Appl Math Optim (2009) 59: 293–318
DOI 10.1007/s00245-008-9054-9

A Damped Gauss-Newton Method for the Second-Order Cone Complementarity Problem

Shaohua Pan · Jein-Shan Chen

Published online: 6 August 2008

Abstract We investigate some properties related to the generalized Newton method for the Fischer-Burmeister (FB) function over second-order cones, which allows us to reformulate the second-order cone complementarity problem (SOCCP) as a semi-smooth system of equations. Specifically, we characterize the B-subdifferential of the FB function at a general point and study the condition for every element of the B-subdifferential at a solution being nonsingular. In addition, for the induced FB merit function, we establish its coerciveness and provide a weaker condition than Chen and Tseng (Math. Program. 104:293–327, 2005) for each stationary point to be a solution, under suitable Cartesian P-properties of the involved mapping. By this, a damped Gauss-Newton method is proposed, and the global and superlinear convergence results are obtained. Numerical results are reported for the second-order cone programs from the DIMACS library, which verify the good theoretical properties of the method.

Keywords Second-order cones · Complementarity · Fischer-Burmeister function · B-subdifferential · Generalized Newton method

---

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

S. Pan (✉)
School of Mathematical Sciences, South China University of Technology, Guangzhou 510640, 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

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, $F, G : \mathbb{R}^n \to \mathbb{R}^m$ are the mapping assumed to be continuously differentiable throughout this paper, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOCs), or called Lorentz cones. In other words,

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

where $q, n_1, \ldots, n_q \geq 1$, $n_1 + \cdots + n_q = m$ and

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

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*). Corresponding to the Cartesian structure of $\mathcal{K}$, in the rest of this paper, we always write $F = (F_1, \ldots, F_q)$ and $G = (G_1, \ldots, G_q)$ with $F_i, G_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$.

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

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

which is a natural extension of the nonlinear complementarity problem (NCP) over the nonnegative orthant cone $\mathbb{R}^n_+$. Another special case corresponds to the Karush-Kuhn-Tucker (KKT) conditions for the convex second-order cone program (CSOCP):

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

where $A \in \mathbb{R}^{p \times m}$ has full row rank, $b \in \mathbb{R}^p$ and $g : \mathbb{R}^m \to \mathbb{R}$ is a twice continuously differentiable convex function. From [6], the KKT conditions of (4), which are sufficient but not necessary for optimality, can be rewritten in the form of (1) with

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

where $\hat{x} \in \mathbb{R}^m$ is any vector satisfying $Ax = b$. When g is a linear function, (4) reduces to the standard second-order cone program which has extensive applications in engineering design, finance, control, and robust optimization; see [1, 14] and references therein.

There have been many methods proposed for solving SOCPs and SOCCPs. They include the interior-point methods [1, 2, 14, 16, 24, 26], the non-interior smoothing Newton methods [7, 11], the smoothing-regularization method [13], and the merit function approach [6]. Among others, the last three kinds of methods are all based on an SOC complementarity function. Specifically, a mapping $\phi : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}^l$ is called an *SOC complementarity function* associated with the cone $\mathcal{K}^l$ $(l \geq 1)$ if

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

A popular choice of ϕ is the vector-valued Fischer-Burmeister (FB) function, defined by

$$\phi(x, y) := (x^2 + y^2)^{1/2} - (x + y) \quad \forall x, y \in \mathbb{R}^l \tag{7}$$

where $x^2 = x \circ x$ denotes the Jordan product of x and itself, $x^{1/2}$ denotes a vector such that $(x^{1/2})^2 = x$, and $x + y$ means the usual componentwise addition of vectors. From the next section, we see that ϕ in (7) is well-defined for all $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$. The function was shown in [11] to satisfy the equivalence (6), and therefore its squared norm

$$\psi(x, y) := \frac{1}{2}\|\phi(x, y)\|^2 \tag{8}$$

is a merit function for the SOCCP, i.e., $\psi(x, y) = 0$ if and only if $x \in \mathcal{K}^l$, $y \in \mathcal{K}^l$ and $\langle x, y\rangle = 0$. The functions ϕ and ψ were studied in the literature [6, 21], where ψ was shown to be continuously differentiable everywhere by Chen and Tseng [6] and ϕ was proved to be strongly semismooth by Sun and Sun [21].

In view of the characterization in (6), clearly, the SOCCP can be reformulated as the following nonsmooth system of equations:

$$\Phi(\zeta) := \begin{pmatrix} \phi(F_1(\zeta), G_1(\zeta)) \\ \vdots \\ \phi(F_i(\zeta), G_i(\zeta)) \\ \vdots \\ \phi(F_q(\zeta), G_q(\zeta)) \end{pmatrix} = 0 \tag{9}$$

where ϕ is defined as in (7) with a suitable dimension l. By Corollary 3.3 of [21], it is not hard to show that the operator $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ in (9) is semismooth. Furthermore, from Proposition 2 of [6], its squared norm induces a smooth merit function, given by

$$\Psi(\zeta) := \frac{1}{2}\|\Phi(\zeta)\|^2 = \sum_{i=1}^{q} \psi(F_i(\zeta), G_i(\zeta)). \tag{10}$$

In this paper, we mainly characterize the B-subdifferential of ϕ at a general point and present an estimate for the B-subdifferential of Φ. By this, a condition is given to guarantee every element of the B-subdifferential of Φ at a solution to be nonsingular, which plays an important role in the local convergence analysis of nonsmooth Newton methods for the SOCCP. In addition, two important results are also presented for the merit function $\Psi(\zeta)$. One of them shows that each stationary point of Ψ is a solution of the SOCCP under a weaker condition than the one used by [6], and the other establishes the coerciveness of Ψ for the SOCCP (3) under the uniform Cartesian P-property of F. Based on these results, we finally propose a damped Gauss-Newton method by applying the generalized Newton method [19, 20] for the system (9), and analyze its global and superlinear (quadratic) convergence. Numerical results are reported for the SOCPs from the DIMACS library [18], which verify the good theoretical properties of the method.

Throughout this paper, I represents an identity matrix of suitable dimension, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors, and $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_q}$ is identified with $\mathbb{R}^{n_1+\cdots+n_q}$. Thus, $(x_1, \ldots, x_q) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_q}$ is viewed as a column vector in $\mathbb{R}^{n_1+\cdots+n_q}$. For any differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, the notation $\nabla F(x) \in \mathbb{R}^{n\times m}$ denotes the transpose of the Jacobian $F'(x)$. For a symmetric matrix A, we write $A \succ O$ (respectively, $A \succeq O$) if A is positive definite (respectively, positive semidefinite). Given a finite number of square matrices $Q_1, \ldots, Q_q$, we denote the block diagonal matrix with these matrices as block diagonals by $\mathrm{diag}(Q_1, \ldots, Q_q)$ or by $\mathrm{diag}(Q_i,\ i = 1, \ldots, q)$. If $\mathcal{J}$ and $\mathcal{B}$ are index sets such that $\mathcal{J}, \mathcal{B} \subseteq \{1, 2, \ldots, q\}$, we denote by $P_{\mathcal{J}\mathcal{B}}$ 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 denote by $x_{\mathcal{B}}$ a vector consisting of sub-vectors $x_i \in \mathbb{R}^{n_i}$ 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 start with the interior and the boundary of $\mathcal{K}^l$ ($l > 1$). 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 any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, we define their *Jordan product* [9] as

$$x \circ y := \big(\langle x, y\rangle, x_1y_2 + y_1x_2\big).$$

The Jordan product "$\circ$", unlike scalar or matrix multiplication, is not associative, which is the main source on complication in the analysis of SOCCP. The identity element under this product is $e := (1, 0, \ldots, 0)^T \in \mathbb{R}^l$. For each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, define the matrix L_x by

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

which can be viewed as a linear mapping from $\mathbb{R}^l$ to $\mathbb{R}^l$ with the following properties.

Property 2.1

(a) $L_x y = x \circ y$ *and* $L_{x+y} = L_x + L_y$ *for any* $x, y \in \mathbb{R}^l$.
(b) $x \in \mathcal{K}^l \Longleftrightarrow L_x \succeq O$ *and* $x \in \text{int}(\mathcal{K}^l) \Longleftrightarrow L_x \succ O$.
(c) L_x *is invertible whenever* $x \in \text{int}(\mathcal{K}^l)$ *with the inverse* L_x^{-1} *given by*

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \frac{\det(x)}{x_1} I + \frac{x_2 x_2^T}{x_1} \end{bmatrix}, \tag{11}$$

where $\det(x) := x_1^2 - \|x_2\|^2$ *denotes the determinant of* x.

In the following, we recall from [9, 11] 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 $\lambda_1(x), \lambda_2(x)$ and $u_x^{(1)}, u_x^{(2)}$ are the spectral values and the associated spectral vectors of x, respectively, defined by

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

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 $x_2 \neq 0$, the factorization is unique. The spectral factorizations of x, x^2 and $x^{1/2}$ have various interesting properties, and some of them are summarized as follows.

Property 2.2 *For any* $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, *let* $\lambda_1(x), \lambda_2(x)$ *and* $u_x^{(1)}, u_x^{(2)}$ *be the spectral values and the associated spectral vectors. Then, the following results hold.*

(a) $x \in \mathcal{K}^l \iff 0 \le \lambda_1(x) \le \lambda_2(x)$ *and* $x \in \text{int}(\mathcal{K}^l) \iff 0 < \lambda_1(x) \le \lambda_2(x)$.
(b) $x^2 = [\lambda_1(x)]^2 u_x^{(1)} + [\lambda_2(x)]^2 u_x^{(2)} \in \mathcal{K}^l$ *for any* $x \in \mathbb{R}^l$.
(c) *If* $x \in \mathcal{K}^l$, *then* $x^{1/2} = \sqrt{\lambda_1(x)}\, u_x^{(1)} + \sqrt{\lambda_2(x)}\, u_x^{(2)} \in \mathcal{K}^l$.

Now we recall the concepts of the B-subdifferential and (strong) semismoothness. Given a mapping $H : \mathbb{R}^n \to \mathbb{R}^m$, if H is locally Lipschitz continuous, then the set

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

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 $\partial H(z) := \text{conv}\partial_B H(z)$ is the generalized Jacobian of Clarke [4]. Semismoothness was originally introduced by Mifflin [15] for functionals. Smooth functions, convex functionals, and piecewise linear functions are examples of semismooth functions. Later, Qi and Sun [20] extended the definition of semismooth functions to a mapping $H : \mathbb{R}^n \to \mathbb{R}^m$. H is called *semismooth at* x if H is directionally differentiable at x and for all $V \in \partial H(x + h)$ and $h \to 0$,

$$Vh - H'(x; h) = o(\|h\|);$$

H is called *strongly semismooth at* x if H is semismooth at x and for all $V \in \partial H(x+h)$ and $h \to 0$,

$$Vh - H'(x;h) = O(\|h\|^2);$$

H is called (*strongly*) *semismooth* if it is (strongly) semismooth everywhere. Here, $o(\|h\|)$ means a function $\alpha : \mathbb{R}^n \to \mathbb{R}^m$ satisfying $\lim_{h\to 0}\alpha(h)/\|h\| = 0$, while $O(\|h\|^2)$ denotes a function $\alpha : \mathbb{R}^n \to \mathbb{R}^m$ satisfying $\|\alpha(h)\| \le C\|h\|^2$ for all $\|h\| \le \delta$ and some $C > 0, \delta > 0$.

Next, we present the definitions of Cartesian P-properties for a matrix $M \in \mathbb{R}^{m\times m}$, which are special cases of those introduced by Chen and Qi [5] for a linear transformation.

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

(a) the Cartesian P-property if for any $0 \neq x = (x_1, \dots, x_q) \in \mathbb{R}^m$ with $x_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \dots, q\}$ such that $\langle x_\nu, (Mx)_\nu\rangle > 0$;
(b) the Cartesian P_0-property if for any $0 \neq x = (x_1, \dots, x_q) \in \mathbb{R}^m$ with $x_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \dots, q\}$ such that $x_\nu \neq 0$ and $\langle x_\nu, (Mx)_\nu\rangle \ge 0$.

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

Definition 2.2 Given a mapping $F = (F_1, \dots, F_q)$ 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 $x = (x_1, \dots, x_q)$, $y = (y_1, \dots, y_q) \in \mathbb{R}^m$, there is an index $\nu \in \{1, 2, \dots, q\}$ and a positive constant $\rho > 0$ such that

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

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

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

From the above definitions, if a continuously differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ has the uniform Cartesian P-property (Cartesian P_0-property), then $\nabla F(x)$ at any $x \in \mathbb{R}^n$ enjoys the Cartesian P-property (Cartesian P_0-property). In addition, we may see that, when $n_1 = \cdots = n_q = 1$, the above concepts reduce to the definitions of P-matrices and P-functions, respectively, for the NCP.

Finally, we introduce some notations which will be used in the rest of this paper. For any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, we define $w, z : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}^l$ by

$$\begin{aligned} w &= (w_1, w_2) = (w_1(x,y), w_2(x,y)) = w(x,y) := x^2 + y^2, \\ z &= (z_1, z_2) = (z_1(x,y), z_2(x,y)) = z(x,y) := (x^2+y^2)^{1/2}. \end{aligned} \tag{12}$$

Clearly, $w \in \mathcal{K}^l$ with $w_1 = \|x\|^2 + \|y\|^2$ and $w_2 = 2(x_1x_2 + y_1y_2)$. 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 $\|\bar{w}_2\| = 1$.

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

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

3 B-Subdifferential of the FB Function

In this section, we characterize the B-subdifferential of the FB function ϕ at a general point $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$. For this purpose, we need several important technical lemmas. The first lemma characterizes the set of the points where $z(x, y)$ is differentiable. Since the proof is direct by [3, Proposition 4] and formula (11), we here omit it.

Lemma 3.1 *The function $z(x, y)$ in (12) is continuously differentiable at a point (x, y) if and only if $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^l)$. Moreover, $\nabla_x z(x, y) = L_x L_z^{-1}$ and $\nabla_y z(x, y) = L_y L_z^{-1}$, where $L_z^{-1} = (1/\sqrt{w_1})I$ if $w_2 = 0$, and otherwise*

$$L_z^{-1} = \begin{pmatrix} b & c\bar{w}_2^T \\ c\bar{w}_2 & aI + (b-a)\bar{w}_2\bar{w}_2^T \end{pmatrix} \tag{13}$$

with

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

The following two lemmas extend the results of Lemmas 2 and 3 of [6], respectively. Since the proofs are direct by using the same technique as [6], we here omit them.

Lemma 3.2 *For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ with $w = x^2 + y^2 \in \mathrm{bd}(\mathcal{K}^l)$, we have*

$$x_1^2 = \|x_2\|^2, \quad y_1^2 = \|y_2\|^2, \qquad x_1 y_1 = x_2^T y_2, \quad x_1 y_2 = y_1 x_2.$$

If, in addition, $w_2 \neq 0$, then $\|w\|^2 = 2w_1^2 = 2\|w_2\|^2 = 4(x_1^2 + y_1^2) \neq 0$ and

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

Lemma 3.3 *For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ with $w_2 = 2(x_1 x_2 + y_1 y_2) \neq 0$, there holds that*

$$\left(x_1 + (-1)^i x_2^T \bar{w}_2\right)^2 \le \|x_2 + (-1)^i x_1 \bar{w}_2\|^2 \le \lambda_i(w) \quad \text{for } i = 1, 2.$$

Based on Lemmas 3.1–3.3, we are now in a position to present the representation for the elements of the B-subdifferential $\partial_B \phi(x, y)$ at a general point $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$.

Proposition 3.1 *Given a general point $(x, y) \in \mathbb{R}^l \times \mathbb{R}^l$, each element in $\partial_B \phi(x, y)$ is given by $[V_x - I \;\; V_y - I]$ with V_x and V_y having the following representation*:

(a) *If $x^2 + y^2 \in \operatorname{int}(\mathcal{K}^l)$, then $V_x = L_z^{-1} L_x$ and $V_y = L_z^{-1} L_y$.*
(b) *If $x^2 + y^2 \in \operatorname{bd}(\mathcal{K}^l)$ and $(x, y) \neq (0, 0)$, then*

$$\begin{aligned} V_x &\in \left\{ \frac{1}{2\sqrt{2w_1}} \begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2 \bar{w}_2^T \end{pmatrix} L_x + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T \right\} \\ V_y &\in \left\{ \frac{1}{2\sqrt{2w_1}} \begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2 \bar{w}_2^T \end{pmatrix} L_y + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T \right\} \end{aligned} \tag{14}$$

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 $\bar{w}_2 = w_2 / \|w_2\|$.
(c) *If $(x, y) = (0, 0)$, then $V_x \in \{L_{\hat{x}}\}, \;\; V_y \in \{L_{\hat{y}}\}$ for some $\hat{x}, \hat{y}$ with $\|\hat{x}\|^2 + \|\hat{y}\|^2 = 1$, or*

$$\begin{aligned} V_x &\in \left\{ \frac{1}{2} \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix} \xi^T + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T + 2 \begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2 \bar{w}_2^T) s_2 & (I - \bar{w}_2 \bar{w}_2^T) s_1 \end{pmatrix} \right\} \\ V_y &\in \left\{ \frac{1}{2} \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix} \eta^T + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T + 2 \begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2 \bar{w}_2^T) \omega_2 & (I - \bar{w}_2 \bar{w}_2^T) \omega_1 \end{pmatrix} \right\} \end{aligned} \tag{15}$$

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

Proof Let D_ϕ denote the set of points where ϕ is differentiable. Recall that this set is characterized by Lemma 3.1 since $\phi(x, y) = z(x, y) - (x + y)$, and moreover,

$$\phi_x'(x, y) = L_z^{-1} L_x - I, \qquad \phi_y'(x, y) = L_z^{-1} L_y - I \quad \forall (x, y) \in D_\phi.$$

(a) In this case, ϕ is continuously differentiable at (x, y) by Lemma 3.1. Hence, $\partial_B \phi(x, y)$ consists of a single element, i.e. $\phi'(x, y) = [L_z^{-1} L_x - I \;\; L_z^{-1} L_y - I]$, and the result is clear.

(b) Assume that $(x, y) \neq (0, 0)$ satisfies $x^2 + y^2 \in \operatorname{bd}(\mathcal{K}^l)$. Then $w \in \operatorname{bd}(\mathcal{K}^l)$ and $w_1 > 0$, which means $\|w_2\| = w_1 > 0$ and $\lambda_2(w) > \lambda_1(w) = 0$. Observe that, when $w_2 \neq 0$, the matrix L_z^{-1} in (13) can be decomposed as the sum of

$$L_1(w) := \frac{1}{2\sqrt{\lambda_1(w)}} \begin{pmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2 \bar{w}_2^T \end{pmatrix} \tag{16}$$

and

$$L_2(w) := \frac{1}{2\sqrt{\lambda_2(w)}} \begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & \frac{4\sqrt{\lambda_2(w)}}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}} (I - \bar{w}_2 \bar{w}_2^T) + \bar{w}_2 \bar{w}_2^T \end{pmatrix} \tag{17}$$

with $\bar{w}_2 = w_2/\|w_2\|$. Consequently, ϕ'_x and ϕ'_y can be rewritten as

$$\begin{aligned} \phi'_x(x,y) &= (L_1(w)+L_2(w))L_x - I, \\ \phi'_y(x,y) &= (L_1(w)+L_2(w))L_y - I. \end{aligned} \tag{18}$$

Let $\{(x^k,y^k)\} \subseteq D_\phi$ be an arbitrary sequence converging to (x,y). Let $w^k = (w_1^k, w_2^k) = w(x^k,y^k)$ and $z^k = z(x^k,y^k)$ for each k, where $w(x,y)$ and $z(x,y)$ are defined by (12). Since $w_2 \neq 0$, we without loss of generality assume $\|w_2^k\| \neq 0$ for each k. Let $\bar{w}_2^k = w_2^k/\|w_2^k\|$ for each k. From (18), it follows that

$$\begin{aligned} \phi'_x(x^k,y^k) &= (L_1(w^k)+L_2(w^k))L_{x^k} - I, \\ \phi'_y(x^k,y^k) &= (L_1(w^k)+L_2(w^k))L_{y^k} - I. \end{aligned} \tag{19}$$

Since $\lim_{k\to\infty}\lambda_1(w^k)=0, \lim_{k\to\infty}\lambda_2(w^k)=2w_1>0$ and $\lim_{k\to\infty}\bar{w}_2^k=\bar{w}_2$, we have

$$\lim_{k\to\infty} L_2(w^k)L_{x^k} = C(w)L_x \quad \text{and} \quad \lim_{k\to\infty} L_2(w^k)L_{y^k} = C(w)L_y \tag{20}$$

where

$$C(w) = \frac{1}{2\sqrt{2w_1}}\begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{pmatrix}. \tag{21}$$

Next we focus on the limit of $L_1(w^k)L_{x^k}$ and $L_1(w^k)L_{y^k}$ as $k\to\infty$. By computing,

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

where

$$\begin{aligned} u_1^k &= \frac{x_1^k - (x_2^k)^T\bar{w}_2^k}{\sqrt{\lambda_1(w^k)}}, & u_2^k &= \frac{x_2^k - x_1^k\bar{w}_2^k}{\sqrt{\lambda_1(w^k)}}, \\ v_1^k &= \frac{y_1^k - (y_2^k)^T\bar{w}_2^k}{\sqrt{\lambda_1(w^k)}}, & v_2^k &= \frac{y_2^k - y_1^k\bar{w}_2^k}{\sqrt{\lambda_1(w^k)}}. \end{aligned}$$

By Lemma 3.3, $|u_1^k| \le \|u_2^k\| \le 1$ and $|v_1^k| \le \|v_2^k\| \le 1$. So, taking the limit (possibly on a subsequence) on $L_1(w^k)L_{x^k}$ and $L_1(w^k)L_{y^k}$, respectively, gives

$$\begin{aligned} L_1(w^k)L_{x^k} &\to \frac{1}{2}\begin{pmatrix} u_1 & u_2 \\ -u_1\bar{w}_2 & -\bar{w}_2u_2^T \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}u^T \\ L_1(w^k)L_{y^k} &\to \frac{1}{2}\begin{pmatrix} v_1 & v_2 \\ -v_1\bar{w}_2 & -\bar{w}_2v_2^T \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}v^T \end{aligned} \tag{22}$$

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 $|v_1|\le\|v_2\|\le 1$. In fact, u and v are some accumulation point of the sequences $\{u^k\}$ and $\{v^k\}$, respectively. From equations (19)–(22), we immediately obtain

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

This shows that $\phi'(x^k,y^k)\to[V_x-I\;\;V_y-I]$ as $k\to\infty$ with V_x, V_y satisfying (14).

(c) Assume that $(x,y)=(0,0)$. Let $\{(x^k,y^k)\}\subseteq D_\phi$ be an arbitrary sequence converging to (x,y). Let $w^k=(w_1^k,w_2^k)=w(x^k,y^k)$ and $z^k=z(x^k,y^k)$ for each k. Since $w=0$, we without any loss of generality assume that $w_2^k=0$ for all k, or $w_2^k\neq 0$ for all k.

Case (1): $w_2^k=0$ for all k. From Lemma 3.1, it follows that $L_{z^k}^{-1}=(1/\sqrt{w_1^k})I$. Therefore,

$$\phi'_x(x^k,y^k)=\frac{1}{\sqrt{w_1^k}}L_{x^k}-I\quad\text{and}\quad\phi'_y(x^k,y^k)=\frac{1}{\sqrt{w_1^k}}L_{y^k}-I.$$

Since $w_1^k=\|x^k\|^2+\|y^k\|^2$, every element in $\phi'_x(x^k,y^k)$ and $\phi'_y(x^k,y^k)$ is bounded. Taking limit (possibly on a subsequence) on $\phi'_x(x^k,y^k)$ and $\phi'_y(x^k,y^k)$, we obtain

$$\phi'_x(x^k,y^k)\;\to\;L_{\hat{x}}-I\quad\text{and}\quad\phi'_y(x^k,y^k)\;\to\;L_{\hat{y}}-I$$

for some vectors $\hat{x},\hat{y}\in\mathbb{R}^l$ satisfying $\|\hat{x}\|^2+\|\hat{y}\|^2=1$, where $\hat{x}$ and $\hat{y}$ are some accumulation point of the sequences $\{\frac{x^k}{\sqrt{w_1^k}}\}$ and $\{\frac{y^k}{\sqrt{w_1^k}}\}$, respectively. Thus, we prove that $\phi'(x^k,y^k)\to[V_x-I\;\;V_y-I]$ as $k\to\infty$ with $V_x\in\{L_{\hat{x}}\}$ and $V_y\in\{L_{\hat{y}}\}$.

Case (2): $w_2^k\neq 0$ for all k. Now $\phi'_x(x^k,y^k)$ and $\phi'_y(x^k,y^k)$ are given as in (19). Using the same arguments as part (b) and noting the boundedness of $\{\bar{w}_2^k\}$, we have

$$L_1(w^k)L_{x^k}\;\to\;\frac{1}{2}\begin{pmatrix}1\\ -\bar{w}_2\end{pmatrix}u^T,\qquad L_1(w^k)L_{y^k}\;\to\;\frac{1}{2}\begin{pmatrix}1\\ -\bar{w}_2\end{pmatrix}v^T\tag{23}$$

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 $|v_1|\le\|v_2\|\le 1$, and $\bar{w}_2\in\mathbb{R}^{l-1}$ satisfying $\|\bar{w}_2\|=1$. We next compute the limit of $L_2(w^k)L_{x^k}$ and $L_2(w^k)L_{y^k}$ as $k\to\infty$. By the definition of $L_2(w)$ in (17),

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

where

$$\xi_1^k = \frac{x_1^k + (x_2^k)^T \bar{w}_2^k}{\sqrt{\lambda_2(w^k)}}, \qquad \xi_2^k = \frac{x_2^k + x_1^k \bar{w}_2^k}{\sqrt{\lambda_2(w^k)}},$$
$$\eta_1^k = \frac{y_1^k + (y_2^k)^T \bar{w}_2^k}{\sqrt{\lambda_2(w^k)}}, \qquad \eta_2^k = \frac{y_2^k + y_1^k \bar{w}_2^k}{\sqrt{\lambda_2(w^k)}}, \tag{24}$$

and

$$s_1^k = \frac{x_1^k}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}, \qquad s_2^k = \frac{x_2^k}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}},$$
$$\omega_1^k = \frac{y_1^k}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}, \qquad \omega_2^k = \frac{y_2^k}{\sqrt{\lambda_2(w^k)} + \sqrt{\lambda_1(w^k)}}. \tag{25}$$

By Lemma 3.3, $|\xi_1^k| \le \|\xi_2^k\| \le 1$ and $|\eta_1^k| \le \|\eta_2^k\| \le 1$. In addition,

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

Hence, taking limit (possibly on a subsequence) on $L_2(w^k)L_{x^k}$ and $L_2(w^k)L_{y^k}$ yields

$$\begin{aligned}
L_2(w^k)L_{x^k} &\to \frac{1}{2}\begin{pmatrix} \xi_1 & \xi_2^T \\ \xi_1\bar{w}_2 + 4(I - \bar{w}_2\bar{w}_2^T)s_2 & \bar{w}_2\xi_2^T + 4(I - \bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\xi^T + 2\begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2\bar{w}_2^T)s_2 & (I - \bar{w}_2\bar{w}_2^T)s_1 \end{pmatrix}, \\
L_2(w^k)L_{y^k} &\to \frac{1}{2}\begin{pmatrix} \eta_1 & \eta_2^T \\ \eta_1\bar{w}_2 + 4(I - \bar{w}_2\bar{w}_2^T)\omega_2 & \bar{w}_2\eta_2^T + 4(I - \bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}\eta^T + 2\begin{pmatrix} 0 & 0 \\ (I - \bar{w}_2\bar{w}_2^T)\omega_2 & (I - \bar{w}_2\bar{w}_2^T)\omega_1 \end{pmatrix}
\end{aligned} \tag{26}$$

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 $|\eta_1| \le \|\eta_2\| \le 1$, $\bar{w}_2 \in \mathbb{R}^{l-1}$ satisfying $\|\bar{w}_2\| = 1$, and $s = (s_1, s_2), \omega = (\omega_1, \omega_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ satisfying $\|s\|^2 + \|\omega\|^2 \le 1/2$. Among others, ξ and η are some accumulation point of the sequences $\{\xi^k\}$ and $\{\eta^k\}$, respectively; and s and ω are some accumulation point of the sequences $\{s^k\}$ and $\{\omega^k\}$, respectively. From (19), (23) and (26), we obtain

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

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

This implies that as $k \to \infty$, $\phi'(x^k, y^k) \to [V_x - I \quad V_y - I]$ with V_x and V_y satisfying (15).

Combining with Case (1) then yields the desired result. □

Remark 3.1 When $x^2 + y^2 \in \mathrm{bd}(\mathcal{K}^l)$ with $(x, y) \neq (0, 0)$, using Lemma 3.2, we can also characterize V_x and V_y in Proposition 3.1(b) by

$$V_x \in \left\{ \frac{1}{\sqrt{2w_1}}\begin{pmatrix} x_1 & x_2^T \\ x_2 & 2x_1 I - \frac{w_2 x_2^T}{w_1} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ \frac{-w_2}{\|w_2\|} \end{pmatrix}u^T \right\}$$

$$V_y \in \left\{ \frac{1}{\sqrt{2w_1}}\begin{pmatrix} y_1 & y_2^T \\ y_2 & 2y_1 I - \frac{w_2 y_2^T}{w_1} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ \frac{-w_2}{\|w_2\|} \end{pmatrix}v^T \right\}$$

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 $|v_1| \le \|v_2\| \le 1$.

4 Properties of the Operator Φ

In this section, we study some properties of Φ related to the generalized Newton method. In particular, we shall present an estimate for the B-subdifferential of Φ and a sufficient condition for all elements of the B-subdifferential of Φ at a solution being nonsingular. For convenience, throughout this section, for any $i \in \{1, 2, \ldots, q\}$ and $\zeta \in \mathbb{R}^n$, we let

$$\begin{aligned}&F_i(\zeta) = (F_{i1}(\zeta), F_{i2}(\zeta)), \qquad G_i(\zeta) = (G_{i1}(\zeta), G_{i2}(\zeta)) \in \mathbb{R} \times \mathbb{R}^{n_i - 1}, \\ &w_i(\zeta) = (w_{i1}(\zeta), w_{i2}(\zeta)) = w(F_i(\zeta), G_i(\zeta)), \\ &z_i(\zeta) = (z_{i1}(\zeta), z_{i2}(\zeta)) = z(F_i(\zeta), G_i(\zeta))\end{aligned}$$

where $w(x, y)$ and $z(x, y)$ are the functions defined as in (12).

First, since Φ 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 [8, Theorem 19], we have the following proposition as an immediate consequence of Corollary 3.3 of [21].

Proposition 4.1 *The operator $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ defined by (9) is semismooth. Furthermore, it is strongly semismooth if F' and G' are locally Lipschitz continuous.*

Let Φ_i denote the i-th component of the function Φ. Notice that, for any $\zeta \in \mathbb{R}^n$,

$$\partial_B \Phi(\zeta)^T \subseteq \partial_B \Phi_1(\zeta)^T \times \partial_B \Phi_2(\zeta)^T \times \cdots \times \partial_B \Phi_q(\zeta)^T \tag{27}$$

where the latter denotes the set of all matrices whose $(n_{i-1}+1)$ to n_i-th columns belong to $\partial_B \Phi_i(\zeta)^T$ for $i=1,2,\ldots,q$ and $n_0=0$. From Proposition 3.1 and Remark 3.1, we immediately obtain the following estimate for $\partial_B \Phi(\zeta)^T$.

Proposition 4.2 *Let $\Phi:\mathbb{R}^n\to\mathbb{R}^m$ be defined by (9). Then, for any $\zeta\in\mathbb{R}^n$,*

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

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

(a) *If $F_i(\zeta)^2+G_i(\zeta)^2\in \operatorname{int}(\mathcal{K}^{n_i})$, then $A_i(\zeta)=L_{F_i(\zeta)}L^{-1}_{z_i(\zeta)}$ and $B_i(\zeta)=L_{G_i(\zeta)}L^{-1}_{z_i(\zeta)}$.*

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

$$A_i(\zeta)\in\left\{\frac{1}{\sqrt{2w_{i1}(\zeta)}}\begin{pmatrix}F_{i1}(\zeta) & F_{i2}(\zeta)^T\\ F_{i2}(\zeta) & 2F_{i1}(\zeta)I-\frac{F_{i2}(\zeta)w_{i2}(\zeta)^T}{w_{i1}(\zeta)}\end{pmatrix}\right.$$
$$\left.+\frac{1}{2}u_i(1,-\bar{w}_{i2}(\zeta)^T)\right\}$$
$$B_i(\zeta)\in\left\{\frac{1}{\sqrt{2w_{i1}(\zeta)}}\begin{pmatrix}G_{i1}(\zeta) & G_{i2}(\zeta)^T\\ G_{i2}(\zeta) & 2G_{i1}(\zeta)I-\frac{G_{i2}(\zeta)w_{i2}(\zeta)^T}{w_{i1}(\zeta)}\end{pmatrix}\right.$$
$$\left.+\frac{1}{2}v_i(1,-\bar{w}_{i2}(\zeta)^T)\right\}$$

for some $u_i=(u_{i1},u_{i2}), v_i=(v_{i1},v_{i2})\in\mathbb{R}\times\mathbb{R}^{n_i-1}$ satisfying $|u_{i1}|\le\|u_{i2}\|\le1$ and $|v_{i1}|\le\|v_{i2}\|\le1$, where $\bar{w}_{i2}(\zeta)=w_{i2}(\zeta)/\|w_{i2}(\zeta)\|$.

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

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

for some $\hat{u}_i,\hat{v}_i\in\mathbb{R}^{n_i}$ satisfying $\|\hat{u}_i\|^2+\|\hat{v}_i\|^2=1$, $u_i=(u_{i1},u_{i2}), v_i=(v_{i1},v_{i2}), \xi_i=(\xi_{i1},\xi_{i2}), \eta_i=(\eta_{i1},\eta_{i2})\in\mathbb{R}\times\mathbb{R}^{n_i-1}$ satisfying $|u_{i1}|\le\|u_{i2}\|\le1$, $|v_{i1}|\le\|v_{i2}\|\le1$, $|\xi_{i1}|\le\|\xi_{i2}\|\le1$ and $|\eta_{i1}|\le\|\eta_{i2}\|\le1$, $\bar{w}_{i2}\in\mathbb{R}^{n_i-1}$ 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 $\|s_i\|^2+\|\omega_i\|^2\le1/2$.

Particularly, for the block matrices $A(\zeta)$ and $B(\zeta)$, we have the following properties.

Lemma 4.1 *For any $\zeta\in\mathbb{R}^n$, let $A(\zeta)$ and $B(\zeta)$ be given as in Proposition 4.2. Then,*

(a) *for all* $i \in \{1, 2, \ldots, q\}$ *such that* $F_i(\zeta)^2 + G_i(\zeta)^2 \in \operatorname{int}(\mathcal{K}^{n_i})$, *there holds that*

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

(b) *for all* $i \in \{1, 2, \ldots, q\}$, *we have* $\langle (A_i(\zeta) - I)\Phi_i(\zeta), (B_i(\zeta) - I)\Phi_i(\zeta) \rangle \geq 0$, *and the inequality holds with equality if and only if* $\Phi_i(\zeta) = 0$.

Proof (a) The proof is similar to that of [6, Lemma 6]. For completeness, we here include it. From Proposition 4.2(a), it follows that for any $\upsilon_i \in \mathbb{R}^{n_i}$,

$$\begin{aligned}\langle (A_i - I)\upsilon_i,\ (B_i - I)\upsilon_i \rangle &= \langle (L_{F_i} L_{z_i}^{-1} - I)\upsilon_i,\ (L_{G_i} L_{z_i}^{-1} - I)\upsilon_i \rangle \\ &= \langle (L_{F_i} - L_{z_i})L_{z_i}^{-1}\upsilon_i,\ (L_{G_i} - L_{z_i})L_{z_i}^{-1}\upsilon_i \rangle \\ &= \langle (L_{G_i} - L_{z_i})(L_{F_i} - L_{z_i})L_{z_i}^{-1}\upsilon_i,\ L_{z_i}^{-1}\upsilon_i \rangle \end{aligned} \tag{29}$$

where, for convenience, we omit the notation ζ in functions. Let S_i be the symmetric part of $(L_{G_i} - L_{z_i})(L_{F_i} - L_{z_i})$. Then, by computing, we have

$$\begin{aligned} S_i &= \frac{1}{2}[(L_{G_i} - L_{z_i})(L_{F_i} - L_{z_i}) + (L_{F_i} - L_{z_i})(L_{G_i} - L_{z_i})] \\ &= \frac{1}{2}(L_{z_i} - L_{F_i} - L_{G_i})^2 + \frac{1}{2}(L_{z_i}^2 - L_{F_i}^2 - L_{G_i}^2). \end{aligned}$$

Notice that $z_i = (F_i^2 + G_i^2)^{1/2} \in \operatorname{int}(\mathcal{K}^{n_i})$ and $z_i^2 - F_i^2 - G_i^2 = 0 \in \mathcal{K}^{n_i}$, and hence we have $L_{z_i}^2 - L_{F_i}^2 - L_{G_i}^2 \succeq O$ by [11, Proposition 3.4]. From (29), it then follows that

$$\begin{aligned}\langle (A_i - I)\upsilon_i,\ (B_i - I)\upsilon_i \rangle &= \langle S_i L_{z_i}^{-1}\upsilon_i, L_{z_i}^{-1}\upsilon_i \rangle \\ &\geq \frac{1}{2}\langle (L_{z_i} - L_{F_i} - L_{G_i})^2 L_{z_i}^{-1}\upsilon_i,\ L_{z_i}^{-1}\upsilon_i \rangle \\ &= \frac{1}{2}\|(L_{z_i} - L_{F_i} - L_{G_i})L_{z_i}^{-1}\upsilon_i\|^2 \geq 0 \end{aligned}$$

for any $\upsilon_i \in \mathbb{R}^{n_i}$, where the first inequality is due to the fact that $L_{z_i}^2 - L_{F_i}^2 - L_{G_i}^2 \succeq O$.

(b) From Theorem 2.6.6 of [4] and the smoothness of $\psi(x, y)$ (see [6]), we have

$$\nabla\psi(x, y) = \partial_B\phi(x, y)^T\phi(x, y) \quad \forall x, y \in \mathbb{R}^l,$$

which, together with Propositions 3.1 and 4.2, implies that for $i = 1, 2, \ldots, q$,

$$\begin{aligned} \nabla_x\psi(F_i(\zeta), G_i(\zeta)) &= (A_i(\zeta) - I)\Phi_i(\zeta), \\ \nabla_y\psi(F_i(\zeta), G_i(\zeta)) &= (B_i(\zeta) - I)\Phi_i(\zeta). \end{aligned} \tag{30}$$

Using Lemma 6(b) of [6], we immediately obtain the desired result. □

In what follows, we study under what conditions all elements of the B-subdifferential $\partial_B\Phi(\zeta)$ at a solution are nonsingular. Given a solution ζ^* of the SOCCP, we call it *non-degeneracy* if $F_i(\zeta^*) + G_i(\zeta^*) \in \operatorname{int}(\mathcal{K}^{n_i})$ for all $i \in \{1, 2, \ldots, q\}$.

Remark 4.1 Let ζ^* be a solution of the SOCCP. From [1], we know that precisely one of the following six cases holds for each block pair $(F_i(\zeta), G_i(\zeta))$:

$F_i(\zeta^*)$	$G_i(\zeta^*)$	SC
$F_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$	$G_i(\zeta^*) = 0$	Yes
$F_i(\zeta^*) = 0$	$G_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$	Yes
$F_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$	$G_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$	Yes
$F_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$	$G_i(\zeta^*) = 0$	No
$F_i(\zeta^*) = 0$	$G_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$	No
$F_i(\zeta^*) = 0$	$G_i(\zeta^*) = 0$	No

where $\text{bd}^+(\mathcal{K}^{n_i}) = \text{bd}(\mathcal{K}^{n_i}) \setminus \{0\}$, and the last column indicates whether the strict complementarity, i.e. $F_i(\zeta^*) + G_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$, holds or not. Particularly, when the i-th block pair satisfies the strict complementarity, $A_i(\zeta^*)$ and $B_i(\zeta^*)$ have an explicit expression as shown by Lemma 4.2 below.

Lemma 4.2 *Let ζ^* be a solution to the SOCCP. For any $i \in \{1, 2, \ldots, q\}$, we have*

(a) $A_i(\zeta^*) = 0$ *and* $B_i(\zeta^*) = I$ *if* $F_i(\zeta^*) = 0$ *and* $G_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$;
(b) $A_i(\zeta^*) = I$ *and* $B_i(\zeta^*) = 0$ *if* $F_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$ *and* $G_i(\zeta^*) = 0$;
(c) $A_i(\zeta^*) = L_{F_i(\zeta^*)} L^{-1}_{z_i(\zeta^*)}$ *and* $B_i(\zeta^*) = L_{G_i(\zeta^*)} L^{-1}_{z_i(\zeta^*)}$ *if* $F_i(\zeta^*), G_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$.

Proof (a) Since $F_i(\zeta^*)^2 + G_i(\zeta^*)^2 = G_i(\zeta^*)^2 \in \text{int}(\mathcal{K}^{n_i})$, by Proposition 4.2(a),

$$A_i(\zeta^*) = L_{F_i(\zeta^*)} L^{-1}_{z_i(\zeta^*)} = 0 \quad \text{and} \quad B_i(\zeta^*) = L_{G_i(\zeta^*)} L^{-1}_{z_i(\zeta^*)} = L_{G_i(\zeta^*)} L^{-1}_{G_i(\zeta^*)} = I.$$

Similarly, we can prove that part (b) holds. Next we consider part (c). We claim that $F_i(\zeta^*)^2 + G_i(\zeta^*)^2 \in \text{int}(\mathcal{K}^{n_i})$. Suppose not, then $F_i(\zeta^*)^2 + G_i(\zeta^*)^2 \in \text{bd}^+(\mathcal{K}^{n_i})$, which by Lemma 3.3 implies that $F_{i1}(\zeta^*) G_{i1}(\zeta^*) = F_{i2}(\zeta^*)^T G_{i2}(\zeta^*)$. On the other hand, since $F_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$ and $G_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})$, we have that

$$F_{i1}(\zeta^*) = \|F_{i2}(\zeta^*)\|, \qquad G_{i1}(\zeta^*) = \|G_{i2}(\zeta^*)\|. \tag{31}$$

Combining the two sides then yields that $\|F_{i2}(\zeta^*)\| \cdot \|G_{i2}(\zeta^*)\| = F_{i2}(\zeta^*)^T G_{i2}(\zeta^*)$. This implies that $F_{i2}(\zeta^*) = \alpha G_{i2}(\zeta^*)$ for some $\alpha > 0$. Combining with (31) then yields $F_{i1}(\zeta^*) = \alpha G_{i1}(\zeta^*)$. Therefore, $F_i(\zeta^*) = \alpha G_i(\zeta^*)$. Noting that $F_i(\zeta^*)^T G_i(\zeta^*) = 0$ since ζ^* is a solution of the SOCCP, we have $F_i(\zeta^*) = G_i(\zeta^*) = 0$. This clearly contradicts the given assumption. Using Proposition 4.2(a), we then obtain the desired result. □

By Remark 4.1, if ζ^* is a nondegenerate solution of the SOCCP, then the index sets

$$\begin{aligned}
\mathcal{I} &:= \left\{ i \in \{1, 2, \ldots, q\} \mid F_i(\zeta^*) = 0,\ G_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i}) \right\}, \\
\mathcal{B} &:= \left\{ i \in \{1, 2, \ldots, q\} \mid F_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i}),\ G_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i}) \right\}, \qquad (32) \\
\mathcal{J} &:= \left\{ i \in \{1, 2, \ldots, q\} \mid F_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i}),\ G_i(\zeta^*) = 0 \right\}
\end{aligned}$$

form a partition of $\{1, 2, \ldots, q\}$. Thus, if $n = m$, by supposing that $\nabla G(\zeta^*)$ is invertible and rearranging the matrices appropriately, $P(\zeta^*) = \nabla G(\zeta^*)^{-1}\nabla F(\zeta^*)$ 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 able to prove the following nonsingularity result under the assumption that the given solution is nondegenerate.

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

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

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

Proof Using (28) and noting that $\nabla G(\zeta^*)$ is invertible, 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. By Lemma 4.2 and Proposition 4.2(a), C can be written in the following partitioned form

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

where $I_{\mathcal{B}} = \mathrm{diag}(I_i,\ i \in \mathcal{B})$ 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 simplicity, we here omit the notation ζ^* in the functions. It is not hard to see that these C are nonsingular if and only if

$$C_r = \begin{pmatrix} -P_{\mathcal{I}\mathcal{I}} & P_{\mathcal{I}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) \\ -P_{\mathcal{B}\mathcal{I}} & P_{\mathcal{B}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) + (B_{\mathcal{B}} - I_{\mathcal{B}}) \end{pmatrix}$$

is nonsingular. Showing that the matrix C_r is nonsingular is equivalent to showing that the only solution of the following system

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

is the zero vector. This system can be rewritten as

$$\begin{cases} P_{\mathcal{I}\mathcal{I}} y_{\mathcal{I}} + P_{\mathcal{I}\mathcal{B}}(I_{\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}} = 0, \\ P_{\mathcal{B}\mathcal{I}} y_{\mathcal{I}} + P_{\mathcal{B}\mathcal{B}}(I_{\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}} = -(I_{\mathcal{B}} - B_{\mathcal{B}}) y_{\mathcal{B}}. \end{cases}$$

Recalling that $P_{\mathcal{I}\mathcal{I}}$ is nonsingular, we obtain from the last system that

$$\begin{cases} y_{\mathcal{I}} = -P_{\mathcal{I}\mathcal{I}}^{-1} P_{\mathcal{I}\mathcal{B}}(I_{\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}}, \\ (P_{\mathcal{B}\mathcal{B}} - P_{\mathcal{B}\mathcal{I}} P_{\mathcal{I}\mathcal{I}}^{-1} P_{\mathcal{I}\mathcal{B}})(I_{\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}} = -(I_{\mathcal{B}} - B_{\mathcal{B}}) y_{\mathcal{B}}. \end{cases} \tag{33}$$

Thus, showing the nonsingularity of C_r is equivalent to showing the unique solution of the second equation is a zero vector. We proceed by contradiction. Suppose that $y_{\mathcal{B}} \neq 0$, and consider the following two cases.

Case (1): $(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}} = 0$. Define $J_{\mathcal{B}} := \{i \in \mathcal{B} : (y_{\mathcal{B}})_i \neq 0\}$. Then $J_{\mathcal{B}} \neq \emptyset$. Moreover,

$$(I - A_i(\zeta^*))(y_{\mathcal{B}})_i = 0 \quad \text{and} \quad (I - B_i(\zeta^*))(y_{\mathcal{B}})_i = 0 \quad \text{for all } i \in J_{\mathcal{B}},$$

where the second equality is from the second equation of (33). This means that

$$[(I - A_i(\zeta^*)) + (I - B_i(\zeta^*))](y_{\mathcal{B}})_i = 0, \quad \forall i \in J_{\mathcal{B}}.$$

Note that $(y_{\mathcal{B}})_i \neq 0$ for all $i \in J_{\mathcal{B}}$, and hence the last equation implies that the matrix

$$[2I - A_i(\zeta^*) - B_i(\zeta^*)] \quad \forall i \in J_{\mathcal{B}} \tag{34}$$

is singular. On the other hand, from Lemma 4.2(c), it follows that

$$\begin{aligned} 2I - A_i(\zeta^*) - B_i(\zeta^*) &= 2I - L_{F_i(\zeta^*)}L^{-1}_{z_i(\zeta^*)} - L_{G_i(\zeta^*)}L^{-1}_{z_i(\zeta^*)} \\ &= [2L_{z_i(\zeta^*)} - L_{F_i(\zeta^*)} - L_{G_i(\zeta^*)}]L^{-1}_{z_i(\zeta^*)} \\ &= [L_{2z_i(\zeta^*)} - L_{F_i(\zeta^*)+G_i(\zeta^*)}]L^{-1}_{z_i(\zeta^*)}, \quad \forall i \in \mathcal{B}. \end{aligned} \tag{35}$$

Notice that $w_i(\zeta^*), z_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$ for each $i \in \mathcal{B}$, and furthermore,

$$4z_i(\zeta^*)^2 - [F_i(\zeta^*) + G_i(\zeta^*)]^2 = 2w_i(\zeta^*) + [F_i(\zeta^*) - G_i(\zeta^*)]^2 \in \text{int}(\mathcal{K}^{n_i}).$$

Using Proposition 3.4 of [11] then yields that $[2z_i(\zeta^*) - (F_i(\zeta^*) + G_i(\zeta^*))] \in \text{int}(\mathcal{K}^{n_i})$, which implies that $L_{2z_i(\zeta^*)} - L_{F_i(\zeta^*)+G_i(\zeta^*)} \succ O$. Combining with (35), we obtain that $2I - A_i(\zeta^*) - B_i(\zeta^*)$ for each $i \in \mathcal{J}_{\mathcal{B}}$ is nonsingular. This leads to a contradiction.

Case (2): $(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}} \neq 0$. Notice that $F_i(\zeta^*)^2 + G_i(\zeta^*)^2 \in \text{int}(\mathcal{K}^{n_i})$ for each $i \in \mathcal{B}$ by Lemma 4.2(c), and hence applying Lemma 4.1(a) yields that

$$\langle [(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}]_i, [(B_{\mathcal{B}} - I_{\mathcal{B}})y_{\mathcal{B}}]_i \rangle \leq 0 \quad \text{for } \forall i \in \mathcal{B}.$$

This together with the second equation in (33) means that

$$\langle [(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}]_i, [(P_{\mathcal{BB}} - P_{\mathcal{BI}}P^{-1}_{\mathcal{II}}P_{\mathcal{IB}})(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}]_i \rangle \leq 0, \quad \forall i \in \mathcal{B}.$$

Since $P_{\mathcal{BB}} - P_{\mathcal{BI}}P^{-1}_{\mathcal{II}}P_{\mathcal{IB}}$ is exactly $\widehat{P}_{\mathcal{II}}$, using the Cartesian P-property of $\widehat{P}_{\mathcal{II}}$, this is only possible if $(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}} = 0$, and again we obtain a contradiction. □

From Theorem 4.1 and [19, Lemma 2.6], we readily obtain the following result.

Corollary 4.1 *Suppose that ζ^* is a nondegenerate solution of the SOCCP with $m = n$ and the mappings F and G at ζ^* satisfy the conditions of Theorem 4.1. Then, there exist a neighborhood $\mathcal{N}(\zeta^*)$ of ζ^* and a constant $C_1 > 0$ such that for any $\zeta \in \mathcal{N}(\zeta^*)$ and any $W \in \partial_B \Phi(\zeta)$, W is nonsingular and satisfies $\|W^{-1}\| \leq C_1$.*

5 Properties of the Merit Function Ψ

This section is mainly concerned with the stationary point property and the coerciveness of the function Ψ. Specifically, we shall provide a weaker condition than the one used by [6, Proposition 3] to guarantee that every stationary point of Ψ is a solution of the SOCCP, and show that the function Ψ for the SOCCP (3) is coercive under the uniform Cartesian P-property of F. For the first result, we need the following technical lemma.

Lemma 5.1 *Let $\psi : \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}_+$ be given by (8). Then, for any $x, y \in \mathbb{R}^l$,*

$$\phi(x, y) \neq 0 \quad \Longleftrightarrow \quad \nabla_x \psi(x, y) \neq 0, \qquad \nabla_y \psi(x, y) \neq 0.$$

Proof The equivalence is direct by Proposition 1 and Lemma 6(b) of [6]. □

Proposition 5.1 *Let $\Psi : \mathbb{R}^n \to \mathbb{R}_+$ be given by (10). Suppose that $n = m$ and ∇G is invertible. If $\nabla G(\zeta)^{-1} \nabla F(\zeta)$ at any $\zeta \in \mathbb{R}^n$ has the Cartesian P_0-property, then every stationary point of Ψ is a solution to the SOCCP.*

Proof Since Ψ is continuously differentiable by Proposition 2 of [6] and Φ is locally Lipschitz continuous, we have by Clarke [4] that for any $\zeta \in \mathbb{R}^n$ and any $V \in \partial \Phi(\zeta)^T$

$$\nabla \Psi(\zeta) = V \Phi(\zeta). \tag{36}$$

Let ζ be an arbitrary stationary point of Ψ and V be an element of $\partial_B \Phi(\zeta)^T (\subseteq \partial \Phi(\zeta)^T)$. From (27), it follows that there exist matrices $V_i \in \partial_B \Phi_i(\zeta)^T$ such that

$$V = V_1 \times V_2 \times \cdots \times V_q.$$

In addition, for each $V_i \in \mathbb{R}^{n \times n_i}$, by Proposition 3.1 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 4.2, such that

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

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

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

which, by the invertibility of ∇G, is equivalent to

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

We next prove that $\Phi(\zeta) = 0$. Suppose not, then there is an index $\nu \in \{1, 2, \ldots, q\}$ such that $\Phi_\nu(\zeta) = \phi(F_\nu(\zeta), G_\nu(\zeta)) \neq 0$. From Propositions 3.1 and 4.2, we notice that

$$\begin{pmatrix} \nabla_x \psi(F_\nu(\zeta), G_\nu(\zeta)) \\ \nabla_y \psi(F_\nu(\zeta), G_\nu(\zeta)) \end{pmatrix} = \begin{pmatrix} (A_\nu(\zeta) - I)\Phi_\nu(\zeta) \\ (B_\nu(\zeta) - I)\Phi_\nu(\zeta) \end{pmatrix}.$$

Therefore, applying Lemma 5.1 yields that

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

In addition, from (37) it follows that

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

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

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

Notice that the first term on the left hand side is nonnegative by (38) and the Cartesian P_0-property of $\nabla G(\zeta)^{-1}\nabla F(\zeta)$, and the second term is positive by Lemma 4.1(b) since $\Phi_\nu(\zeta)\neq 0$. This leads to a contradiction. The proof is thus completed. □

When ∇G is invertible, we know from [6] that the column monotonicity of $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ is equivalent to $\nabla G(\zeta)^{-1}\nabla F(\zeta)\succeq O$, which clearly implies that $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ has the Cartesian P_0-property. Thus, the stationary point condition in Proposition 5.1 is weaker than the one used by [6, Proposition 3]. In addition, for the SOCCP (3), the condition is equivalent to saying that F has the Cartesian P_0-property, which, for the NCP, will reduce to the common stationary point condition that F is a P_0-function.

The following lemma generalizes the result of [6, Lemma 9(a)], which plays a crucial role in establishing the coerciveness of the merit function Ψ.

Lemma 5.2 *Let ψ be defined as in (8). For any sequence $\{(x^k,y^k)\}\subseteq\mathbb{R}^l\times\mathbb{R}^l$, 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.*

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

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

Proof Part (a) is direct by [6, Lemma 9 (a)]. We next prove part (b). Suppose that $\{\phi(x^k,y^k)\}$ is bounded. Let $z^k=[(x^k)^2+(y^k)^2]^{1/2}$ for each k. From the definition of ϕ,

$$x^k+y^k=z^k-\phi(x^k,y^k),\quad \forall\,k.$$

Squaring two sides of the last equality then yields that

$$2x^k\circ y^k=-2z^k\circ\phi(x^k,y^k)+\phi(x^k,y^k)^2,\quad \forall k. \tag{39}$$

Since $\|x^k\|\to+\infty$ and $\|y^k\|\to+\infty$ by the given conditions, we have

$$\lim_{k\to\infty}\frac{z^k}{\|x^k\|\|y^k\|}=\lim_{k\to\infty}\left[\frac{(x^k)^2}{\|x^k\|^2\|y^k\|^2}+\frac{(y^k)^2}{\|x^k\|^2\|y^k\|^2}\right]^{1/2}=0,$$

which, together with the boundedness of $\{\phi(x^k, y^k)\}$, implies that

$$\lim_{k\to\infty} \frac{-2z^k \circ \phi(x^k, y^k) + \phi(x^k, y^k)^2}{\|x^k\|\|y^k\|} = 0.$$

Using the equality (39), we obtain $\lim_{k\to\infty} \frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} = 0$, which clearly contradicts the given assumption. Consequently, the conclusion follows. $\square$

It should be pointed out that in Lemma 5.2(b), the condition $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \nrightarrow 0$ as $k \to \infty$ is necessary, which can be illustrated by the following counterexample.

Example 5.1 Consider the sequences $\{x^k\}$ and $\{y^k\}$ given as follows:

$$x^k = \begin{pmatrix} k \\ -(k+1) \\ 0 \end{pmatrix} \quad \text{and} \quad y^k = \begin{pmatrix} k \\ k-1 \\ 0 \end{pmatrix} \quad \text{for each } k.$$

It is easy to verify that $\lambda_1^k = -1$, $\mu_1^k = 1$ for each k, and $\lambda_2^k \to +\infty$, $\mu_2^k \to +\infty$, but

$$\frac{x^k}{\|x^k\|} \to \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}, \qquad \frac{y^k}{\|y^k\|} \to \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \quad \text{and} \quad \frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \to 0.$$

That is, the sequences $\{x^k\}$ and $\{y^k\}$ do not satisfy the assumption $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \nrightarrow 0$. For such sequences, by a simple computation, we have

$$\phi(x^k, y^k) = \frac{1}{2} \begin{pmatrix} \sqrt{4k^2+2+4k} + \sqrt{4k^2+2-4k} - 4k \\ 4 - (\sqrt{4k^2+2+4k} - \sqrt{4k^2+2-4k}) \\ 0 \end{pmatrix}.$$

Since

$$\lim_{k\to\infty} \sqrt{4k^2+2+4k} + \sqrt{4k^2+2-4k} - 4k = 0,$$

$$\lim_{k\to\infty} 4 - (\sqrt{4k^2+2+4k} - \sqrt{4k^2+2-4k}) = 2,$$

we have $\lim_{k\to\infty} \|\phi(x^k, y^k)\| = 1$, i.e. the conclusion of Lemma 5.2(b) does not hold.

We are now in a position to establish the coerciveness of Ψ for the SOCCP (3) under the uniform Cartesian P-property of F and the following condition.

Condition A For any sequence $\{\zeta^k\} \subseteq \mathbb{R}^n$ satisfying $\|\zeta^k\| \to +\infty$, if there exists an index $i \in \{1, 2, \ldots, q\}$ such that $\{\lambda_1(\zeta_i^k)\}$ and $\{\lambda_1(F_i(\zeta^k))\}$ are bounded below, and $\lambda_2(\zeta_i^k), \lambda_2(F_i(\zeta^k)) \to +\infty$, then

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

Proposition 5.2 *For the SOCCP* (3), *suppose that the mapping F has the uniform Cartesian P-property and satisfies Condition* A. *Then, the merit function Ψ is coercive.*

Proof We shall prove that $\lim_{\|\zeta^k\|\to+\infty}\Psi(\zeta^k)=+\infty$. Let $\{\zeta^k\}\subseteq\mathbb{R}^n$ be a sequence such that $\|\zeta^k\|\to+\infty$, where $\zeta^k=(\zeta_1^k,\ldots,\zeta_q^k)$ with $\zeta_i^k\in\mathbb{R}^{n_i}$. Define the index set

$$J:=\big\{i\in\{1,2,\ldots,q\}\mid\{\zeta_i^k\}\text{ is unbounded}\big\}.$$

Since $\{\zeta^k\}$ is unbounded, $J\neq\emptyset$. Let $\{\xi^k\}$ be a bounded sequence with $\xi^k=(\xi_1^k,\ldots,\xi_q^k)$ and $\xi_i^k\in\mathbb{R}^{n_i}$ for each k, where ξ_i^k is defined as follows:

$$\xi_i^k=\begin{cases}0 & \text{if } i\in J,\\ \zeta_i^k & \text{otherwise,}\end{cases}\qquad i=1,2,\ldots,q.$$

By the uniform Cartesian P-property of F, there is a constant $\rho>0$ such that

$$\begin{aligned}\rho\|\zeta^k-\xi^k\|^2&\le\max_{i=1,\ldots,m}\big\langle\zeta_i^k-\xi_i^k,F_i(\zeta^k)-F_i(\xi^k)\big\rangle\\&=\big\langle\zeta_\nu^k,F_\nu(\zeta^k)-F_\nu(\xi^k)\big\rangle\\&\le\|\zeta_\nu^k\|\|F_\nu(\zeta^k)-F_\nu(\xi^k)\|\quad\text{for each }k,\end{aligned}\tag{40}$$

where ν is an index from $\{1,2,\ldots,q\}$ for which the maximum is attained and we have, without loss of generality, assumed to be independent of k. Clearly, $\nu\in J$, which means that $\{\zeta_\nu^k\}$ is unbounded. Consequently, there exists a subsequence, assumed to be $\{\zeta_\nu^k\}$ without loss of generality, such that $\|\zeta_\nu^k\|\to+\infty$. Notice that

$$\|\zeta^k-\xi^k\|^2\ge\|\zeta_\nu^k-\xi_\nu^k\|^2=\|\zeta_\nu^k\|^2,\quad\text{for each }k.$$

Dividing the both sides of (40) by $\|\zeta_\nu^k\|$ then yields that

$$\rho\|\zeta_\nu^k\|\le\|F_\nu(\zeta^k)-F_\nu(\xi^k)\|\le\|F_\nu(\zeta^k)\|+\|F_\nu(\xi^k)\|,$$

which implies $\|F_\nu(\zeta^k)\|\to+\infty$ since $\|\zeta_\nu^k\|\to+\infty$ and $\{F_\nu(\xi^k)\}$ is bounded. Thus,

$$\|\zeta_\nu^k\|\to+\infty\quad\text{and}\quad\|F_\nu(\zeta^k)\|\to+\infty.\tag{41}$$

If either $\lambda_1(\zeta_\nu^k)\to-\infty$ or $\lambda_1(F_\nu(\zeta^k))\to-\infty$, then using Lemma 5.2(a) readily yields that $\psi(\zeta_\nu^k,F_\nu(\zeta^k))\to+\infty$, and consequently, $\Psi(\zeta^k)\to+\infty$. Otherwise, (41) implies that $\{\lambda_1(\zeta_\nu^k)\}$ and $\{\lambda_1(F_\nu(\zeta^k))\}$ are bounded below, but $\lambda_2(\zeta_\nu^k)\to+\infty$ and $\lambda_2(F_\nu(\zeta^k))\to+\infty$. Using Condition A, it then follows that

$$\limsup_{k\to\infty}\left\langle\frac{\zeta_\nu^k}{\|\zeta_\nu^k\|},\frac{F_\nu(\zeta^k)}{\|F_\nu(\zeta^k)\|}\right\rangle>0,$$

which in turn implies that

$$\limsup_{k\to\infty}\lambda_2\left[\frac{\zeta_\nu^k}{\|\zeta_\nu^k\|}\circ\frac{F_\nu(\zeta^k)}{\|F_\nu(\zeta^k)\|}\right]>0.$$

From this, we have $\frac{\zeta_\nu^k}{\|\zeta_\nu^k\|} \circ \frac{F_\nu(\zeta^k)}{\|F_\nu(\zeta^k)\|} \nrightarrow 0$. This shows that the sequences $\{\zeta_\nu^k\}$ and $\{F_\nu(\zeta^k)\}$ satisfy the conditions of Lemma 5.2(b), and therefore $\Psi(\zeta^k) \to +\infty$. □

When $n_1 = \cdots = n_q = 1$, Condition A automatically holds and the uniform Cartesian P-property of F is equivalent to F being a uniform P-function. Thus, Proposition 5.2 recovers the result of the FB merit function for the NCP; see [12, Theorem 4.2].

6 A Damped Gauss-Newton Method

Based on the previous discussions, we in this section describe a damped Gauss-Newton method for the SOCCP by applying the generalized Newton method for the semismooth system (9). The algorithm is similar to the one proposed by Sun and Womersley [22] for box constrained variational inequality problem.

Algorithm 6.1

Step 0. Choose $\zeta^0 \in \mathbb{R}^n$, $\rho \in (0,1)$, $\sigma \in (0, 1/2)$ and $p_1, p_2 > 0$. Set $k := 0$.
Step 1. If $\|\nabla\Psi(\zeta^k)\| = 0$, then stop.
Step 2. Select an element $W_k \in \partial_B \Phi(\zeta^k)$. Let d^k be the solution of the linear system

$$\big(W_k^T W_k + p_1\|\Phi(\zeta^k)\|^{p_2} I\big)d = -\nabla\Psi(\zeta^k). \tag{42}$$

Step 3. Let l_k be the smallest nonnegative integer l such that

$$\Psi(\zeta^k + \rho^l d^k) \le \Psi(\zeta^k) + \sigma\rho^l \nabla\Psi(\zeta^k)^T d^k. \tag{43}$$

Step 4. Set $\zeta^{k+1} := \zeta^k + \rho^{l_k} d^k$, $k := k+1$ and go to Step 1.

Note that $W_k^T W_k + p_1\|\Phi(\zeta^k)\|^{p_2} I$ is positive definite if ζ^k is not a solution of the SOCCP, and hence $\nabla\Psi(\zeta^k)^T d^k < 0$. This means that Algorithm 6.1 is well defined at the kth iteration. In addition, if $n = m$ and W_k is nonsingular, then $W_k^T W_k \succ O$ and the solution of (42) with $p_1 = 0$ yields a generalized Newton direction $d^k = -W_k^{-1}\Phi(\zeta^k)$.

For the above damped Gauss-Newton method, along the lines of the proof of [10, Theorem 15], we can obtain the following global convergence result.

Theorem 6.1 *Suppose that $\{\zeta^k\}$ is a sequence generated by Algorithm* 6.1. *Then, each accumulation point ζ^* of $\{\zeta^k\}$ is a stationary point of Ψ.*

Using Proposition 4.1 and Corollary 4.1 and the proof of [22, Theorem 7.2], we can prove the following superlinear (quadratic) convergence result.

Theorem 6.2 *Suppose that $\{\zeta^k\}$ is a sequence generated by Algorithm 6.1 and ζ^* is an accumulation point of $\{\zeta^k\}$. If ζ^* is nondegenerate and F and G at ζ^* satisfy the conditions of Theorem 4.1, then the sequence $\{\zeta^k\}$ converges to ζ^* Q-superlinearly.*

Furthermore, if F' and G' are Lipschitz continuous at ζ^* and $p_2 \geq 1$, then the convergence is Q-quadratic.

Corollary 6.1 *For the SOCCP (3), if F has the uniform Cartesian P-property and satisfies Condition A, then the sequence $\{\zeta^k\}$ given by Algorithm 6.1 converges to the unique solution ζ^*. If, in addition, ζ^* is nondegenerate, $\nabla F(\zeta^*)_{\mathcal{I}\mathcal{I}}$ is nonsingular and*

$$\begin{pmatrix} \nabla F(\zeta^*)_{\mathcal{I}\mathcal{I}} & \nabla F(\zeta^*)_{\mathcal{I}\mathcal{B}} \\ \nabla F(\zeta^*)_{\mathcal{B}\mathcal{I}} & \nabla F(\zeta^*)_{\mathcal{B}\mathcal{B}} \end{pmatrix} \tag{44}$$

has the Cartesian P-property, then $\{\zeta^k\}$ converges to ζ^* superlinearly. If F' is also locally Lipschitz continuous around ζ^* and $p_2 \geq 1$, then the convergence is Q-quadratic.

Proof We first show that the SOCCP (3) has the unique solution ζ^*. If not, let $\xi^* \neq \zeta^*$ be another solution of the SOCCP (3). From the uniform Cartesian P-property of F, there is a positive constant ρ such that

$$\langle F(\zeta^*) - F(\xi^*), \zeta^* - \xi^* \rangle \geq \rho \|\zeta^* - \xi^*\|^2 > 0.$$

On the other hand, since $F(\zeta^*)^T \zeta^* = 0$, $F(\xi^*)^T \xi^* = 0$ and $\zeta^*, \xi^* \in \mathcal{K}$, it follows that

$$\langle F(\zeta^*) - F(\xi^*), \zeta^* - \xi^* \rangle = -F(\zeta^*)^T \xi^* - F(\xi^*)^T \zeta^* \leq 0.$$

This gives a contradiction. Consequently, the SOCCP (3) has the unique solution ζ^*. Since $\zeta^k \subseteq \{\zeta \in \mathbb{R}^n \mid \Psi(\zeta) \leq \Psi(\zeta^0)\}$, by Proposition 5.2, the sequence $\{\zeta^k\}$ is bounded. Using Theorem 6.1, we can prove that the sequence $\{\zeta^k\}$ converges to ζ^*.

By Sect. 2 of [17], if ∇F has the Cartesian P-property, then its principal block matrix in (44) also has the Cartesian P-property. If, in addition, $\nabla F(\zeta^*)_{\mathcal{I}\mathcal{I}}$ is nonsingular, then its Schur-complement in the matrix of (44) has the Cartesian P-property by Proposition 2.1 of [17]. Thus, the conditions of Theorem 4.1 are satisfied, and the second part of the conclusion is direct by Theorem 6.2. □

7 Numerical Results

In this section, we report the numerical results with Algorithm 6.1 solving the linear SOCP, i.e. the SOCP (4) with $g(x) = c^T x$ for $c \in \mathbb{R}^m$. We used Algorithm 6.1 to solve the KKT system of (4), which is equivalent to the SOCCP with F and G given by (5). The vector $\hat{x}$ in F was computed as a solution of $\min_x \|Ax - b\|$ by Matlab's least square solver, and F and G were evaluated via the Cholesky factorization of AA^T.

All experiments were done with a PC of 2.8 GHz CPU and 512MB memory. The computer codes were all written in Matlab 6.5. We replaced the monotone line search of Algorithm 6.1 with a nonmonotone version as described by Zhang and Hager [25], i.e., we computed the smallest nonnegative integer l such that

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

Table 1 Numerical results with Algorithm 6.1 and SeDuMi for the SOCPs

Problem	Dimension		Gauss Newton method				SeDuMi	
	p	m	$\Psi(\zeta^k)$	Iter	NF	$c^T x^k$	Iter	$c^T x^k$
nb	123	2383	8.42e–7	34	71	−5.070410e–2	21	−5.070310e–2
nb_L1	915	3176	5.70e–7	109	122	−1.301227e+1	18	−1.301227e+1
nb_L2_bessel	123	2641	2.04e–10	9	14	−1.025697e–1	16	−1.025695e–1

where

$$\mathcal{W}_k = (\eta_{k-1} Q_{k-1} \mathcal{W}_{k-1} + \Psi(\zeta^k))/Q_k$$

with $Q_k = \eta_{k-1} Q_{k-1} + 1$. During the tests, we used $\mathcal{W}_0 = \Psi(\zeta^0)$, $Q_0 = 1$ and $\eta_k \equiv 0.85$ for the line search, and $\rho = 0.5$ and $\sigma = 1.0e - 4$ for Algorithm 6.1. In addition, we adopted a pure Gauss-Newton direction $d^k = -W_k^{-1}\Phi(\zeta^k)$, which is in fact a generalized Newton direction. We started the algorithm with $\zeta^0 = 0$ and solved the linear system involved in the algorithm by Matlab's linear equation solver. The algorithm was stopped whenever one of the following conditions was satisfied: (1) $\max\{\Psi(\zeta^k), |F(\zeta^k)^T G(\zeta^k)|\} \le 10^{-6}$; (2) the steplength is less than 10^{-15}; (3) the number of iteration is over than 150.

We have solved several SOCPs from the DIMACS library [18] and compared the numerical performance of Algorithm 6.1 with SeDuMi [23], a successful interior point method software for the SOCP and the semidefinite programming. Numerical results are listed in Table 1, where **Iter** records the number of iterations, **NF** represents the number of function evaluations required by the algorithm for solving each problem, $\Psi(\zeta^k)$ and $c^T x^k$ denote the value of $\Psi(\zeta)$ and $c^T x$ at the final iteration, respectively.

From Table 1, we see that Algorithm 6.1 yielded a solution with favorable accuracy for all test problems within 110 iterations, and needed less iterations for the problem "nb_L2_bessel". For the more difficult problems "nb" and "nb_L1", Algorithm 6.1 is now not comparable with the sophisticated software SeDuMi in terms of the number of iterations. We observe that the two problems do not satisfy the nondegenerate condition at the optimal solution. It should be pointed out that our code is crude and does not exploit any preprocessing strategy on the problems.

8 Conclusions

The FB function $\phi(x, y)$ is an important SOC complementarity function, by which the SOCCP can be reformulated as a semismooth system involving the operator Φ. In this paper, we have characterized the B-subdifferential of ϕ at a general point and presented an estimate for the B-subdifferential of Φ at any point. A condition was also given to guarantee every element of the B-subdifferential of Φ at a solution to be nonsingular. Although the condition is a little stringent, this is the first article, to our best knowledge, to discover the B-subdifferential of FB function associated with SOCs. In addition, we have established the coerciveness of the merit function

Ψ and provided a weaker condition than [6] for each stationary point of Ψ to be a solution of the SOCCP. With these results, a damped Gauss-Newton method was proposed and the global and local convergence results were obtained under some suitable conditions.

Acknowledgements The authors would like to thank Prof. P. Tseng for the codes of the merit function approach and Prof. D.-F. Sun for his suggestions on this work. The authors are also grateful to the two referees for their valuable comments and suggestions.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**, 3–51 (2003)
2. Andersen, E.D., Roos, C., Terlaky, T.: On implementing a primal-dual interior-point method for conic quadratic optimization. Math. Program. **95**, 249–277 (2003)
3. Chen, J.-S., Chen, X., Tseng, P.: Analysis of nonsmooth vector-valued functions associated with second-order cones. Math. Program. **101**, 95–117 (2004)
4. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983). Reprinted by SIAM, Philadelphia (1990)
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. Fischer, A.: Solution of the monotone complementarity problem with locally Lipschitzian functions. Math. Program. **76**, 513–532 (1997)
9. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs, Oxford University Press, New York (1994)
10. Facchinei, F., Kanzow, C.: A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems. Math. Program. **76**, 493–512 (1997)
11. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2002)
12. Facchinei, F., Soares, J.: A new merit function for nonlinear complementarity problem and a related algorithm. SIAM J. Optim. **7**, 225–247 (1997)
13. 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)
14. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Application of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
15. Mifflin, M.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 957–972 (1977)
16. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions. Math. Program. **88**, 61–83 (2000)
17. Pan, S.-H., Chen, J.-S.: A regularization method for the second-order cone complementarity problems with Cartesian P_0-property, Nonlinear Analysis (to appear)
18. Pataki, G., Schmieta, S.: The DIMACS library of semidefinite-quadratic-linear programs. Preliminary draft, Computational Optimization Research Center, Columbia University, New York. http://dimacs.rutgers.edu/Challenges/Seventh/Instances/
19. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244 (1993)
20. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)
21. Sun, D., Sun, J.: Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions. Math. Program. **103**, 575–581 (2005)
22. Sun, D., Womersley, R.S.: A new unconstrained differential merit function for box constrained variational inequality problems and a damped Gauss-Newton method. SIAM J. Optim. **9**, 388–413 (1999)
23. Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Software **11&12**, 625–653 (1999)

24. 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)
25. Zhang, H.-C., Hager, W.W.: A nonmonotone line search technique and its application to unconstrained optimization. SIAM J. Optim. **14**, 1043–1056 (2004)
26. Yoshise, A.: Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones. SIAM J. Optim. **17**, 1129–1153 (2006)