ \le\ 1,\\
\left\|\hat{c}g_2\bar{w}_2^T + \hat{b}g_1\bar{w}_2\bar{w}_2^T\right\|_F &\le \frac{1}{2\sqrt{\lambda_2(w)}}\|g_2 + g_1\bar{w}_2\| + \frac{1}{2\sqrt{\lambda_1(w)}}\|g_2 - g_1\bar{w}_2\| \ \le\ 1,\\
\left\|\hat{a}g_1I - \hat{a}g_1\bar{w}_2\bar{w}_2^T\right\|_F &\le \frac{2|g_1|}{\sqrt{\|x\|^2+\|y\|^2+(\tau-2)x^Ty}}\cdot\left\|I - \bar{w}_2\bar{w}_2^T\right\|_F \ \le\ 2(l-1).
\end{aligned}$$

The above inequalities imply that the first inequality in (44) holds under this case.

Case (2). $w_2 = 0$. In this case, from Lemma 2.1 it follows that

$$\nabla_x\hat{z}(x,y,\epsilon) = \frac{1}{\sqrt{\hat{w}_1}}\left(L_x + \frac{\tau-2}{2}L_y\right) = \frac{1}{\sqrt{\hat{w}_1}}L_g.$$

Since $\hat{w}_1 = \|x+\frac{\tau-2}{2}y\|^2 + \frac{\tau(4-\tau)}{4}\|y\|^2 + \epsilon$, we have $|g_1|/\sqrt{\hat{w}_1} \le 1$ and $\|g_2\|/\sqrt{\hat{w}_1} \le 1$, which implies the first inequality in (44). Thus, we complete the proof for the first inequality. By the symmetry of $x$ and $y$ in $\hat{z}(x,y,\epsilon)$, the second inequality clearly holds. $\square$

### Proof of Proposition 3.2

**Proof.** Throughout the proof, let $D_{\phi_\tau}$ denote the set of points where $\phi_\tau$ is differentiable. Recall that this set is characterized by Proposition 3.1 (a). Write

$$\phi'_{\tau,x}(x,y) = \nabla_x\phi_\tau(x,y)^T \ \text{ and } \ \phi'_{\tau,y}(x,y) = \nabla_y\phi_\tau(x,y)^T.$$

From Proposition 3.1 (a), it then follows that for any $(x,y) \in D_{\phi_\tau}$,

$$\phi'_{\tau,x}(x,y) = L_z^{-1}L_{x+\frac{\tau-2}{2}y} - I,\ \ \phi'_{\tau,x}(x,y) = L_z^{-1}L_{y+\frac{\tau-2}{2}x} - I. \tag{46}$$

Moreover, we observe from (19) that, when $w_2 \neq 0$, $L_z^{-1}$ can be expressed as the sum of

$$L_1(w) = \frac{1}{2\sqrt{\lambda_1(w)}}\begin{pmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{pmatrix}$$

and

$$L_2(w)=\frac{1}{2\sqrt{\lambda_2(w)}}\left(\begin{array}{cc}1 & \bar{w}_2^T\\ \bar{w}_2 & \dfrac{4\sqrt{\lambda_2(w)}(I-\bar{w}_2\bar{w}_2^T)}{\sqrt{\lambda_2(w)}+\sqrt{\lambda_1(w)}}+\bar{w}_2\bar{w}_2^T\end{array}\right),$$

and consequently $\phi'_{\tau,x}$ and $\phi'_{\tau,y}$ in (46) can be rewritten as

$$\begin{aligned}\phi'_{\tau,x}(x,y) &= (L_1(w)+L_2(w))L_{x+\frac{\tau-2}{2}y}-I,\\ \phi'_{\tau,x}(x,y) &= (L_1(w)+L_2(w))L_{y+\frac{\tau-2}{2}x}-I.\end{aligned}\tag{47}$$

(a) Under the given assumption, $\phi_\tau$ is continuously differentiable at $(x,y)$ by Proposition 3.1 (a). Consequently, the B-subdifferential $\partial_B\phi_\tau(x,y)$ consists of only one element,

$$\phi'_\tau(x,y)=\left[\phi'_{\tau,x}(x,y)\ \ \phi'_{\tau,x}(x,y)\right].$$

Substituting the formulas in (46) into it, we immediately obtain the conclusion.

(b) Assume that $(x,y)\neq(0,0)$ satisfies $(x-y)^2+\tau(x\circ y)\in \mathrm{bd}(\mathcal{K}^l)$. Let $\{(x^k,y^k)\}\subseteq D_{\phi_\tau}$ be an arbitrary sequence converging to $(x,y)$. Let $w^k=(w_1^k,w_2^k)=w(x^k,y^k)$ and $z^k=z(x^k,y^k)$, where $w(x,y)$ and $z(x,y)$ are defined as in (16). From the given assumption on $(x,y)$, we have $w\in\mathrm{bd}(\mathcal{K}^l)$ and $w_1>0$, which means that $\lambda_2(w)>\lambda_1(w)=0$ and $\|w_2\|=w_1>0$. Hence, we assume without loss of generality that $w_2^k\neq 0$ for each $k$. Using the formulas in (47), it then follows that

$$\begin{aligned}\phi'_{\tau,x}(x^k,y^k) &= \left(L_1(w^k)+L_2(w^k)\right)L_{x^k+\frac{\tau-2}{2}y^k}-I,\\ \phi'_{\tau,y}(x^k,y^k) &= \left(L_1(w^k)+L_2(w^k)\right)L_{y^k+\frac{\tau-2}{2}x^k}-I.\end{aligned}\tag{48}$$

Notice that $\lim_{k\to+\infty}\lambda_2(w^k)=2w_1>0$ and $\lim_{k\to+\infty}\lambda_1(w^k)=\lambda_1(w)=0$, which, together with $\lim_{k\to+\infty}L_{x^k}=L_x$, $\lim_{k\to+\infty}L_{y^k}=L_y$ and $\lim_{k\to+\infty}w_2^k=w_2$, yields that

$$\begin{aligned}\lim_{k\to+\infty}L_2(w^k)L_{x^k+\frac{\tau-2}{2}y^k} &= C(w)\left(L_x+\frac{\tau-2}{2}L_y\right),\\ \lim_{k\to+\infty}L_2(w^k)L_{y^k+\frac{\tau-2}{2}x^k} &= C(w)\left(L_y+\frac{\tau-2}{2}L_x\right),\end{aligned}\tag{49}$$

where $C(w)$ is defined as follows:

$$C(w)\ =\ \frac{1}{2\sqrt{2w_1}}\left(\begin{array}{cc}1 & \bar{w}_2^T\\ \bar{w}_2 & 4I-3\bar{w}_2\bar{w}_2^T\end{array}\right)\quad\text{with}\ \ \bar{w}_2=\frac{w_2}{\|w_2\|}.$$

In addition, by a simple computation, we have that

$$\begin{aligned}L_1(w^k)L_{x^k+\frac{\tau-2}{2}y^k} &= \frac{1}{2}\left(\begin{array}{cc}u_1^k & (u_2^k)^T\\ -u_1^k\bar{w}_2^k & -\bar{w}_2^k(u_2^k)^T\end{array}\right),\\ L_1(w^k)L_{y^k+\frac{\tau-2}{2}x^k} &= \frac{1}{2}\left(\begin{array}{cc}v_1^k & (v_2^k)^T\\ -v_1^k\bar{w}_2^k & -\bar{w}_2^k(v_2^k)^T\end{array}\right),\end{aligned}$$

where $\bar{w}_2^k = w_2^k/\|w_2^k\|$ for each $k$, and

$$\begin{aligned}
u_1^k &= \frac{1}{\sqrt{\lambda_1(w^k)}}\left[\left(x_1^k + \frac{\tau-2}{2}y_1^k\right) - \left(x_2^k + \frac{\tau-2}{2}y_2^k\right)^T \bar{w}_2^k\right],\\
u_2^k &= \frac{1}{\sqrt{\lambda_1(w^k)}}\left[\left(x_2^k + \frac{\tau-2}{2}y_2^k\right) - \left(x_1^k + \frac{\tau-2}{2}y_1^k\right) \bar{w}_2^k\right],\\
v_1^k &= \frac{1}{\sqrt{\lambda_1(w^k)}}\left[\left(y_1^k + \frac{\tau-2}{2}x_1^k\right) - \left(y_2^k + \frac{\tau-2}{2}x_2^k\right)^T \bar{w}_2^k\right],\\
v_2^k &= \frac{1}{\sqrt{\lambda_1(w^k)}}\left[\left(y_2^k + \frac{\tau-2}{2}x_2^k\right) - \left(y_1^k + \frac{\tau-2}{2}x_1^k\right) \bar{w}_2^k\right].
\end{aligned}$$

By Lemma 2.3, $|u_1^k| \le \|u_2^k\| \le 1$ and $|v_1^k| \le \|v_2^k\| \le 1$. So, taking the limit (possibly on a subsequence) on $L_1(w^k)L_{x^k+\frac{\tau-2}{2}y^k}$ and $L_1(w^k)L_{y^k+\frac{\tau-2}{2}x^k}$, we have

$$\begin{aligned}
L_1(w^k)L_{x^k+\frac{\tau-2}{2}y^k} &\to \frac{1}{2}\begin{pmatrix} u_1 & u_2^T \\ -u_1\bar{w}_2 & -\bar{w}_2 u_2^T \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T\\
L_1(w^k)L_{y^k+\frac{\tau-2}{2}x^k} &\to \frac{1}{2}\begin{pmatrix} v_1 & v_2^T \\ -v_1\bar{w}_2 & -\bar{w}_2 v_2^T \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T
\end{aligned} \tag{50}$$

for some $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ with $|u_1| \le \|u_2\| \le 1$ and $|v_1| \le \|v_2\| \le 1$, where $\bar{w}_2 = w_2/\|w_2\|$. In fact, $u$ and $v$ are some accumulation point of the sequences $\{u^k\}$ and $\{v^k\}$, respectively. From (48)–(50), we obtain that

$$\begin{aligned}
\phi'_{\tau,x}(x^k, y^k) &\to C(w)\left(L_x + \frac{\tau-2}{2}L_y\right) + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T - I,\\
\phi'_{\tau,y}(x^k, y^k) &\to C(w)\left(L_y + \frac{\tau-2}{2}L_x\right) + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T - I.
\end{aligned}$$

This shows that as $k \to +\infty$, $\phi'_\tau(x^k, y^k) \to [V_x - I \;\; V_y - I]$ with $V_x, V_y$ satisfying (25).

(c) Assume $(x, y) = (0, 0)$. Let $\{(x^k, y^k)\} \subseteq D_{\phi_\tau}$ be an arbitrary sequence converging to $(x, y)$. Let $w^k = (w_1^k, w_2^k)$ and $z^k$ be defined as in Case (b). From the given assumptions, we have $w = 0$. Therefore, we may assume without any loss of generality that $w_2^k = 0$ for all $k$ or $w_2^k \neq 0$ for all $k$. We proceed the arguments by the two cases.

<u>Case (1)</u>: $w_2^k = 0$ for all $k$. From equation (46) and Lemma 2.1, it follows that

$$\begin{aligned}
\phi'_{\tau,x}(x^k, y^k) &= \frac{1}{\sqrt{w_1^k}}\begin{pmatrix} x_1^k + \frac{\tau-2}{2}y_1^k & \left(x_2^k + \frac{\tau-2}{2}y_2^k\right)^T \\ x_2^k + \frac{\tau-2}{2}y_2^k & \left(x_1^k + \frac{\tau-2}{2}y_1^k\right) I \end{pmatrix} - I,\\
\phi'_{\tau,y}(x^k, y^k) &= \frac{1}{\sqrt{w_1^k}}\begin{pmatrix} y_1^k + \frac{\tau-2}{2}x_1^k & \left(y_2^k + \frac{\tau-2}{2}x_2^k\right)^T \\ y_2^k + \frac{\tau-2}{2}x_2^k & \left(y_1^k + \frac{\tau-2}{2}x_1^k\right) I \end{pmatrix} - I.
\end{aligned}$$

Since

$$w_1^k = \|x^k + \frac{\tau-2}{2} y^k\|^2 + \frac{\tau(4-\tau)}{4}\|y^k\|^2 = \|y^k + \frac{\tau-2}{2} x^k\|^2 + \frac{\tau(4-\tau)}{4}\|x^k\|^2,$$

every element in the above $\phi'_{\tau,x}(x^k, y^k)$ and $\phi'_{\tau,y}(x^k, y^k)$ are bounded. Thus, taking limit (possibly on a subsequence) on $\phi'_{\tau,x}(x^k, y^k)$ and $\phi'_{\tau,y}(x^k, y^k)$, respectively, gives

$$\nabla_x \phi_\tau(x^k, y^k) \to \begin{pmatrix} \hat{u}_1 & \hat{u}_2^T \\ \hat{u}_2 & \hat{u}_1 I \end{pmatrix} - I, \quad \nabla_y \phi_\tau(x^k, y^k) \to \begin{pmatrix} \hat{v}_1 & \hat{v}_2^T \\ \hat{v}_2 & \hat{v}_1 I \end{pmatrix} - I$$

for some $\hat{u} = (\hat{u}_1, \hat{u}_2), \hat{v} = (\hat{v}_1, \hat{v}_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ satisfying $\|\hat{u}\| \le 1, \|\hat{v}\| \le 1$ and $\hat{u}_1\hat{v}_2 + \hat{v}_1\hat{u}_2 = 0$. This shows that $\phi'_\tau(x^k, y^k) \to [V_x - I \;\; V_y - I]$ with $V_x \in \{L_{\hat{u}}\}, V_y \in \{L_{\hat{v}}\}$.

Case (2): $w_2^k \neq 0$ for all $k$. Now $\phi'_{\tau,x}(x^k, y^k)$ and $\phi'_{\tau,y}(x^k, y^k)$ are given as in (48). Using the same arguments as part (b) and noting that $\{\bar{w}_2^k\}$ is bounded, we have

$$L_1(w^k) L_{x^k + \frac{\tau-2}{2} y^k} \to \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T, \quad L_1(w^k) L_{y^k + \frac{\tau-2}{2} x^k} \to \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T \qquad (51)$$

for some vectors $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ satisfying $|u_1| \le \|u_2\| \le 1$ and $|v_1| \le \|v_2\| \le 1$, and $\bar{w}_2 \in \mathbb{R}^{l-1}$ satisfying $\|\bar{w}_2\| = 1$. We next compute the limit of $L_2(w^k) L_{x^k + \frac{\tau-2}{2} y^k}$ and $L_2(w^k) L_{y^k + \frac{\tau-2}{2} x^k}$. By the definition of $L_2(w)$,

$$L_2(w^k) L_{x^k + \frac{\tau-2}{2} y^k} = \frac{1}{2} \begin{pmatrix} \xi_1^k & (\xi_2^k)^T \\ \xi_1^k \bar{w}_2^k + 4(I - \bar{w}_2^k(\bar{w}_2^k)^T) s_2^k & \bar{w}_2^k (\xi_2^k)^T + 4(I - \bar{w}_2^k(\bar{w}_2^k)^T) s_1^k \end{pmatrix},$$

$$L_2(w^k) L_{y^k + \frac{\tau-2}{2} x^k} = \frac{1}{2} \begin{pmatrix} \eta_1^k & (\eta_2^k)^T \\ \eta_1^k \bar{w}_2^k + 4(I - \bar{w}_2^k(\bar{w}_2^k)^T) \omega_2^k & \bar{w}_2^k (\eta_2^k)^T + 4(I - \bar{w}_2^k(\bar{w}_2^k)^T) \omega_1^k \end{pmatrix}$$

where

$$\begin{aligned}
\xi_1^k &= \frac{1}{\sqrt{\lambda_2(w^k)}} \left[ \left( x_1^k + \frac{\tau-2}{2} y_1^k \right) + \left( x_2^k + \frac{\tau-2}{2} y_2^k \right)^T \bar{w}_2^k \right], \\
\xi_2^k &= \frac{1}{\sqrt{\lambda_2(w^k)}} \left[ \left( x_2^k + \frac{\tau-2}{2} y_2^k \right) + \left( x_1^k + \frac{\tau-2}{2} y_1^k \right) \bar{w}_2^k \right], \\
\eta_1^k &= \frac{1}{\sqrt{\lambda_2(w^k)}} \left[ \left( y_1^k + \frac{\tau-2}{2} x_1^k \right) + \left( y_2^k + \frac{\tau-2}{2} x_2^k \right)^T \bar{w}_2^k \right], \qquad (52) \\
\eta_2^k &= \frac{1}{\sqrt{\lambda_2(w^k)}} \left[ \left( y_2^k + \frac{\tau-2}{2} x_2^k \right) + \left( y_1^k + \frac{\tau-2}{2} x_1^k \right) \bar{w}_2^k \right],
\end{aligned}$$

and

$$\begin{aligned}
s_1^k &= \frac{\left( x_1^k + \frac{\tau-2}{2} y_1^k \right)}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}, \quad s_2^k = \frac{\left( x_2^k + \frac{\tau-2}{2} y_2^k \right)}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}; \\
\omega_1^k &= \frac{\left( y_1^k + \frac{\tau-2}{2} x_1^k \right)}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}, \quad \omega_2^k = \frac{\left( y_2^k + \frac{\tau-2}{2} x_2^k \right)}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}. \qquad (53)
\end{aligned}$$

By Lemma 2.3, $|\xi_1^k| \le \|\xi_2^k\| \le 1$ and $|\eta_1^k| \le \|\eta_2^k\| \le 1$. In addition,

$$\|s^k\|^2 + \|\omega^k\|^2 = \frac{\|x^k + \frac{\tau-2}{2}y^k\|^2 + \|y^k + \frac{\tau-2}{2}x^k\|^2}{2[\|x^k\|^2 + \|y^k\|^2 + (\tau-2)(x^k)^T y^k] + 2\sqrt{\lambda_2(w^k)}\sqrt{\lambda_1(w^k)}} \le 1.$$

Taking the limit on $L_2(w^k)L_{x^k+\frac{\tau-2}{2}y^k}$ and $L_2(w^k)L_{y^k+\frac{\tau-2}{2}x^k}$, we have

$$\begin{aligned}
L_2(w^k)L_{x^k+\frac{\tau-2}{2}y^k} &\to \frac{1}{2}\begin{pmatrix} \xi_1 & \xi_2 \\ \xi_1\bar{w}_2 + 4(I-\bar{w}_2\bar{w}_2^T)s_2 & \bar{w}_2\xi_2^T + 4(I-\bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\xi^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)s_2 & (I-\bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} \qquad (54) \\
L_2(w^k)L_{y^k+\frac{\tau-2}{2}x^k} &\to \frac{1}{2}\begin{pmatrix} \eta_1 & \eta_2 \\ \eta_1\bar{w}_2^T + 4(I-\bar{w}_2\bar{w}_2^T)\omega_2 & \bar{w}_2\eta_2^T + 4(I-\bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\eta^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)\omega_2 & (I-\bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix} \qquad (55)
\end{aligned}$$

for some vectors $\xi = (\xi_1, \xi_2), \eta = (\eta_1, \eta_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ satisfying $|\xi_1| \le \|\xi_2\| \le 1$ and $|\eta_1| \le \|\eta_2\| \le 1$, and $s = (s_1, s_2), \omega = (\omega_1, \omega_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ satisfying $\|s\|^2 + \|\omega\|^2 \le 1$. From equations (51), (54) and (55), it follows that as $k \to +\infty$,

$$\phi'_{\tau,x}(x^k, y^k) \to \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\xi^T + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}u^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)s_2 & (I-\bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} - I,$$

$$\phi'_{\tau,x}(x^k, y^k) \to \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\eta^T + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}v^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)\omega_2 & (I-\bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix} - I.$$

This shows that as $k \to +\infty$, $\phi'_\tau(x^k, y^k) \to [V_x - I \quad V_y - I]$ with $V_x$ and $V_y$ satisfying (26). Combining with Case (1), the desired result then follows. □