

A least-square semismooth Newton method for the second-order cone complementarity problem

Shaohua Pan ¹

School of Mathematical Sciences
South China University of Technology
Guangzhou 510640, China

Jein-Shan Chen ²

Department of Mathematics
National Taiwan Normal University
Taipei 11677, Taiwan

March 1, 2008

(First revised November 5, 2008)

(Second revised May 10, 2009)

Abstract. We present a nonlinear least-square formulation for the second-order cone complementarity problem based on the Fischer-Burmeister (FB) function and the plus function. The formulation has twofold advantages. Firstly, the operator involved in the overdetermined system of equations inherits the favorable properties of the FB function for local convergence, for example, the (strong) semismoothness. Secondly, the natural merit function of the overdetermined system of equations share all the nice features of the class of merit functions f_{YF} studied in [4] for global convergence. We propose a semismooth Levenberg-Marquardt method to solve the arising overdetermined system of equations, and establish the global and local convergence results. Among others, the superlinear (quadratic) rate of convergence is obtained under the strict complementarity of the solution and a local error bound assumption, respectively. Numerical results verify the advantages of the least-square reformulation for more difficult problems.

Key words. Second-order cone complementarity problem, Fischer-Burmeister function, semismooth, Levenberg-Marquardt method.

¹E-mail: shhpan@scut.edu.cn.

²Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan. E-mail: jschen@math.ntnu.edu.tw, FAX: 886-2-29332342.

1 Introduction

We consider the second-order cone complementarity problem (SOCCP): to find $\zeta \in \mathbb{R}^n$ such that

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

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ are assumed to be continuously differentiable throughout this paper, and \mathcal{K} is the Cartesian product of second-order cones (SOCs), also called Lorentz cones [9]. In other words,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_q}, \quad (2)$$

where $q, n_1, \dots, n_q \geq 1$, $n_1 + n_2 + \cdots + n_q = n$, and

$$\mathcal{K}^{n_i} := \{(x_{i1}, x_{i2}) \in \mathbb{R} \times \mathbb{R}^{n_i-1} \mid x_{i1} \geq \|x_{i2}\|\},$$

with $\|\cdot\|$ denoting the Euclidean norm and \mathcal{K}^1 denoting the set of nonnegative real numbers. In the rest of this paper, corresponding to the Cartesian structure of \mathcal{K} , we write $F = (F_1, \dots, F_q)$ and $G = (G_1, \dots, G_q)$ with F_i and G_i being mappings from \mathbb{R}^n to \mathbb{R}^{n_i} .

An important special case of (1) corresponds to $G(\zeta) \equiv \zeta$, and then (1) reduce to

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

This is a natural extension of the nonlinear complementarity problem (NCP) [10, 11], where $\mathcal{K} = \mathbb{R}_+^n$, the nonnegative orthant in \mathbb{R}^n , corresponds to $n_1 = \cdots = n_q = 1$ and $q = n$. Another important special case of (1) corresponds to the Karush-Kuhn-Tucker (KKT) conditions of the convex second-order cone program (SOCP):

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

where $g: \mathbb{R}^n \rightarrow \mathbb{R}$ is a twice continuously differentiable convex function, $A \in \mathbb{R}^{m \times n}$ has full row rank, and $b \in \mathbb{R}^m$. The KKT conditions of (4) can be rewritten as (1) with

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

where $\hat{x} \in \mathbb{R}^n$ satisfies $Ax = b$; see [4] for details. The convex SOCP arises in many applications from engineering design, finance, and robust optimization; see [1, 21] and references therein. Motivated by [18] where the three-dimensional quasi-static frictional contact was directly reformulated as a linear SOC complementarity problem, we believe that, besides these applications, the SOCCP (1) will be found to have some applications in engineering which can not reduce to SOCPs.

There have been proposed various methods for solving convex SOCPs and SOCCPs. They include the interior point methods [1, 2, 21, 23, 30, 32], the smoothing Newton

methods [6, 13, 16], the merit function method [4] and the semismooth Newton method [17], where the last three kinds of methods are all based on an SOC complementarity function or a merit function. Specifically, we call $\phi : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_i}$ ($\psi : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \rightarrow \mathbb{R}_+$) an *SOC complementarity function (a merit function)* associated with \mathcal{K}^{n_i} if

$$\phi(x_i, y_i) = 0 \quad (\psi(x_i, y_i) = 0) \iff x \in \mathcal{K}^{n_i}, \quad y \in \mathcal{K}^{n_i}, \quad \langle x_i, y_i \rangle = 0. \quad (6)$$

Clearly, when $n_i = 1$, an SOC complementarity function reduces to an NCP function.

A popular choice of ϕ is the Fischer-Burmeister (FB) function $\phi_{\text{FB}} : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_i}$, defined by

$$\phi_{\text{FB}}(x_i, y_i) := (x_i^2 + y_i^2)^{1/2} - (x_i + y_i), \quad (7)$$

where $x_i^2 = x_i \circ x_i$ means the Jordan product of x_i with itself, $x_i^{1/2}$ is a vector such that $(x_i^{1/2})^2 = x_i$, and $x_i + y_i$ denotes the usual componentwise addition of vectors. This function is well defined for all $x_i, y_i \in \mathbb{R}^{n_i}$, and was shown in [13] to satisfy (6). Hence, the SOCCP (1) can be reformulated as the following system of nonsmooth equations

$$\Phi_{\text{FB}}(\zeta) := \begin{pmatrix} \phi_{\text{FB}}(F_1(\zeta), G_1(\zeta)) \\ \vdots \\ \phi_{\text{FB}}(F_q(\zeta), G_q(\zeta)) \end{pmatrix} = 0, \quad (8)$$

which induces a natural merit function $\Psi_{\text{FB}} : \mathbb{R}^n \rightarrow \mathbb{R}_+$ for (1), defined by

$$\Psi_{\text{FB}}(\zeta) := \frac{1}{2} \|\Phi_{\text{FB}}(\zeta)\|^2 = \sum_{i=1}^q \psi_{\text{FB}}(F_i(\zeta), G_i(\zeta)) \quad (9)$$

with

$$\psi_{\text{FB}}(x_i, y_i) := \frac{1}{2} \|\phi_{\text{FB}}(x_i, y_i)\|^2. \quad (10)$$

The function ψ_{FB} was well-studied in [4] and used to develop a merit function approach. Recently, we analyzed in [24] that, to guarantee the boundedness of the level sets of the FB merit function Ψ_{FB} , it requires that the mapping F at least has the uniform Cartesian P -property. This means that ϕ_{FB} has some limitations in handling monotone SOCCPs.

Motivated by the work [19] for the NCPs, in this paper we give a new reformulation for (1) to overcome the disadvantage of ϕ_{FB} . Let $\phi_0 : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \rightarrow \mathbb{R}_+$ be given by

$$\phi_0(x_i, y_i) := \max \{0, x_i^T y_i\}, \quad (11)$$

and define the operator $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^{n+q}$ as

$$\Phi(\zeta) := \begin{pmatrix} \rho_1 \phi_{\text{FB}}(F_1(\zeta), G_1(\zeta)) \\ \vdots \\ \rho_1 \phi_{\text{FB}}(F_q(\zeta), G_q(\zeta)) \\ \rho_2 \phi_0(F_1(\zeta), G_1(\zeta)) \\ \vdots \\ \rho_2 \phi_0(F_q(\zeta), G_q(\zeta)) \end{pmatrix}, \quad (12)$$

where ρ_1, ρ_2 are arbitrary but fixed constants from $(0, 1)$ used as the weights between the first type of terms and the second one. In other words, we define Φ by appending q components to the mapping Φ_{FB} . These additional components, as will be shown later, play a crucial role in overcoming the disadvantage of Ψ_{FB} mentioned above. Noting that

$$\zeta^* \text{ solves } \Phi(\zeta) = 0 \iff \zeta^* \text{ solves (1),} \quad (13)$$

we have the following nonlinear least-square reformulation for the SOCCP (1)

$$\min_{\zeta \in \mathbb{R}^n} \Psi(\zeta) := \frac{1}{2} \|\Phi(\zeta)\|^2 = \sum_{i=1}^q \psi(F_i(\zeta), G_i(\zeta)), \quad (14)$$

where

$$\psi(x_i, y_i) := \rho_1^2 \psi_{\text{FB}}(x_i, y_i) + \frac{1}{2} \rho_2^2 \phi_0(x_i, y_i)^2. \quad (15)$$

The reformulation has the following advantages: Ψ belongs to the class of merit functions f_{YF} introduced in [4], which will be shown to have more desirable properties than Ψ_{FB} ; and Φ inherits the semismoothness of Φ_{FB} even strong semismoothness under some conditions. By this, we propose a semismooth Levenberg-Marquardt type method for solving (14), and establish the superlinear (quadratic) rate of convergence under the strict complementarity of the solution and a local error bound assumption, respectively.

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, $(x_1, \dots, x_q) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_q}$ is viewed as a column vector in $\mathbb{R}^{n_1 + \cdots + n_q}$. For a differentiable mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\nabla F(x)$ denotes the transpose of the Jacobian $F'(x)$. For a (not necessarily symmetric) square 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 matrices Q_1, \dots, Q_n , we denote the block diagonal matrix with these matrices as block diagonals by $\text{diag}(Q_1, \dots, Q_n)$. If \mathcal{J} and \mathcal{B} are index sets such that $\mathcal{J}, \mathcal{B} \subseteq \{1, 2, \dots, q\}$, 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 $j \in \mathcal{J}$ and $k \in \mathcal{B}$. We denote $\text{int}(\mathcal{K}^n)$ and $\text{bd}(\mathcal{K}^n)$ by the interior and the boundary of \mathcal{K}^n , respectively, and denote $\text{bd}^+(\mathcal{K}^n)$ by the boundary of \mathcal{K}^n excluding the origin.

2 Preliminaries

This section recalls some background materials that will be used in the sequel. We start with the definition of Jordan product. For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their *Jordan product* [9] associated with \mathcal{K}^n as

$$x \circ y := (\langle x, y \rangle, x_1 y_2 + y_1 x_2). \quad (16)$$

The Jordan product “ \circ ”, unlike scalar or matrix multiplication, is not associative, which is a main source on complication in the analysis of SOCCPs. The identity element under this product is $e := (1, 0, \dots, 0)^T \in \mathbb{R}^n$. Given a vector $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let

$$L_x := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix}, \quad (17)$$

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

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \frac{\det(x)}{x_1} I + \frac{1}{x_1} x_2 x_2^T \end{bmatrix}, \quad (18)$$

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

From [9, 13], we recall that each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ admits a spectral factorization associated with \mathcal{K}^n , of the form

$$x = \lambda_1(x) \cdot u_x^{(1)} + \lambda_2(x) \cdot u_x^{(2)}, \quad (19)$$

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

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

with $\bar{x}_2 = x_2 / \|x_2\|$ if $x_2 \neq 0$ and otherwise being any vector in \mathbb{R}^{n-1} satisfying $\|\bar{x}_2\| = 1$. If $x_2 \neq 0$, the factorization is unique. The spectral factorization of x , x^2 as well as $x^{1/2}$ have various interesting properties (see [13]). We here list some that will be used later.

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

- (a) $x \in \mathcal{K}^n \iff 0 \leq \lambda_1(x) \leq \lambda_2(x)$, and $x \in \text{int}(\mathcal{K}^n) \iff 0 < \lambda_1(x) \leq \lambda_2(x)$.
- (b) $x^2 = [\lambda_1(x)]^2 \cdot u_x^{(1)} + [\lambda_2(x)]^2 \cdot u_x^{(2)} \in \mathcal{K}^n$.
- (c) $x^{1/2} = \sqrt{\lambda_1(x)} \cdot u_x^{(1)} + \sqrt{\lambda_2(x)} \cdot u_x^{(2)} \in \mathcal{K}^n$ if $x \in \mathcal{K}^n$.

We next present Cartesian P -properties for a matrix and a nonlinear transformation.

Definition 2.1 [5] *A matrix $M \in \mathbb{R}^{n \times n}$ is said to have*

- (a) the Cartesian P -property if for any nonzero $\zeta = (\zeta_1, \dots, \zeta_q) \in \mathbb{R}^n$ with $\zeta_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \dots, q\}$ such that $\langle \zeta_\nu, (M\zeta)_\nu \rangle > 0$;
- (b) the Cartesian P_0 -property if for any nonzero $\zeta = (\zeta_1, \dots, \zeta_q) \in \mathbb{R}^n$ with $\zeta_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \dots, q\}$ such that $\zeta_\nu \neq 0$ and $\langle \zeta_\nu, (M\zeta)_\nu \rangle \geq 0$.

Definition 2.2 [5] The mappings $F = (F_1, \dots, F_q), G = (G_1, \dots, G_q)$ are said to have

- (a) the joint uniform Cartesian P -property if there exists a constant $\rho > 0$ such that, for any $\zeta, \xi \in \mathbb{R}^n$, there exists $\nu \in \{1, 2, \dots, q\}$ such that

$$\langle F_\nu(\zeta) - F_\nu(\xi), G_\nu(\zeta) - G_\nu(\xi) \rangle \geq \rho \|\zeta - \xi\|^2.$$

- (b) the joint Cartesian P -property if for any $\zeta, \xi \in \mathbb{R}^n$ with $G(\zeta) \neq G(\xi)$, there exists $\nu \in \{1, 2, \dots, q\}$ such that $\langle F_\nu(\zeta) - F_\nu(\xi), G_\nu(\zeta) - G_\nu(\xi) \rangle > 0$.
- (c) the joint Cartesian P_0 -property if for any $\zeta, \xi \in \mathbb{R}^n$ with $G(\zeta) \neq G(\xi)$, there exists $\nu \in \{1, 2, \dots, q\}$ such that $G_\nu(\zeta) \neq G_\nu(\xi)$ and $\langle F_\nu(\zeta) - F_\nu(\xi), G_\nu(\zeta) - G_\nu(\xi) \rangle \geq 0$.

When $G(\zeta) \equiv \zeta$, Definition 2.2 gives Cartesian P -properties of the mapping F . Obviously, the uniform Cartesian P -property \implies the Cartesian P -property \implies the Cartesian P_0 -property. Also, a continuously differentiable mapping has the Cartesian P_0 -property if and only if its Jacobian at every point has the Cartesian P_0 -property, and if the Jacobian of a continuously differentiable mapping has the Cartesian P -property at every point, then the mapping has the Cartesian P -property. From Definition 2.1, we also see that the positive semidefiniteness of a matrix implies its Cartesian P_0 -property.

Given a mapping $H : \mathbb{R}^n \rightarrow \mathbb{R}^m$, if H is locally Lipschitz continuous, then the set

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

is nonempty and is called the B-subdifferential of H at ζ , where $D_H \subseteq \mathbb{R}^n$ denotes the set of points at which H is differentiable. The convex hull $\partial H(\zeta) := \text{conv} \partial_B H(\zeta)$ is the generalized Jacobian of H at ζ in the sense of Clarke [7]. For the concepts of (strongly) semismooth functions, please refer to [26, 27] for details.

3 Properties of the operator Φ

In this section, we study several important properties of the operator Φ . We first present two technical lemmas to summarize some properties of ϕ_{FB} and ϕ_0 , respectively. The results of the first lemma can be found in [13, Proposition 4.2], [4, Proposition 2], [28, Corollary 3.3] and [24, Proposition 3.1], and the results of the second lemma are direct.

Lemma 3.1 Let $\phi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be defined by (7). Then the following results hold.

- (a) $\phi_{\text{FB}}(x, y) = 0 \iff x \in \mathcal{K}^n, y \in \mathcal{K}^n$ and $\langle x, y \rangle = 0$.
- (b) The squared norm of ϕ_{FB} , namely ψ_{FB} , is continuously differentiable everywhere.
- (c) ϕ_{FB} is strongly semismooth everywhere.
- (d) For any given $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, each element $[U_x - I \ U_y - I]$ of the B-subdifferential $\partial_B \phi_{\text{FB}}(x, y)$ has the following representation:
 - (d.1) If $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, then $U_x = L_{(x^2+y^2)^{1/2}}^{-1} L_x$ and $U_y = L_{(x^2+y^2)^{1/2}}^{-1} L_y$.
 - (d.2) If $x^2 + y^2 \in \text{bd}(\mathcal{K}^n)$ and $(x, y) \neq (0, 0)$, then $[U_x, U_y]$ belongs to the set

$$\left\{ \left[\begin{array}{l} \frac{1}{2\sqrt{2w_1}} \left(\begin{array}{cc} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - \bar{w}_2 \bar{w}_2^T \end{array} \right) L_x + \frac{1}{2} \left(\begin{array}{c} 1 \\ -\bar{w}_2 \end{array} \right) u^T, \\ \frac{1}{2\sqrt{2w_1}} \left(\begin{array}{cc} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - \bar{w}_2 \bar{w}_2^T \end{array} \right) L_y + \frac{1}{2} \left(\begin{array}{c} 1 \\ -\bar{w}_2 \end{array} \right) v^T \end{array} \right] \mid \\ u = (u_1, u_2), v = (v_1, v_2) \text{ satisfy } |u_1| \leq \|u_2\| \leq 1, |v_1| \leq \|v_2\| \leq 1 \right\},$$

where $w = (w_1, w_2) := x^2 + y^2$, $\bar{w}_2 = w_2 / \|w_2\|$.

- (d.3) If $(x, y) = (0, 0)$, then $[V_x, V_y]$ belongs to $\left\{ [L_{\hat{u}}, L_{\hat{v}}] \mid \|\hat{u}\|^2 + \|\hat{v}\|^2 = 1 \right\}$ or

$$\left\{ \left[\begin{array}{l} \frac{1}{2} \left(\begin{array}{c} 1 \\ \bar{w}_2 \end{array} \right) \xi^T + \frac{1}{2} \left(\begin{array}{c} 1 \\ -\bar{w}_2 \end{array} \right) u^T + 2 \left(\begin{array}{cc} 0 & 0 \\ 0 & (I - \bar{w}_2 \bar{w}_2^T) \end{array} \right) L_s, \\ \frac{1}{2} \left(\begin{array}{c} 1 \\ \bar{w}_2 \end{array} \right) \eta^T + \frac{1}{2} \left(\begin{array}{c} 1 \\ -\bar{w}_2 \end{array} \right) v^T + 2 \left(\begin{array}{cc} 0 & 0 \\ 0 & (I - \bar{w}_2 \bar{w}_2^T) \end{array} \right) L_\omega \end{array} \right] \mid \\ \bar{w}_2 \in \mathbb{R}^{n-1} \text{ satisfies } \|\bar{w}_2\| = 1 \text{ and } u = (u_1, u_2), v = (v_1, v_2), \xi = (\xi_1, \xi_2), \\ \eta = (\eta_1, \eta_2), s = (s_1, s_2), \omega = (\omega_1, \omega_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \text{ satisfy } |\xi_1| \leq \|\xi_2\| \leq 1, \\ |u_1| \leq \|u_2\| \leq 1, |\eta_1| \leq \|\eta_2\| \leq 1, |v_1| \leq \|v_2\| \leq 1, \|s\|^2 + \|\omega\|^2 \leq 1/2 \right\}.$$

Lemma 3.2 Let $\phi_0 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ be defined as in (11). Then,

- (a) the square of ϕ_0 is continuously differentiable everywhere;
- (b) ϕ_0 is strongly semismooth everywhere on $\mathbb{R}^n \times \mathbb{R}^n$;
- (c) the B-subdifferential $\partial_B \phi_0(x, y)$ of ϕ_0 at any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ is given by

$$\partial_B \phi_0(x, y) = [\partial_B(x^T y)_{+y^T} \ \partial_B(x^T y)_{+x^T}],$$

where

$$\partial_B(x^T y)_+ = \begin{cases} \{1\} & \text{if } x^T y > 0, \\ \{1, 0\} & \text{if } x^T y = 0, \\ \{0\} & \text{if } x^T y < 0. \end{cases}$$

Using Lemma 3.1 (b) and Lemma 3.2 (b), we readily get the semismoothness of Φ .

Proposition 3.1 *The operator $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^{n+q}$ defined by (12) is semismooth. If, in addition, F' and G' are Lipschitz continuous, then Φ is strongly semismooth.*

Proof. Let Φ_i denote the i -th component function of Φ for $i = 1, 2, \dots, 2q$, i.e., $\Phi_i(\zeta) = \phi_{\text{FB}}(F_i(\zeta), G_i(\zeta))$ for $i = 1, 2, \dots, q$ and $\Phi_i(\zeta) = \phi_0(F_i(\zeta), G_i(\zeta))$ for $i = q + 1, \dots, 2q$. Then, the mapping Φ is (strongly) semismooth if every Φ_i is (strongly) semismooth. For $i = 1, 2, \dots, q$, $\Phi_i: \mathbb{R}^n \rightarrow \mathbb{R}^{n_i}$ is the composite of the strongly semismooth function ϕ_{FB} and the smooth function $\zeta \mapsto (F_i(\zeta), G_i(\zeta))$, whereas $\Phi_{q+i}: \mathbb{R}^n \rightarrow \mathbb{R}$ is the composite of the strongly semismooth function ϕ_0 and the function $\zeta \mapsto (F_i(\zeta), G_i(\zeta))$. Moreover, when F' and G' are Lipschitz continuous, $\zeta \mapsto (F_i(\zeta), G_i(\zeta))$ is strongly semismooth. By [12, Theorem 19], we have that every component function of Φ is semismooth, and strongly semismooth if F' and G' are Lipschitz continuous. \square

Next we give an estimation for the B-subdifferential of the operator Φ at any $\zeta \in \mathbb{R}^n$.

Proposition 3.2 *Let $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^{n+q}$ be defined by (12). Then, for any given $\zeta \in \mathbb{R}^n$,*

$$\partial_B \Phi(\zeta)^T \subseteq \nabla F(\zeta) [\rho_1(A(\zeta) - I) \quad \rho_2 C(\zeta)] + \nabla G(\zeta) [\rho_1(B(\zeta) - I) \quad \rho_2 D(\zeta)]$$

where $C(\zeta) = \text{diag}(C_1(\zeta), \dots, C_q(\zeta))$ and $D(\zeta) = \text{diag}(D_1(\zeta), \dots, D_q(\zeta))$ with

$$C_i(\zeta) \in G_i(\zeta) \partial_B(F_i(\zeta)^T G_i(\zeta))_+ \quad \text{and} \quad D_i(\zeta) \in F_i(\zeta) \partial_B(F_i(\zeta)^T G_i(\zeta))_+,$$

and $A(\zeta) = \text{diag}(A_1(\zeta), \dots, A_q(\zeta))$ and $B(\zeta) = \text{diag}(B_1(\zeta), \dots, B_q(\zeta))$ with the block diagonals $A_i(\zeta), B_i(\zeta) \in \mathbb{R}^{n_i \times n_i}$ having the following representation:

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

(b) *If $F_i(\zeta)^2 + G_i(\zeta)^2 \in \text{bd}^+(\mathcal{K}^{n_i})$, then $[A_i(\zeta), G_i(\zeta)]$ belongs to the set*

$$\left\{ \left[\frac{1}{2\sqrt{2w_{i1}(\zeta)}} L_{F_i(\zeta)} \begin{pmatrix} 1 & \bar{w}_{i2}(\zeta)^T \\ \bar{w}_{i2}(\zeta) & 4I - \bar{w}_{i2}(\zeta) \bar{w}_{i2}(\zeta)^T \end{pmatrix} + \frac{1}{2} u_i (1, -\bar{w}_{i2}(\zeta)^T), \right. \right. \\ \left. \left. \frac{1}{2\sqrt{2w_{i1}(\zeta)}} L_{G_i(\zeta)} \begin{pmatrix} 1 & \bar{w}_{i2}(\zeta)^T \\ \bar{w}_{i2}(\zeta) & 4I - \bar{w}_{i2}(\zeta) \bar{w}_{i2}(\zeta)^T \end{pmatrix} + \frac{1}{2} v_i (1, -\bar{w}_{i2}(\zeta)^T) \right] \mid \right. \\ \left. u_i = (u_{i1}, u_{i2}), v_i = (v_{i1}, v_{i2}) \text{ satisfy } |u_{i1}| \leq \|u_{i2}\| \leq 1, |v_{i1}| \leq \|v_{i2}\| \leq 1 \right\},$$

where $w_i(\zeta) = (w_{i1}(\zeta), w_{i2}(\zeta)) = F_i(\zeta)^2 + G_i(\zeta)^2$ and $\bar{w}_{i2}(\zeta) = w_{i2}(\zeta) / \|w_{i2}(\zeta)\|$.

(c) If $(F_i(\zeta), G_i(\zeta)) = (0, 0)$, then $[A_i(\zeta), B_i(\zeta)] \in \{[L_{\hat{u}_i}, L_{\hat{v}_i}] \mid \|\hat{u}_i\|^2 + \|\hat{v}_i\|^2 = 1\}$ or

$$\left\{ \left[\begin{array}{c} \frac{1}{2}\xi_i(1, \bar{w}_{i2}^T) - \frac{1}{2}u_i(-1, \bar{w}_{i2}^T) + 2L_{s_i} \begin{pmatrix} 0 & 0 \\ 0 & (I - \bar{w}_{i2}\bar{w}_{i2}^T) \end{pmatrix}, \\ \frac{1}{2}\eta_i(1, \bar{w}_{i2}^T) - \frac{1}{2}v_i(-1, \bar{w}_{i2}^T) + 2L_{\omega_i} \begin{pmatrix} 0 & 0 \\ 0 & (I - \bar{w}_{i2}\bar{w}_{i2}^T) \end{pmatrix} \end{array} \right] \mid \right. \\ \left. \bar{w}_{i2} \in \mathbb{R}^{n_i-1} \text{ satisfies } \|\bar{w}_{i2}\| = 1 \text{ and } \xi_i = (\xi_{i1}, \xi_{i2}), u_i = (u_{i1}, u_{i2}), \eta_i = (\eta_{i1}, \eta_{i2}) \right. \\ \left. v_i = (v_{i1}, v_{i2}), s_i = (s_{i1}, s_{i2}), \omega_i = (\omega_{i1}, \omega_{i2}) \text{ satisfy } |\xi_1| \leq \|\xi_2\| \leq 1, \right. \\ \left. |u_{i1}| \leq \|u_{i2}\| \leq 1, |\eta_{i1}| \leq \|\eta_{i2}\| \leq 1, |v_{i1}| \leq \|v_{i2}\| \leq 1, \|s_i\|^2 + \|\omega_i\|^2 \leq 1/2 \right\}.$$

Proof. Let Φ_i denote the i -th component function of Φ , i.e., $\Phi_i(\zeta) = \phi_{\text{FB}}(F_i(\zeta), G_i(\zeta))$ and $\Phi_{q+i}(\zeta) = \phi_0(F_i(\zeta), G_i(\zeta))$ for $i = 1, \dots, q$. By the definition of the B-subdifferential,

$$\partial_B \Phi(\zeta)^T \subseteq \partial_B \Phi_1(\zeta)^T \times \partial_B \Phi_2(\zeta)^T \times \dots \times \partial_B \Phi_{2q}(\zeta)^T, \quad (20)$$

where the latter means the set of all matrices whose $(n_{i-1} + 1)$ -th to n_i -th columns belong to $\partial_B \Phi_i(\zeta)^T$ with $n_0 = 0$, and $(n + i)$ -th column belongs to $\partial_B \Phi_{q+i}(\zeta)^T$. Notice that

$$\begin{aligned} \partial_B \Phi_i(\zeta)^T &\subseteq \rho_1 [\nabla F_i(\zeta) \quad \nabla G_i(\zeta)] \partial_B \phi_{\text{FB}}(F_i(\zeta), G_i(\zeta))^T, \\ \partial_B \Phi_{q+i}(\zeta)^T &\subseteq \rho_2 [\nabla F_i(\zeta) \quad \nabla G_i(\zeta)] \partial_B \phi_0(F_i(\zeta), G_i(\zeta))^T. \end{aligned} \quad (21)$$

Moreover, using Lemma 3.1 (d) and Lemma 3.2 (c), each element in $\partial_B \phi_{\text{FB}}(F_i(\zeta), G_i(\zeta))^T$ and $\partial_B \phi_0(F_i(\zeta), G_i(\zeta))^T$ has the form of $\begin{pmatrix} A_i(\zeta) - I \\ B_i(\zeta) - I \end{pmatrix}$ and $\begin{pmatrix} C_i(\zeta) \\ D_i(\zeta) \end{pmatrix}$, respectively, with $A_i(\zeta), B_i(\zeta)$ and $C_i(\zeta), D_i(\zeta)$ for $i = 1, 2, \dots, q$ characterized as in the proposition. Therefore, combining with equations (20)–(21) yields the desired result. \square

To prove the fast local convergence of nonsmooth Levenberg-Marquardt methods, we need to know that under what assumptions every element $H \in \partial_B \Phi(\zeta^*)$ has full rank n , where ζ^* is an optimal solution of the SOCCP (1). To the end, define the index sets

$$\begin{aligned} \mathcal{I} &:= \{i \in \{1, 2, \dots, q\} \mid F_i(\zeta^*) = 0, G_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})\}, \\ \mathcal{B} &:= \{i \in \{1, 2, \dots, q\} \mid F_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i}), G_i(\zeta^*) \in \text{bd}^+(\mathcal{K}^{n_i})\}, \\ \mathcal{J} &:= \{i \in \{1, 2, \dots, q\} \mid F_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i}), G_i(\zeta^*) = 0\}. \end{aligned} \quad (22)$$

If ζ^* satisfies *strict complementarity*, i.e., $F_i(\zeta^*) + G_i(\zeta^*) \in \text{int}(\mathcal{K}^{n_i})$ for all i , then $\{1, 2, \dots, q\}$ can be partitioned as $\mathcal{I} \cup \mathcal{B} \cup \mathcal{J}$. 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 have the following results for the full rank of every element $H \in \partial_B \Phi(\zeta^*)$.

Theorem 3.1 *Let ζ^* be a strictly complementary solution of (1). Suppose that $\nabla G(\zeta^*)$ is invertible and let $P(\zeta^*) = \nabla G(\zeta^*)^{-1} \nabla F(\zeta^*)$. If $P(\zeta^*)_{\mathcal{II}}$ is nonsingular and its Schur-complement $\widehat{P}(\zeta^*)_{\mathcal{II}} := P(\zeta^*)_{\mathcal{BB}} - P(\zeta^*)_{\mathcal{BI}} P(\zeta^*)_{\mathcal{II}}^{-1} P(\zeta^*)_{\mathcal{IB}}$, in the matrix*

$$\begin{pmatrix} P(\zeta^*)_{\mathcal{II}} & P(\zeta^*)_{\mathcal{IB}} \\ P(\zeta^*)_{\mathcal{BI}} & P(\zeta^*)_{\mathcal{BB}} \end{pmatrix}$$

has the Cartesian P -property, then every element $H \in \partial_B \Phi(\zeta^)$ has full column rank n .*

Proof. Let $H \in \partial_B \Phi(\zeta^*)$. By Proposition 3.2, $H = \begin{pmatrix} \rho_1 H_1 \\ \rho_2 H_2 \end{pmatrix}$ with H_1^T from the set $\partial_B \Phi_1(\zeta^*)^T \times \cdots \times \partial_B \Phi_q(\zeta^*)^T$. From Theorem 4.1 of [24], it follows that H_1^T is nonsingular under the given assumptions. This implies the desired result $\text{rank}(H) = n$. \square

The proof of Theorem 3.1 is based on the important property of the first block H_1 . However, we see that when the first block H_1 is singular, the second block H_2 may contribute something to guarantee that H has full column rank n .

To close this section, we present a technical lemma that will be used in Section 5.

Lemma 3.3 *Let ζ^* be a solution of (1) such that all elements in $\partial_B \Phi(\zeta^*)$ have full column rank. Then, there exist constants $\varepsilon > 0$ and $c > 0$ such that $\|(H^T H)^{-1}\| \leq c$ for all $\|\zeta - \zeta^*\| < \varepsilon$ and all $H \in \partial_B \Phi(\zeta)$. Furthermore, for any given $\bar{\nu} > 0$, $H^T H + \nu I$ are uniformly positive definite for all $\nu \in [0, \bar{\nu}]$ and $H \in \partial_B \Phi(\zeta)$ with $\|\zeta - \zeta^*\| < \varepsilon$.*

Proof. The proof is similar to [26, Lemma 2.6]. For completeness, we here include it. Suppose that the claim of the lemma is not true. Then there exists a sequence $\{\zeta^k\}$ converging to ζ^* and a corresponding sequence of matrices $\{H_k\}$ with $H_k \in \partial_B \Phi(\zeta^k)$ for all $k \in \mathbb{N}$ such that either $H_k^T H_k$ is singular or $\|(H_k^T H_k)^{-1}\| \rightarrow +\infty$ on a subsequence. Noting that $H_k^T H_k$ is symmetric positive semidefinite, for the nonsingular case we have

$$\|(H_k^T H_k)^{-1}\| = \frac{1}{\lambda_{\min}(H_k^T H_k)},$$

which implies that the condition $\|(H_k^T H_k)^{-1}\| \rightarrow +\infty$ is equivalent to $\lambda_{\min}(H_k^T H_k) \rightarrow 0$. Since $\zeta^k \rightarrow \zeta^*$ and the mapping $\zeta \mapsto \partial_B \Phi(\zeta)$ is upper semicontinuous, it follows that the sequence $\{H_k\}$ is bounded, and hence it has a convergent subsequence. Let H_* be a limit of such a sequence. Then $\lambda_{\min}(H_*^T H_*) = 0$ by the continuity of the minimum eigenvalue. This means that $H_*^T H_*$ is singular. However, from the fact that the mapping $\zeta \mapsto \partial_B \Phi(\zeta)$ is closed, we have $H_* \in \partial_B \Phi(\zeta^*)$, which by the given condition implies that $H_*^T H_*$ is nonsingular. Thus, we obtain a contradiction.

By the definition of matrix norm and the result of the first part, there exist constants $\varepsilon > 0$ and $c > 0$ such that $[\lambda_{\min}(H^T H + \nu I)]^{-1} = \|(H^T H + \nu I)^{-1}\| \leq c$ for all $\nu \in [0, \bar{\nu}]$ and $H \in \partial_B \Phi(\zeta)$ with ζ with $\|\zeta - \zeta^*\| < \varepsilon$. This implies that

$$u^T(H^T H + \nu I)u \geq \lambda_{\min}(H^T H + \nu I)\|u\|^2 \geq \frac{1}{c}\|u\|^2 \quad \forall u \in \mathbb{R}^n.$$

Therefore, the matrices $H^T H + \nu I$ are uniformly positive definite. \square

4 Properties of the merit function Ψ

This section is devoted to the favorable properties of Ψ defined by (14)–(15). For this purpose, we need the following technical lemma which summarizes the properties of ψ .

Lemma 4.1 *Let $\psi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ be defined as in (15). Then, for any $x, y \in \mathbb{R}^n$,*

- (a) $\psi(x, y) = 0 \iff \psi_{\text{FB}}(x, y) = 0 \iff x \in \mathcal{K}^n, y \in \mathcal{K}^n, \langle x, y \rangle = 0$;
- (b) $\psi(x, y)$ is continuously differentiable;
- (c) $\langle x, \nabla_x \psi(x, y) \rangle + \langle y, \nabla_y \psi(x, y) \rangle \geq 2\psi(x, y)$;
- (d) $\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y) \rangle \geq 0$, and the equality holds if and only if $\psi(x, y) = 0$;
- (e) $\psi(x, y) = 0 \iff \nabla \psi(x, y) = 0 \iff \nabla_x \psi(x, y) = 0 \iff \nabla_y \psi(x, y) = 0$.

Proof. Part (a) is direct by the definition of ψ , and part (b) is from Lemma 3.1 (b) and Lemma 3.2 (a). We next consider part (c). By the definition of ψ , for any $x, y \in \mathbb{R}^n$,

$$\begin{aligned} \nabla_x \psi(x, y) &= \rho_1^2 \nabla_x \psi_{\text{FB}}(x, y) + \rho_2^2 \phi_0(x, y)y, \\ \nabla_y \psi(x, y) &= \rho_1^2 \nabla_y \psi_{\text{FB}}(x, y) + \rho_2^2 \phi_0(x, y)x. \end{aligned} \tag{23}$$

From [4, Lemma 6(a)] and the definition of $\phi_0(x, y)$, it then follows that

$$\begin{aligned} &\langle x, \nabla_x \psi(x, y) \rangle + \langle y, \nabla_y \psi(x, y) \rangle \\ &= \rho_1^2 [\langle x, \nabla_x \psi_{\text{FB}}(x, y) \rangle + \langle y, \nabla_y \psi_{\text{FB}}(x, y) \rangle] + 2\rho_2^2 \phi_0(x, y)x^T y \\ &= \rho_1^2 \|\phi_{\text{FB}}(x, y)\|^2 + 2\rho_2^2 \phi_0(x, y)^2 \\ &= 2 \left(\rho_1^2 \psi_{\text{FB}}(x, y) + \frac{1}{2} \rho_2^2 \phi_0(x, y)^2 \right) + \rho_2^2 \phi_0(x, y)^2 \\ &\geq 2\psi(x, y). \end{aligned}$$

(d) Using the formulas in (23) and [4, Lemma 6 (a)], it follows that

$$\begin{aligned}
\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y) \rangle &= \rho_1^4 \langle \nabla_x \psi_{\text{FB}}(x, y), \nabla_y \psi_{\text{FB}}(x, y) \rangle + \rho_2^4 x^T y \phi_0(x, y)^2 \\
&\quad + \rho_1^2 \rho_2^2 \phi_0(x, y) [\langle x, \nabla_x \psi_{\text{FB}}(x, y) \rangle + \langle y, \nabla_y \psi_{\text{FB}}(x, y) \rangle] \\
&= \rho_1^4 \langle \nabla_x \psi_{\text{FB}}(x, y), \nabla_y \psi_{\text{FB}}(x, y) \rangle + \rho_2^4 \phi_0(x, y)^3 \\
&\quad + 2\rho_1^2 \rho_2^2 \phi_0(x, y) \psi_{\text{FB}}(x, y).
\end{aligned} \tag{24}$$

Note that for the second equality, we use the fact

$$x^T y \phi_0(x, y)^2 = x^T y (\max\{0, x^T y\})^2 = \begin{cases} (x^T y)^3 & \text{if } x^T y \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

which says $x^T y \phi_0(x, y) = \phi_0(x, y)^3$. The first term on the right hand side of (24) is nonnegative by [4, Lemma 6 (b)], and the last two terms are also nonnegative. Therefore, $\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y) \rangle \geq 0$, and moreover, $\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y) \rangle = 0$ if and only if

$$\langle \nabla_x \psi_{\text{FB}}(x, y), \nabla_y \psi_{\text{FB}}(x, y) \rangle = 0 \quad \text{and} \quad \phi_0(x, y) = 0,$$

which, together with [4, Lemma 6(b)], implies the desired result.

(e) If $\psi(x, y) = 0$, then from the definition of ψ , we have $\phi_{\text{FB}}(x, y) = 0$ and $\phi_0(x, y) = 0$. From Proposition 1 of [4], we immediately obtain $\nabla_x \psi_{\text{FB}}(x, y) = \nabla_y \psi_{\text{FB}}(x, y) = 0$, and consequently $\nabla_x \psi(x, y) = 0$ and $\nabla_y \psi(x, y) = 0$ by (23). If $\nabla \psi(x, y) = 0$, then by part (c) and the nonnegativity of ψ we get $\psi(x, y) = 0$. Thus we prove the first equivalence. For the second equivalence, it suffices to prove the sufficiency. Suppose that $\nabla_x \psi(x, y) = 0$. From part (d), we readily obtain $\psi(x, y) = 0$, which together with part (a) and (23) implies $\nabla \psi(x, y) = 0$. Consequently, $\nabla \psi(x, y) = 0 \iff \nabla_x \psi(x, y) = 0$. Similarly, $\nabla \psi(x, y) = 0 \iff \nabla_y \psi(x, y) = 0$. This implies the last equivalence. \square

From Lemma 4.1 (b), the function Ψ is continuously differentiable. Also, by Lemma 4.1 (d), we can prove every stationary point of Ψ is a solution of (1) under mild conditions.

Proposition 4.1 *Let $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}_+$ be defined by (14)–(15). Then every stationary point of Ψ is a solution of the SOCCP (1) under one of the following assumptions:*

(a) $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ are column monotone³ for any $\zeta \in \mathbb{R}^n$.

(b) For any $\zeta \in \mathbb{R}^n$, $\nabla G(\zeta)$ is invertible and $\nabla G(\zeta)^{-1} \nabla F(\zeta)$ has Cartesian P_0 -property.

Proof. When the assumption (a) is satisfied, using the same arguments as those of [4, Proposition 3] yields the desired result. Now suppose that the assumption (b) holds. Let $\bar{\zeta}$ be an arbitrary stationary point of Ψ and write

$$\begin{aligned}
\nabla_x \psi(F(\bar{\zeta}), G(\bar{\zeta})) &= (\nabla_{x_1} \psi(F_1(\bar{\zeta}), G_1(\bar{\zeta})), \dots, \nabla_{x_q} \psi(F_q(\bar{\zeta}), G_q(\bar{\zeta}))), \\
\nabla_y \psi(F(\bar{\zeta}), G(\bar{\zeta})) &= (\nabla_{y_1} \psi(F_1(\bar{\zeta}), G_1(\bar{\zeta})), \dots, \nabla_{y_q} \psi(F_q(\bar{\zeta}), G_q(\bar{\zeta}))).
\end{aligned}$$

³ $M_1, M_2 \in \mathbb{R}^{n \times n}$ are column monotone if, for any $u, v \in \mathbb{R}^n$, $M_1 u + M_2 v = 0 \Rightarrow u^T v = 0$.

Then,

$$\nabla\Psi(\bar{\zeta}) = \nabla F(\bar{\zeta})\nabla_x\psi(F(\bar{\zeta}), G(\bar{\zeta})) + \nabla G(\bar{\zeta})\nabla_y\psi(F(\bar{\zeta}), G(\bar{\zeta})) = 0,$$

which, by the invertibility of ∇G , can be rewritten as

$$\nabla G(\bar{\zeta})^{-1}\nabla F(\bar{\zeta})\nabla_x\psi(F(\bar{\zeta}), G(\bar{\zeta})) + \nabla_y\psi(F(\bar{\zeta}), G(\bar{\zeta})) = 0. \quad (25)$$

Suppose that $\bar{\zeta}$ is not the solution of (1). By Lemma 4.1 (e), we necessarily have

$$\nabla_x\psi(F(\bar{\zeta}), G(\bar{\zeta})) \neq 0.$$

From the Cartesian P_0 -property of $\nabla G(\bar{\zeta})^{-1}\nabla F(\bar{\zeta})$, there exists an index $\nu \in \{1, 2, \dots, q\}$ such that $\nabla_{x_\nu}\psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta})) \neq 0$ and

$$\langle \nabla_{x_\nu}\psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta})), [\nabla G(\bar{\zeta})^{-1}\nabla F(\bar{\zeta})\nabla_x\psi(F(\bar{\zeta}), G(\bar{\zeta}))]_\nu \rangle \geq 0. \quad (26)$$

In addition, notice that (25) is equivalent to

$$[\nabla G(\bar{\zeta})^{-1}\nabla F(\bar{\zeta})\nabla_x\psi(F(\bar{\zeta}), G(\bar{\zeta}))]_i + \nabla_{y_i}\psi(F_i(\bar{\zeta}), G_i(\bar{\zeta})) = 0, \quad i = 1, 2, \dots, q.$$

Making the inner product with $\nabla_{x_\nu}\psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta}))$ for the ν th equality, we obtain

$$\begin{aligned} & \langle \nabla_{x_\nu}\psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta})), [\nabla G(\bar{\zeta})^{-1}\nabla F(\bar{\zeta})\nabla_x\psi(F(\bar{\zeta}), G(\bar{\zeta}))]_\nu \rangle \\ & + \langle \nabla_{x_\nu}\psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta})), \nabla_{y_\nu}\psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta})) \rangle = 0. \end{aligned}$$

The first term on the left hand side is nonnegative by (26), whereas the second term is positive by Lemma 4.1 (d) since ζ is not a solution of (1). This leads to a contradiction, and consequently $\bar{\zeta}$ must be a solution of (1). \square

When $\nabla G(\zeta)$ is invertible for any $\zeta \in \mathbb{R}^n$, the assumption in (a) is equivalent to the positive semidefiniteness of $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ at any $\zeta \in \mathbb{R}^n$, which implies the Cartesian P_0 -property of $\nabla G(\zeta)^{-1}\nabla F(\zeta)$. Thus, for the SOCCP (3), the assumption (a) is stronger than the assumption (b) which is now equivalent to the Cartesian P_0 -property of F .

Next we provide a condition to guarantee the boundedness of the level sets of Ψ

$$\mathcal{L}_\Psi(\gamma) := \{\zeta \in \mathbb{R}^n \mid \Psi(\zeta) \leq \gamma\}$$

for all $\gamma \geq 0$. This property is important since it guarantees that the descent sequence of Ψ must have a limit point, and furthermore, the solution set of (1) is bounded if it is nonempty. It turns out that the following condition for F and G is sufficient.

Condition 4.1 *For any sequence $\{\zeta^k\}$ satisfying $\|\zeta^k\| \rightarrow +\infty$, whenever*

$$\limsup \|[-F(\zeta^k)]_+\| < +\infty \quad \text{and} \quad \limsup \|[-G(\zeta^k)]_+\| < +\infty, \quad (27)$$

there exists an index $\nu \in \{1, 2, \dots, q\}$ such that $\limsup \langle F_\nu(\zeta^k), G_\nu(\zeta^k) \rangle = +\infty$.

Proposition 4.2 *If the mappings F and G satisfy Condition 4.1, then the level sets $\mathcal{L}_\Psi(\gamma)$ are bounded for all $\gamma \geq 0$.*

Proof. Assume that there is a unbounded sequence $\{\zeta^k\} \subseteq \mathcal{L}_\Psi(\gamma)$ for some $\gamma \geq 0$. Since $\Psi(\zeta^k) \leq \gamma$ for all k , the sequence $\{\Psi_{\text{FB}}(\zeta^k)\}$ is bounded. By Lemma 8 of [4],

$$\limsup \|[-F_i(x^k)]_+\| < +\infty \quad \text{and} \quad \limsup \|[-G_i(x^k)]_+\| < +\infty$$

hold for all $i \in \{1, 2, \dots, q\}$. This shows that F and G satisfy Condition 4.1, and hence there exists an index ν such that $\limsup \langle F_\nu(\zeta^k), G_\nu(\zeta^k) \rangle = +\infty$. From the definition of Ψ , it follows that the sequence $\{\Psi(\zeta^k)\}$ is unbounded, which clearly contradicts the fact that $\{\zeta^k\} \subseteq \mathcal{L}_\Psi(\gamma)$. The proof is completed. \square

Condition 4.1 is rather weak to guarantee that Ψ has bounded level sets since, as will be shown below, the condition is implied by the jointly monotone functions with a strictly feasible point used in [4] for f_{YF} , the jointly uniform Cartesian P -functions with a feasible point, or the joint \tilde{R}_{01} -functions in the following sense.

Definition 4.1 *The mappings $F, G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are said to have the joint \tilde{R}_{01} -property if for any sequence $\{\zeta^k\}$ with*

$$\|\zeta^k\| \rightarrow +\infty, \quad \frac{[-G(\zeta^k)]_+}{\|\zeta^k\|} \rightarrow 0, \quad \frac{[-F(\zeta^k)]_+}{\|\zeta^k\|} \rightarrow 0, \quad (28)$$

there holds that

$$\liminf_{k \rightarrow +\infty} \frac{\langle F(\zeta^k), G(\zeta^k) \rangle}{\|\zeta^k\|} > 0. \quad (29)$$

Proposition 4.3 *Condition 4.1 is satisfied if one of the following assumptions holds:*

- (a) *F and G are jointly monotone mappings satisfying $\lim_{\|\zeta\| \rightarrow +\infty} \|F(\zeta)\| + \|G(\zeta)\| = +\infty$, and there exists $\hat{\zeta} \in \mathbb{R}^n$ such that $F(\hat{\zeta}), G(\hat{\zeta}) \in \text{int}(\mathcal{K})$.*
- (b) *F and G have jointly uniform Cartesian P -property, and there exists a point $\hat{\zeta} \in \mathbb{R}^n$ such that $F(\hat{\zeta}), G(\hat{\zeta}) \in \mathcal{K}$.*
- (c) *F and G have the joint \tilde{R}_{01} -property.*

Proof. In the proof, let $\{\zeta^k\}$ be a sequence such that $\|\zeta^k\| \rightarrow +\infty$ and (27) holds.

(a) First, $\{\lambda_1[F(\zeta^k)]\}$ and $\{\lambda_1[G(\zeta^k)]\}$ must be bounded from below. If not, using

$$\|[-x]_+\|^2 = (\max\{0, -\lambda_1(x)\})^2 + (\max\{0, -\lambda_2(x)\})^2,$$

we obtain $\limsup \|[-F(\zeta^k)]_+\| = +\infty$ or $\limsup \|[-G(\zeta^k)]_+\| = +\infty$, which contradicts the assumption that $\{\zeta^k\}$ satisfies (27). Noting that $\|F(\zeta^k)\| + \|G(\zeta^k)\| \rightarrow +\infty$ and

$$\|F(\zeta^k)\| + \|G(\zeta^k)\| = \sqrt{\frac{\lambda_1^2[F(\zeta^k)] + \lambda_2^2[F(\zeta^k)]}{2}} + \sqrt{\frac{\lambda_1^2[G(\zeta^k)] + \lambda_2^2[G(\zeta^k)]}{2}},$$

the lower boundness of $\{\lambda_i[F(\zeta^k)]\}$ and $\{\lambda_i[G(\zeta^k)]\}$ for $i = 1, 2$ implies that

$$\limsup \lambda_2[F(\zeta^k)] = +\infty \quad \text{or} \quad \limsup \lambda_2[G(\zeta^k)] = +\infty.$$

From the proof of [4, Lemma 9 (b)] it then follows that

$$\limsup \left\{ \langle F(\zeta^k), G(\hat{\zeta}) \rangle + \langle F(\hat{\zeta}), G(\zeta^k) \rangle \right\} = +\infty. \quad (30)$$

Now suppose that Condition 4.1 is not satisfied. Then, we necessarily have

$$\limsup \langle F_i(\zeta^k), G_i(\zeta^k) \rangle < +\infty \quad \text{for all } i = 1, 2, \dots, q.$$

In addition, from the joint monotonicity of F and G , we have

$$\begin{aligned} \langle F(\zeta^k), G(\hat{\zeta}) \rangle + \langle F(\hat{\zeta}), G(\zeta^k) \rangle &\leq \langle F(\zeta^k), G(\zeta^k) \rangle + \langle F(\hat{\zeta}), G(\hat{\zeta}) \rangle \\ &= \sum_{i=1}^q \langle F_i(\zeta^k), G_i(\zeta^k) \rangle + \langle F(\hat{\zeta}), G(\hat{\zeta}) \rangle. \end{aligned}$$

The last two equations imply $\limsup \{ \langle F(\zeta^k), G(\hat{\zeta}) \rangle + \langle F(\hat{\zeta}), G(\zeta^k) \rangle \} < +\infty$. This clearly contradicts (30), and consequently the desired result follows.

(b) By Definition 2.2 (a), there exists a constant $\rho > 0$ such that

$$\begin{aligned} \rho \|\zeta^k - \hat{\zeta}\|^2 &\leq \max_{i \in \{1, \dots, q\}} \left\{ \langle F_i(\zeta^k) - F_i(\hat{\zeta}), G_i(\zeta^k) - G_i(\hat{\zeta}) \rangle \right\} \\ &= \langle F_\nu(\zeta^k), G_\nu(\zeta^k) \rangle + \langle F_\nu(\hat{\zeta}), -G_\nu(\zeta^k) \rangle \\ &\quad + \langle -F_\nu(\zeta^k), G_\nu(\hat{\zeta}) \rangle + \langle F_\nu(\hat{\zeta}), G_\nu(\hat{\zeta}) \rangle \\ &\leq \langle F_\nu(\zeta^k), G_\nu(\zeta^k) \rangle + \langle F_\nu(\hat{\zeta}), [-G_\nu(\zeta^k)]_+ \rangle \\ &\quad + \langle [-F_\nu(\zeta^k)]_+, G_\nu(\hat{\zeta}) \rangle + \langle F(\hat{\zeta}), G_\nu(\hat{\zeta}) \rangle, \end{aligned}$$

where ν is one of the indices for which the max is attained which we have, without loss of generality, assumed to be independent of k , and the second inequality is since

$$F_\nu(\hat{\zeta}) \in \mathcal{K}^{n_\nu}, \quad G_\nu(\hat{\zeta}) \in \mathcal{K}^{n_\nu}, \quad [-F_\nu(\zeta^k)]_- \in -\mathcal{K}^{n_\nu}, \quad [-G_\nu(\zeta^k)]_- \in -\mathcal{K}^{n_\nu}.$$

Dividing the last inequality by $\|\zeta^k\|^2$ and taking the limit, it follows from (27) that

$$\lim_{k \rightarrow +\infty} \frac{\langle F_\nu(\zeta^k), G_\nu(\zeta^k) \rangle}{\|\zeta^k\|^2} \geq \rho > 0,$$

which immediately implies the result.

(c) Clearly, $\{\zeta^k\}$ satisfies (28), and the result then follows from the following implications:

$$\begin{aligned} \liminf_{k \rightarrow +\infty} \frac{\langle F(\zeta^k), G(\zeta^k) \rangle}{\|\zeta^k\|} > 0 &\implies \liminf_{k \rightarrow +\infty} \frac{\max_i \{\langle F_i(\zeta^k), G_i(\zeta^k) \rangle\}}{\|\zeta^k\|} > 0 \\ &\implies \max_i \{\langle F_i(\zeta^k), G_i(\zeta^k) \rangle\} \rightarrow +\infty. \end{aligned}$$

So far, we complete the proof of this proposition. \square

When $G(\zeta) \equiv \zeta$, if we replace (29) with $\liminf_{k \rightarrow +\infty} \langle F(\zeta^k), G(\zeta^k) \rangle / \|\zeta^k\|^2 > 0$, then Definition 4.1 is saying that F is a R_{01} function. Thus, Proposition 4.2 and Proposition 4.3 (a) show that Ψ has bounded level sets under a weaker condition than the one given by [3, Prop. 4.1 (a)] for the class of merit functions f_{YF} .

To close this section, we show that the function Ψ provides a global error bound for the solution of SOCCP (1) under the jointly uniform Cartesian P -property of F and G . Since the jointly strong monotonicity implies the jointly uniform Cartesian P -property, the global error bound condition is weaker than that of [4, Proposition 5].

Proposition 4.4 *Let ζ^* be a solution of the SOCCP. Suppose that F and G have the jointly uniform Cartesian P -property. Then, there exists a scalar $\kappa > 0$ such that*

$$\|\zeta - \zeta^*\|^2 \leq \kappa \Psi(\zeta)^{1/2} \quad \forall \zeta \in \mathbb{R}^n.$$

Proof. Since F and G have the jointly uniform Cartesian P -property, there exists a scalar $\rho > 0$ such that, for any $\zeta \in \mathbb{R}^n$, there is an index $\nu \in \{1, 2, \dots, q\}$ such that

$$\begin{aligned} \rho \|\zeta - \zeta^*\|^2 &\leq \langle F_\nu(\zeta) - F_\nu(\zeta^*), G_\nu(\zeta) - G_\nu(\zeta^*) \rangle \\ &= \langle F_\nu(\zeta), G_\nu(\zeta) \rangle + \langle -F_\nu(\zeta), G_\nu(\zeta^*) \rangle + \langle F_\nu(\zeta^*), -G_\nu(\zeta) \rangle \\ &\leq \langle F_\nu(\zeta), G_\nu(\zeta) \rangle + \langle [-F_\nu(\zeta)]_+, G_\nu(\zeta^*) \rangle + \langle F_\nu(\zeta^*), [-G_\nu(\zeta)]_+ \rangle \\ &\leq \phi_0(F_\nu(\zeta), G_\nu(\zeta)) + \|[-F_\nu(\zeta)]_+\| \|G_\nu(\zeta^*)\| + \|F_\nu(\zeta^*)\| \|[-G_\nu(\zeta)]_+\| \\ &\leq c \left(\phi_0(F_\nu(\zeta), G_\nu(\zeta)) + \|[-F_\nu(\zeta)]_+\| + \|[-G_\nu(\zeta)]_+\| \right) \\ &\leq c \left(\phi_0(F_\nu(\zeta), G_\nu(\zeta)) + 4\psi_{\text{FB}}(F_\nu(\zeta), G_\nu(\zeta))^{1/2} \right) \\ &\leq c \left(\sqrt{2}/\rho_2 + 4/\rho_1 \right) \Psi(\zeta)^{1/2}, \end{aligned}$$

where $c := \max\{1, \|G_\nu(\zeta^*)\|, \|F_\nu(\zeta^*)\|\}$, the second inequality is using the fact that $G_\nu(\zeta^*) \in \mathcal{K}^{\nu}$ and $F_\nu(\zeta^*) \in \mathcal{K}^{\nu}$, and the next to last inequality is due to [4, Lemma 8]. Letting $\kappa := (c/\rho)(\sqrt{2}/\rho_2 + 4/\rho_1)$, we obtain the desired result. \square

5 Algorithm and convergence

It is known that the Levenberg-Marquardt method using equation (12) has the advantage that it reduces the complementarity gap $\langle x, F(x) \rangle$ for NCP faster than the traditional nonsmooth method using equation (8) does (see [19]). This motivates our employing a Levenberg-Marquardt type method with line search for solving the nonlinear least-square problem (14). We state its iterative scheme as below.

Algorithm 5.1 (Semismooth Levenberg-Marquardt Method)

(S.0) Choose a starting point $\zeta^0 \in \mathbb{R}^n$, the parameters $\rho_1, \rho_2 \in (0, 1)$, $\eta, \beta \in (0, 1)$, and $\sigma \in (0, 1/2)$. Given a tolerance $\varepsilon \geq 0$, and set $k := 0$.

(S.1) If $\|\nabla\Psi(\zeta^k)\| \leq \varepsilon$, then stop.

(S.2) Choose $H_k \in \partial_B\Phi(\zeta^k)$ and $\nu_k > 0$. Find a solution $d^k \in \mathbb{R}^n$ of linear system

$$(H_k^T H_k + \nu_k I)d = -\nabla\Psi(\zeta^k), \quad (31)$$

where $\nu_k > 0$ is the Levenberg-Marquardt parameter.

(S.3) If d^k satisfies

$$\|\Phi(\zeta^k + d^k)\| \leq \eta\|\Phi(\zeta^k)\|, \quad (32)$$

then $\zeta^{k+1} := \zeta^k + d^k$. Otherwise, compute $t_k = \max\{\beta^l \mid l = 0, 1, 2, \dots\}$ such that

$$\Psi(\zeta^k + t_k d^k) \leq \Psi(\zeta^k) + \sigma t_k \nabla\Psi(\zeta^k)^T d^k, \quad (33)$$

and let $\zeta^{k+1} := \zeta^k + t_k d^k$.

(S.4) Set $k := k + 1$, and go to (S.1).

Notice that the above method is different from the classical Levenberg-Marquardt method for nonlinear least-square problems in that Φ is not continuously differentiable. If $\nu_k \equiv 0$, the solution of (31) is exactly the solution of the linear least-square problem

$$\min_{d \in \mathbb{R}^n} \frac{1}{2} \|H_k d + \Phi(\zeta^k)\|^2, \quad (34)$$

since $\nabla\Psi(\zeta^k) = H_k^T \Phi(\zeta^k)$. In this paper, we choose the parameter ν_k by

$$\nu_k := \min \{p_1, p_2 \|\Phi(\zeta^k)\|^\varrho\}, \quad (35)$$

where $p_1, p_2 > 0$ are given constants and ϱ is a real number from $[1, 2]$. Such choice is consistent with the requirements for local superlinear (quadratic) convergence stated in Theorem 5.2 and Theorem 5.3 below, as well as adopted by our numerical experiments.

In what follows, we will study the convergence properties of the algorithm. For this purpose, assume that ε equals to 0. We first state a global convergence result.

Theorem 5.1 *Let $\{\zeta^k\}$ be the sequence generated by Algorithm 5.1 with ν_k updated by (35). Then every accumulation point of $\{\zeta^k\}$ is a stationary point of Ψ .*

Proof. From the steps of Algorithm 5.1, $\{\zeta^k\}$ is well defined since $\nu_k > 0$, and d^k determined by (31) is always a descent direction of Ψ at ζ^k . Let ζ^* be any accumulation point of $\{\zeta^k\}$ and $\{\zeta^k\}_K$ be a subsequence converging to ζ^* . Suppose that $\nabla\Psi(\zeta^*) \neq 0$. Since $\{\Psi(\zeta^k)\}$ is monotonically decreasing and bounded below, and $\{\Psi(\zeta^k)\}_K$ converges to $\Psi(\zeta^*)$, the entire sequence $\{\Psi(\zeta^k)\}$ converges to $\Psi(\zeta^*) > 0$. This implies that (32) holds for only finitely many $k \in K$, and the inequality (33) is satisfied for all sufficiently large k . Since $\Psi(\zeta^{k+1}) - \Psi(\zeta^k) \leq \sigma t_k \nabla\Psi(\zeta^k)^T d^k \leq 0$ for all sufficiently large k , using $\Psi(\zeta^{k+1}) - \Psi(\zeta^k) \rightarrow 0$ yields that

$$\{t_k \nabla\Psi(\zeta^k)^T d^k\}_K \rightarrow 0. \quad (36)$$

We next prove $\{\nabla\Psi(\zeta^k)^T d^k\}_K$ has a nonzero limit as $k \rightarrow +\infty$. By the definition of d^k ,

$$\nabla\Psi(\zeta^k)^T d^k = -\nabla\Psi(\zeta^k)^T (H_k^T H_k + \nu_k I)^{-1} \nabla\Psi(\zeta^k) \quad \forall k. \quad (37)$$

Since the B-subdifferential $\partial_B\Phi(\zeta)$ is a nonempty compact set for any $\zeta \in \mathbb{R}^n$, $\{H_k\}_K$ is bounded. Without loss of generality, assume that $\{H_k\}_K \rightarrow H_*$. Considering that the set-valued mapping $\zeta \mapsto \partial_B\Phi(\zeta)$ is closed and $\{\zeta^k\}_K \rightarrow \zeta^*$, we have $H_* \in \partial_B\Phi(\zeta^*)$. In addition, since $\Phi(\zeta^*) \neq 0$, we have $\nu_k \rightarrow \nu_*$ with $\nu_* = \min\{p_1, p_2 \|\Phi(\zeta^*)\|^e\} > 0$. Thus, $\{H_k^T H_k + \nu_k I\}_{k \in K} \rightarrow H_*^T H_* + \nu_* I \succ O$. This, together with (37) and the continuity of $\nabla\Psi$, implies that $\{\nabla\Psi(\zeta^k)^T d^k\}_K$ has a nonzero limit as $k \rightarrow +\infty$. From (36), it then follows that $\{t_k\}_K \rightarrow 0$. Now, for all sufficiently large k , let $l_k \in \{0, 1, \dots\}$ be the unique exponent such that $t_k = \beta^{l_k}$. Since $\{t_k\}_K \rightarrow 0$, we have $\{l_k\}_{k \in K} \rightarrow \infty$. From the Armijo line search in (S.3), for all $k \in K$ sufficiently large,

$$\frac{\Psi(\zeta^k + \beta^{l_k-1} d^k) - \Psi(\zeta^k)}{\beta^{l_k-1}} > \sigma \nabla\Psi(\zeta^k)^T d^k. \quad (38)$$

Taking the limit $k \rightarrow \infty$ with $k \in K$ and using $\{l_k\}_K \rightarrow \infty$ and $\{\zeta^k\}_K \rightarrow \zeta^*$, we have $\nabla\Psi(\zeta^*)^T d^* \geq \sigma \nabla\Psi(\zeta^*)^T d^*$. This means $\nabla\Psi(\zeta^*)^T d^* \geq 0$. On the other hand, we learn from (31) that $\{d^k\}_K \rightarrow d^*$ with d^* being the solution of

$$(H_*^T H_* + \nu_* I) d = -\nabla\Psi(\zeta^*), \quad (39)$$

which implies that $\nabla\Psi(\zeta^*)^T d^* < 0$ since $(H_*^T H_* + \nu_* I) \succ O$. Thus, we get a contradiction. \square

Observe that the sequence $\{\zeta^k\}$ generated by Algorithm 5.1 always belongs to the level set $\mathcal{L}_\Psi(\Psi(\zeta^0))$. By Propositions 4.2 and 4.3, the existence of accumulation points of $\{\zeta^k\}$ is guaranteed by one of the assumptions of Proposition 4.3. Since, when F and G have the jointly uniform Cartesian P -property, the SOCCP (1) has at most one solution,

$\{\zeta^k\}$ must have a unique accumulation point which is the unique solution of (1) if F and G satisfies the assumption (c) of Proposition 4.3. For the SOCCP (3), the sequence $\{\zeta^k\}$ has accumulation points and each of them is a solution under the assumption that F is monotone and (3) is strictly feasible.

Next we establish the superlinear (or quadratic) rate of convergence of Algorithm 5.1 under the strict complementarity of the solution. This condition seems to be a little rigorous, and later we will replace it with a local error bound assumption.

Theorem 5.2 *Let $\{\zeta^k\}$ be generated by Algorithm 5.1 with ν_k given by (35). Suppose that ζ^* is an accumulation point of $\{\zeta^k\}$ with ζ^* being a strictly complementary solution of (1), and F and G at ζ^* satisfy the condition of Theorem 3.1. Then,*

- (a) *the entire sequence $\{\zeta^k\}$ converges to ζ^* .*
- (b) *The full stepsize $t_k = 1$ is always accepted for sufficiently large k and the rate of convergence is Q -superlinear.*
- (c) *The rate of convergence is Q -quadratic if, in addition, F' and G' are locally Lipschitz continuous around ζ^* and $\nu_k = O(\|\Phi(\zeta^k)\|)$.*

Proof. The proof is similar to the one given by [19]. For completeness, we include it.

(a) By the proof technique of Theorem 3.1 (b) of [22], it suffices to prove that ζ^* is an isolated solution. From Theorem 3.1 and Lemma 3.3, there exist $\varepsilon_1, \kappa_1 > 0$ such that

$$\|H(\zeta - \zeta^*)\|^2 = (\zeta - \zeta^*)H^T H(\zeta - \zeta^*) \geq \kappa_1 \|\zeta - \zeta^*\|^2$$

for all ζ satisfying $\|\zeta - \zeta^*\| < \varepsilon_1$ and all $H \in \partial_B \Phi(\zeta)$. In addition, the semismoothness of Φ implies that there exists $\varepsilon_2 > 0$ such that

$$\|\Phi(\zeta) - \Phi(\zeta^*) - H(\zeta - \zeta^*)\| \leq (\sqrt{\kappa_1}/2)\|\zeta - \zeta^*\|$$

for all $H \in \partial_B \Phi(\zeta)$ with ζ satisfying $\|\zeta - \zeta^*\| < \varepsilon_2$. Set $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Then, we have

$$\begin{aligned} \|\Phi(\zeta)\| &= \|H(\zeta - \zeta^*) + (\Phi(\zeta) - \Phi(\zeta^*) - H(\zeta - \zeta^*))\| \\ &\geq \|H(\zeta - \zeta^*)\| - \|\Phi(\zeta) - \Phi(\zeta^*) - H(\zeta - \zeta^*)\| \\ &\geq (\sqrt{\kappa_1}/2)\|\zeta - \zeta^*\| \end{aligned}$$

for all ζ with $\|\zeta - \zeta^*\| < \varepsilon$. This means that ζ^* is an isolated solution of the SOCCP.

(b) We first prove that for all sufficiently large k ,

$$\|\zeta^k + d^k - \zeta^*\| = o(\|\zeta^k - \zeta^*\|). \quad (40)$$

By part (a), the sequence $\{\zeta^k\}$ converges to a solution ζ^* satisfying the assumptions of Theorem 3.1. From Lemma 3.3, there exists $c > 0$ such that $\|(H_k^T H_k + \nu_k I)^{-1}\| \leq c$ for all

k . Noting that the sequence $\{H_k\}$ is bounded, there exists $c_1 > 0$ such that $\|H_k^T\| \leq c_1$ for all k . Using Theorem 5.1 and the fact that $\Phi(\zeta^*) = 0$, we obtain that

$$\begin{aligned} \|\zeta^k + d^k - \zeta^*\| &= \|\zeta^k - (H_k^T H_k + \nu_k I)^{-1} \nabla \Psi(\zeta^k) - \zeta^*\| \\ &\leq \|(H_k^T H_k + \nu_k I)^{-1}\| \|\nabla \Psi(\zeta^k) - (H_k^T H_k + \nu_k I)(\zeta^k - \zeta^*)\| \\ &\leq c \|H_k^T \Phi(\zeta^k) - H_k^T H_k(\zeta^k - \zeta^*) - \nu_k(\zeta^k - \zeta^*)\| \\ &= c \|H_k^T(\Phi(\zeta^k) - \Phi(\zeta^*) - H_k(\zeta^k - \zeta^*)) - \nu_k(\zeta^k - \zeta^*)\| \\ &\leq c(c_1 \|\Phi(\zeta^k) - \Phi(\zeta^*) - H_k(\zeta^k - \zeta^*)\| + \nu_k \|\zeta^k - \zeta^*\|). \end{aligned}$$

Notice that $\Phi(\zeta^k) - \Phi(\zeta^*) - H_k(\zeta^k - \zeta^*) = o(\|\zeta^k - \zeta^*\|)$ by the semismoothness of Φ , whereas $\nu_k \rightarrow 0$ by part (a) and the continuity of Φ . Thus, the inequality implies (40).

To prove that the full step is eventually accepted, by (32) it suffices to show that

$$\lim_{k \rightarrow \infty} \frac{\Psi(\zeta^k + d^k)}{\Psi(\zeta^k)} = 0. \quad (41)$$

Since all element $V \in \partial_B \Phi_{\text{FB}}(\zeta^*)$ are nonsingular by [24, Theorem 4.1], from Lemma 3.3 and the proof of part (a), there exists a constant $\alpha > 0$ such that

$$\|\Phi(\zeta^k)\| \geq \rho_1 \|\Phi_{\text{FB}}(\zeta^k)\| \geq \alpha \|\zeta^k - \zeta^*\|.$$

Using the locally Lipschitz continuity of Φ and (40) then yields that

$$\frac{\|\Phi(\zeta^k + d^k)\|}{\|\Phi(\zeta^k)\|} \leq \frac{\|\Phi(\zeta^k + d^k) - \Phi(\zeta^*)\|}{\alpha \|\zeta^k - \zeta^*\|} \leq \frac{L \|\zeta^k + d^k - \zeta^*\|}{\alpha \|\zeta^k - \zeta^*\|} \rightarrow 0,$$

where $L > 0$ denotes the locally Lipschitz constant of Φ . Thus, the stepsize $t_k = 1$ is eventually accepted in the line search criterion, i.e., $\zeta^{k+1} = \zeta^k + d^k$ for all k large enough. Consequently, Q-suplinear convergence of $\{\zeta^k\}$ to ζ^* follows from (40).

(c) The proof is essentially same as for the superlinear convergence. We only note that ν_k in (35) satisfies $\nu_k = O(\|\Phi(\zeta^k)\|) = O(\|\zeta^k - \zeta^*\|)$ for k large enough, and

$$\Phi(\zeta^k) - \Phi(\zeta^*) - H_k(\zeta^k - \zeta^*) = O(\|\zeta^k - \zeta^*\|^2)$$

due to the strong semismoothness of Φ by Proposition 3.1. \square

We next establish the superlinear (quadratic) rate of convergence of Algorithm 5.1 under a local error bound assumption, which is stated as follows:

Assumption A. There exist constants $\kappa_2 > 0$ and $0 < \delta < 1$ such that

$$\kappa_2 \text{dist}(\zeta, S^*) \leq \|\Phi(\zeta)\| \quad \forall \zeta \in \mathcal{N}(\zeta^*, \delta), \quad (42)$$

where S^* denotes the solution set of the SOCCP (1) and is assumed to be nonempty.

Lemma 5.1 *Let ζ^k be generated by Algorithm 5.1 with ν_k given by (35). Suppose that F' and G' are Lipschitz continuous on $\mathcal{N}(\zeta^*, \delta)$ and Assumption A holds. If $\nu_k = p_2 \|\Phi(\zeta^k)\|^e$ and $\zeta^k \in \mathcal{N}(\zeta^*, \delta/2)$, then there exists a constant $c_1 > 0$ such that $\|d^k\| \leq c_1 \text{dist}(\zeta^k, S^*)$. If, in addition, $\zeta^k + d^k \in \mathcal{N}(\zeta^*, \delta/2)$, then there exists a constant $c_3 > 0$ such that*

$$\text{dist}(\zeta^k + d^k, S^*) \leq c_3 \text{dist}(\zeta^k, S^*)^{(\varrho+2)/2}.$$

Proof. Let $\bar{\zeta}^k \in S^*$ be such that $\|\zeta^k - \bar{\zeta}^k\| = \text{dist}(\zeta^k, S^*)$. Then, $\bar{\zeta}^k \in \mathcal{N}(\zeta^*, \delta)$ since

$$\|\bar{\zeta}^k - \zeta^*\| \leq \|\bar{\zeta}^k - \zeta^k\| + \|\zeta^k - \zeta^*\| \leq 2\|\zeta^k - \zeta^*\| \leq \delta.$$

Noting that Φ is Lipschitz continuous on $\mathcal{N}(\zeta^*, \delta)$, there is a constant $L_1 > 0$ such that

$$\|\Phi(\zeta^k)\| = \|\Phi(\zeta^k) - \Phi(\bar{\zeta}^k)\| \leq L_1 \|\zeta^k - \bar{\zeta}^k\|.$$

Combining with the inequality (42), we have

$$p_2 \kappa_2^e \|\bar{\zeta}^k - \zeta^k\|^e \leq \nu_k = p_2 \|\Phi(\zeta^k)\|^e \leq p_2 L_1^e \|\zeta^k - \bar{\zeta}^k\|^e. \quad (43)$$

On the other hand, since Φ is strongly semismooth on $\mathcal{N}(\zeta^*, \delta)$ by Proposition 3.1, there exists a constant $\hat{c} > 0$ such that

$$\|\Phi(\zeta^k) + H_k(\bar{\zeta}^k - \zeta^k)\| = \|\Phi(\zeta^k) - \Phi(\bar{\zeta}^k) - H_k(\zeta^k - \bar{\zeta}^k)\| \leq \hat{c} \|\zeta^k - \bar{\zeta}^k\|^2. \quad (44)$$

Define

$$\varphi_k(d) := \|\Phi(\zeta^k) + H_k d\|^2 + \nu_k \|d\|^2. \quad (45)$$

Then, d^k is a minimizer of $\varphi_k(d)$. This together with (44) and (43) yields that

$$\begin{aligned} \|d^k\|^2 &\leq \frac{\varphi_k(d^k)}{\nu_k} \leq \frac{\varphi_k(\bar{\zeta}^k - \zeta^k)}{\nu_k} = \frac{\|\Phi(\zeta^k) + H_k(\bar{\zeta}^k - \zeta^k)\|^2 + \nu_k \|\bar{\zeta}^k - \zeta^k\|^2}{\nu_k} \\ &\leq \frac{\hat{c}^2 p_2^{-1} \kappa_2^{-e} \|\bar{\zeta}^k - \zeta^k\|^{4-e} + \|\bar{\zeta}^k - \zeta^k\|^2}{\nu_k} \\ &= (\hat{c}^2 p_2^{-1} \kappa_2^{-e} + 1) \|\bar{\zeta}^k - \zeta^k\|^2, \end{aligned}$$

which implies the first part with $c_1 = \sqrt{\hat{c}^2 p_2^{-1} \kappa_2^{-e} + 1}$. Noting that

$$\begin{aligned} \varphi_k(d^k) &\leq \varphi_k(\bar{\zeta}^k - \zeta^k) \leq \|\Phi(\zeta^k) + H_k(\bar{\zeta}^k - \zeta^k)\|^2 + \nu_k \|\bar{\zeta}^k - \zeta^k\|^2 \\ &\leq \hat{c}^2 \|\bar{\zeta}^k - \zeta^k\|^4 + p_2 L_1^e \|\zeta^k - \bar{\zeta}^k\|^{2+e} \\ &\leq (\hat{c}^2 + p_2 L_1^e) \|\zeta^k - \bar{\zeta}^k\|^{2+e}, \end{aligned}$$

we have

$$\begin{aligned} \|\Phi(\zeta^k + d^k)\| &= \|\Phi(\zeta^k + d^k) - \Phi(\zeta^k) - H_k d^k + \Phi(\zeta^k) + H_k d^k\| \\ &\leq \|\Phi(\zeta^k + d^k) - \Phi(\zeta^k) - H_k d^k\| + \sqrt{\varphi_k(d^k)} \\ &\leq \hat{c} \|d^k\|^2 + (\hat{c}^2 + p_2 L_1^e)^{1/2} \|\zeta^k - \bar{\zeta}^k\|^{(\varrho+2)/2} \\ &\leq \hat{c} (\hat{c}^2 p_2^{-1} \kappa_2^{-e} + 1) \|\bar{\zeta}^k - \zeta^k\|^2 + (\hat{c}^2 + p_2 L_1^e)^{1/2} \|\zeta^k - \bar{\zeta}^k\|^{(\varrho+2)/2} \\ &\leq c_2 \|\zeta^k - \bar{\zeta}^k\|^{(\varrho+2)/2} \end{aligned}$$

with $c_2 = \hat{c}(\hat{c}^2 p_2^{-1} \kappa_2^{-\varrho} + 1) + (\hat{c}^2 + p_2 L_1^\varrho)^{1/2}$. Consequently,

$$\text{dist}(\zeta^k + d^k, S^*) \leq \frac{1}{\kappa_2} \|\Phi(\zeta^k + d^k)\| \leq \frac{c_2}{\kappa_2} \|\zeta^k - \bar{\zeta}^k\|^{(\varrho+2)/2} = c_3 \text{dist}(\zeta^k, S^*)^{(\varrho+2)/2}.$$

Thus, we complete the proof of the second part. \square

By Lemma 5.1, using the similar arguments to [8, Theorem 2.1] and [31, Theorem 3.1], we get the quadratic rate of convergence of Algorithm 5.1 under Assumption A.

Theorem 5.3 *Let $\{\zeta^k\}$ be generated by Algorithm 5.1 with ν_k given by (35), and ζ^* be an accumulation point of $\{\zeta^k\}$. If ζ^* is a solution of (1), then the sequence $\{\zeta^k\}$ converges to ζ^* superlinearly, and moreover, quadratically when $\varrho = 2$, provided that F' and G' are locally Lipschitz continuous and Assumption A holds.*

Now we do not know whether Assumption A is weaker than the strict complementarity of the solution, although the assumptions of Theorem 5.3 are weaker than those of Theorem 5.2, since the latter implies that each element in $\partial_B \Phi(\zeta^*)$ is nonsingular, and so $\|\Phi(\zeta)\|$ provides a local error bound on some neighborhood of the solution ζ^* , but from [31] the former does not imply the nonsingularity of each element in $\partial_B \Phi(\zeta^*)$. From the proof of Lemma 5.1, we find that the condition (42) cannot be weakened to

$$\kappa_2 \text{dist}(\zeta, S^*) \leq \|\Phi(\zeta)\|^{1/2} \quad \forall \zeta \in \mathcal{N}(\zeta^*, \delta),$$

in order to guarantee the superlinear (or quadratic) convergence of Algorithm 5.1, and therefore the global error bound result of Proposition 4.4 may not be applied for it. If let $\Psi(\zeta) = \|\Phi(\zeta)\|^4/4$ instead of $\Psi(\zeta) = \|\Phi(\zeta)\|^2/2$, then Assumption A holds automatically under the jointly uniform Cartesian P -property of F and G , but this will bring difficulty to numerical implementation due to the bad scaling of Ψ . Thus, it is worthwhile to study what conditions of F and G are sufficient for Assumption A to hold.

6 Numerical results

In this section, we report numerical results with the least-square semismooth method (*LS semismooth method* for short) solving the SOCCP (1), derived from the KKT conditions of convex SOCPs. As one referee pointed out, for the solution of convex SOCPs, the reformulation seems to be circuitous since the KKT conditions can be directly written as a mixed SOCCP. However, since the purpose of this paper is to develop an efficient method for the general SOCCP (1), instead of convex SOCPs, we here adopt such reformulation so as to obtain the corresponding test instances for (1).

All experiments were done with a PC of Intel Pentium Dual CPU E2200 and 2047MB memory, and the computer codes were written in Matlab 7.0. Since the nonmonotone line search [14] is usually superior to the classical monotone line search, we replaced the Armijo line search of Algorithm 4.1 by the nonmonotone version in [14], i.e., we computed t_k such that

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

where

$$\mathcal{W}_k := \max_{j=k-m_k, \dots, k} \Psi(\zeta^j),$$

and where, for a given nonnegative integer \hat{m} and s , $m_k = 0$ if $k \leq s$, and otherwise $m_k = \min\{m_{k-1} + 1, \hat{m}\}$. In our tests, the parameters in Algorithm 5.1 were chosen as

$$\rho_1 = 0.9, \quad \rho_2 = 0.1, \quad \eta = 1.0e - 6, \quad \sigma = 1.0e - 4, \quad \beta = 0.5, \quad \hat{m} = 5 \text{ and } s = 5.$$

The parameter ν_k was chosen as in (35) with $p_1 = 1.0$, $p_2 = 10^{-5}/n$, and $\varrho = 1$. We started Algorithm 5.1 with the initial point $\zeta^0 = 0$ and terminated it whenever

$$\max\{|F(\zeta^k)^T G(\zeta^k)|, \Psi(\zeta^k)\} \leq 10^{-6}, \quad \text{or } k > 150, \quad \text{or } t_k < 10^{-15}. \quad (46)$$

We compared the numerical performance of Algorithm 5.1 with that of the least-square semismooth Newton method based on (8), called the *FB semismooth method*, which corresponds to the special case of $\rho_1 = 1$, $\rho_2 = 0$ of Algorithm 5.1. For the linear SOCPs, we compared the numerical results of the two semismooth methods with those of SeDuMi [29], a successful interior point method software for the linear SOCPs and the semidefinite programming. The parameters of the SeduMi were set as default values.

The first group of test instances is the linear SOCP from the DIMACS Implementation Challenge library [25]. During the tests, we computed $\hat{x} \in \mathbb{R}^n$ in F as a solution of $\min_x \|Ax - b\|$ by Matlab's least square solver "LSQLIN", and evaluated F and G in (5) via the Cholesky factorization of AA^T . The results were reported in Table 1, where **Optval** denotes the objective value of the SOCPs at the final iteration, **Iter** records the number of iteration, and **NF** means the number of function evaluations for each problem.

From Table 1, we see that the two least-square semismooth Newton methods are able to yield a solution with favorable accuracy for all test problems, and requires less iterations for "nb_L2_bessel" than the SeduMi. However, for "nb" and "nb_L1", they are incomparable with the SeduMi in terms of the number of iterations. We also checked that the solutions of the two problems do not satisfy the strict complementarity. In addition, for the two difficult test problems, the LS semismooth method requires less iterations and function evaluations than the FB semismooth method. Moreover, for "nb_L1", the advantage of the LS semismooth method is more remarkable.

Table 1: Numerical results for the DIAMCS linear SOCPs

	LS semismooth Method			FB semismooth method			SeDuMi	
Problem	Optval	Iter	NF	Optval	Iter	NF	Optval	Iter
nb	-5.070456e-2	38	87	-5.070467e-2	39	108	-5.070310e-2	21
nb-L1	-1.301223e+1	90	126	-1.301223e+1	106	187	-1.301227e+1	18
nb-L2-bessel	-1.025697e-1	10	16	-1.025697e-1	10	16	-1.025695e-1	16

The second group of test instances is the nonlinear convex SOCP (4) with sparse A . To generate such test problems, we consider the problem of minimizing a sum of the k largest Euclidean norms with a convex regularization term: $\min_{u \geq 0} \sum_{i=1}^k \|s_{[i]}\| + h(u)$, where $\|s_{[1]}\|, \dots, \|s_{[r]}\|$ are the norms $\|s_1\|, \dots, \|s_r\|$ sorted in nonincreasing order with $r \geq k$ and $s_i = b_i - A_i x$ for $i = 1, \dots, r$ with $A_i \in \mathbb{R}^{m_i \times l}$ and $b_i \in \mathbb{R}^{m_i}$, and $h: \mathbb{R}^l \rightarrow \mathbb{R}$ is a twice continuously differentiable convex function. The problem can be converted to

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

In our tests, we set $h(u) := \frac{1}{3} \|u\|_3^3$ with $\|\cdot\|_3$ denoting the 3-norm, and generated each m_i randomly from $\{2, 3, \dots, 10\}$. All A_i were chosen as sparse matrices with approximately $10\% \cdot m_i \cdot d$ uniformly distributed nonzero entries, and all entries of b_i were chosen from the uniform distribution in $[-1, 0]$. For each (l, r, k) , we generated ten test instances, and then solved the SOCCP (1) derived from the KKT conditions of each problem with the LS semismooth method and the FB semismooth method. The mappings F and G in (5) were evaluated in the same way as above. The first inequality in (46) was replaced by $\max \{|F(\zeta^k)^T G(\zeta^k)|, \Psi(\zeta^k)\} \leq 10^{-8}$. The numerical results were listed in Table 2, in which the second column gives the average dimension (m, n) of A for ten problems, **Gap** denotes the average value of $|F(\zeta^k)^T G(\zeta^k)|$ at final iteration, **NF** means the average function evaluations for solving each instance, and **Iter** denotes the average number of iteration for each instance to satisfy the termination conditions, and **Time** records the average CPU time in second for solving each test problem.

From Table 2, we see that for the second group of test problems which is much easier than “nb” and “nb.L1”, the LS semismooth method does not have remarkable superi-

Table 2: Numerical results for the nonlinear convex SOCPs

	Dim.	LS semismooth Method				FB semismooth method			
(l,r,k)	(m,n)	Gap	NF	Iter	Time	Gap	NF	Iter	Time
(500,10,5)	(77,588)	3.20e-9	45.3	22.4	9.01	3.16e-9	52.3	22.5	6.29
(500,20,5)	(132,653)	2.68e-9	39.8	22.9	12.4	1.99e-9	42.7	21.4	8.00
(500,50,5)	(355,906)	2.15e-9	50.9	28.9	36.1	2.44e-9	41.8	25	22.3
(500,100,2)	(688,1289)	4.08e-9	33.9	26.4	83.1	4.71e-9	33.5	27.5	65.3
(1000,10,5)	(71,1082)	2.74e-9	58.9	24.8	53.4	2.79e-9	65.6	25.5	35.6
(1000,20,5)	(136,1157)	2.44e-9	55.9	22.4	62.4	2.30e-9	57	24.3	41.1
(1000,50,5)	(347,1398)	1.99e-9	49.6	27.9	117.9	2.92e-9	48.7	27.8	81.2
(2000,10,5)	(70,2081)	2.10e-9	92.3	32.5	445.6	1.82e-9	88	31.2	276.8

ority to the FB semismooth method. Among eight groups of test instances, the average number of iteration and the average number of function evaluations required by the LS semismooth method are basically same as that of the FB semismooth method, but the FB semismooth method requires less CPU time due to less computation work at each iteration. Combining with the results in Table 1, we conclude that the LS semismooth method is superior to the FB semismooth method only for those difficult problems.

7 Conclusion

We have presented a nonlinear least-square reformulation for the SOCCP (1) by use of the FB function and the plus function, which was shown to have some advantages over the nonsmooth system reformulation (8). Based on the reformulation, a semismooth Levenberg-Marquardt method was developed, and the superlinear (quadratic) rate of convergence was established under the strict complementarity of the solution and a local error bound assumption, respectively. Although the local error bound assumption makes no requirements for the solution, we do not know what conditions of F and G can guarantee it to hold. We will leave it as a future research topic.

It should be pointed out that other least-square formulations can be constructed by a similar way; for example, appending $(x)_+ \circ (y)_+$ or $(x \circ y)_+$ to the mapping Φ_{FB} . But, it seems that the formulation based on ϕ_0 is the best, since the merit function corresponding to $(x)_+ \circ (y)_+$ is not smooth, whereas the one corresponding to $(x \circ y)_+$ does not have all

the properties of Lemma 4.1. This is completely different from the NCP case. Since the strong semismoothness of the FB function over general symmetric cones is still an open problem, now the method of this paper can not be extended to general symmetric cone complementarity problems.

Acknowledgements. The authors would like to thank two anonymous referees for their valuable comments and suggestions for this paper.

References

- [1] F. ALIZADEH AND D. GOLDFARB (2003), *Second-order cone programming*, Mathematical Programming, vol. 95, pp. 3–51.
- [2] E. D. ANDERSEN, C. ROOS, AND T. TERLAKY (2003), *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming, vol. 95, pp. 249–277.
- [3] J.-S. CHEN (2007), *Conditions for error bounds and bounded level sets of some merit functions for the SOCCP*, Journal of Optimization Theory and Applications, vol. 135, pp. 459–473.
- [4] J.-S. CHEN AND P. TSENG (2005), *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Mathematical Programming, vol. 104, pp. 293–327.
- [5] X. CHEN AND H. QI (2006), *Cartesian P -proeprty and its applications to the semidefinite linear complementarity problem*, Mathematical Programming, vol. 106, pp. 177–201.
- [6] X.-D. CHEN, D. SUN, AND J. SUN (2003), *Complementarity functions and numerical experiments for second-order cone complementarity problems*, Computational Optimization and Applications, vol. 25, pp. 39–56.
- [7] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).
- [8] J. Y. FAN AND Y. X. YUAN (2005), *On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption*, Computing, vol. 74, pp. 23–39.
- [9] J. FARAUT AND A. KORANYI, *Analysis on Symmetric cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.
- [10] F. FACCHINEI AND J.-S. PANG, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York, 2003.

- [11] M. C. FERRIS, O. L. MANGASARIAN, AND J.-S. PANG, editors, *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic Publishers, Dordrecht, 2001.
- [12] A. FISCHER (1997), *Solution of the monotone complementarity problem with locally Lipschitzian functions*, *Mathematical Programming*, vol. 76, pp. 513–532.
- [13] M. FUKUSHIMA, Z.-Q. LUO, AND P. TSENG (2002), *Smoothing functions for second-order cone complementarity problems*, *SIAM Journal on Optimization*, vol. 12, pp. 436–460.
- [14] L. GRIPPO, F. LAMPARIELLO AND S. LUCIDI (1986), *A nonmonotone line search technique for Newton’s method*, *SIAM Journal on Numerical Analysis*, 1986, vol. 23, pp. 707–716.
- [15] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*. Cambridge University Press, Cambridge, 1986.
- [16] S. HAYASHI, N. YAMASHITA, AND M. FUKUSHIMA (2005), *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, *SIAM Journal of Optimization*, vol. 15, pp. 593–615.
- [17] C. KANZOW, I. FERENCZI AND M. FUKUSHIMA, *Semismooth methods for linear and nonlinear second-order cone programs*, Technical Report, Department of Applied Mathematics and Physics, Kyoto University, 2006.
- [18] Y. KANNO, J. A. C. MARTINS AND A. PINTO DA COSTA (2006), *Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem*, *International Journal for Numerical Methods in Engineering*, vol. 65, pp. 62–83.
- [19] C. KANZOW AND S. PETRA (2004), *On a semismooth least squares formulation of complementarity problems with gap reduction*, *Optimization Methods & Software*, vol. 19, pp. 507–525.
- [20] C. KANZOW, N. YAMASHITA AND M. FUKUSHIMA (1997), *New NCP functions and their properties*, *Journal of Optimization Theory and Applications*, vol. 94, pp. 115–135.
- [21] M. S. LOBO, L. VANDENBERGHE, S. BOYD, AND H. LEBRET (1998), *Application of second-order cone programming*, *Linear Algebra and its Applications*, vol. 284, pp. 193–228.
- [22] T. DE LUCA, F. FACCHINEI AND C. KANZOW (1996), *A semismooth equation approach to the solution of nonlinear complementarity problems*, *Mathematical Programming*, vol. 75, pp. 407–439.

- [23] R. D. C. MONTEIRO AND T. TSUCHIYA (2000), *Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions*, *Mathematical Programming*, vol. 88, pp. 61–83.
- [24] S.-H. PAN AND J.-S. CHEN (2009), *A damped Gauss-Newton method for the second-order conecomplementarity problem*, *Applied Mathematics and Optimization*, vol. 59, pp. 293-318.
- [25] G. PATAKI AND S. SCHMIETA, *The DIMACS library of semidefinite-quadratic-linear programs*, Preliminary draft, Computational Optimization Research Center, Columbia University. [http://dimacs.rutgers.edu/Challenges/Seventh Instances/](http://dimacs.rutgers.edu/Challenges/SeventhInstances/).
- [26] L. QI (1993), *Convergence analysis of some algorithms for solving nonsmooth equations*, *Mathematics of Operations Research*, vol. 18, pp. 227–244.
- [27] L. QI AND J. SUN (1993), *A nonsmooth version of Newton’s method*, *Mathematical Programming*, vol. 58, pp. 353–367.
- [28] D. SUN AND J. SUN (2005), *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions*, *Mathematical Programming*, vol. 103, pp. 575–581.
- [29] J. F. STURM (1999), *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, *Optimization Methods and Software*, vol. 11& 12, pp. 625–653.
- [30] T. TSUCHIYA (1999), *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, *Optimization Methods and Software*, vol. 11, pp. 141–182.
- [31] N. YAMASHITA AND M. FUKUSHIMA (2001), *On the rate of convergence of the Levenberg-Marquardt method*, *Computing*, vol. 15, pp. 237–249.
- [32] A. YOSHISE (2006), *Interior point trajectories and a homogenous model for nonlinear complementarity problems over symmetric cones*, *SIAM Journal on Optimization*, vol. 17, pp. 1129–1153.