Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 830698, 16 pages
http://dx.doi.org/10.1155/2013/830698

Research Article

A Smoothing Method with Appropriate Parameter Control Based on Fischer-Burmeister Function for Second-Order Cone Complementarity Problems

Yasushi Narushima,[1] Hideho Ogasawara,[2] and Shunsuke Hayashi[3]

[1] *Department of Management System Science, Yokohama National University, 79-4 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan*
[2] *Department of Mathematical Information Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan*
[3] *Graduate School of Information Sciences, Tohoku University, 6-3-09 Aramaki-aza Aoba, Aoba-ku, Sendai 980-8579, Japan*

Correspondence should be addressed to Yasushi Narushima; narushima@ynu.ac.jp

Received 21 January 2013; Accepted 24 March 2013

Academic Editor: Jein-Shan Chen

We deal with complementarity problems over second-order cones. The complementarity problem is an important class of problems in the real world and involves many optimization problems. The complementarity problem can be reformulated as a nonsmooth system of equations. Based on the smoothed Fischer-Burmeister function, we construct a smoothing Newton method for solving such a nonsmooth system. The proposed method controls a smoothing parameter appropriately. We show the global and quadratic convergence of the method. Finally, some numerical results are given.

1. Introduction

In this paper, we consider the second-order cone complementarity problem (SOCCP) of the following form:

$$\begin{aligned} &\text{find } (x, y, p) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^\ell \quad \text{such that} \\ &\qquad x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad \langle x, y \rangle = 0, \quad F(x, y, p) = 0, \end{aligned} \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $\ell \geq 0$, $n \geq 1$, $F : \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^\ell \to \mathbf{R}^{n+\ell}$ is a continuously differentiable function, and $\mathcal{K}$ denotes the Cartesian product of several second-order cones (SOCs), that is, $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m} \subset \mathbf{R}^n$ with $m, n_1, \ldots, n_m \geq 1$, $n = n_1 + \cdots + n_m$, and

$$\mathcal{K}^{n_i} := \begin{cases} \left\{ (z_1, z_2) \in \mathbf{R} \times \mathbf{R}^{n_i - 1} \mid z_1 \geq \|z_2\|_2 \right\} \subset \mathbf{R}^{n_i} & (n_i \geq 2), \\ \mathbf{R}_+ = \left\{ z_1 \in \mathbf{R} \mid z_1 \geq 0 \right\} & (n_i = 1). \end{cases} \tag{2}$$

The SOCCP is a wide class of complementarity problems. For example, it involves the mixed complementarity problem (MCP) and the nonlinear complementarity problem (NCP) [1] as subclasses, since $\mathcal{K} = \mathbf{R}^n_+$ when $n_i = 1$ for each $i = 1, \ldots, m$ $(=n)$. Moreover, the second-order cone programming (SOCP) problem can be reformulated as an SOCCP by using the Karush-Kuhn-Tucker (KKT) conditions. Apart from them, some practical problems in the game theory [2, 3] and the architecture [4] can be reformulated as the SOCCP.

Much theoretical and algorithmic research has been made so far for solving the SOCCP. Fukushima et al. [5] showed that the *natural residual* function, also called the min function, and the *Fischer-Burmeister* function for the NCP can be extended to the SOCCP by using the Jordan algebra. They further constructed the smoothing functions for those SOC complementarity (C-) functions and analyzed the properties of their Jacobian matrices. Hayashi et al. [6] proposed a smoothing method based on the natural residual and showed its global and quadratic convergence. On the other hand, Chen et al. [7] proposed another smoothing method with the natural residual in which the smoothing parameter is treated as a variable in contrast to [6]. Moreover, they showed the global and quadratic convergence of their method. Similar to Chen et al., Narushima et al. [8] proposed a smoothing method treating a smoothing parameter as a variable. They used the Fischer-Burmeister function instead of the natural residual function and also provided the global and quadratic convergence of the method.

In the present paper, we propose a smoothing method with the Fischer-Burmeister function for solving SOCCP (1). The main difference from the existing methods is twofold.

(i) We do not assume the special structure on the function F in SOCCP (1). In [6, 7, 9, 10], the authors focused on the following type of SOCCP:

$$x \in \mathscr{K}, \quad y \in \mathscr{K}, \quad \langle x, y \rangle = 0, \quad y = f(x), \tag{3}$$

which is a special case of SOCCP (1) with

$$\ell = 0, \quad F(x, y, p) = f(x) - y \tag{4}$$

for some continuously differentiable function $f : \mathbf{R}^n \to \mathbf{R}^n$. In [11–13], the authors studied the following type of SOCCP:

$$\begin{gathered} x \in \mathscr{K}, \quad y \in \mathscr{K}, \quad \langle x, y \rangle = 0, \\ x = f(p), \quad y = g(p), \end{gathered} \tag{5}$$

which is obtained by letting

$$\ell = n, \quad F(x, y, p) = \begin{bmatrix} f(p) - x \\ g(p) - y \end{bmatrix}, \tag{6}$$

where f and $g : \mathbf{R}^n \to \mathbf{R}^n$ are continuously differentiable functions. However, we assume neither (4) nor (6). Therefore, our method is applicable to a wider class of SOCCPs.

(ii) In contrast to [8], we do not incorporate the smoothing parameter into the decision variable. We control the smoothing parameter appropriately in each iteration.

This paper is organized as follows. In Section 2, we give some preliminaries, which will be used in the subsequent analysis. In Section 3, we review the SOC C-function. In particular, we recall the property of the (smoothed) Fischer-Burmeister function. In Section 4, we propose an algorithm for solving the SOCCP and discuss its global and local convergence properties. In Section 5, we report some preliminary numerical results.

Throughout the paper, we use the following notations. Let $\mathbf{R}_+$ and $\mathbf{R}_{++}$ be the sets of nonnegative and positive reals. For a symmetric matrix A, we write $A \succeq O$ (resp., $A \succ O$) if A is positive semidefinite (resp., positive definite). For any $x, y \in \mathbf{R}^n$, we write $x \succeq y$ (resp., $x \succ y$) if $x - y \in \mathscr{K}^n$ (resp., $x - y \in \operatorname{int} \mathscr{K}^n$), and we denote by $\langle x, y \rangle$ the Euclidean inner product, that is, $\langle x, y \rangle := x^\top y$. We use the symbol $\|\cdot\|$ to denote the usual ℓ_2-norm of a vector or the corresponding induced matrix norm. We often write $x = (x_1, x_2) \in \mathbf{R} \times \mathbf{R}^{n-1}$ (possibly x_2 vacuous), instead of $x = (x_1, x_2^\top)^\top \in \mathbf{R}^n$. In addition, we often regard $\mathbf{R}^p \times \mathbf{R}^q$ as $\mathbf{R}^{p+q}$. We sometimes divide a vector $x \in \mathbf{R}^n$ according to the Cartesian structure of $\mathscr{K}$, that is, $x = (x^1, \ldots, x^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m} = \mathbf{R}^n$ with $x^i \in \mathbf{R}^{n_i}$. For any Fréchet-differentiable function $G : \mathbf{R}^n \to \mathbf{R}^m$, we denote its transposed Jacobian matrix at $x \in \mathbf{R}^n$ by $\nabla G(x) \in \mathbf{R}^{n \times m}$. For a given set $S \subset \mathbf{R}^n$, $\operatorname{int} S$, $\operatorname{bd} S$, and $\operatorname{conv} S$ mean the interior, the boundary, and the convex hull of S in $\mathbf{R}^n$, respectively.

2. Some Preliminaries

In this section, we recall some background materials and preliminary results used in the subsequent sections.

First, we review the Jordan algebra associated with SOCs. For any $x = (x_1, x_2) \in \mathbf{R} \times \mathbf{R}^{n-1}$ and $y = (y_1, y_2) \in \mathbf{R} \times \mathbf{R}^{n-1}$ $(n \geq 1)$, the Jordan product associated with $\mathscr{K}^n$ is defined as

$$x \circ y := \left(x^\top y, y_1 x_2 + x_1 y_2\right). \tag{7}$$

When $n = 1$, that is, the second components x_2 and y_2 are vacuous, we interpret that the second component in (7) is also vacuous. We will write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors x and y. For the Jordan product, the identity element $e \in \mathbf{R}^n$ is defined by $e := (1, 0, \ldots, 0)^\top$. It is easily seen that $x \circ e = e \circ x = x$ for any $x \in \mathbf{R}^n$. For any $x \in \mathscr{K}^n$, we define $x^{1/2}$ as

$$x^{1/2} := \begin{cases} \left(\varsigma, \dfrac{x_2}{2\varsigma}\right), & \varsigma = \sqrt{\dfrac{1}{2}\left(x_1 + \sqrt{x_1^2 - \|x_2\|^2}\right)} \quad \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases} \tag{8}$$

Note that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$ and that for $n = 1$, $x \circ y = x_1 y_1$, $e = 1$, and $x^{1/2} = \varsigma = \sqrt{x_1}$. Although the Jordan product is not associative, associativity holds under the inner product in the sense that

$$\langle x, y \circ z \rangle = \langle y, z \circ x \rangle = \langle z, x \circ y \rangle \quad \text{for any } x, y, z \in \mathbf{R}^n. \tag{9}$$

In addition, it follows readily from the definition of $\mathscr{K}^n$ that $\langle x, y \rangle \geq 0$ (resp., $\langle x, y \rangle > 0$) for any $x, y \succeq 0$ (resp., $x, y \succ 0$).

For each $x = (x_1, x_2) \in \mathbf{R} \times \mathbf{R}^{n-1}$ $(n \geq 1)$, we define the symmetric matrix L_x by

$$L_x := \begin{bmatrix} x_1 & x_2^\top \\ x_2 & x_1 I \end{bmatrix}, \tag{10}$$

which can be viewed as a linear mapping having the following properties.

Property 1. There holds that

(a) $L_x y = x \circ y = y \circ x = L_y x$ and $L_{x+y} = L_x + L_y$ for any $x, y \in \mathbf{R}^n$;

(b) $x \succeq 0 \Leftrightarrow L_x \succeq O$, and $x \succ 0 \Leftrightarrow L_x \succ O$;

(c) L_x is invertible whenever $x \in \operatorname{int} \mathscr{K}^n$ with

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

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

An important character of the Jordan algebra is its spectral factorization. By the spectral factorization associated with SOC, any $x = (x_1, x_2) \in \mathbf{R} \times \mathbf{R}^{n-1}$ $(n \geq 1)$ can be decomposed as

$$x = \lambda_1(x) s^{\{1\}} + \lambda_2(x) s^{\{2\}}, \tag{12}$$

where $\lambda_1(x)$, $\lambda_2(x)$ and $s^{\{1\}}$, $s^{\{2\}}$ are the spectral values and the associated spectral vectors of x given by

$$\lambda_j(x) := x_1 + (-1)^j \|x_2\|,$$
$$s^{\{j\}} := \begin{cases} \frac{1}{2}\left(1, (-1)^j \frac{x_2}{\|x_2\|}\right) & \text{if } x_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^j \bar{s}_2\right) & \text{if } x_2 = 0, \end{cases} \tag{13}$$

for $j = 1, 2$, with $\bar{s}_2$ being any vector in $\mathbf{R}^{n-1}$ satisfying $\|\bar{s}_2\| = 1$. If $x_2 \neq 0$, the decomposition (12) is unique. We note again that when $n = 1$ (viz., $x = x_1$), we have $\lambda_1(x) = \lambda_2(x) = x_1$, $s^{\{1\}} = s^{\{2\}} = 1/2$. The spectral factorization associated with SOC leads to a number of interesting properties, some of which are as follows.

Property 2. For any $x = (x_1, x_2) \in \mathbf{R} \times \mathbf{R}^{n-1}$ $(n \geq 1)$, let $\lambda_1(x)$, $\lambda_2(x)$ and $s^{\{1\}}$, $s^{\{2\}}$ be the spectral values and the associated spectral vectors at x. Then the following statements hold.

(a) $x \in \mathscr{K}^n \Leftrightarrow \lambda_1(x) \geq 0$, $x \in \operatorname{int} \mathscr{K}^n \Leftrightarrow \lambda_1(x) > 0$, $x \in \operatorname{bd} \mathscr{K}^n \Leftrightarrow \lambda_1(x) = 0$.

(b) $x^2 = \lambda_1(x)^2 s^{\{1\}} + \lambda_2(x)^2 s^{\{2\}} \in \mathscr{K}^n$.

(c) Let $x \in \mathscr{K}^n$. Then $x^{1/2} = \sqrt{\lambda_1(x)} s^{\{1\}} + \sqrt{\lambda_2(x)} s^{\{2\}} \in \mathscr{K}^n$. Moreover, $x \in \operatorname{int} \mathscr{K}^n \Leftrightarrow x^{1/2} \in \operatorname{int} \mathscr{K}^n$, $x \in \operatorname{bd} \mathscr{K}^n \Leftrightarrow x^{1/2} \in \operatorname{bd} \mathscr{K}^n$.

In what follows, we recall some definitions for functions and matrices. The semismoothness is a generalized concept of the smoothness, which was originally introduced by Mifflin [14] for functionals, and extended to vector-valued functions by Qi and Sun [15]. For vector-valued functions associated with SOC, see also the work of Chen et al. [16]. Now we give the definition of the Clarke subdifferential [17].

Definition 1. Let $H : \mathbf{R}^n \to \mathbf{R}^m$ be a locally Lipschitzian function. The Clarke subdifferential of H at x is defined by

$$\partial H(x) := \operatorname{conv}\left\{\lim_{\hat{x} \to x} \nabla H(\hat{x}) \mid \hat{x} \in \mathscr{D}_H\right\}, \tag{14}$$

where $\mathscr{D}_H$ is the set of points at which H is differentiable.

Note that if H is continuously differentiable at x, then $\partial H(x) = \{\nabla H(x)\}$. We next give the definitions of the semismoothness and the strong semismoothness.

Definition 2. A directionally differentiable and locally Lipschitzian function $H : \mathbf{R}^n \to \mathbf{R}^m$ is said to be semismooth at x if

$$H'(x; d) - V^\top d = o(\|d\|) \tag{15}$$

for any sufficiently small $d \in \mathbf{R}^n \setminus \{0\}$ and $V \in \partial H(x + d)$, where

$$H'(x; d) := \lim_{\tau \to +0} \frac{H(x + \tau d) - H(x)}{\tau} \tag{16}$$

is the directional derivative of H at x along the direction d. In particular, if $o(\|d\|)$ can be replaced by $O(\|d\|^2)$, then function H is said to be strongly semismooth.

It is known that if H is (strongly) semismooth, then

$$\left\|H'(x; d) - (H(x + d) - H(x))\right\| = o(\|d\|)\left(O\left(\|d\|^2\right)\right) \tag{17}$$

holds (see [18], e.g.).

The definitions below for a function can be found in [10, 13, 19].

Definition 3 (see [10, 13]). A function $F = (F^1, \ldots, F^m)$ with $F^i : \mathbf{R}^n \to \mathbf{R}^{n_i}$ is said to have

(a) the Cartesian P_0-property, if for any $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m}$ with $x \neq y$, there exists an index $\nu \in \{1, \ldots, m\}$ such that $x^\nu \neq y^\nu$ and $\langle x^\nu - y^\nu, F^\nu(x) - F^\nu(y)\rangle \geq 0$;

(b) the uniform Cartesian P-property, if there exists a constant $\rho > 0$ such that, for any $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m}$, there exists an index $\nu \in \{1, \ldots, m\}$ such that $\langle x^\nu - y^\nu, F^\nu(x) - F^\nu(y)\rangle \geq \rho \|x - y\|^2$.

By the definitions, it is clear that the Cartesian P-property implies the Cartesian P_0-property. Definition 3 is associated with SOCCP (3), while the following definitions are associated with SOCCP (5).

Definition 4 (see [19]). Let $F : \mathbf{R}^n \to \mathbf{R}^n$ and $G : \mathbf{R}^n \to \mathbf{R}^n$ be functions such that $F = (F^1, \ldots, F^m)$, $G = (G^1, \ldots, G^m)$ with $F^i : \mathbf{R}^n \to \mathbf{R}^{n_i}$ and $G^i : \mathbf{R}^n \to \mathbf{R}^{n_i}$. Then, F and G are said to have

(a) the joint uniform Jordan P-property, if there exists a constant $\rho > 0$ such that

$$\lambda_2\left((F(x) - F(y)) \circ (G(x) - G(y))\right) \geq \rho \|x - y\|^2 \quad \text{for any } x, y \in \mathbf{R}^n; \tag{18}$$

(b) the joint Cartesian weak coerciveness, if there exists a vector $\bar{x} \in \mathbf{R}^n$ such that

$$\lim_{\|x\| \to \infty} \max_{1 \leq i \leq m} \frac{\left\langle F^i(x), G^i(x) - G^i(\bar{x})\right\rangle}{\|x - \bar{x}\|} = +\infty. \tag{19}$$

Next we recall the concept of linear growth of a function, which is weaker than the global Lipschitz continuity.

Definition 5 (see [19]). A function $F : \mathbf{R}^n \to \mathbf{R}^n$ is said to have linear growth, if there exists a constant $C > 0$ such that $\|F(x)\| \leq \|F(0)\| + C\|x\|$ for any $x \in \mathbf{R}^n$.

The following definitions for a matrix are originally given in [8], which is a generalization of the mixed P_0-property [1].

Definition 6 (see [8]). Let $M \in \mathbf{R}^{(n+\ell)\times(2n+\ell)}$ be a matrix partitioned as follows:

$$M = [M_1 \ M_2 \ M_3], \tag{20}$$

where $M_1, M_2 \in \mathbf{R}^{(n+\ell)\times n}$ and $M_3 \in \mathbf{R}^{(n+\ell)\times \ell}$. Then, M is said to have

(a) the Cartesian mixed P_0-property, if the following statements hold:

(1) M_3 has full column rank;
(2) for any $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m}$ with $(x, y) \neq 0$ and $p \in \mathbf{R}^\ell$ such that $M_1x + M_2y + M_3p = 0$, there exists an index $\nu \in \{1, \ldots, m\}$ such that $(x^\nu, y^\nu) \neq 0$ and $\langle x^\nu, y^\nu \rangle \geq 0$;

(b) the Cartesian mixed P-property, if (1) of (a) and the following statement hold:

(2′) for any $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m}$ with $(x, y) \neq 0$ and $p \in \mathbf{R}^\ell$ such that $M_1x+M_2y+M_3p = 0$, there exists an index $\nu \in \{1, \ldots, m\}$ such that $\langle x^\nu, y^\nu \rangle > 0$.

In the case $\ell = 0$ (i.e., M_3 is vacuous) and $M = [M_1 \ \ -I]$, M has the Cartesian mixed P_0 (P)-property if and only if M_1 has the Cartesian P_0 (P)-property (see [10, 13], e.g.). By the definitions, it is clear that the Cartesian mixed P-property implies the Cartesian mixed P_0-property. Moreover, when $n_1 = \cdots = n_m = 1$, the Cartesian mixed P_0-property reduces to the mixed P_0-property (see [1, page 1013]).

We now introduce the Cartesian mixed Jordan P_0 (P)-property.

Definition 7. Let $M \in \mathbf{R}^{(n+\ell)\times(2n+\ell)}$ be a matrix partitioned as follows:

$$M = [M_1 \ M_2 \ M_3], \tag{21}$$

where $M_1, M_2 \in \mathbf{R}^{(n+\ell)\times n}$ and $M_3 \in \mathbf{R}^{(n+\ell)\times \ell}$. Then, M is said to have

(a) the Cartesian mixed Jordan P_0-property, if the following statements hold:

(1) M_3 has full column rank;
(2) for any $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m}$ with $(x, y) \neq 0$ and $p \in \mathbf{R}^\ell$ such that $M_1x+M_2y+M_3p = 0$, there exists an index $\nu \in \{1, \ldots, m\}$ such that $(x^\nu, y^\nu) \neq 0$ and $x^\nu \circ y^\nu \succeq 0$;

(b) the Cartesian mixed Jordan P-property, if (1) of (a) and the following statement hold:

(2′) for any $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m}$ with $(x, y) \neq 0$ and $p \in \mathbf{R}^\ell$ such that $M_1x+M_2y+M_3p = 0$, there exists an index $\nu \in \{1, \ldots, m\}$ such that $x^\nu \circ y^\nu \succ 0$.

Note that the relation $x^\nu \circ y^\nu \succeq 0$ (resp., $x^\nu \circ y^\nu \succ 0$) can be rewritten as $\lambda_1(x^\nu \circ y^\nu) \geq 0$ (resp., $\lambda_1(x^\nu \circ y^\nu) > 0$). By the definitions, it is clear that the Cartesian mixed Jordan P-property implies the Cartesian mixed Jordan P_0-property, and that the Cartesian mixed Jordan P_0 (P)-property implies the Cartesian mixed P_0 (P)-property. Similar to the Cartesian mixed P_0-property, in the case $n_1 = \cdots = n_m = 1$, the Cartesian mixed Jordan P_0-property also reduces to the mixed P_0-property.

3. SOC C-Function and Its Smoothing Function

In this section, we introduce the SOC C-function and its smoothing function. In Section 3.1, we give the concept of the SOC C-function to transform the SOCCP into a system of equations. We focus on the Fischer-Burmeister function as an SOC C-function and review some properties of the smoothed Fischer-Burmeister function in Section 3.2.

3.1. SOC C-Function. First, we recall the concept of the SOC C-function.

Definition 8. A function $\phi : \mathbf{R}^r \times \mathbf{R}^r \to \mathbf{R}^r$ is said to be an SOC complementarity (C-) function, if the following holds:

$$\phi(x, y) = 0 \Longleftrightarrow x \in \mathscr{K}^r, \ \ y \in \mathscr{K}^r, \ \ \langle x, y \rangle = 0. \tag{22}$$

Let $\widehat{\phi} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$ be defined as

$$\widehat{\phi}(x, y) := \begin{bmatrix} \phi^1(x^1, y^1) \\ \vdots \\ \phi^m(x^m, y^m) \end{bmatrix}, \tag{23}$$

where x and $y \in \mathbf{R}^n$ are divided as $x = (x^1, \ldots, x^m)$ and $y = (y^1, \ldots, y^m)$ with $x^i, y^i \in \mathbf{R}^{n_i}$, $i = 1, \ldots, m$, and $\phi^i : \mathbf{R}^{n_i} \times \mathbf{R}^{n_i} \to \mathbf{R}^{n_i}$ are SOC C-functions. Then it follows from (22) that

$$\widehat{\phi}(x, y) = 0 \Longleftrightarrow x \in \mathscr{K}, \ \ y \in \mathscr{K}, \ \ \langle x, y \rangle = 0. \tag{24}$$

Accordingly, SOCCP (1) is reformulated as a system of equations $\widehat{H}(x, y, p) = 0$, where $\widehat{H} : \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^\ell \to \mathbf{R}^{2n+\ell}$ is defined by

$$\widehat{H}(x, y, p) := \begin{bmatrix} \widehat{\phi}(x, y) \\ F(x, y, p) \end{bmatrix}. \tag{25}$$

Moreover, we also give a merit function $\widehat{\Psi} : \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^\ell \to \mathbf{R}$ defined by

$$\begin{aligned} \widehat{\Psi}(x, y, p) &:= \frac{1}{2}\left\|\widehat{H}(x, y)\right\|^2 \\ &= \frac{1}{2}\left\|\widehat{\phi}(x, y)\right\|^2 + \frac{1}{2}\left\|F(x, y, p)\right\|^2. \end{aligned} \tag{26}$$

Note that $\widehat{\Psi}(x,y,p) \geq 0$ for any $(x,y,p) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^\ell$, and that $\widehat{\Psi}(x,y,p) = 0$ if and only if (x,y,p) is a solution of SOCCP (1).

There are many kinds of SOC C-functions. The natural residual function $\phi^i_{\mathrm{NR}} : \mathbf{R}^{n_i} \times \mathbf{R}^{n_i} \to \mathbf{R}^{n_i}$ and the Fischer-Burmeister function $\phi^i_{\mathrm{FB}} : \mathbf{R}^{n_i} \times \mathbf{R}^{n_i} \to \mathbf{R}^{n_i}$ are respectively defined by

$$\phi^i_{\mathrm{NR}}(x^i, y^i) := x^i - [x^i - y^i]_+, \tag{27}$$

$$\phi^i_{\mathrm{FB}}(x^i, y^i) := x^i + y^i - \left((x^i)^2 + (y^i)^2\right)^{1/2}, \tag{28}$$

where $[z]_+$ denotes the projection of z onto the SOC $\mathcal{K}^{n_i}$. Fukushima et al. [5] showed that (22) holds for functions ϕ^i_{NR} and ϕ^i_{FB}. Chen et al. [7] and Hayashi et al. [6] proposed methods for solving SOCCP based on the natural residual function (27), whereas Narushima et al. [8] proposed methods for solving SOCCP based on the Fischer-Burmeister function (28).

In what follows, functions ϕ_{FB}, H_{FB}, and Ψ_{FB} denote $\widehat{\phi}$, $\widehat{H}$, and $\widehat{\Psi}$ with ϕ^i_{FB}, respectively. Also, functions ϕ_{NR}, H_{NR}, and Ψ_{NR} denote $\widehat{\phi}$, $\widehat{H}$, and $\widehat{\Psi}$ with ϕ^i_{NR}, respectively.

Recently, Bi et al. [20] showed the following inequality:

$$\begin{aligned}(2-\sqrt{2})&\left\|\phi^i_{\mathrm{NR}}(x^i,y^i)\right\| \\ &\leq \left\|\phi^i_{\mathrm{FB}}(x^i,y^i)\right\| \leq (2+\sqrt{2})\left\|\phi^i_{\mathrm{NR}}(x^i,y^i)\right\| \\ &\qquad \text{for any } x^i, y^i \in \mathbf{R}^{n_i}.\end{aligned} \tag{29}$$

We see from (29) that the level-boundedness of Ψ_{FB} is equivalent to that of Ψ_{NR}.

3.2. Smoothed FB Function and Its Properties. In this section, we consider the smoothing function associated with the Fischer-Burmeister function and give its properties and Jacobian matrix.

Since H_{FB} is not differentiable in general, we cannot apply conventional methods such as Newton's method or Newton-based methods. We therefore consider the smoothed Fischer-Burmeister function $\phi_t : \mathbf{R}^{2n} \to \mathbf{R}^n$, which was originally proposed by Kanzow [21] for solving NCP and generalized by Fukushima et al. [5] to SOCCP. Let $\phi^i_t : \mathbf{R}^{2n_i} \to \mathbf{R}^{n_i}$ be defined by

$$\phi^i_t(x,y) := x^i + y^i - \left(2t^2 e^i + (x^i)^2 + (y^i)^2\right)^{1/2} \tag{30}$$

for each $i = 1,\dots,m$, where t is a smoothing parameter and $e^i := (1,0,\dots,0)^\top \in \mathbf{R}^{n_i}$. Then, the smoothed Fischer-Burmeister function $\phi_t : \mathbf{R}^{2n} \to \mathbf{R}^n$ is defined as

$$\phi_t(x,y) := \begin{bmatrix}\phi^1_t(x^1,y^1) \\ \vdots \\ \phi^m_t(x^m,y^m)\end{bmatrix}. \tag{31}$$

Also, the smoothing function $H_t : \mathbf{R}^{2n+\ell} \to \mathbf{R}^{2n+\ell}$ and the merit function $\Psi_t : \mathbf{R}^{2n+\ell} \to \mathbf{R}$ are defined as

$$H_t(x,y,p) := \begin{bmatrix}\phi_t(x,y) \\ F(x,y,p)\end{bmatrix}, \tag{32}$$

$$\begin{aligned}\Psi_t(x,y,p) &:= \frac{1}{2}\left\|H_t(x,y,p)\right\|^2 \\ &= \frac{1}{2}\left\|\phi_t(x,y)\right\|^2 + \frac{1}{2}\left\|F(x,y,p)\right\|^2,\end{aligned} \tag{33}$$

respectively. Clearly, $\phi_0(x,y) \equiv \phi_{\mathrm{FB}}(x,y)$, and so $H_0(x,y,p) \equiv H_{\mathrm{FB}}(x,y,p)$ and $\Psi_0(x,y,p) \equiv \Psi_{\mathrm{FB}}(x,y,p)$. We note that

$$\left\|\phi^i_t(x^i,y^i) - \phi^i_{\mathrm{FB}}(x,y)\right\| \leq \sqrt{2}\,|t| \tag{34}$$

holds for any $t \in \mathbf{R}$ and $(x^i, y^i) \in \mathbf{R}^{2n_i}$ (see [5] or [22]). From definition (32) of H_t and (34), it follows that

$$\begin{aligned}&\left\|H_t(x,y,p) - H_{\mathrm{FB}}(x,y,p)\right\| \\ &\quad = \left\|\phi_t(x,y) - \phi_{\mathrm{FB}}(x,y)\right\| \\ &\quad = \left(\sum_{i=1}^m \left\|\phi^i_t(x^i,y^i) - \phi^i_{\mathrm{FB}}(x^i,y^i)\right\|^2\right)^{1/2} \\ &\quad \leq \sqrt{2m}\,|t|\end{aligned} \tag{35}$$

for any $t \in \mathbf{R}$ and $(x,y,p) \in \mathbf{R}^{2n+\ell}$.

In what follows, we write $x^i = ((x^i)_1, (x^i)_2) \in \mathbf{R} \times \mathbf{R}^{n_i - 1}$ for any vector $x^i \in \mathbf{R}^{n_i}$. Moreover, for convenience, we use the following notation. For any x^i, $y^i \in \mathbf{R}^{n_i}$ and any $t \in \mathbf{R}$, we write $z^i = (x^i, y^i) \in \mathbf{R}^{2n_i}$ and define the functions w^i_t, $u^i_t : \mathbf{R}^{2n_i} \to \mathbf{R} \times \mathbf{R}^{n_i - 1}$ by

$$\begin{aligned}w^i_t &= \left((w^i_t)_1, (w^i_t)_2\right) := 2t^2 e^i + (x^i)^2 + (y^i)^2, \\ u^i_t &= \left((u^i_t)_1, (u^i_t)_2\right) := \left(2t^2 e^i + (x^i)^2 + (y^i)^2\right)^{1/2}.\end{aligned} \tag{36}$$

Furthermore, we drop the subscript for $t = 0$ for simplicity, and thus,

$$\begin{aligned}w^i &= \left((w^i)_1, (w^i)_2\right) := (x^i)^2 + (y^i)^2, \\ u^i &= \left((u^i)_1, (u^i)_2\right) := \left((x^i)^2 + (y^i)^2\right)^{1/2}.\end{aligned} \tag{37}$$

Direct calculation yields

$$\begin{aligned}(w^i_t)_1 &= 2t^2 + \left\|x^i\right\|^2 + \left\|y^i\right\|^2 = 2t^2 + (w^i)_1, \\ (w^i_t)_2 &= 2\left((x^i)_1 (x^i)_2 + (y^i)_1 (y^i)_2\right) = (w^i)_2.\end{aligned} \tag{38}$$

Note that $(w_t^i)_2$ is actually independent of t, so that hereafter we will write $w_t^i = ((w_t^i)_1, (w^i)_2)$. We also easily get, for $j = 1, 2$,

$$\lambda_j\left(w_t^i\right) = 2t^2 + \left\|x^i\right\|^2 + \left\|y^i\right\|^2 + 2(-1)^j \left\|\left(x^i\right)_1\left(x^i\right)_2 + \left(y^i\right)_1\left(y^i\right)_2\right\| \qquad (39)$$
$$s^{\{j\}} = \frac{1}{2}\left(1, (-1)^j\left(\overline{w}^i\right)_2\right),$$

where $\lambda_1(w_t^i)$, $\lambda_2(w_t^i)$ and $s^{\{1\}}$, $s^{\{2\}}$ are the spectral values and the associated spectral vectors of w_t^i, respectively, with $(\overline{w}^i)_2 := (w^i)_2/\|(w^i)_2\|$ if $(w^i)_2 \neq 0$, and otherwise, $(\overline{w}^i)_2$ being any vector in $\mathbf{R}^{n_i-1}$ satisfying $\|(\overline{w}^i)_2\| = 1$.

Now we review some propositions needed to establish convergence properties of the smoothing Newton method. The following proposition gives explicit expression of the transposed Jacobian matrix with $t \neq 0$.

Proposition 9 (see [5]). *For any $t \neq 0$, H_t is continuously differentiable on $\mathbf{R}^{2n+\ell}$, and its transposed Jacobian matrix is given by*

$$\nabla H_t(x, y, p) = \begin{bmatrix} \operatorname{diag}\left\{\nabla_{x^i}\phi_t^i\left(x^i, y^i\right)\right\}_{i=1}^m & \nabla_x F(x, y, p) \\ \operatorname{diag}\left\{\nabla_{y^i}\phi_t^i\left(x^i, y^i\right)\right\}_{i=1}^m & \nabla_y F(x, y, p) \\ O & \nabla_p F(x, y, p) \end{bmatrix}, \qquad (40)$$

where $\operatorname{diag}\{D_i\}_{i=1}^m$ *denotes the block-diagonal matrix with block elements $D_i \in \mathbf{R}^{n_i \times n_i}$, and*

$$\nabla_{x^i}\phi_t^i\left(x^i, y^i\right) = I - L_{x^i}L_{u_t^i}^{-1},$$
$$\nabla_{y^i}\phi_t^i\left(x^i, y^i\right) = I - L_{y^i}L_{u_t^i}^{-1}, \qquad (41)$$
$$L_{u_t^i}^{-1} = \begin{bmatrix} b_t^i & -c_t^i\left(\overline{w}^i\right)_2^{\top} \\ -c_t^i\left(\overline{w}^i\right)_2 & a_t^i I + \left(b_t^i - a_t^i\right)\left(\overline{w}^i\right)_2\left(\overline{w}^i\right)_2^{\top} \end{bmatrix}$$

with

$$a_t^i := \frac{2}{\sqrt{\lambda_1\left(w_t^i\right)} + \sqrt{\lambda_2\left(w_t^i\right)}},$$
$$b_t^i := \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_1\left(w_t^i\right)}} + \frac{1}{\sqrt{\lambda_2\left(w_t^i\right)}}\right), \qquad (42)$$
$$c_t^i := \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_1\left(w_t^i\right)}} - \frac{1}{\sqrt{\lambda_2\left(w_t^i\right)}}\right).$$

In order to obtain the Newton step, the nonsingularity of ∇H_t is important. The next proposition establishes the nonsingularity of ∇H_t.

Proposition 10 (see [8]). *Let t be an arbitrary nonzero number and let $(x, y, p) \in \mathbf{R}^{2n+\ell}$ be an arbitrary triple such that the Jacobian matrix $\nabla F(x, y, p)^{\top}$ has the Cartesian mixed P_0-property at (x, y, p), that is, $\nabla_p F(x, y, p)$ satisfies*

$$\operatorname{rank} \nabla_p F(x, y, p) = \ell, \qquad (43)$$

and

$$\left.\begin{matrix} \nabla F(x, y, p)^{\top}(\xi, \eta, \varphi) = 0, \quad (\xi, \eta) \neq 0, \ \varphi \in \mathbf{R}^{\ell} \\ \xi = \left(\xi^1, \ldots, \xi^m\right), \quad \eta = \left(\eta^1, \ldots, \eta^m\right) \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_m} \end{matrix}\right\}$$
$$\Longrightarrow \text{there exists an index } \nu \in \{1, \ldots, m\}$$
$$\text{such that } \left(\xi^{\nu}, \eta^{\nu}\right) \neq 0, \ \left\langle \xi^{\nu}, \eta^{\nu}\right\rangle \geq 0. \qquad (44)$$

Then, the matrix $\nabla H_t(x, y, p)$ given by (40) *is nonsingular.*

The local Lipschitz continuity and the (strong) semismoothness of H_{FB} play a significant role in establishing locally rapid convergence.

Proposition 11. *The function H_{FB} is locally Lipschitzian on $\mathbf{R}^{2n+\ell}$ and, moreover, is semismooth on $\mathbf{R}^{2n+\ell}$. In addition, if ∇F is locally Lipschitzian, then H_{FB} is strongly semismooth on $\mathbf{R}^{2n+\ell}$.*

Proof. It follows from [23, Corollary 3.3] that ϕ_{FB}^i is globally Lipschitzian and strongly semismooth. Since F is a continuously differentiable function, H_{FB} is locally Lipschitzian on $\mathbf{R}^{2n+\ell}$. Also, the (strong) semismoothness of H_{FB} can be easily shown from the strong semismoothness of ϕ_{FB}^i and the (local Lipschitz) continuity of ∇F. □

We define function $\Theta_i : \mathbf{R}^{2n_i} \to \mathbf{R}_+$ by $\Theta_i(x^i, y^i) := \|L_{x^i} \ L_{y^i}\|$. It is easily seen that $\Theta_i(x^i, y^i) = 0$ if and only if $(x^i, y^i) = (0, 0)$. Now we partition $\mathbf{R}^{2n_i}$ as $\mathbf{R}^{2n_i} = \mathscr{Z}_1^i \cup \mathscr{Z}_2^i \cup \{(0, 0)\}$, where

$$\mathscr{Z}_1^i := \left\{z^i = \left(x^i, y^i\right) \in \mathbf{R}^{2n_i} \mid w^i = \left(x^i\right)^2 + \left(y^i\right)^2 \in \operatorname{int}\mathscr{K}^{n_i}\right\},$$
$$\mathscr{Z}_2^i := \left\{z^i = \left(x^i, y^i\right) \in \mathbf{R}^{2n_i} \mid w^i = \left(x^i\right)^2 + \left(y^i\right)^2 \in \operatorname{bd}\mathscr{K}^{n_i}, w^i \neq 0\right\}. \qquad (45)$$

In order to achieve locally rapid convergence of the method, we need to control the parameter t so that the distance between ∇H_t and ∂H_{FB} is sufficiently small. The following proposition is helpful to control the parameter t appropriately.

Proposition 12 (see [22]). *Let (x, y, p) be any point in $\mathbf{R}^{2n+\ell}$. Let $\theta_i(x^i, y^i)$ be any function such that $\Theta_i(x^i, y^i) \le \theta_i(x^i, y^i)$. Let $\delta > 0$ be given. Let $\bar{t} : \mathbf{R}^{2n} \times \mathbf{R}_+ \to \mathbf{R}_+ \cup \{+\infty\}$ be defined by*

$$\bar{t}(x, y, \delta) := \min\left\{\bar{t}_i\left(x^i, y^i, \delta\right) \mid i = 1, \dots, m\right\}, \tag{46}$$

where

$$\bar{t}_i\left(x^i, y^i, \delta\right) := \begin{cases} \dfrac{\lambda_1\left(w^i\right)\delta}{\sqrt{2\left(\theta_i(z^i)^2 - \lambda_1\left(w^i\right)\delta^2\right)}} & \text{if } z^i = \left(x^i, y^i\right) \in \mathscr{Z}_1^i,\ \delta < \dfrac{\theta_i\left(z^i\right)}{\sqrt{\lambda_1\left(w^i\right)}}, \\ \dfrac{\left(w^i\right)_1\delta}{2\sqrt{\theta_i\left(z^i\right)\left(2\theta_i\left(z^i\right) - \delta\sqrt{2(w^i)_1}\right)}} & \text{if } z^i = \left(x^i, y^i\right) \in \mathscr{Z}_2^i,\ \delta < \dfrac{2\theta_i\left(z^i\right)}{\sqrt{2(w^i)_1}}, \\ +\infty & \text{otherwise}. \end{cases} \tag{47}$$

Then, for any $t \in \mathbf{R}$ such that $0 < |t| \le \bar{t}(x, y, \delta)$,

$$\operatorname{dist}\left(\nabla H_t\left(x, y, p\right), \partial H_{FB}\left(x, y, p\right)\right) < \delta, \tag{48}$$

where $\operatorname{dist}(X, S)$ *denotes* $\min\{\|X - Y\| \mid Y \in S\}$.

4. Smoothing Newton Method and Its Convergence Properties

In this section, we first propose an algorithm of the smoothing Newton method based on the Fischer-Burmeister function and its smoothing function. We then prove its global and Q-superlinear (Q-quadratic) convergence.

4.1. Algorithm. We provide the smoothing Newton algorithm based on the Fischer-Burmeister function. In what follows, we write $v^{(k)} = (x^{(k)}, y^{(k)}, p^{(k)})$ and $z^{(k)} = (x^{(k)}, y^{(k)})$ for simplicity.

Algorithm 13.

Step 0. Choose $\eta, \rho \in (0, 1)$, $\overline{\eta} \in (0, \eta]$, $\sigma \in (0, 1/2)$, $r > 1$, $\kappa > 0$, and $\widehat{\kappa} > 0$.

Choose $v^{(0)} = (x^{(0)}, y^{(0)}, p^{(0)}) \in \mathbf{R}^{2n+\ell}$ and $\beta_0 \in (0, \infty)$. Let $t_0 := \| H_{\mathrm{FB}}(v^{(0)}) \|$. Set $k := 0$.

Step 1. If a stopping criterion, such as $\|H_{\mathrm{FB}}(v^{(k)})\| = 0$, is satisfied, then stop.

Step 2.

Step 2.0. Set $\widehat{v}^{(0)} := v^{(k)}$ and $j := 0$.

Step 2.1. Compute $\widehat{d}^{(j)} \in \mathbf{R}^{2n+\ell}$ by solving

$$H_{t_k}\left(\widehat{v}^{(j)}\right) + \nabla H_{t_k}\left(\widehat{v}^{(j)}\right)^{\top}\widehat{d}^{(j)} = 0. \tag{49}$$

Step 2.2. If $\|H_{t_k}(\widehat{v}^{(j)} + \widehat{d}^{(j)})\| \le \beta_k$, then let $v^{(k+1)} := \widehat{v}^{(j)} + \widehat{d}^{(j)}$ and go to Step 3. Otherwise, go to Step 2.3.

Step 2.3. Let l_j be the smallest nonnegative integer l satisfying

$$\Psi_{t_k}\left(\widehat{v}^{(j)} + \rho^l\widehat{d}^{(j)}\right) \le \left(1 - 2\sigma\rho^l\right)\Psi_{t_k}\left(\widehat{v}^{(j)}\right). \tag{50}$$

Let $\tau_j := \rho^{l_j}$ and $\widehat{v}^{(j+1)} := \widehat{v}^{(j)} + \tau_j\widehat{d}^{(j)}$.

Step 2.4. If

$$\left\|H_{t_k}\left(\widehat{v}^{(j+1)}\right)\right\| \le \beta_k, \tag{51}$$

then let $v^{(k+1)} := \widehat{v}^{(j+1)}$ and go to Step 3. Otherwise, set $j := j + 1$ and go back to Step 2.1.

Step 3. Update the parameters as follows:

$$\begin{aligned} \delta_{k+1} &:= \left\|H_{\mathrm{FB}}\left(v^{(k+1)}\right)\right\|, \\ t_{k+1} &:= \min\left\{\kappa\delta_{k+1}^r, t_0\overline{\eta}^{k+1}, \bar{t}\left(z^{(k+1)}, \widehat{\kappa}\delta_{k+1}\right)\right\}, \\ \beta_{k+1} &:= \beta_0\eta^{k+1}. \end{aligned} \tag{52}$$

Step 4. Set $k := k + 1$. Go back to Step 1.

Note that the proposed algorithm consists of the outer iteration steps and the inner iteration steps. Step 2 is the inner iteration steps with the variable $\widehat{v}$ and the counter j, while the outer iteration steps have the variable v and the counter k.

From Step 3 of Algorithm 13 and (48), the following inequality holds:

$$\begin{aligned} &\min\left\{\left\|\nabla H_{t_k}\left(v^{(k)}\right) - Y\right\| \mid Y \in \partial H_{\mathrm{FB}}\left(v^{(k)}\right)\right\} \\ &\quad\le \widehat{\kappa}\left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\|. \end{aligned} \tag{53}$$

Letting v^* be a solution of SOCCP (1), we have $H_{\mathrm{FB}}(v^*) = 0$. Therefore, from (53) and the local Lipschitz continuity of H_{FB}, the following holds:

$$\begin{aligned} &\min\left\{\left\|\nabla H_{t_k}\left(v^{(k)}\right) - Y\right\| \mid Y \in \partial H_{\mathrm{FB}}\left(v^{(k)}\right)\right\} \\ &\quad\le \widehat{\kappa}\left\|H_{\mathrm{FB}}\left(v^{(k)}\right) - H_{\mathrm{FB}}\left(v^*\right)\right\| \\ &\quad= O\left(\left\|v^{(k)} - v^*\right\|\right). \end{aligned} \tag{54}$$

In the rest of this section, we consider convergence properties of Algorithm 13. In Section 4.2, we prove the global convergence of the algorithm, and in Section 4.3, we investigate local behavior of the algorithm. For this purpose, we make the following assumptions.

Assumption 1.

(A1) The solution set $\mathcal{S}$ of SOCCP (1) is nonempty and bounded.

(A2) The function Ψ_{FB} is level-bounded, that is, for any $\overline{v} \in \mathbf{R}^{2n+\ell}$, the level set $\{v \in \mathbf{R}^{2n+\ell} \mid \Psi_{\mathrm{FB}}(v) \le \Psi_{\mathrm{FB}}(\overline{v})\}$ is bounded.

(A3) For any $t > 0$ and $v \in \mathbf{R}^{2n+\ell}$, $\nabla H_t(v)$ is nonsingular.

From Proposition 10, Assumption (A3) holds if $\nabla F(v)^\top$ has the Cartesian mixed P_0-property for any $v \in \mathbf{R}^{2n+\ell}$. The following remarks correspond to SOCCPs (3) and (5).

Remark 14. The case of SOCCP (3). If $\nabla f(x)^\top$ has the Cartesian P_0-property, then $\nabla F(x, y, p)^\top$ with F in (4) has the Cartesian mixed P_0-property and vice versa. Note that $\nabla f(x)^\top$ has the Cartesian P_0-property at any $x \in \mathbf{R}^n$ if f has the Cartesian P_0-property (see [10, 13] for the definition of the Cartesian P_0-property for a matrix).

Remark 15. The case of SOCCP (5). If $\nabla f(p)$ is nonsingular and $(\nabla f(p)^{-1}\nabla g(p))^\top$ has the Cartesian P_0-property, then $\nabla F(x, y, p)^\top$ with F in (6) has the Cartesian mixed P_0-property.

Note that Assumption (A2) is equivalent to the coerciveness in the sense that

$$\lim_{\|v\|\to\infty} \Psi_{\mathrm{FB}}(v) = +\infty. \tag{55}$$

We now give some sufficient conditions for Assumption (A2) in the case of SOCCP (3) or (5).

Lemma 16. *Consider SOCCP* (5). *Let* $\widetilde{\Psi}_{FB} : \mathbf{R}^n \to \mathbf{R}$ *be a function such that* $\widetilde{\Psi}_{FB}(p) := (1/2)\|\phi_{FB}(f(p), g(p))\|^2$ *for any* $p \in \mathbf{R}^n$. *Then* $\widetilde{\Psi}_{FB}$ *is level-bounded if and only if* $\widetilde{\Psi}_{FB}$ *is level-bounded.*

Proof. We use below condition (55) equivalent to the level-boundedness. We first assume that $\widetilde{\Psi}_{\mathrm{FB}}$ is level-bounded and claim that Ψ_{FB} is level-bounded. Suppose the contrary. Then, we can find a sequence $\{v^{(k)}\} = \{(x^{(k)}, y^{(k)}, p^{(k)})\} \subset \mathbf{R}^{3n}$ such that the sequence $\{\Psi_{\mathrm{FB}}(v^{(k)})\}$ is bounded and $\|v^{(k)}\| \to \infty$. If $\{\|p^{(k)}\|\}$ is bounded, then we must have $\|(x^{(k)}, y^{(k)})\| \to \infty$. Thus, from the inequality

$$\begin{aligned}\sqrt{2\Psi_{\mathrm{FB}}(v^{(k)})} &\geq \left\|\left(f\left(p^{(k)}\right) - x^{(k)}, g\left(p^{(k)}\right) - y^{(k)}\right)\right\| \\ &\geq \left\|\left(x^{(k)}, y^{(k)}\right)\right\| - \left\|\left(f\left(p^{(k)}\right), g\left(p^{(k)}\right)\right)\right\|\end{aligned} \tag{56}$$

and from the continuity of f and g, we have $\Psi_{\mathrm{FB}}(v^{(k)}) \to \infty$. Since this is not possible, $\{\|p^{(k)}\|\}$ is unbounded. By taking a subsequence if necessary, we may assume that $\|p^{(k)}\| \to \infty$ as $k \to \infty$. Since ϕ_{FB} is globally Lipschitzian by [23, Corollary 3.3], we have from (33) that, for any $(x, y, p) \in \mathbf{R}^{3n}$,

$$\begin{aligned}&\sqrt{2\widetilde{\Psi}_{\mathrm{FB}}(p)} \\ &= \|\phi_{\mathrm{FB}}(f(p), g(p))\| \\ &\leq \|\phi_{\mathrm{FB}}(f(p), g(p)) - \phi_{\mathrm{FB}}(x, y)\| + \|\phi_{\mathrm{FB}}(x, y)\| \\ &\leq L\|(f(p), g(p)) - (x, y)\| + \|\phi_{\mathrm{FB}}(x, y)\| \\ &\leq \sqrt{L^2+1}\sqrt{\|(f(p) - x, g(p) - y)\|^2 + \|\phi_{\mathrm{FB}}(x, y)\|^2} \\ &= \sqrt{L^2+1}\sqrt{2\Psi_{\mathrm{FB}}(x, y, p)},\end{aligned} \tag{57}$$

where $L > 0$ is a Lipschitz constant. Then, it follows from (57) and the level-boundedness of $\widetilde{\Psi}_{\mathrm{FB}}$ that $\lim_{k\to\infty}\Psi_{\mathrm{FB}}(v^{(k)}) = +\infty$, contradicting the boundedness of $\{\Psi_{\mathrm{FB}}(v^{(k)})\}$. This proves the level-boundedness of Ψ_{FB}.

We next assume that Ψ_{FB} is level-bounded. Let $\{p^{(k)}\} \subset \mathbf{R}^n$ be an arbitrary sequence such that $\|p^{(k)}\| \to \infty$ and let $x^{(k)} := f(p^{(k)})$ and $y^{(k)} := g(p^{(k)})$. Then we have

$$\begin{aligned}&2\Psi_{\mathrm{FB}}\left(x^{(k)}, y^{(k)}, p^{(k)}\right) \\ &= \left\|\phi_{\mathrm{FB}}\left(x^{(k)}, y^{(k)}\right)\right\|^2 + \left\|f\left(p^{(k)}\right) - x^{(k)}\right\|^2 \\ &\quad + \left\|g\left(p^{(k)}\right) - y^{(k)}\right\|^2 \\ &= \left\|\phi_{\mathrm{FB}}\left(f\left(p^{(k)}\right), g\left(p^{(k)}\right)\right)\right\|^2 \\ &= 2\widetilde{\Psi}_{\mathrm{FB}}\left(p^{(k)}\right).\end{aligned} \tag{58}$$

Thus, from $\|(x^{(k)}, y^{(k)}, p^{(k)})\| \geq \|p^{(k)}\| \to \infty$ and the level-boundedness of Ψ_{FB}, we have $\lim_{k\to\infty}\widetilde{\Psi}_{\mathrm{FB}}(p^{(k)}) = +\infty$. Therefore, $\widetilde{\Psi}_{\mathrm{FB}}$ is level-bounded. □

Remark 17. Consider SOCCP (3). Let $\widetilde{\Psi}_{\mathrm{FB}} : \mathbf{R}^n \to \mathbf{R}$ be a function such that $\widetilde{\Psi}_{\mathrm{FB}}(x) := (1/2)\|\phi_{\mathrm{FB}}(x, f(x))\|^2$ for any $x \in \mathbf{R}^n$. Then, in the same way as in Lemma 16, we can show that $\widetilde{\Psi}_{\mathrm{FB}}$ is level-bounded if and only if $\widetilde{\Psi}_{\mathrm{FB}}$ is level-bounded (we have only to consider $g(p) \equiv p$).

We now provide some sufficient conditions under which Assumption (A2) holds.

Proposition 18. *Consider SOCCP* (5). *Assume that f and g have linear growth. Assume further that f and g satisfy one of the following statements:*

(a) *f and g have the joint uniform Jordan P-property;*

(b) *f and g have the joint Cartesian weak coerciveness.*

Then Ψ_{FB} is level-bounded.

Proof. It follows from [19] that $\widetilde{\Psi}_{\mathrm{NR}}(p) := (1/2)\|\phi_{\mathrm{NR}}(f(p), g(p))\|^2$ is level-bounded for each case. Thus from Lemma 16 and (29), we have desired results. □

The following condition was given by Pan and Chen [13] to establish the level-boundedness property of the merit function $\widetilde{\Psi}_{\mathrm{FB}}(x)$ defined in Remark 17.

Condition A. Consider SOCCP (3). For any sequence $\{\xi_k\} \subset \mathbf{R}^n$ satisfying $\|\xi_k\| \to \infty$ with $\xi_k^i \in \mathbf{R}^{n_i}$, if there exists an index $\nu \in \{1, \ldots, m\}$ such that $\{\lambda_1(\xi_k^\nu)\}$ and $\{\lambda_1(f^\nu(\xi_k))\}$ are bounded below, and $\lambda_2(\xi_k^\nu)$, $\lambda_2(f^\nu(\xi_k)) \to \infty$, then

$$\limsup_{k\to\infty}\left\langle \frac{\xi_k^\nu}{\|\xi_k^\nu\|}, \frac{f^\nu(\xi_k)}{\|f^\nu(\xi_k)\|}\right\rangle > 0. \tag{59}$$

Under Condition A, we have the following proposition, which corresponds to Proposition 5.2 of [13].

Proposition 19. *Consider SOCCP* (3). *Assume that f has the uniform Cartesian P-property and satisfies Condition A. Then Ψ_{FB} is level-bounded.*

4.2. Global Convergence. In this section, we show the global convergence of Algorithm 13. We first give the well-definedness of the algorithm.

Lemma 20. *Suppose that Assumption (A3) holds. Let t be any fixed positive number. Every stationary point v_t^* of Ψ_t satisfies $\Psi_t(v_t^*) = 0$.*

Proof. For each stationary point v_t^* of Ψ_t, $\nabla\Psi_t(v_t^*) = \nabla H_t(v_t^*)H(v_t^*) = 0$ holds. Since, from $t > 0$ and Assumption (A3), $\nabla H_t(v_t^*)$ is nonsingular, we have $H_t(v_t^*) = 0$, and thus, $\Psi_t(v_t^*) = 0$. □

It follows from (35) that

$$\begin{aligned}\sqrt{2\Psi_{\mathrm{FB}}(\hat{v})} &\le \|H_t(\hat{v})\| + \|H_{\mathrm{FB}}(\hat{v}) - H_t(\hat{v})\| \\ &\le \sqrt{2\Psi_t(\hat{v})} + \sqrt{2mt}\end{aligned} \tag{60}$$

for any $\hat{v} \in \mathbf{R}^{2n+\ell}$, and hence we have, from Assumption (A2) and (55), that Ψ_t is level-bounded for any fixed $t > 0$. Therefore, there exists at least one stationary point of Ψ_t. Thus from Lemma 20, the system $H_t(v) = 0$ has at least one solution, and hence, there exists a point v satisfying $\|H_{t_k}(v)\| < \beta_k$ in Step 2 at each iteration.

We are now ready to show the well-definedness of the algorithm.

Proposition 21. *Suppose that Assumptions (A2) and (A3) hold. Then Algorithm 13 is well-defined.*

Proof. To establish the well-definedness of Algorithm 13, we only need to prove the well-definedness and the finite termination property of Step 2 at each iteration. Now we fix k and $t_k > 0$. Since $\nabla H_{t_k}(\hat{v}^{(j)})$ is nonsingular for any $\hat{v}^{(j)} \in \mathbf{R}^{2n+\ell}$ by $t_k > 0$ and Assumption (A3), $\hat{d}^{(j)}$ is uniquely determined for any $j \ge 0$. In addition, we have

$$\begin{aligned}&\nabla\Psi_{t_k}\left(\hat{v}^{(j)}\right)^\top \hat{d}^{(j)} \\ &\quad= -H_{t_k}\left(\hat{v}^{(j)}\right)^\top \nabla H_{t_k}\left(\hat{v}^{(j)}\right)^\top \nabla H_{t_k}\left(\hat{v}^{(j)}\right)^{-\top} H_{t_k}\left(\hat{v}^{(j)}\right) \\ &\quad= -\left\|H_{t_k}\left(\hat{v}^{(j)}\right)\right\|^2 \\ &\quad\le 0.\end{aligned} \tag{61}$$

If $\|H_{t_k}(\hat{v}^{(j)} + \hat{d}^{(j)})\| \le \beta_k$, then Step 2 terminates in Step 2.2. If $\|H_{t_k}(\hat{v}^{(j)} + \hat{d}^{(j)})\| > \beta_k$, then integer l satisfying (50) can be found at Step 2.3, because $\hat{d}^{(j)} \ne 0$ and $\nabla\Psi_{t_k}(\hat{v}^{(j)})^\top \hat{d}^{(j)} < 0$. Thus, Step 2 is well-defined at each iteration.

Next we prove the finite termination property of Step 2. To prove by contradiction, we assume that Step 2 never stops and then

$$\left\|H_{t_k}\left(\hat{v}^{(j)}\right)\right\| > \beta_k > 0 \tag{62}$$

holds for all $j \ge 0$. We consider two cases.

(i) The case where there exists a subsequence J such that $\lim_{j\in J, j\to\infty}\tau_j = 0$. From the boundedness of the level set of Ψ_{t_k} at $\hat{v}^{(0)}$ and the line search rule (50), $\{\hat{v}^{(j)}\}$ is bounded. In addition, from the continuous differentiability of Ψ_{t_k} and Assumption (A3), $\{\hat{d}^{(j)}\}$ is also bounded. Thus, there exists a subsequence $J' \subset J$ such that

$$\lim_{j\in J', j\to\infty} \hat{v}^{(j)} = \hat{v}_*, \qquad \lim_{j\in J', j\to\infty} \hat{d}^{(j)} = \hat{d}_*. \tag{63}$$

Now $\tau_j \ne 1$ holds for all sufficiently large $j \in J'$, and hence, we have

$$\frac{\Psi_{t_k}\left(\hat{v}^{(j)} + \rho^{l_{j-1}}\hat{d}^{(j)}\right) - \Psi_{t_k}\left(\hat{v}^{(j)}\right)}{\rho^{l_{j-1}}} > -2\sigma\Psi_{t_k}\left(\hat{v}^{(j)}\right). \tag{64}$$

Passing to the limit $j \to \infty$ with $j \in J'$ on the above inequality and taking (62) into account, we have

$$\nabla\Psi_{t_k}(\hat{v}_*)^\top \hat{d}_* \ge -2\sigma\Psi_{t_k}(\hat{v}_*) > -2\Psi_{t_k}(\hat{v}_*). \tag{65}$$

On the other hand, it follows from (61) that $\nabla\Psi_{t_k}(\hat{v}_*)^\top \hat{d}_* = -2\Psi_{t_k}(\hat{v}_*)$, which contradicts (65).

(ii) The case where there exists $\overline{\tau} > 0$ such that $\tau_j > \overline{\tau}$ for all j. It follows from (50) that

$$\Psi_{t_k}\left(\hat{v}^{(j)}\right) \le (1 - 2\sigma\overline{\tau})^j \Psi_{t_k}\left(\hat{v}^{(0)}\right), \tag{66}$$

which implies $\|H_{t_k}(\hat{v}^{(j)})\| \le \beta_k$ holds for sufficiently large j. This contradicts (62). Therefore, the proof is complete. □

In order to show the global convergence of the proposed method, we recall the mountain pass theorem (see [24], e.g.), which is as follows.

Lemma 22. *Let $\varphi : \mathbf{R}^n \to \mathbf{R}$ be a continuously differentiable and level-bounded function. Let $\mathscr{C} \subset \mathbf{R}^n$ be a nonempty and compact set and let $\overline{\varphi}$ be the minimum value of φ on $\mathrm{bd}\,\mathscr{C}$, that is,*

$$\overline{\varphi} := \min_{x\in \mathrm{bd}\,\mathscr{C}} \varphi(x). \tag{67}$$

Assume that there exist vectors $\xi \in \mathscr{C}$ and $\eta \notin \mathscr{C}$ such that $\varphi(\xi) < \overline{\varphi}$ and $\varphi(\eta) < \overline{\varphi}$. Then, there exists a vector $\zeta \in \mathbf{R}^n$ such that $\nabla\varphi(\zeta) = 0$ and $\varphi(\zeta) \ge \overline{\varphi}$.

By using the mountain pass theorem, we can show the following global convergence property.

Theorem 23. *Suppose that Assumptions (A1)–(A3) hold. Then, any accumulation point of the sequence $\{v^{(k)}\}$ generated by Algorithm 13 is bounded, and hence, at least one accumulation point exists, and any such point is a solution of SOCCP (1).*

Proof. From the choices of t_k and β_k in Step 3 of Algorithm 13, t_k and β_k converge to zero. Since $\beta_k \to 0$, we have

$\lim_{k\to\infty}\|H_{t_k}(v^{(k)})\| = 0$. Thus, it follows from $t_k \to 0$ and (35) that $\lim_{k\to\infty}\|H_{\rm FB}(v^{(k)})\| = 0$. It implies from the continuity of $H_{\rm FB}$ that any accumulation point of the sequence $\{v^{(k)}\}$ is a solution of $H_{\rm FB}(v) = 0$, and hence, it suffices to prove the boundedness of $\{v^{(k)}\}$. To the contrary, we assume $\{v^{(k)}\}$ is unbounded. Then there exist an index set K and a subsequence $\{v^{(k)}\}_{k\in K}$ such that $\lim_{k\in K,k\to\infty}\|v^{(k)}\| = \infty$. Since, by Assumption (A1), the solution set $\mathcal{S}$ is bounded, there exists a compact neighborhood $\mathcal{C}$ of $\mathcal{S}$ such that $\mathcal{S}\subset \operatorname{int}\mathcal{C}$. From the boundedness of $\mathcal{C}$, $v^{(k)} \notin \mathcal{C}$ for all k sufficiently large $k\in K$. In addition, from $\mathcal{S}\subset\operatorname{int}\mathcal{C}$, we have

$$\hat{\varphi} := \min_{v\in \operatorname{bd}\mathcal{C}} \Psi_{\rm FB}(v) > 0, \tag{68}$$

for otherwise, there would exist $v\in\operatorname{bd}\mathcal{C}$ with $\Psi_{\rm FB}(v) = 0$, that is, $v\in\mathcal{S}\cap\operatorname{bd}\mathcal{C}$, which contradicts $\mathcal{S}\subset\operatorname{int}\mathcal{C}$. Since t_k is small enough for all sufficiently large k, it follows from (35) that

$$-\frac{\hat{\varphi}}{2} < \Psi_{t_k}(v) - \Psi_{\rm FB}(v) < \frac{\hat{\varphi}}{2} \tag{69}$$

holds for any $v\in\mathcal{C}$. Now we take $\bar{v}\in\mathcal{S}\subset\mathcal{C}$. Then (69) yields

$$\left|\Psi_{t_k}(\bar{v}) - \Psi_{\rm FB}(\bar{v})\right| = \Psi_{t_k}(\bar{v}) < \frac{\hat{\varphi}}{2}. \tag{70}$$

Letting

$$\tilde{v}^{(k)} \in \arg\min_{v\in\operatorname{bd}\mathcal{C}} \Psi_{t_k}(v), \tag{71}$$

we have from (68) and $\tilde{v}^{(k)}\in\operatorname{bd}\mathcal{C}$ that $\hat{\varphi}\le\Psi_{\rm FB}(\tilde{v}^{(k)})$. Therefore, it follows from (69) and (70) that, for all sufficiently large k,

$$\begin{aligned}\min_{v\in\operatorname{bd}\mathcal{C}}\Psi_{t_k}(v) &= \Psi_{t_k}\left(\tilde{v}^{(k)}\right)\\ &> -\frac{\hat{\varphi}}{2} + \Psi_{\rm FB}\left(\tilde{v}^{(k)}\right)\\ &\ge -\frac{\hat{\varphi}}{2} + \hat{\varphi} = \frac{\hat{\varphi}}{2}.\end{aligned} \tag{72}$$

On the other hand, since $0\le \|H_{t_k}(v^{(k+1)})\|\le\beta_k$ and $\beta_k\to 0$, we get

$$\Psi_{t_k}\left(v^{(k+1)}\right) < \frac{\hat{\varphi}}{2} \tag{73}$$

for all k sufficiently large.

Now we choose sufficiently large $\hat{k}$ satisfying all the above arguments with $\hat{k}+1\in K$ and apply Lemma 22 with

$$\begin{aligned}&\varphi = \Psi_{t_{\hat{k}}}, \qquad \xi = \bar{v}\in\mathcal{C},\\ &\eta = v^{(\hat{k}+1)} \notin \mathcal{C}, \qquad \bar{\varphi} = \min_{v\in\operatorname{bd}\mathcal{C}}\Psi_{t_{\hat{k}}}(v) > \frac{\hat{\varphi}}{2}.\end{aligned} \tag{74}$$

Then there exists $\zeta\in\mathbf{R}^{2n+\ell}$ satisfying

$$\nabla\Psi_{t_{\hat{k}}}(\zeta) = 0, \qquad \Psi_{t_{\hat{k}}}(\zeta) \ge \bar{\varphi} > \frac{\hat{\varphi}}{2} > 0, \tag{75}$$

which contradicts Lemma 20, and therefore the proof is complete. □

4.3. Local Q-Superlinear and Q-Quadratic Convergence. In Section 4.2, we have shown that the sequence $\{v^{(k)}\}$ is bounded, and any accumulation point of $\{v^{(k)}\}$ is a solution of SOCCP (1). In this section, we prove that $\{v^{(k)}\}$ is superlinearly convergent, or more strongly, quadratically convergent if ∇F is locally Lipschitzian. In order to establish the superlinear (quadratic) convergence of the algorithm, we need an assumption that every accumulation point of $\{\nabla H_{t_k}(v^{(k)})\}$ is nonsingular. We first consider a sufficient condition for this assumption to hold.

Let $\{v^{(k)}\}$ and $\{t_k\}$ be the sequences generated by Algorithm 13, and let $v^* = (x^*, y^*, p^*)$ be any accumulation point of $\{v^{(k)}\}$. Then, by Theorem 23, v^* is a solution of SOCCP (1). We call the following condition nondegeneracy of a solution of the SOCCP (see also [13, 25]).

Definition 24. Let $v^* = (x^*, y^*, p^*)\in\mathbf{R}^{2n+\ell}$ be a solution of SOCCP (1) with $x^* = (x^{*1},\dots,x^{*m})$, $y^* = (y^{*1},\dots,y^{*m})\in\mathbf{R}^{n_1}\times\cdots\times\mathbf{R}^{n_m}$. Then we say that v^* is nondegenerate if $x^*+y^*\succ 0$, or equivalently, $x^{*i}+y^{*i}\succ 0$ for all $i=1,\dots,m$.

For a nondegenerate solution, we have the next lemma.

Lemma 25. *Let $v^* = (x^*, y^*, p^*)\in\mathbf{R}^{2n+\ell}$ be a nondegenerate solution of SOCCP* (1)*, and put $z^i = (x^i, y^i)$, $z^{*i} = (x^{*i}, y^{*i})\in\mathbf{R}^{2n_i}$ for $i = 1,\dots,m$. Let $t\in\mathbf{R}$ be a nonzero number. Then, for each i, the following holds:*

$$\begin{aligned}\lim_{(t,z^i)\to(0,z^{*i})}\nabla_{x^i}\phi_t^i\left(x^i,y^i\right) &= I - L_{x^{*i}}L_{u^{*i}}^{-1},\\ \lim_{(t,z^i)\to(0,z^{*i})}\nabla_{y^i}\phi_t^i\left(x^i,y^i\right) &= I - L_{y^{*i}}L_{u^{*i}}^{-1},\end{aligned} \tag{76}$$

where

$$L_{u^{*i}}^{-1} = \begin{bmatrix} b^{*i} & -c^{*i}\left(\overline{w}^{*i}\right)_2^\top \\ -c^{*i}\left(\overline{w}^{*i}\right)_2 & a^{*i}I + \left(b^{*i}-a^{*i}\right)\left(\overline{w}^{*i}\right)_2\left(\overline{w}^{*i}\right)_2^\top\end{bmatrix} \tag{77}$$

with

$$\begin{aligned}a^{*i} &:= \frac{2}{\sqrt{\lambda_1\left(w^{*i}\right)}+\sqrt{\lambda_2\left(w^{*i}\right)}},\\ b^{*i} &:= \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_1\left(w^{*i}\right)}}+\frac{1}{\sqrt{\lambda_2\left(w^{*i}\right)}}\right),\\ c^{*i} &:= \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_1\left(w^{*i}\right)}}-\frac{1}{\sqrt{\lambda_2\left(w^{*i}\right)}}\right).\end{aligned} \tag{78}$$

Here w^{*i} and u^{*i} are defined by (37) *with x^{*i} and y^{*i}. We also write $w^{*i} = ((w^{*i})_1, (w^{*i})_2)\in\mathbf{R}\times\mathbf{R}^{n_i-1}$, and set $(\overline{w}^{*i})_2 := (w^{*i})_2/\|(w^{*i})_2\|$ if $(w^{*i})_2\ne 0$, and otherwise, set $(\overline{w}^{*i})_2$ to any vector in $\mathbf{R}^{n_i-1}$ satisfying $\|(\overline{w}^{*i})_2\| = 1$.*

Proof. Since v^* is a solution of SOCCP (1), it follows from [5, Propsition 4.2] that $x^{*i} + y^{*i} - ((x^{*i})^2 + (y^{*i})^2)^{1/2} = 0$ for all $i = 1, \ldots, m$. Hence, from the nondegeneracy of v^*, we have

$$u^{*i} = \left(\left(x^{*i}\right)^2 + \left(y^{*i}\right)^2\right)^{1/2} = x^{*i} + y^{*i} \succ 0. \quad (79)$$

By Property 1(c), this implies that $L_{u^{*i}}$ is nonsingular. In order to prove this lemma, it suffices to show

$$\lim_{(t,z^i)\to(0,z^{*i})} L_{u_t^i}^{-1} = L_{u^{*i}}^{-1}. \quad (80)$$

Since (36) yields $\lim_{(t,z^i)\to(0,z^{*i})} u_t^i = u^{*i}$, (80) follows from Property 1(c). □

The following proposition gives a sufficient condition for the nonsingularity of accumulation points of $\{\nabla H_{t_k}(v^{(k)})\}$.

Proposition 26. *Suppose that Assumptions (A1)–(A3) hold. Let $\{v^{(k)}\}$ be a sequence generated by Algorithm 13 and let v^* be any accumulation point of it. Moreover, assume that v^* is nondegenerate and the Jacobian matrix $\nabla F(v^*)^\top$ has the Cartesian mixed Jordan P-property, that is,* rank $\nabla_p F(v^*) = \ell$ *and*

$$\begin{gathered}\nabla F(v^*)^\top(\xi,\eta,\varphi) = 0, \quad (\xi,\eta) \neq 0,\ \varphi \in \mathbf{R}^\ell \\ \xi = \left(\xi^1,\ldots,\xi^m\right), \quad \eta = \left(\eta^1,\ldots,\eta^m\right) \in \mathbf{R}^{n_1}\times\cdots\times\mathbf{R}^{n_m} \\ \Longrightarrow \text{there exists an index } \nu \text{ such that } \xi^\nu \circ \eta^\nu \succ 0.\end{gathered} \quad (81)$$

Then, every accumulation point of $\{\nabla H_{t_k}(v^{(k)})\}$ is nonsingular.

Proof. By Theorem 23, the sequence $\{v^{(k)}\}$ is bounded and has at least one accumulation point v^*. Hence, we may assume that $\{v^{(k)}\}$ converges to v^* without loss of generality. It follows from Lemma 25 and $t_k \to 0$ that any accumulation point of $\{\nabla H_{t_k}(v^{(k)})\}$, say J_0, is given in the following form:

$$J_0 := \begin{bmatrix} \operatorname{diag}\left\{I - L_{x^{*i}}L_{u^{*i}}^{-1}\right\}_{i=1}^m & \nabla_x F\left(v^*\right) \\ \operatorname{diag}\left\{I - L_{y^{*i}}L_{u^{*i}}^{-1}\right\}_{i=1}^m & \nabla_y F\left(v^*\right) \\ O & \nabla_p F\left(v^*\right)\end{bmatrix}. \quad (82)$$

In order to prove that J_0 is nonsingular, suppose that $J_0^\top(\xi,\eta,\varphi) = 0$, where $(\xi,\eta,\varphi) \in \mathbf{R}^{2n+\ell}$. We will show that $(\xi,\eta,\varphi) = 0$. It follows from (82) that

$$\nabla F(v^*)^\top(\xi,\eta,\varphi) = 0, \quad (83)$$

$$\left(I - L_{u^{*i}}^{-1}L_{x^{*i}}\right)\xi^i + \left(I - L_{u^{*i}}^{-1}L_{y^{*i}}\right)\eta^i = 0, \quad i = 1,\ldots,m, \quad (84)$$

where $\xi = (\xi^1,\ldots,\xi^m)$, $\eta = (\eta^1,\ldots,\eta^m) \in \mathbf{R}^{n_1}\times\cdots\times\mathbf{R}^{n_m}$. Multiplying both sides of the above equations by $L_{u^{*i}}$ from the left-hand side, we get

$$\begin{aligned}\left(L_{u^{*i}} - L_{x^{*i}}\right)\xi^i + \left(L_{u^{*i}} - L_{y^{*i}}\right)\eta^i &= L_{u^{*i}-x^{*i}}\xi^i + L_{u^{*i}-y^{*i}}\eta^i \\ &= L_{y^{*i}}\xi^i + L_{x^{*i}}\eta^i \\ &= 0\end{aligned} \quad (85)$$

for all $i = 1,\ldots,m$, where the second equality uses the fact $u^{*i} = x^{*i} + y^{*i}$ (see (79)). Suppose on the contrary that $(\xi,\eta) \neq 0$. Then from (83) and the assumption (81), we have that

$$\xi^\nu \circ \eta^\nu \succ 0 \quad (86)$$

for some $\nu \in \{1,\ldots,m\}$. Since $x^{*\nu}$, $y^{*\nu} \succeq 0$, by Property 1(b), we have $L_{x^{*\nu}} \succeq O$ and $L_{y^{*\nu}} \succeq O$. By using Property 1(a), (85) can be rewritten equivalently as

$$L_{y^{*\nu}}\xi^\nu + x^{*\nu}\circ\eta^\nu = 0 \quad \text{or} \quad y^{*\nu}\circ\xi^\nu + L_{x^{*\nu}}\eta^\nu = 0. \quad (87)$$

Multiplying both sides of the first equation in (87) by $(\xi^\nu)^\top$ from the left, we have

$$\begin{aligned}&(\xi^\nu)^\top L_{y^{*\nu}}\xi^\nu + \langle\xi^\nu, x^{*\nu}\circ\eta^\nu\rangle \\ &\quad = (\xi^\nu)^\top L_{y^{*\nu}}\xi^\nu + \langle x^{*\nu}, \xi^\nu\circ\eta^\nu\rangle = 0.\end{aligned} \quad (88)$$

Since $L_{y^{*\nu}}$ is positive semidefinite, we have $\langle x^{*\nu}, \xi^\nu\circ\eta^\nu\rangle \le 0$. Similarly, multiplying both sides of the second equation in (87) by $(\eta^\nu)^\top$ from the left, we have $\langle y^{*\nu}, \xi^\nu\circ\eta^\nu\rangle \le 0$. Adding these two inequalities yields

$$\langle x^{*\nu} + y^{*\nu}, \xi^\nu\circ\eta^\nu\rangle \le 0. \quad (89)$$

On the other hand, by the nondegeneracy of v^*, we have $x^{*\nu} + y^{*\nu} \succ 0$. This together with (86) yields

$$\langle x^{*\nu} + y^{*\nu}, \xi^\nu\circ\eta^\nu\rangle > 0, \quad (90)$$

which contradicts (89), and hence, we must have $(\xi,\eta) = 0$. Then, since the matrix $\nabla_p F(v^*)^\top \in \mathbf{R}^{(n+\ell)\times\ell}$ has full column rank, we also have from (83) that $\varphi = 0$. Therefore, J_0 is nonsingular. □

Next, we show the local convergence properties of the sequence $\{v^{(k)}\}$ generated by Algorithm 13. The following lemma plays a key role in proving such properties.

Lemma 27. *Suppose that Assumptions (A1)–(A3) hold. Let $\{v^{(k)}\}$ be a sequence generated by Algorithm 13 and let v^* be any accumulation point of it. In addition, assume that every accumulation point of $\{\nabla H_{t_k}(v^{(k)})\}$ is nonsingular. Then, for $v^{(k)}$ sufficiently close to v^*,*

$$\left\|v^{(k)} + d^{(k)} - v^*\right\| = o\left(\left\|v^{(k)} - v^*\right\|\right) \quad (91)$$

holds, where $d^{(k)}$ *is defined by* $-\nabla H_{t_k}(\nu^{(k)})^{-\top} H_{t_k}(\nu^{(k)})$. *Moreover, if* ∇F *is locally Lipschitzian and* $r \geq 2$, *then*

$$\left\|\nu^{(k)} + d^{(k)} - \nu^*\right\| = O\left(\left\|\nu^{(k)} - \nu^*\right\|^2\right) \tag{92}$$

holds.

Proof. From Theorem 23, $\{\nu^{(k)}\}$ is bounded, and hence, there exists at least one accumulation point, and any such point ν^* satisfies $H_{\mathrm{FB}}(\nu^*) = 0$. Since every accumulation point of $\{\nabla H_{t_k}(\nu^{(k)})\}$ is nonsingular, we have, from Assumption (A3), that there exists a constant $c_1 > 0$ such that

$$\left\|\nabla H_{t_k}\left(\nu^{(k)}\right)^{-\top}\right\| \leq c_1 \tag{93}$$

for all k. Let $V_k \in \partial H_{\mathrm{FB}}(\nu^{(k)})$ such that

$$V_k \in \underset{V \in \partial H_{\mathrm{FB}}(\nu^{(k)})}{\arg\min} \left\|\nabla H_{t_k}\left(\nu^{(k)}\right) - V\right\|. \tag{94}$$

Note that V_k exists for all k, because $\partial H_{\mathrm{FB}}(\nu^{(k)})$ is compact [17, page 70]. We have from (93) and $H_{\mathrm{FB}}(\nu^*) = 0$ that

$$\begin{aligned}
&\left\|\nu^{(k)} + d^{(k)} - \nu^*\right\| \\
&\quad = \left\|\nu^{(k)} - \nabla H_{t_k}\left(\nu^{(k)}\right)^{-\top} H_{t_k}\left(\nu^{(k)}\right) - \nu^*\right\| \\
&\quad \leq \left\|\nabla H_{t_k}\left(\nu^{(k)}\right)^{-\top}\right\| \left\|\nabla H_{t_k}\left(\nu^{(k)}\right)^{\top}\left(\nu^{(k)} - \nu^*\right) - H_{t_k}\left(\nu^{(k)}\right)\right\| \\
&\quad \leq c_1 \left\{\left\|\left(\nabla H_{t_k}\left(\nu^{(k)}\right) - V_k\right)^{\top}\left(\nu^{(k)} - \nu^*\right)\right\|\right. \\
&\qquad + \left\|V_k^{\top}\left(\nu^{(k)} - \nu^*\right) - H_{\mathrm{FB}}'\left(\nu^*; \nu^{(k)} - \nu^*\right)\right\| \\
&\qquad + \left\|H_{\mathrm{FB}}'\left(\nu^*; \nu^{(k)} - \nu^*\right) - \left(H_{\mathrm{FB}}\left(\nu^{(k)}\right) - H_{\mathrm{FB}}\left(\nu^*\right)\right)\right\| \\
&\qquad \left. + \left\|H_{\mathrm{FB}}\left(\nu^{(k)}\right) - H_{t_k}\left(\nu^{(k)}\right)\right\|\right\}.
\end{aligned} \tag{95}$$

It follows from (54) that, for $\nu^{(k)}$ sufficiently close to ν^*,

$$\begin{aligned}
&\left\|\left(\nabla H_{t_k}\left(\nu^{(k)}\right) - V_k\right)^{\top}\left(\nu^{(k)} - \nu^*\right)\right\| \\
&\quad \leq \left\|\nabla H_{t_k}\left(\nu^{(k)}\right) - V_k\right\| \left\|\nu^{(k)} - \nu^*\right\| \\
&\quad = O\left(\left\|\nu^{(k)} - \nu^*\right\|^2\right).
\end{aligned} \tag{96}$$

From (35), the inequality $t_k \leq \kappa\delta_k^r = \kappa\|H_{\mathrm{FB}}(\nu^{(k)})\|^r$ by Step 3 of Algorithm 13, and the local Lipschitz continuity of H_{FB}, we get

$$\begin{aligned}
\left\|H_{\mathrm{FB}}\left(\nu^{(k)}\right) - H_{t_k}\left(\nu^{(k)}\right)\right\| &\leq \sqrt{2m}t_k \\
&\leq \sqrt{2m}\kappa\left\|H_{\mathrm{FB}}\left(\nu^{(k)}\right)\right\|^r \\
&= \sqrt{2m}\kappa\left\|H_{\mathrm{FB}}\left(\nu^{(k)}\right) - H_{\mathrm{FB}}\left(\nu^*\right)\right\|^r \\
&= O\left(\left\|\nu^{(k)} - \nu^*\right\|^r\right).
\end{aligned} \tag{97}$$

Since, by Proposition 11, H_{FB} is semismooth, we have from (15) and (17) that

$$\begin{aligned}
&\left\|V_k^{\top}\left(\nu^{(k)} - \nu^*\right) - H_{\mathrm{FB}}'\left(\nu^*; \nu^{(k)} - \nu^*\right)\right\| \\
&\qquad + \left\|H_{\mathrm{FB}}'\left(\nu^*; \nu^{(k)} - \nu^*\right) - \left(H_{\mathrm{FB}}\left(\nu^{(k)}\right) - H_{\mathrm{FB}}\left(\nu^*\right)\right)\right\| \\
&\quad = o\left(\left\|\nu^{(k)} - \nu^*\right\|\right).
\end{aligned} \tag{98}$$

Therefore, from (95)–(97) and $r > 1$, we obtain (91). Moreover, if ∇F is locally Lipschitzian, then, by Proposition 11, H_{FB} is strongly semismooth, and hence, we have

$$\begin{aligned}
&\left\|V_k^{\top}\left(\nu^{(k)} - \nu^*\right) - H_{\mathrm{FB}}'\left(\nu^*; \nu^{(k)} - \nu^*\right)\right\| \\
&\qquad + \left\|H_{\mathrm{FB}}'\left(\nu^*; \nu^{(k)} - \nu^*\right) - \left(H_{\mathrm{FB}}\left(\nu^{(k)}\right) - H_{\mathrm{FB}}\left(\nu^*\right)\right)\right\| \\
&\quad = O\left(\left\|\nu^{(k)} - \nu^*\right\|^2\right).
\end{aligned} \tag{99}$$

Therefore, from (95)–(97) and $r \geq 2$, we obtain (92). □

Using Lemma 27, we obtain the local Q-superlinear (Q-quadratic) convergence result.

Theorem 28. *Suppose that Assumptions (A1)–(A3) hold. Let* $\{\nu^{(k)}\}$ *be a sequence generated by Algorithm 13. If every accumulation point of* $\{\nabla H_{t_k}(\nu^{(k)})\}$ *is nonsingular, then the following statements hold.*

(a) *For all k sufficiently large,* $\|H_{FB}(\hat{\nu}^{(j)} + \hat{d}^{(j)})\| \leq \beta_k$ *is satisfied at* $j = 0$ *in Step 2.2 of Algorithm 13. Moreover, for all k sufficiently large,* $\nu^{(k+1)} = \nu^{(k)} + d^{(k)}$ *holds, where* $d^{(k)} = -\nabla H_{t_k}(\nu^{(k)})^{-\top} H_{t_k}(\nu^{(k)})$.

(b) *The whole sequence* $\{\nu^{(k)}\}$ *converges Q-superlinearly to a solution* ν^* *of SOCCP (1). Moreover, if* ∇F *is locally Lipschitz continuous and* $r \geq 2$, *then the sequence* $\{\nu^{(k)}\}$ *converges Q-quadratically.*

Proof. Since (b) is directly obtained from (a) and Lemma 27, it suffices to prove (a). Namely, we prove that $\|H_{t_k}(\nu^{(k)} + d^{(k)})\| \leq \beta_k = \beta_0\eta^k$ for all k sufficiently large. We have from (93) that

$$\left\|d^{(k)}\right\| = \left\|-\nabla H_{t_k}\left(\nu^{(k)}\right)^{-\top} H_{t_k}\left(\nu^{(k)}\right)\right\| \leq c_1 \left\|H_{t_k}\left(\nu^{(k)}\right)\right\|, \tag{100}$$

and hence, it follows from the boundedness of $\{\nu^{(k)}\}$ and the continuity of H_{t_k} that $\{d^{(k)}\}$ is bounded. Let ν^* be any accumulation point of $\nu^{(k)}$, and let $\nu^{(k)}$ be sufficiently close to ν^*. By Theorem 23, ν^* is a solution of SOCCP (1), and thus, $H_{\mathrm{FB}}(\nu^*) = 0$. From the local Lipschitz continuity of H_{FB}, we may assume that $\|H_{\mathrm{FB}}(v^{(k)})\|$ is small enough. By Lemma 27 and (100), there exists a constant $c_2 \in (0,1)$ such that

$$\begin{aligned} \left\|v^{(k)} - v^*\right\| &\le \left\|-d^{(k)}\right\| + \left\|v^{(k)} + d^{(k)} - v^*\right\| \\ &\le c_1 \left\|H_{t_k}\left(v^{(k)}\right)\right\| + o\left(\left\|v^{(k)} - v^*\right\|\right) \\ &\le c_1 \left\|H_{t_k}\left(v^{(k)}\right)\right\| + c_2 \left\|v^{(k)} - v^*\right\|. \end{aligned} \quad (101)$$

Therefore, we have from (35) with $t_k \le \kappa\|H_{\mathrm{FB}}(v^{(k)})\|^r \le \kappa\|H_{\mathrm{FB}}(v^{(k)})\|$ that

$$\begin{aligned} &\left\|v^{(k)} - v^*\right\| \\ &\le \frac{c_1}{1-c_2}\left\|H_{t_k}\left(v^{(k)}\right)\right\| \\ &\le \frac{c_1}{1-c_2}\left(\left\|H_{t_k}\left(v^{(k)}\right) - H_{\mathrm{FB}}\left(v^{(k)}\right)\right\| + \left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\|\right) \\ &\le \frac{c_1}{1-c_2}\left(\sqrt{2m}t_k + \left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\|\right) \\ &= O\left(\left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\|\right). \end{aligned} \quad (102)$$

This together with Lemma 27 and the local Lipschitz continuity of H_{FB} yields that

$$\begin{aligned} \left\|H_{\mathrm{FB}}\left(v^{(k)} + d^{(k)}\right)\right\| &= \left\|H_{\mathrm{FB}}\left(v^{(k)} + d^{(k)}\right) - H_{\mathrm{FB}}\left(v^*\right)\right\| \\ &= O\left(\left\|v^{(k)} + d^{(k)} - v^*\right\|\right) \\ &= o\left(\left\|v^{(k)} - v^*\right\|\right) = o\left(\left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\|\right). \end{aligned} \quad (103)$$

Therefore, it follows again from (35) with $t_k \le \kappa\|H_{\mathrm{FB}}(v^{(k)})\|^r$ that

$$\begin{aligned} &\left\|H_{t_k}\left(v^{(k)} + d^{(k)}\right)\right\| \\ &\le \left\|H_{t_k}\left(v^{(k)} + d^{(k)}\right) - H_{\mathrm{FB}}\left(v^{(k)} + d^{(k)}\right)\right\| \\ &\quad + \left\|H_{\mathrm{FB}}\left(v^{(k)} + d^{(k)}\right)\right\| \\ &\le \sqrt{2m}t_k + \left\|H_{\mathrm{FB}}\left(v^{(k)} + d^{(k)}\right)\right\| \\ &= o\left(\left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\|\right). \end{aligned} \quad (104)$$

From (35), the choices of t_k and β_k in Step 3 of Algorithm 13, and $0 < \overline{\eta} \le \eta < 1$, we get

$$\begin{aligned} \left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\| &\le \left\|H_{\mathrm{FB}}\left(v^{(k)}\right) - H_{t_{k-1}}\left(v^{(k)}\right)\right\| + \left\|H_{t_{k-1}}\left(v^{(k)}\right)\right\| \\ &\le \sqrt{2m}t_{k-1} + \beta_{k-1} \\ &\le \sqrt{2m}t_0\eta^{k-1} + \beta_0\eta^{k-1} \\ &= c_3\eta^k, \end{aligned} \quad (105)$$

where $c_3 = \eta^{-1}(\sqrt{2m}t_0 + \beta_0)$, and hence, (104) yields $\|H_{t_k}(v^{(k)}+d^{(k)})\| = o(\eta^k)$. This implies that $\|H_{t_k}(v^{(k)}+d^{(k)})\| \le \beta_k = \beta_0\eta^k$. Taking into account $\widehat{v}^{(0)} = v^{(k)}$ and $\widehat{d}^{(0)} = d^{(k)}$, we have $\|H_{\mathrm{FB}}(\widehat{v}^{(0)} + \widehat{d}^{(0)})\| \le \beta_k$ and $v^{(k+1)} = v^{(k)} + d^{(k)}$. Since, by Lemma 27, $v^{(k+1)}$ remains in the neighborhood of v^*, the desired results are obtained. □

5. Numerical Experiments

In this section, we show some numerical results for Algorithm 13. The program was coded in MATLAB 7, and computations were carried out on a machine with Intel Core i7-3770 K CPU (3.50 GHz×2) and 8.0 GB RAM. We set the parameters $\eta = \overline{\eta} = 0.1$, $\rho = 0.66$, $\sigma = 0.1$, $r = 2$, $\kappa = 1$, $\widehat{\kappa} = 0.2$, and $\beta_0 = 2$. We also set the function $\overline{t}_i$ as follows (see [22] for details):

$$\overline{t}_i\left(z^i,\delta\right) := \begin{cases} \dfrac{\left\|z^i\right\|\delta}{\sqrt{2\left(1-\delta^2\right)}} & \text{if } z^i \ne (0,0) \in \mathbf{R}^2,\ \delta < 1, \\ 10^{10} & \text{otherwise} \end{cases} \quad (106)$$

for $n_i = 1$, and

$$\overline{t}_i\left(z^i,\delta\right) := \begin{cases} \dfrac{\lambda_1\left(w^i\right)\delta}{\sqrt{2\left(2\left(w^i\right)_1 - \lambda_1\left(w^i\right)\delta^2\right)}} & \text{if } z^i \in \mathcal{Z}_1^i,\ \delta < \sqrt{\dfrac{2\left(w^i\right)_1}{\lambda_1\left(w^i\right)}}, \\ \dfrac{\delta\sqrt{\left(w^i\right)_1}}{2\sqrt{2\left(2-\delta\right)}} & \text{if } z^i \in \mathcal{Z}_2^i,\ \delta < 2, \\ 10^{10} & \text{otherwise} \end{cases} \quad (107)$$

for $n_i \ge 2$. For all problems, we randomly chose the initial point $(x^{(0)}, y^{(0)}, p^{(0)}) \in \mathbf{R}^{2n+\ell}$ whose components were distributed on the interval $[0,1]$, by using **rand** command of MATLAB. The stopping criterion in Step 1 is relaxed to

$$\left\|H_{\mathrm{FB}}\left(v^{(k)}\right)\right\| \le 10^{-8}. \quad (108)$$

We first solve the following second-order cone programming (SOCP) problem:

$$\begin{aligned} &\text{minimize} \quad c^\top x \\ &\text{subject to} \quad Ax + b = 0, \quad x \in \mathcal{K}, \end{aligned} \quad (109)$$

which is reformulated as SOCCP (1) with

$$F(x,y,p) := \begin{bmatrix} O & -I & A^\top \\ A & O & O \end{bmatrix}\begin{bmatrix} x \\ y \\ p \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix} \quad (110)$$

equivalently. We generate one hundred test problems randomly such that there exist primal and dual strictly feasible

Table 1: Numerical comparison of our method with SDPT3.

$\mathcal{K}$	n	ℓ	Our method		SDPT3	
			Iter	CPU	Iter	CPU
$(\mathcal{K}^5)^3 \times (\mathcal{K}^2)^2 \times \mathbf{R}_+$	20	5	8.99	0.063	8.84	0.070
$(\mathcal{K}^{10})^5$	50	10	8.28	0.058	9.82	0.078
$(\mathcal{K}^{100})^3 \times (\mathcal{K}^{50})^2$	400	100	7.02	1.521	10.19	0.291
$\mathcal{K}^{500} \times \mathcal{K}^{200} \times (\mathcal{K}^{100})^3$	1000	200	7.01	8.253	14.85	2.512

solutions. Specifically, we first choose matrix $A \in \mathbf{R}^{\ell \times n}$ whose components are distributed on $[-100, 100]$, vectors $\overline{x}, \overline{y} \in \operatorname{int} \mathcal{K}$ and $\overline{p} \in \mathbf{R}^{\ell}$ randomly, and then set $b := -A\overline{x}$ and $c := -A^{\top}\overline{p} + \overline{y}$. Here, each component of $(\overline{x}^i)_2$ is distributed on $[-100, 100]$ and $(\overline{x}^i)_1$ is set to $(\overline{x}^i)_1 := \|(\overline{x}^i)_2\| + \alpha$, where $\alpha \in (0, 100]$ is chosen randomly, and $\overline{y}$ is also generated similarly, while each component of $\overline{p}$ is distributed on $[0, 1]$. In order to compare our method with another method, we solve SOCP (109) by SDPT3 [26, 27], which is the software of interior point methods for solving semidefinite, second-order cone, and linear programming problems. We use SDPT3 with the default parameter and option settings. The obtained results are shown in Table 1, in which "Iter" and "CPU" denote the average values of the number of iterations and the CPU time in seconds, respectively. In particular, the value of "Iter" in "our method" denotes the number of times that the Newton equations (49) have been solved. In the column of $\mathcal{K}$, $(\mathcal{K}^5)^3$ denotes $\mathcal{K}^5 \times \mathcal{K}^5 \times \mathcal{K}^5$, for example. Since SDPT3 failed to solve some test problems when $\mathcal{K} = (\mathcal{K}^5)^3 \times (\mathcal{K}^2)^2 \times \mathbf{R}_+$, the average values in this case were taken over the successful trials only. We see from Table 1 that our method is superior to or at least comparable with SDPT3 from the viewpoint of the number of iterations. On the other hand, from the viewpoint of CPU time, our method is also superior to SDPT3 for small-scale problems. However, SDPT3 outperforms our method for middle- or large-scale problems. We believe that this is because SDPT3 is coded to reduce the computational costs by means of some fundamental techniques on matrix computation and so forth. For further development of our method, we will need more appropriate tuning of our code. However, it is not the purpose of this paper.

In order to confirm the local behaviors of the sequence generated by Algorithm 13, we list the value of $\|H_{\mathrm{FB}}(\widehat{v}^{(j)})\|$ at each outer iteration k in Table 2. In addition, to investigate how the parameter r affects the rate of convergence, we performed the algorithm with $r = 1, 1.5, 2$. We also investigate the relation between the choices of r and the behavior of $\{t_k\}$; we list the behaviors of $\{t_k\}$. We chose one of the above test problems in the case $\mathcal{K} = \mathcal{K}^{500} \times \mathcal{K}^{200} \times (\mathcal{K}^{100})^3$. We note that 2.66e − 09 means 2.66×10^{-9}, for example. We see from Table 2 that the sequence generated by Algorithm 13 seems to converge Q-quadratically and the parameter r does not affect the convergence of the sequence. On the other hand, we find that the choices of r affect the behavior of $\{t_k\}$.

The next experiment is an application of Algorithm 13 to the robust Nash equilibrium problem in the game theory. The robust Nash equilibrium [2, 3, 28, 29] is a new solution concept for noncooperative games with uncertain information. In this model, it is assumed that each player's cost (payoff) function and/or the opponents' strategies are uncertain, but they belong to some uncertainty sets and each player chooses his strategy by taking the worst possible case into consideration. In other words, each player makes decision according to the robust optimization policy. In this experiment, we focus on the following 2-person robust Nash game with quadratic cost functions:

Player 1:

$$\begin{aligned} &\underset{x^1}{\text{minimize}} \quad \max_{\delta x^2 \in D_2} \left\{ \frac{1}{2} \left(x^1\right)^{\top} A_{11} x^1 + \left(x^1\right)^{\top} A_{12} \left(x^2 + \delta x^2\right) \right\} \\ &\text{subject to} \quad x^1 \geq 0, \quad \widehat{e}^{\top} x^1 = 1; \end{aligned} \tag{111}$$

Player 2:

$$\begin{aligned} &\underset{x^2}{\text{minimize}} \quad \max_{\delta x^1 \in D_1} \left\{ \frac{1}{2} \left(x^2\right)^{\top} A_{22} x^2 + \left(x^2\right)^{\top} A_{21} \left(x^1 + \delta x^1\right) \right\} \\ &\text{subject to} \quad x^2 \geq 0, \quad \widehat{e}^{\top} x^2 = 1, \end{aligned} \tag{112}$$

where $A_{ij} \in \mathbf{R}^{m_i \times m_j}$ for $(i, j) \in \{1, 2\} \times \{1, 2\}$ are given matrices, $\widehat{e}$ is the vector of ones of appropriate dimension, and $x^1 \in \mathbf{R}^{m_1}$ and $x^2 \in \mathbf{R}^{m_2}$ denote the mixed strategies for Players 1 and 2, respectively. Moreover, δx^1 and δx^2 mean the estimation error or noise, and each player knows that they belong to the uncertainty sets D_1 and D_2, respectively. Under this situation, the tuple (x^1, x^2) is called a *robust Nash equilibrium* when x^1 and x^2 solve (111) and (112) simultaneously. In this experiment, we set

$$\begin{gathered} A_{11} := \begin{bmatrix} 30 & -12 & -3 \\ -5 & 13 & 15 \\ 1 & 9 & 23 \end{bmatrix}, \qquad A_{12} := \begin{bmatrix} -11 & 3 & -10 \\ -8 & -15 & -2 \\ 3 & -1 & -6 \end{bmatrix}, \\ A_{21} := \begin{bmatrix} 3 & -1 & 1 \\ 10 & 0 & 6 \\ -1 & 6 & 6 \end{bmatrix}, \qquad A_{22} := \begin{bmatrix} 31 & 9 & 6 \\ 1 & 24 & 6 \\ 4 & 8 & 29 \end{bmatrix}, \\ D_1 := \left\{ \delta x_1 \mid \left\|\delta x_1\right\| \leq \rho_1, \widehat{e}^{\top} \delta x_1 = 0 \right\}, \\ D_2 := \left\{ \delta x_2 \mid \left\|\delta x_2\right\| \leq \rho_2, \widehat{e}^{\top} \delta x_2 = 0 \right\} \end{gathered} \tag{113}$$

Table 2: Numerical behaviors of $\|H_{\mathrm{FB}}(\hat{v}^{(j)})\|$.

		$r = 1$				$r = 1.5$				$r = 2$	
k	j	t_k	$\|H_{\mathrm{FB}}(\hat{v}^{(j)})\|$	k	j	t_k	$\|H_{\mathrm{FB}}(\hat{v}^{(j)})\|$	k	j	t_k	$\|H_{\mathrm{FB}}(\hat{v}^{(j)})\|$
1	0	1.00e + 00	1.62e + 06	1	0	1.00e + 00	1.62e + 06	1	0	1.00e + 00	1.62e + 06
1	1	1.00e + 00	4.35e + 03	1	1	1.00e + 00	3.80e + 03	1	1	1.00e + 00	3.77e + 03
1	2	1.00e + 00	8.03e + 02	1	2	1.00e + 00	7.38e + 02	1	2	1.00e + 00	7.32e + 02
1	3	1.00e + 00	1.37e + 02	1	3	1.00e + 00	1.27e + 02	1	3	1.00e + 00	1.27e + 02
1	4	1.00e + 00	1.38e + 01	1	4	1.00e + 00	1.30e + 01	1	4	1.00e + 00	1.30e + 01
1	5	1.00e + 00	1.38e + 01	1	5	1.00e + 00	1.30e + 01	1	5	1.00e + 00	1.30e + 01
2	1	1.00e − 01	2.85e − 01	2	1	1.00e − 01	2.54e − 01	2	1	6.57e − 02	2.56e − 01
3	1	1.17e − 04	1.17e − 04	3	1	8.80e − 07	9.19e − 05	3	1	9.71e − 09	9.86e − 05
4	1	2.66e − 09	2.66e − 09	4	1	1.20e − 13	2.43e − 09	4	1	5.39e − 18	2.32e − 09

Table 3: Robust Nash equilibria with various choices of uncertainty radiuses.

Iter	CPU	ρ_1	ρ_2	x^1	x^2
8	0.078	0.2	0.2	(0.3916, 0.6083, 0)	(0.3247, 0.3625, 0.3128)
8	0.094	0.4	0.4	(0.4168, 0.5832, 0)	(0.3354, 0.3152, 0.3495)
7	0.078	0.6	0.6	(0.4292, 0.5708, 0)	(0.3477, 0.2821, 0.3702)
7	0.078	0.8	0.8	(0.4017, 0.5065, 0.0918)	(0.3686, 0.2503, 0.3812)
9	0.078	1.0	1.0	(0.3627, 0.4589, 0.1784)	(0.3891, 0.2258, 0.3851)

and change the values of ρ_1 and ρ_2 variously. Since D_1 and D_2 are defined by means of Euclidean norm, the robust Nash equilibrium problem can be reformulated as an SOCCP equivalently (the reformulated SOCCP is explicitly written in Section 5.1.1 of [3]. We thus omit the details here). Here, we emphasize that the reformulated SOCCP cannot be expressed as any SOCP, and hence existing software such as SDPT3 cannot be applied. Moreover, if the reformulated SOCCP is rewritten of the form (1), then it satisfies neither (4) nor (6). The obtained results are summarized in Table 3, in which x^1 and x^2 denote the obtained robust Nash equilibria for various choices of uncertainty radiuses ρ_1 and ρ_2. For all problems, we could calculate the robust Nash equilibria correctly. Moreover, as is discussed in the existing papers, we can observe that the robust Nash equilibria move smoothly as the values of ρ_1 and ρ_2 change gradually.

6. Conclusion

In this paper, we have proposed a smoothing Newton method with appropriate parameter control based on the Fischer-Burmeister function for solving the SOCCP. We have shown its global and Q-quadratic convergence properties under some assumptions. In addition, we have considered some sufficient conditions for the assumptions. In numerical experiments, we have confirmed the effectiveness of the proposed methods.

Acknowledgments

The authors would like to thank the academic editor Prof. Jein-Shan Chen and the anonymous reviewers who helped them improve the original paper. The first author is supported in part by the Grant-in-Aid for Scientific Research (C) 25330030 and Young Scientists (B) 25870239 of Japan Society for the Promotion of Science. The third author is supported in part by the Grant-in-Aid for Young Scientists (B) 22760064 of Japan Society for the Promotion of Science.

References

[1] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, NY, USA, 2003.

[2] S. Hayashi, N. Yamashita, and M. Fukushima, "Robust Nash equilibria and second-order cone complementarity problems," *Journal of Nonlinear and Convex Analysis*, vol. 6, no. 2, pp. 283–296, 2005.

[3] R. Nishimura, S. Hayashi, and M. Fukushima, "Robust Nash equilibria in N-person non-cooperative games: uniqueness and reformulation," *Pacific Journal of Optimization*, vol. 5, no. 2, pp. 237–259, 2009.

[4] 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," *International Journal for Numerical Methods in Engineering*, vol. 65, no. 1, pp. 62–83, 2006.

[5] M. Fukushima, Z.-Q. Luo, and P. Tseng, "Smoothing functions for second-order-cone complementarity problems," *SIAM Journal on Optimization*, vol. 12, no. 2, pp. 436–460, 2002.

[6] S. Hayashi, N. Yamashita, and M. Fukushima, "A combined smoothing and regularization method for monotone second-order cone complementarity problems," *SIAM Journal on Optimization*, vol. 15, no. 2, pp. 593–615, 2005.

[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," *Computational Optimization and Applications*, vol. 25, no. 1–3, pp. 39–56, 2003.

[8] Y. Narushima, N. Sagara, and H. Ogasawara, "A smoothing Newton method with Fischer-Burmeister function for second-order cone complementarity problems," *Journal of Optimization Theory and Applications*, vol. 149, no. 1, pp. 79–101, 2011.

[9] J.-S. Chen, "A new merit function and its related properties for the second-order cone complementarity problem," *Pacific Journal of Optimization*, vol. 2, no. 1, pp. 167–179, 2006.

[10] S. Pan and J.-S. Chen, "A regularization method for the second-order cone complementarity problem with the Cartesian P_0-property," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 70, no. 4, pp. 1475–1491, 2009.

[11] J.-S. Chen and P. Tseng, "An unconstrained smooth minimization reformulation of the second-order cone complementarity problem," *Mathematical Programming*, vol. 104, no. 2-3, pp. 293–327, 2005.

[12] J.-S. Chen, "Two classes of merit functions for the second-order cone complementarity problem," *Mathematical Methods of Operations Research*, vol. 64, no. 3, pp. 495–519, 2006.

[13] S. Pan and J.-S. Chen, "A damped Gauss-Newton method for the second-order cone complementarity problem," *Applied Mathematics and Optimization*, vol. 59, no. 3, pp. 293–318, 2009.

[14] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization," *SIAM Journal on Control and Optimization*, vol. 15, no. 6, pp. 959–972, 1977.

[15] L. Qi and J. Sun, "A nonsmooth version of Newton's method," *Mathematical Programming*, vol. 58, no. 3, pp. 353–367, 1993.

[16] J.-S. Chen, X. Chen, and P. Tseng, "Analysis of nonsmooth vector-valued functions associated with second-order cones," *Mathematical Programming*, vol. 101, no. 1, pp. 95–117, 2004.

[17] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, NY, USA, 1983, Reprinted by SIAM, Philadelphia, Pa, USA, 1990.

[18] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations," *Mathematics of Operations Research*, vol. 18, no. 1, pp. 227–244, 1993.

[19] J.-S. Chen and S. Pan, "A survey on SOC complementarity functions and solution methods for SOCPS and SOCCPS," *Pacific Journal of Optimization*, vol. 8, no. 1, pp. 33–74, 2012.

[20] S. Bi, S. Pan, and J.-S. Chen, "The same growth of FB and NR symmetric cone complementarity functions," *Optimization Letters*, vol. 6, no. 1, pp. 153–162, 2012.

[21] C. Kanzow, "Some noninterior continuation methods for linear complementarity problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 4, pp. 851–868, 1996.

[22] H. Ogasawara and Y. Narushima, "The Jacobian consistency of a smoothed Fischer-Burmeister function associated with second-order cones," *Journal of Mathematical Analysis and Applications*, vol. 394, no. 1, pp. 231–247, 2012.

[23] D. Sun and J. Sun, "Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions," *Mathematical Programming*, vol. 103, no. 3, pp. 575–581, 2005.

[24] F. Facchinei and C. Kanzow, "Beyond monotonicity in regularization methods for nonlinear complementarity problems," *SIAM Journal on Control and Optimization*, vol. 37, no. 4, pp. 1150–1161, 1999.

[25] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Mathematical Programming*, vol. 95, no. 1, pp. 3–51, 2003.

[26] K. C. Toh, R. H. T. Tütüncü, and M. J. Todd, SDPT3 version 4.0—a MATLAB software for semidefinite-quadratic-linear programming, http://www.math.nus.edu.sg/~mattohkc/sdpt3.html.

[27] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, vol. 95, no. 2, pp. 189–217, 2003.

[28] M. Aghassi and D. Bertsimas, "Robust game theory," *Mathematical Programming*, vol. 107, no. 1-2, pp. 231–273, 2006.

[29] R. Nishimura, S. Hayashi, and M. Fukushima, "Semidefinite complementarity reformulation for robust Nash equilibrium problems with Euclidean uncertainty sets," *Journal of Global Optimization*, vol. 53, no. 1, pp. 107–120, 2012.