Comput Optim Appl (2010) 45: 59–88
DOI 10.1007/s10589-008-9166-9

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

Shaohua Pan · Jein-Shan Chen

Received: 29 March 2007 / Revised: 29 October 2007 / Published online: 7 February 2008

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 their B-subdifferential. Moreover, for the merit functions induced by them for the second-order cone complementarity problem (SOCCP), we provide a condition for each stationary point to be a solution of the SOCCP and establish the boundedness of their level sets, by exploiting Cartesian P-properties. We also propose a semismooth Newton type 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.

Keywords Second-order cone · Complementarity · B-subdifferential · Semismooth · Newton's method

S. Pan work is partially supported by the Doctoral Starting-up Foundation (B13B6050640) of GuangDong Province.
J.-S. Chen member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

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

J.-S. Chen (✉)
Department of Mathematics, National Taiwan Normal University, Taipei 11677, 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*

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

where $\langle \cdot, \cdot \rangle$ represents 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 [10]. 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

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

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of nonnegative real numbers $\mathbb{R}_+$. We refer to (1)–(2) as the *second-order cone complementarity problem (SOCCP)*. In the sequel, corresponding to the Cartesian structure of $\mathcal{K}$, we write $x = (x_1, \ldots, x_m)$ with $x_i \in \mathbb{R}^{n_i}$ for any $x \in \mathbb{R}^n$, and $F = (F_1, \ldots, F_m)$ 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

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

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):

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

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 convex twice continuously differentiable function. From [6], the KKT conditions for (4), which are sufficient but not necessary for optimality, can be written in the form of (1) and (2) with

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

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 ∇F and ∇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, which include interior-point methods [1–3, 18, 19, 23, 26], non-interior smoothing Newton methods [7, 13], smoothing-regularization methods [15], merit function methods [6] and semismooth Newton methods [16]. 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 $\phi : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}^l$, we call ϕ an *SOC complementarity function* associated with the cone $\mathcal{K}^l$ if for any $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$,

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

Clearly, when $l = 1$, an SOC complementarity function reduces to an NCP function, which plays an important role in the solution of NCPs; see [24] and references therein. A popular choice of ϕ is the Fischer-Burmeister (FB) function [11, 12], defined by

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

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 "$\circ$", 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 $e := (1, 0, \ldots, 0)^T \in \mathbb{R}^l$. 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 $x^2 \in \mathcal{K}^l$ 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, ϕ_{FB} in (7) is well-defined for all $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$ and maps $\mathbb{R}^l \times \mathbb{R}^l$ to $\mathbb{R}^l$. The function ϕ_{FB} was proved in [13] to satisfy the equivalence (6), and therefore its squared norm, denoted by

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

is a merit function for the SOCCP. The merit function is shown to be continuously differentiable by Chen and Tseng [6], and a merit function approach was proposed by use of it.

Another popular choice of ϕ is the natural residual function $\phi_{\mathrm{NR}} : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}^l$ given by

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

where $[\cdot]_+$ means the minimum Euclidean distance projection onto $\mathcal{K}^l$. The function was studied in [13, 15] which is involved in smoothing methods for the SOCCP, recently it was used to develop a semismooth Newton method for nonlinear SOCPs by Kanzow and Fukushima [16]. We note that ϕ_{NR} induces a natural residual merit function

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

but, compared to ψ_{FB}, it has a remarkable drawback, i.e. the non-differentiability.

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

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

with τ being any but fixed parameter in $(0, 4)$. The class of functions is a natural extension of the family of NCP functions proposed by Kanzow and Kleinmichel [17], and has been shown in [4] to satisfy the characterization (6). It is not hard to see that as $\tau = 2$, ϕ_τ reduces to the FB function ϕ_{FB} in (7) while it becomes a multiple of the natural residual function ϕ_{NR} as $\tau \to 0^+$. With the class of SOC complementarity functions, clearly, 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 ψ_τ being the natural merit function associated with ϕ_τ, i.e.,

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

In [4], we studied the continuous differentiability of ψ_τ and showed that each stationary point of Ψ_τ is a solution of the SOCCP if ∇F and $-\nabla G$ are column monotone. In this paper, we concentrate on the properties of ϕ_τ, including the globally Lipschitz continuity, the strong semismoothness, and the characterization of the B-subdifferential. Particularly, we provide a weaker condition than [4] for each stationary point of Ψ_τ to be a solution of the SOCCP and establish the boundedness of the level sets of Ψ_τ, by using Cartesian P-properties. We also propose a semismooth Newton method based on the system (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 $Q_1, \ldots, Q_n$, we denote the block diagonal matrix with these matrices as block diagonals by $\mathrm{diag}(Q_1, \ldots, Q_n)$ or by $\mathrm{diag}(Q_i,\ i = 1, \ldots, n)$. 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 sub-matrices $P_{jk} \in \mathbb{R}^{n_j\times n_k}$ of P with $j \in \mathcal{J}, k \in \mathcal{B}$, and by $x_{\mathcal{B}}$ a vector consisting of sub-vectors $x_i \in \mathbb{R}^{n_i}$ with $i \in \mathcal{B}$.

2 Preliminaries

In this section, we recall some background materials and preliminary results that will be used in the subsequent sections. We begin with the interior and the boundary of $\mathcal{K}^l$. It is known that $\mathcal{K}^l$ is a closed convex self-dual cone with nonempty interior given by

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

and the boundary given by

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

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

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

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 $L_x \succ O$. Then L_x is invertible with

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

We recall from [13] 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

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

with $\bar{x}_2 = x_2/\|x_2\|$ if $x_2 \neq 0$, and otherwise $\bar{x}_2$ being any vector in $\mathbb{R}^{l-1}$ satisfying $\|\bar{x}_2\| = 1$. If $x_2 \neq 0$, then the factorization is unique. The spectral decompositions of x, x^2 and $x^{1/2}$ have some basic properties as below, whose proofs can be found in [13].

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, we have that*

(a) $x \in \mathcal{K}^l$ *if and only if* $\lambda_1(x) \geq 0$, *and* $x \in \operatorname{int}(\mathcal{K}^l)$ *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) $\det(x) = \lambda_1(x)\lambda_2(x)$, $\operatorname{tr}(x) = \lambda_1(x) + \lambda_2(x)$ *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 let

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

for any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$. It is easy to compute

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

Moreover, $w \in \mathcal{K}^l$ and $z \in \mathcal{K}^l$ hold by considering 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 what follows, we present several important technical lemmas. Since their proofs can be found in [4], we here omit them for simplicity.

Lemma 2.1 [4, Lemma 3.4] *For any* $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ *and* $\tau \in (0, 4)$, *let* $w = (w_1, w_2)$ *be defined as in* (16). *If* $\|w_2\| \neq 0$, *then*

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

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

Lemma 2.2 [4, Lemma 3.2] *For any* $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ *and* $\tau \in (0, 4)$, *let* $w = (w_1, w_2)$ *be given as in* (16). *If* $w \notin \operatorname{int}(\mathcal{K}^l)$, *then*

$$x_1^2 = \|x_2\|^2, \qquad y_1^2 = \|y_2\|^2, \qquad x_1 y_1 = x_2^T y_2, \qquad x_1 y_2 = y_1 x_2; \tag{18}$$

$$\begin{aligned} x_1^2 + y_1^2 + (\tau - 2)x_1 y_1 &= \|x_1 x_2 + y_1 y_2 + (\tau - 2)x_1 y_2\| \\ &= \|x_2\|^2 + \|y_2\|^2 + (\tau - 2)x_2^T y_2. \end{aligned} \tag{19}$$

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

$$x_2^T \frac{w_2}{\|w_2\|} = x_1, \qquad x_1 \frac{w_2}{\|w_2\|} = x_2, \qquad y_2^T \frac{w_2}{\|w_2\|} = y_1, \qquad y_1 \frac{w_2}{\|w_2\|} = y_2. \tag{20}$$

Lemma 2.3 [4, Proposition 3.2] *For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, let $z(x, y)$ be defined by (16). Then $z(x, y)$ 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,*

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

where

$$L_z^{-1} = \begin{cases} \begin{pmatrix} b & c\frac{w_2^T}{\|w_2\|} \\ c\frac{w_2}{\|w_2\|} & aI + (b-a)\frac{w_2 w_2^T}{\|w_2\|^2} \end{pmatrix} & \text{if } w_2 \neq 0; \\ (1/\sqrt{w_1})I & \text{if } w_2 = 0, \end{cases} \tag{21}$$

with

$$\begin{aligned} a &= \frac{2}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}}, \qquad b = \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2(w)}} + \frac{1}{\sqrt{\lambda_1(w)}}\right), \\ c &= \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2(w)}} - \frac{1}{\sqrt{\lambda_1(w)}}\right). \end{aligned} \tag{22}$$

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

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

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 [8]. For the concepts of (strongly) semismooth functions, please refer to [21, 22] 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 [5] 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

$$x_\nu \neq 0 \quad \text{and} \quad \langle x_\nu, (Mx)_\nu \rangle \geq 0.$$

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

Definition 2.2 Given a mapping $F = (F_1, \dots, F_m)$ 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 $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m) \in \mathbb{R}^n$, there exists an index $\nu \in \{1, 2, \dots, m\}$ and a positive 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 $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m) \in \mathbb{R}^n$, there exists an index $\nu \in \{1, 2, \dots, m\}$ such that

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

If a continuously differentiable mapping F has the Cartesian P-properties, then the matrix $\nabla F(x)$ at any $x \in \mathbb{R}^n$ enjoys the corresponding Cartesian P-properties.

3 Properties of the functions ϕ_τ and Φ_τ

This section is devoted to investigating the favorable properties of ϕ_τ, which include the globally Lipschitz continuity, the strong semismoothness and the characterization of the B-subdifferential at any point. Based on these results, we also present some properties of the operator Φ_τ related to the generalized Newton method.

From the definition of ϕ_τ and $z(x, y)$ given as in (9) and (16), respectively, we have

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

for any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$. Recall that the vectors $w = (w_1, w_2)$ and $z = (z_1, z_2)$ in (16) satisfy $w, z \in \mathcal{K}^l$, and hence, from Property 2.1(b) and (c),

$$z = \left(\frac{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}}{2}, \frac{\sqrt{\lambda_2(w)} - \sqrt{\lambda_1(w)}}{2} \bar{w}_2 \right), \tag{24}$$

where $\bar{w}_2 = \frac{w_2}{\|w_2\|}$ if $w_2 \neq 0$, and otherwise $\bar{w}_2$ is any vector in $\mathbb{R}^{l-1}$ satisfying $\|\bar{w}_2\| = 1$. The following proposition states some favorable properties possessed by ϕ_τ.

Proposition 3.1 *The function ϕ_τ defined as in (9) has the following properties.*

(a) *ϕ_τ is continuously differentiable at a point $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$ if and only if $(x - y)^2 + \tau(x \circ y) \in \mathrm{int}(\mathcal{K}^l)$. Moreover,*

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

(b) *ϕ_τ is globally Lipschitz continuous with the Lipschitz constant independent of τ.*
(c) *ϕ_τ is strongly semismooth at any $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$.*
(d) *ψ_τ defined by (12) is continuously differentiable everywhere.*

Proof (a) The proof directly follows from Lemma 2.3 and (23).

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

$$\hat{z} = (\hat{z}_1, \hat{z}_2) = \hat{z}(x, y, \epsilon) := [(x - y)^2 + \tau(x \circ y) + \epsilon e]^{1/2} \tag{25}$$

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 Appendix and the Mean-Value Theorem, we have

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

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

(c) From the definition of ϕ_τ and ϕ_{FB}, it is not hard to check that

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

Note that ϕ_{FB} is strongly semismooth everywhere by Corollary 3.3 of [25], 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 at any $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$. Therefore, ϕ_τ is a strongly semismooth function since by [12, Theorem 19] the composition of strongly semismooth functions is strongly semismooth.

(d) The proof can be found in Proposition 3.3 of the literature [4]. □

Proposition 3.1(c) indicates that, when a smoothing or nonsmooth Newton method is employed to solve the system (10), a fast convergence rate (at least superlinear) can 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 ϕ_{FB} was given in [20]. Here, we generalize it to ϕ_τ for any $\tau \in (0, 4)$. The detailed derivation process is included in Appendix for completeness.

Proposition 3.2 *Given a general point $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, each element in $\partial_B \phi_\tau(x, y)$ is of the form $V = [V_x - I \quad V_y - I]$ 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*

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

$$V_y \in \left\{ \frac{1}{2\sqrt{2w_1}} \begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{pmatrix} \left(L_y + \frac{\tau-2}{2} L_x \right) + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T \right\} \tag{26}$$

for some $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ *satisfying* $|u_1| \le \|u_2\| \le 1$ *and* $|v_1| \le \|v_2\| \le 1$, *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* $\hat{u} = (\hat{u}_1, \hat{u}_2)$, $\hat{v} = (\hat{v}_1, \hat{v}_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ *satisfying* $\|\hat{u}\|, \|\hat{v}\| \le 1$ *and* $\hat{u}_1\hat{v}_2 + \hat{v}_1\hat{u}_2 = 0$, *or*

$$V_x \in \left\{ \frac{1}{2} \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix} \xi^T + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T + 2 \begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2\bar{w}_2^T)s_2 & (I - \bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} \right\},$$

$$V_y \in \left\{ \frac{1}{2} \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix} \eta^T + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T + 2 \begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2\bar{w}_2^T)\omega_2 & (I - \bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix} \right\}$$

for some $\|\bar{w}_2\| = 1$, $u = (u_1, u_2)$, $v = (v_1, v_2)$, $\xi = (\xi_1, \xi_2)$, $\eta = (\eta_1, \eta_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ *satisfying* $|u_1| \le \|u_2\| \le 1$, $|v_1| \le \|v_2\| \le 1$, $|\xi_1| \le \|\xi_2\| \le 1$ *and* $|\eta_1| \le \|\eta_2\| \le 1$, *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 focus on the properties of the operator Φ_τ defined in (10). We start with the semismoothness of Φ_τ. Since Φ_τ 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 [12, Theorem 19], we obtain the following conclusion as an immediate consequence of Proposition 3.1(c).

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

To characterize the B-subdifferential of Φ_τ, in the rest of this paper, we let

$$F_i(\zeta) = (F_{i1}(\zeta), F_{i2}(\zeta)), \qquad G_i(\zeta) = (G_{i1}(\zeta), G_{i2}(\zeta)) \in \mathbb{R} \times \mathbb{R}^{n_i - 1}$$

and $w_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ and $z_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ for $i = 1, 2, \ldots, m$ be given as follows:

$$\begin{aligned} w_i &= (w_{i1}(\zeta), w_{i2}(\zeta)) = w(F_i(\zeta), G_i(\zeta)), \\ z_i &= (z_{i1}(\zeta), z_{i2}(\zeta)) = z(F_i(\zeta), G_i(\zeta)). \end{aligned} \tag{27}$$

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

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

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*

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

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

$$A_i(\zeta)\in\{L_{\hat{u}_1}\}\cup\left\{\frac{1}{2}\xi_i(1,\bar{w}_{i2}^T)+\frac{1}{2}u_i(1,-\bar{w}_{i2}^T)+\begin{pmatrix}0 & 2s_{i2}^T(I-\bar{w}_{i2}\bar{w}_{i2}^T)\\ 0 & 2s_{i1}(I-\bar{w}_{i2}\bar{w}_{i2}^T)\end{pmatrix}\right\}$$

$$B_i(\zeta)\in\{L_{\hat{v}_1}\}\cup\left\{\frac{1}{2}\eta_i(1,\bar{w}_{i2}^T)+\frac{1}{2}v_i(1,-\bar{w}_{i2}^T)+\begin{pmatrix}0 & 2\omega_{i2}^T(I-\bar{w}_{i2}\bar{w}_{i2}^T)\\ 0 & 2\omega_{i1}(I-\bar{w}_{i2}\bar{w}_{i2}^T)\end{pmatrix}\right\}$$

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* $\|\hat{u}_i\|,\|\hat{v}_i\|\le 1$ *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* $|u_{i1}|\le\|u_{i2}\|\le 1$, $|v_{i1}|\le\|v_{i2}\|\le 1$, $|\xi_{i1}|\le\|\xi_{i2}\|\le 1$ *and* $|\eta_{i1}|\le\|\eta_{i2}\|\le 1$, $\bar{\omega}_{i2}\in\mathbb{R}^{n_i-1}$ *satisfying* $\|\bar{\omega}_{i2}\|=1$, *and* $s_i=(s_{i1},s_{i2})$, $\omega_i=(\omega_{i1},\omega_{i2})\in\mathbb{R}\times\mathbb{R}^{n_i-1}$ *such that* $\|s_i\|^2+\|\omega_i\|^2\le 1$.

Proof Let $\Phi_{\tau,i}(\zeta)$ denote the ith subvector of Φ_τ, i.e. $\Phi_{\tau,i}(\zeta)=\phi_\tau(F_i(\zeta),G_i(\zeta))$ for all $i=1,2,\dots,m$. From Proposition 2.6.2 of [8], 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_ith 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

$$\partial_B\Phi_{\tau,i}(\zeta)^T=[\nabla F_i(\zeta)\ \ \nabla G_i(\zeta)]\partial_B\phi_\tau(F_i(\zeta),G_i(\zeta))^T,\quad i=1,2,\dots,m. \tag{30}$$

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

Proposition 3.5 *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* $i\in\{1,2,\dots,m\}$,

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

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

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

Proof From Theorem 2.6.6 of [8] and Proposition 3.1(d), we have that

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

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

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

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 conclusions is a direct consequence of Proposition 4.1 of [4]. Notice that for any $i\in\mathcal{O}(\zeta)$ and $\upsilon_i\in\mathbb{R}^{n_i}$,

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

Using the same argument as Case (2) of [4, Proposition 4.1] then yields the second part. □

4 Nonsingularity conditions

In this section, we show that all elements of the B-subdifferential $\partial_B\Phi_\tau(\zeta)$ at a solution ζ^* of the SOCCP are nonsingular if ζ^* 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 multi-valued matrix $(A_i(\zeta^*)-I)+(B_i(\zeta^*)-I)$ is nonsingular if the i-th block component satisfies strict complementarity.

Lemma 4.1 *Let* ζ^* *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* ζ^* *and* $(A_i(\zeta^*)-I)+(B_i(\zeta^*)-I)$ *is nonsingular.*

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

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

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 ζ^*. Since $z_i(\zeta^*) = w_i(\zeta^*)^{1/2} = F_i(\zeta^*)$, from Proposition 3.4(a) it follows that

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

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 ζ^* by Proposition 3.1(a). Since

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

applying Proposition 3.4(a) yields that

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

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* $G_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i})$, *where* $\mathrm{bd}^+(\mathcal{K}^{n_i}) := \mathrm{bd}(\mathcal{K}^{n_i}) \setminus \{0\}$. By Proposition 3.1(a), it suffices to prove $w_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$. Suppose that $w_i(\zeta^*) \in \mathrm{bd}(\mathcal{K}^{n_i})$. Then, from (18) in Lemma 2.2, it follows that

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

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 $\alpha > 0$. Consequently, $F_i(\zeta^*) = \alpha G_i(\zeta^*)$. Noting that $\langle F_i(\zeta^*), G_i(\zeta^*)\rangle = 0$, we then get $F_i(\zeta^*) = G_i(\zeta^*) = 0$. This clearly contradicts the assumptions that $F_i(\zeta^*) \neq 0$ and $G_i(\zeta^*) \neq 0$. So, $w_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$.

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. Note that

$$(2z_i(\zeta^*))^2 = 4w_i(\zeta^*) = 2\left[\left(F_i(\zeta^*)+\frac{\tau-2}{2}G_i(\zeta^*)\right)^2+\frac{\tau(4-\tau)}{4}G_i(\zeta^*)^2\right] + 2\left[\left(G_i(\zeta^*)+\frac{\tau-2}{2}F_i(\zeta^*)\right)^2+\frac{\tau(4-\tau)}{4}F_i(\zeta^*)^2\right],$$

which means that

$$\begin{aligned}(2z_i(\zeta^*))^2-\frac{\tau^2}{4}(F_i(\zeta^*)+G_i(\zeta^*))^2 \\ = \frac{\tau(4-\tau)}{2}[G_i(\zeta^*)^2+F_i(\zeta^*)^2]+\frac{(4-\tau)^2}{4}(F_i(\zeta^*)-G_i(\zeta^*))^2. \qquad (34)\end{aligned}$$

On the other hand, $w_i(\zeta^*)\in \mathrm{int}(\mathcal{K}^{n_i})$ implies that $(F_i(\zeta^*)-G_i(\zeta^*))^2\in \mathrm{int}(\mathcal{K}^{n_i})$ since $F_i(\zeta^*)\circ G_i(\zeta^*)=0$. From the above two sides, we immediately obtain that

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

Since $z_i(\zeta^*)=w_i(\zeta^*)^{1/2}\in \mathrm{int}(\mathcal{K}^{n_i})$, using Proposition 3.4 of [13] 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 so it is nonsingular. □

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

$$\begin{aligned}\mathcal{I}&:=\{i\in\{1,2,\ldots,m\}\mid F_i(\zeta^*)\in \mathrm{int}(\mathcal{K}^{n_i}),\, G_i(\zeta^*)=0\},\\ \mathcal{B}&:=\{i\in\{1,2,\ldots,m\}\mid F_i(\zeta^*)\in \mathrm{bd}^+(\mathcal{K}^{n_i}),\, G_i(\zeta^*)\in \mathrm{bd}^+(\mathcal{K}^{n_i})\}, \qquad (35)\\ \mathcal{J}&:=\{i\in\{1,2,\ldots,m\}\mid F_i(\zeta^*)=0,\, G_i(\zeta^*)\in \mathrm{int}(\mathcal{K}^{n_i})\}\end{aligned}$$

forms a partition of the set $\{1,2,\ldots,m\}$. Thus, suppose that $\nabla G(\zeta^*)$ is invertible, then by rearrangement the matrix $P(\zeta^*)=\nabla G(\zeta^*)^{-1}\nabla F(\zeta^*)$ can be rewritten as

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

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

Theorem 4.1 *Let ζ^* 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 ζ^* is a strict complementarity solution, it follows from Lemma 4.1 that the matrix C can be written in the following partitioned form

$$C = \begin{pmatrix} \frac{\tau-4}{2} I_{\mathcal{I}\mathcal{I}} & P_{\mathcal{I}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}\mathcal{B}}) & \frac{\tau-4}{2} P_{\mathcal{I}\mathcal{J}} \\ 0_{\mathcal{B}\mathcal{I}} & P_{\mathcal{B}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}\mathcal{B}}) + (B_{\mathcal{B}} - I_{\mathcal{B}\mathcal{B}}) & \frac{\tau-4}{2} P_{\mathcal{B}\mathcal{J}} \\ 0_{\mathcal{J}\mathcal{I}} & P_{\mathcal{J}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}\mathcal{B}}) & \frac{\tau-4}{2} P_{\mathcal{J}\mathcal{J}} \end{pmatrix},$$

where $I_{\mathcal{I}\mathcal{I}} = \mathrm{diag}(I_{ii},\ i \in \mathcal{I})$ with I_{ii} 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 ζ^* in the functions. It is not hard to see that these C are nonsingular if and only if

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

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

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

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

$$\begin{cases} y_{\mathcal{J}} = -\frac{2}{4-\tau} P_{\mathcal{J}\mathcal{J}}^{-1} P_{\mathcal{J}\mathcal{B}}(I_{\mathcal{B}\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}}, \\ (P_{\mathcal{B}\mathcal{B}} - P_{\mathcal{B}\mathcal{J}} P_{\mathcal{J}\mathcal{J}}^{-1} P_{\mathcal{J}\mathcal{B}})(I_{\mathcal{B}\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}} = -(I_{\mathcal{B}\mathcal{B}} - B_{\mathcal{B}}) y_{\mathcal{B}}. \end{cases} \tag{36}$$

Thus, by Lemma 4.1 and Proposition 3.5, using the same arguments as Theorem 4.1 of [20] yields the desired result. □

Observe that, when $n_i = 1$ for all $i = 1, 2, \ldots, m$, the assumption for $\widehat{P}_{\mathcal{J}\mathcal{J}}$ in Theorem 4.1 is actually equivalent to requiring that $\widehat{P}_{\mathcal{J}\mathcal{J}}$ 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 ζ^* is a strict complementarity solution of the SOCCP and the mappings F and G at the ζ^* satisfy the conditions of Theorem 4.1. Then, there exists a neighborhood $\mathcal{N}(\zeta^*)$ of ζ^* 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 conditions and bounded level sets

In general a stationary point of a merit function is not a solution of the underlying problem. In [4], we showed that, when ∇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 using the Cartesian P_0-property of a matrix, which, as shown later, is weaker than that of [4] when ∇G is invertible. We also establish the boundedness of the level sets of Ψ_τ for the SOCCP (3) under the condition that $F(\zeta)$ 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 $x, y \in \mathbb{R}^l$,*

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

Proof From Proposition 3.2 of [4], 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 [4], 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 ∇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 Ψ_τ is a solution of the SOCCP.*

Proof Let ζ be an arbitrary stationary point of $\Psi_\tau(\zeta)$. Since Ψ_τ is continuously differentiable by Proposition 3.1(d) and Φ_τ is locally Lipschitz continuous, applying Theorem 2.6.6 of Clarke [8] then gives that for any $V \in \partial_B \Phi_\tau(\zeta)^T$

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

Let V be an element of $\partial_B \Phi_\tau(\zeta)^T (\subseteq \partial \Phi_\tau(\zeta))$. 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 $B(\zeta) = \mathrm{diag}(B_1(\zeta), \ldots, B_m(\zeta))$. From the last three equations, it then follows that

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

By the invertibility of $\nabla G(\zeta)$, the last equation is equivalent to

$$[\nabla G(\zeta)^{-1}\nabla F(\zeta)(A(\zeta) - I) + (B(\zeta) - I)]\Phi_\tau(\zeta) = 0. \tag{37}$$

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 (31) then yields

$$(A_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta) \neq 0 \quad \text{and} \quad (B_\nu(\zeta) - I)\Phi_{\tau,\nu}(\zeta) \neq 0. \tag{38}$$

In addition, from (37) it follows that

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

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

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

Notice that the first term of the left hand side is nonnegative due to (38) 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 Proposition 3.5 since $\Phi_{\tau,\nu}(\zeta) \neq 0$. This leads to a contradiction. Consequently, the proof is completed. □

Remark 5.1 (i) It is not hard to verify that $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ has the Cartesian P_0-property if $\nabla G(\zeta)^{-1}\nabla F(\zeta) \succeq O$. While, if $\nabla G(\zeta)$ is invertible, the column monotonicity of $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ is equivalent to the positive semidefiniteness of $\nabla G(\zeta)^{-1}\nabla F(\zeta)$. Thus, the condition in Proposition 5.1 is weaker than that of [4].

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

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

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 [6] immediately yields the desired result. □

Lemma 5.3 *Let ψ_τ be defined as in (9) and $\{(x^k, y^k)\} \subseteq \mathbb{R}^l \times \mathbb{R}^l$ be a sequence satisfying $\|x^k\| \to +\infty$ and $\|y^k\| \to +\infty$. If $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \not\to 0$, then $\psi_\tau(x^k, y^k) \to +\infty$.*

Proof For each k, 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. Since $\|x^k\|^2 = \frac{1}{2}[(\lambda_1^k)^2 + (\lambda_2^k)^2]$ and $\|y^k\|^2 = \frac{1}{2}[(\mu_1^k)^2 + (\mu_2^k)^2]$, by the given conditions, we only need consider the two cases shown as below.

Case (1): $\lambda_1^k \to -\infty$ or $\mu_1^k \to -\infty$. From Lemma 5.2 and the following fact that

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

we immediately have $\psi_\tau(x^k, y^k) \to +\infty$.

Case (2): $\{\lambda_1^k\}$ and $\{\mu_1^k\}$ are bounded below, but $\lambda_2^k, \mu_2^k \to +\infty$. We will proceed the arguments by contradiction. Suppose that $\{\psi_\tau(x^k, y^k)\}$ is bounded. Since

$$x^k + y^k = z^k - \phi_\tau(x^k, y^k) \quad \text{for each } k,$$

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

$$(4-\tau)x^k \circ y^k = -2z^k \circ \phi_\tau(x^k, y^k) + (\phi_\tau(x^k, y^k))^2. \tag{39}$$

Noting that, for each k,

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

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

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

Using (39) 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. □

Now using Lemma 5.3 and the same arguments as Proposition 5.2 of [20], 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 Ψ_τ.

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 function* Ψ_τ *has bounded level sets.*

6 Algorithm and numerical results

From the previous discussions, we see that the class of SOC complementarity function ϕ_τ with $\tau \in (0, 4)$ 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 [9], which is described as follows.

Algorithm 6.1

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

Step 1. If $\|\nabla\Psi_\tau(\zeta^k)\| \le \varepsilon$, 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

$$W_k d = -\Phi_\tau(\zeta^k). \tag{40}$$

If the system is not solvable or if the descent condition

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

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

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

$$\Psi_\tau(\zeta^k + \rho^m d^k) \le \Psi_\tau(\zeta^k) + \sigma\rho^m\nabla\Psi_\tau(\zeta^k)^T d^k, \tag{41}$$

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 ε is equal 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* Ψ_τ.

Table 1 Set of test problems

No.	Problem names	n	m	# of nonzero elements of matrix A	Structure of SOCs
1	nb	2383	123	192439	$[4\times 1; 793\times 3]$
2	nb-L1	3176	915	193104	$[797\times 1; 793\times 3]$
3	nb-L2-bessel	2641	123	209924	$[4\times 1; 1\times 123; 838\times 3]$

(b) *If ζ^* is an isolated accumulation point of $\{\zeta^k\}$, then the entire sequence $\{\zeta^k\}$ converges to ζ^*.*
(c) *If ζ^* is an accumulation point such that ζ^* is a strict complementarity solution and $F(\zeta)$ and $G(\zeta)$ at ζ^* satisfy the conditions of Theorem 4.1. Then,*
  (i) *the whole sequence $\{\zeta^k\}$ converges to ζ^*;*
  (ii) *the search direction d^k is eventually given by the solution of* (40);
  (iii) *the sequence $\{\zeta^k\}$ converges to ζ^* Q-superlinearly*;
  (iv) *if, in addition, F' and G' are Lipschitz continuous at ζ^*, then the rate of convergence is Q-quadratic.*

Proof Since the proofs are similar to that of [17, Theorem 4.2] or [9, Theorem 3.1] by the results obtained in Sects. 3–5, we here omit them. □

Note that Theorem 6.1(a) and (b) only gives global convergence results to stationary points of the merit function Ψ_τ whereas we are much concerned with finding a global minimizer of Ψ_τ 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 Ψ_τ 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 experiments in the next section also verify this fact.

In what follows, we report the computational experience with solving some linear SOCPs, which correspond to the SOCP (4) with $g(x)=c^Tx$, 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 2.8 GHz CPU and 512 MB memory. The computer codes were all written in Matlab 6.5. During the experiments, we replaced the standard Armijo linesearch rule in Algorithm 6.1 with a nonmonotone linesearch as described in [14]. 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 [14] to have better numerical performance for the unconstrained minimization of smooth functions. Specifically, we computed the smallest nonnegative integer m such that

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

Table 2 Numerical results of Algorithm 6.1 for linear SOCPs with a different τ

No.	τ	Obj.	NF	k	Time	τ	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

where

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

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

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

Throughout the experiments, the following parameters were used in the algorithm:

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

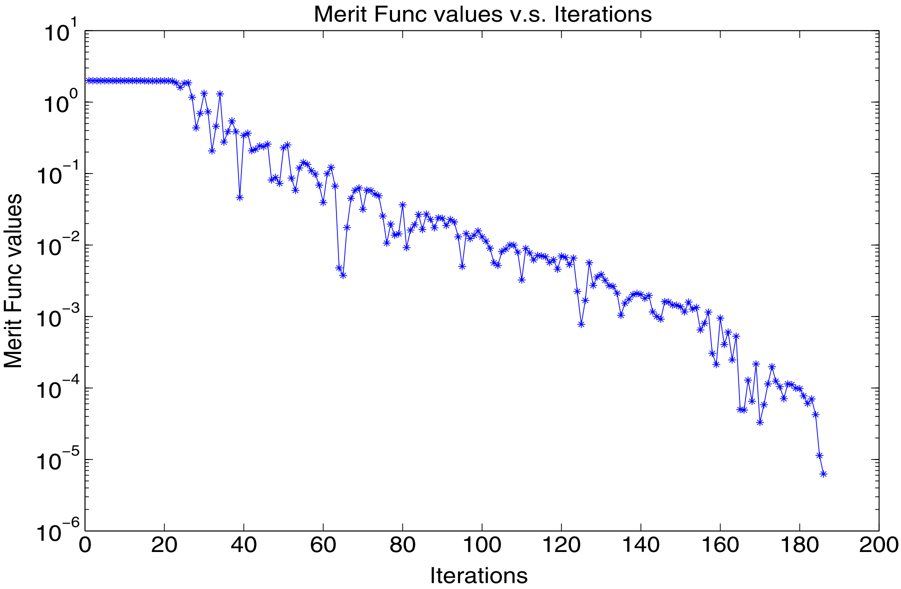
The starting point was chosen to be $\zeta^0 = 0$. The method terminates whenever one of the following conditions is satisfied

$$\max\{|F(\zeta^k)^T G(\zeta^k)|, \Psi_\tau(\zeta^k)\} \le 10^{-5}, \quad k > 200, \ \alpha_k < 10^{-15}. \tag{42}$$

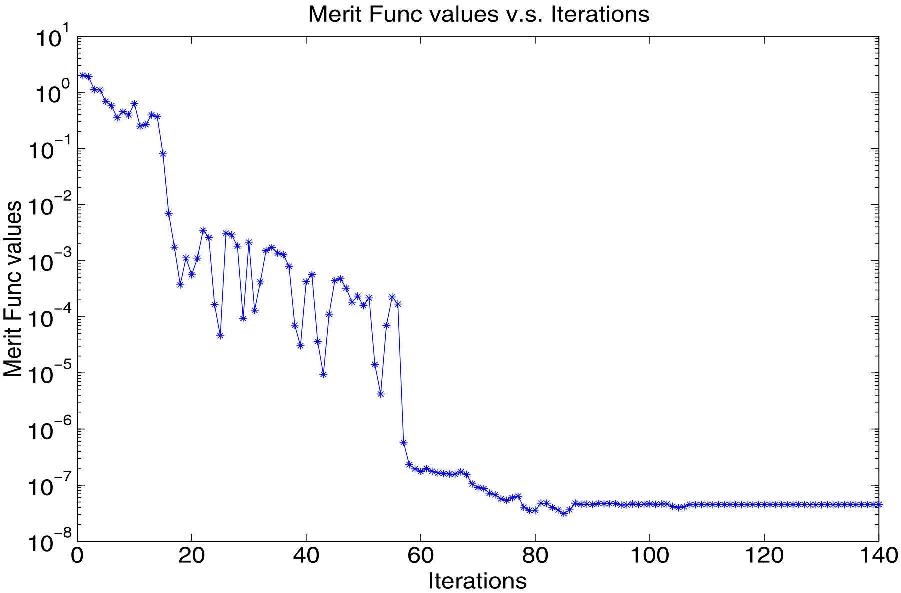
The term $|F(\zeta^k)^T G(\zeta^k)|$ in the first condition aims to obtain a solution with a favorable dual gap. In addition, it also helps to stop the algorithm when the decrease of $\Psi_\tau(\zeta)$ has little advantage in reducing the dual gap.

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 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 τ 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 τ close to 4 often gives a better global convergence, whereas the parameter τ close to 0 leads to a fast local convergence. Figure 1



(a) $\tau = 0.1$



(b) $\tau = 3.9$

Fig. 1 The convergence of Algorithm 6.1 with different τ for 'nb'

below displays the convergence of Ψ_τ for problem "nb" with $\tau = 0.1$ and $\tau = 3.9$, respectively. The performance of Ψ_τ coincides with the case described by [17] for the NCPs, which is very important for the use of the class of SOC complementarity functions. Based on this feature of ϕ_τ, we may adopt a dynamic choice of τ in the algorithm by following a line similar to [17].

7 Conclusions

In this paper, we gave a detailed investigation for the properties of a one-parametric class of SOC complementarity functions ϕ_τ, which includes the FB SOC complementarity function and the natural residual SOC complementarity function as a special case. We showed that ϕ_τ is globally Lipschitz continuous and strongly semismooth and characterized its B-subdifferential at any point. Furthermore, for the induced natural merit functions Ψ_τ, we provided a weaker condition than [4] to guarantee every stationary point to be a solution of the SOCCP, and proved that it has bounded level sets for the SOCCPs with the uniform Cartesian P-property. Thus, combining with the results of [4], we extended most of favorable properties of the class of complementarity functions for the NCPs to the setting of the SOCCPs.

A semismooth Newton method was also proposed by the nonsmooth reformulation (10) involving the class of SOC complementarity functions. The superlinear convergence of the algorithm was 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.

Appendix

Lemma A.1 *Let $\hat{z}(x,y,\epsilon)$ be defined as in (25) for any $\epsilon>0$. Then $\hat{z}(x,y,\epsilon)$ is continuously differentiable everywhere, and there exists a scalar $C>0$ such that*

$$\|\nabla_x\hat{z}(x,y,\epsilon)\|_F\le C,\qquad \|\nabla_y\hat{z}(x,y,\epsilon)\|_F\le C \tag{43}$$

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 $(x-y)^2+\tau(x\circ y)+\epsilon e\in\operatorname{int}(\mathcal{K}^l)$ for any $(x,y)\in\mathbb{R}^l\times\mathbb{R}^l$ and $\epsilon>0$, by Lemma 2.3 the function $\hat{z}(x,y,\epsilon)$ is continuously differentiable everywhere and

$$\nabla_x\hat{z}(x,y,\epsilon)=\left(L_x+\frac{\tau-2}{2}L_y\right)L_{\hat{z}}^{-1},\qquad \nabla_y\hat{z}(x,y,\epsilon)=\left(L_y+\frac{\tau-2}{2}L_x\right)L_{\hat{z}}^{-1}. \tag{44}$$

We next prove the bound in (43). For convenience, write

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

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 (44) and the formula of $L_{\hat{z}}^{-1}$ given by (21), we can compute that

$$\nabla_x \hat{z}(x,y,\epsilon) = \begin{pmatrix} \hat{b}g_1 + \hat{c}\frac{g_2^T w_2}{\|w_2\|} & \hat{c}\frac{g_1 w_2^T}{\|w_2\|} + \hat{a}g_2^T + (\hat{b}-\hat{a})g_2^T\frac{w_2 w_2^T}{\|w_2\|^2} \\ \hat{b}g_2 + \hat{c}g_1\frac{w_2}{\|w_2\|} & \hat{c}\frac{g_2 w_2^T}{\|w_2\|} + \hat{a}g_1 I + (\hat{b}-\hat{a})g_1\frac{w_2 w_2^T}{\|w_2\|^2} \end{pmatrix},$$

where $\hat{a}, \hat{b}$ and $\hat{c}$ are defined as in (22) with $w = \hat{w}$. Notice that

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

Using the expression of $\hat{a}, \hat{b}$ and $\hat{c}$ and the result of Lemma 2.1 then yields that

$$\begin{aligned} &\left|\hat{b}g_1 + \hat{c}g_2^T\frac{w_2}{\|w_2\|}\right| \\ &\quad \le \frac{1}{2\sqrt{\lambda_2(w)}}\left|g_1 + g_2^T\frac{w_2}{\|w_2\|}\right| + \frac{1}{2\sqrt{\lambda_1(w)}}\left|g_1 - g_2^T\frac{w_2}{\|w_2\|}\right| \le 1, \\ &\left\|\hat{c}\frac{g_1 w_2^T}{\|w_2\|} + \hat{b}\frac{g_2^T w_2 w_2^T}{\|w_2\|^2}\right\| \\ &\quad \le \frac{1}{2\sqrt{\lambda_2(w)}}\left|g_1 + g_2^T\frac{w_2}{\|w_2\|}\right| + \frac{1}{2\sqrt{\lambda_1(w)}}\left|g_1 - g_2^T\frac{w_2}{\|w_2\|}\right| \le 1, \\ &\left\|\hat{a}g_2^T - \hat{a}g_2^T w_2\frac{w_2^T}{\|w_2\|^2}\right\| \\ &\quad \le \frac{\|2g_2\|}{\sqrt{\|x\|^2 + \|y\|^2 + (\tau-2)x^T y}} \cdot \left\|I - \frac{w_2 w_2^T}{\|w_2\|^2}\right\|_F \le 2(l-1), \\ &\left\|\hat{b}g_2 + \hat{c}g_1\frac{w_2}{\|w_2\|}\right\| \\ &\quad \le \frac{1}{2\sqrt{\lambda_2(w)}}\left\|g_2 + g_1\frac{w_2}{\|w_2\|}\right\| + \frac{1}{2\sqrt{\lambda_1(w)}}\left\|g_2 - g_1\frac{w_2}{\|w_2\|}\right\| \le 1, \\ &\left\|\hat{c}\frac{g_2 w_2^T}{\|w_2\|} + \hat{b}\frac{g_1 w_2 w_2^T}{\|w_2\|^2}\right\|_F \\ &\quad \le \frac{1}{2\sqrt{\lambda_2(w)}}\left\|g_2 + g_1\frac{w_2}{\|w_2\|}\right\| + \frac{1}{2\sqrt{\lambda_1(w)}}\left\|g_2 - g_1\frac{w_2}{\|w_2\|}\right\| \le 1, \\ &\left\|\hat{a}g_1 I - \hat{a}g_1\frac{w_2 w_2^T}{\|w_2\|^2}\right\|_F \\ &\quad \le \frac{2|g_1|}{\sqrt{\|x\|^2 + \|y\|^2 + (\tau-2)x^T y}} \cdot \left\|I - \frac{w_2 w_2^T}{\|w_2\|^2}\right\|_F \le 2(l-1). \end{aligned}$$

The above inequalities imply that the first inequality in (43) holds under this case.

Case (2) $w_2 = 0$. In this case, from Lemma 2.3 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 (43). 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_{ϕ_τ} denote the set of points where ϕ_τ is differentiable. Recall that this set is characterized by Proposition 3.1(a). Write

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

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

$$\phi'_{\tau,x}(x, y) = L_z^{-1} L_{x+\frac{\tau-2}{2}y} - I, \qquad \phi'_{\tau,x}(x, y) = L_z^{-1} L_{y+\frac{\tau-2}{2}x} - I. \tag{45}$$

Moreover, we observe from (21) 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

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

where $\hat{w}_2 = w_2/|w_2\|$, and consequently $\phi'_{\tau,x}$ and $\phi'_{\tau,y}$ in (45) can be rewritten as

$$\begin{aligned} \phi'_{\tau,x}(x, y) &= (L_1(w) + L_2(w)) L_{x+\frac{\tau-2}{2}y} - I, \\ \phi'_{\tau,y}(x, y) &= (L_1(w) + L_2(w)) L_{y+\frac{\tau-2}{2}x} - I. \end{aligned} \tag{46}$$

(a) Under the given assumption, ϕ_τ 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,

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

Substituting the formulas in (45) 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 (46), it then follows that

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

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

$$
\begin{aligned}
\lim_{k\to+\infty} L_2(w^k)L_{x^k + \frac{\tau-2}{2} y^k} &= C(w)\left(L_x + \frac{\tau-2}{2} L_y\right), \\
\lim_{k\to+\infty} L_2(w^k)L_{y^k + \frac{\tau-2}{2} x^k} &= C(w)\left(L_y + \frac{\tau-2}{2} L_x\right),
\end{aligned} \tag{48}
$$

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

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

In addition, by a simple computation, we have that

$$
L_1(w^k)L_{x^k + \frac{\tau-2}{2} y^k} = \frac{1}{2} \begin{pmatrix} u_1^k & (u_2^k)^T \\ -u_1^k \bar{w}_2^k & -\bar{w}_2^k (u_2^k)^T \end{pmatrix},
$$

$$
L_1(w^k)L_{y^k + \frac{\tau-2}{2} x^k} = \frac{1}{2} \begin{pmatrix} v_1^k & (v_2^k)^T \\ -v_1^k \bar{w}_2^k & -\bar{w}_2^k (v_2^k)^T \end{pmatrix},
$$

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 2.1, $|u_1^k| \le \|u_2^k\| \le 1$ 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

$$\begin{aligned}
L_1(w^k)L_{x^k+\frac{\tau-2}{2}y^k} &\to \frac{1}{2}\begin{pmatrix} u_1 & u_2^T \\ -u_1\bar{w}_2 & -\bar{w}_2u_2^T \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T, \\
L_1(w^k)L_{y^k+\frac{\tau-2}{2}x^k} &\to \frac{1}{2}\begin{pmatrix} v_1 & v_2^T \\ -v_1\bar{w}_2 & -\bar{w}_2v_2^T \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T
\end{aligned} \tag{49}$$

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 (47)–(49), 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 (26).

(c) Assume $(x, y) = (0, 0)$. Let $\{(x^k, y^k)\} \subset D_{\phi_\tau}$ 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 $w_2^k \ne 0$ for all k. We proceed the arguments by the two cases.

Case (1): $w_2^k = 0$ for all k. From (45) and Lemma 2.3, it follows that

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

Since

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

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

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

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 $\phi'_{\tau,x}(x^k,y^k)$ are given as in (47). Using the same arguments as part (b) and noting that $\{\bar{w}_2^k\}$ is bounded, we have

$$L_1(w^k)L_{x^k+\frac{\tau-2}{2}y^k}\to\frac{1}{2}\begin{pmatrix}1\\-\bar{w}_2\end{pmatrix}u^T,\qquad L_1(w^k)L_{y^k+\frac{\tau-2}{2}x^k}\to\frac{1}{2}\begin{pmatrix}1\\-\bar{w}_2\end{pmatrix}v^T \tag{50}$$

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 $L_2(w)$,

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

where

$$\begin{aligned}
\xi_1^k&=\frac{1}{\sqrt{\lambda_2(w^k)}}\left[\left(x_1^k+\frac{\tau-2}{2}y_1^k\right)+-\left(x_2^k+\frac{\tau-2}{2}y_2^k\right)^T\bar{w}_2^k\right],\\
\xi_2^k&=\frac{1}{\sqrt{\lambda_2(w^k)}}\left[\left(x_2^k+\frac{\tau-2}{2}y_2^k\right)+-\left(x_1^k+\frac{\tau-2}{2}y_1^k\right)\bar{w}_2^k\right],\\
\eta_1^k&=\frac{1}{\sqrt{\lambda_2(w^k)}}\left[\left(y_1^k+\frac{\tau-2}{2}x_1^k\right)+\left(y_2^k+\frac{\tau-2}{2}x_2^k\right)^T\bar{w}_2^k\right],\\
\eta_2^k&=\frac{1}{\sqrt{\lambda_2(w^k)}}\left[\left(y_2^k+\frac{\tau-2}{2}x_2^k\right)+\left(y_1^k+\frac{\tau-2}{2}x_1^k\right)\bar{w}_2^k\right],
\end{aligned}\tag{51}$$

and

$$\begin{aligned}
s_1^k&=\frac{(x_1^k+\frac{\tau-2}{2}y_1^k)}{\sqrt{\lambda_2(w^k)}+\sqrt{\lambda_1(w^k)}},\qquad s_2^k=\frac{(x_2^k+\frac{\tau-2}{2}y_2^k)}{\sqrt{\lambda_2(w^k)}+\sqrt{\lambda_1(w^k)}};\\
\omega_1^k&=\frac{(y_1^k+\frac{\tau-2}{2}x_1^k)}{\sqrt{\lambda_2(w^k)}+\sqrt{\lambda_1(w^k)}},\qquad \omega_2^k=\frac{(y_2^k+\frac{\tau-2}{2}x_2^k)}{\sqrt{\lambda_2(w^k)}+\sqrt{\lambda_1(w^k)}}.
\end{aligned}\tag{52}$$

By Lemma 2.1, $|\xi_1^k| \le \|\xi_2^k\| \le 1$ 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

$$\begin{aligned} L_2(w^k)&L_{x^k+\frac{\tau-2}{2}y^k} \\ &\to \frac{1}{2}\begin{pmatrix} \xi_1 & \xi_2 \\ \xi_1\bar{w}_2 + 4(I-\bar{w}_2\bar{w}_2^T)s_2 & \bar{w}_2\xi_2^T + 4(I-\bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\xi^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)s_2 & (I-\bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix}, \end{aligned} \tag{53}$$

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

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 (50), (53) and (54), it follows that as $k \to +\infty$,

$$\begin{aligned} \phi'_{\tau,x}&(x^k, y^k) \\ &\to \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\xi^T + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}u^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)s_2 & (I-\bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} - I, \\ \phi'_{\tau,x}&(x^k, y^k) \\ &\to \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\eta^T + \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}v^T + 2\begin{pmatrix} 0 & 0 \\ (I-\bar{w}_2\bar{w}_2^T)\omega_2 & (I-\bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix} - I. \end{aligned}$$

This shows that as $k \to +\infty$, $\phi'_\tau(x^k, y^k) \to [V_x - I \;\; V_y - I]$ with V_x and V_y satisfying the characterization in Proposition 3.2(c). Combining with Case (1), the desired result then follows. □

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**, 3–51 (2003)
2. Alizadeh, F., Schmieta, S.: Symmetric cones, potential reduction methods, and word-by-word extensions. In: Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.) Handbook of Semidefinite Programming, pp. 195–233. Kluwer Academic, Boston (2000)
3. Andersen, E.D., Roos, C., Terlaky, T.: On implementing a primal-dual interior-point method for conic quadratic optimization. Math. Program. Ser. B **95**, 249–277 (2003)

4. Chen, J.-S., Pan, S.-H.: A one-parametric class of merit functions for the second-order cone complementarity problem. Comput. Optim. Appl. (2008, accepted)
5. Chen, X., Qi, H.: Cartesian P-property and its applications to the semidefinite linear complementarity problem. Math. Program. **106**, 177–201 (2006)
6. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. Math. Program. **104**, 293–327 (2005)
7. Chen, X.-D., Sun, D., Sun, J.: Complementarity functions and numerical experiments for second-order cone complementarity problems. Comput. Optim. Appl. **25**, 39–56 (2003)
8. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983). Reprinted by SIAM, Philadelphia (1990)
9. De Luca, T., Facchinei, F., Kanzow, C.: A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 407–439 (1996)
10. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs. Oxford University Press, New York (1994)
11. Fischer, A.: A special Newton-type optimization methods. Optimization **24**, 269–284 (1992)
12. Fischer, A.: Solution of the monotone complementarity problem with locally Lipschitzian functions. Math. Program. **76**, 513–532 (1997)
13. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2002)
14. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
15. Hayashi, S., Yamashita, N., Fukushima, M.: A combined smoothing and regularization method for monotone second-order cone complementarity problems. SIAM J. Optim. **15**, 593–615 (2005)
16. Kanzow, C., Fukushima, M.: Semismooth methods for linear and nonlinear second-order cone programs. Technical Report, Department of Applied Mathematics and Physics, Kyoto University (2006)
17. Kanzow, C., Kleinmichel, H.: A new class of semismooth Newton-type methods for nonlinear complementarity problems. Comput. Optim. Appl. **11**, 227–251 (1998)
18. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Application of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
19. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions. Math. Program. **88**, 61–83 (2000)
20. Pan, S.-H., Chen, J.-S.: A damped Gauss-Newton method for the second-order cone complementarity problem. Appl. Math. Optim. (2008, accepted)
21. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244 (1993)
22. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)
23. Schmieta, S., Alizadeh, F.: Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. Math. Oper. Res. **26**, 543–564 (2001)
24. Sun, D., Qi, L.-Q.: On NCP-functions. Comput. Optim. Appl. **13**, 201–220 (1999)
25. Sun, D., Sun, J.: Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions. Math. Program. **103**, 575–581 (2005)
26. Tsuchiya, T.: A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim. Methods Softw. **11**, 141–182 (1999)