A Regularized Smoothing Newton Method for Symmetric Cone Complementarity Problems*

Lingchen Kong[1], Jie Sun[2] and Naihua Xiu[3]

(December 5, 2006; Revised August 28, 2007)

Abstract

This paper extends the regularized smoothing Newton method in vector optimization to symmetric cone optimization, which provide a unified framework for dealing with the nonlinear complementarity problem, the second-order cone complementarity problem, and the semidefinite complementarity problem (SCCP). In particular, we study strong semismoothness and Jacobian nonsingularity of the total natural residual function and show that the algorithm of Hayashi, Yamashita and Fukushima [SIAM J. Optim., 15 (2005), pp. 593-615] for the monotone SOCCP can be extended to the monotone SCCP.

Keywords: Symmetric cone complementarity problem; Monotonicity; Natural residual function; Regularized smoothing method; Quadratic convergence.

AMS Subject Classification: 65K05, 90C33

Abbreviated Title: Regularized Smoothing Method for SCCP

1 Introduction

We are interested in the following symmetric cone complementarity problem (SCCP): Find vectors $x, y \in \mathcal{J}$ such that

$$x \in K, \quad y = F(x) \in K, \quad \langle x, y \rangle = 0, \tag{1.1}$$

where $\mathcal{J}$ is an n-dimensional real Euclidean space, $\mathcal{A} := (\mathcal{J}, \langle \cdot, \cdot \rangle, \circ)$ is a Euclidean Jordan algebra, K is a symmetric cone in $\mathcal{A}$ (see Section 2), and $F : \mathcal{J} \to \mathcal{J}$ is a continuously differentiable mapping. Problem (1.1) includes the semidefinite complementarity problem (SDCP), the second-order cone complementarity problem (SOCCP), and the nonlinear complementarity problem (NCP) as special cases. The SCCP has wide applications, in particular, it may arise from the Karush-Kuhn-Tucker (KKT) system of a nonlinear optimization problem. The SCCP has been the focus of several recent studies, see, e.g., [12, 13, 14, 22, 23, 31, 35].

we intend to design an algorithm for SCCPs, which is called the regularized smoothing Newton method. In the setting of NCP, various regularized smoothing methods have been

*The work was partly supported by the National Natural Science Foundation of China (70471002, 10671010), the Singapore-MIT Alliance, and Grant RP314000-057-112 of National University of Singapore.

[1] Lingchen Kong, Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail:konglchen@126.com).

[2] Jie Sun, Department of Decision Sciences and Singapore-MIT Alliance, National University of Singapore, Republic of Singapore (e-mail: jsun@nus.edu.sg).

[3] Naihua Xiu, Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail:nhxiu@center.njtu.edu.cn).

tested, which, in addition to their simplicity of implementation, have the advantage of being able to solve some ill-posed problems. Recently, there are some papers studying the smoothing Newton methods with or without regularization for SOCCP and SDCP, see, e.g., [3, 4, 5, 6, 7, 11, 15, 16, 30, 33]. These papers either address the case of SOCCP or that of SDCP, but to our knowledge, there is no discussions on this type of algorithms under the general framework of SCCP.

In this paper, with the help of the Euclidean Jordan algebra we analyze strong semismoothness and Jacobian nonsingularity of the total natural residual function (the total NR-function) for SCCP. We construct the Chen-Mangasarian smoothing function for the natural residual, which provides a unified and computable formula in the more general setting. Moreover, we study the uniform approximation property and Jacobian consistency of this smoothing function. These nice properties play an important role in establishing quadratic convergence of this algorithm. Finally, by using a regularized smoothing technique we extend the globally and quadratically convergent algorithm proposed by Hayashi, Yamashita and Fukushima [15] for solving the monotone SOCCP to SCCP. The major analytic tools we used in this paper are taken from the recent work by Sun and Sun [29] in which differentiability and semismoothness of Löwner operator and spectral function are studied under the framework of Euclidean Jordan algebras.

This paper is organized as follows. In Section 2, we briefly describe Euclidean Jordan algebra and some of its properties used in our analysis. We also derive new results on the Jacobian and Clarke generalized Jacobian of Löwner operator. In Section 3, we introduce and characterize the total NR-function for SCCP. In Section 4, we present the Chen-Mangasarian smoothing function in the context of SCCPs and discuss its properties. In Section 5, we introduce the regularized smoothing Newton method for SCCP and discuss its convergence.

The following notations will be used throughout this paper. Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite dimensional real Euclidean spaces. For a given set $S \subseteq \mathcal{X}$, int$S$ and conv$S$ denote the interior and convex hull of S, respectively. Let dist(a, S) be $\min\{\|a-b\| : b \in S\}$ for $a \in \mathcal{X}$, where $\|\cdot\|$ is the norm on $\mathcal{X}$ induced by the inner product $\langle\cdot,\cdot\rangle$. We write $x \succeq_K y$ (respectively, $x \succ_K y$) to mean $x - y \in K$ (respectively, $x - y \in \text{int}K$) for vectors $x, y \in \mathcal{J}$. Also, we write $A \succeq B$ ($A \succ B$) to mean $A - B$ being positive semidefinite (positive definite) for operators A and B from $\mathcal{J}$ into itself. Let I be the identity operator, i.e., $Ix = x$ for all $x \in \mathcal{J}$. We say that the operator A is invertible (or nonsingular) if the equation $Ax = 0$ has a unique solution $x = 0$. For an operator A, A^T denotes the adjoint operator of A.

2 Preliminaries

2.1 Euclidean Jordan algebras

We give a brief introduction to Euclidean Jordan algebras. Details on Euclidean Jordan algebras can be found in Koecher's lecture note [19] and the monograph by Faraut and Korányi [10].

A *Euclidean Jordan algebra (EJA)* is a triple $(\mathcal{J}, \langle\cdot,\cdot\rangle, \circ) \stackrel{\Delta}{=} \mathcal{A}$ for short), where $(\mathcal{J}, \langle\cdot,\cdot\rangle)$ is a real n-dimensional inner product space and $(x, y) \mapsto x \circ y : \mathcal{J} \times \mathcal{J} \to \mathcal{J}$ is a bilinear mapping which satisfies the following conditions:

(i) $x \circ y = y \circ x$ for all $x, y \in \mathcal{J}$,
(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathcal{J}$ where $x^2 := x \circ x$,
(iii) $\langle x \circ y, z\rangle = \langle x, y \circ z\rangle$ for all $x, y, z \in \mathcal{J}$.

We call $x \circ y$ the *Jordan product* of x and y. In general, the Jordan product is not associative, i.e., $(x \circ y) \circ z \neq x \circ (y \circ z)$ for all $x, y, z \in \mathcal{J}$. In addition, we assume that there exists an element e (called the *identity* element) such that $x \circ e = e \circ x = x$ for all $x \in \mathcal{J}$.

- Given a Euclidean Jordan algebra $\mathcal{A}$, define *the set of squares* as $K := \{x^2 : x \in \mathcal{J}\}$. From Theorem III 2.1 in [10], K is a *symmetric cone* in $\mathcal{A}$. In other words, K is a self-dual closed convex cone and for any two elements $x, y \in \text{int}K$, there exists an invertible linear transformation $\Gamma : \mathcal{J} \to \mathcal{J}$ such that $\Gamma(K) = K$ and $\Gamma(x) = y$.
- For $x \in \mathcal{J}$, let $m := m(x)$ be the smallest positive integer such that the set $\{e, x, x^2, \cdots, x^m\}$ is linearly dependent. Then m is said to be the *degree* of x, which is denoted by $\deg(x)$.
- The *rank* of $\mathcal{A}$ denoted by $\text{rk}(\mathcal{A})$ is defined as $\text{rk}(\mathcal{A}) \triangleq \max\{\deg(x) : x \in \mathcal{J}\}$. Let $\dim(\mathcal{J})$ denote the dimension of $\mathcal{J}$. Obviously, $\text{rk}(\mathcal{A}) \leq \dim(\mathcal{J})$.
- An element $c \in \mathcal{J}$ is an *idempotent* if $c^2 = c \neq 0$. An idempotent element is *primitive* if it cannot be written as a sum of two idempotents.
- A *complete system of orthogonal idempotents* in $\mathcal{A}$ is a finite set $\{c_1, c_2, \cdots, c_k\}$ of idempotents where $c_i \circ c_j = 0$ for all $i \neq j$, and $c_1 + c_2 + \cdots + c_k = e$.
- A *Jordan frame* in $\mathcal{A}$ is a complete system of orthogonal primitive idempotents. The number of elements of any Jordan frame equals the positive integer $\text{rk}(\mathcal{A})$.

Example 2.1 Let $\mathbb{R}^n$ denote the space of n-dimensional real column vectors, $\mathbb{R}^n_+$ be the nonnegative orthant. Consider $\mathbb{R}^n$ with the (usual) inner product and Jordan product defined respectively by $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$ and $x \circ y := x * y$, where x_i denotes the i-th component of x etc., and $x * y$ denotes the componentwise product of vectors x and y. Then $(\mathbb{R}^n, \langle\cdot,\cdot\rangle, *)$ forms a Euclidean Jordan algebra with $\text{rk}((\mathbb{R}^n, \langle\cdot,\cdot\rangle, *)) = \dim(\mathbb{R}^n) = n$ and $\mathbb{R}^n_+$ as its cone of squares. The identity element is the n-vector of ones, and the set $\{e_1, e_2, \cdots, e_n\}$ is the unique Jordan frame where e_i is the ith coordinate vector for $i \in \{1, 2, \cdots, n\}$.

Example 2.2 Consider $\mathbb{R}^n (n \geq 2)$ where any x is written as $x = (x_1, x_2^T)^T$ with $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}^{n-1}$. The inner product is the same as usual and the Jordan product is defined by

$$x \circ y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \circ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} := \begin{pmatrix} \langle x, y\rangle \\ x_1 y_2 + y_1 x_2 \end{pmatrix}.$$

Then $\Lambda^n := (\mathbb{R}^n, \langle\cdot,\cdot\rangle, \circ)$ forms a Euclidean Jordan algebra, and its cone of squares (*Lorentz cone or second-order cone*) is specified by $\Lambda^n_+ := \{(x_1, x_2^T)^T : \ x_1 \geq \|x_2\| \ \}$, where $\|\cdot\|$ denotes the 2-norm. The identity element in Λ^n is $e = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. The set $\{c_1, c_2\}$ is a Jordan frame given by $c_i = \frac{1}{2}\begin{pmatrix} 1 \\ (-1)^i\ \omega \end{pmatrix}$ for $i = 1, 2$ with any $\omega \in \mathbb{R}^{n-1}$ satisfying $\|\omega\| = 1$.

Example 2.3 Let $\mathbb{S}^n$ denote the set of all $n \times n$ real symmetric matrices with the inner product and Jordan product defined respectively by $\langle X, Y\rangle := \text{Trace}(XY)$ and $X \circ Y := (XY + YX)/2$. Thus $(\mathbb{S}^n, \langle\cdot,\cdot\rangle, \circ)$ forms a Euclidean Jordan algebra, and its cone of squares $\mathbb{S}^n_+$ is the set of all positive semidefinite symmetric matrices. The identity element in this setting is the identity matrix E. The set $\{E_1, E_2, \cdots, E_n\}$ is a Jordan frame where E_i is the diagonal matrix with one in the ii-entry and zeros elsewhere for $i \in \{1, 2, \cdots, n\}$.

We now review the following spectral decomposition theorem of an element in a Euclidean Jordan algebra.

Theorem 2.1 *(Spectral Decomposition Type II (Theorem III.1.2, [10])) Let $\mathcal{A}$ be a Euclidean Jordan algebra with rank r. Then for $x \in \mathcal{J}$ there exist a Jordan frame $\{c_1, c_2, \cdots, c_r\}$ and real numbers $\lambda_1(x), \lambda_2(x), \cdots, \lambda_r(x)$ such that*

$$x = \lambda_1(x)c_1 + \lambda_2(x)c_2 + \cdots + \lambda_r(x)c_r. \tag{2.1}$$

The numbers $\lambda_i(x)$ $(i = 1, 2, \cdots, r)$ are the eigenvalues of x. We call (2.1) the spectral decomposition (or the spectral expansion) of x.

Note that the Jordan frame $\{c_1, c_2, \cdots, c_r\}$ in (2.1) depends on x. We do not write this dependence explicitly for simplicity in notation. Let $\sigma(x)$ be the set of all eigenvalues of x. Then $\sigma(x)$ contains at least one element and at most r. For each $\mu_i(x) \in \sigma(x)$, denote $N_i(x) := \{j : \lambda_j(x) = \mu_i(x)\}$ and $b_i \triangleq \sum_{j\in N_i(x)} c_j$. Then the set $\{b_i : \mu_i(x) \in \sigma(x)\}$ is a complete system of orthogonal idempotents, and its uniqueness is guaranteed by Theorem III.1.1 in [10]. Let $\bar{r}$ be the number of elements in this set. We then have the spectral decomposition of type I stated in [10], i.e.,

$$x = \mu_1(x)b_1 + \mu_2(x)b_2 + \cdots + \mu_{\bar{r}}(x)b_{\bar{r}}.$$

Next, we recall the Peirce decomposition theorem on the space $\mathcal{J}$, where the Jordan frame $\{c_1, c_2, \cdots, c_r\}$ can be fixed beforehand. In this case, define the following subspaces

$$J_{ii} \triangleq \{x \in \mathcal{J} : x \circ c_i = x\} \text{ and } J_{ij} \triangleq \{x \in \mathcal{J} : x \circ c_i = \frac{1}{2}x = x \circ c_j\},\ i \neq j, \tag{2.2}$$

for $i, j \in \{1, 2, \cdots, r\}$. In the SOC case, we have $J_{12} \triangleq \{x \in \mathbb{R}^n : x_1 = 0, \langle x_2, w\rangle = 0\}$, where w is characterized by the Jordan frame as in Example 2.2.

Theorem 2.2 *(Theorem IV.2.1 in [10]) Let $\{c_1, c_2, \cdots, c_r\}$ be a given Jordan frame in a Euclidean Jordan algebra $\mathcal{A}$ of rank r. Then $\mathcal{J}$ is the orthogonal direct sum of spaces J_{ij} $(i \leq j)$. Furthermore,*

(i) $J_{ij} \circ J_{ij} \subseteq J_{ii} + J_{jj}$;

(ii) $J_{ij} \circ J_{jk} \subseteq J_{ik}$, *if* $i \neq k$;

(iii) $J_{ij} \circ J_{kl} = \{0\}$, *if* $\{i, j\} \bigcap \{k, l\} = \emptyset$.

For each $x \in \mathcal{J}$, we define the *Lyapunov transformation* $L(x) : \mathcal{J} \to \mathcal{J}$ by $L(x)y = x \circ y$ for all $y \in \mathcal{J}$, which is a symmetric operator in the sense that $\langle L(x)y, z\rangle = \langle y, L(x)z\rangle$ for all $y, z \in \mathcal{J}$. Meanwhile, the operator $Q(x) \triangleq 2L^2(x) - L(x^2)$ is called the *quadratic representation* of x. We say two elements $x, y \in \mathcal{J}$ *operator commute* if $L(x)L(y) = L(y)L(x)$. Lemma X.2.2 in [10] gives the following characterization of operator commutativity.

Theorem 2.3 *Two elements x, y of a Euclidean Jordan algebra of rank r operator commute if and only if they share a common Jordan frame.*

Thus, for a given Jordan frame $\{c_1, c_2, \cdots, c_r\}$, it is easy to see that c_i, c_j operator commute and $L(c_i)L(c_j) = L(c_j)L(c_i)$ for any $i, j \in \{1, 2, \cdots, r\}$. So do $b_i(x)$ and $b_j(x)$ for any $i, j \in \{1, 2, \cdots, \bar{r}\}$ in view of the argument after Theorem 2.1.

2.2 Jacobian of Löwner operator

We review differentiability and semismoothness of a vector-valued function which was called the Löwner operator by Sun and Sun [29] in recognition of Löwner's contribution [21]. We also present some new results on the Jacobian and the Clarke generalized Jacobian of the Löwner operator, which are basic and useful in the subsequent analysis.

Definition 2.4 *Let $x = \sum_{j=1}^{r} \lambda_j(x)c_j$ and $g : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We define the Löwner operator (function) $G : \mathcal{J} \to \mathcal{J}$ as*

$$G(x) \triangleq \sum_{j=1}^{r} g(\lambda_j(x))c_j = g(\lambda_1(x))c_1 + g(\lambda_2(x))c_2 + \cdots + g(\lambda_r(x))c_r. \tag{2.3}$$

When $g(t) = t_+ = \max\{0, t\}(t \in \mathbb{R})$, this becomes the *metric projection operator*

$$P_K(x) \triangleq (\lambda_1(x))_+c_1 + (\lambda_2(x))_+c_2 + \cdots + (\lambda_r(x))_+c_r \tag{2.4}$$

onto the symmetric cone K. Note that $x \in K$ (respectively, $x \in \mathrm{int}K$) if and only if $\lambda_i(x) \geq 0$ (respectively, $\lambda_i(x) > 0$) $(i = 1, 2, \cdots, r)$. For any $x \in K$, we define $\sqrt{x} \triangleq \sum_{j=1}^{r} \sqrt{\lambda_j(x)}c_j$ for $x \in K$.

Consider the differentiability of the Löwner operator $G(\cdot)$. Suppose that g is differentiable at $\tau_i, i = 1, 2, \cdots, r$. Define *the first divided difference* $g^{[1]}$ of g at $\tau \triangleq (\tau_1, \tau_2, \cdots, \tau_r)^T \in \mathbb{R}^r$ as the $r \times r$ symmetric matrix with the ij-th entry being $(g^{[1]}(\tau))_{ij}$, given by

$$[\tau_i, \tau_j]_g \triangleq \begin{cases} \frac{g(\tau_i)-g(\tau_j)}{\tau_i-\tau_j} & \text{if } \tau_i \neq \tau_j, \\ g'(\tau_i) & \text{if } \tau_i = \tau_j, \end{cases} \quad i, j = 1, 2, \cdots, r. \tag{2.5}$$

Based on Theorem 13 in [29], the following Jacobian properties of Löwner operator $G(\cdot)$ is obvious.

Theorem 2.5 *Let $x = \sum_{j=1}^{r} \lambda_j(x)c_j = \sum_{i=1}^{\bar{r}} \mu_i(x)b_i(x)$. Then, $G(\cdot)$ is (continuously) differentiable at x if and only if for each $j \in \{1, 2, \cdots, r\}$, g is (continuously) differentiable at $\lambda_j(x)$. In this case, the Jacobian $\nabla G(x)$ is given by*

$$\nabla G(x) = 2 \sum_{i \neq j, i,j=1}^{r} [\lambda_i(x), \lambda_j(x)]_g L(c_i)L(c_j) + \sum_{i=1}^{r} g'(\lambda_i(x))Q(c_i) \tag{2.6}$$

or equivalently

$$\nabla G(x) = 2 \sum_{i \neq j,\ i,j=1}^{\bar{r}} [\mu_i(x), \mu_j(x)]_g L(b_i(x))L(b_j(x)) + \sum_{i=1}^{\bar{r}} g'(\mu_i(x))Q(b_i(x)). \tag{2.7}$$

Furthermore, $\nabla G(x)$ is a linear and symmetric operator from $\mathcal{J}$ into itself.

As a consequence of Theorem 2.5, we have the following result in the case of rk($\mathcal{A}$)=dim($\mathcal{J}$).

Corollary 2.6 *Suppose that rk($\mathcal{A}$)=dim($\mathcal{J}$) $= n$ and $x = \sum_{j=1}^{n} \lambda_j(x)c_j = \sum_{i=1}^{\bar{n}} \mu_i(x)b_i(x)$. If $G(\cdot)$ is (continuously) differentiable at x, then it holds that*

$$\nabla G(x) = \sum_{i=1}^{n} g'(\lambda_i(x))L(c_i) = \sum_{i=1}^{\bar{n}} g'(\mu_i(x))L(b_i(x)). \tag{2.8}$$

Proof. Since rk($\mathcal{A}$)=dim($\mathcal{J}$) $= n$, it follows from Theorem 3.5 in [20] that there is a unique Jordan frame $\{c_1, c_2, \cdots, c_n\}$ in $\mathcal{A}$. Thus, through Theorem 2.2, any element $h \in \mathcal{J}$ can be expressed by $h = \sum_{i=1}^n h_i c_i$ with $h_i \in \mathbb{R}$ $(i = 1, 2, \cdots, n)$. Therefore,

$$L(c_i)L(c_j)h = L(c_j)L(c_i)h = c_i \circ (c_j \circ h) = \begin{cases} c_i \circ (h_j c_j) = 0 & \text{if } \ i \neq j, \\ c_i \circ (h_i c_i) = c_i \circ h & \text{if } \ i = j. \end{cases}$$

This implies that $L(c_i)L(c_j) = 0$ $(i \neq j)$ and $L(c_i)L(c_i) = L(c_i)$ for any $i, j \in \{1, 2, \cdots, n\}$. Hence $Q(c_i) = L(c_i)$. Formula (2.8) is then an implication of Theorem 2.5. □

As an application of Corollary 2.6, we obtain the Jacobian of the Löwner operator on $\mathbb{R}^n$.

Example 2.4 Suppose that $\mathcal{A} = (\mathbb{R}^n, \langle \cdot, \cdot \rangle, *)$ as in Example 2.1. Let $x = \sum_{i=1}^n x_i e_i$. One can easily verify that $L(e_i) = e_i e_i^T = E_i$ $(i = 1, 2, \cdots, n)$. Note that rk$((\mathbb{R}^n, \langle \cdot, \cdot \rangle, *))$=dim$(\mathbb{R}^n) = n$. It is obvious via Corollary 2.6 that

$$\nabla G(x) = \sum_{i=1}^n g'(x_i) L(e_i) = \text{Diag}\{g'(x_1), g'(x_2), \cdots, g'(x_n)\}.$$

The next theorem gives a sufficient condition which guarantees that the Jacobian $\nabla G(x)$ is positive semidefinite (respectively, positive definite). Here $\nabla G(x)$ is called *positive semidefinite* (*respectively, positive definite*) if $\langle h, \nabla G(x) h \rangle \geq 0$ for all $h \in \mathcal{J}$ (respectively, $\langle h, \nabla G(x) h \rangle > 0$ for all $0 \neq h \in \mathcal{J}$).

Theorem 2.7 *Let $x = \sum_{j=1}^r \lambda_j(x) c_j$. If g is (continuously) differentiable at $\lambda_j(x)$ for each $j \in \{1, 2, \cdots, r\}$ and $g'(t) \geq 0$ for all $t \in \mathbb{R}$, then $G(\cdot)$ is (continuously) differentiable at x and the Jacobian $\nabla G(x)$ is positive semidefinite. Moreover, the Jacobian is positive definite if the condition $g'(t) \geq 0$ is replaced by $g'(t) \geq 0$.*

Proof. Let $x = \sum_{j=1}^r \lambda_j(x) c_j$. For the Jordan frame $\{c_1, c_2, \cdot, c_r\}$, by Theorem 2.2, any element $h \in \mathcal{J}$ can be expressed by $h = \sum_{i=1}^r h_i c_i + \sum_{1 \leq k < l \leq r} h_{kl}$ where $h_i \in \mathbb{R}$ $(i = 1, 2, \cdots, r)$ and $h_{kl} \in J_{kl}$ $(1 \leq k < l \leq r)$. Theorem 2.2 also implies that $c_j \circ \sum_{i=1}^r h_i c_i = h_j c_j$ and $c_j \circ \sum_{1 \leq k < l \leq r} h_{kl} = \frac{1}{2}(\sum_{k<j} h_{kj} + \sum_{l>j} h_{jl})$. It therefore holds that

$$c_j \circ h = h_j c_j + \frac{1}{2}(\sum_{k=1}^{j-1} h_{kj} + \sum_{l=j+1}^{r} h_{jl}) \tag{2.9}$$

where $\sum_{k=1}^{j-1} h_{kj} \triangleq 0$ if $j = 1$ and $\sum_{l=j+1}^r h_{jl} \triangleq 0$ if $j = r$. Furthermore, Theorem 2.2 derives

$$\langle h, c_j \circ (c_i \circ h) \rangle = \langle c_j \circ h, c_i \circ h \rangle = \frac{1}{4} \langle h_{ji}, h_{ji} \rangle = \frac{1}{4} \|h_{ji}\|^2, \ j \neq i, \tag{2.10}$$

and

$$Q(c_j)h = 2c_j \circ (c_j \circ h) - c_j \circ h = h_j c_j, \qquad j = 1, 2, \cdots, r. \tag{2.11}$$

Meanwhile, noting that $c_j^2 = c_j$, one has $\langle h, c_j \rangle = \langle c_j \circ h, c_j \rangle = h_j \langle c_j, c_j \rangle = h_j \|c_j\|^2$. Combining this with (2.6), (2.10), (2.11) and $L(c_j)L(c_i)h = c_j \circ (c_i \circ h)$, one has

$$\langle h, \nabla G(x) h \rangle \;\; = \;\; \langle h, \sum_{j \neq i, j, i=1}^{r} 2(g^{[1]}(\lambda(x)))_{ji} c_j \circ (c_i \circ h) + \sum_{j=1}^{r} (g^{[1]}(\lambda(x)))_{jj} h_j c_j \rangle$$

$$= \sum_{j\neq i,j,i=1}^{r} 2(g^{[1]}(\lambda(x)))_{ji}\langle h, c_j \circ (c_i \circ h)\rangle + \sum_{j=1}^{r}(g^{[1]}(\lambda(x)))_{jj} h_j \langle h, c_j\rangle$$

$$= \frac{1}{2}\sum_{j\neq i,j,i=1}^{r} (g^{[1]}(\lambda(x)))_{ji}\|h_{ji}\|^2 + \sum_{j=1}^{r}(g^{[1]}(\lambda(x)))_{jj} h_j^2 \|c_j\|^2.$$

If $g'(t) \geq 0$ ($g'(t) > 0$) for all $t \in \mathbb{R}$, then by (2.5) we can easily get $(g^{[1]}(\lambda(x)))_{ji} \geq 0$ ($(g^{[1]}(\lambda(x)))_{ji} > 0$) for all $j \neq i, j, i = 1, 2, \cdots, r$. Hence, $\langle h, \nabla G(x)h\rangle \geq 0$ for all $h \in \mathcal{J}$ ($\langle h, \nabla G(x)h\rangle > 0$ for all $0 \neq h \in \mathcal{J}$) through the above equation. □

We proceed with (strong) semismoothness of Löwner operator $G(\cdot)$. Semismoothness was originally introduced by Mifflin [24] for functionals. Qi and Sun [26] extended the concept of semismoothness to vector-valued functions and developed a systematic theory that employs semismoothness in the analysis of the superlinear convergence of Newton methods for solving systems of nondifferentiable equations.

We now briefly review some concepts and results of the semismoothness from [26]. Let $\digamma : C \subseteq \mathcal{X} \to \mathcal{Y}$ be a locally Lipschitz function on an open set C. By Rademacher's theorem, $\digamma$ is almost everywhere differentiable (in the sense of Fréchet) in C. Suppose $D_\digamma$ is the set of points where $\digamma$ is differentiable. Let $\digamma'(x)$, which is a linear mapping from $\mathcal{X}$ to $\mathcal{Y}$, denote the *derivative* of $\digamma$ at x if $\digamma$ is differentiable at x, and $\nabla \digamma(x)$ denote the *Jacobian* of $\digamma$ at x specified by $\nabla \digamma(x) \triangleq \digamma'(x)^T$, in the sense of $\langle y, \nabla \digamma(x)z\rangle = \langle \digamma'(x)y, z\rangle$ for all $y \in \mathcal{X}$ and $z \in \mathcal{Y}$. Then, the *Clarke generalized Jacobian* of $\digamma$ at x is defined by $\partial \digamma(x) \triangleq \text{conv}\{\partial_B \digamma(x)\}$, where $\partial_B \digamma(x) \triangleq \{\lim_{\bar{x}\to x, \bar{x}\in D_\digamma} \nabla \digamma(\bar{x})\}$. Observe that $\partial \digamma(x) = \{\nabla \digamma(x)\}$ if $\digamma$ is smooth (continuously differentiable) at x. We say $\digamma$ is *directionally differentiable* at x along the direction d if

$$\digamma'(x, d) \triangleq \lim_{t\downarrow 0} \frac{\digamma(x + td) - \digamma(x)}{t} \quad \text{exists,}$$

where $\digamma'(x, d)$ is called the *directional derivative* of $\digamma$ at x along the direction d; and $\digamma$ is directionally differentiable at x if $\digamma$ is directionally differentiable at x along any direction $d \neq 0$.

Employing the above concepts, we can define (strong) semismoothness of a function $\digamma$.

Definition 2.8 *A directionally differentiable and locally Lipschitz function $\digamma : C \subseteq \mathcal{X} \to \mathcal{Y}$ is semismooth at $x \in C$ if $V^T d - \digamma'(x; d) = o(\|d\|)$ for any $d \neq 0, d \in \mathcal{X}$ sufficiently small and $V \in \partial \digamma(x + d)$. In particular, if $o(\|d\|)$ can be replaced by $O(\|d\|^2)$, $\digamma$ is strongly semismooth.*

By combining Theorem 17 with Proposition 15 in [29], we have the following result on (strong) semismoothness of Löwner operator $G(\cdot)$.

Lemma 2.9 *Let $x = \sum_{j=1}^{r} \lambda_j(x)c_j$. Then $G(\cdot)$ is (strongly) semismooth at x if and only if for each $j \in \{1, 2, \cdots, r\}$, g is (strongly) semismooth at $\lambda_j(x)$. In particular, the metric projection operator P_K is strongly semismooth on $\mathcal{J}$.*

We are ready to extend Theorems 2.5 and 2.7 to the case of semismooth Löwner operator $G(\cdot)$. Let g be semismooth at τ_i ($i = 1, 2, \cdots, r$) and ∂g denote the subderivative of g in the sense of Clarke. Define *the first generalized divided difference* $g^{[1,\partial]}$ of g at τ as the set of all $r \times r$ symmetric matrices, where the ij-th entry $(g^{[1,\partial]}(\tau))_{ij}$ of the element $g^{[1,\partial]}(\tau) \in g^{[1,\partial]}$ is given by a set $\{[\tau_i, \tau_j]_g\}$ for $i, j = 1, 2, \cdots, r$, where

$$(2.5)' \qquad \{[\tau_i,\tau_j]_g\} = \begin{cases} \{\frac{g(\tau_i)-g(\tau_j)}{\tau_i-\tau_j}\} & \text{if } \tau_i \neq \tau_j, \\ \partial g(\tau_i) & \text{if } \tau_i = \tau_j. \end{cases}$$

Theorem 2.10 *Let $x \in \mathcal{J}$. Then $G(\cdot)$ is (strongly) semismooth at x if and only if g is (strongly) semismooth at every eigenvalue of x. In this case, the Clarke generalized Jacobian $\partial G(x)$ satisfies*

$$\overline{\partial} G(x) \supseteq \partial G(x) \supseteq \underline{\partial} G(x)$$

with the sets $\overline{\partial}G(x)$ and $\underline{\partial}G(x)$ being given respectively by

$$\overline{\partial}G(x) \triangleq \operatorname{conv}\left[\bigcup_{\{c_1,\cdots,c_r\}\in\mathcal{C}(x)} \partial_{c_1,\cdots,c_r}G(x)\right],$$

$$\underline{\partial}G(x) \triangleq \left\{2\sum_{i\neq j,\ i,j=1}^{\bar{r}} [\mu_i(x),\mu_j(x)]_g L(b_i(x))L(b_j(x)) + \sum_{i=1}^{\bar{r}} \partial g(\mu_i(x))Q(b_i(x))\right\},$$

where $\mathcal{C}(x)$ is the set consisting of all Jordan frames in the spectral decomposition type II of x, and $\partial_{c_1,\cdots,c_r}G(x) \triangleq \{2\sum_{i\neq j,\ i,j=1}^{r}\{[\lambda_i(x),\lambda_j(x)]_g\}L(c_i)L(c_j) + \sum_{i=1}^{r}\partial g(\lambda_i(x))Q(c_i)\}$.

Proof. We first show $\overline{\partial}G(x) \supseteq \partial G(x)$. By the definitions of $\overline{\partial}G$ and ∂G we need to only prove that $\bigcup_{\{c_1,\cdots,c_r\}\in\mathcal{C}(x)} \partial_{c_1,\cdots,c_r}G(x) \supseteq \partial_B G(x)$. Taking any $V \in \partial_B G(x)$, by the definition of $\partial_B G(x)$ there exists a vector $h \triangleq h(V) \in \mathcal{J}$ such that $V = \lim_{h\to 0, x+h\in D_G} \nabla G(x+h)$. In order to show $V \in \bigcup_{\{c_1,\cdots,c_r\}\in\mathcal{C}(x)} \partial_{c_1,\cdots,c_r}G(x)$, we proceed as follows.

Take any $\{c_1,\cdots,c_r\} \in \mathcal{C}(x)$ and let $x = \sum_{j=1}^{r}\lambda_j(x)c_j = \sum_{i=1}^{\bar{r}}\mu_i(x)b_i(x)$ with $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_r(x)$ and $\mu_1(x) > \mu_2(x) > \cdots > \mu_{\bar{r}}(x)$. For the above $h \in \mathcal{J}$, let $x+h \triangleq \sum_{j=1}^{r}\lambda_j(x+h)c_j(x+h)$ with $\lambda_1(x+h) \geq \lambda_2(x+h) \geq \cdots \geq \lambda_r(x+h)$. By Theorem 2.1 and the argument after it, in the sense of set convergence (see [27]), one has

$$\lim_{h\to 0, x+h\in D_G} \{\lambda(x+h)\} = \{\lambda(x)\}$$

where $\lambda(x+h) \triangleq (\lambda_1(x+h),\lambda_2(x+h),\cdots,\lambda_r(x+h))^T$ and $\lambda(x) \triangleq (\lambda_1(x),\lambda_2(x),\cdots,\lambda_r(x))^T$, and

$$(2.12) \qquad \limsup_{h\to 0, x+h\in D_G} \{(c_1(x+h),c_2(x+h),\cdots,c_r(x+h))\} \subseteq \mathcal{C}(x).$$

Notice that for any $i,j = 1,2,\cdots,r$,

$$\limsup_{h\to 0, x+h\in D_G} \{[\lambda_i(x+h),\lambda_j(x+h)]_g\} \begin{cases} = \{\frac{g(\lambda_i(x))-g(\lambda_j(x))}{\lambda_i(x)-\lambda_j(x)}\} & \text{if } \lambda_i(x) \neq \lambda_j(x), \\ \subseteq \partial g(\lambda_i(x)) & \text{if } \lambda_i(x) = \lambda_j(x). \end{cases}$$

Thus,

$$(2.13) \qquad \limsup_{h\to 0, x+h\in D_G} \{[\lambda_i(x+h),\lambda_j(x+h)]_g\} \subseteq \{[\lambda_i(x),\lambda_j(x)]_g\}.$$

Also, it holds by (2.6) that for $x+h \in D_G$

$$\begin{aligned}\nabla G(x+h) &= 2\sum_{i\neq j,\ i,j=1}^{r}[\lambda_i(x+h),\lambda_j(x+h)]_g L(c_i(x+h))L(c_j(x+h)) \\ &\quad + \sum_{i=1}^{r} g'(\lambda_i(x+h))Q(c_i(x+h)).\end{aligned}$$

This, together with (2.12), (2.13) and the continuity property of $L(x)$ and $Q(x)$, leads to

$$\begin{aligned}&\limsup_{h\to 0, x+h\in D_G}\{\nabla G(x+h)\} \\ \subseteq &\bigcup_{\{c_1,\cdots,c_r\}\in\mathcal{C}(x)}\left\{2\sum_{i\neq j,\ i,j=1}^{r}\{[\lambda_i(x),\lambda_j(x)]_g\}L(c_i)L(c_j)+\sum_{i=1}^{r}\partial g(\lambda_i(x))Q(c_i)\right\}.\end{aligned}$$

Clearly, $V=\lim_{h\to 0, x+h\in D_G}\nabla G(x+h)\in \limsup_{h\to 0, x+h\in D_G}\{\nabla G(x+h)\}$. This implies that $V\in\bigcup_{\{c_1,\cdots,c_r\}\in\mathcal{C}(x)}\partial_{c_1,\cdots,c_r}G(x)$ by the definition of $\partial_{c_1,\cdots,c_r}G(x)$.

We next prove $\partial G(x)\supseteq \underline{\partial} G(x)$. For any $W(x)\in\underline{\partial} G(x)$ with $x=\sum_{i=1}^{\bar r}\mu_i(x)b_i(x)$ and $\mu_1(x)>\mu_2(x)>\cdots>\mu_{\bar r}(x)$, by the definition of $\underline{\partial} G(x)$ we have

$$W(x)=2\sum_{i\neq j,\ i,j=1}^{\bar r}[\mu_i(x),\mu_j(x)]_g L(b_i(x))L(b_j(x))+\sum_{i=1}^{\bar r}w_iQ(b_i(x))$$

with $w_i\in\partial g(\mu_i(x))$ $(i=1,2,\cdots,\bar r)$. Since g is semismooth at $\mu_i(x)$ and $\dim(\partial g(\mu_i(x)))=1$, by Carathéodory Theorem (see [27]), for any given $w_i\in\partial g(\mu_i(x))$ there exist elements

(2.14) $$w_{i,l_i}\triangleq \lim_{h_{i,l_i}\to 0,\mu_i(x)+h_{i,l_i}\in D_g} g'(\mu_i(x)+h_{i,l_i})\in\partial g(\mu_i(x)),\quad l_i\in\{0,1\}$$

and $t_i\in[0,1]$ such that

(2.15) $$w_i=t_iw_{i,0}+(1-t_i)w_{i,1}$$

where D_g is the set consisting of differentiable points of g. Based on the set $\{h_{i,l_i}\ :\ l_i\in\{0,1\},\ i=1,2,\cdots,\bar r\}$, we construct a set $\mathcal{H}$ by

$$\mathcal{H}\triangleq\left\{\sum_{i=1}^{\bar r}h_{i,l_i}b_i(x)\ :\ l_i\in\{0,1\}\right\}.$$

Let $l\triangleq(l_1,l_2,\cdots,l_{\bar r})$ and $h_l\triangleq\sum_{i=1}^{\bar r}h_{i,l_i}b_i(x)$ with $l_i\in\{0,1\}$. Then the set $\mathcal{H}$ can be rewritten as $\mathcal{H}\triangleq\{h_l\ :\ l\in\{0,1\}^{\bar r}\}$, which includes $2^{\bar r}$ elements. Meanwhile, for each element h_l, we have

$$x+h_l=\sum_{i=1}^{\bar r}(\mu_i(x)+h_{i,l_i})b_i(x).$$

Moreover, taking sufficiently small $\|h_l\|$, we have $\mu_1(x)+h_{1,l_1}>\mu_2(x)+h_{2,l_2}>\cdots>\mu_{\bar r}(x)+h_{\bar r,l_{\bar r}}$, and hence $\mu_i(x+h_l)=\mu_i(x)+h_{i,l_i},\ b_i(x+h_l)=b_i(x)$ by the uniqueness of spectral decomposition type I. Thus, $x+h_l \in D_G$ by $\mu_i(x)+h_{i,l_i} \in D_g$, and from (2.7) and (2.14) we obtain

$$\begin{aligned} W_l(x) &\triangleq \lim_{h_l\to 0, x+h_l\in D_G} \nabla G(x+h_l) \\ &= 2\sum_{i\neq j,\ i,j=1}^{\bar r} [\mu_i(x),\mu_j(x)]_g L(b_i(x))L(b_j(x)) + \sum_{i=1}^{\bar r} w_{i,l_i} Q(b_i(x)). \end{aligned}$$

Therefore, $W_l(x) \in \partial G(x)$ for every $l \in \{0,1\}^{\bar r}$. This implies that

$$\mathcal{W}(x) \triangleq \mathrm{conv}\{W_l(x)\ :\ l\in\{0,1\}^{\bar r}\} \subseteq \partial G(x). \tag{2.16}$$

To prove $W(x) \in \partial G(x)$, it suffices to claim that $W(x) \in \mathcal{W}(x)$. In fact, from expressions of $W(x)$ and $W_l(x)$, it is easy to see that $w \triangleq (w_1, w_2, \cdots, w_{\bar r})$ given above lies in the hypercube whose extreme points are defined by w_{i,l_i} with $l_i \in \{0,1\}, i=1,2,\cdots,\bar r$. Hence, $W(x)$ must be a convex combination of points $\{W_l(x)\ :\ l\in\{0,1\}^{\bar r}\}$. The proof is completed. □

Remark 2.1 From Theorem 2.10, we easily observe that if $x \in \mathcal{J}$ has distinct eigenvalues $\lambda_1(x),\cdots,\lambda_r(x)$ and $\mathcal{C}(x)$ has an element, then $\underline{\partial} G(x) = \partial G(x) = \overline{\partial} G(x)$. However, if x has the multiple eigenvalues or $\mathcal{C}(x)$ contains many elements, the sets $\underline{\partial} G(x)$, $\partial G(x)$ and $\overline{\partial} G(x)$ may be different as the following example shows.

Let $\mathcal{A} = \Lambda^n$ $(n \geq 3)$ and $x = \sum_{i=1}^2 \lambda_i c_i$ as in Example 2.2. Take $G(x) = P_K(x)$ where $g(t) = t_+$, and let $x=0$. Then $\lambda_1 = \lambda_2 = 0$ and there are infinitely many Jordan frames at $x=0$. The direct calculation yields $\overline{\partial} P_{\Lambda^n_+}(0) = \mathrm{conv}\{4[0,1]L(c_1)L(c_2)+[0,1]Q(c_1)+[0,1]Q(c_2)\}$ and $\underline{\partial} P_{\Lambda^n_+}(0) = \mathrm{conv}\{0,E\}$ where $\mathrm{conv}\{0,E\} = \{\alpha E : 0 \leq \alpha \leq 1\}$. Note that $\partial P_{\Lambda^n_+}(0) = \mathrm{conv}\{0,E,S\}$ by Proposition 4.8 in [15] where S satisfies

$$S = 4\times \frac{1+\beta}{2} L(c_1)L(c_2) + 0\times Q(c_1) + Q(c_2),$$

where $\frac{1+\beta}{2} \in [0,1]$. A simple calculation checks that $\underline{\partial} P_{\Lambda^n_+}(0) \subset \partial P_{\Lambda^n_+}(0) \subset \overline{\partial} P_{\Lambda^n_+}(0)$. □

Remark 2.2 Suppose that rk($\mathcal{A}$)=dim($\mathcal{J}$) $= n$ and $x = \sum_{j=1}^n \lambda_j(x)c_j = \sum_{i=1}^{\bar n} \mu_i(x) b_i(x)$ as in the case of Corollary 2.6. If $G(\cdot)$ is (strong) semismooth at x, we derive by Theorem 2.10 that $\underline{\partial} G(x) \subseteq \partial G(x) \subseteq \overline{\partial} G(x)$, where

$$\overline{\partial} G(x) = \sum_{i=1}^{n} \partial g(\lambda_i(x)) L(c_i), \qquad \underline{\partial} G(x) = \sum_{i=1}^{\bar n} \partial g(\mu_i(x)) L(b_i(x)).$$

Especially, when $\mathcal{A} = (\mathbb{R}^n, \langle\cdot,\cdot\rangle, *)$ as in Example 2.4 and $x = \sum_{i=1}^n x_i e_i = \sum_{i=1}^{\bar n} y_i (\sum_{j\in N(i)} e_j)$, in the similar way of the second part in the proceeding proof, one has $\overline{\partial} G(x) \subseteq \partial G(x)$. Hence,

$$\overline{\partial} G(x) = \partial G(x) = \sum_{i=1}^{n} \partial g(x_i) E_i = \mathrm{Diag}\ \{\partial g(x_1),\cdots,\partial g(x_n)\},$$

$$\underline{\partial} G(x) = \sum_{i=1}^{\bar n} \partial g(y_i) (\sum_{j\in N(i)} E_j) = \mathrm{Diag}\ \{\partial g(y_1) I_1,\cdots,\partial g(y_{\bar n}) I_{\bar n}\},$$

where I_i is the $|N(i)|\times|N(i)|$ identity matrix for $i=1,2,\cdots,\bar{n}$. Moreover, letting $G(x)=P_K(x)$ and $x=0$, we derive

$$\overline{\partial}P_{\mathbb{R}^n_+}(0)=\partial P_{\mathbb{R}^n_+}(0)=\left\{\begin{pmatrix}[0,1] & 0 & \cdots & 0\\ 0 & [0,1] & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & [0,1]\end{pmatrix}\right\}\supset\{\alpha E, 0\le\alpha\le 1\}=\underline{\partial}P_{\mathbb{R}^n_+}(0).$$

□

Although Theorem 2.10 provides some approximations of the Clarke generalized Jacobian, it can be successfully employed to prove the positive semidefiniteness of $\partial G(\cdot)$.

Theorem 2.11 *Let $x\in\mathcal{J}$. If g is (strongly) semismooth at every eigenvalue of x and $\partial g(t)\subseteq\mathbb{R}_+$ ($\partial g(t)\subseteq\mathbb{R}_{++}$) for all $t\in\mathbb{R}$, then the function $G(\cdot)$ is (strongly) semismooth at x and the element $V\in\partial G(x)$ is positive semidefinite (positive definite). Moreover, when $\partial g(t)\subseteq\mathbb{R}_{++}$, there exists a scalar $\alpha(x)>0$ such that $V\succeq\alpha(x)I\succ 0$.*

Proof. By Theorem 2.10 and the definition of $\overline{\partial}G$, it suffices to demonstrate that if $\partial g(t)\subseteq\mathbb{R}_{++}$ for all $t\in\mathbb{R}$, then for any $\{c_1,\cdots,c_r\}\in\mathcal{C}(x)$ and $V\in\partial_{c_1,\cdots,c_r}G(x)$ there is a scalar $\alpha(x)$ such that $V\succeq\alpha(x)I\succ 0$. In this case, one has $x=\sum_{i=1}^r\lambda_i(x)c_i$ and

$$V=2\sum_{i\neq j,\ i,j=1}^{r}\nu_{ij}(x)L(c_i)L(c_j)+\sum_{i=1}^{r}\nu_{ii}(x)Q(c_i)$$

with $\nu_{ij}(x)\in\{[\lambda_i(x),\lambda_j(x)]_g\}$. Note that $\partial g(\lambda_j(x))\subseteq\mathbb{R}_{++}$ is a closed convex set for every $j=1,\cdots,r$. Taking

$$\alpha(x)\triangleq\min_{i,j}\{[\lambda_i(x),\lambda_j(x)]_g\},$$

by the definition (2.5)$'$ and the given assumptions we have $\alpha(x)>0$ and hence $\alpha(x)I\succ 0$.

We now prove $\bar{V}\triangleq V-\alpha(x)I\succeq 0$, that is, $\langle h,\bar{V}h\rangle\ge 0$ for any $h\in\mathcal{J}$. Note that

$$\bar{V}=2\sum_{i\neq j,\ i,j=1}^{r}[\nu_{ij}(x)-\alpha(x)]L(c_i)L(c_j)+\sum_{i=1}^{r}[\nu_{ii}(x)-\alpha(x)]Q(c_i)$$

with $[\nu_{ij}(x)-\alpha(x)]\ge 0$ for any $i,j=1,\cdots,r$. Modelling the proof of Theorem 2.7, we immediately derive the desired result. □

Furthermore, we can obtain the bounded property of ∂G if ∂g is a bounded set.

Corollary 2.12 *Under the assumptions of Theorem 2.11, for any $V\in\partial G(x)$ and scalars $a,b\in\mathbb{R}$ with $a\le b$, the following conclusions hold:*

(i) If $\partial g(t)\subseteq[a,b]$, then $aI\preceq V\preceq bI$.

(ii) If $\partial g(t)\subseteq(a,b)$ with $a<b$, then $aI\prec V\prec bI$.

Proof. Let $f(t)=g(t)-at$. Note that $\partial g(t)\subseteq[a,b]$, then $\partial f(t)\subseteq\mathbb{R}_+$. By Theorem 2.11, one has $V-aI\succeq 0$ for any $V\in\partial G(x)$. On the other hand, letting $\bar{f}(t)=bt-g(t)$, one has $\partial\bar{f}(t)\subseteq\mathbb{R}_+$ and hence $bI-V\succeq 0$ for any $V\in\partial G(x)$. These two arguments show part (i). Similarly, we can verity Part (ii). □

3 Total NR-function

For problem (1.1), we define the *natural residual function* (NR-function) $\Phi_{NR} : \mathcal{J} \times \mathcal{J} \to \mathcal{J}$ by

$$\Phi_{NR}(x, y) \triangleq x - P_K(x - y), \tag{3.1}$$

and the *total NR-function* $H_{NR} : \mathcal{J} \times \mathcal{J} \to \mathcal{J} \times \mathcal{J}$ by

$$H_{NR}(x, y) \triangleq \begin{pmatrix} \Phi_{NR}(x, y) \\ F(x) - y \end{pmatrix}. \tag{3.2}$$

Moreover, we specify function $\Psi_{NR} : \mathcal{J} \times \mathcal{J} \to \mathbb{R}$ by

$$\Psi_{NR}(x, y) \triangleq \frac{1}{2}\|H_{NR}(x, y)\|^2 = \frac{1}{2}\|\Phi_{NR}(x, y)\|^2 + \frac{1}{2}\|F(x) - y\|^2. \tag{3.3}$$

From Proposition 6 in [12], we know that

$$\Phi_{NR}(x, y) = 0 \iff x \in K, \ y \in K, \ \langle x, y \rangle = 0.$$

Therefore, problem (1.1) can be reformulated as a nonsmooth system of nonlinear equations: $H_{NR}(x, y) = 0$. Based on this system, we may establish various solution methods, such as nonsmooth and smoothing Newton methods, see, e.g., [9, 17] for the case of NCP. In this paper, our aim is to present a globally and quadratically convergent regularized smoothing Newton method for SCCP. For this purpose, we need to investigate strong semismoothness of H_{NR}, nonsingularity of ∂H_{NR}, and level-boundedness of Ψ_{NR}.

First, we show strong semismoothness of H_{NR}. Since the proof is similar to that of Theorem 4.6 in [15], it is omitted.

Theorem 3.1 *Let $F : \mathcal{J} \to \mathcal{J}$ be continuously differentiable. Then the function H_{NR} defined by (3.2) is semismooth at any $(x, y) \in \mathcal{J} \times \mathcal{J}$. Moreover, if ∇F is locally Lipschitzian, then H_{NR} is strongly semismooth at any $(x, y) \in \mathcal{J} \times \mathcal{J}$.*

Next, we address Clarke generalized Jacobian ∂H_{NR}. Let $T \in \partial H_{NR}(x, y)$ for any $(x, y) \in \mathcal{J} \times \mathcal{J}$. Then T has the following form:

$$T = \begin{pmatrix} I - V & \nabla F(x) \\ V & -I \end{pmatrix}, \tag{3.4}$$

where $V \in \partial P_K(x - y)$. Since ∂t_+ equals $\{1\}$ for $t > 0$, $[0, 1]$ for $t = 0$, and $\{0\}$ for $t < 0$, by Corollary 2.12 (i) we have $0 \preceq V \preceq I$.

The nonsingularity result on T is well-known for NCP (see, e.g., [9]) or SOCCP (see, e.g., [11]). In a similar manner, we can easily show that it is still true for SCCP, which does not need a further proof. We say that $F : \mathcal{J} \to \mathcal{J}$ is *monotone* (*strongly monotone*) if for all $(x, y) \in \mathcal{J} \times \mathcal{J}$, $\langle x - y, F(x) - F(y) \rangle \geq 0$ ($\langle x - y, F(x) - F(y) \rangle \geq \varepsilon\|x - y\|^2$ with some $\varepsilon > 0$).

Theorem 3.2 *Let $F : \mathcal{J} \to \mathcal{J}$ be continuously differentiable, and T be given by (3.4).*

(a) If F is monotone and $0 \prec V \prec I$, then T is invertible for any $(x, y) \in \mathcal{J} \times \mathcal{J}$.

(b) If F is strongly monotone and $0 \preceq V \preceq I$, then T is invertible for any $(x, y) \in \mathcal{J} \times \mathcal{J}$.

A remark on Theorem 3.2 is here given. If V is a linear and symmetric operator from $\mathcal{J}$ into itself, then the results in this theorem are still true.

We end this section by stating a well-known result on the boundedness of the level sets $\mathrm{Lev}_\alpha(\Psi_{NR}) \triangleq \{(x,y) \in \mathcal{J} \times \mathcal{J} : \Psi_{NR}(x,y) \le \alpha\}$ for $\alpha \in \mathbb{R}$, which can ensure that the sequence generated by a descent method for solving $\min \Psi_{NR}(x,y)$ has at least one accumulation point. For more details, see, e.g., [25, 34].

Theorem 3.3 *Let Ψ_{NR} be defined by (3.3). If $F(x)$ is strongly monotone and locally Lipschitzian, then the level sets $Lev_\alpha(\Psi_{NR})$ are bounded for all $\alpha \in \mathbb{R}$.*

Notice that Gowda, Sznajder and Tao [12] introduced the P and P_0 properties for linear transformations on Euclidean Jordan algebras, and Tao and Gowda [31] studied the P and P_0 properties for nonlinear transformations. Do Theorem 3.2 (b) and Theorem 3.3 hold for the SCCP with P property? These questions are yet to be answered.

4 Chen-Mangasarian Smoothing Function

In the literature on NCP, there are two well-known classes of the smoothing functions, i.e., the Chen-Mangasarian smoothing function and the smoothed Fischer-Burmeister function. Recently, they were successfully extended to SDCP [6] and SOCCP [11]. To smoothen the NR-function established in previous section, we shall focus on an extension of the Chen-Mangasarian smoothing function and analyze the corresponding properties.

Definition 4.1 *Let $\digamma : \mathcal{X} \to \mathcal{Y}$ be a nondifferentiable function. A function $\digamma_u : \mathcal{X} \to \mathcal{Y}$ with a parameter vector $u \in \mathbb{R}^q_+$ is called a smoothing function of $\digamma$ if it has the following properties:*

(a) $\digamma_u$ is continuously differentiable for any $u \in \mathbb{R}^q_{++}$;

(b) $\lim_{u\downarrow 0} \digamma_u(x) = \digamma(x)$ for any $x \in \mathcal{X}$, where $u \downarrow 0$ means $u \in \mathbb{R}^q_{++}, u \to 0$.

We say $\digamma_u$ is a uniformly smooth approach function of $\digamma$ if there is a scalar $\kappa > 0$ such that

$$\|\digamma_u(x) - \digamma(x)\| \le \kappa \|u\|, \quad \forall u \in \mathbb{R}^q_{++},\ \forall x \in \mathcal{X}.$$

Let $\mu \in \mathbb{R}_{++}$. For NR-function Φ_{NR} as in (3.1), we define *Chen-Mangasarian smoothing function* $\Phi_\mu : \mathcal{J} \times \mathcal{J} \to \mathcal{J}$ by

$$\Phi_\mu(x,y) = x - \Pi_\mu(x-y), \tag{4.1}$$

where $\Pi_\mu : \mathcal{J} \to \mathcal{J}$ is specified by $\Pi_\mu(z) \triangleq \mu G(z/\mu)$ and $G \in \mathcal{CM}$. Here, $\mathcal{CM}$ denotes the set of Löwner operators defined by (2.3) with $g : \mathbb{R} \to \mathbb{R}_+$, a continuously differentiable convex function satisfying

$$\lim_{t\to-\infty} g(t) = 0,\ \lim_{t\to\infty}(g(t)-t) = 0 \ \text{ and } \ 0 < g'(t) < 1 \ \text{ for all } \ t \in \mathbb{R}. \tag{4.2}$$

Two known cases of function g are as follows: One is the CHKS function $g(t) = (\sqrt{t^2+4}+t)/2$, which was proposed by Chen and Harker [1], Kanzow [18] and Smale [28], and the other is the neural network function $g(t) = \ln(e^t+1)$, which was used in neural networks [2]. Based on the above definitions and Theorem 2.1, we below derive formulae for Φ_μ.

Proposition 4.2 *Let Φ_μ be given by (4.1). Then it holds that $\Phi_\mu(x,y) = x - \mu\sum_{i=1}^r g(\lambda_i/\mu)c_i$ where $\lambda_i,\ c_i\ (i = 1, 2, \cdots, r)$ are given by $x - y = \sum_{i=1}^r \lambda_i c_i$. Moveover, the pointwise limit*

$$\Phi_0(x,y) \triangleq \lim_{\mu\downarrow 0} \Phi_\mu(x,y) = x - P_K(x-y).$$

Proof. The first part is trivial. Note that $\lim_{\mu\downarrow 0}\mu g(\lambda_i/\mu) = (\lambda_i)_+$ by (4.2). This derives that $\lim_{\mu\downarrow 0}\Phi_\mu(x,y) = x - \sum_{i=1}^r(\lambda_i)_+c_i$. The second part holds by (2.4). □

4.1 Uniformly smooth approximation

The following proposition claims that Φ_μ is a uniformly smooth approximation of Φ_{NR}.

Proposition 4.3 *Let Φ_μ be given by (4.1). Then, for any scalars $\mu > \nu \geq 0$, we have*

$$g(0)(\mu-\nu)e \succeq_K \Phi_\nu(x,y) - \Phi_\mu(x,y) \succ_K 0, \quad \forall x, y \in \mathcal{J}. \tag{4.3}$$

Proof. In order to prove the proposition, we first consider the case where $\mu > \nu > 0$. By Proposition 4.2, it is easy to verify $\Phi_\nu(x,y) - \Phi_\mu(x,y) = \sum_{i=1}^r(\mu g(\lambda_i/\mu) - \nu g(\lambda_i/\nu))c_i$ where λ_i and c_i are given by $x - y = \sum_{i=1}^r \lambda_i c_i$. Noting that for every $i = 1, 2, \cdots, r$, $0 < \mu g(\lambda_i/\mu) - \nu g(\lambda_i/\nu) \leq g(0)(\mu - \nu)$ by Lemma 3.1 in [32], we have

$$g(0)(\mu-\nu)e = \sum_{i=1}^r g(0)(\mu-\nu)c_i \succeq_K \Phi_\nu(x,y) - \Phi_\mu(x,y) \succ_K 0. \tag{4.4}$$

This shows that (4.3) holds in the case of $\mu > \nu > 0$, and that $-\Phi_\nu$ is monotone in $\nu > 0$ with respect to the partial ordering $\succ_K$. Taking $\nu \to 0^+$ in (4.4), one has $g(0)\mu e \succeq_K \Phi_0(x,y) - \Phi_\mu(x,y) \succ_K 0$. That is, (4.4) also holds for $\mu > \nu = 0$. The proof is completed. □

4.2 Differentiability

Let $g : \mathbb{R} \to \mathbb{R}_+$ be a continuously differentiable convex function satisfying (4.2). As in [32] for the setting of NCP and in [15] for the context of SOCCP, we define for any $\mu > 0$,

$$\gamma_\mu(t) \triangleq \mu g(t/\mu), \tag{4.5}$$

$$\gamma_0(t) \triangleq \lim_{\mu\downarrow 0}\gamma_\mu(t) = \max\{0, t\}, \tag{4.6}$$

$$\gamma_0^+(t) \triangleq \lim_{\mu\downarrow 0}\gamma'_\mu(t) = \begin{cases} 0 & \text{for } t < 0, \\ g'(0) & \text{for } t = 0, \\ 1 & \text{for } t > 0. \end{cases} \tag{4.7}$$

Let $z = \sum_{j=1}^r \lambda_j(z)c_j(z)$. By $\Pi_\mu(z) = \mu G(z/\mu)$ with $G \in \mathcal{CM}$, Theorem 2.5 leads to

$$\nabla\Pi_\mu(z) = \nabla G(z/\mu) = 2\sum_{i\neq j, i,j=1}^r a_{ij}L(c_i(z))L(c_j(z)) + \sum_{i=1}^r a_{ii}Q(c_i(z)), \tag{4.8}$$

where for all $i, j = 1, 2, \cdots, r$,

$$a_{ij} = [\lambda_i(z)/\mu, \lambda_j(z)/\mu]_g = \begin{cases} \frac{g(\lambda_i(z)/\mu) - g(\lambda_j(z)/\mu)}{\lambda_i(z)/\mu - \lambda_j(z)/\mu} & \text{if } \lambda_i(z) \neq \lambda_j(z), \\ g'(\lambda_i(z)/\mu) & \text{if } \lambda_i(z) = \lambda_j(z). \end{cases}$$

By (4.5), we have $\gamma'_\mu(t) = g'(t/\mu)$. Therefore

$$a_{ij} = [\lambda_i(z), \lambda_j(z)]_{\gamma_\mu} = \begin{cases} \frac{\gamma_\mu(\lambda_i(z)) - \gamma_\mu(\lambda_j(z))}{\lambda_i(z) - \lambda_j(z)} & \text{if } \lambda_i(z) \neq \lambda_j(z), \\ \gamma'_\mu(\lambda_i(z)) & \text{if } \lambda_i(z) = \lambda_j(z). \end{cases} \tag{4.9}$$

By (4.2) and (4.9), one has $0 < a_{ij} < 1$. Thus, by Corollary 2.12 (ii), it holds $I \succ \nabla\Pi_\mu(z) \succ 0$. In summary, we have the following conclusion.

Proposition 4.4 *The function* Π_μ *is continuously differentiable, and* $I \succ \nabla\Pi_\mu(z) \succ 0$.

Furthermore, by applying Theorem 2.5 and the chain rule, we immediately obtain the differential property of the Chen-Mangasarian smoothing function Φ_μ. The proof is not needed.

Proposition 4.5 *For any* $\mu > 0$, *the Chen-Mangasarian smoothing function* Φ_μ, *defined by (4.1), is continuously differentiable and its Jacobian is given by*

$$\nabla\Phi_\mu(x,y) = \begin{pmatrix} I - \nabla\Pi_\mu(x-y) \\ \nabla\Pi_\mu(x-y) \end{pmatrix} = \begin{pmatrix} I - \nabla G((x-y)/\mu) \\ \nabla G((x-y)/\mu) \end{pmatrix}.$$

4.3 Jacobian consistency

Like strong semismoothness, Jacobian consistency plays an important role in establishing rapid convergence of smoothing Newton methods. This concept was originally introduced by Chen, Qi and Sun [8] for variational inequalities, and was recently used by Hayashi, Yamashita and Fukushima [15] analyzing the regularized smoothing method for SOCCP, where their Jacobian consistency contains two parameters. We state more general definition as follows.

Definition 4.6 *Suppose that* $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ *is a continuous function and* $\partial\mathcal{F}$ *exists. Let* $\mathcal{F}_u$ *be a smoothing function of* $\mathcal{F}$. *We say that* $\mathcal{F}_u$ *satisfies the Jacobian consistency if*

$$\lim_{u\downarrow 0} \operatorname{dist}(\nabla\mathcal{F}_u(x), \partial\mathcal{F}(x)) = 0, \quad \text{for any } x \in \mathcal{X}. \tag{4.10}$$

To show Jacobian consistency of Chen-Mangasarian smoothing function Φ_μ, we first look at the function $\Pi_\mu(z)$. Define $b_{ij} \triangleq \lim_{\mu\downarrow 0} a_{ij}$ for all $i, j = 1, 2, \cdots, r$. From (4.5)-(4.7) and (4.9), we derive that

$$b_{ij} = \begin{cases} \frac{\gamma_0(\lambda_i(z)) - \gamma_0(\lambda_j(z))}{\lambda_i(z) - \lambda_j(z)} & \text{if } \lambda_i(z) \neq \lambda_j(z), \\ \gamma_0^+(\lambda_i(z)) & \text{if } \lambda_i(z) = \lambda_j(z). \end{cases} \tag{4.11}$$

Obviously, by (4.2), $0 \le b_{ij} \le 1$. By the direct calculation, one has

$$\lim_{\mu\downarrow 0} \nabla\Pi_\mu(z) = 2 \sum_{i\neq j, i,j=1}^{r} b_{ij} L(c_i(z)) L(c_j(z)) + \sum_{i=1}^{r} b_{ii} Q(c_i(z)). \tag{4.12}$$

Rewriting z as $z = \sum_{i=1}^{\bar{r}} \mu_i(z) b_i(z)$, from Theorem 2.5 we deduce

$$\nabla\Pi_\mu(z) = 2 \sum_{i\neq j,\ i,j=1}^{\bar{r}} [\mu_i(z), \mu_j(z)]_{\gamma_\mu} L(b_i(z)) L(b_j(z)) + \sum_{i=1}^{\bar{r}} \gamma'_\mu(\mu_i(z)) Q(b_i(z)).$$

In a similar manner as in (4.12), we derive that

$$\lim_{\mu\downarrow 0} \nabla\Pi_\mu(z) = 2 \sum_{i\neq j,\ i,j=1}^{\bar{r}} [\mu_i(z), \mu_j(z)]_{\gamma_0} L(b_i(z)) L(b_j(z)) + \sum_{i=1}^{\bar{r}} \gamma_0^+(\mu_i(z)) Q(b_i(z)).$$

Take $\partial^0_\Pi(z) \triangleq \lim_{\mu\downarrow 0} \nabla\Pi_\mu(z)$. It follows from Theorem 2.10 that $\partial^0_\Pi(z) \in \underline{\partial} P_K(z) \subseteq \partial P_K(z)$. Summarizing the preceding argument, we have the following.

Lemma 4.7 *Let* $\partial^0_\Pi(z) = \lim_{\mu\downarrow 0} \nabla\Pi_\mu(z)$. *Then* $\partial^0_\Pi(z) \in \partial P_K(z)$ *for any* $z \in \mathcal{J}$. *Thus* Π_μ *satisfies the Jacobian consistency.*

Combining Lemma 4.7 with Proposition 4.5, the Jacobian consistency of Φ_μ is immediate.

Proposition 4.8 *Φ_μ satisfies the Jacobian consistency.*

In the end of this section, we further consider the function g satisfying both (4.2) and the following

$$g(t) - t = g(-t), \quad \forall t \in \mathbb{R}. \tag{4.13}$$

For instance, $(\sqrt{t^2+4}+t)/2$ and $\ln(e^t+1)$ are such two functions. Can we get more specific result than Proposition 4.8 in this case? To settle this question, we need the following lemma from [15].

Lemma 4.9 *(Lemma 4.10, [15]) Let g be a continuously differentiable convex function satisfying (4.2) and (4.13). Let γ_μ, γ_0 and γ_0^+ be given by (4.5)-(4.7). Then it holds that*

(a) $\gamma_\mu(t) - \gamma_0(t) = \gamma_\mu(-t) - \gamma_0(-t)$ for any $t \in \mathbb{R}$;

(b) $|\gamma'_\mu(t) - \gamma_0^+(t)| = |\gamma'_\mu(|t|) - \gamma_0^+(|t|)|$ for any $t \in \mathbb{R}$;

(c) $|\gamma'_\mu(0) - \gamma_0^+(0)| = 0 < |\gamma'_\mu(t_2) - \gamma_0^+(t_2)| \le |\gamma'_\mu(t_1) - \gamma_0^+(t_1)|$ for any $t_i \in \mathbb{R}$ (i=1,2) such that $0 < |t_1| \le |t_2|$.

For $z = \sum_{j=1}^r \lambda_j(z)c_j(z)$, let $N(z)$ be the index set specified by $N(z) \triangleq \{i : \lambda_i(z) \neq 0\}$. Define the function $\tilde{\lambda} : \mathcal{J} \to \mathbb{R}_+$ by

$$\tilde{\lambda}(z) \triangleq \begin{cases} \min_{i\in N(z)} |\lambda_i(z)| & \text{for } N(z) \neq \emptyset, \\ 0 & \text{for } N(z) = \emptyset. \end{cases} \tag{4.14}$$

Obviously, $\tilde{\lambda}(z) = 0$ if and only if $z = 0$. When $z \neq 0$, by (4.5) and the continuous differentiability of g, there is a scalar $\varsigma \in (0, \tilde{\lambda}(z))$ such that $\gamma'_\mu(\varsigma) = \frac{\gamma_\mu(\tilde{\lambda}(z))-\gamma_\mu(0)}{\tilde{\lambda}(z)}$; meanwhile, noting that g is convex, one has $\gamma'_\mu(t) \le \frac{\gamma_\mu(\tilde{\lambda}(z))-\gamma_\mu(0)}{\tilde{\lambda}(z)}$ for any $t \in (0, \varsigma)$. So, in the case of $z \neq 0$, there exists a positive integer l such that $\frac{1}{2^l}\tilde{\lambda}(z) \in (0, \varsigma)$.

Based on the preceding argument, we define the function $\lambda^\star : \mathcal{J} \to \mathbb{R}_+$ by

$$\lambda^\star(z) \triangleq \begin{cases} \frac{1}{2^l}\tilde{\lambda}(z) & \text{for } N(z) \neq \emptyset, \\ 0 & \text{for } N(z) = \emptyset, \end{cases} \tag{4.15}$$

where l is the smallest positive integer such that

$$\gamma'_\mu(\frac{1}{2^l}\tilde{\lambda}(z)) \le \frac{\gamma_\mu(\tilde{\lambda}(z)) - \gamma_\mu(0)}{\tilde{\lambda}(z)}. \tag{4.16}$$

Then $\lambda^\star(z)$ is well-defined and $0 < \lambda^\star(z) < \tilde{\lambda}(z)$. Thus, it holds by Lemma 4.9 (c) that

$$|\gamma'_\mu(\lambda_i(z)) - \gamma_0^+(\lambda_i(z))| \le |\gamma'_\mu(\tilde{\lambda}(z)) - \gamma_0^+(\tilde{\lambda}(z))| \le |\gamma'_\mu(\lambda^\star(z)) - \gamma_0^+(\lambda^\star(z))|, \; i = 1, 2, \cdots, r. \tag{4.17}$$

Now we are ready to claim that $\Pi_\mu(z)$ not only satisfies the Jacobian consistency but also has the stronger Jacobian property.

Theorem 4.10 *Let $\partial^0_\Pi(z) = \lim_{\mu\downarrow 0} \nabla\Pi_\mu(z)$. Suppose g is a continuously differentiable convex function satisfying (4.2) and (4.13). Let $\gamma_\mu, \gamma_0, \gamma_0^+$ and $\lambda^\star$ be given by (4.5)-(4.7) and (4.15), respectively. Then there exists a scalar $\bar{M} > 0$ such that*

$$\|\nabla\Pi_\mu(z) - \partial^0_\Pi(z))\| \le \bar{M}|\gamma'_\mu(\lambda^\star(z)) - \gamma_0^+(\lambda^\star(z))|, \quad \forall \mu \in \mathbb{R}_{++}, \; \forall\, z \in \mathcal{J}.$$

Proof. Let $z = \sum_{j=1}^{r} \lambda_j(z) c_j(z)$. Then from (4.8) and (4.12) we obtain

$$\nabla \Pi_\mu(z) - \partial_\Pi^0(z) = 2 \sum_{i \neq j, i,j=1}^{r} (a_{ij} - b_{ij}) L(c_i(z)) L(c_j(z)) + \sum_{i=1}^{r} (a_{ii} - b_{ii}) Q(c_i(z)).$$

To prove the theorem, it is enough to show $|a_{ij} - b_{ij}| \leq |\gamma'_\mu(\lambda^\star(z)) - \gamma_0^+(\lambda^\star(z))|$ for every $i, j = 1, 2, \cdots, r$. We below consider two cases.

Case (i): $0 = \lambda_i(z) < |\lambda_j(z)|$. By (4.9) and (4.11), the direct calculation yields

$$\begin{aligned} |a_{ij} - b_{ij}| &= \left| \frac{\gamma_\mu(0) - \gamma_\mu(\lambda_j(z))}{0 - \lambda_j(z)} - \frac{\gamma_0(0) - \gamma_0(\lambda_j(z))}{0 - \lambda_j(z)} \right| \\ &= \left| \frac{\gamma_\mu(\lambda_j(z)) - \gamma_\mu(0)}{\lambda_j(z)} - 1 \right| \\ &= 1 - \frac{\gamma_\mu(\lambda_j(z)) - \gamma_\mu(0)}{\lambda_j(z)} \\ &\leq 1 - \gamma'_\mu(\lambda^\star(z)) \\ &= |\gamma'_\mu(\lambda^\star(z)) - \gamma_0^+(\lambda^\star(z))|, \end{aligned}$$

where the second equality follows from the fact $\frac{\gamma_0(0) - \gamma_0(\lambda_j(z))}{0 - \lambda_j(z)} = 1$ by (4.6), the third one from $0 < \frac{\gamma_\mu(\lambda_j(z)) - \gamma_\mu(0)}{\lambda_j(z)} = \frac{g(\lambda_j(z)/\mu) - g(0)}{\lambda_j(z)/\mu} < 1$ by (4.2), the inequality from (4.15), and the last equality from $\gamma_0^+(\lambda^\star(z)) = 1$ by (4.15) and (4.7).

Case (ii): Otherwise, one has $|a_{ij} - b_{ij}| \leq |\gamma'_\mu(\tilde{\lambda}(z)) - \gamma_0^+(\tilde{\lambda}(z))|$, whose proof is perfectly similar to that in [15] and is omitted for brevity. □

5 Regularized Smoothing Function and Algorithm

Based on the proceeding results, we shall develop the Chen-Mangasarian class of regularized smoothing functions for SCCP, and derive the regularized smoothing Newton method for solving the monotone SCCP.

For the given F in (1.1) and a parameter $\varepsilon > 0$, we define a new function $F_\varepsilon : \mathcal{J} \to \mathcal{J}$ as

(5.1) $$F_\varepsilon(x) \triangleq F(x) + \varepsilon x.$$

Again, define functions $H_{\mu,\varepsilon} : \mathcal{J} \times \mathcal{J} \to \mathcal{J} \times \mathcal{J}$ and $\Psi_{\mu,\varepsilon} : \mathcal{J} \times \mathcal{J} \to \mathbb{R}$ by

(5.2) $$H_{\mu,\varepsilon}(x, y) \triangleq \begin{pmatrix} \Phi_\mu(x, y) \\ F_\varepsilon(x) - y \end{pmatrix},$$

(5.3) $$\Psi_{\mu,\varepsilon}(x, y) \triangleq \frac{1}{2} \|H_{\mu,\varepsilon}(x, y)\|^2 = \frac{1}{2} \|\Phi_\mu(x, y)\|^2 + \frac{1}{2} \|F_\varepsilon(x) - y\|^2.$$

Then, $H_{\mu,\varepsilon}$ is a smoothing approximation of the regularized SCCP involving F_ε with $\varepsilon > 0$. Obviously, if F is monotone, then F_ε is strongly monotone for any $\varepsilon > 0$. In addition, if F is also locally Lipschitzian, then $\Psi_{\mu,\varepsilon}$ is level-bounded for any $\mu \geq 0$ and $\varepsilon > 0$ via Theorem 3.3.

The proposed method applies the Newton algorithm to the system $H_{\mu,\varepsilon}(x, y) = 0$ with μ and ε properly adjusted at each iteration, so that a solution of the original SCCP is eventually obtained by taking the limits as $\mu \downarrow 0$ and $\varepsilon \downarrow 0$.

For this purpose, we deal with $H_{\mu,\varepsilon}$. From Proposition 4.5, we obtain

$$\nabla H_{\mu,\varepsilon}(x,y) = \begin{pmatrix} I - \nabla\Pi_\mu(x-y) & \nabla F(x) + \varepsilon I \\ \nabla\Pi_\mu(x-y) & -I \end{pmatrix}, \tag{5.4}$$

where $\nabla\Pi_\mu(\cdot)$ is specified by (4.8). From (5.4) and Proposition 4.4, one easily get the nonsingularity of $\nabla H_{\mu,\varepsilon}$. The proof is omitted.

Theorem 5.1 *Let $F : \mathcal{J} \to \mathcal{J}$ be continuously differentiable. For parameters $\mu > 0$ and $\varepsilon > 0$, let $\Phi_\mu(x,y), F_\varepsilon(x)$ and $H_{\mu,\varepsilon}(x,y)$ be defined by (4.1), (5.1) and (5.2), respectively. If F is monotone, then $\nabla H_{\mu,\varepsilon}$, given by (5.4), is invertible for any $(x,y) \in \mathcal{J} \times \mathcal{J}$.*

In view of (5.4), we also deduce the Jacobian consistency of $H_{\mu,\varepsilon}$.

Theorem 5.2 *Let $F : \mathcal{J} \to \mathcal{J}$ be continuously differentiable. For parameters $\mu > 0$ and $\varepsilon > 0$, let $\Phi_\mu(x,y), F_\varepsilon(x)$ and $H_{\mu,\varepsilon}(x,y)$ be defined by (4.1), (5.1) and (5.2), respectively. Then $H_{\mu,\varepsilon}$ satisfies the Jacobian consistency.*

Proof. It holds by (5.4) and $\partial^0_\Pi(z) = \lim_{\mu\downarrow 0} \nabla\Pi_\mu(z)$ that

$$\partial^0_\Pi H(x,y) \triangleq \lim_{(\mu,\varepsilon)\downarrow(0,0)} \nabla H_{\mu,\varepsilon}(x,y) = \begin{pmatrix} I - \partial^0_\Pi(x-y) & \nabla F(x) \\ \partial^0_\Pi(x-y) & -I \end{pmatrix}. \tag{5.5}$$

This implies from (3.4) and Lemma 4.7 that $\partial^0_\Pi H(x,y) \in \partial H_{NR}(x,y)$ for any $(x,y) \in \mathcal{J} \times \mathcal{J}$. The desired holds obviously. □

Furthermore, applying Theorems 4.10 and 5.2, we estimate the upper bound of the distance $\mathrm{dist}(\nabla H_{\mu,\varepsilon}(x,y), \partial H_{NR}(x,y))$.

Theorem 5.3 *Let $F : \mathcal{J} \to \mathcal{J}$ be continuously differentiable, and g be a continuously differentiable convex function satisfying (4.2) and (4.13). Suppose γ_μ, γ_0 and γ_0^+ are given by (4.5)-(4.7), and let $\lambda^\star$ be defined by (4.15). Then, there exists a scalar $M > 0$ such that*

$$\mathrm{dist}(\nabla H_{\mu,\varepsilon}(x,y), \partial H_{NR}(x,y)) \le M(|\gamma'_\mu(\lambda^\star(x-y)) - \gamma_0^+(\lambda^\star(x-y))| + \varepsilon),$$

for any $\mu > 0, \varepsilon \ge 0$ and any $(x,y) \in \mathcal{J} \times \mathcal{J}$.

Proof. By (5.4), (5.5) and the fact $\partial^0_\Pi H(x,y) \in \partial H_{NR}(x,y)$, one has for any $\mu > 0, \varepsilon \ge 0$ and any $(x,y) \in \mathcal{J} \times \mathcal{J}$,

$$\begin{aligned} \mathrm{dist}(\nabla H_{\mu,\varepsilon}(x,y), \partial H_{NR}(x,y)) &\le ||\nabla H_{\mu,\varepsilon}(x,y) - \partial^0_\Pi H(x,y)|| \\ &\le \tilde{M}(\|\nabla\Pi_\mu(x-y)) - \partial^0_\Pi(x-y))\| + \varepsilon) \\ &\le \tilde{M}(\bar{M}|\gamma'_\mu(\lambda^\star(x-y)) - \gamma_0^+(\lambda^\star(x-y))| + \varepsilon) \end{aligned}$$

where $\tilde{M}$ in the second inequality is a positive scalar, the third follows from Theorem 4.10. The desired holds immediately. □

In the end of this paper, we describe the desired algorithm which is a word-for-word extension of the one by Hayashi, Yamashita and Fukushima [15] for SOCCP, and state the corresponding convergence theorem which can be obtained by Theorems 5.1-5.3 and following the proof of Theorem 4.13 in [15].

ALGORITHM Set $w \triangleq (x, y)$ and $w^{(k)} \triangleq (x^{(k)}, y^{(k)})$. Choose $\eta, \rho \in (0,1), \bar{\eta} \in (0, \eta], \sigma \in (0, 1/2), \kappa > 0$ and $\hat{\kappa} > 0$.

Step 0 Choose $w^{(0)} \in \mathcal{J} \times \mathcal{J}$ and $\beta_0 \in (0, \infty)$. Let $\mu_0 \triangleq \|H_{NR}(w^{(0)})\|$ and $\varepsilon_0 \triangleq \|H_{NR}(w^{(0)})\|$. Set $k \triangleq 0$.

Step 1 Terminate if $\|H_{NR}(w^{(k)})\| = 0$.

Step 2

Step 2.0 Set $v^{(0)} \triangleq w^{(0)}$ and $j \triangleq 0$.

Step 2.1 Find a vector $\hat{d}^{(j)}$ such that

$$H_{\mu_k, \varepsilon_k}(v^{(j)}) + \nabla H_{\mu_k, \varepsilon_k}(v^{(j)})^T \hat{d}^{(j)} = 0.$$

Step 2.2 If $\|H_{\mu_k, \varepsilon_k}(v^{(j)} + \hat{d}^{(j)})\| \leq \beta_k$, then let $w^{(k+1)} \triangleq v^{(j)} + \hat{d}^{(j)}$ and go to Step 3. Otherwise, go to Step 2.3.

Step 2.3 Find the smallest nonnegative integer m such that

$$\Psi_{\mu_k, \varepsilon_k}(v^{(j)} + \rho^m \hat{d}^{(j)}) \leq (1 - 2\sigma\rho^m) \Psi_{\mu_k, \varepsilon_k}(v^{(j)}).$$

Let $m_j \triangleq m,\ \tau_j \triangleq \rho^{m_j}$ and $v^{(j+1)} \triangleq v^{(j)} + \tau_j \hat{d}^{(j)}$.

Step 2.4 If $\|H_{\mu_k, \varepsilon_k}(v^{(j+1)})\| \leq \beta_k$, then let $w^{(k+1)} \triangleq v^{(j+1)}$ and go to Step 3. Otherwise, set $j \triangleq j + 1$ and go back to Step 2.1.

Step 3 Update the parameters as follows:

$$\begin{aligned}
\mu_{k+1} &:= \min\{\kappa \|H_{NR}(w^{(k+1)})\|^2, \mu_0 \bar{\eta}^{k+1}, \bar{\mu}(\lambda^\star(x^{(k+1)} - y^{(k+1)}), \hat{\kappa}\|H_{NR}(w^{(k+1)})\|)\}, \\
\varepsilon_{k+1} &:= \min\{\kappa \|H_{NR}(w^{(k+1)})\|^2, \varepsilon_0 \bar{\eta}^{k+1}\}, \\
\beta_{k+1} &:= \beta_0 \eta^{k+1},
\end{aligned}$$

where $\lambda^\star$ is given by (4.15), and $\bar{\mu}(t, \delta)$ is determined so that $|\gamma'_\mu(t) - \gamma_0^+(t)| < \delta$ for any $\mu \in (0, \bar{\mu}(t, \delta))$.

Step 4 Set $k \triangleq k + 1$. Go back to Step 1.

Theorem 5.4 *Let $F : \mathcal{J} \to \mathcal{J}$ be a continuously differentiable and monotone function, and $\{w^{(k)}\}$ be a sequence generated by Algorithm. If the solution set of SCCP (1.1) is nonempty and bounded, then $\{w^{(k)}\}$ is bounded, and every accumulation point is a solution of SCCP(1.1). In addition, if ∇F is locally Lipschitzian and every accumulation point of $\{\nabla H_{\mu_k, \varepsilon_k}(w^{(k)})\}$ is nonsingular, then the sequence $\{w^{(k)}\}$ converges to a solution w^* of SCCP(1.1) quadratically.*

Acknowledgments The authors are very grateful to the anonymous referees for their constructive comments and valuable suggestions which helped us refine the main content and improve the presentation of this paper. The authors would like to thank Levent Tunçel from University of Waterloo and Defeng Sun from National University of Singapore for helpful discussions.

References

[1] B. Chen and P.T. Harker, A non-interior-point continuation method for linear complementarity problems, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168-1190.

[2] C. Chen and O.L. Mangasarian, Smoothing methods for convex inequalities and linear complementarity problems, Math. Program., 71 (1995), pp. 51-69.

[3] J.-S. Chen and P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Math. Program. Ser. B, 104 (2005), pp. 293-327.

[4] X. Chen and H.D. Qi, Cartesian P-property and its applications to the semidefinite linear complementarity problem, Math. Program., 106 (2006), pp. 177-201.

[5] X. Chen, H.D. Qi and P. Tseng, Analysis of nonsmooth symmetric matrix functions with applications to semidefinite complementarity problems, SIAM J. Optim., 13 (2003), pp. 960-985.

[6] X. Chen and P. Tseng, Non-interior continuation methods for solving semidefinite complementarity problems, Math. Program., 95 (2003), pp. 431-474.

[7] X.D. Chen, D. Sun and J. Sun, Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems, Comput. Optim. Appl., 25 (2003), pp. 39-56.

[8] X.J. Chen, L. Qi and D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, Math. Comput., 67 (1998), pp. 519-540.

[9] F. Facchinei and J.-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I and II, Springer-Verlag, New York, 2003.

[10] J. Faraut and A. Korányi, Analysis on Symmetric Cones, Oxford University Press, New York, 1994.

[11] M. Fukushima, Z.Q. Luo and P. Tseng, Smoothing functions for second-order cone complementarity problems, SIAM J. Optim., 12 (2001), pp. 436-460.

[12] M.S. Gowda, R. Sznajder and J. Tao, Some P-properties for linear transformations on Euclidean Jordan algebras, Linear Algebra and Its Applications, 393 (2004), pp. 203-232.

[13] M.S. Gowda and R. Sznajder, Automorphism invariance of P and GUS properties of linear transformations on Euclidean Jordan algebras, Math. Oper. Res., 31 (2006), pp. 109-123.

[14] M.S. Gowda and R. Sznajder, Some global uniqueness and solvability results for linear complementarity problems over symmetric cones, SIAM J. Optim., 18 (2007), pp 461-481.

[15] S. Hayashi, N. Yamashita and M. Fukuahima, A combined smoothing and regularization method for monotone second-order cone complementarity problems, SIAM J. Optim., 15 (2005), pp. 593-615.

[16] Z.H. Huang and J. Han, Non-interior continuation method for solving the monotone semidefinite complementarity problem, Appl. Math. Optim., 47 (2003), pp. 195-211.

[17] G. Isac, Topological Methods in Complementarity Theory, Kluwer Academic Publishers, Dordrecht, 2000.

[18] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851-868.

[19] M. Koecher, The Minnesota Notes on Jordan Algebras and Their Applications, edited and annotated by A. Brieg and S. Walcher, Springer, Berlin, 1999.

[20] L.C. Kong and N.H. Xiu, On uniqueness of the Jordan frame in Euclidean Jordan algebras, Journal of Beijing Jiaotong University, 3 (2007), pp. 54-57.

[21] K. Löwner, Über monotone matrixfunctionen, Mathematische Zeitschrift, 38 (1934), pp. 177-216.

[22] Y. Liu, L. Zhang and Y. Wang, Some properties of a class of merit functions for symmetric cone complementarity problems, Asia-Pacific Journal of Operational Research, 23 (2006), pp. 473-496.

[23] M. Malik and S.R. Mohan, Cone complementarity problems with finite solution sets, Operations Research Letters, 34 (2006), pp. 121-126.

[24] R. Mifflin, Semismooth and semiconvex functions in constrained optimization. SIAM J. Cont. Optim., 15 (1977), pp. 957-972.

[25] J. M. Peng and M. Fukushima, A hybrid Newton method for solving the variational inequality problem via the D-gap function, Math. Program., 86 (1999), pp. 367-386.

[26] L. Qi and J. Sun, A nonsmooth version of Newton's method, Math. Program., 58 (1993), pp. 353-367.

[27] R.T. Rockafellar and R.J.-B. Wets, Variational Analysis, Springer, Second Version 2004.

[28] S. Smale, Algorithms for solving equations, in Proceedings of the International Congress of Mathematicians, American Mathematical Society, Providence, RI, 1987, pp. 172-195.

[29] D. Sun and J. Sun, Löwner's operator and spectral functions on Euclidean Jordan algebras, submitted to Math. Oper. Res. (2004).

[30] J. Sun, D. Sun and L. Qi, A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems, SIAM J. Optim., 14 (2004), pp. 783-806.

[31] J. Tao and M.S. Gowda, Some P-properties for nonlinear transformations on Euclidean Jordan algebras, Math. Oper. Res., 30 (2005), pp. 985-1004.

[32] P. Tseng, Analysis of a non-interior continuation method based on Chen-Mangasarian smoothing functions for complementarity problems. In: Reformulation-Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods,M. Fukushima, L. Qi (eds), Kluwer Academic Publishers, Boston, pp. 381-404, 1999.

[33] Y. Xia and J.M. Peng, A continuation method for the linear second-order cone complementarity Problem, Computational Science and Its Applications-ICCSA 2005, VOL 4, Proceedings Lecture Notes in Computer Science 3483, Berlin: Springer, 2005, pp. 290-300.

[34] N. Yamashita and M. Fukushima, On the level-boundedness of the natural residual function for variational inequality problems, Technique report 2002-005, Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan, 2002.

[35] A. Yoshise, Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones, SIAM J. Optim., 17 (2006), pp. 1129-1153.