〔論　説〕

A Smoothing Newton Method by Fischer-Burmeister Function with an outside parameter for Second-Order-Cone Complementarity Problems

Nobuko Sagara

Abstract The second-order cone complementarity problem (SOCCP) is an important class of problems containing a lot of optimization problems. The SOCCP can be transformed into a system of nonsmooth equations. To solve this nonsmooth system by a smoothing Newton method, there are mainly two ways to use the Chen-Mangasarian class, that is, the smoothed natural residual and the smoothed Fischer-Burmeister function. Fukushima, Luo and Tseng (2001) [13] studied practical and concrete theories and properties of the above smoothing functions for SOCCP. Recently, a practical computational method using the natural residual function to solve SOCCP was given by Hayashi, Yamashita and Fukushima (2005) [14]. In this paper we propose an algorithm to solve SOCCP by using the Fischer-Burmeister function instead of the natural residual function in [14].

1. Introduction

In this paper we consider the second-order cone complementarity problem (SOCCP), which is to find $(x, y) \in R^n \times R^n$ such that

$$x \in \mathcal{K}, \qquad y \in \mathcal{K}, \qquad \langle x, y\rangle = 0, \qquad y = f(x) \tag{1.1}$$

where $\langle \cdot\ ,\ \cdot \rangle$ denotes the Euclidean inner product, f is a continuously differentiable function from R^n to R^n, and $\mathcal{K} \subset R^n$ is the Cartesian product of second-order cones, that is, $\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m}$ with $n = n_1 + \cdots + n_m$ and $\mathcal{K}^{n_i}$ is the n_i-dimensional second-order cone defined by

$$\mathcal{K}^{n_i} := \left\{(z_1, z_2) \in R \times R^{n_i - 1} \quad | \quad \|z_2\|_2 \leq z_1\right\} \subset R^{n_i}.$$

The KKT conditions for any second-order cone program (SOCP) is written as the SOCCP. The theoretical research on primal-dual path-following algorithms for solving second-order cone programs is done by Tsuchiya and others [1, 2, 17, 22, 23]. On the other hand, the research on SOCCP may be found in [6, 7, 8, 9, 10, 14, 13, 18]. The theory of solving SOCCP by smoothing functions including natural residual and Fischer-Burmeister functions was studied by Fukushima et al. [13]. Recently, using this result, the practical computational method using the natural residual function was given by Hayashi et al. [14]. In this paper, we propose an algorithm to solve SOCCP by using the Fischer-Burmeister function instead of the natural residual.

In datail, Fukushima, Luo and Tseng [13] showed that the "min" function and Fischer-Burmeister function for the NCP can be extended to the SOCCP by using the Jordan algebra. Also, the SOCCP function associated with the min function is called the natural residual function. Furthermore, Fukushima et al. [13] constructed smoothing functions for those functions and analyzed the properties of their Jacobians. Hayashi et al. [14] proposed a smoothing method based on the smoothed natural residual function, and showed its global and quadratic convergence. On the other hand, Chen, Sun and Sun [6] proposed another smoothing method based on natural residual function that is called the CHKS(Chen-Harker-Kanzow-Smale) smoothing function in X.Chen [5]. Following the idea of Qi, Sun and Zhou [21], Chen et al. [6] treated a smoothing parameter as a variable, in contrast with Hayashi et al. [14]. Moreover, they gave global and quadratic convergence of their method, which follows results of Qi et al. [21].

This paper is organized as follows. In Section 2, we review some concepts of semismoothness and some properties of the spectral factorization with respect to SOC, which will be used in the subsequent analysis. Also, we give a merit function by means of Fischer-Burmeister function for the SOCCP. In Section 3, we introduce a smoothing function with the smoothed Fischer-Burmeister function. In Section 4, we have the analysis that we get information how to update for an outside parameter "t" on the practical computation. In section 5, we propose an algorithm for solving the SOCCP and discuss its convergence properties, and then we give some numerical experiences of the proposed method.

Throughout this paper, we let R_+ and R_{++} denote the nonnegative and positive reals.

2. Some Preliminaries

2.1 Semismoothness and strong semismoothness

Semismoothness is a generalized concept of the smoothness, which was originally introduced by Mifflin [15] for functionals, and extended to vector-valued functions by Qi and Sun [20].

Definition 2..1 *[12] Let $H : R^n \to R^m$ be a locally Lipschitzian function. Then H is differentiable almost everywhere by Rademacher's Theorem [11]. Let D_H be the set of differentiable points of H. The B(ouligant)-subdifferential and Clarke subdifferential of H at x is respectively defined by*

$$\begin{aligned} \partial_B H(x) &:= \{lim_{\hat{x}\to x} \nabla H(\hat{x}) \ | \ \hat{x} \in D_H\} \\ \partial H(x) &:= co\, \partial_B H(x) \end{aligned}$$

where $\nabla H(x)$ is the Jacobian of H at x and $co\ S$ is the convex hull of S.

Note that if H is continuously differentiable at x, then $\partial H(x) = \{\nabla H(x)\}$.

Definition 2..2 *[12] A directionally differentiable and locally Lipschitizian function $H : R^n \to R^m$ is said to be semismooth at x if*

$$H'(x; d) - V^T d = o(\|d\|)$$

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

$$H'(x; d) := lim_{\tau \downarrow 0} \left(H(x + \tau d) - H(x) \right) / \tau$$

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.

2.2 Jordan algebra associated with SOCCP

We first recall the spectral factorization of a vector in R^n asoociated with $\mathcal{K}^n$. Let $z = (z_1,\ z_2) \in R \times R^{n-1}$. Then z can be decomposed as

$$z = \lambda_1 u^{(1)} + \lambda_2 u^{(2)} \tag{2.1}$$

where λ_1, λ_2 and $u^{(1)}$, $u^{(2)}$ are the spectral values and the associated spectral vectors of z given by

$$\lambda_i = z_1 + (-1)^i \|z_2\|, \tag{2.2}$$

$$u^{(i)} = \begin{cases} \frac{1}{2}\left(1,\ (-1)^i \frac{z_2}{\|z_2\|}\right) & \text{if } z_2 \neq 0 \\ \frac{1}{2}\left(1,\ (-1)^i w\right) & \text{if } z_2 = 0 \end{cases} \tag{2.3}$$

for $i = 1, 2$, with w being any vector in R^{n-1} satisfying $\|w\| = 1$. If $z_2 \neq 0$, the decomposition (2.1) is unique.

The Jordan product of $x = (x_1,\ x_2) \in R \times R^{n-1}$ and $y = (y_1,\ y_2) \in R \times R^{n-1}$ is defined as

$$x \cdot y = (x^T y,\ y_1 x_2 + x_1 y_2). \tag{2.4}$$

We will write z^2 to mean $z \cdot z$ and write $x + y$ to mean the usual componentwise addition of vectors x and y. We define $z^{1/2}$ as

$$z^{1/2} = \left(s,\ \frac{z_2}{2s}\right), \quad \text{where } s = \sqrt{\left(z_1 + \sqrt{z_1^2 - \|z_2\|^2}\right)/2}. \tag{2.5}$$

Note that $(z^{1/2})^2 = z^{1/2} \cdot z^{1/2} = z$. Moreover, for any z we define the symmetric matrix L_z as

$$L_z = \begin{bmatrix} z_1 & z_2^T \\ z_2 & z_1 I \end{bmatrix}. \tag{2.6}$$

Property 2.1 For any $x = (x_1, x_2) \in R \times R^{n-1}$, let λ_1, λ_2 and $u^{(1)}, u^{(2)}$ be the spectral values and the associated spectral vectors at x. Then the following hold.

1. $x \in \mathcal{K}^n \Longleftrightarrow 0 \le \lambda_1 \le \lambda_2$ and $x \in \mathrm{int}\mathcal{K}^n. \Longleftrightarrow 0 < \lambda_1 \le \lambda_2$
2. $x^2 = \lambda_1^2 u^{(1)} + \lambda_2^2 u^{(2)} \in \mathcal{K}^n$.
3. If $x \in \mathcal{K}^n$, then $x^{1/2} = \sqrt{\lambda_1}u^{(1)} + \sqrt{\lambda_2}u^{(2)} \in \mathcal{K}^n$.

2.3 Merit function

In order to construct a merit function for SOCCP (1.1), it is convenient to introduce a function $\hat{\Phi} : R^n \times R^n \to R^n$ satisfying

$$\hat{\Phi}(x, y) = 0 \quad \Leftrightarrow \quad x \in \mathcal{K}, \qquad y \in \mathcal{K}, \qquad \langle x, y \rangle = 0. \tag{2.7}$$

By using such a function, we define $\hat{H} : R^n \times R^n \to R^{2n}$ by

$$\hat{H}(x, y) := \begin{pmatrix} \hat{\Phi}(x, y) \\ f(x) - y \end{pmatrix}.$$

It is obvious that SOCCP (1.1) is equivalent to the equation $\hat{H}(x, y) = 0$. Moreover, we define function $\hat{\Psi} : R^n \times R^n \to R$ by

$$\hat{\Psi}(x, y) := \frac{1}{2} \| \hat{H}(x, y) \|^2 = \frac{1}{2} \| \hat{\Phi}(x, y) \|^2 + \frac{1}{2} \| f(x) - y \|^2 . \tag{2.8}$$

Then, it is easy to see that $\hat{\Psi}(x, y) \ge 0$ for any $(x, y) \in R^n \times R^n$, and that $\hat{\Psi}(x, y) = 0$ if and only if (x, y) is a solution of (1.1).

For a general SOCCP on $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}$, since

$$x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad \langle x, y \rangle = 0 \quad \Leftrightarrow \quad x^i \in \mathcal{K}^{n_i}, \quad y^i \in \mathcal{K}^{n_i}, \quad \langle x^i, y^i \rangle = 0 \quad (i = 1, ..., m), \tag{2.9}$$

where $x = (x^1, \cdots, x^m)$ and $y = (y^1, \cdots, y^m)$, we can define

$$\hat{\Phi}(x, y) := \begin{pmatrix} \hat{\phi}^1(x^1, y^1) \\ \vdots \\ \hat{\phi}^m(x^m, y^m) \end{pmatrix},$$

where $\hat{\phi}^i : R^{n_i} \times R^{n_i} \to R^{n_i}$ is any function satisfying

$$\hat{\phi}^i(x^i, y^i) = 0 \quad \Leftrightarrow \quad x^i \in \mathcal{K}^{n_i}, \quad y^i \in \mathcal{K}^{n_i}, \quad \langle x^i, y^i \rangle = 0 \tag{2.10}$$

for $i = 1, ..., m$. Fukushima et al. [13] showed that (2.10) holds for the Fischer-Burmeister function ϕ^i_{FB} : $R^{n_i} \times R^{n_i} \to R^{n_i}$ defined by

$$\phi^i_{FB}(x^i, y^i) := x^i + y^i - ((x^i)^2 + (y^i)^2)^{1/2}.$$

Using this function, we define the function $\Phi_{FB} : R^n \times R^n \to R^n$ by

$$\Phi_{FB}(x, y) := \begin{pmatrix} \phi^1_{FB}(x^1, y^1) \\ \vdots \\ \phi^m_{FB}(x^m, y^m) \end{pmatrix}$$

and $H_{FB} : R^n \times R^n \to R^{2n}$ by

$$H_{FB}(x, y) := \begin{pmatrix} \Phi_{FB}(x, y) \\ f(x) - y \end{pmatrix}. \tag{2.11}$$

Then, we can construct a merit function $\Psi_{FB} : R^n \times R^n \to R$ for SOCCP (1.1) by

$$\Psi_{FB}(x, y) := \frac{1}{2} \parallel H_{FB}(x, y) \parallel^2 = \frac{1}{2} \sum_{i=1}^{m} \parallel \phi^i_{FB}(x^i, y^i) \parallel^2 + \frac{1}{2} \parallel f(x) - y \parallel^2 .$$

3. Smoothing functions and its properties

Since H_{FB} is not differentiable, we cannot apply conventional methods such as the steepest descent method and Newton's method that use the gradient of the function. Therefore, we consider the smoothing function that is a generalization of the proposal of Kanzow [16] for the Fischer-Burmeister function with an outside parameter t:

$$\phi_t(x, y) := x + y - (x^2 + y^2 + 2t^2 e)^{1/2}, \tag{3.1}$$

where $e = (1, 0, ..., 0)^T$. Notice that $\phi_t(x, y) = 0$ if and only if $x \in \text{int}\, \mathcal{K}^n, y \in \text{int}\, \mathcal{K}^n$ and $x \cdot y = t^2 e$.
In the remainder of the paper, we assume $\mathcal{K} = \mathcal{K}^n$. Then we can rewrite SOCCP (1.1) as follows: Find $(x, y) \in R^n \times R^n$ such that

$$x \in \mathcal{K}^n, \qquad y \in \mathcal{K}^n, \qquad \langle x, y \rangle = 0, \qquad y = f(x). \tag{3.2}$$

The assumption $\mathcal{K} = \mathcal{K}^n$ is only for simplicity of presentation. The subsequent analysis may be extended to the general case $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}$ without difficulty. We define functions H_{FB} and Ψ_{FB} by

$$\begin{aligned} H_{FB}(x, y) &:= \begin{pmatrix} \phi_{FB}(x, y) \\ f(x) - y \end{pmatrix}, \qquad (3.3) \\ \Psi_{FB}(x, y) := \frac{1}{2} \parallel H_{FB}(x, y) \parallel^2 &= \frac{1}{2} \parallel \phi_{FB}(x, y) \parallel^2 + \frac{1}{2} \parallel f(x) - y \parallel^2, \end{aligned}$$

where $\phi_{FB}(x,y) = x + y - (x^2+y^2)^{1/2}$.

3.1 Smoothing functions

For a nondifferentiable function $h : R^n \to R^m$, we consider a function $h_t : R^n \to R^m$ with a parameter $t > 0$ that has the following properties:

(a) h_t is differentiable for any $t > 0$.

(b) $\lim_{t\downarrow 0} h_t(x) = h(x)$ for any $x \in R^n$.

Such a function h_t is called a smoothing function of h. Instead of handling the nonsmooth equation $h(x) = 0$ directly, the smoothing method solves a family of smoothed subproblems $h_t(x) = 0$ for $t > 0$, and obtain a solution of the original problem by letting $t \downarrow 0$. It can be shown [13] that $\phi_t(x,y)$ may be regarded as a smoothing function of the Fischer-Burmeister function ϕ_{FB}.
Recall that $\phi_t(x,y) = 0$ if and only if $x \in \text{int } \mathcal{K}^n, y \in \text{int } \mathcal{K}^n, x \cdot y = t^2 e$. From Proposition 5.1 in [13], we have the following relations

$$\begin{aligned} \|\phi_{t_2}(x,y) - \phi_{t_1}(x,y)\| &\leq \sqrt{2}(t_1 - t_2) \quad \text{for} \quad t_1 > t_2 > 0 \\ \|\phi_t(x,y) - \phi_{FB}(x,y)\| &\leq \sqrt{2}t \quad \text{for} \quad t > 0. \end{aligned} \tag{3.4}$$

In this paper, we treat the parameter t as a positive real variable that be controled from outside. Specifically, for the function $H_{FB}(x,y)$, we consider the following smoothing function by putting $F(x,y) = f(x) - y$:

$$H_t(x,y) := \begin{pmatrix} \phi_t(x,y) \\ F(x,y) \end{pmatrix}. \tag{3.5}$$

The merit function for this smoothing function Ψ is given by

$$\Psi_t(x,y) := \frac{1}{2}\left\{\|\phi_t(x,y)\|^2 + \|F(x,y)\|^2\right\}. \tag{3.6}$$

3.2 The smoothed Fischer-Burmeister function

We next give the explicit expressions of the Jacobian $H_t(w)$. When we define $w^t = w^t(x,y) = x^2+y^2+2t^2e = (w_1^t, w_2^t) \in R \times R^{n-1}$, we have

$$w_1^t = \|x\|^2 + \|y\|^2 + 2t^2 \quad \text{and} \quad w_2^t = 2(x_1x_2 + y_1y_2), \tag{3.7}$$

where $x = (x_1,x_2), y = (y_1,y_2) \in R \times R^{n-1}$. The spectral factorization of w^t is as follws:

$$w^t = \lambda_1(w^t)u^{(1)} + \lambda_2(w^t)u^{(2)},$$

where $\lambda_1(w^t), \lambda_2(w^t)$ and u_1^t, u_2^t are the spectral value and the associated spectral vectors of w^t given by

$$\lambda_i(w^t) = \|x\|^2 + \|y\|^2 + 2t^2 + 2(-1)^i \|x_1 x_2 + y_1 y_2\| \tag{3.8}$$

and

$$u_i^t = \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{w_2^t}{\|w_2^t\|}\right) & \text{if} \quad w_2^t \neq 0 \\ \frac{1}{2}\left(1, (-1)^i \nu\right) & \text{if} \quad w_2^t = 0, \end{cases}$$

for $i = 1, 2$, with ν being any vector in R^{n-1} satisfing $\|\nu\| = 1$. Since $w^t \in \mathcal{K}^n$, from Property 2.1 $u^t = (w^t)^{1/2}$ is given by

$$u^t = \sqrt{\lambda_1(w^t)}u^{(1)} + \sqrt{\lambda_2(w^t)}u^{(2)}. \tag{3.9}$$

Now we consider the Jacobian of H_t with $t > 0$.

In the rest of the paper, we use the following notation:

$$\begin{aligned} w = w(x, y) = x^2 + y^2 = (w_1, w_2), \qquad & w^t = w^t(x, y) = x^2 + y^2 + 2t^2 e = (w_1^t, w_2^t), \\ u = u(x, y) = w(x, y)^{1/2} = (u_1, u_2), \qquad & u^t = u^t(x, y) = (w^t)^{1/2} = (u_1^t, u_2^t). \end{aligned}$$

Proposition 3.1 *Let $w^t = (w_1^t, w_2^t) = x^2 + y^2 + 2t^2 e \in R \times R^{n-1}$ and $\lambda_1(w^t), \lambda_2(w^t)$ be in (3.8). If $t > 0$, then the function H_t is continuously differentiable on R^{2n}, and its Jacobian is given by*

$$\nabla H_t(x, y) = \begin{bmatrix} I - L_x L_{u^t}^{-1} & \nabla f(x) \\ I - L_y L_{u^t}^{-1} & -I \end{bmatrix}, \tag{3.10}$$

where

$$L_{u^t}^{-1} = \begin{cases} \frac{1}{\sqrt{w_1^t}} I & \text{if} \quad w_2^t = 0 \\ \begin{bmatrix} b^t & c^t \bar{w}_2^T \\ c^t \bar{w}_2 & a^t I + (b^t - a^t)\bar{w}_2 \bar{w}_2^T \end{bmatrix} & \text{if} \quad w_2^t \neq 0 \end{cases}$$

with

$$\bar{w}_2 = w_2 / \|w_2\| = w_2^t / \|w_2^t\|$$

and

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

Proof. It follows from Corollary 5.4 of [13].

q.e.d

We propose a smoothing Newton method for solving $H_t(x,y)=0$. In order to obtain the Newton step, nonsingularity of the Jacobian of H_t is important. To establish the nonsingularity of the Jacobian of H_t, we consider the following rank and monotonicity assumptions on $\nabla F(x,y,\zeta)$. Our case is $F(x,y,\zeta)=f(x)-y$. The two assumption (6.2) and (6.3) in [13] are necessary. Since our case $F(x,y,\zeta)=f(x)-y$ does not include ζ, we don't need to consider about (6.2) in [13]. Thus, the assumption (6.3) in [13] says that $\nabla f(x)$ is positive semidefinite, and hence, f is monotone if x is allowed to be any point in R^n.

Proposition 3.2 *For each $t \neq 0$ and $(x,y)\in R^{2n}$ satisfying the assumption (6.3) in [13], that is,*

$$(u,v)\in R^n\times R^n, \quad \nabla F(x,y)^T(u,v)=0 \Longrightarrow u^Tv\geq 0,$$

the matrix $\nabla H_t(x,y)$ given (3.10) is nonsingular.

3.3 Jacobian consistency

In this section, we consider Jacobian consistency, which was introduced by Chen,Qi and Sun [4]

Lemma 3.1 *[[19], Lemma 3.2] For any $x=(x_1,x_2), y=(y_1,y_2)\in R\times R^{n-1}$ with $w=(w_1,w_2)=x^2+y^2\in bd\,\mathcal{K}^n$, we have*

$$x_1^2=\|x_2\|^2,\ \ y_1^2=\|y_2\|^2,\quad x_1y_1=x_2^Ty_2,\quad x_1y_2=y_1x_2. \tag{3.11}$$

In addition, if $w_2\neq 0$, then $\|w_2\|=w_1=\|x\|^2+\|y\|^2=2(x_1^2+y_1^2)=2(\|x_2\|^2+\|y_2\|^2)=\frac{1}{\sqrt{2}}\|w\|$ and

$$\begin{aligned} x_1w_2=\|w_2\|x_2=w_1x_2, &\qquad x_2^Tw_2=x_1\|w_2\|=w_1x_1,\\ y_1w_2=\|w_2\|y_2=w_1y_2, &\qquad y_2^Tw_2=y_1\|w_2\|=w_1y_1. \end{aligned} \tag{3.12}$$

Also, from the results, we obtain by putting $\bar{w}_2=w_2/\|w_2\|$,

$$\begin{aligned} x_1-x_2^T\bar{w}_2=0, &\qquad x_2-x_1\bar{w}_2=0,\\ y_1-y_2^T\bar{w}_2=0, &\qquad y_2-y_1\bar{w}_2=0. \end{aligned}$$

Note that from (3.11) $w=x^2+y^2=0$ implies $(x,y)=(0,0)$.

Now we define the following three sets

$$
\begin{aligned}
S_1 &= \left\{(x,y)\in R^{2n} \mid x^2+y^2\in \text{int}\,\mathcal{K}^n\right\},\\
S_2 &= \left\{(x,y)\in R^{2n} \mid x^2+y^2\in \text{bd}\,\mathcal{K}^n,\ \ (x,y)\neq(0,0)\right\},\\
S_3 &= \left\{(x,y)\in R^{2n} \mid (x,y)=(0,0)\right\}.
\end{aligned}
$$

Clearly, we see that $S_1\cup S_2\cup S_3=R^{2n}$.

Lemma 3.2

$$
J_H^0(x,y) := lim_{t\to 0}\nabla H_t(x,y) = \begin{bmatrix} I-J_x & \nabla F(x)\\ I-J_y & -I \end{bmatrix}
$$

where

$$
J_x = \begin{cases} L_xL_u^{-1} & \textit{if } (x,y)\in S_1\\ \frac{1}{2\sqrt{2(\|x\|^2+\|y\|^2)}}L_x\begin{bmatrix} 1 & \bar{w}_2^T\\ \bar{w}_2 & 4I-3\bar{w}_2\bar{w}_2^T\end{bmatrix} & \textit{if } (x,y)\in S_2\\ 0 & \textit{if } (x,y)\in S_3 \end{cases}
$$

and

$$
J_y = \begin{cases} L_yL_u^{-1} & \textit{if } (x,y)\in S_1\\ \frac{1}{2\sqrt{2(\|x\|^2+\|y\|^2)}}L_y\begin{bmatrix} 1 & \bar{w}_2^T\\ \bar{w}_2 & 4I-3\bar{w}_2\bar{w}_2^T\end{bmatrix} & \textit{if } (x,y)\in S_2\\ 0 & \textit{if } (x,y)\in S_3 \end{cases}
$$

with $\bar{w}_2=w_2/\|w_2\|$.

Proof. We only might prove that $\lim_{t\to 0}L_xL_{u^t}^{-1}=J_x$ and $\lim_{t\to 0}L_yL_{u^t}^{-1}=J_y$.

- The case $(x,y)\in S_1$. From the fact that $\lim_{t\to 0}L_{u^t}^{-1}=L_u^{-1}$, the results are clearly.
- The case $(x,y)\in S_2$. It follows from $w=x^2+y^2\in \text{bd}\,\mathcal{K}^n$ that we have $\|x\|^2+\|y\|^2=2\|x_1x_2+y_1y_2\|\neq 0$. Therefore, we have

$$
\lambda_1(w^t) = \|x\|^2+\|y\|^2-2\|x_1x_2+y_1y_2\|+2t^2 = 2t^2 \tag{3.13}
$$

$$
\lambda_2(w^t) = \|x\|^2+\|y\|^2+2\|x_1x_2+y_1y_2\|+2t^2 = 2t^2+2(\|x\|^2+\|y\|^2). \tag{3.14}
$$

By using the technique in the proof of Proposition 3.1 in [19], we can rewite $L_{u^t}^{-1}$ as

$$
L_{u^t}^{-1} = \frac{1}{2\sqrt{\lambda_1(w^t)}}\begin{bmatrix} 1 & -\bar{w}_2^T\\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix} + \frac{1}{2\sqrt{\lambda_2(w^t)}}\begin{bmatrix} 1 & \bar{w}_2^T\\ \bar{w}_2 & \frac{4\sqrt{\lambda_2(w^t)}}{\sqrt{\lambda_1(w^t)}+\sqrt{\lambda_2(w^t)}}(I-\bar{w}_2\bar{w}_2^T)+\bar{w}_2\bar{w}_2^T\end{bmatrix} \tag{3.15}
$$

Now we define the first and second terms by $L_1(w^t)$ and $L_2(w^t)$ respectively. It follows from (3.13) and (3.14), that we have $\lambda_1(w^t) \to 0$ and $\lambda_2(w^t) \to 2(\|x\|^2 + \|y\|^2)$ as $t \to 0$. Therefore,we have

$$\lim_{t\to 0} L_2(w^t) = \frac{1}{2\sqrt{2(\|x\|^2 + \|y\|^2)}} \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix}$$

Since from $x^2 + y^2 \in \mathrm{bd}\,\mathcal{K}^n$ and Lemma 3.1, we have $x_1 - x_2^T\bar{w}_2 = 0$ and $x_2 - x_1\bar{w}_2 = 0$,

$$\begin{aligned} L_x L_1(w^t) &= \frac{1}{2\sqrt{\lambda_1(w^t)}} \begin{bmatrix} x_1 - x_2^T\bar{w}_2 \\ x_2 - x_1\bar{w}_2 \end{bmatrix} [1 \quad -\bar{w}_2^T] \\ &= 0. \end{aligned}$$

In a similar way, we see $L_y L_1(w^t) = 0$. Therefore, we obtain

$$\begin{aligned} \lim_{t\to 0} L_x L_{u^t}^{-1} &= \lim_{t\to 0}(L_x L_1(w^t) + L_x L_2(w^t)) \\ &= \lim_{t\to 0} L_x L_2(w^t) \\ &= \frac{1}{2\sqrt{2(\|x\|^2 + \|y\|^2)}} L_x \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix} \end{aligned}$$

and

$$\lim_{t\to 0} L_y L_{u^t}^{-1} = \frac{1}{2\sqrt{2(\|x\|^2 + \|y\|^2)}} L_y \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix}.$$

- The case $(x, y) \in S_3$. From $w^t = 2t^2 e$, we have $L_{u^t}^{-1} = \frac{1}{\sqrt{2}t} I$, which implies

$$\begin{aligned} L_x L_{u^t}^{-1} &= 0 \cdot \frac{1}{\sqrt{2}t} I = 0, \\ L_y L_{u^t}^{-1} &= 0 \cdot \frac{1}{\sqrt{2}t} I = 0. \end{aligned}$$

Therefore, $J_x = J_y = 0$.

q.e.d.

Lemma 3.3

$$\partial_B H_{FB}(x, y) \ni \begin{bmatrix} I - V_x & \nabla F(x) \\ I - V_y & -I \end{bmatrix}$$

where

$$V_x = \begin{cases} L_x L_u^{-1} & \textit{if } (x,y) \in S_1 \\ \frac{1}{2\sqrt{2(\|x\|^2+\|y\|^2)}} L_x \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix} \pm \frac{1}{2} \begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} & \textit{if } (x,y) \in S_2 \\ \pm I & \textit{if } (x,y) \in S_3 \end{cases} \tag{3.16}$$

and

$$V_y = \begin{cases} L_y L_u^{-1} & \textit{if } (x,y) \in S_1 \\ \frac{1}{2\sqrt{2(\|x\|^2+\|y\|^2)}} L_y \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix} & \textit{if } (x,y) \in S_2 \\ \pm I & \textit{if } (x,y) \in S_3 \end{cases} \tag{3.17}$$

Proof. Since the Fisher-Burmeister is continuously differentiable on $(x,y) \in S_1$, it suffices to prove the cases $(x,y) \in S_2$ and $(x,y) \in S_3$.

Now we consider the point $\hat{z} = (\hat{x}, \hat{y}) = (x + \epsilon e, y)$ where $\epsilon \neq 0$ is sufficiently small. In this proof, we use the next notations,

$$\begin{aligned} w &= (w_1, w_2) = x^2 + y^2 = \left(\|x\|^2 + \|y\|^2,\ 2(x_1x_2 + y_1y_2)\right), \\ \hat{w} &= (\hat{w}_1, \hat{w}_2) = \hat{x}^2 + \hat{y}^2, \quad \hat{u} = (\hat{w})^{1/2}, \quad \tilde{w}_2 = \hat{w}_2/\|\hat{w}_2\|, \\ \hat{\lambda}_i &= \lambda_i(\hat{w}) = \hat{w}_1 + (-1)^i\|\hat{w}_2\| \quad (i = 1, 2). \end{aligned}$$

By simple calculation, we have

$$\hat{w}_1 = w_1 + 2\epsilon x_1 + \epsilon^2, \quad \hat{w}_2 = w_2 + 2\epsilon x_2, \tag{3.18}$$

$$\hat{\lambda}_i = w_1 + 2\epsilon x_1 + \epsilon^2 + (-1)^i\|w_2 + 2\epsilon x_2\| \quad (i = 1, 2). \tag{3.19}$$

- The case $(x,y) \in S_2$. Since $(x,y) \in S_2$, it follows from Lemma 3.1 that

$$\begin{aligned} \|\hat{w}_2\|^2 &= \|w_2 + 2\epsilon x_2\|^2 \\ &= \|w_2\|^2 + 4\epsilon x_2^T w_2 + 4\epsilon^2\|x_2\|^2 \\ &= w_1^2 + 4\epsilon x_1 w_1 + 4\epsilon^2 x_1^2 \\ &= (w_1 + 2\epsilon x_1)^2. \end{aligned}$$

  Because $w_1 > 0$ and ϵ is sufficiently small, we have

$$\|\hat{w}_2\| = w_1 + 2\epsilon x_1. \tag{3.20}$$

  Therefore from (3.19), we have

$$\begin{aligned} \hat{\lambda}_1 &= w_1 + 2\epsilon x_1 + \epsilon^2 - \|w_2 + 2\epsilon x_2\| = \epsilon^2 > 0 \\ \hat{\lambda}_2 &= w_1 + 2\epsilon x_1 + \epsilon^2 + \|w_2 + 2\epsilon x_2\| = 2(w_1 + 2\epsilon x_1) + \epsilon^2 > 0. \end{aligned} \tag{3.21}$$

So that $\hat{\lambda}_1 > 0$ implies that H_{FB} is continuously differtiable at $\hat{z}$.

Next we consider $\lim_{\epsilon\to\pm 0} L_{\hat{x}} L_{\hat{u}}^{-1}$ and $\lim_{\epsilon\to\pm 0} L_{\hat{y}} L_{\hat{u}}^{-1}$. We can rewrite $L_{\hat{u}}^{-1}$ as

$$L_{\hat{u}}^{-1} = \frac{1}{2\sqrt{\hat{\lambda}_1}} \begin{bmatrix} 1 & -\tilde{w}_2^T \\ -\tilde{w}_2 & \tilde{w}_2\tilde{w}_2^T \end{bmatrix} + \frac{1}{2\sqrt{\hat{\lambda}_2}} \begin{bmatrix} 1 & \tilde{w}_2^T \\ \tilde{w}_2 & \frac{4\sqrt{\hat{\lambda}_2}}{\sqrt{\hat{\lambda}_1}+\sqrt{\hat{\lambda}_2}}(I - \tilde{w}_2\tilde{w}_2^T) + \tilde{w}_2\tilde{w}_2^T \end{bmatrix}. \tag{3.22}$$

Now we define the first and second term by $\hat{L}_1$ and $\hat{L}_2$ respectively like previously. In the similar way of the proof of Lemma 3.2, we have

$$\lim_{\epsilon\to\pm 0} \hat{L}_2 = \frac{1}{2\sqrt{2(\|x\|^2 + \|y\|^2)}} \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix}, \tag{3.23}$$

since $\lim_{\epsilon\to\pm 0}\hat{w}_2 = \lim_{\epsilon\to\pm 0}(w_2 + 2\epsilon x_2) = w_2$, so that $\lim_{\epsilon\to\pm 0}\tilde{w}_2 = \lim_{\epsilon\to\pm 0}\hat{w}_2/\|\hat{w}_2\| = w_2/\|w_2\| = \bar{w}_2$.

It follows from the definition of $\hat{L}_1$ that

$$\begin{aligned} L_{\hat{x}}\hat{L}_1 &= \frac{1}{2\sqrt{\hat{\lambda}_1}} \begin{bmatrix} x_1 + \epsilon & x_2^T \\ x_2 & (x_1+\epsilon)I \end{bmatrix} \begin{bmatrix} 1 \\ -\tilde{w}_2 \end{bmatrix} [1 \quad -\tilde{w}_2^T] \\ &= \frac{1}{2|\epsilon|} \left\{ L_x \begin{bmatrix} 1 \\ -\tilde{w}_2 \end{bmatrix} [1 \quad -\tilde{w}_2^T] + \begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon I \end{bmatrix} \begin{bmatrix} 1 \\ -\tilde{w}_2 \end{bmatrix} [1 \quad -\tilde{w}_2^T] \right\}. \end{aligned} \tag{3.24}$$

Here, we consider the first term:

$$L_x \begin{bmatrix} 1 \\ -\tilde{w}_2 \end{bmatrix} = x \circ \begin{bmatrix} 1 \\ -\frac{\hat{w}_2}{\|\hat{w}_2\|} \end{bmatrix} = \begin{bmatrix} x_1 - \frac{x_2^T\hat{w}_2}{\|\hat{w}_2\|} \\ x_2 - \frac{x_1\hat{w}_2}{\|\hat{w}_2\|} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

since by using $\|\hat{w}_2\| = w_1 + 2\epsilon x_1$, $\hat{w}_2 = w_2 + 2\epsilon x_2$ and Lemma 3.1,

$$\begin{aligned} x_1 - \frac{x_2^T\hat{w}_2}{\|\hat{w}_2\|} &= \frac{1}{\|\hat{w}_2\|}\left\{x_1\|\hat{w}_2\| - x_2^T\hat{w}_2\right\} \\ &= \frac{1}{\|\hat{w}_2\|}\left\{x_1(w_1 + 2\epsilon x_1) - x_2^T(w_2 + 2\epsilon x_2)\right\} \\ &= \frac{1}{\|\hat{w}_2\|}\left\{x_1 w_1 - x_2^T w_2 + 2\epsilon(x_1^2 - \|x_2\|^2)\right\} = 0 \end{aligned}$$

and

$$\begin{aligned} x_2 - \frac{x_1^T\hat{w}_2}{\|\hat{w}_2\|} &= \frac{1}{\|\hat{w}_2\|}\{x_2\|\hat{w}_2\| - x_1\hat{w}_2\} \\ &= \frac{1}{\|\hat{w}_2\|}\{x_2(w_1 + 2\epsilon x_1) - x_1(w_2 + 2\epsilon x_2)\} \\ &= \frac{1}{\|\hat{w}_2\|}\{w_1 x_2 - x_1 w_2\} = 0. \end{aligned}$$

Therefore, we obtain

$$\lim_{\epsilon\to\pm 0} L_{\hat{x}}\hat{L}_1 = \lim_{\epsilon\to\pm 0}\frac{\mathrm{sgn}(\epsilon)}{2} \begin{bmatrix} 1 & -\tilde{w}_2^T \\ -\tilde{w}_2 & \tilde{w}_2\tilde{w}_2^T \end{bmatrix} = \pm\frac{1}{2} \begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix}.$$

Consequently, we have from the above results,

$$\begin{aligned}
\lim_{\epsilon\to\pm 0} L_{\hat{x}} L_{\hat{u}}^{-1} &= \lim_{\epsilon\to\pm 0}(L_{\hat{x}}\hat{L}_1 + L_{\hat{x}}\hat{L}_2)\\
&= \lim_{\epsilon\to\pm 0} L_{\hat{x}}\hat{L}_1 + \lim_{\epsilon\to\pm 0} L_{\hat{x}}\hat{L}_2\\
&= \pm\frac{1}{2}\begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} + \frac{1}{2\sqrt{2(\|x\|^2+\|y\|^2)}} L_x \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix}.
\end{aligned}$$

- The case $(x, y) \in S_3$. Because $(x, y) = (0, 0)$, we have $\hat{w} = \epsilon^2 e \in \mathrm{int}\,\mathcal{K}^n$, and hence ∇H is continuously differentiable on $\hat{z} = (\hat{x}, \hat{y})$.From the fact that $\hat{u} = (\hat{w})^{1/2} = |\epsilon| e$, we have

$$L_{\hat{x}} L_{\hat{u}}^{-1} = \mathrm{sgn}(\epsilon) I$$

which implies

$$\lim_{\epsilon\to\pm 0} L_{\hat{x}} L_{\hat{u}}^{-1} = \pm I$$

q.e.d

Definition 3..1 *Suppose that $\mathcal{F}$ is a continuous function and $\partial\mathcal{F}$ exists. Let $\mathcal{F}_u$ with $u \geq 0$ be a smoothing function. We say that $\mathcal{F}_u$ satisfies* **the Jacobian consistency** *if*

$$lim_{u\to 0}\, dist\,(\nabla\mathcal{F}_u(x), \partial\mathcal{F}(x)) = 0 \quad \textit{for any } x,$$

where $dist(X, S) = min\{\|X - Y\|,\ Y \in S\}$.

Theorem 3.1 *The smoothed Fischer-Burmeister function H_t satisfies the Jacobian consistency property.*

Proof. In Lemma 3.3, for the case S_2, let

$$V_x^{(i)} = L_x J + (-1)^i Z \quad \text{for } i = 1, 2$$

where $J := \frac{1}{2\sqrt{2(\|x\|^2+\|y\|^2)}} \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 4I - 3\bar{w}_2\bar{w}_2^T \end{bmatrix}$ and $Z := \frac{1}{2}\begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix}$.

Next,let

$$W^{(i)} = \begin{bmatrix} I - V_x^{(i)} & \nabla F(x) \\ I - V_y & -I \end{bmatrix}.$$

Then, from Lemma 3.3 we have $W^{(i)} \in \partial_B H_{FB}(x, y)$ for $i = 1, 2$. Therefore, $W := (W^{(1)} + W^{(2)})/2 \in \partial H_{FB}(x, y)$. From Lemma 3.2, we can prove the Jacobian consistency of H_t.

q.e.d

4. Update of an outside parameter

For updating of an outside parameter t, we need to estimate $\mathrm{dist}(\nabla H_t(x,y), \partial H_{FB}(x,y))$ in terms of t more preceisely, because we want to inform how small we may choose a parameter t.

Lemma 4.1 *Let $\alpha, \beta, \gamma \in R$ and $w \in R^{n-1}$ with $\|w\| = 1$. Let*

$$A = \begin{bmatrix} \beta & \gamma w^T \\ \gamma w & \alpha I + (\beta - \alpha) w w^T \end{bmatrix}$$

Then eigenvalues of the symmetric matrix A are α of multiplicity $n-2$ and $\beta \pm \gamma$.

Proof. If let $\delta := \beta - \alpha$, A becomes

$$G = \begin{bmatrix} \beta & \gamma w^T \\ \gamma w & \alpha I + \delta w w^T \end{bmatrix}.$$

$$\begin{aligned}
\det(\lambda I - G) &= \begin{vmatrix} \lambda - \beta & -\gamma w^T \\ -\gamma w & (\lambda - \alpha) I - \delta w w^T \end{vmatrix} \\
&= (\lambda - \beta)|(\lambda - \alpha) I - \delta w w^T - \gamma^2 (\lambda - \beta)^{-1} w w^T| \\
&= (\lambda - \beta)|(\lambda - \alpha)[I - (\lambda - \alpha)^{-1}(\delta + \gamma^2(\lambda - \beta)^{-1}) w w^T]| \\
&= (\lambda - \beta)(\lambda - \alpha)^{n-1}|I - (\lambda - \alpha)^{-1}(\delta + \gamma^2(\lambda - \beta)^{-1}) w w^T| \\
&= (\lambda - \beta)(\lambda - \alpha)^{n-2}[(\lambda - \alpha) - (\delta + \gamma^2(\lambda - \beta)^{-1})] \\
&= (\lambda - \beta)(\lambda - \alpha)^{n-2}[(\lambda - \alpha - \delta) - \gamma^2(\lambda - \beta)^{-1}] \\
&= (\lambda - \alpha)^{n-2}((\lambda - \beta)(\lambda - \alpha - \delta) - \gamma^2),
\end{aligned}$$

since $\det(I + uv^T) = 1 + u^T v$ for $u,\ v \in R^n$ and then $\|w\| = 1$ and $|tA| = t^p |A|$ for $A \in R^{p \times p}$ and $t \in R$. Therefore,

$$\begin{aligned}
\det(\lambda I - A) &= (\lambda - \alpha)^{n-2}[(\lambda - \beta)(\lambda - \alpha - \beta + \alpha) - \gamma^2] \\
&= (\lambda - \alpha)^{n-2}[(\lambda - \beta)^2 - \gamma^2] \\
&= (\lambda - \alpha)^{n-2}(\lambda - (\beta + \gamma))(\lambda - (\beta - \gamma)).
\end{aligned}$$

q.e.d

In the following lemma, we describe the property of some functions used in the proof of the next proposition.

Lemma 4.2 *Let $0 < \lambda_1 \le \lambda_2$. Then the functions $g_1,\ g_2 : (-\lambda_1, \infty) \to R$ defined by*

$$
\begin{aligned}
g_1(\tau) &:= \frac{1}{\sqrt{\lambda_1+\tau}} - \frac{2}{\sqrt{\lambda_1+\tau}+\sqrt{\lambda_2+\tau}} &(4.1)\\
g_2(\tau) &:= \frac{2}{\sqrt{\lambda_1+\tau}+\sqrt{\lambda_2+\tau}} - \frac{1}{\sqrt{\lambda_2+\tau}} &(4.2)
\end{aligned}
$$

are decreasing in $\tau > -\lambda_1$. In particular, if $\lambda_1 = \lambda_2$, then g_1 and g_2 are identically rero for all $\tau > -\lambda_1$. If $\lambda_1 < \lambda_2$, then g_1 and g_2 are strictly decreasing , and hence $g_1(0) > g_1(\tau)$, and $g_2(0) > g_2(\tau)$ for all $\tau > 0$.

Proof. The case $\lambda_1 = \lambda_2$ is evident. So, we let $\lambda_1 < \lambda_2$. We prove our assertion only for g_1. The proof for g_2 is similarly. We have

$$
\begin{aligned}
g_1'(\tau) &= -\frac{1}{2}\frac{1}{(\lambda_1+\tau)^{3/2}} + \frac{\frac{1}{\sqrt{\lambda_1+\tau}}+\frac{1}{\sqrt{\lambda_2+\tau}}}{(\sqrt{\lambda_1+\tau}+\sqrt{\lambda_2+\tau})^2}\\
&= \frac{1}{\sqrt{(\lambda_1+\tau)(\lambda_2+\tau)}(\sqrt{\lambda_1+\tau}+\sqrt{\lambda_2+\tau})} - \frac{1}{2}\frac{1}{(\lambda_1+\tau)^{3/2}}\\
&= \frac{1}{2\sqrt{\lambda_1+\tau}}\left(\frac{2}{\sqrt{(\lambda_1+\tau)(\lambda_2+\tau)}+\lambda_2+\tau} - \frac{1}{\lambda_1+\tau}\right)\\
&< \frac{1}{2\sqrt{\lambda_1+\tau}}\left(\frac{2}{\sqrt{(\lambda_1+\tau)(\lambda_1+\tau)}+\lambda_1+\tau} - \frac{1}{\lambda_1+\tau}\right) = 0
\end{aligned}
$$

for all $\tau > -\lambda_1$, where the last inequality follows from $\lambda_1 + \tau < \lambda_2 + \tau$. q.e.d

Proposition 4.1 *Let $z = (x, y) \in R^{2n}$. The z is used temporarily in this proposition. For any $t > 0$, we have*

$$
dist(\nabla H_t(z), \partial H_{FB}(z)) \le \Gamma(z)[h_0(z) - h_t(z)] \tag{4.3}
$$

where $\Gamma(z) := \|(L_x\ \ L_y)^T\|$ and $h_t : R^{2n} \to R_+$ be a function defined by

$$
h_t(z) = \begin{cases} \frac{1}{\sqrt{\lambda_1+2t^2}} & \text{if } z \in S_1\\ \frac{\sqrt{2}}{t+\sqrt{t^2+w_1}} & \text{if } z \in S_2\\ 0 & \text{if } z \in S_3 \end{cases}
$$

where $w_1 \in R$ is defined as $(w_1, w_2) = x^2 + y^2$ previously. Here $dist(X, S)$ denotes $min\{\|X - Y\| | Y \in S\}$.

Proof.

$$
\begin{aligned}
\mathrm{dist}(\nabla H_t(z), \partial H_{FB}(z)) &= \min\{\|\nabla H_t(z) - V\| \mid V \in \partial H_{FB}(z)\}\\
&\le \|\nabla H_t(z) - J_H^0(z)\|
\end{aligned}
$$

$$
\begin{aligned}
&= \left\| \begin{bmatrix} I - L_x L_{u^t}^{-1} & \nabla F(x) \\ I - L_y L_{u^t}^{-1} & -I \end{bmatrix} - \begin{bmatrix} I - J_x & \nabla F(x) \\ I - J_y & -I \end{bmatrix} \right\| \\
&= \left\| \begin{bmatrix} J_x - L_x L_{u^t}^{-1} & 0 \\ J_y - L_y L_{u^t}^{-1} & 0 \end{bmatrix} \right\| \\
&= \|J_x - L_x L_{u^t}^{-1}\| + \|J_y - L_y L_{u^t}^{-1}\|.
\end{aligned}
$$

Case 1. $z \in S_1$: Let $G := L_u^{-1} - L_{u^t}^{-1}$. Since $J_x = L_x L_u^{-1}$ and $J_y = L_y L_u^{-1}$ from Lemma 3.2, we have

$$
\begin{aligned}
\left\| \begin{bmatrix} J_x - L_x L_{u^t}^{-1} \\ J_y - L_y L_{u^t}^{-1} \end{bmatrix} \right\| &= \left\| \begin{matrix} L_x L_u^{-1} - L_x L_{u^t}^{-1} \\ L_y L_u^{-1} - L_y L_{u^t}^{-1} \end{matrix} \right\| \\
&= \left\| \begin{bmatrix} L_x \\ L_y \end{bmatrix} (L_u^{-1} - L_{u^t}^{-1}) \right\| \\
&\le \left\| \begin{matrix} L_x \\ L_y \end{matrix} \right\| \left\| (L_u^{-1} - L_{u^t}^{-1}) \right\| = \Gamma(z)\|G\|,
\end{aligned}
$$

where $\Gamma(z) := \left\| \begin{matrix} L_x \\ L_y \end{matrix} \right\|$ and $G := L_u^{-1} - L_{u^t}^{-1}$.

$$
\begin{aligned}
\|G\| = \|L_u^{-1} - L_{u^t}^{-1}\| &= \begin{bmatrix} b^0 - b^t & (c^0 - c^t)\bar{w}_2^T \\ (c^0 - c^t)\bar{w}_2 & (a^0 - a^t)I + [(b^0 - b^t) - (a^0 - a^t)]\bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&= \begin{bmatrix} \beta & \gamma\bar{w}_2^T \\ \gamma\bar{w}_2 & \alpha I + (\beta - \alpha)\bar{w}_2\bar{w}_2^T \end{bmatrix},
\end{aligned}
$$

where $\alpha := a^0 - a^t$, $\beta := b^0 - b^t$ and $\gamma := c^0 - c^t$. From Lemma 4.1, eigenvalues of G are α and $\beta \pm \gamma$.

$$
\begin{aligned}
\beta + \gamma &= \frac{1}{2}\left[\left(\frac{1}{\sqrt{\lambda_1}} + \frac{1}{\sqrt{\lambda_2}}\right) - \left(\frac{1}{\sqrt{\lambda_1^t}} + \frac{1}{\sqrt{\lambda_2^t}}\right)\right] + \frac{1}{2}\left[\left(\frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_1}}\right) - \left(\frac{1}{\sqrt{\lambda_2^t}} - \frac{1}{\sqrt{\lambda_1^t}}\right)\right] \\
&= \frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_2^t}} > 0.
\end{aligned}
$$

$$
\begin{aligned}
\beta - \gamma &= \frac{1}{2}\left[\left(\frac{1}{\sqrt{\lambda_1}} + \frac{1}{\sqrt{\lambda_2}}\right) - \left(\frac{1}{\sqrt{\lambda_1^t}} + \frac{1}{\sqrt{\lambda_2^t}}\right)\right] - \frac{1}{2}\left[\left(\frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_1}}\right) - \left(\frac{1}{\sqrt{\lambda_2^t}} - \frac{1}{\sqrt{\lambda_1^t}}\right)\right] \\
&= \frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_1^t}} > 0.
\end{aligned}
$$

$$
\begin{aligned}
\alpha &= a^0 - a^t \\
&= \frac{2}{\sqrt{\lambda_1} + \sqrt{\lambda_2}} - \frac{2}{\sqrt{\lambda_1^t} + \sqrt{\lambda_2^t}} \\
&= \frac{\sqrt{\lambda_1^t} - \sqrt{\lambda_1} + (\sqrt{\lambda_2^t} - \sqrt{\lambda_2})}{(\sqrt{\lambda_1} + \sqrt{\lambda_2})(\sqrt{\lambda_1^t} + \sqrt{\lambda_2^t})} > 0.
\end{aligned}
$$

Thus, the G is a positive definite matrix with the norm $\|G\| = \max\{\alpha, \beta \pm \gamma\}$. We have the relation $\beta-\gamma \geq \alpha \geq \beta+\gamma$. We first show that $\beta-\gamma \geq \alpha$ holds.

$$\begin{aligned}
\beta-\gamma-\alpha &= \frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_1^t}} - \left(\frac{2}{\sqrt{\lambda_1}+\sqrt{\lambda_2}} - \frac{2}{\sqrt{\lambda_1^t}+\sqrt{\lambda_2^t}}\right) \\
&= \frac{1}{\sqrt{\lambda_1}} - \frac{2}{\sqrt{\lambda_1}+\sqrt{\lambda_2}} - \left(\frac{1}{\sqrt{\lambda_1^t}} - \frac{2}{\sqrt{\lambda_1^t}+\sqrt{\lambda_2^t}}\right) \\
&= g_1(0) - g_1(2t^2) \geq 0,
\end{aligned}$$

since the last inequarity follows by Lemma 4.2. We next show that $\alpha \geq \beta+\gamma$.

$$\begin{aligned}
\alpha-(\beta+\gamma) &= \frac{2}{\sqrt{\lambda_1}+\sqrt{\lambda_2}} - \frac{2}{\sqrt{\lambda_1^t}+\sqrt{\lambda_2^t}} - \left(\frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_2^t}}\right) \\
&= \frac{2}{\sqrt{\lambda_1}+\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_2}} - \left(\frac{2}{\sqrt{\lambda_1^t}+\sqrt{\lambda_2^t}} - \frac{1}{\sqrt{\lambda_2^t}}\right) \\
&= g_2(0) - g_2(2t^2) \geq 0,
\end{aligned}$$

since the last inequarity follows by Lemma 4.2. From those results, we have $\beta-\gamma \geq \alpha \geq \beta+\gamma$ and hence

$$\|G\| = \beta-\gamma = \frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_1^t}} = h_0(z) - h_t(z).$$

Case 2. $z \in S_2$: In this case, $2w_1 = \lambda_2 > \lambda_1 = 0$. In the proof of Lemma 3.2, we defined $L_{u^t}^{-1} = L_1(w^t) + L_2(w^t)$. And we got $L_xL_1(w^t) = 0$, $L_yL_1(w^t) = 0$, $L_xL_1(w) = 0$, $L_yL_1(w) = 0$ since $\lim_{t\to 0} L_1(w^t) = L_1(w)$ and $\lim_{t\to 0} L_2(w^t) = L_2(w)$. Thus, since we have $L_yL_{u^t}^{-1} = L_yL_2(w^t)$ and $L_xL_{u^t}^{-1} = L_xL_2(w^t)$,

$$\begin{aligned}
\left\| \begin{matrix} J_x - L_xL_{u^t}^{-1} \\ J_y - L_yL_{u^t}^{-1} \end{matrix} \right\| &= \left\| \begin{matrix} L_xL_2(w) - L_xL_2(w^t) \\ L_yL_2(w) - L_yL_2(w^t) \end{matrix} \right\| \\
&= \left\| \begin{bmatrix} L_x \\ L_y \end{bmatrix} (L_2(w) - L_2(w^t)) \right\| = \Gamma(z)\|G\|,
\end{aligned}$$

where $G = L_2(w) - L_2(w^t)$. Here, by using $a^0 = 2/\sqrt{\lambda_2}$ and (3.15),

$$\begin{aligned}
G &= L_2(w) - L_2(w^t) \\
&= \frac{1}{2\sqrt{\lambda_2}} \begin{bmatrix} 1 \\ \bar{w}_2 \end{bmatrix} [1\ \bar{w}_2^T] + a^0 \begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix} - \left(\frac{1}{2\sqrt{\lambda_2(w^t)}} \begin{bmatrix} 1 \\ \bar{w}_2 \end{bmatrix} [1\ \bar{w}_2^T] + a^t \begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \right) \\
&= \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_2(w^t)}}\right) \begin{bmatrix} 1 \\ \bar{w}_2 \end{bmatrix} [1\ \bar{w}_2^T] + (a^0 - a^t) \begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&= \beta \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} + \alpha \begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} \beta & \beta \bar{w}_2^T \\ \beta \bar{w}_2 & \alpha I + (\beta - \alpha)\bar{w}_2 \bar{w}_2^T \end{bmatrix},$$

where $\alpha := a^0 - a^t$ and $\beta := (1/2)(1/\sqrt{\lambda_2} - 1/\sqrt{\lambda_2^t})$. Again by Lemma 4.1, eigenvalues of G are α and $\beta \pm \beta$ i.e., 0 and 2β. By the same reason as in Case 1, we have $\alpha > 0$ and $\beta > 0$. Thus, G is a positive semidefinite matrix with norm $\|G\| = \max\{\alpha,\ 2\beta\}$. We show that $\alpha > 2\beta$ holds. By using $2w_1 = \lambda_2 > \lambda_1 = 0$,

$$\begin{aligned} \alpha = a^0 - a^t &= \frac{2}{\sqrt{\lambda_1} + \sqrt{\lambda_2}} - \frac{2}{\sqrt{\lambda_1(w^t)} + \sqrt{\lambda_2(w^t)}} \\ &= \frac{2}{\sqrt{2w_1}} - \frac{2}{\sqrt{2t^2} + \sqrt{2t^2 + 2w_1}} \\ &= \sqrt{2}\left(\frac{1}{\sqrt{w_1}} - \frac{1}{\sqrt{t^2} + \sqrt{t^2 + w_1}}\right). \end{aligned}$$

$$\begin{aligned} 2\beta &= \frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_2(w^t)}} = \frac{1}{\sqrt{2w_1}} - \frac{1}{\sqrt{2t^2 + 2w_1}} \\ &= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{w_1}} - \frac{1}{\sqrt{t^2 + w_1}}\right). \end{aligned}$$

From this result, we have

$$\begin{aligned} \sqrt{2}(\alpha - 2\beta) &= 2\left(\frac{1}{\sqrt{w_1}} - \frac{1}{\sqrt{t^2} + \sqrt{t^2 + w_1}}\right) - \left(\frac{1}{\sqrt{w_1}} - \frac{1}{\sqrt{t^2 + w_1}}\right) \\ &= \frac{1}{\sqrt{w_1}} - \frac{2}{\sqrt{t^2} + \sqrt{t^2 + w_1}} + \frac{1}{\sqrt{t^2 + w_1}} \\ &= \left(\frac{1}{\sqrt{w_1}} - \frac{1}{\sqrt{t^2} + \sqrt{t^2 + w_1}}\right) + \left(\frac{1}{\sqrt{t^2 + w_1}} - \frac{1}{\sqrt{t^2} + \sqrt{t^2 + w_1}}\right) > 0, \end{aligned}$$

where the last inequality follows since each term is positive for $t > 0$. Hence, we have $\|G\| = \alpha = \sqrt{2}/\sqrt{w_1} - \sqrt{2}/(t + \sqrt{t^2 + w_1}) = h_0(z) - h_t(z)$.

Case 3. $z \in S_3$. In Lemma 3.2, we have $J_x = 0$ and $J_y = 0$ on S_3. And in the proof of Lemma 3.2, we have $L_x L_{u^t}^{-1} = 0$ and $L_y L_{u^t}^{-1} = 0$ on S_3. Thus, we obtain

$$\left\| \begin{matrix} J_x - L_x L_{u^t}^{-1} \\ J_y - L_y L_{u^t}^{-1} \end{matrix} \right\| = 0.$$

q.e.d

Proposition 4.2 *Let $z = (x, y) \in R^{2n}$. And let $\gamma(z)$ be any function such that $\Gamma(z) \leq \gamma(z)$. For a given $\delta > 0$, we have*

$$\exists \bar{t}(z, \delta) > 0 : \qquad dist(\nabla H_t(z), \partial H_{FB}(z)) < \delta \tag{4.4}$$

for any $t > 0$ *such that* $0 < t < \bar{t}(z,\delta)$. *The* $\bar{t}(z,\delta)$ *is defined by*

$$\bar{t}(z,\delta) = \begin{cases} \frac{\lambda_1 \delta}{\sqrt{2(\gamma(z)^2 - \lambda_1 \delta^2)}} & \text{for } z \in S_1 \quad \text{and} \quad \delta < \gamma(z)/\sqrt{\lambda_1} \\ \frac{w_1 \delta}{2\sqrt{\gamma(z)(2\gamma(z) - \delta^2\sqrt{2w_1})}} & \text{for } z \in S_2 \quad \text{and} \quad \delta < 2\gamma(z)/\sqrt{2w_1} \\ \infty & \text{otherwise.} \end{cases}$$

Proof. We put $\gamma := \gamma(z)$ and $\bar{t} := \bar{t}(z,\delta)$ for simplicity.

Case 1. $z \in S_1$. If $\delta \geq \gamma/\sqrt{\lambda_1}$, then $h_0(z) - h_t(z) < h_0(z) = 1/\sqrt{\lambda_1} \leq \delta/\gamma$. Let $\delta < \gamma/\sqrt{\lambda_1}$. Since $\sqrt{a} - \sqrt{b} < \sqrt{a-b}$ for all $a > b > 0$, we have

$$\begin{aligned} h_0(z) - h_t(z) &= \frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_1 + 2t^2}} = \frac{\sqrt{\lambda_1 + 2t^2} - \sqrt{\lambda_1}}{\sqrt{\lambda_1}\sqrt{\lambda_1 + 2t^2}} \\ &< \frac{\sqrt{2}t}{\sqrt{\lambda_1}\sqrt{\lambda_1 + 2t^2}} = \frac{\sqrt{2}}{\sqrt{\lambda_1}\sqrt{\lambda_1/t^2 + 2}} \\ &< \frac{\sqrt{2}}{\sqrt{\lambda_1}\sqrt{\lambda_1/\bar{t}^2 + 2}}, \end{aligned}$$

since the last function is a strictly increasing one for $t > 0$. Therefore, we obtain

$$\begin{aligned} h_0(z) - h_t(z) &< \frac{\sqrt{2}}{\sqrt{\lambda_1}\sqrt{\lambda_1/\bar{t}^2 + 2}} = \frac{\sqrt{2}}{\sqrt{\lambda_1}\sqrt{\lambda_1\left(\frac{2(\gamma^2 - \lambda_1\delta^2)}{(\lambda_1\delta)^2}\right) + 2}} \\ &= \frac{1}{\sqrt{\lambda_1}\sqrt{\frac{\gamma^2}{\lambda_1\delta^2}}} = \frac{\delta}{\gamma}. \end{aligned}$$

Case 2. $z \in S_2$. If $\delta \geq 2\gamma/\sqrt{2w_1}$, then

$$\begin{aligned} h_0(z) - h_t(z) &= \frac{\sqrt{2}}{\sqrt{w_1}} - \frac{\sqrt{2}}{t + \sqrt{t^2 + w_1}} \\ &= \frac{\sqrt{2}(t + \sqrt{t^2 + w_1} - \sqrt{w_1})}{\sqrt{w_1}(t + \sqrt{t^2 + w_1})} < \frac{\sqrt{2}(t + t)}{\sqrt{w_1}(t + \sqrt{t^2 + w_1})} \\ &= \frac{2\sqrt{2}t}{\sqrt{w_1}(t + \sqrt{t^2 + w_1})} = \frac{2\sqrt{2}}{\sqrt{w_1}(1 + \sqrt{1 + \frac{w_1}{t^2}})} \\ &< \frac{2\sqrt{2}}{\sqrt{w_1}(1 + \sqrt{1 + \frac{w_1}{\bar{t}^2}})} = \frac{2\sqrt{2}}{\sqrt{w_1}\left(1 + \sqrt{1 + \frac{w_1}{\left(\frac{w_1\delta^2}{2\sqrt{\gamma(2\gamma - \delta\sqrt{2w_1})}}\right)^2}}\right)} \\ &= \frac{2\sqrt{2}}{\sqrt{w_1}\left(1 + \sqrt{\frac{(2\sqrt{2}\gamma - \delta\sqrt{w_1})^2}{(\sqrt{w_1}\delta)^2}}\right)} = \frac{2\sqrt{2}}{\sqrt{w_1}\left(1 + \frac{2\sqrt{2}\gamma - \delta\sqrt{w_1}}{\sqrt{w_1}\delta}\right)} = \frac{\delta}{\gamma}. \end{aligned}$$

q.e.d

5. Algorithm and its Properties

5.1 Algorithm

In this section, we propose an algorithm of our smoothing method. For simplicity, we put $z^{(k)} := (x^{(k)}, y^{(k)})$.

Algorithm 5.1

Step 0. Choose $\eta, \rho \in (0,1), \bar{\eta} \in (0, \eta],\ \sigma \in (0, 1/2),\ \kappa > 0$, and $\hat{\kappa} > 0$. Choose $z^{(0)} = (x^{(0)}, y^{(0)}) \in R^{2n}$ and $\beta_0 \in (0, \infty)$. Let $t_0 := \|H_{FB}(z^{(0)})\|$. Set $k := 0$.

Step 1. If $\|H_{FB}(z^{(k)})\| = 0$ is satisfied, then stop.

Step 2.

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

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

$$H_{t_k}(v^{(j)}) + \nabla H_{t_k}(v^{(j)})^T d^{(j)} = 0. \tag{5.1}$$

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

Step 2.3. Let m_j be the smallest nonnegative integer m satisfying

$$\psi_{t_k}(v^{(j)} + \rho^m d^{(j)}) \le (1 - 2\sigma\rho^m)\Psi_{t_k}(v^{(j)}). \tag{5.2}$$

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

Step 2.4. If

$$\|H_{t_k}(v^{(j+1)})\| \le \beta_k, \tag{5.3}$$

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

Step 3. Update the parameters as follows:

$$\begin{aligned} \delta_{k+1} &:= \|H_{FB}(z^{(k+1)})\| \\ t_{k+1} &:= \min\left\{\kappa\delta_{k+1}^2,\ t_0\bar{\eta}^{k+1},\ \bar{t}(z^{(k+1)}, \hat{\kappa}\delta_{k+1})\right\} \\ \beta_{k+1} &:= \beta_0\eta^{k+1}. \end{aligned}$$

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

5.2 Convergence

To show the global convergence property of Algorithm 5.1, we introduce the following lemma.

Lemma 5.1 (Mountain Pass Theorem) *Let $\theta : R^n \to R$ be a continuously differentiable and level-bounded function. Let $C \subset R^n$ be a nonempty and compact set and ξ be a minimum value of θ on the boundary of C, that is,*

$$\xi := min_{x\in\partial C}\theta(x).$$

Assume that there exit points $p \in C$ and $q \notin C$ such that $\theta(p) < \xi$ and $\theta(q) < \xi$. Then, there exists a point $r \in R^n$ such that $\nabla\theta(r) = 0$ and $\theta(r) \geq \xi$.

Lemma 5.2 *If $f : R^n \to R^n$ is monotone, then for any $t > 0$, every satationary point $(\bar{x}, \bar{y})$ of the function Ψ_t satisfies $\Psi_t(\bar{x}, \bar{y}) = 0$.*

Proof.

$$\nabla\Psi_t(\bar{x}, \bar{y}) = 2\nabla H_t(\bar{x}, \bar{y})H_t(\bar{x}, \bar{y}) = 0.$$

Since $\nabla H_T(\bar{x}, \bar{y})$ is nonsingular, we have $H_t(\bar{x}, \bar{y}) = 0$, that is, $\Psi_t(\bar{x}, \bar{y}) = 0$.

q.e.d.

From (3.4), we have the following lemma:

Lemma 5.3 *Let $C \subset R^n \times R^n$ be a compact set. Then, for any given $\delta > 0$, there exist $t' > 0$ such that*

$$|\Psi_t(x, y) - \Psi_{FB}(x, y)| \leq \delta$$

for any $(x, y) \in C$ and $t \in [0, t']$.

By using Mountain Pass Theorem, Lemma 5.2 and Lemma 5.3, we can establish the following convergence theorem of Algorithm 5.1.
Note that for using Mountain Pass Theorem, we need the level boundness of $\Psi_{FB}(\cdot)$. The fact is proved by S. Pan and J-S. Chen [19] for a function with the uniform Cartesian P-property (is defined below) and satisfies some condition (that is Condition A).

Definition 5..1 *Given a mapping $f = (f_1, ..., f_q)$ with $f_i : R^n \to R^{n_i}$, where $n_1 + ... + n_q = n$, f is said to have* **the uniform Cartesian P-property** *if for any $x = (x_1, ..., x_q)$, $y = (y_1, ..., y_q) \in R^n$, there is an index $v \in \{1, 2, ..., q\}$ and a positive constant $\rho > 0$ such that*

$$< x_v - y_v,\ f_v(x) - f_v(y) > \geq\ \rho\|x - y\|^2.$$

Theorem 5.1 *Let $f : R^n \to R^n$ be a function with the uniform Cartesian P-property and satisfies Condition A. Assume that the solution set S of SOCCP (3.2) is nonempty and bounded. Let $\{(x^k, y^k)\}$ be a sequence generated by Algorithm 5.1. Then, $\{(x^k, y^k)\}$ is bounded, and every accumulation point is a solution of SOCCP (3.2).*

Proof. It is sufficient to show only the boundedness of $\{(x^k, y^k)\}$. Assume that $\{(x^k, y^k)\}$ by Algorithm 5.1 is not bounded.

- $\exists\, K:$ subsequence $\{(x^k, y^k)\}_{k\in K}$ such that $\lim_{k\to\infty, k\in K}\|(x^k, y^k)\| = \infty$.
- $\exists$ compact set $C \subset R^n \times R^n$ such that $S \subset \text{int}\, C$ because of boundness of S.

Thus, we have

(a) $(x^k, y^k) \notin C$ for $\forall k \in K$ and $k \gg 0$.

(b) $\bar{\xi} := \min_{(x,y)\in\partial C} \Psi_{FB}(x, y) > 0$.

From Lemma 5.3 with $\delta := \bar{\xi}/4 > 0$,

$$\Psi_{t_k}(x, y) - \Psi_{FB}(x, y) \le \frac{1}{4}\bar{\xi} \tag{5.4}$$

$$\Psi_{t_k}(x, y) - \Psi_{FB}(x, y) \ge -\frac{1}{4}\bar{\xi} \tag{5.5}$$

for $\forall(x, y) \in C$ and $k \gg 0$. Let $(\bar{x}, \bar{y}) \in S \subset C$. From (5.4),

$$\Psi_{t_k}(\bar{x}, \bar{y}) - \Psi_{FB}(\bar{x}, \bar{y}) = \Psi_{t_k}(\bar{x}, \bar{y}) \le \frac{1}{4}\bar{\xi} \tag{5.6}$$

for $\forall k \in K,\ \ k \gg 0$. On the other hand, letting $(\bar{x^k}, \bar{y^k})$ be min of $\Psi_{t_k}(x, y)$ on ∂C, for $\forall k \in K$ and $k \gg 0$,

$$\begin{aligned}\min_{(x,y)\in\partial C}\Psi_{t_k}(x, y) &= \Psi_{t_k}(\bar{x^k}, \bar{y^k}) \\ &\ge -\frac{1}{4}\bar{\xi} + \Psi_{FB}(\bar{x^k}, \bar{y^k}) \\ &\ge -\frac{1}{4}\bar{\xi} + \bar{\xi} = \frac{3}{4}\bar{\xi},\end{aligned} \tag{5.7}$$

since (5.5) and (b). Furthermore, from $\|H_{t_k}(\cdot)\| \le \beta_k$ in Step 2 of Algorithm 5.1 and $\Psi_{t_k}(\cdot) = \frac{1}{2}\|H_{t_k}(\cdot)\|^2$, we can put

$$\Psi_{t_k}(x^{k+1}, y^{k+1}) \le \frac{1}{4}\bar{\xi} \tag{5.8}$$

for $\forall k \in K,\ \ k \gg 0$. Now let $k \in K,\ \ k \gg 0$ satisfying (a), (5.6), (5.7) and (5.8). Then by applying Mountain Pass Theorem to Ψ_{t_k} with $(\bar{x}, \bar{y}) \in C$ and $(x^{k+1}, y^{k+1}) \notin C$ and $\xi := \min_{(x,y)\in\partial C}\Psi_{t_k}(x, y) \ge (3/4)\bar{\xi} > 0$, we obtain

$$\exists r := (\hat{x}, \hat{y}) \quad \text{such that} \quad \nabla\Psi_{t_k}(r) = 0 \quad \text{and} \quad \Psi_{t_k}(r) \ge \frac{3}{4}\bar{\xi}.$$

This contradicts Lemma 5.2. Therefore, $\{(x^k, y^k)\}$ is bounded.

q.e.d.

Theorem 5.2 *f is monotone. Let (x^k, y^k) by Algorithm 5.1. Assume that*

(i) *The solution set of SOCCP (3.2) is nonempty and bounded*

(ii) *Any accumulation point of $\{\nabla H_{t_k}(x^k, y^k)\}$ is nonsigular*

then the sequence $\{(x^k, y^k)\}$ converges to a solution (x^*, y^*) of SOCCP (3.2) quadratically.

Proof. Let $z := (x, y)$. Let $H_{FB}(z^*) = 0$. Note

$$\exists\, C > 0 \qquad \text{such that} \qquad \|\nabla H_{t_k}(z^k)^{-T}\| \le C \quad \text{for} \;\; k \gg 0.$$

We show $\|z^k + d^k - z^*\| = O(\|z^k - z^*\|^2)$ for $k \gg 0$.

$$\begin{aligned}
\|z^k + d^k - z^*\| &= \|z^k - \nabla H_{t_k}(z^k)^{-T} H_{t_k}(z^k) - z^*\| \\
&\le \|\nabla H_{t_k}(z^k)^{-T}\| \|\nabla H_{t_k}(z^k)^T (z^k - z^*) - H_{t_k}(z^k)\| \\
&\le C\left\{\|(\nabla H_{t_k}(z^k) - V_k)^T (z^k - z^*)\| + \|V_k^T (z^k - z^*) - H'_{FB}(z^*; z^k - z^*)\|\right\} \\
&+ C\|H'_{FB}(z^*; z^k - z^*) - H_{FB}(z^k) + H_{FB}(z^*)\| \\
&+ C\|H_{FB}(z^k) - H_{t_k}(z^k)\|, \qquad (5.9)
\end{aligned}$$

where $V_k \in \partial H_{FB}(z)$.

From the fact that H_t satisfies the Jacobian consistency, the definition of $\bar{t} := \bar{t}(z^{k+1}, \hat{\kappa}\|H_{FB}(z^{k+1})\|)$ in Step 3 of Algorithm 5.1 and the local Lipchitz continuity of H_{FB},

$$\begin{aligned}
\|\nabla H_{t_k}(z^k) - V_k\| &\le \hat{\kappa}\|H_{FB}(z^k)\| \\
&= \hat{\kappa}\|H_{FB}(z^k) - H_{FB}(z^*)\| \\
&\le \hat{\kappa} L\|z^k - z^*\|. \qquad (5.10)
\end{aligned}$$

By (5.10), the first term of the second inequarity of (5.9) is $O(\|z^k - z^*\|^2)$. The second and third term of the second inequarity of (5.9) are also $O(\|z^k - z^*\|^2)$, since H_{FB} is strongly semismooth and directionary differentiable, because the strongly semismooth of ϕ_{FB} has been proved in Corollary 3.3 in [18]. For the last term of the second inequarity of (5.9),

$$\begin{aligned}
\|H_{t_k}(z^k) - H_{FB}(z^k)\| &\le \sqrt{2}\|t_k| = O(\|H_{FB}(z^k)\|^2) \\
&= O(\|z^k - z^*\|^2)
\end{aligned}$$

since $\|H_t(z) - H_{FB}(z)\| \le \sqrt{2}t$ because we get $\|H_t(z) - H_{FB}(z)\| = \|\phi_t(z) - \phi_{FB}(z)\| \le \sqrt{2}t$ by Fukushima et al. [13]. Step 3 of Algorithm 5.1, $H_{FB}(z^*) = 0$ and local Lipsctitz continuity of H_{FB}. Consequently, we have

$$\|z^k + d^k - z^*\| = O(\|z^k - z^*\|^2).$$

Thus, we can get the result.

q.e.d.

5.3 Numerical experiments

We executed numerical experiments to compare the algorithm proposed with Algorithm 2 in Hayashi et al. [14]. The program was coded in MATLAB 7. The computation was carried out on a Compaq nx 9030.

5.3.1 Linear case

The problem is to find $(x, y) \in R^n \times R^n$ such that

$$x \in \mathcal{K}, \qquad y \in \mathcal{K}, \qquad x^T y = 0, \qquad y = Mx + q,$$

where $M \in R^{n\times n}$ is a rank-deficient positive semidefinite matrix, $q \in R^n$, and $\mathcal{K} \subset R^n$ is the Cartesian product of second-order cones, that is, $\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m}$ with $n = n_1 + \cdots + n_m$. We choose $\gamma = 0.4$, $\rho = 0.5$, $\sigma = 0.4$ κ and $\bar{\kappa} = 1.0$ in the algorithm.

In order to obtain a positive semidefinite matrix M with rank $M = r < n$, we let $M := nBB^T/\|BB^T\|$, where $B \in R^{n\times r}$ is a matrix of which components are randomly chosen from the interval $[-1, 1]$. Furthermore, we let $q := 10^\alpha n^{1/2} p - Me$, where $e = (1, 0, ...0)^T \in \mathrm{int}\mathcal{K}^n$, and α is randomly chosen from the interval $[-1, 1]$. The vector $p := 2^{-1/2}\cos\theta(1, w/\|w\|) + 2^{-1/2}\sin\theta(1, -w/\|w\|)$ is chosen as a vector such that $p \in \mathrm{int}\mathcal{K}^n$ and $\|p\| = 1$, where the components of w are randomly chosen from the interval $[-1, 1]$, and $\theta = \pi/5$. This technique is due to Hayashi et al. [14] except for using fixed $\theta = \pi/5$.

When we have $r = 0.7 * n$ and $\mathcal{K} = [2, 2, n-6, 1, 1]$, Table 5.1 shows the number of iterations (Iter) and CPU time in second (CPU) required to solve test problems of size n for the proposed method and Hayashi method [14]. The results are those for the average of some runs for each n.

Table 5.1

	Proposed	Method	Hayashi	Method
n	Iter	CPU	Iter	CPU
500	4	12.9186	6	7.5308
800	4	45.7157	6	25.4065
1000	4	74.7775	6	46.1363

Acknowledgment

The author would like to express my appreciation to Associate Professor H. Ogasawara of Tokyo University of science for the idea of the Jacobian consistency property and update of an outside parameter.

References

[1] F. Alizadeh and D. Goldfarb, Second-order cone programmming, *Mathematical Programming* Vol. 95 (2003) 3-51.

[2] Z. Cai and K.-C. Toh, Solving second order cone programming via a reduced augmented system approach, *SIAM Journal on Optimization* Vol. 17 (2006) 711-737.

[3] C. Chen and O. L. Mangasarian, Smoothing methods for convex inequalities and linear complementarity problems, *Mathematical Programming* Vol. 71 (1995) 51-69.

[4] X. Chen, L. Qi and D. Sun, Global and superlinear convergence of smoothing Newton method and its application to general box constrained variational inequarities, *Mathematics of Computation* Vol. 67 (1998) 519-540.

[5] X. Chen, Smoothing methods for complementarity problems and their applications: a survey, *Journal of Operations Research Society of Japan* Vol. 43 (2000) 32-47.

[6] X. D. Chen, D. Sun and J. Sun, Complemenarity functions and numerical experiments on some smoothing Newton Methods for second-order-cone complementartiy problems, *Computaional Optimization and Applications* Vol.25 (2003) 39-56.

[7] J. S. Chen, X. Chen and P. Tseng, Analysis of nonsmooth vector-valued functions associated with second-order cones, *Mathematical Programming* Vol. 101 (2004) 95-117.

[8] J. S. Chen and P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, *Mathematical Programming* Vol. 104 (2005) 293-327.

[9] J. S. Chen, Two classes of merit functions for the second-order cone complementarity problem, *Mathematical Methods of Operations Research* Vol. 64 (2006) 495-519.

[10] J. S. Chen, A new merit function and its related properties for the second-order cone complementarity problem, *Pacific Journal of Optimization* Vol. 2 (2006) 167-179.

[11] F. H. Clarke, *Optimization and Nonsmooth Analysis,* (Wiley, New York, 1983).

[12] F. Facchinei and J. - S. Pang, *Finite-Dimensional Variational Inequalities and Complementartiy Problems,* (Springer-Verlag, New York, 2003).

[13] M. Fukushima, Z. Q. Luo and P. Tseng, Smoothing functions for second-order-cone complementarity problems, *SIAM Journal on Optimization* Vol. 12 (2001) 436-460.

[14] S. Hayashi, N. Yamashita and M. Fukushima, A combined smoothing and regularation method for monotone second-order-cone complementarity problems, *SIAM Journal on Optimization* Vol. 15 (2005) 593-615.

[15] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, *SIAM Journal on Control Opimization* Vol. 15 (1977) 959-972.

[16] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, *SIAM Journal on Matrix Analysis and Applications* Vol. 17 (1996) 851-868.

[17] R. D. C. Monteiro and T. Tsuchiya, Polynomial convergence of primal - dual algorithms for the second - order cone program based on the MZ - family of directions, *Mathematical Programming* Vol. 88 (2000) 61-83.

[18] D. Sun and J. Sun, Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions, *Mathematical Programming* Vol. 103 (2005) 575-581.

[19] S. Pan and J-S. Chen, A damped Gauss - Newton method for the second - order cone complementarity problem, *Applied Mathematics and Optimization* Vol. 59 (2009) 293-318.

[20] L. Qi and D. Sun, A nonsmooth version of Newton's method, *Mathematical Programming* Vol. 58 (1993) 353-367.

[21] L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Mathematical Programming* Vol. 87 (2000) 1-35.

[22] T. Tsuchiya, A polynomiyal primal - dual path - following algorithm for second - order cone programming, *Reseach Memorandum* No. 649 (1997) The Institute of Statistical Mathematics, Tokyo, Japan, October (Reviesed: December 1997).

[23] T. Tsuchiya, A convergence analysis of the scaling - invariant primal - dual path - following algorithms for second - order cone programming, *Optim. Methods Software* Vol. 11 and 12 (1999) 141-181.

Appendix

A Level-boundednes of the merit function and its smoothing function

We prove that if f has the uniform Cartesian P-property and satisfies some condation(Condition A), then the merit function ψ_{FB} and its smoothing function ψ_t are level-bounded. Since the level-boundedness of a function $\psi : R^n \to R$ is equivalent to

$$\lim_{\|z\|\to\infty}\psi(z) = +\infty, \tag{1.1}$$

we show (1.1)instead of the boundedness of the level set $L_\alpha := \{z|\psi(z) \le \alpha\}$. Now, let the function $\tilde{\psi}_{FB} : R^n \to R$ be defined by

$$\tilde{\psi}_{FB}(x) := \|\phi_{FB}(x, f(x))\|.$$

Condition A[19] For any $\{x^k\} \subseteq R^n$ such that $\|x^k\| \to +\infty$, if there exists $i \in \{1, 2, .., m\}$ such that $\lambda_1(x_i^k),\ \lambda_1(f_i(x^k)) > -\infty$ and $\lambda_2(x_i^k),\ \lambda_2(f_i(x^k)) \to \infty$, then

$$\limsup_{k\to+\infty} \left\langle \frac{x_i^k}{\|x_i^k\|}, \frac{f_i(x^k)}{\|f_i(x^k)\|} \right\rangle > 0.$$

Lemma A.1[19] If f has the uniform Cartesian P-property and satisfies Condition A, then the merit function $\tilde{\psi}_{FB}(x)$ is level-bounded.

Lemma A.2 $\tilde{\psi}_{FB}(x)$ is level-bounded if and only if $\psi_{FB}(x)$ is level-bounded.

Proof. $\Longleftarrow$ Suppose that

$$\lim_{\|(x,y)\|\to\infty}\psi_{FB}(x, y) = +\infty. \tag{1.2}$$

Let $\{x^{(k)}\}$ be an arbitrary sequence such that $\|x^{(k)}\| \to \infty$ and $\{y^{(k)}\}$ be the corresponding sequence such that $y^{(k)} = f(x^{(k)})$ for all k. Then we have

$$\begin{aligned} \frac{1}{2}\tilde{\psi}_{FB}(x^{(k)})^2 &= \frac{1}{2}\|\phi_{FB}(x^{(k)}, f(x^{(k)})\|^2 \\ &= \frac{1}{2}\|\phi_{FB}(x^{(k)}, y^{(k)})\|^2 + \frac{1}{2}\|f(x^{(k)}) - y^{(k)}\|^2 \\ &= \psi_{FB}(x^{(k)}, y^{(k)}), \end{aligned}$$

where the second equality follows from $y^{(k)} = f(x^{(k)})$. From that $\|(x^{(k)}, y^{(k)})\| \to \infty$ and (1.2), we have $\lim_{k\to\infty}\tilde{\psi}_{FB}(x^{(k)}) = +\infty$.

$\Longrightarrow$ Suppose that $\tilde{\psi}_{FB}(x)$ is level-bounded. From $\sqrt{2\|\xi\|+2\|\eta\|} \geq \|\xi\|+\|\eta\|$ for any $\xi,\ \eta \in R^n$, we have

$$\begin{aligned}
2\sqrt{\psi_{FB}(x,y)} &\geq \|x+y-(x^2+y^2)^{1/2}\| + \|f(x)-y\| \\
&\geq \|x+f(x)-(x^2+f(x)^2)^{1/2}\| - \|f(x)-(x^2+f(x)^2)^{1/2}\| + \|y-(x^2+y^2)^{1/2}\| + \|f(x)-y\| \\
&= \tilde{\psi}_{FB}(x),
\end{aligned}$$

since the last equation follows from $y = f(x)$. Hence, ψ_{FB} is level-bounded from the assumption that $\tilde{\psi}_{FB}$ is level-bounded.

q.e.d

Theorem A.1 If f has the uniform Cartesian P-property and satisfies Condition A, then ψ_t is level-bounded.

Proof. From (3.4) of the body, we have

$$\begin{aligned}
\sqrt{2\psi_{FB}(x,y)} &= \left\| \begin{pmatrix} \phi_{FB}(x,y) \\ f(x)-y \end{pmatrix} \right\| \\
&\leq \left\| \begin{pmatrix} \phi_t(x,y) \\ f(x)-y \end{pmatrix} \right\| + \left\| \begin{pmatrix} \phi_{FB}(x,y)-\phi_t(x,y) \\ 0 \end{pmatrix} \right\| \\
&= \sqrt{2\psi_t(x,y)} + \sqrt{2}t.
\end{aligned}$$

Therefore, since ψ_{FB} is level-bounded, ψ_t is level-bounded.

q.e.d

B An estimation $\gamma(z)$ in Proposition 4.2

An estimation $\gamma(z)$ for $\Gamma(z) \leq \gamma(z)$ where $\Gamma(z) := \left\| \begin{matrix} L_x \\ L_y \end{matrix} \right\|$ with $z := (x, y)$.

$$\begin{aligned}
\Gamma(z) := \left\| \begin{matrix} L_x \\ L_y \end{matrix} \right\| &\leq \sqrt{\|L_x\|^2 + \|L_y\|^2} \\
&\leq \|L_x\| + \|L_y\| \\
&= \|x\| + \|y\|,
\end{aligned}$$

because $\sqrt{a^2+b^2} \leq a+b$ for $a,\ b \geq 0$.

Therefore we obtain $\gamma(z) = \|x\| + \|y\|$.