Optimization Methods & Software
Vol. 26, No. 1, February 2011, 1–22

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

Shaohua Pan[a] and Jein-Shan Chen[b]*

[a]*Department of Mathematics, South China University of Technology, Guangzhou 510640, People's Republic of China;* [b]*Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan, Republic of China*

(*Received 11 March 2008; final version received 30 June 2009*)

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. This formulation has two-fold advantages. First, the operator involved in the over-determined system of equations inherits the favourable properties of the FB function for local convergence, for example, the (strong) semi-smoothness; second, the natural merit function of the over-determined system of equations share all the nice features of the class of merit functions f_{YF} studied in [J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Math. Program. 104 (2005), pp. 293–327] for global convergence. We propose a semi-smooth Levenberg–Marquardt method to solve the arising over-determined system of equations, and establish the global and local convergence results. Among others, the superlinear (quadratic) rate of convergence is obtained under strict complementarity of the solution and a local error bound assumption, respectively. Numerical results verify the advantages of the least-square reformulation for difficult problems.

Keywords: second-order cone complementarity problem; Fischer–Burmeister function; semi-smooth; Levenberg–Marquardt method

1. Introduction

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

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

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $F : \mathbb{R}^n \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathbb{R}^n$ 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 [11], i.e.

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

with $n_1 + \cdots + n_q = n$ and $\mathcal{K}^{n_i} := \{(x_{i1}, x_{i2}) \in \mathbb{R} \times \mathbb{R}^{n_i-1} | x_{i1} \geq ||x_{i2}||\}$. In this paper, corresponding to the Cartesian structure of the cone $\mathcal{K}$, we will write $F = (F_1, \ldots, F_q)$ and $G = (G_1, \ldots, G_q)$ with $F_i, G_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$.

*Corresponding author. Email: jschen@math.ntnu.edu.tw

ISSN 1055-6788 print/ISSN 1029-4937 online

DOI: 10.1080/10556780903180366
http://www.informaworld.com

An important special case of problem (1) corresponds to $G(\zeta) \equiv \zeta$, i.e.

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

This is a natural extension of the non-linear complementarity problem (NCP) [9,12], where $\mathcal{K} = \mathbb{R}^n_+$, the non-negative 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):

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

where $g : \mathbb{R}^n \to \mathbb{R}$ 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

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

where $\hat{x} \in \mathbb{R}^n$ satisfies $Ax = b$; see [5] for details. The convex SOCP arises in many applications from engineering design, finance, and robust optimization; see [1,20] and references therein. Motivated by Kanno *et al.* [17] 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 cannot reduce to SOCPs.

Various methods have been proposed for solving convex SOCPs and SOCCPs, including the interior point methods [1,2,20,21,28,30], the smoothing Newton methods [6,14,16], the merit function method [5] and the semi-smooth Newton method [19]. Among others, the last three kinds of methods are all based on an SOC complementarity function or a merit function. Specifically, $\phi : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$ is called an *SOC complementarity function* associated with $\mathcal{K}^{n_i}$ if

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

Clearly, when $n_i = 1$, an SOC complementarity function becomes an NCP function. A popular choice of ϕ is the Fischer–Burmeister (FB) function defined by

$$\phi_{\mathrm{FB}}(x_i, y_i) := (x_i^2 + y_i^2)^{1/2} - (x_i + y_i) \quad \forall x_i, y_i \in \mathbb{R}^{n_i}, \tag{7}$$

where $x_i^2 = x_i \circ x_i$ means the Jordan product of x_i with itself (the definition of Jordan product is given in Section 2), and $(x_i)^{1/2}$ means a vector such that $[(x_i)^{1/2}]^2 = x_i$. The function ϕ_{FB} is well-defined and satisfies (6); see [14]. Hence, the SOCCP (1) can be reformulated as the following non-smooth system

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

The system (8) induces a natural merit function $\Psi_{\mathrm{FB}} : \mathbb{R}^n \to \mathbb{R}_+$ for (1), given by

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

with

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

The function ψ_{FB} was studied in [5] and used to develop a merit function method. Recently, we analysed in [22] 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 Kanzow and Petra [18] for the NCPs, we present a new reformulation for (1) in this paper to overcome the disadvantage of ϕ_{FB}. Let $\phi_0 : \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \to \mathbb{R}_+$ be given by

$$\phi_0(x_i, y_i) := \max\{0, \langle x_i, y_i\rangle\}, \tag{11}$$

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

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

where ρ_1, ρ_2 are arbitrary but fixed constants from (0,1) used as the weights for the first type of terms and the second one, respectively. 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

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

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

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

where

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

The reformulation has the following advantages: on the one hand, Ψ belongs to the class of merit functions f_{YF} introduced in [5], which will be shown to have more desirable properties than Ψ_{FB}; on the other hand, Φ inherits the semi-smoothness of Φ_{FB} even strong semi-smoothness under some conditions. By this, we propose a semi-smooth Levenberg–Marquardt type method for solving (14), and establish the superlinear (quadratic) rate of convergence under strict complementarity and a local error bound assumption of the solution, respectively.

Throughout this paper, I represents an identity matrix of suitable dimension, $||\cdot||$ denotes the Euclidean norm, $\mathbb{R}^n$ 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 a differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x)$ denotes the transpose of the Jacobian $F'(x)$. For a (not necessarily symmetric) square matrix $A \in \mathbb{R}^{n\times n}$, we write $A \succeq 0$ (respectively, $A \succ 0$) to mean A is positive semi-definite (respectively, positive definite). Given a finite number of matrices $Q_1, \ldots, Q_n$, we denote the block diagonal matrix with these matrices as block diagonals by diag $(Q_1, \ldots, Q_n)$. If $\mathcal{J}$ and $\mathcal{B}$ are index sets such that $\mathcal{J}, \mathcal{B} \subseteq \{1, 2, \ldots, q\}$, 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 $\mathrm{int}(\mathcal{K}^n)$, $\mathrm{bd}(\mathcal{K}^n)$ and $\mathrm{bd}^+(\mathcal{K}^n)$ by the interior, the boundary of $\mathcal{K}^n$, and the boundary of $\mathcal{K}^n$ excluding the origin, respectively.

2. Preliminaries

This section recalls some background materials that will be used in the sequel. We start with the definition of the 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* [11] associated with $\mathcal{K}^n$ as

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

The Jordan product '$\circ$', unlike the 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 $e := (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbb{R}^n$. Given a vector $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let $L_x := \begin{bmatrix} x_1 & x_2^{\mathrm{T}} \\ x_2 & x_1 I \end{bmatrix}$ which can be viewed as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. It is easy to verify that $L_x y = x \circ y$ 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 $L_x \succeq 0$, and $x \in \mathrm{int}(\mathcal{K}^n)$ if and only if $L_x \succ 0$. When $x \in \mathrm{int}(\mathcal{K}^n)$, the inverse of L_x is given by

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^{\mathrm{T}} \\ -x_2 & \dfrac{\det(x)}{x_1} I + \dfrac{1}{x_1} x_2 x_2^{\mathrm{T}} \end{bmatrix}, \tag{17}$$

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

From [11,14], 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

$$x = \lambda_1(x) \cdot u_x^{(1)} + \lambda_2(x) \cdot u_x^{(2)}, \tag{18}$$

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

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

with $\bar{x}_2 = x_2/||x_2||$ if $x_2 \neq 0$ and otherwise being any vector in $\mathbb{R}^{n-1}$ with $||\bar{x}_2|| = 1$. If $x_2 \neq 0$, the factorization is unique. The spectral factorizations of x, x^2 as well as $x^{1/2}$ have various interesting properties [14]; for example, $x \in \mathcal{K}^n$ if and only if $0 \leq \lambda_1(x) \leq \lambda_2(x)$, and $x \in \mathrm{int}(\mathcal{K}^n)$ if and only if $0 < \lambda_1(x) \leq \lambda_2(x)$.

We next recall from Chen and Qi, [4] the definition of Cartesian P-property for a matrix and a nonlinear transformation.

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

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

(b) the Cartesian P_0-property if for any non-zero $\zeta = (\zeta_1, \ldots, \zeta_q) \in \mathbb{R}^n$ with $\zeta_i \in \mathbb{R}^{n_i}$, there exists an index $\nu \in \{1, 2, \ldots, q\}$ such that

$$\zeta_\nu \neq 0 \ \text{ and } \ \langle \zeta_\nu, (M\zeta)_\nu\rangle \geq 0.$$

DEFINITION 2.2 *The mappings $F = (F_1, \ldots, F_q)$ and $G = (G_1, \ldots, G_q)$ are said to have*

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

$$G_\nu(\zeta) \neq G_\nu(\xi) \ \text{ and } \ \langle F_\nu(\zeta) - F_\nu(\xi), G_\nu(\zeta) - G_\nu(\xi)\rangle \geq 0,$$

When $G(\zeta) \equiv \zeta$, Definition 2.2 gives the Cartesian P-properties of F. Obviously, the uniform Cartesian P-property $\Rightarrow$ the Cartesian P-property $\Rightarrow$ the Cartesian P_0-property. Also, a continuously differentiable mapping has the Cartesian P_0-property if and only if its Jacobian matrix at every point has the Cartesian P_0-property, and if the Jacobian matrix 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, the positive semi-definitness implies the Cartesian P_0-property.

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

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

is non-empty and 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 $\partial H(\zeta) := \mathrm{conv}\partial_{\mathrm{B}} H(\zeta)$ is the generalized Jacobian of H at ζ in the sense of Clarke [4]. For the concepts of (strongly) semi-smooth functions, please refer to [24,25] for details.

3. Properties of the operator Φ

To study the favourable properties of the operator Φ, we first give two technical lemmas to summarize some properties of ϕ_0 and ϕ_{FB}, respectively. The results of the first lemma are direct, and the results of the second lemma can be found in [14, Prop. 4.2], [5, Prop. 2], [27, Cor. 3.3] and [22, Prop. 3.1].

LEMMA 3.1 *Let $\phi_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ be defined as in Equation* (11)*. Then*,

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

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

where

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

LEMMA 3.2 *Let $\phi_{\mathrm{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be defined as in Equation* (7)*. Then, for any given $x = (x_1, x_2),\ y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the following results hold.*

(a) $\phi_{\mathrm{FB}}(x, y) = 0 \iff x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \quad \langle x, y\rangle = 0.$
(b) *ϕ_{FB} is strongly semismooth at (x, y).*

(c) *Each element* $[U_x - I \;\; U_y - I]$ *of* $\partial_B \phi_{FB}(x, y)$ *has the following representation:*
   (c.1) *If* $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$, *then* $U_x = L^{-1}_{(x^2+y^2)^{1/2}} L_x$ and $U_y = L^{-1}_{(x^2+y^2)^{1/2}} L_y$.
   (c.2) *If* $x^2 + y^2 \in \mathrm{bd}^+(\mathcal{K}^n)$, *then* $[U_x, \; U_y]$ *belongs to the set*

$$\left\{ \left[\frac{1}{2\sqrt{2w_1}} \begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{pmatrix} L_x + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T, \right. \right.$$

$$\left. \frac{1}{2\sqrt{2w_1}} \begin{pmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{pmatrix} L_y + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T \right] \Big| \, u = (u_1, u_2),$$

$$\left. v = (v_1, v_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \text{ satisfy } |u_1| \le \|u_2\| \le 1, |v_1| \le \|v_2\| \le 1 \right\},$$

where $w = (w_1, w_2) = x^2 + y^2$ and $w_2 = w_2/\|w_2\|$.

   (c.3) *If* $(x, y) = (0, 0)$, $[U_x, \; U_y]$ *belongs to* $\{[L_{\hat{u}}, L_{\hat{v}}] | \; \|\hat{u}\|^2 + \|\hat{v}\|^2 = 1\}$ *or*

$$\left\{ \left[\frac{1}{2} \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix} \xi^T + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T + 2 \begin{pmatrix} 0 & 0 \\ 0 & (I - \bar{w}_2\bar{w}_2^T) \end{pmatrix} L_s, \right. \right.$$

$$\left. \frac{1}{2} \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix} \eta^T + \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T + 2 \begin{pmatrix} 0 & 0 \\ 0 & (I - \bar{w}_2\bar{w}_2^T) \end{pmatrix} L_\omega \right] \; || \; \bar{w}_2 \text{ satisfies}$$

$$\|\bar{w}_2\| = 1 \text{ and } u = (u_1, u_2), v = (v_1, v_2), \; \xi = (\xi_1, \xi_2), \eta = (\eta_1, \eta_2),$$

$$s = (s_1, s_2), \; \omega = (\omega_1, \omega_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \text{ satisfy } |\xi_1| \le \|\xi_2\| \le 1,$$

$$\left. |u_1| \le \|u_2\| \le 1, |\eta_1| \le \|\eta_2\| \le 1, |v_1| \le \|v_2\| \le 1, \|s\|^2 + \|\omega\|^2 \le \frac{1}{2} \right\}.$$

(d) *The squared norm of* ϕ_{FB}, *i.e.* Ψ_{FB}, *is continuously differentiable at* (x, y).

From Lemma 3.1 (b) and Lemma 3.2 (b), we obtain the semi-smoothness of Φ.

PROPOSITION 3.3 *The operator* $\Phi : \mathbb{R}^n \to \mathbb{R}^{n+q}$ *defined by* (12) *is semi-smooth. If, in addition,* F' *and* G' *are Lipschitz continuous, then* Φ *is strongly semi-smooth.*

Proof Let Φ_i denote the ith component function of Φ for $i = 1, 2, \ldots, 2q$, i.e., $\Phi_i(\zeta) = \phi_{FB}(F_i(\zeta), G_i(\zeta))$ for $i = 1, 2, \ldots, q$ and $\Phi_i(\zeta) = \phi_0(F_i(\zeta), G_i(\zeta))$ for $i = q + 1, \ldots, 2q$. Then, the mapping Φ is (strongly) semi-smooth if every Φ_i is (strongly) semi-smooth. Note that $\Phi_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ for $i = 1, 2, \ldots, q$ is the composite of the strongly semi-smooth function ϕ_{FB} and the smooth function $\zeta \mapsto (F_i(\zeta), G_i(\zeta))$, whereas $\Phi_{q+i} : \mathbb{R}^n \to \mathbb{R}$ is the composite of the strongly semi-smooth 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 semi-smooth. By [13, Theorem 19], we have that every component function of Φ is semi-smooth, and strongly semi-smooth if, in addition, F' and G' are Lipschitz continuous. ■

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

PROPOSITION 3.4 *Let $\Phi : \mathbb{R}^n \to \mathbb{R}^{n+q}$ be given by* (12)*. Then, for any $\zeta \in \mathbb{R}^n$,*

$$\partial_{\mathrm{B}}\Phi(\zeta)^{\mathrm{T}} \subseteq \nabla F(\zeta)[\rho_1(A(\zeta)-I)\ \ \rho_2 C(\zeta)] + \nabla G(\zeta)[\rho_1(B(\zeta)-I)\ \ \rho_2 D(\zeta)]$$

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

$$C_i(\zeta) \in G_i(\zeta)\partial_{\mathrm{B}}(F_i(\zeta)^{\mathrm{T}}G_i(\zeta))_+ \ \text{ and } \ D_i(\zeta) \in F_i(\zeta)\partial_{\mathrm{B}}(F_i(\zeta)^{\mathrm{T}}G_i(\zeta))_+,$$

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

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

$$A_i(\zeta) = L_{F_i(\zeta)}L^{-1}_{[F_i(\zeta)^2+G_i(\zeta)^2]^{1/2}} \text{ and } B_i(\zeta) = L_{G_i(\zeta)}L^{-1}_{[F_i(\zeta)^2+G_i(\zeta)^2]^{1/2}}.$$

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

$$\left\{\left[\frac{1}{2\sqrt{2w_{i1}(\zeta)}}L_{F_i(\zeta)}\begin{pmatrix}1 & \bar{w}_{i2}(\zeta)^{\mathrm{T}}\\ \bar{w}_{i2}(\zeta) & 4I-3\bar{w}_{i2}(\zeta)\bar{w}_{i2}(\zeta)^{\mathrm{T}}\end{pmatrix} + \frac{1}{2}u_i(1,-\bar{w}_{i2}(\zeta)^{\mathrm{T}}),\right.\right.$$
$$\left.\frac{1}{2\sqrt{2w_{i1}(\zeta)}}L_{G_i(\zeta)}\begin{pmatrix}1 & \bar{w}_{i2}(\zeta)^{\mathrm{T}}\\ \bar{w}_{i2}(\zeta) & 4I-3\bar{w}_{i2}(\zeta)\bar{w}_{i2}(\zeta)^{\mathrm{T}}\end{pmatrix} + \frac{1}{2}v_i(1,-\bar{w}_{i2}(\zeta)^{\mathrm{T}})\right] \Bigg|$$
$$\left. u_i = (u_{i1}, u_{i2}),\ v_i = (v_{i1}, v_{i2}) \text{ satisfy } |u_{i1}| \le \|u_{i2}\| \le 1,\ |v_{i1}| \le \|v_{i2}\| \le 1\right\},$$

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

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

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

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

$$\partial_{\mathrm{B}}\Phi(\zeta)^{\mathrm{T}} \subseteq \partial_{\mathrm{B}}\Phi_1(\zeta)^{\mathrm{T}} \times \partial_{\mathrm{B}}\Phi_2(\zeta)^{\mathrm{T}} \times \cdots \times \partial_{\mathrm{B}}\Phi_{2q}(\zeta)^{\mathrm{T}}, \tag{19}$$

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

$$\begin{aligned}\partial_{\mathrm{B}}\Phi_i(\zeta)^{\mathrm{T}} &\subseteq \rho_1[\nabla F_i(\zeta)\ \ \nabla G_i(\zeta)]\ \partial_{\mathrm{B}}\phi_{\mathrm{FB}}(F_i(\zeta), G_i(\zeta))^{\mathrm{T}},\\ \partial_{\mathrm{B}}\Phi_{q+i}(\zeta)^{\mathrm{T}} &\subseteq \rho_2[\nabla F_i(\zeta)\ \ \nabla G_i(\zeta)]\ \partial_{\mathrm{B}}\phi_0(F_i(\zeta), G_i(\zeta))^{\mathrm{T}}.\end{aligned} \tag{20}$$

Also, by Lemmas 3.1(c) and 3.2(c), each element in $\partial_{\mathrm{B}}\phi_{\mathrm{FB}}(F_i(\zeta), G_i(\zeta))^{\mathrm{T}}$ and $\partial_{\mathrm{B}}\phi_0(F_i(\zeta), G_i(\zeta))^{\mathrm{T}}$ 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 $i = 1, \ldots, q$ characterized as in the proposition. Combining with Equations (19) and (20), we obtain the desired result. ■

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

$$\begin{aligned}
\mathcal{I} &:= \{i \in \{1, 2, \ldots, q\} \mid F_i(\zeta^*) = 0,\ G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})\},\\
\mathcal{B} &:= \{i \in \{1, 2, \ldots, q\} \mid F_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i}),\ G_i(\zeta^*) \in \mathrm{bd}^+(\mathcal{K}^{n_i})\},\\
\mathcal{J} &:= \{i \in \{1, 2, \ldots, q\} \mid F_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i}),\ G_i(\zeta^*) = 0\}.
\end{aligned} \tag{21}$$

If ζ^* satisfies *strict complementarity*, i.e. $F_i(\zeta^*) + G_i(\zeta^*) \in \mathrm{int}(\mathcal{K}^{n_i})$ for all i, then $\{1, 2, \ldots, q\}$ can be partitioned as $\mathcal{I} \cup \mathcal{B} \cup \mathcal{J}$. Thus, if $\nabla G(\zeta^*)$ is invertible, then by rearrangement the matrix $P(\zeta^*) = \nabla G(\zeta^*)^{-1}\nabla F(\zeta^*)$ can be rewritten as

$$P(\zeta^*) = \begin{pmatrix} P(\zeta^*)_{\mathcal{II}} & P(\zeta^*)_{\mathcal{IB}} & P(\zeta^*)_{\mathcal{IJ}} \\ P(\zeta^*)_{\mathcal{BI}} & P(\zeta^*)_{\mathcal{BB}} & P(\zeta^*)_{\mathcal{BJ}} \\ P(\zeta^*)_{\mathcal{JI}} & P(\zeta^*)_{\mathcal{JB}} & P(\zeta^*)_{\mathcal{JJ}} \end{pmatrix}.$$

Now we have the following results for the full rank of every element $H \in \partial_B\Phi(\zeta^*)$.

THEOREM 3.5 *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 non-singular 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 $\left(\begin{smallmatrix} P(\zeta^*)_{\mathcal{II}} & P(\zeta^*)_{\mathcal{IB}} \\ P(\zeta^*)_{\mathcal{BI}} & P(\zeta^*)_{\mathcal{BB}} \end{smallmatrix}\right)$ has the Cartesian P-property, then every element H in the B-subdifferential $\partial_B\Phi(\zeta^*)$ has full column rank n.*

Proof Let $H \in \partial_B\Phi(\zeta^*)$. By Proposition 3.4, $H = \left(\begin{smallmatrix} \rho_1 H_1 \\ \rho_2 H_2 \end{smallmatrix}\right)$ with H_1^{T} from the set $\partial_B\Phi_1(\zeta^*)^{\mathrm{T}} \times \cdots \times \partial_B\Phi_q(\zeta^*)^{\mathrm{T}}$. From Theorem 4.1 of [22], H_1^{T} is non-singular under the given assumptions. This implies the desired result $\mathrm{rank}(H) = n$. ■

The proof of Theorem 3.5 is based on the important property of the first block of H. Nevertheless, when the first block H_1 is singular, the second block H_2 may contribute something to guarantee that H has a full column rank n.

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

LEMMA 3.6 *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 $\|(H^{\mathrm{T}}H)^{-1}\| \le c$ for all $\|\zeta - \zeta^*\| < \varepsilon$ and all $H \in \partial_B\Phi(\zeta)$. Furthermore, for any given $\bar{\nu} > 0$, $H^{\mathrm{T}}H + \nu I$ 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 [24, Lemma 2.6]. For completeness, we include it here. 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 $H_k^{\mathrm{T}}H_k$ is singular or $\|(H_k^{\mathrm{T}}H_k)^{-1}\| \to +\infty$ on a subsequence. Noting that $H_k^{\mathrm{T}}H_k$ is symmetric positive semi-definite, for the non-singular case, we have $\|(H_k^{\mathrm{T}}H_k)^{-1}\| = 1/\lambda_{\min}(H_k^{\mathrm{T}}H_k)$, which implies that the condition $\|(H_k^{\mathrm{T}}H_k)^{-1}\| \to +\infty$ is equivalent to $\lambda_{\min}(H_k^{\mathrm{T}}H_k) \to 0$. Since $\zeta^k \to \zeta^*$ and the mapping $\zeta \mapsto \partial_B\Phi(\zeta)$ is upper semi-continuous, 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 $\lambda_{\min}(H_*^{\mathrm{T}}H_*) = 0$

by the continuity of the minimum eigenvalue. This means that $H_*^{\mathrm{T}} H_*$ is singular. However, from the fact that the mapping $\zeta \mapsto \partial_{\mathrm{B}}\Phi(\zeta)$ is closed, we have $H_* \in \partial_{\mathrm{B}}\Phi(\zeta^*)$, which by the given condition implies that $H_*^{\mathrm{T}} H_*$ is non-singular. Thus, we obtain a contradiction, and the first part follows. By the result of the first part and the definition of matrix norm, there exist constants $\varepsilon > 0$ and $c > 0$ such that

$$[\lambda_{\min}(H^{\mathrm{T}}H + \nu I)]^{-1} = \|(H^{\mathrm{T}}H + \nu I)^{-1}\| \leq c$$

for all $\nu \in [0, \bar{\nu}]$ and $H \in \partial_{\mathrm{B}}\Phi(\zeta)$ with $\|\zeta - \zeta^*\| < \varepsilon$. This implies that

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

Therefore, all the matrices $H^{\mathrm{T}}H + \nu I$ are uniformly positive definite. ■

4. Properties of the merit function Ψ

This section is devoted to the favourable properties of Ψ defined by (14) and (15). To this end, we need the following lemma which summarizes the properties of ψ.

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

(a) $\psi(x, y) = 0 \iff \Psi_{\mathrm{FB}}(x, y) = 0 \iff x \in \mathcal{K}^n,\ y \in \mathcal{K}^n,\ \langle x, y\rangle = 0$*;*
(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 Lemmas 3.1(a) and 3.2(d). We next consider part (c). By the definition of ψ,

$$\begin{aligned}\nabla_x\psi(x, y) &= \rho_1^2\nabla_x\psi_{\mathrm{FB}}(x, y) + \rho_2^2\phi_0(x, y)y,\\ \nabla_y\psi(x, y) &= \rho_1^2\nabla_y\psi_{\mathrm{FB}}(x, y) + \rho_2^2\phi_0(x, y)x.\end{aligned} \tag{22}$$

From Lemma 6 (a) of [5] and the definition of $\phi_0(x, y)$, it then follows that

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

(d) Using the formulas in (22) and [5, Lemma 6(a)], it follows that

$$\begin{aligned}\langle \nabla_x\psi(x, y), \nabla_y\psi(x, y)\rangle &= \rho_1^4\langle \nabla_x\psi_{\mathrm{FB}}(x, y), \nabla_y\psi_{\mathrm{FB}}(x, y)\rangle + \rho_2^4 x^{\mathrm{T}}y\phi_0(x, y)^2\\ &\quad + \rho_1^2\rho_2^2\phi_0(x, y)[\langle x, \nabla_x\psi_{\mathrm{FB}}(x, y)\rangle + \langle y, \nabla_y\psi_{\mathrm{FB}}(x, y)\rangle]\\ &= \rho_1^4\langle \nabla_x\psi_{\mathrm{FB}}(x, y), \nabla_y\psi_{\mathrm{FB}}(x, y)\rangle + \rho_2^4\phi_0(x, y)^3\\ &\quad + 2\rho_1^2\rho_2^2\phi_0(x, y)\psi_{\mathrm{FB}}(x, y).\end{aligned} \tag{23}$$

The first term on the right-hand side of (23) is non-negative by [5, Lemma 6(b)], and the last two terms are also non-negative. 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

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

which, together with Lemma 6(b) of [5], implies the desired result.

(e) If $\psi(x, y) = 0$, then from the definition of ψ it follows that $\phi_{\rm FB}(x, y) = 0$ and $\phi_0(x, y) = 0$. From Proposition 1 of [5], we immediately obtain $\nabla_x \psi_{\rm FB}(x, y) = \nabla_y \psi_{\rm FB}(x, y) = 0$, and consequently $\nabla_x \psi(x, y) = 0$ and $\nabla_y \psi(x, y) = 0$ by (22). If $\nabla \psi(x, y) = 0$, then by part (c) and the non-negativity 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 get $\psi(x, y) = 0$, which together with part (a) and (22) implies $\nabla \psi(x, y) = 0$. Consequently, $\nabla \psi(x, y) = 0 \Longleftrightarrow \nabla_x \psi(x, y) = 0$. Similarly, $\nabla \psi(x, y) = 0 \Longleftrightarrow \nabla_y \psi(x, y) = 0$. This implies the last equivalence. ∎

Lemma 4.1(b) shows that Ψ is continuously differentiable. By Lemma 4.1(d), we can prove every stationary point of Ψ is a solution of Equation (1) under mild conditions.

Proposition 4.2 *Let* $\Psi : \mathbb{R}^n \to \mathbb{R}_+$ *be given by* (14) *and* (15). *Then every stationary point of* Ψ *is a solution of* (1) *under one of the following assumptions:*

(a) $\nabla F(\zeta)$ *and* $-\nabla G(\zeta)$ *are column monotone*[1] *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 [5, Prop. 3] yields the desired result. Now suppose that the assumption (b) holds. Let $\bar{\zeta}$ be an arbitrary stationary point of Ψ and write

$$\nabla_x \psi(F(\zeta), G(\zeta)) = (\nabla_{x_1} \psi(F_1(\zeta), G_1(\zeta)), \ldots, \nabla_{x_q} \psi(F_q(\zeta), G_q(\zeta))),$$
$$\nabla_y \psi(F(\zeta), G(\zeta)) = \left(\nabla_{y_1} \psi(F_1(\zeta), G_1(\zeta)), \ldots, \nabla_{y_q} \psi(F_q(\zeta), G_q(\zeta))\right).$$

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

$$\nabla G(\bar{\zeta})^{-1} \nabla F(\bar{\zeta}) \nabla_x \psi(F(\bar{\zeta}), G(\bar{\zeta})) + \nabla_y \psi(F(\bar{\zeta}), G(\bar{\zeta})) = 0. \tag{24}$$

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

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

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

$$\langle \nabla_{x_\nu} \psi(F_\nu(\bar{\zeta}), G_\nu(\bar{\zeta})), [\nabla G(\bar{\zeta})^{-1} \nabla F(\bar{\zeta}) \nabla_x \psi(F(\bar{\zeta}), G(\bar{\zeta}))]_\nu \rangle \geq 0. \tag{25}$$

In addition, note that (24) is equivalent to

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

Making the inner product with $\nabla_{x_\nu}\psi(F(\bar{\zeta}), G(\bar{\zeta}))$ for the νth equality, we obtain

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

The first term on the left-hand side is non-negative by (25), 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). ■

When $\nabla G(\zeta)$ is invertible for any $\zeta \in \mathbb{R}^n$, the assumption in (a) is equivalent to the positive semi-definiteness 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

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

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

Condition A. For any sequence $\{\zeta^k\}$ satisfying $\|\zeta^k\| \to +\infty$, whenever

$$\limsup \|[-F(\zeta^k)]_+\| < +\infty \quad\text{and}\quad \limsup \|[-G(\zeta^k)]_+\| < +\infty, \tag{26}$$

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

Proposition 4.3 *If the mappings F and G satisfy Condition A, then the level sets $\mathcal{L}_\Psi(\gamma)$ are bounded for all $\gamma \ge 0$.*

Proof Assume that there is a unbounded sequence $\{\zeta^k\} \subseteq \mathcal{L}_\Psi(\gamma)$ for some $\gamma \ge 0$. Since $\Psi(\zeta^k) \le \gamma$ for all k, the sequence $\{\Psi_{\mathrm{FB}}(\zeta^k)\}$ is bounded. By Lemma 8 of [5],

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

hold for all $i \in \{1, 2, \ldots, q\}$. This shows that F and G satisfy Condition A, 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. ■

Condition A is rather weak to guarantee that Ψ has bounded level sets since, as will be shown below, the condition is implied by the joint monotonicity of F and G with the strict feasibility of (1) used in [5] for f_{YF}, the jointly uniform Cartesian P-functions with a feasible point, and the joint $\tilde{R}_{01}$-property in the following sense.

Definition 4.4 *The mappings $F, G : \mathbb{R}^n \to \mathbb{R}^n$ are said to have the joint $\tilde{R}_{01}$-property if for any sequence $\{\zeta^k\}$ with*

$$\|\zeta^k\| \to +\infty, \quad \frac{\|[-G(\zeta^k)]_+\|}{\|\zeta^k\|} \to 0, \quad \frac{\|[-F(\zeta^k)]_+\|}{\|\zeta^k\|} \to 0, \tag{27}$$

there holds that

$$\liminf_{k\to+\infty} \frac{\langle F(\zeta^k), G(\zeta^k)\rangle}{\|\zeta^k\|} > 0. \tag{28}$$

PROPOSITION 4.5 *Condition A is satisfied if one of the following conditions holds:*

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

Proof In the proof, let $\{\zeta^k\}$ be a sequence satisfying $\|\zeta^k\|\to+\infty$ and (26) 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 $\limsup\|[-G(\zeta^k)]_+\|=+\infty$, contradicting the assumption that $\{\zeta^k\}$ satisfies Equation (26). Noting that $\|F(\zeta^k)\|+\|G(\zeta^k)\|\to+\infty$ 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 $\{\lambda_1[F(\zeta^k)]\}$ and $\{\lambda_1[G(\zeta^k)]\}$ implies that

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

From the proof of [5, Lemma 9(b)] it then follows that

$$\limsup\{\langle F(\zeta^k),G(\hat{\zeta})\rangle+\langle F(\hat{\zeta}),G(\zeta^k)\rangle\}=+\infty. \tag{29}$$

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

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

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

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

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 (29), and consequently the desired result follows.

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

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

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 (26) that

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

which immediately implies the result.

(c) Clearly, $\{\zeta^k\}$ satisfies (27), and the result follows by the following implications:

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

So far, we complete the proof of this proposition. ■

When $G(\zeta) \equiv \zeta$, if we replace (28) with $\liminf_{k\to+\infty} \langle F(\zeta^k), G(\zeta^k)\rangle/\|\zeta^k\|^2 > 0$, then Definition 4.4 is saying that F is a R_{01}-function. Thus, Propositions 4.3 and 4.5 (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 [5, Prop. 5].

PROPOSITION 4.6 *Let ζ^* be a solution of* (1). *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 $\nu \in \{1, \ldots, q\}$ such that

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

where $c := \max\{1, \|G_\nu(\zeta^*)\|, \|F_\nu(\zeta^*)\|\}$, the second inequality is using the fact that $G_\nu(\zeta^*) \in \mathcal{K}^{n_\nu}$ and $F_\nu(\zeta^*) \in \mathcal{K}^{n_\nu}$, and the next to last inequality is due to [5, Lemma 8]. Letting $\kappa := (c/\rho)(\sqrt{2}/\rho_2 + 4/\rho_1)$, we obtain the desired result. ■

5. Algorithm and convergence

It is known that the Levenberg–Marquardt method using (12) has the advantage that it reduces the complementarity gap $\langle \zeta, F(\zeta)\rangle$ for the NCPs faster than the traditional non-smooth method based on (8) does [18]. This motivates us to employ 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 (*semi-smooth Levenberg–Marquardt method*)

(S.0) *Choose* $\zeta^0 \in \mathbb{R}^n$, $\rho_1, \rho_2 \in (0,1)$, $\eta, \beta \in (0,1)$, $\sigma \in (0, 1/2)$ *and* $\varepsilon \geq 0$. *Set* $k := 0$.
(S.1) *If* $\|\nabla\Psi(\zeta^k)\| \leq \varepsilon$, *then stop. Otherwise, go to the next step.*
(S.2) *Choose* $H_k \in \partial_{\mathrm{B}}\Phi(\zeta^k)$ *and* $\nu_k > 0$. *Find a solution* $d^k \in \mathbb{R}^n$ *of the linear system*

$$(H_k^{\mathrm{T}} H_k + \nu_k I)d = -\nabla\Psi(\zeta^k), \tag{30}$$

where $\nu_k > 0$ *is the Levenberg–Marquardt parameter.*
(S.3) *If* d^k *satisfies*

$$\|\Phi(\zeta^k + d^k)\| \leq \eta\|\Phi(\zeta^k)\|, \tag{31}$$

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

$$\Psi(\zeta^k + t_k d^k) \leq \Psi(\zeta^k) + \sigma t_k \nabla\Psi(\zeta^k)^{\mathrm{T}} d^k. \tag{32}$$

(S.4) *Let* $\zeta^{k+1} := \zeta^k + t_k d^k$, $k := k+1$, *and go to* (*S.1*).

The above method is different from the classical Levenberg–Marquardt method for non-linear least-square problems in that Φ is not continuously differentiable. If $\nu_k \equiv 0$, the solution of (30) is exactly the solution of the linear least-square problem

$$\min_{d\in\mathbb{R}^n} \frac{1}{2}\|H_k d + \Phi(\zeta^k)\|^2, \tag{33}$$

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

$$\nu_k := \min\{p_1, p_2\|\Phi(\zeta^k)\|^{\varrho}\}, \tag{34}$$

where $p_1, p_2 > 0$ are given constants and ϱ is a real number from [1,2]. Such a choice is consistent with the requirements for local superlinear (quadratic) convergence stated in Theorems 5.3 and 5.6 below, as well as adopted by numerical experiments.

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

THEOREM 5.2 *Let* $\{\zeta^k\}$ *be the sequence generated by Algorithm 5.1 with* ν_k *updated by* (34). *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 (30) is always a descent direction of Ψ at ζ^k. Let ζ^* be an arbitrary 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 (31) holds for only finitely many $k \in K$, and the inequality (32) is satisfied for all sufficiently large k. Since

$$\Psi(\zeta^{k+1}) - \Psi(\zeta^k) \leq \sigma t_k \nabla\Psi(\zeta^k)^{\mathrm{T}} d^k \leq 0$$

for sufficiently large k, using $\Psi(\zeta^{k+1}) - \Psi(\zeta^k) \to 0$ yields that

$$\{t_k \nabla\Psi(\zeta^k)^{\mathrm{T}} d^k\}_K \to 0. \tag{35}$$

We next prove that $\{\nabla\Psi(\zeta^k)^{\mathrm{T}} d^k\}_K$ has a non-zero limit. By the definition of d^k,

$$\nabla\Psi(\zeta^k)^{\mathrm{T}} d^k = -\nabla\Psi(\zeta^k)^{\mathrm{T}}(H_k^{\mathrm{T}} H_k + \nu_k I)^{-1}\nabla\Psi(\zeta^k) \quad \forall k. \tag{36}$$

Since the B-subdifferential $\partial_{\mathrm{B}}\Phi(\zeta)$ is a non-empty compact set for any $\zeta \in \mathbb{R}^n$, $\{H_k\}_K$ is bounded. Without loss of generality, assume that $\{H_k\}_K \to H_*$. Taking into account that the set-valued mapping $\zeta \mapsto \partial_{\mathrm{B}}\Phi(\zeta)$ is closed and $\{\zeta^k\}_K \to \zeta^*$, we have $H_* \in \partial_{\mathrm{B}}\Phi(\zeta^*)$. In addition, since $\Phi(\zeta^*) \neq 0$, we have $\nu_k \to \nu_*$ with $\nu_* = \min\{p_1, p_2\|\Phi(\zeta^*)\|^{\varrho}\} > 0$. Thus, $\{H_k^{\mathrm{T}} H_k + \nu_k I\}_{k\in K} \to H_*^{\mathrm{T}} H_* + \nu_* I \succ 0$. This, together with (36) and the continuity of $\nabla\Psi$, implies that $\{\nabla\Psi(\zeta^k)^{\mathrm{T}} d^k\}_K$ has a non-zero limit as $k \to +\infty$. From (35), it then follows that $\{t_k\}_K \to 0$. Now, for sufficiently large k, let $l_k \in \{0, 1, \ldots\}$ be the unique exponent such that $t_k = \beta^{l_k}$. Since $\{t_k\}_K \to 0$, we have $\{l_k\}_{k\in K} \to \infty$. From the Armijo line search in (S.3), for sufficiently large $k \in K$,

$$\frac{\Psi(\zeta^k + \beta^{l_k-1} d^k) - \Psi(\zeta^k)}{\beta^{l_k-1}} > \sigma\nabla\Psi(\zeta^k)^{\mathrm{T}} d^k. \tag{37}$$

Taking the limit $k \to \infty$ with $k \in K$ and using $\{l_k\}_K \to \infty$ and $\{\zeta^k\}_K \to \zeta^*$, we have $\nabla\Psi(\zeta^*)^{\mathrm{T}} d^* \geq \sigma\nabla\Psi(\zeta^*)^{\mathrm{T}} d^*$. This means $\nabla\Psi(\zeta^*)^{\mathrm{T}} d^* \geq 0$. On the other hand, we learn from Equation (30) that $\{d^k\}_K \to d^*$ with d^* being the solution of

$$(H_*^{\mathrm{T}} H_* + \nu_* I)d = -\nabla\Psi(\zeta^*), \tag{38}$$

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

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 4.3 and 4.5, the existence of accumulation points of $\{\zeta^k\}$ is guaranteed by one of the assumptions of Proposition 4.5. 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 4.5. 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) 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 5.3 *Let $\{\zeta^k\}$ be generated by Algorithm 5.1 with ν_k given by* (34). *Suppose that ζ^* is an accumulation point of $\{\zeta^k\}$ with ζ^* being a strictly complementary solution of Equation* (1), *and F and G at ζ^* satisfy the condition of Theorem* 3.5. *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 [18]. We include it for completeness.

(a) By the proof technique of Theorem 3.1(b) of [8], it suffices to prove that ζ^* is an isolated solution. From Theorem 3.5 and Lemma 3.6, there exist $\varepsilon_1, \kappa_1 > 0$ such that

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

for all ζ satisfying $\|\zeta - \zeta^*\| < \varepsilon_1$ and all $H \in \partial_{\mathrm{B}}\Phi(\zeta)$. In addition, the semi-smoothness of Φ implies that there exists $\varepsilon_2 > 0$ such that

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

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

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

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

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

$$\|\zeta^k + d^k - \zeta^*\| = o(\|\zeta^k - \zeta^*\|). \tag{39}$$

By part (a), the sequence $\{\zeta^k\}$ converges to a solution ζ^* satisfying the assumptions of Theorem 3.5. By Lemma 3.6, there exists $c > 0$ such that $\|(H_k^{\mathrm{T}}H_k + \nu_k I)^{-1}\| \leq c$ for all k. Noting that the sequence $\{H_k\}$ is bounded, there exists $c_1 > 0$ such that $\|H_k^{\mathrm{T}}\| \leq c_1$ for all k. Using Theorem 5.2 and the fact that $\Phi(\zeta^*) = 0$, we obtain

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

Note that $\Phi(\zeta^k) - \Phi(\zeta^*) - H_k(\zeta^k - \zeta^*) = o(\|\zeta^k - \zeta^*\|)$ by the semi-smoothness of Φ, whereas $\nu_k \to 0$ by part (a) and the continuity of Φ. Thus, the inequality above implies Equation (39). To prove that the full step is eventually accepted, by (31) it suffices to show that

$$\lim_{k\to\infty}\frac{\Psi(\zeta^k + d^k)}{\Psi(\zeta^k)} = 0. \tag{40}$$

Since all elements $V \in \partial_{\mathrm{B}}\Phi_{\mathrm{FB}}(\zeta^*)$ are non-singular by [22, Theorem 4.1], from Lemma 3.6 and the proof of part (a), there exists a constant $\alpha > 0$ such that

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

Using the locally Lipschitz continuity of Φ and Equation (39) then yields

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

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 (39).

(c) The proof is essentially same as for the superlinear convergence. We only note that ν_k in Equation (34) 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 semi-smoothness of Φ by Proposition 3.3. ■

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

ASSUMPTION 5.4 *There exist constants* $\kappa_2 > 0$ *and* $0 < \delta < 1$ *such that*

$$\kappa_2 \mathrm{dist}(\zeta, S^*) \leq \|\Phi(\zeta)\| \quad \forall \zeta \in \mathcal{N}(\zeta^*, \delta), \tag{41}$$

where S^* *denotes the solution set of* (1) *and is assumed to be non-empty.*

LEMMA 5.5 *Let* ζ^k *be generated by Algorithm 5.1 with* ν_k *given by* (34). *Suppose that* F' *and* G' *are Lipschitz continuous on* $\mathcal{N}(\zeta^*, \delta)$ *and Assumption 5.4 holds. If* $\nu_k = p_2\|\Phi(\zeta^k)\|^{\varrho}$ *and* $\zeta^k \in \mathcal{N}(\zeta^*, \delta/2)$, *then there exists a constant* $c_1 > 0$ *such that* $\|d^k\| \leq c_1 \mathrm{dist}(\zeta^k, S^*)$. *If, in addition,* $\zeta^k + d^k \in \mathcal{N}(\zeta^*, \delta/2)$, *then there exists a constant* $c_3 > 0$ *such that*

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

Proof Let $\bar{\zeta}^k \in S^*$ be such that $\|\zeta^k - \bar{\zeta}^k\| = \mathrm{dist}(\zeta^k, S^*)$. 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 exists a $L_1 > 0$ such that

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

Combining with the inequality (41), we have

$$p_2\kappa_2^{\varrho}\|\bar{\zeta}^k - \zeta^k\|^{\varrho} \leq \nu_k = p_2\|\Phi(\zeta^k)\|^{\varrho} \leq p_2 L_1^{\varrho}\|\zeta^k - \bar{\zeta}^k\|^{\varrho}. \tag{42}$$

On the other hand, since Φ is strongly semi-smooth on $\mathcal{N}(\zeta^*, \delta)$ by Proposition 3.3, there exists a constant $\hat{c} > 0$ such that

$$\|\Phi(\zeta^k) + H_k(\bar{\zeta}^k - \zeta^k)\| = \|\Phi(\zeta^k) - \Phi(\bar{\zeta}^k) - H_k(\zeta^k - \bar{\zeta}^k)\| \leq \hat{c}\|\zeta^k - \bar{\zeta}^k\|^2. \tag{43}$$

Define

$$\varphi_k(d) := \|\Phi(\zeta^k) + H_k d\|^2 + \nu_k\|d\|^2. \tag{44}$$

Then, d^k is a minimizer of $\varphi_k(d)$. This, together with (43) and (42), yields that

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

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

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

we have

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

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,

$$\begin{aligned}\mathrm{dist}(\zeta^k + d^k, S^*) \le \frac{1}{\kappa_2}\|\Phi(\zeta^k + d^k)\| &\le \frac{c_2}{\kappa_2}\|\zeta^k - \bar{\zeta}^k\|^{(\varrho+2)/2} \\ &= c_3\mathrm{dist}(\zeta^k, S^*)^{(\varrho+2)/2},\end{aligned}$$

where $c_3 = c_2/\kappa_2$. Thus, we complete the proof of the second part. ■

By Lemma 5.5, using arguments similar to [10, Theorem 2.1] and [29, Theorem 3.1], we get the quadratic rate of convergence of Algorithm 5.1 under Assumption 5.4.

THEOREM 5.6 *Let $\{\zeta^k\}$ be generated by Algorithm 5.1 with ν_k given by* (34), *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 5.4 holds.*

Now, we do not know whether Assumption 5.4 is weaker than the strict complementarity of the solution, although the assumptions of Theorem 5.6 are weaker than those of Theorem 5.3, since the latter implies that each element in $\partial_B\Phi(\zeta^*)$ is non-singular, and so $\|\Phi(\zeta)\|$ provides a local error bound on some neighbourhood of the solution ζ^*, but from [29] the former does not imply the non-sigularity of each element in $\partial_B\Phi(\zeta^*)$. From the proof of Lemma 5.5, we find that the condition (41) cannot be weakened to

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

in order to guarantee the superlinear (or quadratic) convergence of Algorithm 5.1, and therefore the global error bound result of Proposition 4.6 may not be applied for it. If let $\Psi(\zeta) = \|\Phi(\zeta)\|^4/4$ instead of $\Psi(\zeta) = \|\Phi(\zeta)\|^2/2$, then Assumption 5.4 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 5.4 to hold.

6. Numerical results

This section will report numerical results with the least-square semi-smooth method (LS *semi-smooth 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 to get the corresponding test instances for (1).

All experiments were done with a PC of Intel Pentium Dual CPU E2200 and 2047 MB memory, and the computer codes were written in Matlab 7.0. Since the non-monotone line search [15] is usually superior to the classical monotone line search, we replaced the Armijo line search of Algorithm 5.1 by the non-monotone version in [15], i.e. we computed t_k such that

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

where

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

and where, for a given non-negative integer $\hat{m}$ and s, $m_k = 0$ if $k \le s$, 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 (34) 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

$$\max\{|F(\zeta^k)^{\mathrm{T}} G(\zeta^k)|, \Psi(\zeta^k)\} \le 10^{-6}, \quad \text{or} \quad k > 150, \quad \text{or } t_k < 10^{-15}. \tag{45}$$

We compared the numerical performance of Algorithm 5.1 with that of the least-square semi-smooth Newton method based on (8), called the FB *semi-smooth 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 semi-smooth methods with those of SeDuMi [26], a successful interior point method software for the linear SOCPs and the semi-definite programming. The parameters of the SeduMi were set as default values.

The first group of test instances is the linear SOCPs from the DIMACS Implementation Challenge library [23]. 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 semi-smooth Newton methods are able to yield a solution with favourable accuracy for all test problems, and require 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. For the two difficult test problems, the LS semi-smooth method requires less iterations and function evaluations than the FB semi-smooth method. Also, for 'nb_L1', the advantage of the LS semi-smooth method is more remarkable.

The second group of test instances is the non-linear 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: $\min_{u\ge 0} \sum_{i=1}^{k} \|s_{[i]}\| + h(u)$, where $\|s_{[1]}\|, \ldots, \|s_{[r]}\|$

Table 1. Numerical results for the DIAMCS linear SOCPs.

	LS semi-smooth method			FB semi-smooth 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

are the norms $\|s_1\|, \ldots, \|s_r\|$ sorted in non-increasing order with $r \geq k$ and $s_i = b_i - A_i x$ for $i = 1, \ldots, r$ with $A_i \in \mathbb{R}^{m_i \times l}$ and $b_i \in \mathbb{R}^{m_i}$, and $h : \mathbb{R}^l \to \mathbb{R}$ is a twice continuously differentiable convex function. The problem can be converted to

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

In the tests, we set $h(u) := 1/3\|u\|_3^3$ with $\|\cdot\|_3$ denoting the 3-norm, and generated each m_i randomly from $\{2, 3, \ldots, 10\}$. All A_i were chosen as sparse matrices with approximately $10\% \cdot m_i \cdot d$ uniformly distributed non-zero entries, and all entries of b_i were chosen from the uniform distribution in $[-1, 0]$. For each (l, r, k), we generated 10 test instances, and then solved the SOCCP (1) derived from the KKT conditions of each problem with the LS semi-smooth method and the FB semi-smooth method. The mappings F and G in (5) were evaluated in the same way as above. The first inequality in (45) was replaced by

$$\max\{|F(\zeta^k)^{\mathrm{T}} G(\zeta^k)|, \Psi(\zeta^k)\} \leq 10^{-8}.$$

The numerical results were listed in Table 2, in which the second column gives the average dimension (m, n) of A for 10 problems, **Gap** denotes the average value of $|F(\zeta^k)^{\mathrm{T}}\mathrm{G}(\zeta^k)|$ at the final iteration, **NF** means the average function evaluations for solving each instance, and **Iter** denotes the average number of iterations for each instance to satisfy the termination conditions, and **Time** records the average CPU time in seconds 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 semi-smooth method does not have a remarkable superiority to the FB semi-smooth method. Among eight groups of test instances, the average number of iterations and the average number of function evaluations required by the LS semi-smooth method are basically same as that of the FB semi-smooth method, but the FB semi-smooth 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 semi-smooth method is superior to the FB semi-smooth method only for those difficult problems.

Table 2. Numerical results for the non-linear convex SOCPs.

	Dim.	LS semi-smooth method				FB semi-smooth 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

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 non-smooth system reformulation (8). Based on the reformulation, a semi-smooth 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 in 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 semi-smoothness of the FB function over general symmetric cones is still an open problem, now the method of this paper cannot be extended to general symmetric cone complementarity problems.

Acknowledgements

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

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

Note

1. $M_1, M_2 \in \mathbb{R}^{n\times n}$ are column monotone if, for any $u, v \in \mathbb{R}^n$, $M_1u + M_2v = 0 \Rightarrow u^{\mathrm{T}}v = 0$.

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Math. Program. 95 (2003), pp. 3–51.
[2] E.D. Andersen, C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Math. Program. 95 (2003), pp. 249–277.
[3] J.-S. Chen, *Conditions for error bounds and bounded level sets of some merit functions for the SOCCP*, J. Optim. Theory Appl. 135 (2007), pp. 459–473.
[4] X. Chen and H. Qi, *Cartesian P-proeprty and its applications to the semidefinite linear complementarity problem*, Math. Program. 106 (2006), pp. 177–201.
[5] J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Math. Program. 104 (2005), pp. 293–327.

[6] X.-D. Chen, D. Sun, and J. Sun, *Complementarity functions and numerical experiments for second-order cone complementarity problems*, Comput. Optim. Appl. 25 (2003), pp. 39–56.
[7] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).
[8] T. De Luca, F. Facchinei, and C. Kanzow, *A semi-smooth equation approach to the solution of nonlinear complementarity problems*, Math. Program. 75 (1996), pp. 407–439.
[9] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. I and II, Springer-Verlag, New York, 2003.
[10] J.Y. Fan and Y.X. Yuan, *On the quadratic convergence of the Levenberg–Marquardt method without nonsingularity assumption*, Computing 74 (2005), pp. 23–39.
[11] J. Faraut and A. Koranyi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.
[12] M.C. Ferris, O.L. Mangasarian, and J.-S. Pang (eds), *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic Publishers, Dordrecht, 2001.
[13] A. Fischer, *Solution of the monotone complementarity problem with locally Lipschitzian functions*, Math. Program. 76 (1997), pp. 513–532.
[14] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM J. Optim. 12 (2002), pp. 436–460.
[15] L. Grippo, F. Lampariello, and S. Lucidi, *A nonmonotone line search technique for Newton's method*, SIAM J. Numer. Anal. 23 (1986), pp. 707–716.
[16] S. Hayashi, N. Yamashita, and M. Fukushima, *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM J. Optim. 15 (2005), pp. 593–615.
[17] Y. Kanno, J.A.C. Martins, and A. Pinto da Costa, *Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem*, Intern. J. Numer. Methods Eng. 65 (2006), pp. 62–83.
[18] C. Kanzow and S. Petra, *On a semi-smooth least squares formulation of complementarity problems with gap reduction*, Optim. Methods Softw. 19 (2004), pp. 507–525.
[19] C. Kanzow, I. Ferenczi, and M. Fukushima, *Semi-smooth methods for linear and nonlinear second-order cone programs*, Tech. Rep. Department of Applied Mathematics and Physics, Kyoto University, 2006.
[20] M.S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Application of second-order cone programming*, Linear Algebra Appl. 284 (1998), pp. 193–228.
[21] R.D.C. Monteiro and T. Tsuchiya, *Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions*, Math. Program. 88 (2000), pp. 61–83.
[22] S.-H. Pan and J.-S. Chen, *A damped Guass-Newton method for the second-order conecomplementarity problem*, Appl. Math. Optim. 59 (2009), pp. 293–318.
[23] G. Pataki and S. Schmieta, *The DIMACS library of semidefinite-quadratic-linear programs*, Preliminary draft, Computational Optimization Research Center, Columbia University (1999). Available at, http://dimacs.rutgers.edu/Challenges/Seventh Instances/.
[24] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res. 18 (1993), pp. 227–244.
[25] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Math. Program. 58 (1993), pp. 353–367.
[26] J.F. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optim. Methods Softw. 11&12 (1999), pp. 625–653.
[27] D. Sun and J. Sun, *Strong semi-smoothness of the Fischer–Burmeister SDC and SOC complmentarity functions*, Math. Program. 103 (2005), pp. 575–581.
[28] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optim. Methods Softw. 11 (1999), pp. 141–182.
[29] N. Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg–Marquardt method*, Computing 15 (2001), pp. 237–249.
[30] A. Yoshise, *Interior point trajectories and a homogenous model for nonlinear complementarity problems over symmetric cones*, SIAM J. Optim. 17 (2006), pp. 1129–1153.