Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 11, No. 1, pp. 3–27, January 2015

THE JACOBIAN CONSISTENCY OF A SMOOTHED GENERALIZED FISCHER-BURMEISTER FUNCTION FOR THE SECOND-ORDER CONE COMPLEMENTARITY PROBLEM*

Xiaoni Chi, Zhongping Wan† and Zhibin Zhu

Abstract: Second-order cone (SOC) complementarity functions and their smoothing functions play an important role in the solution of the second-order cone complementarity problem. In this paper, we propose a smoothing function of the generalized Fischer-Burmeister function, which is also called the smoothed generalized Fischer-Burmeister function, and show its Jacobian consistency. Moreover, we estimate the distance between the subgradient of the generalized Fischer-Burmeister function and the gradient of its smoothing function. These reults will play an important role for achieving the rapid convergence of smoothing methods.

Key words: *second-order cone complementarity problem, generalized Fischer-Burmeister function, smoothing function, Jacobian consistency*

Mathematics Subject Classification: *90C33, 90C30*

1 Introduction

The second-order cone complementarity problem (SOCCP) [12] is to find $(x, y, \zeta) \in R^n \times R^n \times R^l$ such that

$$x \in K,\ y \in K,\ x^T y = 0,\ F(x, y, \zeta) = 0, \tag{1.1}$$

where $F : R^n \times R^n \times R^l \to R^n \times R^l$ is a continuously differentiable mapping with $l \geq 0$ and $n \geq 1$. Here $K \subset R^n$ is the Cartesian product of second-order cones (SOC), i.e., $K = K^{n_1} \times K^{n_2} \times \cdots \times K^{n_m}$ with $m, n_1, n_2, \cdots, n_m \geq 1$, $n = n_1 + n_2 + \cdots + n_m$ and the n_i-dimensional SOC defined by

$$K^{n_i} = \{x^i = (x_1^i; x_2^i) \in R \times R^{n_i-1} :\ x_1^i - \|x_2^i\| \geq 0\} \subset R^{n_i}.$$

The SOCCP has a variety of engineering and management applications, such as filter design, antenna array weight design, truss design, and grasping force optimization in robotics [15,16].

*This research is supported by the National Natural Science Foundation of China (Nos. 11401126, 71471140, 11361018), Guangxi Natural Science Foundation (No. 2014GXNSFFA118001), and Guangxi Key Laboratory of Automatic Detecting Technology and Instruments (No.YQ15112).

†Corresponding author.

Moreover, the SOCCP contains a wide class of problems, such as the nonlinear complementarity problems (NCP) [2], the second-order cone programming (SOCP) [1, 12, 23], and so on. For example, the SOCCP with $n_1 = n_2 = \cdots = n_m = 1$ and $F(x, y, \zeta) = f(x) - y$ is the NCP, and the KKT conditions for the SOCP (possibly including nonlinear functions) reduce to the SOCCP [13].

Without loss of generality we may assume that $m = 1$ and $K = K^n$ in the following analysis, since our analysis can be easily extended to the general case.

Recently great attention has been paid to smoothing methods, partially due to their superior theoretical and numerical performances [7, 10, 14, 19]. Based on the smoothing functions of the SOC complementarity functions, smoothing methods usually reformulate the SOCCP as a system of equations [3, 4, 22, 24]. The smoothing parameter involved in smoothing functions may be treated as a variable [21] or a parameter with an appropriate parameter control [13]. In the later case, the Jacobian consistency plays an important role for achieving the rapid convergence of Newton methods or Newton-like methods. Chen, Qi and Sun [6] proposed a smoothing Newton method for the general box constrained variational inequalities, and showed its global and superlinear convergence if the smoothing function f is semismooth at the solution and satisfies the Jacobian consistency. Hayashi, Yamashita and Fukushima [13] proposed a combined smoothing and regularized method for the monotone SOCCP, and showed its global and quadratic convergence based on the Jacobian consistency of the smoothed natural residual function. Ogasawara and Narushima [18] showed the Jacobian consistency of a smoothed Fischer-Burmeister (FB) function for the SOCCP. Based on the result in [18], Narushima, Ogasawara and Hayashi [17] constructed a smoothing Newton method with appropriate parameter control for the SOCCP, which is shown to possess the global and quadratic convergence with good numerical performances. Chi, Wan and Hao [8] studied the directional derivative and B-subdifferential of a one-parametric class of the SOC complementarity functions, proposed their smoothing functions, and showed the Jacobian consistency of the one-parametric class of smoothing functions for the SOCCP. Chen, Pan and Lin [5] presented a smoothing function of generalized FB function in the context of the NCP, and proposed a global and local superlinear (or quadratic) convergent smoothing algorithm for the mixed complementarity problem via the Jacobian consistency of the smoothed generalized FB function.

In [20], Pan, Kum, Lim and Chen were concerned with the generalized Fischer-Burmeister (FB) function $\phi_p : R^n \times R^n \to R^n$ defined by

$$\phi_p(x, y) := \sqrt[p]{|x|^p + |y|^p} - x - y, \tag{1.2}$$

where $p \in (1, 4)$ is an arbitrary but fixed parameter, and $|x|^p$ is the vector-valued function (see Section 2 for the definition). When $p = 2$, ϕ_p reduces to the FB function for the SOCCP given by

$$\phi_{\text{FB}}(x, y) := \sqrt{x^2 + y^2} - x - y,$$

where $x^2 = x \circ x$ being the Jordan product of x with itself, and $\sqrt{x}$ with $x \in K^n$ being the unique vector such that $\sqrt{x} \circ \sqrt{x} = x$. They [20] showed that the generalized FB function ϕ_p defined by (2) is an SOC complementarity function, the generalized FB merit function

$$\psi_p(x, y) := \frac{1}{2}\|\phi_p(x, y)\|^2$$

is smooth for $p \in (1, 4)$ and provided the condition for its coerciveness.

In this paper, we study the smoothing function of ϕ_p, i.e., the smoothed generalized FB function for the SOCCP. The motivations for us to study the smoothed generalized FB function are as follows. (i) The SOC complementarity function ϕ_p is typically nonsmooth, and therefore it can not be directly used to smoothing methods for solving the SOCCP. Thus it is necessary to present a smoothing function of the generalized FB function for the SOCCP. (ii) In the setting of the NCP, the smoothed generalized FB function is shown to share some favorable properties as the smoothed FB function holds [5] -for example, the Jacobian consistency. Thus, it is natural to ask whether a smoothed generalized FB function has the desirable properties of the smoothed FB function or not in the setting of the SOCCP. This work is the first step to resolve these questions, and makes it possible to design the smoothing methods for the SOCCP based on the generalized FB function ϕ_p.

The main contribution of this paper is to propose a smoothing function of the generalized FB function ϕ_p, and study its favorable properties. We show the Jacobian consistency of the smoothed generalized FB function, which will play an important role for achieving the rapid convergence of smoothing methods. Moreover, we estimate the distance between the subgradient of the generalized FB function and the gradient of its smoothing function, which will help to adjust a parameter appropriately in smoothing methods.

The organization of this paper is as follows. In Section 2, we review some preliminaries including the Euclidean Jordan algebra associated with the SOC, subdifferentials and Jacobian consistency. In Section 3, we propose a smoothing function of the generalized FB function for the SOCCP, and derive the computable formula for its Jacobian. In Section 4, we show the Jacobian consistency of the smoothed generalized FB function for the SOCCP. And in Section 5, we estimate the distance between the subgradient of the generalized FB function and the gradient of its smoothing function for the SOCCP. Finally, we close this paper with some conclusions in Section 6.

In what follows, we denote the nonnegative orthant of R by R_+. The symbol $\|\cdot\|$ means the Euclidean norm defined by $\|x\| := \sqrt{x^T x}$ for a vector x or the corresponding induced matrix norm. For simplicity, we often use $x = (x_1; x_2)$ for the column vector $x = (x_1, x_2^T)^T$. For the SOC K^n, intK^n and bdK^n mean the topological interior and the boundary of K^n, respectively. For a given set $S \subset R^{m\times n}$, conv$S$ stands for the convex hull of S in $R^{m\times n}$, and dist(X, S) denotes $\inf\{\|X - Y\| : Y \in S\}$ for a matrix $X \in R^{m\times n}$.

2 Preliminaries

In this section, we review some concepts and results, which include the Euclidean Jordan algebra [1,11] associated with the SOC K^n, subdifferentials [9] and Jacobian consistency [6].

First, we recall the Euclidean Jordan algebra associated with the SOC and some useful results. For any $x = (x_1; x_2), y = (y_1; y_2) \in R \times R^{n-1} (n \geq 1)$, the Euclidean Jordan algebra associated with the SOC K^n is defined by

$$x \circ y = (x^T y; x_1 y_2 + y_1 x_2),$$

with $e = (1, 0, \cdots, 0) \in R^n$ being its unit element. Given an element $x = (x_1; x_2) \in R \times R^{n-1}$, we define

$$L(x) = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix},$$

where I represents the $(n-1) \times (n-1)$ identity matrix. It is not difficult to see that $x \circ y = L(x)y$ for any $y \in R^n$. Moreover, $L(x)$ is symmetric positive definite (and hence invertible) if and only if $x \in \text{int}K^n$.

Now we give the spectral factorization of vectors in R^n associated with the SOC K^n. Each $x = (x_1; x_2) \in R \times R^{n-1}$ has a spectral factorization associated with K^n given by

$$x = \lambda_1(x)u_1(x) + \lambda_2(x)u_2(x),$$

where $\lambda_1(x)$, $\lambda_2(x)$ and $u_1(x)$, $u_2(x)$ are the spectral values, and the associated spectral vectors of x given by

$$\lambda_i(x) = x_1 + (-1)^i\|x_2\|,$$

$$u_i(x) = \begin{cases} \frac{1}{2}(1; (-1)^i \frac{x_2}{\|x_2\|}) & \text{if } x_2 \neq 0, \\ \frac{1}{2}(1; (-1)^i \omega) & \text{otherwise,} \end{cases}$$

for $i = 1, 2$, with any $\omega \in R^{n-1}$ such that $\|\omega\| = 1$. By the spectral factorization, a scalar function can be extended to a vector-valued function for the SOC. For any given $p \in (1, +\infty)$, we define the vector-valued functions

$$|x|^p = |\lambda_1(x)|^p u_1(x) + |\lambda_2(x)|^p u_2(x), \ \forall x \in R^n,$$

$$\sqrt[p]{x} = \sqrt[p]{\lambda_1(x)}u_1(x) + \sqrt[p]{\lambda_2(x)}u_2(x), \ \forall x \in K^n.$$

Definition 2.1 ([13])**.** For a nondifferentiable function $\varphi : R^m \to R^n$, we consider a function $\varphi_t : R^m \to R^n$ with a parameter $t > 0$ that has the following properties:
(i) φ_t is differentiable for any $t > 0$;
(ii) $\lim_{t \downarrow 0} \varphi_t(x) = \varphi(x)$ for any $x \in R^m$.
Such a function φ_t is called a smoothing function of φ.

Let $G : R^m \to R^n$ be a locally Lipschitzian function. Then G is differentiable almost everywhere by Rademacher's theorem [9]. The Bouligand (B-) subdifferential and the Clarke subdifferential of G at z are defined by

$$\partial G_B(z) := \{\lim_{\hat{z} \to z} \nabla G(\hat{z}) : \hat{z} \in D_G\} \quad \text{and} \quad \partial G(z) = \text{conv}\partial G_B(z)$$

respectively, where D_G denotes the set of points at which G is differentiable. It is obvious that $\partial G(z) = \{\nabla G(z)\}$ if G is continuously differentiable at z.

Based on the concepts of subdifferentials, we give the definition of Jacobian consistency, which was first introduced by Chen, Qi and Sun [6]. It is a concept relating the generalized Jacobian of a nonsmooth function with the Jacobian of a smoothing function [13].

Definition 2.2 ([6])**.** Let $G : R^m \to R^n$ be a locally Lipschitzian function. Let $G_t : R^m \to R^n$ be a continuously differentiable function for any $t > 0$ such that $\lim_{t \downarrow 0} G_t(z) = G(z)$ for any $z \in R^m$. We say that G_t satisfies the Jacobian consistency if for any $z \in R^m$, $\lim_{t \downarrow 0} \text{dist}(\nabla G_t(z), \partial G(z)) = 0$.

It should be noted that the "inf" appearing in the definition of $\text{dist}(\nabla G_t(z), \partial G(z))$ can be replaced by "min", since the set $\partial G(z)$ is compact at all $z \in R^m$ [9].

3 Smoothing Function

In this section, we propose a smoothing function of the generalized FB function for the SOCCP, and derive the computable formula for its Jacobian.

Since the generalized FB function ϕ_p given by (1.2) is nonsmooth, we consider the function $\phi_{p,t} : R^n \times R^n \to R^n$ defined by

$$\phi_{p,t}(x, y) := \sqrt[p]{|x|^p + |y|^p + |t|^p e} - x - y \tag{3.1}$$

with $p \in (1, 4)$ and the smoothing parameter $t \in R$.

In the following, we will show that the function $\phi_{p,t}$ given by (3.1) is a smoothing function of ϕ_p. Thus we can solve a family of smoothing subproblems $\phi_{p,t}(x, y) = 0$ for $t \neq 0$ and obtain a solution of $\phi_p(x, y) = 0$ by letting $|t| \downarrow 0$.

For convenience, we give some notations. For any $p \in (1, 4)$, let $q = (1 - p^{-1})^{-1}$. For any $x = (x_1; x_2), y = (y_1; y_2) \in R \times R^{n-1}$ and any $t \in R$, we define the mapping $w^t : R^{2n} \to R \times R^{n-1}$ by

$$w^t = (w_1^t; w_2^t) = w^t(x, y) := |x|^p + |y|^p + |t|^p e,$$

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

$$w = (w_1; w_2) = w(x, y) := |x|^p + |y|^p.$$

By the definitions of $|x|^p$ and $|y|^p$, we obtain

$$w_1^t = \frac{|\lambda_2(x)|^p + |\lambda_1(x)|^p}{2} + \frac{|\lambda_2(y)|^p + |\lambda_1(y)|^p}{2} + |t|^p = w_1 + |t|^p,$$

$$w_2^t = \frac{|\lambda_2(x)|^p - |\lambda_1(x)|^p}{2}\bar{x}_2 + \frac{|\lambda_2(y)|^p - |\lambda_1(y)|^p}{2}\bar{y}_2 = w_2,$$

where $\bar{x}_2 = \frac{x_2}{\|x_2\|}$ if $x_2 \neq 0$, and otherwise $\bar{x}_2$ is an arbitrary vector in R^{n-1} such that $\|\bar{x}_2\| = 1$, and $\bar{y}_2$ has the similar definition. Therefore $w^t = (w_1^t; w_2)$, and the spectral factorization of w^t is

$$w^t = \lambda_1(w^t)u_1(w) + \lambda_2(w^t)u_2(w),$$

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

$$\lambda_i(w^t) = w_1 + |t|^p + (-1)^i\|w_2\|, \tag{3.2}$$

$$u_i(w) = \frac{1}{2}\left(1; (-1)^i\bar{w}_2\right) \tag{3.3}$$

for $i = 1, 2$. Here $\bar{w}_2 := \frac{w_2}{\|w_2\|}$ if $w_2 \neq 0$, and otherwise $\bar{w}_2$ is any vector in R^{n-1} such that $\|\bar{w}_2\| = 1$.

For any $x = (x_1; x_2), y = (y_1; y_2) \in R \times R^{n-1}$ and any $t \in R$, we can also define

$$z^t = (z_1^t; z_2^t) = z^t(x, y) := \sqrt[p]{w^t} = \sqrt[p]{|x|^p + |y|^p + |t|^p e},$$

and for $t = 0$,

$$z = (z_1; z_2) = z(x, y) := \sqrt[p]{|x|^p + |y|^p}.$$

The spectral factorization of z^t and z are given by respectively,

$$z^t = \sqrt[p]{\lambda_1(w^t)}u_1(w) + \sqrt[p]{\lambda_2(w^t)}u_2(w),$$

$$z = \sqrt[p]{\lambda_1(w)}u_1(w) + \sqrt[p]{\lambda_2(w)}u_2(w),$$

and therefore

$$z_1^t = \frac{\sqrt[p]{\lambda_2(w^t)} + \sqrt[p]{\lambda_1(w^t)}}{2}, \quad z_2^t = \frac{\sqrt[p]{\lambda_2(w^t)} - \sqrt[p]{\lambda_1(w^t)}}{2}\bar{w}_2,$$

$$z_1 = \frac{\sqrt[p]{\lambda_2(w)} + \sqrt[p]{\lambda_1(w)}}{2}, \quad z_2 = \frac{\sqrt[p]{\lambda_2(w)} - \sqrt[p]{\lambda_1(w)}}{2}\bar{w}_2.$$

Thus we can partition R^{2n} as $R^{2n} = S_1 \cup S_2 \cup \{(0,0)\}$, where

$$\begin{aligned} S_1 &:= \{(x,y) \in R^{2n} : w \in \mathrm{int}K^n\} = \{(x,y) \in R^{2n} : \lambda_2(w) \ge \lambda_1(w) > 0\}, \\ S_2 &:= \{(x,y) \in R^{2n} : w \in \mathrm{bd}K^n \backslash \{0\}\} = \{(x,y) \in R^{2n} : 2w_1 = \lambda_2(w) > \lambda_1(w) = 0\}. \end{aligned}$$

Theorem 3.1. *For any $(x,y) \in R^n \times R^n$ and $t \neq 0$, let ϕ_p and $\phi_{p,t}$ with $p \in (1,4)$ be defined by (1.2) and (3.1) respectively. Then the following results hold.*

(i) *The function $\phi_{p,t}$ is continuously differentiable everywhere and its Jacobian is given by*

$$\nabla\phi_{p,t}(x,y) = \begin{bmatrix} \nabla g^{soc}(x)\nabla g^{soc}(z^t)^{-1} - I \\ \nabla g^{soc}(y)\nabla g^{soc}(z^t)^{-1} - I \end{bmatrix}, \tag{3.4}$$

where

$$\nabla g^{soc}(x) = \begin{cases} p\,sign(x_1)|x_1|^{p-1}I & \text{if } x_2 = 0, \\ \begin{bmatrix} b(x) & c(x)\bar{x}_2^T \\ c(x)\bar{x}_2 & a(x)I + (b(x) - a(x))\bar{x}_2\bar{x}_2^T \end{bmatrix} & \text{if } x_2 \neq 0, \end{cases} \tag{3.5}$$

with

$$\bar{x}_2 = \frac{x_2}{\|x_2\|}, \quad a(x) = \frac{|\lambda_2(x)|^p - |\lambda_1(x)|^p}{\lambda_2(x) - \lambda_1(x)}, \tag{3.6}$$

$$b(x) = \frac{p}{2}\left[sign(\lambda_2(x))|\lambda_2(x)|^{p-1} + sign(\lambda_1(x))|\lambda_1(x)|^{p-1}\right], \tag{3.7}$$

$$c(x) = \frac{p}{2}\left[sign(\lambda_2(x))|\lambda_2(x)|^{p-1} - sign(\lambda_1(x))|\lambda_1(x)|^{p-1}\right]; \tag{3.8}$$

and $\nabla g^{soc}(z^t)^{-1} = \left(p\sqrt[q]{w_1 + |t|^p}\right)^{-1} I$ if $w_2 = 0$, and otherwise

$$\begin{aligned} \nabla g^{soc}(z^t)^{-1} &= L_1(w^t) + L_2(w^t) + L_3(w^t) \\ &= \begin{bmatrix} e(w^t) & -f(w^t)\bar{w}_2^T \\ -f(w^t)\bar{w}_2 & d(w^t)I + (e(w^t) - d(w^t))\bar{w}_2\bar{w}_2^T \end{bmatrix} \end{aligned} \tag{3.9}$$

with

$$L_1(w^t) = \frac{1}{2p\sqrt[q]{\lambda_1(w^t)}}\begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix}, \quad L_2(w^t) = \frac{1}{2p\sqrt[q]{\lambda_2(w^t)}}\begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix}, \tag{3.10}$$

$$L_3(w^t) = d(w^t)\begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix}, \quad d(w^t) = \frac{\sqrt[p]{\lambda_2(w^t)} - \sqrt[p]{\lambda_1(w^t)}}{\lambda_2(w^t) - \lambda_1(w^t)}, \tag{3.11}$$

$$e(w^t) = \frac{1}{2p}\left(\frac{1}{\sqrt[q]{\lambda_1(w^t)}} + \frac{1}{\sqrt[q]{\lambda_2(w^t)}}\right), \quad f(w^t) = \frac{1}{2p}\left(\frac{1}{\sqrt[q]{\lambda_1(w^t)}} - \frac{1}{\sqrt[q]{\lambda_2(w^t)}}\right). \tag{3.12}$$

(ii) *For any* $(x, y) \in R^n \times R^n$, $\lim_{t\to 0} \phi_{p,t}(x, y) = \phi_p(x, y)$. *Thus,* $\phi_{p,t}$ *is a smoothing function of* ϕ_p.

Proof. (i) Since $w^t \in \text{int}K^n$ for any $(x, y) \in R^n \times R^n$ and any $t \neq 0$, it follows from the proof of Lemma 3.2 [20] and the chain rule for that relation (3.4) holds. Relations (3.9)-(3.12) are due to Proposition 5.2 and its proof in [12].

(ii) Fix any $x = (x_1; x_2), y = (y_1; y_2) \in R \times R^{n-1}$. For any $t > 0$, it follows from the spectral factorization of w and w^t that

$$\phi_p(x, y) = \sqrt[p]{\lambda_1(w)}u_1(w) + \sqrt[p]{\lambda_2(w)}u_2(w) - x - y,$$

$$\phi_{p,t}(x, y) = \sqrt[p]{\lambda_1(w^t)}u_1(w) + \sqrt[p]{\lambda_2(w^t)}u_2(w) - x - y,$$

where

$$\lambda_i(w) = w_1 + (-1)^i\|w_2\|,$$

and $\lambda_i(w^t)$ and $u_i(w)$ are respectively given by (3.2) and (3.3) for $i = 1, 2$. It is not difficult to see that

$$\lim_{t\to 0} \lambda_i(w^t) = \lim_{t\to 0}(\lambda_i(w) + |t|^p) = \lambda_i(w)$$

for $i = 1, 2$, and hence $\lim_{t\to 0} \phi_{p,t}(x, y) = \phi_p(x, y)$. Therefore, it follows from (i) and Definition 2.1 that $\phi_{p,t}$ is a smoothing function of ϕ_p. ☐

Remark It follows from $\lambda_2(w^t) > \lambda_1(w^t) > 0$ for any $w_2 \neq 0$ and $t \neq 0$, and $\lambda_i(z^t) = \sqrt[p]{\lambda_i(w^t)}(i = 1, 2)$ that

$$d(w^t) = \frac{\sqrt[p]{\lambda_2(w^t)} - \sqrt[p]{\lambda_1(w^t)}}{\lambda_2(w^t) - \lambda_1(w^t)} = \frac{\lambda_2(z^t) - \lambda_1(z^t)}{|\lambda_2(z^t)|^p - |\lambda_1(z^t)|^p} = \frac{1}{a(z^t)},$$

where the functions $a(\cdot)$ and $d(\cdot)$ are given by (3.6) and (3.11), respectively.

Next we give some properties of ϕ_p (see Lemma 3.1 in [20]), which will be used in the subsequent analysis.

Lemma 3.2. *For any* $x = (x_1; x_2), y = (y_1; y_2) \in R \times R^{n-1}$ *with* $w(x, y) \in bdK^n$, *we have the following equalities:*

$$w_1(x, y) = \|w_2(x, y)\| = 2^{p-1}(|x_1|^p + |y_1|^p) \tag{3.13}$$

$$x_1^2 = \|x_2\|^2,\ y_1^2 = \|y_2\|^2,\ x_1y_1 = x_2^Ty_2,\ x_1y_2 = y_1x_2. \tag{3.14}$$

If, in addition, $w_2(x, y) \neq 0$, *the following equalities hold with* $\bar{w}_2(x, y) = \frac{w_2(x,y)}{\|w_2(x,y)\|}$*:*

$$x_2^T\bar{w}_2(x, y) = x_1,\ x_1\bar{w}_2(x, y) = x_2,\ y_2^T\bar{w}_2(x, y) = y_1,\ y_1\bar{w}_2(x, y) = y_2. \tag{3.15}$$

4 Jacobian Consistency

In this section, we show the Jacobian consistency of the smoothing function $\phi_{p,t}$ with $p \in (1,4)$, which will play an important role for establishing the rapid convergence of smoothing methods.

It has been shown in Lemma 2.5 [20] that the function ϕ_p with any $p \in (1,4)$ satisfies

$$\phi_p(x,y) = 0 \Leftrightarrow x \in K^n,\ y \in K^n,\ \langle x, y\rangle = 0. \tag{4.1}$$

Let $(x, y, \zeta) \in R^n \times R^n \times R^l$ and define

$$\Phi_{p,t}(x,y,\zeta) := \begin{bmatrix} \phi_{p,t}(x,y) \\ F(x,y,\zeta) \end{bmatrix}, \tag{4.2}$$

$$\Phi_p(x,y,\zeta) := \begin{bmatrix} \phi_p(x,y) \\ F(x,y,\zeta) \end{bmatrix}. \tag{4.3}$$

It is easy to see that $\Phi_{p,t}(x,y,\zeta) = 0$ is the perturbation of the system of equations $\Phi_p(x,y,\zeta) = 0$. On account of (1.1), (4.1) and (4.3), we have

$$\Phi_p(x,y,\zeta) = 0 \Leftrightarrow (x,y,\zeta) \text{ solves (1.1)}.$$

Since $\Phi_p(x,y,\zeta)$ is typically nonsmooth, we can solve approximately the smooth system $\Phi_{p,t}(x,y,\zeta) = 0$ by using Newton's method at each iteration, and then obtain a solution of $\Phi_p(x,y,\zeta) = 0$ by reducing the parameter t to zero.

First, we show that the function $\Phi_{p,t}(x,y,\zeta)$ satisfies the Jacobian consistency.

Lemma 4.1. *For any $(x,y) \in R^n \times R^n$, we have*

$$J^0_{\phi_p}(x,y) := \lim_{t\to 0} \nabla\phi_{p,t}(x,y) = \begin{bmatrix} J_x - I \\ J_y - I \end{bmatrix} \tag{4.4}$$

with $p \in (1,4)$, where

$$J_x := \begin{cases} \nabla g^{soc}(x)\nabla g^{soc}(z)^{-1} & \text{if } (x,y) \in S_1, \\ \dfrac{sign(x_1)|x_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}} \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 2I - \bar{w}_2\bar{w}_2^T \end{bmatrix} & \text{if } (x,y) \in S_2, \\ O & \text{if } (x,y) = (0,0), \end{cases} \tag{4.5}$$

$$J_y := \begin{cases} \nabla g^{soc}(y)\nabla g^{soc}(z)^{-1} & \text{if } (x,y) \in S_1, \\ \dfrac{sign(y_1)|y_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}} \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 2I - \bar{w}_2\bar{w}_2^T \end{bmatrix} & \text{if } (x,y) \in S_2, \\ O & \text{if } (x,y) = (0,0). \end{cases} \tag{4.6}$$

Proof. By (3.4) and the symmetry of x and y, it suffices to show that

$$\lim_{t\to 0} \nabla g^{\rm soc}(x)\nabla g^{\rm soc}(z^t)^{-1} = J_x.$$

Case (1) If $(x,y) \in S_1$, we have from (3.2) that

$$\lim_{t\to 0} \lambda_i(w^t) = \lim_{t\to 0}[\lambda_i(w) + |t|^p] = \lambda_i(w) > 0$$

for $i = 1, 2$, and then

$$\begin{aligned}\lim_{t\to 0} z^t &= \lim_{t\to 0}\left[\sqrt[p]{\lambda_1(w^t)}u_1(w) + \sqrt[p]{\lambda_2(w^t)}u_2(w)\right]\\ &= \sqrt[p]{\lambda_1(w)}u_1(w) + \sqrt[p]{\lambda_2(w)}u_2(w)\\ &= z \in \text{int}K^n.\end{aligned}$$

Therefore,

$$\lim_{t\to 0}\nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z^t)^{-1} = \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z)^{-1} = J_x.$$

Case (2) If $(x, y) \in S_2$, we obtain $w \in \text{bd}K^n\backslash\{0\}$ and Lemma 3.2 holds, and thus

$$w_1 = \|w_2\| = 2^{p-1}(|x_1|^p + |y_1|^p),$$

$$2w_1 = \lambda_2(w) > \lambda_1(w) = 0,$$

which imply $w_2 \neq 0$. Then for any $t \neq 0$, we have from (3.2) that

$$\lambda_1(w^t) = \lambda_1(w) + |t|^p = |t|^p > 0, \tag{4.7}$$

$$\lambda_2(w^t) = \lambda_2(w) + |t|^p = 2^p(|x_1|^p + |y_1|^p) + |t|^p > 0. \tag{4.8}$$

If $x_2 = 0$, we have $x = 0$ from (3.14) and $\nabla g^{\text{soc}}(x) = O$. Therefore,

$$\begin{aligned}\lim_{t\to 0}\nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z^t)^{-1} &= \lim_{t\to 0} O\cdot\nabla g^{\text{soc}}(z^t)^{-1} = O\\ &= \frac{\text{sign}(x_1)|x_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 1 & \bar{w}_2^T\\ \bar{w}_2 & 2I - \bar{w}_2\bar{w}_2^T\end{bmatrix}\\ &= J_x.\end{aligned}$$

Hence, we consider $x_2 \neq 0$ in the following analysis of Case (2). For any $t \neq 0$, it follows from (3.9) that

$$\nabla g^{\text{soc}}(z^t)^{-1} = L_1(w^t) + L_2(w^t) + L_3(w^t).$$

We first show that $\nabla g^{\text{soc}}(x)L_1(w^t) = O$ for any $t \neq 0$. Since $w \in \text{bd}K^n\backslash\{0\}$ implies (3.14) holds, i.e., $|x_1| = \|x_2\|$, we have from (3.6), (3.7) and (3.8) that

$$a(x) = 2^{p-1}\text{sign}(x_1)|x_1|^{p-1},\quad b(x) = 2^{p-2}p\,\text{sign}(x_1)|x_1|^{p-1},\quad c(x) = 2^{p-2}p|x_1|^{p-1}. \tag{4.9}$$

Taking into account the fact that $w_2 \neq 0, x_2 \neq 0$ and $w \in \text{bd}K^n\backslash\{0\}$, we obtain from (3.15) that

$$\bar{x}_2^T\bar{w}_2 = \text{sign}(x_1),\quad \bar{w}_2 = \text{sign}(x_1)\bar{x}_2. \tag{4.10}$$

It follows from (4.9) and (4.10) that

$$\begin{aligned}&\nabla g^{\text{soc}}(x)L_1(w^t)\\ &= \frac{1}{2p\sqrt[q]{\lambda_1(w^t)}}\begin{bmatrix} b(x) & c(x)\bar{x}_2^T\\ c(x)\bar{x}_2 & a(x)I + (b(x)-a(x))\bar{x}_2\bar{x}_2^T\end{bmatrix}\begin{bmatrix} 1\\ -\bar{w}_2\end{bmatrix}\begin{bmatrix} 1 & -\bar{w}_2^T\end{bmatrix}\\ &= \frac{1}{2p|t|^{p-1}}\begin{bmatrix} b(x) - c(x)\bar{x}_2^T\bar{w}_2\\ c(x)\bar{x}_2 - a(x)\bar{w}_2 - (b(x)-a(x))\bar{x}_2\bar{x}_2^T\bar{w}_2\end{bmatrix}\begin{bmatrix} 1 & -\bar{w}_2^T\end{bmatrix}\\ &= \frac{1}{2p|t|^{p-1}}\begin{bmatrix} b(x) - \text{sign}(x_1)c(x)\\ (c(x) - \text{sign}(x_1)b(x))\bar{x}_2 - a(x)(\bar{w}_2 - \text{sign}(x_1)\bar{x}_2)\end{bmatrix}\begin{bmatrix} 1 & -\bar{w}_2^T\end{bmatrix}\\ &= \frac{b(x) - \text{sign}(x_1)c(x)}{2p|t|^{p-1}}\begin{bmatrix} 1 & -\bar{w}_2^T\\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix}\\ &= O.\end{aligned} \tag{4.11}$$

We next show

$$\lim_{t\to 0} \nabla g^{\rm soc}(x)[L_2(w^t)+L_3(w^t)] = J_x.$$

In fact, we obtain from (3.5), (3.10), (4.8), (4.9) and (4.10) that

$$\begin{aligned}
&\lim_{t\to 0} \nabla g^{\rm soc}(x)L_2(w^t)\\
&= \lim_{t\to 0}\frac{1}{2p\sqrt[q]{\lambda_2(w^t)}}\begin{bmatrix} b(x) & c(x)\bar{x}_2^T \\ c(x)\bar{x}_2 & a(x)I+(b(x)-a(x))\bar{x}_2\bar{x}_2^T\end{bmatrix}\begin{bmatrix}1\\ \bar{w}_2\end{bmatrix}\begin{bmatrix}1 & \bar{w}_2^T\end{bmatrix}\\
&= \frac{1}{2p\sqrt[q]{2^p(|x_1|^p+|y_1|^p)}}\begin{bmatrix} b(x)+\mathrm{sign}(x_1)c(x) \\ (c(x)+\mathrm{sign}(x_1)b(x))\bar{x}_2 + a(x)(\bar{w}_2-\mathrm{sign}(x_1)\bar{x}_2)\end{bmatrix}\begin{bmatrix}1 & \bar{w}_2^T\end{bmatrix}\\
&= \frac{b(x)+\mathrm{sign}(x_1)c(x)}{2p\sqrt[q]{2^p(|x_1|^p+|y_1|^p)}}\begin{bmatrix}1 & \bar{w}_2^T\\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix} \qquad (4.12)\\
&= \frac{\mathrm{sign}(x_1)|x_1|^{p-1}}{2\sqrt[q]{|x_1|^p+|y_1|^p}}\begin{bmatrix}1 & \bar{w}_2^T\\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix}.
\end{aligned}$$

Moreover, it follows from (3.5), (3.11), (4.7), (4.8), (4.9) and (4.10) that

$$\begin{aligned}
&\lim_{t\to 0} \nabla g^{\rm soc}(x)L_3(w^t)\\
&= \lim_{t\to 0} d(w^t)\begin{bmatrix} b(x) & c(x)\bar{x}_2^T \\ c(x)\bar{x}_2 & a(x)I+(b(x)-a(x))\bar{x}_2\bar{x}_2^T\end{bmatrix}\begin{bmatrix}0 & 0^T\\ 0 & I-\bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \lim_{t\to 0} d(w^t)\begin{bmatrix} 0 & c(x)(\bar{x}_2^T-\mathrm{sign}(x_1)\bar{w}_2^T) \\ 0 & a(x)(I-\bar{w}_2\bar{w}_2^T)+(b(x)-a(x))(\bar{x}_2\bar{x}_2^T-\mathrm{sign}(x_1)\bar{x}_2\bar{w}_2^T)\end{bmatrix}\\
&= \lim_{t\to 0} d(w^t)\begin{bmatrix} 0 & 0^T \\ 0 & a(x)(I-\bar{w}_2\bar{w}_2^T)\end{bmatrix} \qquad (4.13)\\
&= \lim_{t\to 0}\frac{\sqrt[p]{2^p(|x_1|^p+|y_1|^p)+|t|^p}-|t|}{2^p(|x_1|^p+|y_1|^p)}\begin{bmatrix} 0 & 0^T \\ 0 & 2^{p-1}\mathrm{sign}(x_1)|x_1|^{p-1}(I-\bar{w}_2\bar{w}_2^T)\end{bmatrix}\\
&= \frac{\mathrm{sign}(x_1)|x_1|^{p-1}}{\sqrt[q]{|x_1|^p+|y_1|^p}}\begin{bmatrix} 0 & 0^T \\ 0 & I-\bar{w}_2\bar{w}_2^T\end{bmatrix}.
\end{aligned}$$

Combining (4.11), (4.12) and (4.13) yields

$$\lim_{t\to 0} \nabla g^{\rm soc}(x)\nabla g^{\rm soc}(z^t)^{-1} = \lim_{t\to 0} \nabla g^{\rm soc}(x)[L_2(w^t)+L_3(w^t)] = J_x. \qquad (4.14)$$

Case (3) If $(x,y)=(0,0)$, then we obtain from Theorem 3.1 that $\nabla g^{\rm soc}(z^t)^{-1} = (p\sqrt[q]{|t|^p})^{-1}I$, $\nabla g^{\rm soc}(0)=O$, and therefore

$$\lim_{t\to 0} \nabla g^{\rm soc}(x)\nabla g^{\rm soc}(z^t)^{-1} = \lim_{t\to 0} O\cdot(p\sqrt[q]{|t|^p})^{-1}I = O = J_x.$$

This completes the proof. □

Lemma 4.2. *For any $(x,y)\in R^n\times R^n$, we have*

$$\begin{bmatrix} U_x - I \\ U_y - I\end{bmatrix} \in \partial_B \phi_p(x,y)$$

with $p\in(1,4)$, where

$$U_x = \pm Z + J_x, \quad U_y = J_y, \qquad (4.15)$$

with J_x and J_y defined by (4.5) and (4.6) respectively, and

$$Z = \begin{cases} O & \text{if } (x,y) \in S_1, \\ \frac{1}{2}\begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} & \text{if } (x,y) \in S_2, \\ I & \text{if } (x,y) = (0,0). \end{cases} \tag{4.16}$$

Proof. It follows from Lemma 3.2 [20] and the chain rule for differentiation that the generalized FB function ϕ_p for the SOC complementarity problem is continuously differentiable at any $(x, y) \in S_1$ with

$$\nabla\phi_p(x,y) = \begin{bmatrix} \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z)^{-1} - I \\ \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(z)^{-1} - I \end{bmatrix} \in \partial_B\phi_p(x,y).$$

Therefore, it suffices to consider the two cases: $(x, y) \in S_2$ and $(x, y) = (0, 0)$.

For any $(x, y) \in S_2$ or $(x, y) = (0, 0)$, let $(\hat{x}, y) = (x + \varepsilon e, y)$ with sufficiently small $\varepsilon \neq 0$, and define

$$\hat{w} = (\hat{w}_1; \hat{w}_2) := |\hat{x}|^p + |y|^p = |x + \varepsilon e|^p + |y|^p,$$

$$\hat{z} = (\hat{z}_1; \hat{z}_2) := \sqrt[p]{\hat{w}}, \quad \tilde{w}_2 := \frac{\hat{w}_2}{\|\hat{w}_2\|},$$

$$\lambda_i(\hat{w}) := \hat{w}_1 + (-1)^i\|\hat{w}_2\|, \quad i = 1, 2.$$

Direct calculations yield

$$\lambda_i(\hat{x}) = x_1 + \varepsilon + (-1)^i\|x_2\| = \lambda_i(x) + \varepsilon, \quad i = 1, 2,$$

$$\hat{w}_1 = \frac{|\lambda_2(\hat{x})|^p + |\lambda_1(\hat{x})|^p}{2} + \frac{|\lambda_2(y)|^p + |\lambda_1(y)|^p}{2}, \tag{4.17}$$

$$\hat{w}_2 = \frac{|\lambda_2(\hat{x})|^p - |\lambda_1(\hat{x})|^p}{2}\bar{x}_2 + \frac{|\lambda_2(y)|^p - |\lambda_1(y)|^p}{2}\bar{y}_2. \tag{4.18}$$

Therefore, as $\varepsilon \to 0$ we have $(\hat{x}, y) \to (x, y)$, $\hat{w} \to w$, $\hat{z} \to z$ and $\lambda_i(\hat{w}) \to \lambda_i(w)$ for $i = 1, 2$.

By the definition of B-subdifferential and (1.2), it suffices to show

$$\lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})\nabla g^{\text{soc}}(\hat{z})^{-1} = U_x, \tag{4.19}$$

$$\lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(\hat{z})^{-1} = U_y, \tag{4.20}$$

if ϕ_p is differentiable at $(\hat{x}, y)$.

Case (1) If $(x, y) = (0, 0)$, it is easy to see that $\hat{w} = |\varepsilon|^p e \in \text{int}K^n$, $\hat{z} = |\varepsilon|e$ and ϕ_p is differentiable at $(\hat{x}, y)$. Then we have

$$\begin{aligned} \lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})\nabla g^{\text{soc}}(\hat{z})^{-1} &= \lim_{\varepsilon\to 0} p\text{sign}(\varepsilon)|\varepsilon|^{p-1}I \cdot \frac{1}{p|\varepsilon|^{p-1}}I = \lim_{\varepsilon\to 0} \text{sign}(\varepsilon)I \\ &= \pm I = U_x, \end{aligned}$$

$$\begin{aligned} \lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(\hat{z})^{-1} &= \lim_{\varepsilon\to 0} O \cdot \frac{1}{p|\varepsilon|^{p-1}}I \\ &= O = U_y. \end{aligned}$$

Case (2) If $(x, y) \in S_2$, we obtain $w \in \text{bd}K^n\backslash\{0\}$. It follows from (3.14) and $w \in \text{bd}K^n$ that $|x| + |y| \in \text{bd}K^n$. Therefore, by Lemma 2.4 [20] we have that one of the following subcases must hold: (i) $x = 0, |y| \in \text{bd}K^n\backslash\{0\}$; (ii) $|x| \in \text{bd}K^n\backslash\{0\}, y = 0$; (iii) $|x| \in \text{bd}K^n\backslash\{0\}, |y| \in \text{bd}K^n\backslash\{0\}$.

Subcase (2.1) If $x = 0, |y| \in \text{bd}K^n\backslash\{0\}$, we have by (3.5), (4.17) and (4.18) that

$$\nabla g^{\text{soc}}(\hat{x}) = p\text{sign}(\varepsilon)|\varepsilon|^{p-1}I,$$

$$\hat{w}_1 = \frac{|\lambda_2(y)|^p + |\lambda_1(y)|^p}{2} + |\varepsilon|^p, \quad \hat{w}_2 = \frac{|\lambda_2(y)|^p - |\lambda_1(y)|^p}{2}\bar{y}_2.$$

We first consider the case $y_1 = \|y_2\| > 0$, and thus

$$\lambda_1(\hat{w}) = \hat{w}_1 - \|\hat{w}_2\| = |\varepsilon|^p > 0,$$

$$\lambda_2(\hat{w}) = \hat{w}_1 + \|\hat{w}_2\| = |2y_1|^p + |\varepsilon|^p > 0,$$

$$\tilde{w}_2 = \frac{\hat{w}_2}{\|\hat{w}_2\|} = \bar{y}_2 = \bar{w}_2.$$

Thus we obtain $\hat{w} \in \text{int}K^n$, which together with Lemma 3.2 [20] implies that ϕ_p is differentiable at $(\hat{x}, y)$.

Now we will show

$$\lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})\nabla g^{\text{soc}}(\hat{z})^{-1} = U_x.$$

By Theorem 3.1, we have $\nabla g^{\text{soc}}(\hat{z})^{-1} = L_1(\hat{w}) + L_2(\hat{w}) + L_3(\hat{w})$, where $L_1(\hat{w}), L_2(\hat{w})$ and $L_3(\hat{w})$ are given by (3.10) and (3.11) with $\hat{w}$ and $\tilde{w}_2$ replacing w^t and $\bar{w}_2$, respectively. Then we obtain

$$\begin{aligned} \lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})L_1(\hat{w}) &= \lim_{\varepsilon\to 0} \frac{p\text{sign}(\varepsilon)|\varepsilon|^{p-1}}{2p\sqrt[q]{|\varepsilon|^p}} \begin{bmatrix} 1 & -\tilde{w}_2^T \\ -\tilde{w}_2 & \tilde{w}_2\tilde{w}_2^T \end{bmatrix} \\ &= \pm\frac{1}{2} \begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix}, \end{aligned} \tag{4.21}$$

$$\begin{aligned} \lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})L_2(\hat{w}) &= \lim_{\varepsilon\to 0} \frac{p\text{sign}(\varepsilon)|\varepsilon|^{p-1}}{2p\sqrt[q]{|2y_1|^p + |\varepsilon|^p}} \begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} \\ &= O, \end{aligned} \tag{4.22}$$

$$\begin{aligned} \lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})L_3(\hat{w}) &= \lim_{\varepsilon\to 0} \frac{p\text{sign}(\varepsilon)|\varepsilon|^{p-1}\left(\sqrt[p]{|2y_1|^p + |\varepsilon|^p} - |\varepsilon|\right)}{|2y_1|^p} \begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \\ &= O. \end{aligned} \tag{4.23}$$

Combining (4.21), (4.22) and (4.23) yields

$$\lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(\hat{x})\nabla g^{\text{soc}}(\hat{z})^{-1} = \pm Z + J_x = U_x.$$

Next we will show

$$\lim_{\varepsilon\to 0} \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(\hat{z})^{-1} = U_y.$$

By following the proof of Case (2) in Lemma 4.1, we have

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(y)\nabla g^{\rm soc}(\hat z)^{-1}\\
&=\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(y)[L_1(\hat w)+L_2(\hat w)+L_3(\hat w)]\\
&=O+\frac{\mathrm{sign}(y_1)|y_1|^{p-1}}{2\sqrt[q]{|y_1|^p}}\begin{bmatrix}1 & \bar w_2^T\\ \bar w_2 & \bar w_2\bar w_2^T\end{bmatrix}+\frac{\mathrm{sign}(y_1)|y_1|^{p-1}}{\sqrt[q]{|y_1|^p}}\begin{bmatrix}0 & 0^T\\ 0 & I-\bar w_2\bar w_2^T\end{bmatrix}\\
&=\frac{\mathrm{sign}(y_1)|y_1|^{p-1}}{2\sqrt[q]{|x_1|^p+|y_1|^p}}\begin{bmatrix}1 & \bar w_2^T\\ \bar w_2 & 2I-\bar w_2\bar w_2^T\end{bmatrix}\\
&=J_y=U_y.
\end{aligned}$$

By using the same arguments, we can prove that (4.19) and (4.20) hold under the case $y_1=-\|y_2\|<0$.

Subcase (2.2) If $|x|\in \mathrm{bd}K^n\backslash\{0\}, y=0$, we have from (4.17) and (4.18) that

$$\hat w_1=\frac{|\lambda_2(\hat x)|^p+|\lambda_1(\hat x)|^p}{2},\quad \hat w_2=\frac{|\lambda_2(\hat x)|^p-|\lambda_1(\hat x)|^p}{2}\bar x_2.$$

We first consider the case $x_1=\|x_2\|>0$, and thus

$$\lambda_1(\hat w)=\hat w_1-\|\hat w_2\|=|\varepsilon|^p>0,$$

$$\lambda_2(\hat w)=\hat w_1+\|\hat w_2\|=|2x_1+\varepsilon|^p>0,$$

$$\tilde w_2=\frac{\hat w_2}{\|\hat w_2\|}=\bar x_2=\bar w_2.$$

Then we obtain $\hat w\in \mathrm{int}K^n$, which together with Lemma 3.2 [20] implies that ϕ_p is differentiable at $(\hat x, y)$. It follows from (3.5), (3.6), (3.7) and (3.8) that

$$\nabla g^{\rm soc}(\hat x)=\begin{bmatrix} b(\hat x) & c(\hat x)\bar x_2^T\\ c(\hat x)\bar x_2 & a(\hat x)I+(b(\hat x)-a(\hat x))\bar x_2\bar x_2^T\end{bmatrix},$$

where

$$a(\hat x)=\frac{|2x_1+\varepsilon|^p-|\varepsilon|^p}{2x_1},\quad d(\hat w)=\frac{|2x_1+\varepsilon|-|\varepsilon|}{|2x_1+\varepsilon|^p-|\varepsilon|^p},$$

$$b(\hat x)=\frac{p}{2}\left[\mathrm{sign}(2x_1+\varepsilon)|2x_1+\varepsilon|^{p-1}+\mathrm{sign}(\varepsilon)|\varepsilon|^{p-1}\right],$$

$$c(\hat x)=\frac{p}{2}\left[\mathrm{sign}(2x_1+\varepsilon)|2x_1+\varepsilon|^{p-1}-\mathrm{sign}(\varepsilon)|\varepsilon|^{p-1}\right].$$

Now we will show

$$\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(\hat x)\nabla g^{\rm soc}(\hat z)^{-1}=U_x.$$

By using (3.10), (3.11) and (4.10), and following the proof of Case (2) in Lemma 4.1, we obtain

$$\begin{aligned}
\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(\hat x)L_1(\hat w)&=\lim_{\varepsilon\to 0}\frac{b(\hat x)-\mathrm{sign}(x_1)c(\hat x)}{2p\sqrt[q]{\lambda_1(\hat w)}}\begin{bmatrix}1 & -\bar w_2^T\\ -\bar w_2 & \bar w_2\bar w_2^T\end{bmatrix}\\
&=\lim_{\varepsilon\to 0}\frac{p\,\mathrm{sign}(\varepsilon)|\varepsilon|^{p-1}}{2p\sqrt[q]{|\varepsilon|^p}}\begin{bmatrix}1 & -\bar w_2^T\\ -\bar w_2 & \bar w_2\bar w_2^T\end{bmatrix}\\
&=\pm\frac{1}{2}\begin{bmatrix}1 & -\bar w_2^T\\ -\bar w_2 & \bar w_2\bar w_2^T\end{bmatrix},
\end{aligned}\tag{4.24}$$

$$\begin{aligned}
\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(\hat{x})L_2(\hat{w}) &= \lim_{\varepsilon\to 0}\frac{b(\hat{x})+\mathrm{sign}(x_1)c(\hat{x})}{2p\sqrt[q]{\lambda_2(\hat{w})}}\begin{bmatrix}1 & \bar{w}_2^T\\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \lim_{\varepsilon\to 0}\frac{p\mathrm{sign}(2x_1+\varepsilon)|2x_1+\varepsilon|^{p-1}}{2p\sqrt[q]{|2x_1+\varepsilon|^p}}\begin{bmatrix}1 & \bar{w}_2^T\\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \frac{\mathrm{sign}(x_1)}{2}\begin{bmatrix}1 & \bar{w}_2^T\\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \frac{\mathrm{sign}(x_1)|x_1|^{p-1}}{2\sqrt[q]{|x_1|^p+|y_1|^p}}\begin{bmatrix}1 & \bar{w}_2^T\\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T\end{bmatrix},
\end{aligned}\tag{4.25}$$

$$\begin{aligned}
\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(\hat{x})L_3(\hat{w}) &= \lim_{\varepsilon\to 0}a(\hat{x})d(\hat{w})\begin{bmatrix}0 & 0^T\\ 0 & I-\bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \lim_{\varepsilon\to 0}\frac{|2x_1+\varepsilon|-|\varepsilon|}{2x_1}\begin{bmatrix}0 & 0^T\\ 0 & I-\bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \mathrm{sign}(x_1)\begin{bmatrix}0 & 0^T\\ 0 & I-\bar{w}_2\bar{w}_2^T\end{bmatrix}\\
&= \frac{\mathrm{sign}(x_1)|x_1|^{p-1}}{\sqrt[q]{|x_1|^p+|y_1|^p}}\begin{bmatrix}0 & 0^T\\ 0 & I-\bar{w}_2\bar{w}_2^T\end{bmatrix}.
\end{aligned}\tag{4.26}$$

Combining (4.24), (4.25) and (4.26) yields

$$\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(\hat{x})\nabla g^{\rm soc}(\hat{z})^{-1} = \pm Z + J_x = U_x.$$

Under this subcase, we have from (3.5)

$$\lim_{\varepsilon\to 0}\nabla g^{\rm soc}(y)\nabla g^{\rm soc}(\hat{z})^{-1} = \lim_{\varepsilon\to 0} O\cdot\nabla g^{\rm soc}(\hat{z})^{-1} = O = J_y = U_y.$$

By using the same arguments, we can prove that (4.19) and (4.20) hold under the case $x_1 = -\|x_2\| < 0$.

Subcase (2.3) If $|x| \in \mathrm{bd}K^n\backslash\{0\}, |y| \in \mathrm{bd}K^n\backslash\{0\}$, we first consider the case $x_1 = \|x_2\| > 0, y_1 = -\|y_2\| < 0$, and thus from Lemma 3.2

$$\bar{w}_2 = \mathrm{sign}(x_1)\bar{x}_2 = \bar{x}_2, \quad \bar{w}_2 = \mathrm{sign}(y_1)\bar{y}_2 = -\bar{y}_2.$$

Then we have by (3.5), (4.17) and (4.18) that

$$\hat{w}_1 = \frac{|2x_1+\varepsilon|^p+|\varepsilon|^p}{2} + \frac{|2y_1|^p}{2},$$

$$\hat{w}_2 = \frac{|2x_1+\varepsilon|^p-|\varepsilon|^p}{2}\bar{x}_2 - \frac{|2y_1|^p}{2}\bar{y}_2 = \frac{|2x_1+\varepsilon|^p-|\varepsilon|^p}{2}\bar{w}_2 + \frac{|2y_1|^p}{2}\bar{w}_2,$$

and thus

$$\lambda_1(\hat{w}) = \hat{w}_1 - \|\hat{w}_2\| = |\varepsilon|^p > 0,$$

$$\lambda_2(\hat{w}) = \hat{w}_1 + \|\hat{w}_2\| = |2x_1+\varepsilon|^p + |2y_1|^p > 0,$$

$$\tilde{w}_2 = \frac{\hat{w}_2}{\|\hat{w}_2\|} = \bar{x}_2 = \bar{w}_2.$$

Then we have $\hat{w} \in \text{int}K^n$, which together with Lemma 3.2 [20] implies that ϕ_p is differentiable at $(\hat{x}, y)$. It follows from (3.5), (3.6), (3.7) and (3.8) that

$$\nabla g^{\text{soc}}(\hat{x}) = \begin{bmatrix} b(\hat{x}) & c(\hat{x})\bar{x}_2^T \\ c(\hat{x})\bar{x}_2 & a(\hat{x})I + (b(\hat{x}) - a(\hat{x}))\bar{x}_2\bar{x}_2^T \end{bmatrix},$$

where

$$a(\hat{x}) = \frac{|2x_1 + \varepsilon|^p - |\varepsilon|^p}{2x_1}, \quad d(\hat{w}) = \frac{\sqrt[p]{|2x_1 + \varepsilon|^p + |2y_1|^p} - |\varepsilon|}{|2x_1 + \varepsilon|^p + |2y_1|^p - |\varepsilon|^p},$$

$$b(\hat{x}) = \frac{p}{2}\left[\text{sign}(2x_1 + \varepsilon)|2x_1 + \varepsilon|^{p-1} + \text{sign}(\varepsilon)|\varepsilon|^{p-1}\right],$$

$$c(\hat{x}) = \frac{p}{2}\left[\text{sign}(2x_1 + \varepsilon)|2x_1 + \varepsilon|^{p-1} - \text{sign}(\varepsilon)|\varepsilon|^{p-1}\right].$$

By using (3.10), (3.11) and (4.10), and following the proof of Case (2) in Lemma 4.1, we obtain

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \nabla g^{\text{soc}}(\hat{x})\nabla g^{\text{soc}}(\hat{z})^{-1} \\
&= \lim_{\varepsilon \to 0} \nabla g^{\text{soc}}(\hat{x})[L_1(\hat{w}) + L_2(\hat{w}) + L_3(\hat{w})] \\
&= \pm\frac{1}{2}\begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} + \frac{\text{sign}(x_1)|x_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&\quad + \frac{\text{sign}(x_1)|x_1|^{p-1}}{\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&= \pm\frac{1}{2}\begin{bmatrix} 1 & -\bar{w}_2^T \\ -\bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} + \frac{\text{sign}(x_1)|x_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 2I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&= \pm Z + J_x = U_x.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(\hat{z})^{-1} \\
&= \lim_{\varepsilon \to 0} \nabla g^{\text{soc}}(y)[L_1(\hat{w}) + L_2(\hat{w}) + L_3(\hat{w})] \\
&= O + \frac{\text{sign}(y_1)|y_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{bmatrix} + \frac{\text{sign}(y_1)|y_1|^{p-1}}{\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 0 & 0^T \\ 0 & I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&= \frac{\text{sign}(y_1)|y_1|^{p-1}}{2\sqrt[q]{|x_1|^p + |y_1|^p}}\begin{bmatrix} 1 & \bar{w}_2^T \\ \bar{w}_2 & 2I - \bar{w}_2\bar{w}_2^T \end{bmatrix} \\
&= J_y = U_y.
\end{aligned}$$

By using the same arguments, we can prove that (4.19) and (4.20) hold under the cases: $x_1 = \|x_2\| > 0, y_1 = \|y_2\| > 0$; $x_1 = -\|x_2\| < 0, y_1 = -\|y_2\| < 0$; and $x_1 = -\|x_2\| < 0, y_1 = \|y_2\| > 0$. This completes the proof. □

By Lemma 4.1 and Lemma 4.2, we obtain the following result.

Theorem 4.3. *The function $\Phi_{p,t}$ defined by (4.2) with $p \in (1, 4)$ and $t > 0$ satisfies the Jacobian consistency.*

Proof. By (4.2), it suffices to show the Jacobian consistency of the function $\phi_{p,t}$ with $p \in (1,4)$ and $t > 0$. Let

$$V^i := \begin{bmatrix} U_x^i - I \\ U_y - I \end{bmatrix},$$

where

$$U_x^i = (-1)^i Z + J_x, \quad U_y = J_y,$$

for $i = 1, 2$, and J_x, J_y and Z are given by (4.5), (4.6) and (4.16). Define

$$V := \frac{1}{2}(V^1 + V^2) = \begin{bmatrix} J_x - I \\ J_y - I \end{bmatrix}.$$

On the one hand, we obtain from (4.4) that $V = J^0_{\phi_p}(x,y) = \lim_{t \to 0} \nabla \phi_{p,t}(x,y)$. On the other hand, we have by Lemma 4.2 that $V^1, V^2 \in \partial_B \phi_p(x,y)$ and therefore $V = \frac{1}{2}(V^1 + V^2) \in \partial \phi_p(x,y)$. This together with Theorem 3.1 and Definition 2.2 implies the Jacobian consistency of $\phi_{p,t}$ with $p \in (1,4)$ and $t > 0$. □

5 Estimation of the Distance Between $\nabla\Phi_{p,t}$ and $\partial\Phi_p$

In this section, we estimate the distance between the subgradient of the generalized FB function Φ_p and the gradient of its smoothing function $\Phi_{p,t}$ in terms of t precisely. Therefore, we may control the parameter t so that $\mathrm{dist}(\nabla\Phi_{p,t}(x,y,\zeta), \partial\Phi_p(x,y,\zeta))$ becomes small appropriately, which will help to establish a rapid convergence in smoothing methods. We begin with two technical lemmas that will be used in the subsequent analysis.

Lemma 5.1 ([18])**.** *Let $\alpha, \beta, \gamma \in R$ and $v \in R^{n-1}$ with $\|v\| = 1 (n \geq 2)$. Let*

$$G = \begin{bmatrix} \beta & -\gamma v^T \\ -\gamma v & \alpha I + (\beta - \alpha) v v^T \end{bmatrix}.$$

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

Lemma 5.2. *Let $\lambda_2 > \lambda_1 > 0$ and $w_1 > 0$ be arbitrarily fixed. Then the functions $g_1, g_2, g_3 : [0, \infty) \to R$ defined by*

$$g_1(\tau) := \frac{1}{p\sqrt[q]{\lambda_1 + \tau}} - \frac{\sqrt[p]{\lambda_2 + \tau} - \sqrt[p]{\lambda_1 + \tau}}{\lambda_2 - \lambda_1}, \tag{5.1}$$

$$g_2(\tau) := \frac{\sqrt[p]{\lambda_2 + \tau} - \sqrt[p]{\lambda_1 + \tau}}{\lambda_2 - \lambda_1} - \frac{1}{p\sqrt[q]{\lambda_2 + \tau}}, \tag{5.2}$$

$$g_3(\tau) := \frac{\sqrt[p]{2w_1 + \tau} - \sqrt[p]{\tau}}{2w_1} - \frac{1}{p\sqrt[q]{2w_1 + \tau}}, \tag{5.3}$$

are all decreasing, and hence $g_1(0) \geq g_1(\tau), g_2(0) \geq g_2(\tau)$ and $g_3(0) \geq g_3(\tau)$ for all $\tau \geq 0$.

Proof. For any fixed $\lambda_2 > \lambda_1 > 0$ and $w_1 > 0$, we prove our assertion only for g_1 and g_2. The claim for g_3 can be proved similarly, since g_3 can be given by (5.2) with $2w_1$ and 0 replacing λ_2 and λ_1, respectively.

We first show that the function $g_1(\tau)$ is decreasing in $\tau \in [0, \infty)$. Direct calculations yield

$$\begin{aligned} g_1'(\tau) &= -\frac{1}{pq(\lambda_1+\tau)^{\frac{1}{q}+1}} - \frac{1}{p(\lambda_2-\lambda_1)}\left[\frac{1}{(\lambda_2+\tau)^{\frac{1}{q}}} - \frac{1}{(\lambda_1+\tau)^{\frac{1}{q}}}\right] \\ &= -\frac{\lambda_2+\tau}{pq(\lambda_2-\lambda_1)(\lambda_1+\tau)^{\frac{1}{q}+1}}\left\{\frac{\lambda_2-\lambda_1}{\lambda_2+\tau} + \frac{q(\lambda_1+\tau)}{\lambda_2+\tau}\left[\left(\frac{\lambda_1+\tau}{\lambda_2+\tau}\right)^{\frac{1}{q}} - 1\right]\right\} \\ &= -\frac{\lambda_2+\tau}{pq(\lambda_2-\lambda_1)(\lambda_1+\tau)^{\frac{1}{q}+1}} f_1\left(\frac{\lambda_1+\tau}{\lambda_2+\tau}\right), \end{aligned}$$

where the function $f_1 : [0, 1] \to R$ is defined by

$$f_1(s) := qs^{1+\frac{1}{q}} - (1+q)s + 1.$$

Since

$$f_1'(s) = (1+q)(s^{\frac{1}{q}} - 1) \le 0$$

for any $0 \le s \le 1$, the function f_1 is decreasing and $0 = f_1(1) \le f_1(s) \le f_1(0) = 1$ for any $s \in [0, 1]$. For any $\tau \in [0, \infty)$, we have $0 < s = \frac{\lambda_1+\tau}{\lambda_2+\tau} < 1$, and thus $0 \le f_1(\frac{\lambda_1+\tau}{\lambda_2+\tau}) \le 1$, and therefore $g_1'(\tau) \le 0$.

We next show that the function $g_2(\tau)$ is decreasing in $\tau \in [0, \infty)$. By direct calculations, we have

$$\begin{aligned} g_2'(\tau) &= \frac{1}{p(\lambda_2-\lambda_1)}\left[\frac{1}{(\lambda_2+\tau)^{\frac{1}{q}}} - \frac{1}{(\lambda_1+\tau)^{\frac{1}{q}}}\right] + \frac{1}{pq(\lambda_2+\tau)^{\frac{1}{q}+1}} \\ &= \frac{1}{pq(\lambda_2-\lambda_1)(\lambda_1+\tau)^{\frac{1}{q}}}\left\{q\left[\left(\frac{\lambda_1+\tau}{\lambda_2+\tau}\right)^{\frac{1}{q}} - 1\right] + \left(\frac{\lambda_1+\tau}{\lambda_2+\tau}\right)^{\frac{1}{q}}\frac{\lambda_2-\lambda_1}{\lambda_2+\tau}\right\} \\ &= \frac{1}{pq(\lambda_2-\lambda_1)(\lambda_1+\tau)^{\frac{1}{q}}} f_2\left(\frac{\lambda_1+\tau}{\lambda_2+\tau}\right), \end{aligned}$$

where the function $f_2 : [0, 1] \to R$ is defined by

$$f_2(s) := -s^{1+\frac{1}{q}} + (1+q)s^{\frac{1}{q}} - q.$$

Since

$$f_2'(s) = \left(1+\frac{1}{q}\right)(-s^{\frac{1}{q}} + s^{-\frac{1}{p}}) \ge 0$$

for any $0 \le s \le 1$, the function f_2 is increasing and $-q = f_2(0) \le f_2(s) \le f_2(1) = 0$ for any $s \in [0, 1]$. For any $\tau \in [0, \infty)$, we obtain $0 < s = \frac{\lambda_1+\tau}{\lambda_2+\tau} < 1$, and thus $-q \le f_2(\frac{\lambda_1+\tau}{\lambda_2+\tau}) \le 0$, and therefore $g_2'(\tau) \le 0$. This completes our proof. □

Theorem 5.3. *Let $p \in (1, 4)$ be given, and $(x, y, \zeta) \in R^{2n+l}$ be any point. Let the function $\Theta : R^{2n} \to R_+$ be defined by*

$$\Theta(x, y) := \|\nabla g^{soc}(x) \quad \nabla g^{soc}(y)\| = \sqrt{\|\nabla g^{soc}(x)^2 + \nabla g^{soc}(y)^2\|},$$

and for any $t \in R$, *the function* $h_{p,t} : R^{2n} \to R_+$ *be defined by*

$$h_{p,t}(x,y) := \begin{cases} e(w^t) + f(w^t) = \dfrac{1}{p\sqrt[q]{\lambda_1(w) + |t|^p}} & \text{if } (x,y) \in S_1, \\ d(w^t) = \dfrac{\sqrt[p]{2w_1 + |t|^p} - |t|}{2w_1} & \text{if } (x,y) \in S_2, \\ 0 & \text{if } (x,y) = (0,0). \end{cases} \tag{5.4}$$

Then for any nonzero $t \in R$,

$$dist(\nabla\Phi_{p,t}(x,y,\zeta), \partial\Phi_p(x,y,\zeta)) \le \Theta(x,y)[h_{p,0}(x,y) - h_{p,t}(x,y)]. \tag{5.5}$$

Proof. Since it follows from the proof of Theorem 4.3 that $J^0_{\phi_p}(x,y) = V \in \partial\phi_p(x,y)$ for any $(x,y,\zeta) \in R^{2n+l}$, we obtain

$$J^0_{\Phi_p}(x,y,\zeta) := \begin{pmatrix} J^0_{\phi_p}(x,y) & \nabla_{x,y}F(x,y,\zeta) \\ O & \nabla_\zeta F(x,y,\zeta) \end{pmatrix} \in \partial\Phi_p(x,y,\zeta).$$

Thus, we have from (3.4), (4.2) and (4.4) that

$$\begin{aligned} \text{dist}(\nabla\Phi_{p,t}(x,y,\zeta), \partial\Phi_p(x,y,\zeta)) &= \min\{\|\nabla\Phi_{p,t}(x,y,\zeta) - W\| : W \in \partial\Phi_p(x,y,\zeta)\} \\ &\le \|\nabla\Phi_{p,t}(x,y,\zeta) - J^0_{\Phi_p}(x,y,\zeta)\| \\ &= \|\nabla\phi_{p,t}(x,y,\zeta) - J^0_{\phi_p}(x,y,\zeta)\| \\ &= \left\| \begin{matrix} J_x - \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z^t)^{-1} \\ J_y - \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(z^t)^{-1} \end{matrix} \right\|. \end{aligned} \tag{5.6}$$

By (5.6), it suffices to estimate $\left\| \begin{matrix} J_x - \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z^t)^{-1} \\ J_y - \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(z^t)^{-1} \end{matrix} \right\|$. For simplicity, we write $\lambda_i^t := \lambda_i(w^t)$ and $\lambda_i := \lambda_i(w)$ for $i = 1,2$ in the following analysis.

Case (1) If $(x,y) \in S_1$, we have $\lambda_2 \ge \lambda_1 > 0$. Then it follows from (4.5) and (4.6) that

$$\begin{aligned} \left\| \begin{matrix} J_x - \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z^t)^{-1} \\ J_y - \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(z^t)^{-1} \end{matrix} \right\| &= \left\| \begin{matrix} \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z)^{-1} - \nabla g^{\text{soc}}(x)\nabla g^{\text{soc}}(z^t)^{-1} \\ \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(z)^{-1} - \nabla g^{\text{soc}}(y)\nabla g^{\text{soc}}(z^t)^{-1} \end{matrix} \right\| \\ &\le \left\| \begin{matrix} \nabla g^{\text{soc}}(x) \\ \nabla g^{\text{soc}}(y) \end{matrix} \right\| \|\nabla g^{\text{soc}}(z)^{-1} - \nabla g^{\text{soc}}(z^t)^{-1}\| \\ &= \Theta(x,y)\|G\|, \end{aligned}$$

where $G := \nabla g^{\text{soc}}(z)^{-1} - \nabla g^{\text{soc}}(z^t)^{-1}$. Now we will show $\|G\| = h_{p,0}(x,y) - h_{p,t}(x,y)$.

Subcase (1.1) If $(x,y) \in S_1$ and $w_2 = 0$, we have $\lambda_1 = \lambda_2 = w_1 > 0$. By Theorem 3.1, we have

$$\begin{aligned} \|G\| &= \|\nabla g^{\text{soc}}(z)^{-1} - \nabla g^{\text{soc}}(z^t)^{-1}\| = \left\| \frac{1}{p\sqrt[q]{w_1}} I - \frac{1}{p\sqrt[q]{w_1 + |t|^p}} I \right\| \\ &= \frac{1}{p\sqrt[q]{\lambda_1}} - \frac{1}{p\sqrt[q]{\lambda_1^t}} \\ &= h_{p,0}(x,y) - h_{p,t}(x,y). \end{aligned}$$

Subcase (1.2) If $(x,y) \in S_1$ and $w_2 \ne 0$, we have $\lambda_2 > \lambda_1 > 0$. Then by (3.9), the matrix G can be written as

$$\begin{aligned} G &= \nabla g^{\text{soc}}(z)^{-1} - \nabla g^{\text{soc}}(z^t)^{-1} \\ &= \begin{bmatrix} e_0 - e_t & -(f_0 - f_t)\bar{w}_2^T \\ -(f_0 - f_t)\bar{w}_2 & (d_0 - d_t)I + [(e_0 - e_t) - (d_0 - d_t)]\bar{w}_2\bar{w}_2^T \end{bmatrix} \\ &= \begin{bmatrix} \beta & -\gamma\bar{w}_2^T \\ -\gamma\bar{w}_2 & \alpha I + (\beta - \alpha)\bar{w}_2\bar{w}_2^T \end{bmatrix}, \end{aligned}$$

where $d_t := d(w^t), e_t := e(w^t), f_t := f(w^t)$ are given by (3.11) and (3.12), and $\alpha := d_0 - d_t, \beta := e_0 - e_t, \gamma := f_0 - f_t$. From Lemma 5.1, the eigenvalues of the symmetric matrix G are α and $\beta \pm \gamma$. Since $\lambda_2 > \lambda_1 > 0$ and $\lambda_i^t = \lambda_i + |t|^p$ $(i = 1, 2)$, it is not difficult to verify that $d_t = d_{|t|}, e_t + f_t = e_{|t|} + f_{|t|}$ and $e_t - f_t = e_{|t|} - f_{|t|}$ are all strictly decreasing functions in $|t| \geq 0$, i.e., for any $|t_1| > |t_2|$, we have $d_{t_1} < d_{t_2}, e_{t_1} + f_{t_1} < e_{t_2} + f_{t_2}$ and $e_{t_1} - f_{t_1} < e_{t_2} - f_{t_2}$. Therefore for any $t \neq 0$,

$$\alpha = d_0 - d_t = \frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1} - \frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2^t - \lambda_1^t} = \frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1} - \frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2 - \lambda_1} > 0, \tag{5.7}$$

$$\beta + \gamma = (e_0 - e_t) + (f_0 - f_t) = (e_0 + f_0) - (e_t + f_t) = \frac{1}{p}\left(\frac{1}{\sqrt[q]{\lambda_1}} - \frac{1}{\sqrt[q]{\lambda_1^t}}\right) > 0, \tag{5.8}$$

$$\beta - \gamma = (e_0 - e_t) - (f_0 - f_t) = (e_0 - f_0) - (e_t - f_t) = \frac{1}{p}\left(\frac{1}{\sqrt[q]{\lambda_2}} - \frac{1}{\sqrt[q]{\lambda_2^t}}\right) > 0. \tag{5.9}$$

Thus, for any $t \neq 0$, the matrix G is a positive definite matrix with $\|G\| = \max\{\alpha, \beta \pm \gamma\}$.

We will show below that $\beta + \gamma \geq \alpha \geq \beta - \gamma$ holds. For any $t \in R$, it follows from (5.1), (5.7), (5.8) and Lemma 5.2 that

$$\begin{aligned}(\beta + \gamma) - \alpha &= \frac{1}{p}\left(\frac{1}{\sqrt[q]{\lambda_1}} - \frac{1}{\sqrt[q]{\lambda_1^t}}\right) - \left(\frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1} - \frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2 - \lambda_1}\right) \\ &= \left(\frac{1}{p\sqrt[q]{\lambda_1}} - \frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1}\right) - \left(\frac{1}{p\sqrt[q]{\lambda_1^t}} - \frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2 - \lambda_1}\right) \\ &= g_1(0) - g_1(|t|^p) \geq 0.\end{aligned}$$

By (5.2), (5.7), (5.9) and Lemma 5.2, we obtain for any $t \in R$ that

$$\begin{aligned}\alpha - (\beta - \gamma) &= \left(\frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1} - \frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2 - \lambda_1}\right) - \frac{1}{p}\left(\frac{1}{\sqrt[q]{\lambda_2}} - \frac{1}{\sqrt[q]{\lambda_2^t}}\right) \\ &= \left(\frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1} - \frac{1}{p\sqrt[q]{\lambda_2}}\right) - \left(\frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2 - \lambda_1} - \frac{1}{p\sqrt[q]{\lambda_2^t}}\right) \\ &= g_2(0) - g_2(|t|^p) \geq 0.\end{aligned}$$

Therefore, we have proved $\beta + \gamma \geq \alpha \geq \beta - \gamma$, and hence for any $t \neq 0$,

$$\|G\| = \beta + \gamma = \frac{1}{p\sqrt[q]{\lambda_1}} - \frac{1}{p\sqrt[q]{\lambda_1^t}} = h_{p,0}(x, y) - h_{p,t}(x, y).$$

Case (2) If $(x, y) \in S_2$, we have $2w_1 = \lambda_2 > \lambda_1 = 0$. Then from (4.14), we obtain

$$J_x = \nabla g^{\mathrm{soc}}(x)[L_2(w) + L_3(w)]$$

with $L_2(w) = \lim_{t \to 0} L_2(w^t), L_3(w) = \lim_{t \to 0} L_3(w^t)$. Therefore it follows from (4.11) and (4.14) that

$$\left\| \begin{array}{l} J_x - \nabla g^{\rm soc}(x)\nabla g^{\rm soc}(z^t)^{-1} \\ J_y - \nabla g^{\rm soc}(y)\nabla g^{\rm soc}(z^t)^{-1} \end{array} \right\| = \left\| \begin{array}{l} \nabla g^{\rm soc}(x)[L_2(w)+L_3(w)] - \nabla g^{\rm soc}(x)[L_2(w^t)+L_3(w^t)] \\ \nabla g^{\rm soc}(y)[L_2(w)+L_3(w)] - \nabla g^{\rm soc}(y)[L_2(w^t)+L_3(w^t)] \end{array} \right\|$$

$$\begin{aligned} &\le \left\| \begin{array}{l} \nabla g^{\rm soc}(x) \\ \nabla g^{\rm soc}(y) \end{array} \right\| \|L_2(w)+L_3(w)-L_2(w^t)-L_3(w^t)\| \\ &= \Theta(x,y)\|G\|, \end{aligned}$$

where $G := L_2(w)+L_3(w)-L_2(w^t)-L_3(w^t)$. Now we will show again $\|G\| = h_{p,0}(x,y) - h_{p,t}(x,y)$.

From (3.10) and (3.11), the matrix G can be written as

$$\begin{aligned} G &= L_2(w)+L_3(w)-L_2(w^t)-L_3(w^t) \\ &= \frac{1}{2p\sqrt[q]{\lambda_2}}\left[\begin{array}{cc} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{array}\right] + d_0\left[\begin{array}{cc} 0 & 0^T \\ 0 & I-\bar{w}_2\bar{w}_2^T \end{array}\right] \\ &\quad -\frac{1}{2p\sqrt[q]{\lambda_2^t}}\left[\begin{array}{cc} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{array}\right] - d_t\left[\begin{array}{cc} 0 & 0^T \\ 0 & I-\bar{w}_2\bar{w}_2^T \end{array}\right] \\ &= \left(\frac{1}{2p\sqrt[q]{\lambda_2}} - \frac{1}{2p\sqrt[q]{\lambda_2^t}}\right)\left[\begin{array}{cc} 1 & \bar{w}_2^T \\ \bar{w}_2 & \bar{w}_2\bar{w}_2^T \end{array}\right] + (d_0-d_t)\left[\begin{array}{cc} 0 & 0^T \\ 0 & I-\bar{w}_2\bar{w}_2^T \end{array}\right] \\ &= \left[\begin{array}{cc} \delta & \delta\bar{w}_2^T \\ \delta\bar{w}_2 & \alpha I + (\delta-\alpha)\bar{w}_2\bar{w}_2^T \end{array}\right], \end{aligned}$$

with $\alpha := d_0 - d_t$ and $\delta := \frac{1}{2}[(e_0-f_0)-(e_t-f_t)]$. By Lemma 5.1, the eigenvalues of the symmetric matrix G are α, 0 and 2δ. Since $2w_1 = \lambda_2 > \lambda_1 = 0$ and $\lambda_i^t = \lambda_i + |t|^p$ $(i=1,2)$, it is not difficult to verify that $d_t = d_{|t|}$ and $e_t - f_t = e_{|t|} - f_{|t|}$ are all strictly decreasing functions in $|t| \ge 0$, i.e., for any $|t_1| > |t_2|$, we have $d_{t_1} < d_{t_2}$ and $e_{t_1} - f_{t_1} < e_{t_2} - f_{t_2}$. Then for any $t \ne 0$,

$$\begin{aligned} \alpha &= d_0 - d_t = \frac{\sqrt[p]{\lambda_2} - \sqrt[p]{\lambda_1}}{\lambda_2 - \lambda_1} - \frac{\sqrt[p]{\lambda_2^t} - \sqrt[p]{\lambda_1^t}}{\lambda_2^t - \lambda_1^t} \\ &= \frac{\sqrt[p]{2w_1}}{2w_1} - \frac{\sqrt[p]{2w_1+|t|^p} - |t|}{2w_1} > 0, \end{aligned}$$

$$\begin{aligned} 2\delta &= (e_0-f_0)-(e_t-f_t) = \left(\frac{1}{p\sqrt[q]{\lambda_2}} - \frac{1}{p\sqrt[q]{\lambda_2^t}}\right) \\ &= \frac{1}{p}\left(\frac{1}{\sqrt[q]{2w_1}} - \frac{1}{\sqrt[q]{2w_1+|t|^p}}\right) > 0. \end{aligned}$$

Therefore, it follows from (5.3) and Lemma 5.2 that for any $t \in R$,

$$\begin{aligned} \alpha - 2\delta &= \frac{\sqrt[p]{2w_1}}{2w_1} - \frac{\sqrt[p]{2w_1+|t|^p} - |t|}{2w_1} - \frac{1}{p}\left(\frac{1}{\sqrt[q]{2w_1}} - \frac{1}{\sqrt[q]{2w_1+|t|^p}}\right) \\ &= \frac{\sqrt[p]{2w_1}}{2w_1} - \frac{1}{p\sqrt[q]{2w_1}} - \left(\frac{\sqrt[p]{2w_1+|t|^p} - \sqrt[p]{|t|^p}}{2w_1} - \frac{1}{p\sqrt[q]{2w_1+|t|^p}}\right) \\ &= g_3(0) - g_3(|t|^p) \ge 0. \end{aligned}$$

Hence, for any $t \neq 0$, the matrix G is a positive definite matrix with

$$\begin{aligned}\|G\| &= \alpha = \frac{\sqrt[p]{2w_1}}{2w_1} - \frac{\sqrt[p]{2w_1 + |t|^p} - |t|}{2w_1} \\ &= h_{p,0}(x,y) - h_{p,t}(x,y).\end{aligned}$$

Case (3) If $(x,y) = (0,0)$, we obtain from (3.5), (4.5) and (4.6) that $J_x = J_y = O$, $\nabla g^{\rm soc}(x) = \nabla g^{\rm soc}(y) = O$, and therefore

$$\left\| \begin{array}{c} J_x - \nabla g^{\rm soc}(x)\nabla g^{\rm soc}(z^t)^{-1} \\ J_y - \nabla g^{\rm soc}(y)\nabla g^{\rm soc}(z^t)^{-1} \end{array} \right\| = 0.$$

Hence it is obvious that $\text{dist}(\nabla\Phi_{p,t}(x,y,\zeta), \partial\Phi_p(x,y,\zeta)) = 0$ and (5.5) holds. This completes the proof. □

Now we are in a position to estimate an upper bound of the parameter $t > 0$ for the predicted accuracy of the distance between the gradient of $\Phi_{p,t}$ and the subgradient of Φ_p.

Theorem 5.4. *Let $p \in (1,4)$, $\delta > 0$ be given, and $(x,y,\zeta) \in R^{2n+l}$ be any point. Let $\theta(x,y)$ be any function such that $\theta(x,y) \geq \Theta(x,y)$, and the function $\bar{t} : R^{2n} \times R_+ \to R_+ \cup \{+\infty\}$ be defined by*

$$\bar{t}(x,y,\delta) := \begin{cases} \sqrt[p]{\dfrac{\lambda_1(w)^2(p\delta)^q}{\theta(x,y)^q - \lambda_1(w)(p\delta)^q}} & \text{if } (x,y) \in S_1 \text{ and } \delta < \dfrac{\theta(x,y)}{p\sqrt[q]{\lambda_1(w)}}, \\ \dfrac{2w_1\delta}{\theta(x,y)} & \text{if } (x,y) \in S_2 \text{ and } \delta < \dfrac{\theta(x,y)}{\sqrt[q]{2w_1}}, \\ +\infty & \text{otherwise.} \end{cases}$$

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

$$dist(\nabla\Phi_{p,t}(x,y,\zeta), \partial\Phi_p(x,y,\zeta)) < \delta.$$

Moreover, one of the easily computable functions $\theta(x,y)$ that dominate $\Theta(x,y)$ is

$$\theta(x,y) := p\sqrt{2^{p-1}(\|x\|^{2p-2} + \|y\|^{2p-2})}.$$

Proof. **(i)** Without loss of generality, we may assume that $\theta(x,y) > 0$, since otherwise $\Theta(x,y) = 0$, and thus $\nabla g^{\rm soc}(x) = O, \nabla g^{\rm soc}(y) = O$, which together with (3.5) imply $(x,y) = (0,0) \notin S_1 \cup S_2$. Therefore

$$\text{dist}(\nabla\Phi_{p,t}(0,0,\zeta), \partial\Phi_p(0,0,\zeta)) = 0 < \delta$$

for any $t \in R$. Then by Theorem 5.3, it suffices to show that

$$\theta(x,y)[h_{p,0}(x,y) - h_{p,t}(x,y)] < \delta$$

for any $(x,y) \in S_1 \cup S_2$. We will show below that

$$h_{p,0}(x,y) - h_{p,t}(x,y) < \delta/\theta(x,y)$$

for any $t \in R$ such that $0 < |t| \leq \bar{t}(x,y,\delta)$. For simplicity, we write $\lambda_1 := \lambda_1(w), \theta := \theta(x,y)$ and $\bar{t} := \bar{t}(x,y,\delta)$ in the following analysis.

Case (1) If $(x,y) \in S_1$, we have by (5.4)

$$h_{p,0}(x,y) - h_{p,t}(x,y) < h_{p,0}(x,y) = \frac{1}{p\sqrt[q]{\lambda_1}} \le \frac{\delta}{\theta}$$

for any given $\delta \ge \frac{\theta}{p\sqrt[q]{\lambda_1}}$ and any $t \in R \cup \{+\infty\}$. Now let $\delta < \frac{\theta}{p\sqrt[q]{\lambda_1}}$ and $\bar{t} = \sqrt[p]{\frac{\lambda_1^2(p\delta)^q}{\theta^q - \lambda_1(p\delta)^q}}$. By (5.4), we obtain for any $t \in R$ such that $0 < |t| \le \bar{t}$,

$$\begin{aligned} h_{p,0}(x,y) - h_{p,t}(x,y) &= \frac{1}{p\sqrt[q]{\lambda_1}} - \frac{1}{p\sqrt[q]{\lambda_1 + |t|^p}} = \frac{\sqrt[q]{\lambda_1 + |t|^p} - \sqrt[q]{\lambda_1}}{p\sqrt[q]{\lambda_1}\sqrt[q]{\lambda_1 + |t|^p}} \\ &< \frac{\sqrt[q]{|t|^p}}{p\sqrt[q]{\lambda_1}\sqrt[q]{\lambda_1 + |t|^p}} = \frac{1}{p\sqrt[q]{\lambda_1}\sqrt[q]{1 + \lambda_1/|t|^p}} \\ &\le \frac{1}{p\sqrt[q]{\lambda_1}\sqrt[q]{1 + \lambda_1/|\bar{t}|^p}} = \frac{1}{p\sqrt[q]{\lambda_1}\sqrt[q]{1 + \lambda_1 \frac{\theta^q - \lambda_1(p\delta)^q}{\lambda_1^2(p\delta)^q}}} \\ &= \frac{\delta}{\theta}, \end{aligned}$$

where the first inequality is due to the fact that $\sqrt[q]{a} - \sqrt[q]{b} < \sqrt[q]{a-b}$ for any $a > b > 0$.

Case (2) If $(x,y) \in S_2$, we obtain from (5.4)

$$h_{p,0}(x,y) - h_{p,t}(x,y) < h_{p,0}(x,y) = \frac{\sqrt[p]{2w_1}}{2w_1} \le \frac{\delta}{\theta}$$

for any given $\delta \ge \theta/\sqrt[q]{2w_1}$ and any $t \in R \cup \{+\infty\}$. Now let $\delta < \theta/\sqrt[q]{2w_1}$ and $\bar{t} = 2w_1\delta/\theta$. From (5.4), we have

$$\begin{aligned} h_{p,0}(x,y) - h_{p,t}(x,y) &= \frac{\sqrt[p]{2w_1}}{2w_1} - \frac{\sqrt[p]{2w_1 + |t|^p} - |t|}{2w_1} \\ &< \frac{\sqrt[p]{2w_1}}{2w_1} - \frac{\sqrt[p]{2w_1} - |t|}{2w_1} = \frac{|t|}{2w_1} \\ &\le \frac{|\bar{t}|}{2w_1} = \frac{2w_1\delta}{2w_1\theta} \\ &= \frac{\delta}{\theta} \end{aligned}$$

for any $t \in R$ such that $0 < |t| \le \bar{t}$.

(ii) Since $\Theta(x,y) = \|\nabla g^{\rm soc}(x) \quad \nabla g^{\rm soc}(y)\|$ is not easy to compute in general, we estimate $\Theta(x,y)$ as

$$\begin{aligned} \Theta(x,y) &= \|\nabla g^{\rm soc}(x) \quad \nabla g^{\rm soc}(y)\| = \left\| [\nabla g^{\rm soc}(x) \quad \nabla g^{\rm soc}(y)] \begin{bmatrix} \nabla g^{\rm soc}(x) \\ \nabla g^{\rm soc}(y) \end{bmatrix} \right\|^{1/2} \\ &= \sqrt{\|\nabla g^{\rm soc}(x)^2 + \nabla g^{\rm soc}(y)^2\|} \\ &\le \sqrt{\|\nabla g^{\rm soc}(x)\|^2 + \|\nabla g^{\rm soc}(y)\|^2}. \end{aligned}$$

If $x_2 = 0$, we obtain from (3.5) that

$$\|\nabla g^{\rm soc}(x)\| = \|p{\rm sign}(x_1)|x_1|^{p-1}I\| = p|x_1|^{p-1} = p(|x_1| + \|x_2\|)^{p-1}.$$

If $x_2 \ne 0$, we obtain from (3.5), (3.6), (3.7), (3.8) and Lemma 5.1 that the eigenvalues of the matrix $\nabla g^{\rm soc}(x)$ are

$$\begin{aligned} b(x) + c(x) &= p{\rm sign}(\lambda_2(x))|\lambda_2(x)|^{p-1}, \\ b(x) - c(x) &= p{\rm sign}(\lambda_1(x))|\lambda_1(x)|^{p-1}, \end{aligned}$$

$$|a(x)| = p|\nu\lambda_2(x) + (1-\nu)\lambda_1(x)|^{p-1}$$

for some $\nu \in (0,1)$, where the last relation is due to (3.6) and the mean-value theorem. Then, the norm of the matrix $\nabla g^{\rm soc}(x)$ is given by

$$\begin{aligned}\|\nabla g^{\rm soc}(x)\| &= \max\{|b(x)+c(x)|, |b(x)-c(x)|, |a(x)|\} \\ &= p\max\{|\lambda_2(x)|^{p-1}, |\lambda_1(x)|^{p-1}\} \\ &= p(|x_1| + \|x_2\|)^{p-1},\end{aligned}$$

and thus

$$\begin{aligned}\Theta(x,y) &\le \sqrt{\|\nabla g^{\rm soc}(x)\|^2 + \|\nabla g^{\rm soc}(y)\|^2} \\ &= p\sqrt{(|x_1| + \|x_2\|)^{2p-2} + (|y_1| + \|y_2\|)^{2p-2}} \\ &\le p\sqrt{(2|x_1|^2 + 2\|x_2\|^2)^{p-1} + (2|y_1|^2 + 2\|y_2\|^2)^{p-1}} \\ &= p\sqrt{2^{p-1}(\|x\|^{2p-2} + \|y\|^{2p-2})}.\end{aligned}$$

Therefore, we can take $\theta(x,y) := p\sqrt{2^{p-1}(\|x\|^{2p-2} + \|y\|^{2p-2})}$ as one of the easily computable functions $\theta(x,y)$ that dominate $\Theta(x,y)$. □

6 Conclusions

In this paper, we show the Jacobian consistency of the smoothed generalized FB function $\phi_{p,t}$ with $p \in (1,4)$. Moreover, we estimate the distance between the subgradient of the generalized FB function ϕ_p and the gradient of its smoothing function $\phi_{p,t}$. These results will help to adjust a parameter appropriately, and play an important role for establishing the rapid convergence of the smoothing methods for the SOCCP via the smoothed generalized FB function.

Acknowledgements

The authors are grateful to the editor and the anonymous referees for their careful reading and valuable comments, which have greatly improved this paper.

References

[1] F. Alizadeh and D. Goldfarb, Second-order cone Programming, *Math. Program.* Ser. B 95 (2003) 3–51.

[2] H. Che, Y. Wang and M. Li, A smoothing inexact Newton method for P_0 nonlinear complementarity problem, *Front. Math. China* 7 (2012) 1043–1058.

[3] J.S. Chen, X. Chen and P. Tseng, Analysis of nonsmooth vector-valued functions associated with second-order cones, Math. Program. Ser. B, 101 (2004) 95–117.

[4] J.S. Chen and S. Pan, A survey on SOC complementarity functions and solution methods for SOCPs and SOCCPs, *Pacific. J. Optim.* 8 (2012) 33–74.

[5] J.S. Chen, S. Pan and T.C. Lin, A smoothing Newton method based on the generalized Fischer-Burmeister function for MCPs, *Nonlinear Anal.* 72 (2010) 3739–3758.

[6] X. Chen, L. Qi and D. Sun, Global and superlinear convergence of the smooothing Newton method and its application to general box constrained variational inequalities, *Math. Comput.* 67 (1998) 519–540.

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

[8] X. Chi, Z. Wan and Z. Hao, The Jacobian consistency of a one-parametric class of smoothing functions for SOCCP, Abstr. Appl. Anal. 2013 (2013), Article ID 965931.

[9] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.

[10] L. Fang and C. Han, A new one-step smoothing newton method for the second-order cone complementarity problem, *Math. Method Appl. Sci.* 34 (2011) 347–359.

[11] U. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, New York, Oxford Mathematical Monographs, 1994.

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

[13] S. Hayashi, N. Yamashita and M. Fukushima, A combined smoothing and regularized method for monontone second-order cone complementarity problems, *SIAM J. Optim.* 15 (2005) 593–615.

[14] Z. Huang and T. Ni, Smoothing algorithms for complementarity problems over symmetric cones, *Comput. Optim. Appl.* 45 (2010) 557–579.

[15] Y.J. Kuo and H.D. Mittelmann, Interior point methods for second-order cone programming and OR applications, *Comput. Optim. Appl.* 28 (2004) 255–285.

[16] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, Applications of second-order cone programming, *Linear Algebra Appl.* 284 (1998) 193–228.

[17] Y. Narushima, H. Ogasawara and S. Hayashi, A smoothing method with appropriate parameter control based on Fischer-Burmeister function for second-order cone complementarity problems, *Abstr. Appl. Anal.* 2013 (2013), Article ID 830698.

[18] H. Ogasawara and Y. Narushima, The Jacobian consistency of a smoothed Fischer-Burmeister function associated with second-order cones, *J. Math. Anal. Appl.* 394 (2012) 231–247.

[19] S. Pan and J.S. Chen, A damped Gauss-Newton method for the second-order cone complementarity problem, *Appl. Math. Optim.* 59 (2009) 293–318.

[20] S. Pan, S. Kum, Y. Lim and J.S. Chen, On the generalized Fischer-Burmeister merit function for the second-order cone complementarity problem, *Math. Comput.* 83 (2014) 1143–1171.

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

[22] D. Sun and J. Sun, Strong semismoothness of Fischer-Burmeister SDC and SOC complementarity functions, *Math. Program.* 103 (2005) 575–581.

[23] J. Tang, G. He and L. Fang, A new kernel function and its related properties for second-order cone optimization, *Pacific J. Optim.* 8 (2012) 321–346.

[24] L. Han, S. Bi and S. Pan, Nonsingularity of FB system and constraint nondegeneracy in semidefinite programming, *Numer. Algorithms* 62 (2013) 79–113.

Manuscript received 8 February 2014
revised 23 May 2014, 2 November 2014
accepted for publication 5 December 2014

Xiaoni Chi
School of Mathematics and Computing Science
Guilin University of Electronic Technology, Guilin 541004, China
Guangxi Key Laboratory of Automatic Detecting Technology and Instruments
(Guilin University of Electronic Technology), China
E-mail address: chixiaoni@126.com

Zhongping Wan
School of Mathematics and Statistics, Wuhan University
Wuhan 430072, China
E-mail address: mathwanzhp@whu.edu.cn

Zhibin Zhu
School of Mathematics and Computing Science
Guilin University of Electronic Technology
Guilin 541004, China
E-mail address: zhuzbma@hotmail.com