# 國立臺灣師範大學數學系碩士班碩士論文 

## 指導教授：陳 界 山 博士

A smoothing Newton method based on the generalized Fischer－Burmeister function for MCPs

研 究 生：林 子 靖

中 華民國九十八年六月

## 致 謝

在這份論文完成的同時，十分感謝所有陪我一路走過這兩年研究所生涯的許多人。首先，要感謝我的指導老師一陳界山教授，在專業的領域上，給予我適時地引領與啓發；在論交的撰寫上，給予我不斷地指導與幫助。

研究所尋晦與探索的過程中，最重要的莫過於有指導老師的從旁協助，引領著而不迷失於研究的瓶頸與窠臼，給予我許多寶貴的經驗傳承，在與指導老師不斷的討論中，點醒了我所忽略的細節與關鍵，同時爲我指引出未來明確的方向。

念研究所的這段日子中，無論是生活上，課業上，都受到老師許多的照顧與鼓勵，儘管人生中總是充滿著一關一關的挑戰與關卡，但老師總在許多面向給予我建議與支持，引領著我突破難關，不斷地向前邁進，逐步地實現自我的目標。在此，十分感謝老師這些年來的含辛茹苦與諄諄教誨，再次獻上最誠㨍的謝意。

此外，也要感謝口試委員一柯春旭教授與朱亮儒教授，感謝兩位教授在百忙之中抽空給予指導，有教授寶貴的指正與建議，才能使得本論文的內容更加精實與完備。另外，感謝同窗一淳格與晉宇在課業上的協助，有你們的陪伴與相互砥礪，研究努力的過程，不只有汗水與辛苦，還有共同攜手扶持的誠摯友誼與互相幫助的溫馨。最後，要感謝我的家人，給予我最強力的支持；並且是我最強力的後盾，無論遇到任何困難，總是對我深具信心，有爸媽在背後支持我，我才能無憂無慮的撰寫論文。

最後，由表地感謝所有協助過我的老師，同學，家人與朋友們，感謝你們的幫助，這份論文才能順利地完成，謝謝大家。

林子靖 敬致

## 目 次

1 Introduction. ..... 1
2 Preliminary ..... 3
3 The smoothing function and its properties ..... 6
4 Conclusions ..... 15
5 References ..... 15

# A smoothing Newton method based on the generalized Fischer-Burmeister function for MCPs 

June 5, 2009


#### Abstract

We present a smooth approximation for the generalized Fischer-Burmeister function where the 2-norm in the FB function is relaxed to a general $p$-norm $(p>1)$, and establish some favorable properties for it, for example, the Jacobian consistency. With the smoothing function, we transform the mixed complementarity problem (MCP) into solving a sequence of smooth system of equations.


Key Words. Mixed complementarity problem, the generalized FB function, smoothing approximation.

## 1 Introduction

The mixed complementarity problem (MCP) arises in many applications including the fields of economics, engineering, and operations research [11, 17, 18, 21] and has attracted much attention in last decade $[1,2,16,23,24,25]$. A collection of nonlinear mixed complementarity problems called MCPLIB can be found in [13] and two excellent books $[14,15]$ are good sources for seeking theoretical backgrounds and numerical methods.

Let $l_{i} \in \mathbb{R} \cup\{-\infty\}$ and $u_{i} \in \mathbb{R} \cup\{+\infty\}$ be given lower and upper bounds with $l_{i}<u_{i}$ for $i=1,2, \ldots, n$. Define $l=\left(l_{1}, l_{2}, \ldots, l_{n}\right)^{T}$ and $u=\left(u_{1}, u_{2}, \ldots, u_{n}\right)^{T}$. Given a mapping $F:[l, u] \rightarrow \mathbb{R}^{n}$ with $F=\left(F_{1}, F_{2}, \ldots, F_{n}\right)^{T}$. The MCP is to find a vector $x^{*} \in[l, u]$ such that each component $x_{i}^{*}$ satisfies exactly one of the following implications:

$$
\begin{align*}
x_{i}^{*}=l_{i} & \Longrightarrow F_{i}\left(x^{*}\right) \geq 0, \\
x_{i}^{*} \in\left(l_{i}, u_{i}\right) & \Longrightarrow F_{i}\left(x^{*}\right)=0,  \tag{1}\\
x_{i}^{*}=u_{i} & \Longrightarrow F_{i}\left(x^{*}\right) \leq 0 .
\end{align*}
$$

It is not hard to see that, when $l_{i}=-\infty$ and $u_{i}=+\infty$ for all $i=1,2, \ldots, n$, the MCP (1) is equivalent to solving the nonlinear system of equations

$$
\begin{equation*}
F(x)=0 \tag{2}
\end{equation*}
$$

whereas when $l_{i}=0$ and $u_{i}=+\infty$ for all $i=1,2, \ldots, n$, it reduces to the nonlinear complementarity problems (NCP) which is to find a point $x \in \mathbb{R}^{n}$ such that

$$
\begin{equation*}
x \geq 0, \quad F(x) \geq 0, \quad\langle x, F(x)\rangle=0 \tag{3}
\end{equation*}
$$

In fact, from Theorem 2 of [12], the MCP (1) is also equivalent to the famous variational inequality problem (VIP) which is to find a vector $x^{*} \in[l, u]$ such that

$$
\begin{equation*}
\left\langle F\left(x^{*}\right), x-x^{*}\right\rangle \geq 0 \quad \forall x \in[l, u] . \tag{4}
\end{equation*}
$$

In the rest of this paper, we assume the mapping $F$ to be continuously differentiable.

It is well-known that NCP functions play an important role in the design of algorithms for the MCP (1). With an NCP function $\phi$, the MCP (1) can be reformulated as a nonsmooth system of equations $\Phi(x)=0$, and consequently nonsmooth Newton methods or smoothing Newton methods can be applied for solving the system $\Phi(x)=0$. Among others, the latter is based on a smooth approximation of $\phi$. In the past two decades, many smooth approximation functions and Newton-type methods using smoothing NCP functions for complementarity problems have been developed [3, 4, 9, 10, 19, 20, 23]. Most of these methods focus on the Chen-Mangasarian class of smoothing functions of the minimum NCP function or the smoothing function of the Fischer-Burmeister NCP function.

Recently, an extension of the Fischer-Burmeister (FB) NCP function was considered in $[5,6,7]$ by two of the authors. Specifically, they define the generalized FB function by

$$
\begin{equation*}
\phi_{p}(a, b):=\|(a, b)\|_{p}-(a+b) \quad \forall a, b \in \mathbb{R}, \tag{5}
\end{equation*}
$$

where $p$ is an arbitrary fixed real number from the interval $(1,+\infty)$ and $\|(a, b)\|_{p}$ denotes the $p$-norm of $(a, b)$, i.e., $\|(a, b)\|_{p}=\sqrt[p]{|a|^{p}+|b|^{p}}$. In other words, in the function $\phi_{p}$, they replace the 2-norm of $(a, b)$ involved in the FB function by a more general $p$-norm. The function $\phi_{p}$ is still an NCP-function, that is, it satisfies the equivalence

$$
\begin{equation*}
\phi_{p}(a, b)=0 \quad \Longleftrightarrow \quad a \geq 0, \quad b \geq 0, \quad a b=0 \tag{6}
\end{equation*}
$$

Moreover, it turns out that $\phi_{p}$ possesses all favorable properties of the FB function; see $[5,6,7]$. For example, $\phi_{p}$ is strongly semismooth and its square is continuously differentiable everywhere on $\mathbb{R}^{2}$.

In this paper, we are concerned with the smoothing Newton method [10] based on the generalized FB function. In Section 2, we review some definitions and preliminary results to be used in the subsequent analysis. In Section 3, we present a smooth approximation function of the generalized FB function, and studied some favorable properties for it, including the Jacobian consistency property. In Section 4, we make concluding remarks.

Throughout this paper, $\mathbb{R}^{n}$ denotes the space of $n$-dimensional real column vectors and $e_{i}$ means a unit vector with $i$ th component being 1 and the others being 0 . For a differentiable mapping $F, F^{\prime}(x)$ and $\nabla F(x)$ denote the Jacobian of $F$ at $x$ and the transposed Jacobian of $F$, respectively. Given an index set $\mathcal{I}$, the notation $\left[F^{\prime}(x)\right]_{\mathcal{I I}}$ denotes
the submatrix consisting of the $i$ th row and the $j$ th column of $F^{\prime}(x)$ with $i \in \mathcal{I}$ and $j \in \mathcal{I}$.

## 2 Preliminary

In this section, we review some basic concepts and results that will be used in subsequent analysis. We start with introducing the concept of generalized Jacobian of a mapping. Let $G: \mathbb{R}^{n} \rightarrow \mathbb{R}^{m}$ be a locally Lipschitz continuous mapping. Then, $G$ is almost everywhere differentiable by Rademacher's Theorem (see [8]). In this case, the generalized Jacobian $\partial G(x)$ of $G$ at $x$ (in the Clarke sense) is defined as the convex hull of the B-subdifferential

$$
\partial_{B} G(x):=\left\{V \in \mathbb{R}^{m \times n} \mid \exists\left\{x^{k}\right\} \subseteq D_{G}:\left\{x^{k}\right\} \rightarrow x \text { and } G^{\prime}\left(x^{k}\right) \rightarrow V\right\}
$$

where $D_{G}$ is the set of differentiable points of $G$. In other words, $\partial G(x)=\operatorname{conv} \partial_{B} G(x)$. If $m=1$, we call $\partial G(x)$ the generalized gradient of $G$ at $x$. The calculation of $\partial G(x)$ is usually difficult in practice, so Qi proposed so-called $C$-subdifferential of $G$ :

$$
\begin{equation*}
\partial_{C} G(x)^{T}:=\partial G_{1}(x) \times \partial G_{2}(x) \times \cdots \times \partial G_{m}(x) \tag{7}
\end{equation*}
$$

which is easier to compute than the generalized Jacobian $\partial G(x)$. Here, the right-hand side of (7) denotes the set of matrices in $\mathbb{R}^{n \times m}$ whose $i$-th column is given by the generalized gradient of the $i$-th component function $G_{i}$. In fact, by Proposition 2.6.2 of [8],

$$
\begin{equation*}
\partial G(x)^{T} \subseteq \partial_{C} G(x)^{T} \tag{8}
\end{equation*}
$$

In addition, we also need the $P$-functions and $P$-matrices in the subsequent sections.
Definition 2.1 Let $F=\left(F_{1}, F_{2}, \ldots, F_{n}\right)^{T}$ with $F_{i}: \mathbb{R}^{n} \rightarrow \mathbb{R}$ for $i=1,2, \ldots, n$. Then,
(a) the mapping $F$ is monotone if

$$
\langle x-y, F(x)-F(y)\rangle \geq 0 \text { for all } x, y \in \mathbb{R}^{n}
$$

(b) the mapping $F$ is strictly monotone if

$$
\langle x-y, F(x)-F(y)\rangle>0 \text { for all } x, y \in \mathbb{R}^{n} \text { and } x \neq y
$$

(c) the mapping $F$ is strong monotone with modulus $\mu>0$ if

$$
\langle x-y, F(x)-F(y)\rangle \geq \mu\|x-y\|^{2} \text { for all } x, y \in \mathbb{R}^{n}
$$

(d) the mapping $F$ is called a $P_{0}$-function if for all $x, y \in \mathbb{R}^{n}$ and $x \neq y$, there is an index $i \in\{1,2, \ldots, n\}$ such that

$$
x_{i} \neq y_{i} \quad \text { and } \quad\left(x_{i}-y_{i}\right)\left(F_{i}(x)-F_{i}(y)\right) \geq 0
$$

(e) the mapping $F$ is called a $P$-function if for all $x, y \in \mathbb{R}^{n}$ and $x \neq y$, there is an index $i \in\{1,2, \ldots, n\}$ such that

$$
\left(x_{i}-y_{i}\right)\left(F_{i}(x)-F_{i}(y)\right)>0
$$

(f) the mapping $F$ is called a uniform $P$-function with modulus $\mu>0$ if there is an index $i \in\{1,2, \ldots, n\}$ such that

$$
\left(x_{i}-y_{i}\right)\left(F_{i}(x)-F_{i}(y)\right) \geq \mu\|x-y\|^{2} \quad \text { for all } x, y \in \mathbb{R}^{n} .
$$

From above definition, we know that

| $F$ is strong monotone | $\Rightarrow$ | $F$ is strictly monotone | $\Rightarrow$ | $F$ is monotone |
| :---: | :---: | :---: | :---: | :---: |
| $\Downarrow$ |  | $\Downarrow$ <br> $\Downarrow$ |  |  |
| F is uniform $P$-function | $\Rightarrow$ | F is $P$-function | $\Rightarrow \mathrm{F}$ is $P_{0}$-function |  |

Definition 2.2 $A$ matrix $M \in \mathbb{R}^{n \times n}$ is called an
(a) $P_{0}$-matrix if each of its principal minors is nonnegative.
(b) P-matrix if each of its principal minors is positive.

From above definition, we know that

$$
\mathrm{M} \text { is } P \text {-matrix } \Rightarrow \mathrm{M} \text { is } P_{0} \text {-matrix }
$$

From Definition 2.1 and 2.2, we see that a continuously differentiable mapping $F$ is a $P_{0}$-function if and only if $\nabla F(x)$ is $P_{0}$-matrix for all $x \in \mathbb{R}^{n}$. For the $P_{0}$-matrix, we also have the following important property.

Lemma 2.1 A matrix $M \in \mathbb{R}^{n \times n}$ is a $P_{0}$-matrix if and only if for every nonzero vector $x$, there exists an index $i$ such that $x_{i} \neq 0$ and $x_{i}(M x)_{i} \geq 0$.

Next we present some favorable properties of $\phi_{p}$ whose proofs can be found in $[5,6,7]$.
Lemma 2.2 Let $\phi_{p}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be defined by (5). Then, the following results hold.
(a) $\phi_{p}$ is a strongly semismooth NCP-function.
(b) Given any point $(a, b) \in \mathbb{R}^{2}$, each element in the generalized gradient $\partial \phi_{p}(a, b)$ has the representation $(\xi-1, \zeta-1)$ where, if $(a, b) \neq(0,0)$,

$$
(\xi, \zeta)=\left(\frac{\operatorname{sgn}(a) \cdot|a|^{p-1}}{\|(a, b)\|_{p}^{p-1}}, \frac{\operatorname{sgn}(b) \cdot|b|^{p-1}}{\|(a, b)\|_{p}^{p-1}}\right)
$$

and otherwise $(\xi, \zeta)$ is an arbitrary vector in $\mathbb{R}^{2}$ satisfying $|\xi|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1$.
(c) The square of $\phi_{p}$ is a continuously differentiable NCP function.
(d) If $\left\{\left(a^{k}, b^{k}\right)\right\} \subseteq \mathbb{R}^{2}$ satisfies $\left(a^{k} \rightarrow-\infty\right)$ or $\left(b^{k} \rightarrow-\infty\right)$ or $\left(a^{k} \rightarrow \infty\right.$ and $\left.b^{k} \rightarrow \infty\right)$, then we have $\left|\phi_{p}\left(a^{k}, b^{k}\right)\right| \rightarrow \infty$ as $k \rightarrow \infty$.

Lemma 2.3 Let $\phi_{p}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be defined by (5). Then, the following limits hold.
(a) $\lim _{l_{i} \rightarrow-\infty} \phi_{p}\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right)=-\phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)$.
(b) $\lim _{u_{i} \rightarrow \infty} \phi_{p}\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right)=\phi_{p}\left(x_{i}-l_{i}, F_{i}(x)\right)$.
(c) $\lim _{l_{i} \rightarrow-\infty} \lim _{u_{i} \rightarrow \infty} \phi_{p}\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right)=-F_{i}(x)$.

Proof. Let $\left\{a^{k}\right\} \subseteq \mathbb{R}$ be any sequence converging to $+\infty$ as $k \rightarrow \infty$ and $b \in \mathbb{R}$ be any fixed number. We will prove $\lim _{k \rightarrow \infty} \phi_{p}\left(a^{k}, b\right)=-b$, and part (a) then follows by continuity arguments. Without loss of generality, assume that $a^{k}>0$ for each $k$. Then,

$$
\begin{aligned}
\phi_{p}\left(a^{k}, b\right)= & \left(\left|a^{k}\right|^{p}+|b|^{p}\right)^{1 / p}-\left(a^{k}+b\right) \\
= & a^{k}\left(1+\left(|b| / a^{k}\right)^{p}\right)^{1 / p}-a^{k}-\bar{b} \\
= & a^{k}\left[1+\frac{1}{p}\left(\frac{|b|}{a^{k}}\right)^{p}+\frac{1-p}{2 p^{2}}\left(\frac{|b|}{a^{k}}\right)^{2 p}+\cdots+\right. \\
& \left.\frac{(1-p) \cdots(1-p n+p)}{n!p^{n}}\left(\frac{|b|}{a^{k}}\right)^{n p}+o\left(\left(\frac{|b|}{a^{k}}\right)^{p n}\right)\right]-a^{k}-b \\
= & \frac{1}{p} \frac{|b|^{p}}{\left(a^{k}\right)^{p-1}}+\frac{1-p}{2 p^{2}} \frac{|b|^{2 p}}{\left(a^{k}\right)^{2 p-1}}+\cdots+\frac{(1-p) \cdots(1-p n+p)}{n!p^{n}} \frac{|b|^{n p}}{\left(a^{k}\right)^{n p-1}} \\
& +\frac{\left(a^{k}\right)|b|^{n p}}{\left(a^{k}\right)^{n p}} \frac{o\left(\left(|b| / a^{k}\right)\right)^{p n}}{\left(|b| / a^{k}\right)^{p n}}-b
\end{aligned}
$$

where the third equality is using the Taylor expansion of the function $(1+t)^{1 / p}$ and the notation $o(t)$ means $\lim _{t \rightarrow 0} o(t) / t=0$. Since $a^{k} \rightarrow+\infty$ as $k \rightarrow \infty$, we have $\frac{|b|^{n p}}{\left(a^{k}\right)^{n p-1}} \rightarrow 0$ for all $n$. This together with the last equation implies $\lim _{k \rightarrow \infty} \phi_{p}\left(a^{k}, b\right)=-b$. This proves part (a). Part (b) and (c) are direct by part (a) and the continuity of $\phi_{\mathrm{FB}}$.

Lemma 2.4 [22, 1.3]Let $x \in \mathbb{R}^{n}$ and $1<p_{1}<p_{2}$. Then

$$
\|x\|_{p_{2}} \leq\|x\|_{p_{1}} \leq n^{\left(1 / p_{1}-1 / p_{2}\right)}\|x\|_{p_{2}} .
$$

## 3 The smoothing function and its properties

For convenience, in the rest of this paper, we adopt the following index sets:

$$
\begin{align*}
& I_{l}:=\left\{i \in\{1,2, \ldots, n\} \mid-\infty<l_{i}<u_{i}=+\infty\right\}, \\
& I_{u}:=\left\{i \in\{1,2, \ldots, n\} \mid-\infty=l_{i}<u_{i}<+\infty\right\}, \\
& I_{l u}:=\left\{i \in\{1,2, \ldots, n\} \mid-\infty<l_{i}<u_{i}<+\infty\right\},  \tag{9}\\
& I_{f}:=\left\{i \in\{1,2, \ldots, n\} \mid-\infty=l_{i}<u_{i}=+\infty\right\} .
\end{align*}
$$

With the generalized FB function, we define a operator $\Phi_{p}: \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$ componentwise as

$$
\Phi_{p, i}(x):= \begin{cases}\phi_{p}\left(x_{i}-l_{i}, F_{i}(x)\right) & \text { if } i \in I_{l},  \tag{10}\\ -\phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right) & \text { if } i \in I_{u} \\ \phi_{p}\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right) & \text { if } i \in I_{l u} \\ -F_{i}(x) & \text { if } i \in I_{f}\end{cases}
$$

where the minus sign for $i \in I_{u}$ and $i \in I_{f}$ is motivated by Lemma 2.3. In fact, all results of this paper would be true without the minus sign. Using the equivalence (6), it is not hard to verify that the following result holds.

Proposition $3.1 x^{*} \in \mathbb{R}^{n}$ is a solution of the $M C P$ (1) if and only if $x^{*}$ solves the nonlinear system of equations $\Phi_{p}(x)=0$.

Since $\phi_{p}$ is not differentiable at the origion, the system $\Phi_{p}(x)=0$ is nonsmooth. In this paper, we will find a solution of nonsmooth system $\Phi_{p}(x)=0$ by solving a sequence of smooth approximations $\Psi_{p}(x, \varepsilon)=0$, where $\varepsilon>0$ is a smoothing parameter and the operator $\Psi_{p}: \mathbb{R}^{n} \times(0, \infty) \rightarrow \mathbb{R}^{n}$ is defined componentwise as

$$
\Psi_{p, i}(x, \varepsilon):= \begin{cases}\psi_{p}\left(x_{i}-l_{i}, F_{i}(x), \varepsilon\right) & \text { if } i \in I_{l}  \tag{11}\\ -\psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right) & \text { if } i \in I_{u} \\ \psi_{p}\left(x_{i}-l_{i}, \psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right), \varepsilon\right) & \text { if } i \in I_{l u} \\ -F_{i}(x) & \text { if } i \in I_{f}\end{cases}
$$

with

$$
\begin{equation*}
\psi_{p}(a, b, \varepsilon):=\sqrt[p]{|a|^{p}+|b|^{p}+|\varepsilon|^{p}}-(a+b) \tag{12}
\end{equation*}
$$

In what follows, we concentrate on the favorable properties of the smoothing function $\psi_{p}$ and the operator $\Psi_{p}$. First, let us state the favorable properties of $\psi_{p}$.

Lemma 3.1 Let $\psi_{p}: \mathbb{R}^{3} \rightarrow \mathbb{R}$ be defined by (12). Then, the following result holds.
(a) For any fixed $\varepsilon>0, \psi_{p}(a, b, \varepsilon)$ is continuously differentiable for all $(a, b) \in \mathbb{R}^{2}$ with

$$
\begin{equation*}
-2<\frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial a}<0, \quad-2<\frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial b}<0 \tag{13}
\end{equation*}
$$

(b) For any fixed $(a, b) \in \mathbb{R}^{2}, \psi_{p}(a, b, \varepsilon)$ is continuously differentiable, strictly increasing and convex with respect to $\varepsilon>0$. Moreover, for any $0 \leq \varepsilon_{1} \leq \varepsilon_{2}$,

$$
\begin{equation*}
0 \leq \psi_{p}\left(a, b, \varepsilon_{2}\right)-\psi_{p}\left(a, b, \varepsilon_{1}\right) \leq \varepsilon_{2}-\varepsilon_{1} \tag{14}
\end{equation*}
$$

In particular, $\left|\psi_{p}(a, b, \varepsilon)-\phi_{p}(a, b)\right| \leq \varepsilon$ for all $\varepsilon \geq 0$.
(c) For any fixed $(a, b) \in \mathbb{R}^{2}$, let $\psi_{p}^{0}(a, b):=\lim _{\varepsilon \downarrow 0}\left(\frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial a}, \frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial b}\right)$. Then,

$$
\lim _{h=\left(h_{1}, h_{2}\right) \rightarrow(0,0)} \frac{\phi_{p}\left(a+h_{1}, b+h_{2}\right)-\phi_{p}(a, b)-\psi_{p}^{0}\left(a+h_{1}, b+h_{2}\right)^{T} h}{\|h\|}=0 .
$$

(d) For any given $\varepsilon>0$, if $\psi_{p}(a, b, \varepsilon)=0$, then $a>0, b>0$, $\min \{a, b\} \leq \frac{\varepsilon}{\sqrt[p]{2^{p}-2}}$.

In particular, if $\psi_{p}(a, b, \varepsilon)=0$, then $a>0, b>0, a b \leq \frac{\varepsilon^{2}}{2}$ when $p \geq 2$.
Proof. (a) Using an elementary calculation, we immediately obtain that

$$
\begin{align*}
& \frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial a}=\frac{\sqrt{\operatorname{sgn}}(a)|a|^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}+|\varepsilon|^{p}}\right)^{p-1}}-1 \\
& \frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial b}=\frac{\operatorname{sgn}(b)|b|^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}+|\varepsilon|^{p}}\right)^{p-1}}-1 \tag{15}
\end{align*}
$$

For any fixed $\varepsilon>0$, since $\frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial a}$ and $\frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial b}$ are continuous at every $(a, b) \in \mathbb{R}^{2}$, we have that $\psi_{p}(a, b, \varepsilon)$ is continuously differentiable for all $(a, b) \in \mathbb{R}^{2}$. Noting that

$$
\left|\frac{\operatorname{sgn}(a)|a|^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}+\varepsilon^{p}}\right)^{p-1}}\right|<1 \quad \text { and } \quad\left|\frac{\operatorname{sgn}(b)|b|^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}+\varepsilon^{p}}\right)^{p-1}}\right|<1
$$

we readily get the inequality (13).
(b) For any $\varepsilon>0$, from an elementary calculation, we have that

$$
\begin{gathered}
\frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial \varepsilon}=\frac{\varepsilon^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}+\varepsilon^{p}}\right)^{p-1}}>0 \\
\frac{\partial^{2} \psi_{p}(a, b, \varepsilon)}{\partial \varepsilon^{2}}=\frac{(p-1) \varepsilon^{p-2}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}+\varepsilon^{p}}\right)^{p-1}}\left(1-\frac{\varepsilon^{p}}{|a|^{p}+|b|^{p}+\varepsilon^{p}}\right) \geq 0 .
\end{gathered}
$$

Therefore, for any fixed $(a, b) \in \mathbb{R}^{2}, \psi_{p}(a, b, \varepsilon)$ is continuously differentiable, strictly increasing and convex with respect to $\varepsilon>0$. By the mean-value theorem, for any $0<\varepsilon_{1} \leq \varepsilon_{2}$, there exists some $\varepsilon_{0} \in\left(\varepsilon_{1}, \varepsilon_{2}\right)$ such that

$$
\psi_{p}\left(a, b, \varepsilon_{2}\right)-\psi_{p}\left(a, b, \varepsilon_{1}\right)=\frac{\partial \psi_{p}}{\partial \varepsilon}\left(a, b, \varepsilon_{0}\right)\left(\varepsilon_{2}-\varepsilon_{1}\right)
$$

Together with $\frac{\partial \psi_{p}}{\partial \varepsilon}\left(a, b, \varepsilon_{0}\right) \leq 1$, we have that (14) holds for all $0<\varepsilon_{1} \leq \varepsilon_{2}$. Letting $\varepsilon_{1} \downarrow 0$, the desired result then follows.
(c) Using the formula (15), it is easy to calculate that

$$
\begin{align*}
& \lim _{\varepsilon \downarrow 0} \frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial a}=\left\{\begin{array}{cl}
\frac{\operatorname{sgn}(a)|a|^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}}\right)^{p-1}}-1 & \text { if }(a, b) \neq(0,0), \\
-1 & \text { if }(a, b)=(0,0) ;
\end{array}\right. \\
& \lim _{\varepsilon \downarrow 0} \frac{\partial \psi_{p}(a, b, \varepsilon)}{\partial b}=\left\{\begin{array}{cc}
\frac{\operatorname{sgn}(b)|b|^{p-1}}{\left(\sqrt[p]{|a|^{p}+|b|^{p}}\right)^{p-1}}-1 & \text { if }(a, b) \neq(0,0), \\
-1 & \text { if }(a, b)=(0,0) .
\end{array}\right. \tag{16}
\end{align*}
$$

From this, we see that $\psi_{p}^{0}(a, b)=\left(\frac{\partial \phi_{p}(a, b)}{\partial a}, \frac{\partial \phi_{p}(a, b)}{\partial b}\right)$ at $(a, b) \neq(0,0)$. Therefore, we only need to check the case $(a, b)=(0,0)$. The desired result follows by

$$
\begin{aligned}
& \phi_{p}\left(h_{1}, h_{2}\right)-\phi_{p}(0,0)-\psi_{p}^{0}\left(h_{1}, h_{2}\right)^{T} h \\
= & \sqrt[p]{\left|h_{1}\right|^{p}+\left|h_{2}\right|^{p}}-\frac{\left|h_{1}\right|^{p}+\left|h_{2}\right|^{p}}{\left(\sqrt[p]{\left|h_{1}\right|^{p}+\left|h_{2}\right|^{p}}\right)^{p-1}} \\
= & \sqrt[p]{\left|h_{1}\right|^{p}+\left|h_{2}\right|^{p}}-\sqrt[p]{\left|h_{1}\right|^{p}+\left|h_{2}\right|^{p}} \\
= & 0 .
\end{aligned}
$$

(d) From the definition of $\psi_{p}(a, b, \varepsilon)$, clearly, $\psi_{p}(a, b, \varepsilon)=0$ implies $a+b \geq 0$, and hence $a \geq 0$ or $b \geq 0$. In addition, from the monotonicity of $p$-norm, if $a \geq 0, b \leq 0$ or $a \leq 0, b \geq 0$, we have

$$
\sqrt[p]{|a|^{p}+|b|^{p}+\varepsilon^{p}}>\sqrt[p]{|a|^{p}+|b|^{p}} \geq \max \{|a|,|b|\} \geq a+b
$$

which implies $\psi_{p}(a, b, \varepsilon)>0$. The two sides show that for any given $\varepsilon>0, \psi_{p}(a, b, \varepsilon)=0$ implies $a>0$ and $b>0$. Without loss of generality, we let $0<a \leq b$. For any fixed $a>0$, consider the function $f(t)=(t+a)^{p}-t^{p}-a^{p}-\varepsilon^{p}(t \geq 0)$. It is easy to verify that $f$ is strictly increasing on $[0,+\infty)$. Moreover, since $\psi_{p}(a, b, \varepsilon)=0$, we have $f(b)=0$. Hence $f(a)=\left(2^{p}-2\right) a^{p}-\varepsilon^{p} \leq 0$, we get that $a \leq \frac{\varepsilon}{\sqrt[p]{2^{p}-2}}$. Therefore, $\min \{a, b\} \leq \frac{\varepsilon}{\sqrt[p]{2^{p}-2}}$. Moreover, if $p \geq 2$, let $x=(a, b, \varepsilon) \in \mathbb{R}^{3}$, by lemma 2.4 we have $\|x\|_{p} \leq\|x\|_{2}$. Hence

$$
\begin{aligned}
& a+b=\sqrt[p]{|a|^{p}+|b|^{p}+\varepsilon^{p}} \leq \sqrt{|a|^{2}+|b|^{2}+\varepsilon^{2}} \\
\Rightarrow & (a+b)^{2} \leq a^{2}+b^{2}+\varepsilon^{2} \\
\Rightarrow & a b \leq \frac{\varepsilon^{2}}{2}
\end{aligned}
$$

The proof is thus complete.

Using Lemma 3.1 and the expression of $\Psi_{p}$, we readily obtain the following result.

Proposition 3.2 Let $\Psi_{p}: \mathbb{R}^{n} \times(0, \infty) \rightarrow \mathbb{R}^{n}$ be defined by (11). Then,
(a) for any fixed $\varepsilon>0, \Psi_{p}(x, \varepsilon)$ is continuously differentiable on $\mathbb{R}^{n}$ with

$$
\nabla_{x} \Psi_{p}(x, \varepsilon)=D_{a}(x, \varepsilon)+\nabla F(x) D_{b}(x, \varepsilon)
$$

where $D_{a}(x, \varepsilon)$ and $D_{b}(x, \varepsilon)$ are $n \times n$ diagonal matrices with the diagonal elements $\left(D_{a}\right)_{i i}(x, \varepsilon)$ and $\left(D_{b}\right)_{i i}(x, \varepsilon)$ defined as follows:
(a1) For $i \in I_{l}$,

$$
\begin{aligned}
\left(D_{a}\right)_{i i}(x, \varepsilon) & =\frac{\operatorname{sgn}\left(x_{i}-l_{i}\right)\left|x_{i}-l_{i}\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, F_{i}(x), \varepsilon\right)\right\|_{p}^{p-1}}-1 \\
\left(D_{b}\right)_{i i}(x, \varepsilon) & =\frac{\operatorname{sgn}\left(F_{i}(x)\right)\left|F_{i}(x)\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, F_{i}(x), \varepsilon\right)\right\|_{p}^{p-1}}-1
\end{aligned}
$$

(a2) For $i \in I_{u}$,

$$
\begin{aligned}
& \left(D_{a}\right)_{i i}(x, \varepsilon)=\frac{\operatorname{sgn}\left(u_{i}-x_{i}\right)\left|u_{i}-x_{i}\right|^{p-1}}{\left\|\left(u_{i}-x_{i}, F_{i}(x), \varepsilon\right)\right\|_{p}^{p-1}}-1 \\
& \left(D_{b}\right)_{i i}(x, \varepsilon)=\frac{-\operatorname{sgn}\left(F_{i}(x)\right)\left|F_{i}(x)\right|^{p-1}}{\left\|\left(u_{i}-x_{i}, F_{i}(x), \varepsilon\right)\right\|_{p}^{p-1}}-1 .
\end{aligned}
$$

(a3) For $i \in I_{l u}$,

$$
\left(D_{a}\right)_{i i}(x, \varepsilon)=a_{i}(x, \varepsilon)+b_{i}(x, \varepsilon) c_{i}(x, \varepsilon) \quad \text { and } \quad\left(D_{b}\right)_{i i}(x, \varepsilon)=b_{i}(x, \varepsilon) d_{i}(x, \varepsilon)
$$

with

$$
\begin{aligned}
a_{i}(x, \varepsilon) & =\frac{\operatorname{sgn}\left(x_{i}-l_{i}\right)\left|x_{i}-l_{i}\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, \psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right), \varepsilon\right)\right\|_{p}^{p-1}}-1, \\
b_{i}(x, \varepsilon) & =\frac{\operatorname{sgn}\left(\psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right)\right)\left|\psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right)\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, \psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right), \varepsilon\right)\right\|_{p}^{p-1}}-1, \\
c_{i}(x, \varepsilon) & =-\frac{\operatorname{sgn}\left(u_{i}-x_{i}\right)\left|u_{i}-x_{i}\right|^{p-1}}{\left\|\left(u_{i}-x_{i}, F_{i}(x), \varepsilon\right)\right\|_{p}^{p-1}}+1, \\
d_{i}(x, \varepsilon) & =\frac{\operatorname{sgn}\left(F_{i}(x)\right)\left|F_{i}(x)\right|^{p-1}}{\left\|\left(u_{i}-x_{i}, F_{i}(x), \varepsilon\right)\right\|_{p}^{p-1}}+1 .
\end{aligned}
$$

(a4) For $i \in I_{f},\left(D_{a}\right)_{i i}(x, \varepsilon)=0$ and $\left(D_{b}\right)_{i i}(x, \varepsilon)=-1$.
Moreover, $-2<\left(D_{a}\right)_{i i}(x, \varepsilon)<0$ and $-2<\left(D_{b}\right)_{i i}(x, \varepsilon)<0$ for all $i \in I_{l} \cup I_{u}$ and $-6<\left(D_{a}\right)_{i i}(x, \varepsilon)<0$ and $-4<\left(D_{b}\right)_{i i}(x, \varepsilon)<0$ for all $i \in I_{l u}$.
(b) For any given $\varepsilon_{1} \geq 0$ and $\varepsilon_{2} \geq 0$, we have

$$
\left\|\Psi_{p}\left(x, \varepsilon_{2}\right)-\Psi_{p}\left(x, \varepsilon_{1}\right)\right\| \leq \sqrt{n}(\sqrt[p]{2}+1)\left|\varepsilon_{2}-\varepsilon_{1}\right|, \quad \forall x \in \mathbb{R}^{n}
$$

Particularly, for any given $\varepsilon \geq 0$,

$$
\left\|\Psi_{p}(x, \varepsilon)-\Phi_{p}(x)\right\| \leq \sqrt{n}(\sqrt[p]{2}+1) \varepsilon, \quad \forall x \in \mathbb{R}^{n}
$$

To show that the smoothing operator $\Psi_{p}$ satisfies the Jacobian consistency property, we need the following characterization of the generalized Jacobian $\partial_{C} \Phi_{p}(x)$, which is direct by Lemma 2.2 (b).

Proposition 3.3 For any given $x \in \mathbb{R}^{n}$, we have $\partial_{C} \Phi_{p}(x)^{T}=\left\{D_{a}(x)+\nabla F(x) D_{b}(x)\right\}$, where $D_{a}(x), D_{b}(x)$ are $n \times n$ diagonal matrices whose diagonal elements are given below:
(a) For $i \in I_{l}$, if $\left(x_{i}-l_{i}, F_{i}(x)\right) \neq(0,0)$, then

$$
\begin{aligned}
\left(D_{a}\right)_{i i}(x) & =\frac{\operatorname{sgn}\left(x_{i}-l_{i}\right) \cdot\left|x_{i}-l_{i}\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, F_{i}(x)\right)\right\|_{p}^{p-1}}-1 \\
\left(D_{b}\right)_{i i}(x) & =\frac{\operatorname{sgn}\left(F_{i}(x)\right) \cdot\left|F_{i}(x)\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, F_{i}(x)\right)\right\|_{p}^{p-1}}-1
\end{aligned}
$$

and otherwise

$$
\left(\left(D_{a}\right)_{i i}(x),\left(D_{b}\right)_{i i}(x)\right) \in\left\{\left.(\xi-1, \zeta-1) \in \mathbb{R}^{2}| | \xi\right|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1\right\}
$$

(b) For $i \in I_{u}$, if $\left(u_{i}-x_{i},-F_{i}(x)\right) \neq(0,0)$, then

$$
\begin{aligned}
\left(D_{a}\right)_{i i}(x) & =\frac{\operatorname{sgn}\left(u_{i}-x_{i}\right) \cdot\left|u_{i}-x_{i}\right|^{p-1}}{\left\|\left(u_{i}-x_{i},-F_{i}(x)\right)\right\|_{p}^{p-1}}-1 \\
\left(D_{b}\right)_{i i}(x) & =-\frac{\operatorname{sgn}\left(F_{i}(x)\right) \cdot\left|F_{i}(x)\right|^{p-1}}{\left\|\left(u_{i}-x_{i},-F_{i}(x)\right)\right\|_{p}^{p-1}}-1
\end{aligned}
$$

and otherwise

$$
\left(\left(D_{a}\right)_{i i}(x),\left(D_{b}\right)_{i i}(x)\right) \in\left\{\left.(\xi-1, \zeta-1) \in \mathbb{R}^{2}| | \xi\right|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1\right\}
$$

(c) For $i \in I_{l u},\left(D_{a}\right)_{i i}(x)=a_{i}(x)+b_{i}(x) c_{i}(x)$ and $\left(D_{b}\right)_{i i}(x)=b_{i}(x) d_{i}(x)$, where if $\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right) \neq(0,0)$, then

$$
\begin{aligned}
a_{i}(x) & =\frac{\operatorname{sgn}\left(x_{i}-l_{i}\right) \cdot\left|x_{i}-l_{i}\right|^{p-1}}{\|\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right) \|_{p}^{p-1}\right.}-1 \\
b_{i}(x) & =\frac{\operatorname{sgn}\left(\phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right) \cdot\left|\phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right|^{p-1}}{\|\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right) \|_{p}^{p-1}\right.}-1
\end{aligned}
$$

and otherwise

$$
\left(a_{i}(x), b_{i}(x)\right) \in\left\{\left.(\xi-1, \zeta-1) \in \mathbb{R}^{2}| | \xi\right|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1\right\}
$$

and if $\left(u_{i}-x_{i},-F_{i}(x)\right) \neq(0,0)$, then

$$
\begin{aligned}
c_{i}(x) & =\frac{-\operatorname{sgn}\left(u_{i}-x_{i}\right) \cdot\left|u_{i}-x_{i}\right|^{p-1}}{\left\|\left(u_{i}-x_{i},-F_{i}(x)\right)\right\|_{p}^{p-1}}+1, \\
d_{i}(x) & =\frac{\operatorname{sgn}\left(F_{i}(x)\right) \cdot\left|F_{i}(x)\right|^{p-1}}{\left\|\left(u_{i}-x_{i},-F_{i}(x)\right)\right\|_{p}^{p-1}}+1,
\end{aligned}
$$

and otherwise

$$
\left(c_{i}(x), d_{i}(x)\right) \in\left\{\left.(\xi+1, \zeta+1) \in \mathbb{R}^{2}| | \xi\right|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1\right\} .
$$

(d) For $i \in I_{f},\left(D_{a}\right)_{i i}(x)=0$ and $\left(D_{b}\right)_{i i}(x)=-1$.

Proposition 3.4 Let $\Psi_{p}$ be defined by (11). Then, for any fixed $x \in \mathbb{R}^{n}$,

$$
\lim _{\varepsilon \downarrow 0} \operatorname{dist}\left(\nabla_{x} \Psi_{p}(x, \varepsilon)^{T}, \partial_{C} \Phi_{p}(x)\right)=0
$$

Proof. For the sake of notation, for any given $x \in \mathbb{R}^{n}$, we define the index sets:

$$
\begin{align*}
& \beta_{1}(x):=\left\{i \in I_{l} \mid\left(x_{i}-l_{i}, F_{i}(x)\right)=(0,0)\right\}, \\
& \bar{\beta}_{1}(x):=\left\{i \in I_{l} \mid\left(x_{i}-l_{i}, F_{i}(x)\right) \neq(0,0)\right\}, \\
& \beta_{2}(x):=\left\{i \in I_{u} \mid\left(u_{i}-x_{i}, F_{i}(x)\right)=(0,0)\right\}, \\
& \bar{\beta}_{2}(x):=\left\{i \in I_{u} \mid\left(u_{i}-x_{i}, F_{i}(x)\right) \neq(0,0)\right\},  \tag{17}\\
& \beta_{3}(x):=\left\{i \in I_{l u} \mid\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right)=(0,0)\right\}, \\
& \bar{\beta}_{3}(x):=\left\{i \in I_{l u} \mid\left(x_{i}-l_{i}, \phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)\right) \neq(0,0)\right\}, \\
& \beta_{4}(x):=\left\{i \in \bar{\beta}_{3}(x) \mid\left(u_{i}-x_{i}, F_{i}(x)\right)=(0,0)\right\}, \\
& \bar{\beta}_{4}(x):=\left\{i \in \bar{\beta}_{3}(x) \mid\left(u_{i}-x_{i}, F_{i}(x)\right) \neq(0,0)\right\} .
\end{align*}
$$

We proceed the arguments by the cases $i \in I_{l} \cup I_{u}, i \in I_{l u}$ and $i \in I_{f}$, respectively.

Case 1: $i \in I_{l} \cup I_{u}$. When $i \in \beta_{1}(x) \cup \beta_{2}(x)$, it is easy to see that

$$
\left(D_{a}\right)_{i i}(x, \varepsilon)=-1 \text { and }\left(D_{b}\right)_{i i}(x, \varepsilon)=-1
$$

From Proposition 3.2 (a1) and (a2), it then follows that

$$
\nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T}=-e_{i}^{T}-F_{i}^{\prime}(x) \text { for all } \varepsilon>0
$$

Since

$$
\begin{equation*}
(-1,-1) \in\left\{\left.(\xi-1, \zeta-1) \in \mathbb{R}^{2}| | \xi\right|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1\right\} \tag{18}
\end{equation*}
$$

by Proposition 3.3 (a) and (b) we get $\nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T} \in \partial_{C} \Phi_{p, i}(x)$. When $i \in \bar{\beta}_{1}(x) \cup \bar{\beta}_{2}(x)$,

$$
\lim _{\varepsilon \downarrow 0}\left(D_{a}\right)_{i i}(x, \varepsilon)=\left(D_{a}\right)_{i i}(x) \text { and } \lim _{\varepsilon \downarrow 0}\left(D_{b}\right)_{i i}(x, \varepsilon)=\left(D_{b}\right)_{i i}(x),
$$

which by Proposition 3.2 (a1) and (a2) implies

$$
\lim _{\varepsilon \downarrow 0} \nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T}=\left(D_{a}\right)_{i i}(x) e_{i}^{T}+\left(D_{b}\right)_{i i}(x) F_{i}^{\prime}(x) \in \partial_{C} \Phi_{p, i}(x)
$$

Since $I_{l} \cup I_{u}=\beta_{1}(x) \cup \beta_{2}(x) \cup \bar{\beta}_{1}(x) \cup \bar{\beta}_{2}(x)$, the last two subcases show that

$$
\begin{equation*}
\lim _{\varepsilon \downarrow 0} \nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T} \in \partial_{C} \Phi_{p, i}(x), \quad \forall i \in I_{l} \cup I_{u} \tag{19}
\end{equation*}
$$

Case 2: $i \in I_{l u}$. When $i \in \beta_{3}(x)$, clearly, $a_{i}(x, \varepsilon)=-1$. Notice that $\phi_{p}\left(u_{i}-x_{i},-F_{i}(x)\right)=$ 0 and $x_{i}-l_{i}=0$ imply $u_{i}-x_{i}>0$ and $F_{i}(x)=0$. Therefore,

$$
\begin{aligned}
& \lim _{\varepsilon \downarrow 0} b_{i}(x, \varepsilon) \\
= & \lim _{\varepsilon \downarrow 0} \frac{\operatorname{sgn}\left(\psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right)\right)\left|\psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right)\right|^{p-1}}{\left\|\left(x_{i}-l_{i}, \psi_{p}\left(u_{i}-x_{i},-F_{i}(x), \varepsilon\right), \varepsilon\right)\right\|_{p}^{p-1}}-1 \\
= & \lim _{\varepsilon \downarrow 0} \frac{\psi_{p}\left(u_{i}-x_{i}, 0, \varepsilon\right)^{p-1}}{\left\|\left(0, \psi_{p}\left(u_{i}-x_{i}, 0, \varepsilon\right), \varepsilon\right)\right\|_{p}^{p-1}}-1 \\
= & \lim _{\varepsilon \downarrow 0} \frac{1}{\left(\sqrt[p]{\left.1+\left(\frac{\varepsilon}{\psi_{p}\left(u_{i}-x_{i}, 0, \varepsilon\right)}\right)^{p}\right)^{p-1}}-1\right.} \\
= & -1
\end{aligned}
$$

where the last equality is by

$$
\begin{aligned}
& \lim _{\varepsilon \downarrow 0} \frac{\varepsilon}{\psi_{p}\left(u_{i}-x_{i}, 0, \varepsilon\right)} \\
= & \lim _{\varepsilon \downarrow 0} \frac{\left(\sqrt[p]{\left.\left(u_{i}-x_{i}\right)^{p}+\varepsilon^{p}\right)^{p-1}}\right.}{\varepsilon^{p-1}} \\
= & \infty
\end{aligned}
$$

which used L'Hospital's rule and

$$
\lim _{\varepsilon \downarrow 0} c_{i}(x, \varepsilon)=0, \quad d_{i}(x, \varepsilon)=1 \quad \text { and } \quad c_{i}(x)=0, \quad d_{i}(x)=1
$$

From Proposition 3.2 (a3) and Proposition 3.3 (c) and (18), it follows that

$$
\lim _{\varepsilon \downarrow 0} \nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T}=-e_{i}^{T}-F_{i}^{\prime}(x) \in \partial_{C} \Phi_{p, i}(x), \quad i \in \beta_{3}(x) .
$$

When $i \in \bar{\beta}_{3}(x)$, we have $\lim _{\varepsilon \downarrow 0} a_{i}(x, \varepsilon)=a_{i}(x)$ and $\lim _{\varepsilon \downarrow 0} b_{i}(x, \varepsilon)=b_{i}(x)$. Also,

$$
c_{i}(x, \varepsilon)=1, \quad d_{i}(x, \varepsilon)=1 \quad \text { for } \quad i \in \beta_{4}(x)
$$

and

$$
\lim _{\varepsilon \downarrow 0} c_{i}(x, \varepsilon)=c_{i}(x), \quad \lim _{\varepsilon \downarrow 0} d_{i}(x, \varepsilon)=d_{i}(x) \quad \text { for } \quad i \in \bar{\beta}_{4}(x) \text {. }
$$

Using Proposition 3.3 (c) and noting that

$$
(1,1) \in\left\{\left.(\xi+1, \zeta+1) \in \mathbb{R}^{2}| | \xi\right|^{\frac{p}{p-1}}+|\zeta|^{\frac{p}{p-1}} \leq 1\right\}
$$

we get $\lim _{\varepsilon \downarrow 0} \nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T} \in \partial_{C} \Phi_{p, i}(x)$ for $i \in \bar{\beta}_{3}(x)$. Along with the above discussions,

$$
\begin{equation*}
\lim _{\varepsilon \downarrow 0} \nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T} \in \partial_{C} \Phi_{p, i}(x) \text { for } i \in I_{l u} . \tag{20}
\end{equation*}
$$

Case 3: $i \in I_{f}$. By Proposition 3.2 (a4) and Proposition 3.3 (d), it is obvious that

$$
\begin{equation*}
\lim _{\varepsilon \downarrow 0} \nabla_{x} \Psi_{p, i}(x, \varepsilon)^{T} \in \partial_{C} \Phi_{p, i}(x) \quad \text { for } i \in I_{f} \tag{21}
\end{equation*}
$$

Now the desired result follows from (19)-(21) and $\{1,2, \ldots, n\}=I_{f} \cup I_{l} \cup I_{u} \cup I_{l u}$.

In order to use Newton method, we need the Jacobian matrix of $\Psi_{p}$ is nonsingular.
Proposition 3.5 For any fixed $\varepsilon>0$, the Jacobian matrix of $\Psi_{p}$ at any $x \in \mathbb{R}^{n}$ is nonsingular if $F$ is a $P_{0}$-function and the submatrix $\left[F^{\prime}(x)\right]_{I_{f} I_{f}}$ is nonsingular. Particularly, if $I_{f}=\emptyset$, the Jacobian matrix of $\Psi_{p}$ at any $x \in \mathbb{R}^{n}$ is nonsingular if and only if $F$ is a $P_{0}$-function.

Proof. For any given $\varepsilon>0$, the Jacobian matrix of $\Psi_{p}$ at any $x \in \mathbb{R}^{n}$ is

$$
\nabla_{x} \Psi_{p}(x, \varepsilon)^{T}=D_{a}(x, \varepsilon)+D_{b}(x, \varepsilon) F^{\prime}(x)
$$

where $D_{a}(x, \varepsilon)$ and $D_{b}(x, \varepsilon)$ are $n \times n$ diagonal matrices whose diagonal elements $\left(D_{a}\right)_{i i}(x, \varepsilon)$ and $\left(D_{b}\right)_{i i}(x, \varepsilon)$ are negative for $i \in I_{l} \cup I_{u} \cup I_{l u}$, and $\left(D_{a}\right)_{i i}(x, \varepsilon)=0,\left(D_{b}\right)_{i i}(x, \varepsilon)=-1$ for $i \in I_{f}$. Now suppose that $\nabla_{x} \Psi_{p}(x, \varepsilon)^{T} z=0$. Then,

$$
\begin{equation*}
z_{i}=-\frac{\left(D_{b}\right)_{i i}(x, \varepsilon)}{\left(D_{a}\right)_{i i}(x, \varepsilon)}\left(F^{\prime}(x) z\right)_{i}, \quad \text { for } i \in I_{l} \cup I_{u} \cup I_{l u} \tag{22}
\end{equation*}
$$

and

$$
\begin{equation*}
\left(F^{\prime}(x) z\right)_{i}=0, \quad \text { for } i \in I_{f} . \tag{23}
\end{equation*}
$$

Since $F$ is a continuously differentiable $P_{0}$-function, $F^{\prime}(x)$ is a $P_{0}$-matrix. From Lemma 2.1, we get $z_{i}=0$ for $i \in I_{l} \cup I_{u} \cup I_{l u}$. Substituting this into (23) then gives

$$
\left[F^{\prime}(x)_{I_{f} I_{f}}\right] z_{I_{f}}=0
$$

where $z_{I_{f}}$ is a vector consisting of $z_{i}$ with $i \in I_{f}$. This along with the nonsingularity of $\left[F^{\prime}(x)\right]_{I_{f} I_{f}}$ implies $z_{i}=0$ for $i \in I_{f}$. Thus, we prove $z=0$, and consequently the first part of the conclusions follows. The second part is implied by the above arguments.

Remark 3.1 We want to point out when $p \rightarrow+\infty$, the diagonal elements $\left(D_{a}\right)_{i i}(x, \varepsilon)$ and $\left(D_{b}\right)_{i i}(x, \varepsilon)$ for $i \in I_{l} \cup I_{u} \cup I_{l u}$ will tend to 0 , though $\left(D_{a}\right)_{i i}(x, \varepsilon)+\left(D_{b}\right)_{i i}(x, \varepsilon)<0$. This implies that for a larger $p$ the nonsingularity of $\nabla \Psi_{p}(x, \varepsilon)$ actually requires stronger conditions than those given by Proposition 3.5.

The boundedness of level sets of $\left\|\Phi_{p}(x)\right\|$ is also important since it ensures that the sequences generated by a descent method has at least one accumulation point. The following proposition is to prove that

$$
\begin{equation*}
\mathcal{L}(\gamma):=\left\{x \in \mathbb{R}^{n} \mid\left\|\Phi_{p}(x)\right\| \leq \gamma\right\} \tag{24}
\end{equation*}
$$

are bounded.

Proposition 3.6 The level sets $\mathcal{L}(\gamma)$ are bounded for all $\gamma>0$ if one of the following two conditions is satisfied:
(a) If $l_{i}$ and $u_{i}$ are bounded for all $i \in\{1,2, \ldots, n\}$.
(b) $F$ is a uniform $P$-function.

Proof. Under the condition (a), we have $\{1,2, \ldots, n\}=I_{l u}$. The result is clear by the definition of $\Phi_{p}$ and Lemma $2.2(\mathrm{~d})$. Next we prove the boundedness of $\mathcal{L}(\gamma)$ under the condition (b). Suppose that there exists some $\gamma>0$ such that $\mathcal{L}(\gamma)$ is unbounded, i.e., there exists a sequence $\left\{x^{k}\right\} \subseteq \mathcal{L}(\gamma)$ such that $\left\|x^{k}\right\| \rightarrow \infty$. Define the index set

$$
J:=\left\{i \in\{1,2, \ldots, n\} \mid\left\{x_{i}^{k}\right\} \text { is unbounded }\right\} .
$$

Then $J \neq \emptyset$. We choose a bounded sequence $y^{k}$ with

$$
y_{i}^{k}= \begin{cases}0 & \text { if } i \in J \\ x_{i}^{k} & \text { otherwise }\end{cases}
$$

Since F is a uniform $P$-function, there is a constant $\mu>0$ such that

$$
\begin{aligned}
\mu\left\|x^{k}-y^{k}\right\|^{2} & \leq \max _{1 \leq i \leq n}\left(x_{i}^{k}-y_{i}^{k}\right)\left(F_{i}\left(x^{k}\right)-F_{i}\left(y^{k}\right)\right) \\
& =\max _{i \in J}\left(x_{i}^{k}\right)\left(F_{i}\left(x^{k}\right)-F_{i}\left(y^{k}\right)\right) \\
& \leq\left|x_{j_{0}}^{k} \| F_{j_{0}}\left(x^{k}\right)-F_{j_{0}}\left(y^{k}\right)\right|
\end{aligned}
$$

where $j_{0}$ is an index from $\{1,2, \cdots, n\}$ for which the maximum is attained. Here we have, without loss of generality, assumed to be independent of $k$. Clearly, $j_{0} \in J$, which means that $\left\{x_{j_{0}}^{k}\right\}$ is unbounded. Consequently, there exists a subsequence, assumed to be $\left\{x_{j_{0}}^{k}\right\}$ without loss of generality, such that $\left|x_{j_{0}}^{k}\right| \rightarrow \infty$. Notice that

$$
\left\|x^{k}-y^{k}\right\|^{2} \geq\left|x_{j_{0}}^{k}-y_{j_{0}}^{k}\right|^{2}=\left|x_{j_{0}}^{k}\right|^{2} \quad \text { for each } k
$$

Therefore, $\mu\left|x_{j_{0}}^{k}\right|^{2} \leq\left|x_{j_{0}}^{k}\right|\left|F_{j_{0}}\left(x^{k}\right)-F_{j_{0}}\left(y^{k}\right)\right|$ and

$$
\mu\left|x_{j_{0}}^{k}\right| \leq\left|F_{j_{0}}\left(x^{k}\right)-F_{j_{0}}\left(y^{k}\right)\right| \leq\left|F_{j_{0}}\left(x^{k}\right)\right|+\left|F_{j_{0}}\left(y^{k}\right)\right|,
$$

which implies $\left|F_{j_{0}}\left(x^{k}\right)\right| \rightarrow \infty$ as $\left|x_{j_{0}}^{k}\right| \rightarrow \infty$. Thus, we prove that

$$
\left|x_{j_{0}}^{k}\right| \rightarrow+\infty \text { and }\left|F_{j_{0}}\left(x^{k}\right)\right| \rightarrow+\infty .
$$

Using the last equation and Lemma 2.2 (d), we have $\left|\Phi_{p, j_{0}}\left(x^{k}\right)\right| \rightarrow+\infty$ from the definition of $\Phi_{p}$. This contradicts the fact that $\left\{x^{k}\right\} \subseteq \mathcal{L}(\gamma)$.

## 4 Conclusions

In this paper, we have studied the smoothing Newton method [10] based on the smooth approximation $\psi_{p}$ of the generalized FB function, and smooth operator $\Psi_{p}$ is shown to possess the Jacobian consistence. We also believe both Proposition 3.5 and Proposition 3.6 may be useful in general smoothing algorithms for MCP.

## References

[1] S. C. Billups, S. P. Dirkse and M. C. Soares, A comparison of algorithms for large scale mixed complementarity problems, Computational Optimization and Applications, vol. 7, pp. 3-25, 1997.
[2] S. C. Billups, and M. C. Soares, QPCOMP: A quadratic programming based solver for mixed complementarity problems, Mathematical Programming, vol. 76, pp. 533-562, 1997.
[3] B. Chen and P. T. Harker, A non-intertior-point continuation method for linear complementarity problem, SIAM Journal on Matrix Analysis and Applications, vol. 14, pp. 1168-1190, 1993.
[4] C. Chen and O. L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, Computational Optimization and Applications, vol. 5, pp. 97-138, 1995.
[5] J.-S. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, Journal of Global Optimization, vol. 36, pp. 565-580, 2006.
[6] J.-S. Chen, On some NCP-functions based on the generalized Fischer-Burmeister function, Asia-Pacific Journal of Opertional Research, vol. 24, pp. 401-420, 2007.
[7] J.-S. Chen and S. Pan, A family of NCP-functions and a descent method for the nonlinear complementarity problem, Computational Optimization and Applications, vol. 40, pp. 389-404, 2008.
[8] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[9] X. Chen and L. Qi, A parameterized Newton method and a Broyden-like method for solving nonsmooth equations, Computational Optimization and Applications, vol. 3, pp. 157-179, 1994.
[10] X. Chen, L. Qi and D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, Mathematics of Computation, vol. 67, pp. 519-540, 1998.
[11] R.W. Cottle, J.-S. Pang and R.-E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.
[12] S. Dirkse, Robusr solution of mixed complementarity problems, Ph.D. Thesis, Computer Sciences Department, University of Wisconsin, 1994.
[13] S. Dirkse and M. Ferris, MCPLIB: a collection of nonlinear mixed complementarity problems, Optimization Methods and Softwares, vol. 5, pp. 319-345, 1995.
[14] F. Facchinei, and J-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I, Springer, New York, 2003.
[15] F. Facchinei, and J-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II, Springer, New York, 2003.
[16] M. Ferris, C. Kanzow, and T. Munson, Feasible descent algorithms for mixed complementarity problems, Mathematical Programming, vol. 86, pp. 475-497, 1999.
[17] M. Ferris, and J-S. Pang, Engineering and economic applications of complementarity problems, SIAM Review, vol. 39, pp. 669-713, 1997.
[18] M. Ferris and K. Sinapiromsaran, Formulating and solving nonlinear programs as mixed complementarity problems, Optimization, volume 481 of Lecture Notes in Economics and Mathematical Systems, edited by V. Nguyen, J. Strodiot, and P. Tossings, 2000.
[19] A. Fischer, A special Newton-type optimization method, Optimization, vol. 24, pp. 269-284, 1992.
[20] S. A. Gabriel and J. J. Moré, Smoothing of mixed complementarity probblems. In: M.C. Ferris and J. S. Pang (eds.): Complementarity and Variational Problems: State of the Art (SIAM, Pliladelphia, Pennsylvania, 1997)105-116.
[21] P. T. Harker and J.-S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications, Mathematical Programming, vol. 48, pp. 161-220, 1990.
[22] N. J. Higham, Estimating the matrix p-norm, Numerical Mathematics, vol. 62, pp. 539-555, 1992.
[23] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, SIAM Journal on Matrix Analysis and Applications, vol. 7, pp. 851-868, 1996.
[24] C. Kanzow and S. Petra, On a semismooth least squares formulation of mixed complementarity problems with gap reduction, Optimization Methods and Softwares, vol. 19, pp. 507-525, 2004.
[25] C. Kanzow and S. Petra, Projected filter trust region methods for a semismooth least squares formulation of mixed complementarity problems, Optimization Methods and Softwares, vol. 22, pp. 713-735, 2007.

