Nonlinear Analysis 72 (2010) 3739–3758

Contents lists available at ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

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

Jein-Shan Chen[a,*,1], Shaohua Pan[b], Tzu-Ching Lin[a]

[a] Department of Mathematics, National Taiwan Normal University, Taipei, 11677, Taiwan

[b] School of Mathematical Sciences, South China University of Technology, Guangzhou 510640, China

ARTICLE INFO

Article history:
Received 5 June 2009
Accepted 12 January 2010

Keywords:
Mixed complementarity problem
The generalized FB function
Smoothing approximation
Convergence rate

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, and then trace a smooth path generated by the smoothing algorithm proposed by Chen (2000) [28] to the solution set. In particular, we investigate the influence of p on the numerical performance of the algorithm by solving all MCPLIP test problems, and conclude that the smoothing algorithm with $p \in (1, 2]$ has better numerical performance than the one with $p > 2$.

1. Introduction

The mixed complementarity problem (MCP) arises in many applications including the fields of economics, engineering, and operations research [1–4] and has attracted much attention in last decade [5–10]. A collection of nonlinear mixed complementarity problems called MCPLIB can be found in [11] and the excellent book [12] is a good source for seeking theoretical backgrounds and numerical methods.

Given a mapping $F : [l, u] \to \mathbb{R}^n$ with $F = (F_1, \dots, F_n)^T$, where $l = (l_1, \dots, l_n)^T$ and $u = (u_1, \dots, u_n)^T$ with $l_i \in \mathbb{R} \cup \{-\infty\}$ and $u_i \in \mathbb{R} \cup \{+\infty\}$ being given lower and upper bounds satisfying $l_i < u_i$ for $i = 1, 2, \dots, n$. The MCP is to *find a vector* $x^* \in [l, u]$ *such that each component* x_i^* *satisfies exactly one of the following implications*:

$$\begin{aligned} &x_i^* = l_i \Longrightarrow F_i(x^*) \geq 0,\\ &x_i^* \in (l_i, u_i) \Longrightarrow F_i(x^*) = 0,\\ &x_i^* = u_i \Longrightarrow F_i(x^*) \leq 0. \end{aligned} \tag{1}$$

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

$$F(x) = 0; \tag{2}$$

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

$$x \geq 0, \qquad F(x) \geq 0, \qquad \langle x, F(x)\rangle = 0. \tag{3}$$

* Corresponding author. Tel.: +886 2 29325417; fax: +886 2 29332342.
E-mail addresses: jschen@math.ntnu.edu.tw (J.-S. Chen), shhpan@scut.edu.cn (S. Pan), cashplayer35@yahoo.com.tw (T.-C. Lin).

1 Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office.

0362-546X/$ – see front matter
doi:10.1016/j.na.2010.01.012

In fact, from Theorem 2 of [13], 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

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

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). Specifically, $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is called an NCP function if

$$\phi(a, b) = 0 \Longleftrightarrow a \geq 0, \quad b \geq 0, \qquad ab = 0. \tag{5}$$

With such a function, the MCP (1) can be reformulated as a nonsmooth system $\Phi(x) = 0$, 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 ϕ. In the past two decades, many smooth approximations and Newton-type methods using smoothing NCP functions for complementarity problems have been proposed (see, e.g., [14–18,8,19]). Most of these methods focus on the Chen–Mangasarian class of smooth approximations of the minimum NCP function or the smoothing function of the Fischer–Burmeister (FB) NCP function. It is worthwhile to mention that the smoothing Newton method developed by Chen et al. [19] has global and superlinear (even quadratic) convergence by solving only one linear system of equations at each iteration.

Recently, an extension of the FB NCP function was considered in [20–22] by two of the authors. Specifically, they define the generalized FB function as

$$\phi_p(a, b) := \|(a, b)\|_p - (a + b) \quad \forall a, b \in \mathbb{R}, \tag{6}$$

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 ϕ_p, they replace the 2-norm of (a, b) involved in the FB function by a more general p-norm. The function ϕ_p is still an NCP-function — that is, it satisfies the equivalence in (5). Moreover, it turns out that ϕ_p possesses all favorable properties of the FB function; see [20–22]. For example, ϕ_p is strongly semismooth and its square is a continuously differentiable NCP function. In particular, numerical results in [23] for all MCPLIB problems indicate that the least-square semismooth Newton method with p close to 1 has better performance than the case of $p = 2$. Thus, it is natural to ask whether the smoothing Newton method based on ϕ_p has similar a numerical performance.

In this paper, we are concerned with the smoothing Newton method [19,28] based on the generalized FB function, motivated by the inexpensive computation work of the method at each iteration, and the fact that there are no corresponding numerical experiments to verify the effectiveness of this algorithm. We investigate the influence of the parameter p on the numerical performance of the smoothing method for solving the MCPLIB test problems. Specifically, in Section 3, we present a smoothing function of the generalized FB function, and studied some of its favorable properties, including the Jacobian consistency property; in Section 4, we describe the iterative steps of the smoothing algorithm and provide the corresponding conditions for the global convergence and local superlinear (or quadratic) convergence; in Section 5, we report the numerical results of the smoothing algorithm for solving the MCPLIB test problems.

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors and e_i means a unit vector with ith component being 1 and the others being 0. For a differentiable mapping F, $F'(x)$ and $\nabla F(x)$ to denote the Jacobian of F at x and the transposed Jacobian of F, respectively. Given an index set $\mathcal{I}$, the notation $[F'(x)]_{\mathcal{I}\mathcal{I}}$ denotes the submatrix consisting of the ith row and the jth column 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 $G : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz continuous mapping. Then, G is almost everywhere differentiable by Rademacher's Theorem (see [24]). 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

$$\partial_B G(x) := \left\{ V \in \mathbb{R}^{m\times n} \,|\, \exists \{x^k\} \subseteq D_G : \{x^k\} \to x \text{ and } G'(x^k) \to V \right\},$$

where D_G is the set of differentiable points of G. In other words, $\partial G(x) = \text{conv}\partial_B G(x)$. 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, and Qi [25] proposed so-called *C-subdifferential* of G:

$$\partial_C G(x)^T := \partial G_1(x) \times \cdots \times \partial G_m(x) \tag{7}$$

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 [24], $\partial G(x)^T \subseteq \partial_C G(x)^T$. We assume that the reader is familiar with the concepts of (strongly) semismooth functions, and refer to [26,27] for details.

We also need the definitions of P-functions and P-matrices in the subsequent sections.

Definition 2.1. Let $F = (F_1, \ldots, F_n)^T$ with $F_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, 2, \ldots, n$. Then,

(a) the mapping F is called a P_0-function if, for every x and y in $\mathbb{R}^n$ with $x \neq y$, there is an index $i \in \{1, 2, \ldots, n\}$ such that

$$x_i \neq y_i \quad \text{and} \quad (x_i - y_i)(F_i(x) - F_i(y)) \geq 0;$$

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

$$x_i \neq y_i \quad \text{and} \quad (x_i - y_i)(F_i(x) - F_i(y)) > 0.$$

(c) the mapping F is called a uniform P-function if there exists a positive constant $\mu > 0$ such that, for every x and y in $\mathbb{R}^n$, there is an index $i \in \{1, 2, \ldots, n\}$ such that

$$(x_i - y_i)(F_i(x) - F_i(y)) \geq \mu \|x - y\|^2.$$

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 Definitions 2.1 and 2.2, it is not hard to 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 ([12]). *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 $x_i(Mx)_i \geq 0$.*

Next we recall some favorable properties of ϕ_p whose proofs can be found in [20–22].

Lemma 2.2. *Let $\phi_p : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by (6). Then, the following results hold.*

(a) *ϕ_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 $(a, b) \neq (0, 0)$,*

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

and otherwise (ξ, ζ) 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 ϕ_p is a continuously differentiable NCP function.*
(d) *If $\{(a^k, b^k)\} \subseteq \mathbb{R}^2$ satisfies $(a^k \to -\infty)$ or $(b^k \to -\infty)$ or $(a^k \to \infty$ and $b^k \to \infty)$, then we have $|\phi_p(a^k, b^k)| \to \infty$ as $k \to \infty$.*

The following lemma establishes another property of ϕ_p, which plays a key role in the nonsmooth system reformulation of the MCP (1) with the generalized FB function.

Lemma 2.3. *Let $\phi_p : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by (6). Then, the following limits hold.*

(a) $\lim\limits_{l_i \to -\infty} \phi_p\left(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))\right) = -\phi_p\,(u_i - x_i, -F_i(x))$.
(b) $\lim\limits_{u_i \to \infty} \phi_p\left(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))\right) = \phi_p\,(x_i - l_i, F_i(x))$.
(c) $\lim\limits_{l_i \to -\infty} \lim\limits_{u_i \to \infty} \phi_p\left(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))\right) = -F_i(x)$.

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

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

where the second equality is using the Taylor expansion of the function $(1 + t)^{1/p}$ and the notation $o(t)$ means $\lim_{t\to 0} o(t)/t = 0$. Since $a^k \to +\infty$ as $k \to \infty$, we have $\frac{|b|^{np}}{(a^k)^{np-1}} \to 0$ for all n. This together with the last equation implies $\lim_{k\to\infty} \phi_p(a^k, b) = -b$. This proves part (a). Part (b) and (c) are direct by part (a) and the continuity of ϕ_{FB}. $\square$

To close this section, we summarize the monotonicity of two scalar-valued functions that will be used in the subsequent section. Since the proof is direct, we omit it here.

Lemma 2.4. *For any fixed $0 \le \mu_1 < \mu_2$, the following functions*

$$f_1(t) := (t+\mu_1)^{-\frac{p-1}{p}} - (t+\mu_2)^{-\frac{p-1}{p}} \quad (t > 0)$$

and

$$f_2(t) := (t+\mu_2)^{\frac{p-1}{p}} - (t+\mu_1)^{\frac{p-1}{p}} \quad (t \ge 0)$$

are decreasing on $(0, +\infty)$, and furthermore, $f_2(t) \le f_2(0) = \mu_2^{(p-1)/p} - \mu_1^{(p-1)/p}$.

3. The smoothing function and its properties

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

$$\begin{aligned} I_l &:= \{i \in \{1,2,\dots,n\} \mid -\infty < l_i < u_i = +\infty\}, \\ I_u &:= \{i \in \{1,2,\dots,n\} \mid -\infty = l_i < u_i < +\infty\}, \\ I_{lu} &:= \{i \in \{1,2,\dots,n\} \mid -\infty < l_i < u_i < +\infty\}, \\ I_f &:= \{i \in \{1,2,\dots,n\} \mid -\infty = l_i < u_i = +\infty\}. \end{aligned} \tag{8}$$

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

$$\Phi_{p,i}(x) := \begin{cases} \phi_p(x_i - l_i, F_i(x)) & \text{if } i \in I_l, \\ -\phi_p(u_i - x_i, -F_i(x)) & \text{if } i \in I_u, \\ \phi_p(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))) & \text{if } i \in I_{lu}, \\ -F_i(x) & \text{if } i \in I_f, \end{cases} \tag{9}$$

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 in (5), it is not difficult to verify that the following result holds.

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

We want to point out that, unlike for the nonlinear complementarity problem, when writing the generalized FB function ϕ_p as $\phi_p(a,b) = (a+b) - \|(a,b)\|_p$, the conclusion of Proposition 3.1 does not necessarily hold since, if $I_l = \{1,2,\dots,n\}$, then $\bar{x} = l$ satisfies $\Phi_p(\bar{x}) = 0$, but $F(\bar{x}) \ge 0$ does not necessarily hold. Similar phenomenon also appears when replacing ϕ_p by the minimum NCP function.

Since ϕ_p is not differentiable at the origin, 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 $\Psi_p : \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}^n$ is defined componentwise as

$$\Psi_{p,i}(x,\varepsilon) := \begin{cases} \psi_p(x_i - l_i, F_i(x), \varepsilon) & \text{if } i \in I_l, \\ -\psi_p(u_i - x_i, -F_i(x), \varepsilon) & \text{if } i \in I_u, \\ \psi_p\left(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon\right) & \text{if } i \in I_{lu}, \\ -F_i(x) & \text{if } i \in I_f, \end{cases} \tag{10}$$

with

$$\psi_p(a,b,\varepsilon) := \sqrt[p]{|a|^p + |b|^p + \varepsilon^p} - (a+b). \tag{11}$$

In what follows, we concentrate on the favorable properties of the smoothing function ψ_p and the operator Ψ_p. First, let us state the favorable properties of ψ_p.

Lemma 3.1. *Let $\psi_p : \mathbb{R}^3 \to \mathbb{R}$ be defined by (11). Then, the following result holds.*

(a) *For any fixed $\varepsilon > 0$, $\psi_p(a,b,\varepsilon)$ is continuously differentiable at all $(a,b) \in \mathbb{R}^2$ with*

$$-2 < \frac{\partial \psi_p(a,b,\varepsilon)}{\partial a} < 0, \qquad -2 < \frac{\partial \psi_p(a,b,\varepsilon)}{\partial b} < 0. \tag{12}$$

(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 $0 < \varepsilon_1 \le \varepsilon_2$,*

$$0 \le \psi_p(a,b,\varepsilon_2) - \psi_p(a,b,\varepsilon_1) \le (\varepsilon_2 - \varepsilon_1). \tag{13}$$

In particular, $|\psi_p(a,b,\varepsilon) - \phi_p(a,b)| \le \varepsilon$ for all $\varepsilon \ge 0$.

(c) *For any fixed $(a,b) \in \mathbb{R}^2$, let $\psi_p^0(a,b) := \left(\lim_{\varepsilon \downarrow 0} \frac{\partial \psi_p(a,b,\varepsilon)}{\partial a}, \lim_{\varepsilon \downarrow 0} \frac{\partial \psi_p(a,b,\varepsilon)}{\partial b}\right)$. Then,*

$$\lim_{h=(h_1,h_2)\to(0,0)} \frac{\phi_p(a+h_1, b+h_2) - \phi_p(a,b) - \psi_p^0(a+h_1, b+h_2)^T h}{\|h\|} = 0.$$

(d) *For any given* $\varepsilon > 0$, *if* $p \geq 2$, *then* $\psi_p(a, b, \varepsilon) = 0 \Longrightarrow a > 0,\ b > 0,\ 2ab \leq \varepsilon^2$, *and whenever* $p > 1$, $\psi_p(a, b, \varepsilon) = 0 \Longrightarrow a > 0,\ b > 0,\ \min\{a, b\} \leq \frac{\varepsilon}{\sqrt[p]{2^p-2}}$.

Proof. (a) Using an elementary calculation, we immediately obtain that

$$\frac{\partial \psi_p(a,b,\varepsilon)}{\partial a} = \frac{\operatorname{sign}(a)|a|^{p-1}}{\left(\sqrt[p]{|a|^p+|b|^p+\varepsilon^p}\right)^{p-1}} - 1,$$
$$\frac{\partial \psi_p(a,b,\varepsilon)}{\partial b} = \frac{\operatorname{sign}(b)|b|^{p-1}}{\left(\sqrt[p]{|a|^p+|b|^p+\varepsilon^p}\right)^{p-1}} - 1. \tag{14}$$

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 all $(a, b) \in \mathbb{R}^2$, it follows that $\psi_p(a, b, \varepsilon)$ is continuously differentiable at all $(a, b) \in \mathbb{R}^2$. Noting that

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

we readily get the inequality (12).

(b) For any $\varepsilon > 0$, an elementary calculation yields that

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

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 $\varepsilon_0 \in (\varepsilon_1, \varepsilon_2)$ such that

$$\psi_p(a, b, \varepsilon_2) - \psi_p(a, b, \varepsilon_1) = \frac{\partial \psi_p}{\partial \varepsilon}(a, b, \varepsilon_0)(\varepsilon_2 - \varepsilon_1).$$

Since $\frac{\partial \psi_p}{\partial \varepsilon}(a, b, \varepsilon_0) \leq 1$ by the proof of part (a), inequality (13) holds for all $0 < \varepsilon_1 \leq \varepsilon_2$. Letting $\varepsilon_1 \downarrow 0$, the desired result then follows.

(c) Using the formula (14), it is easy to calculate that

$$\lim_{\varepsilon\downarrow 0}\frac{\partial \psi_p(a,b,\varepsilon)}{\partial a} = \begin{cases} \dfrac{\operatorname{sign}(a)|a|^{p-1}}{\left(\sqrt[p]{|a|^p+|b|^p}\right)^{p-1}} - 1 & \text{if } (a,b) \neq (0,0), \\ -1 & \text{if } (a,b) = (0,0); \end{cases}$$
$$\lim_{\varepsilon\downarrow 0}\frac{\partial \psi_p(a,b,\varepsilon)}{\partial b} = \begin{cases} \dfrac{\operatorname{sign}(b)|b|^{p-1}}{\left(\sqrt[p]{|a|^p+|b|^p}\right)^{p-1}} - 1 & \text{if } (a,b) \neq (0,0), \\ -1 & \text{if } (a,b) = (0,0). \end{cases} \tag{15}$$

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

$$\begin{aligned}\phi_p(h_1, h_2) - \phi_p(0, 0) - \psi_p^0(h_1, h_2)^T h &= \sqrt[p]{|h_1|^p+|h_2|^p} - \frac{|h_1|^p+|h_2|^p}{(\sqrt[p]{|h_1|^p+|h_2|^p})^{p-1}} \\ &= \sqrt[p]{|h_1|^p+|h_2|^p} - \sqrt[p]{|h_1|^p+|h_2|^p} \\ &= 0.\end{aligned}$$

(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 $b \geq 0$. Note that, whenever $a \geq 0, b \leq 0$ or $a \leq 0, b \geq 0$, there holds that

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

i.e., $\psi_p(a, b, \varepsilon) > 0$. Hence, for any given $\varepsilon > 0$, $\psi_p(a, b, \varepsilon) = 0$ implies $a > 0$ and $b > 0$.

(i) If $p \geq 2$, using the nonincreasing of p-norm with respect to p leads to

$$\begin{aligned}\psi_p(a, b, \varepsilon) = 0 &\Longleftrightarrow a + b = \sqrt[p]{|a|^p+|b|^p+\varepsilon^p} \leq \sqrt{|a|^2+|b|^2+\varepsilon^2} \\ &\Longrightarrow\ (a+b)^2 \leq a^2 + b^2 + \varepsilon^2 \Longrightarrow 2ab \leq \varepsilon^2.\end{aligned}$$

(ii) For $p > 1$, without loss of generality, we assume $0 < a \le b$. For any fixed $a \ge 0$, consider $f(t) = (t+a)^p - t^p - a^p - \varepsilon^p$ $(t \ge 0)$. It is easy to verify that the function f is strictly increasing on $[0, +\infty)$. Since $\psi_p(a, b, \varepsilon) = 0$, we have $f(b) = 0$ which says $f(a) = (2^p - 2)a^p - \varepsilon^p \le f(b) = 0$. From this inequality, we get $\min\{a, b\} = a \le \frac{\varepsilon}{\sqrt[p]{2^p-2}}$. $\square$

Using Lemma 3.1 and the expression of Ψ_p, we readily obtain the following result.

Proposition 3.2. *Let Ψ_p be defined by (10). Then, the following results hold.*

(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 $(D_a)_{ii}(x, \varepsilon)$ and $(D_b)_{ii}(x, \varepsilon)$ defined as follows:

(a1) *For $i \in I_l$,*

$$(D_a)_{ii}(x, \varepsilon) = \frac{\operatorname{sign}(x_i - l_i)|x_i - l_i|^{p-1}}{\|(x_i - l_i, F_i(x), \varepsilon)\|_p^{p-1}} - 1,$$

$$(D_b)_{ii}(x, \varepsilon) = \frac{\operatorname{sign}(F_i(x))|F_i(x)|^{p-1}}{\|(x_i - l_i, F_i(x), \varepsilon)\|_p^{p-1}} - 1.$$

(a2) *For $i \in I_u$,*

$$(D_a)_{ii}(x, \varepsilon) = \frac{\operatorname{sign}(u_i - x_i)|u_i - x_i|^{p-1}}{\|(u_i - x_i, F_i(x), \varepsilon)\|_p^{p-1}} - 1,$$

$$(D_b)_{ii}(x, \varepsilon) = \frac{-\operatorname{sign}(F_i(x))|F_i(x)|^{p-1}}{\|(u_i - x_i, F_i(x), \varepsilon)\|_p^{p-1}} - 1.$$

(a3) *For $i \in I_{lu}$,*

$$(D_a)_{ii}(x, \varepsilon) = a_i(x, \varepsilon) + b_i(x, \varepsilon)c_i(x, \varepsilon) \quad \text{and} \quad (D_b)_{ii}(x, \varepsilon) = b_i(x, \varepsilon)d_i(x, \varepsilon)$$

with

$$a_i(x, \varepsilon) = \frac{\operatorname{sign}(x_i - l_i)|x_i - l_i|^{p-1}}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{p-1}} - 1,$$

$$b_i(x, \varepsilon) = \frac{\operatorname{sign}(\psi_p(u_i - x_i, -F_i(x), \varepsilon))|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{p-1}} - 1,$$

$$c_i(x, \varepsilon) = -\frac{\operatorname{sign}(u_i - x_i)|u_i - x_i|^{p-1}}{\|(u_i - x_i, F_i(x), \varepsilon)\|_p^{p-1}} + 1,$$

$$d_i(x, \varepsilon) = \frac{\operatorname{sign}(F_i(x))|F_i(x)|^{p-1}}{\|(u_i - x_i, F_i(x), \varepsilon)\|_p^{p-1}} + 1.$$

(a4) *For $i \in I_f$, $(D_a)_{ii}(x, \varepsilon) = 0$ and $(D_b)_{ii}(x, \varepsilon) = -1$.*

Moreover, $-2 < (D_a)_{ii}(x, \varepsilon) < 0$ and $-2 < (D_b)_{ii}(x, \varepsilon) < 0$ for all $i \in I_l \cup I_u$, and $-6 < (D_a)_{ii}(x, \varepsilon) < 0$ and $-4 < (D_b)_{ii}(x, \varepsilon) < 0$ for $i \in I_{lu}$.

(b) *For any given $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, we have*

$$\|\Psi_p(x, \varepsilon_2) - \Psi_p(x, \varepsilon_1)\| \le \sqrt{n}\left(\sqrt[p]{2} + 1\right)|\varepsilon_2 - \varepsilon_1|, \quad \forall x \in \mathbb{R}^n.$$

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

$$\|\Psi_p(x, \varepsilon) - \Phi_p(x)\| \le \sqrt{n}\left(\sqrt[p]{2} + 1\right)\varepsilon, \quad \forall x \in \mathbb{R}^n.$$

The Jacobian consistency property plays a crucial role in the analysis of local fast convergence of the smoothing algorithm [19]. To show that the smoothing operator Ψ_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 $x \in \mathbb{R}^n$, $\partial_C \Phi_p(x)^T = \{D_a(x) + \nabla F(x) D_b(x)\}$, where $D_a(x)$, $D_b(x)$ are $n \times n$ diagonal matrices whose diagonal elements are given as below:*

(a) *For $i \in I_l$, if $(x_i - l_i, F_i(x)) \neq (0, 0)$, then*

$$(D_a)_{ii}(x) = \frac{\operatorname{sign}(x_i - l_i) \cdot |x_i - l_i|^{p-1}}{\|(x_i - l_i, F_i(x))\|_p^{p-1}} - 1,$$

$$(D_b)_{ii}(x) = \frac{\operatorname{sign}(F_i(x)) \cdot |F_i(x)|^{p-1}}{\|(x_i - l_i, F_i(x))\|_p^{p-1}} - 1;$$

and otherwise

$$((D_a)_{ii}(x), (D_b)_{ii}(x)) \in \left\{ (\xi - 1, \zeta - 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1 \right\}.$$

(b) *For $i \in I_u$, if $(u_i - x_i, -F_i(x)) \neq (0, 0)$, then*

$$(D_a)_{ii}(x) = \frac{\mathrm{sign}(u_i - x_i) \cdot |u_i - x_i|^{p-1}}{\|(u_i - x_i, -F_i(x))\|_p^{p-1}} - 1,$$
$$(D_b)_{ii}(x) = -\frac{\mathrm{sign}(F_i(x)) \cdot |F_i(x)|^{p-1}}{\|(u_i - x_i, -F_i(x))\|_p^{p-1}} - 1;$$

and otherwise

$$((D_a)_{ii}(x), (D_b)_{ii}(x)) \in \left\{ (\xi - 1, \zeta - 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1 \right\}.$$

(c) *For $i \in I_{lu}$, $(D_a)_{ii}(x) = a_i(x) + b_i(x)c_i(x)$ and $(D_b)_{ii}(x) = b_i(x)d_i(x)$ where, if $(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))) \neq (0, 0)$, then*

$$a_i(x) = \frac{\mathrm{sign}(x_i - l_i) \cdot |x_i - l_i|^{p-1}}{\left\| \left(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))\right) \right\|_p^{p-1}} - 1,$$
$$b_i(x) = \frac{\mathrm{sign}\left(\phi_p(u_i - x_i, -F_i(x))\right) \cdot \left|\phi_p(u_i - x_i, -F_i(x))\right|^{p-1}}{\left\| \left(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))\right) \right\|_p^{p-1}} - 1,$$

and otherwise

$$(a_i(x), b_i(x)) \in \left\{ (\xi - 1, \zeta - 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1 \right\};$$

and if $(u_i - x_i, -F_i(x)) \neq (0, 0)$, then

$$c_i(x) = \frac{-\mathrm{sign}(u_i - x_i) \cdot |u_i - x_i|^{p-1}}{\|(u_i - x_i, -F_i(x))\|_p^{p-1}} + 1,$$
$$d_i(x) = \frac{\mathrm{sign}\,(F_i(x)) \cdot |F_i(x)|^{p-1}}{\|(u_i - x_i, -F_i(x))\|_p^{p-1}} + 1,$$

and otherwise

$$(c_i(x), d_i(x)) \in \left\{ (\xi + 1, \zeta + 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1 \right\}.$$

(d) *For $i \in I_f$, $(D_a)_{ii}(x) = 0$ and $(D_b)_{ii}(x) = -1$.*

Now we are in a position to establish the Jacobian consistency of the operator Ψ_p.

Proposition 3.4. *Let Ψ_p be defined by (10). Then, for any fixed $x \in \mathbb{R}^n$,*

$$\lim_{\varepsilon \downarrow 0} \mathrm{dist}(\nabla_x \Psi_p(x, \varepsilon)^T, \partial_C \Phi_p(x)) = 0.$$

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

$$\begin{aligned}
&\beta_1(x) := \{i \in I_l \mid (x_i - l_i, F_i(x)) = (0, 0)\}, \qquad \bar\beta_1(x) := I_l \setminus \beta_1(x),\\
&\beta_2(x) := \{i \in I_u \mid (u_i - x_i, F_i(x)) = (0, 0)\}, \qquad \bar\beta_2(x) := I_u \setminus \beta_2(x),\\
&\beta_3(x) := \{i \in I_{lu} \mid \left(x_i - l_i, \phi_p(u_i - x_i, -F_i(x))\right) = (0, 0)\}, \qquad \bar\beta_3(x) := I_{lu} \setminus \beta_3(x),\\
&\beta_4(x) := \{i \in \bar\beta_3(x) \mid (u_i - x_i, F_i(x)) = (0, 0)\}, \qquad \bar\beta_4(x) := \bar\beta_3(x) \setminus \beta_4(x).
\end{aligned} \tag{16}$$

We proceed the arguments by the cases $i \in I_l \cup I_u$, $i \in I_{lu}$ 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

$$(D_a)_{ii}(x, \varepsilon) = -1 \quad \text{and} \quad (D_b)_{ii}(x, \varepsilon) = -1.$$

By Proposition 3.2(a1) and (a2), $\nabla_x \Psi_{p,i}(x, \varepsilon)^T = -e_i^T - F_i'(x)$ for all $\varepsilon > 0$. Since

$$(-1, -1) \in \left\{ (\xi - 1, \zeta - 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1 \right\}, \tag{17}$$

from Proposition 3.3(a) and (b) we get $\nabla_x \Psi_{p,i}(x, \varepsilon)^T \in \partial_C \Phi_{p,i}(x)$. When $i \in \bar\beta_1(x) \cup \bar\beta_2(x)$,

$$\lim_{\varepsilon \downarrow 0} (D_a)_{ii}(x, \varepsilon) = (D_a)_{ii}(x) \quad \text{and} \quad \lim_{\varepsilon \downarrow 0} (D_b)_{ii}(x, \varepsilon) = (D_b)_{ii}(x),$$

which together with Proposition 3.2(a1) and (a2) implies that

$$\lim_{\varepsilon\downarrow 0} \nabla_x \Psi_{p,i}(x,\varepsilon)^T = (D_a)_{ii}(x)e_i^T + (D_b)_{ii}(x)F_i'(x) \in \partial_C \Phi_{p,i}(x).$$

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

$$\lim_{\varepsilon\downarrow 0} \nabla_x \Psi_{p,i}(x,\varepsilon)^T \in \partial_C \Phi_{p,i}(x), \quad \forall i \in I_l \cup I_u. \tag{18}$$

Case 2: $i \in I_{lu}$. When $i \in \beta_3(x)$, we have $x_i - l_i = 0$, $\phi_p(u_i - x_i, -F_i(x)) = 0$, $u_i - x_i > 0$ and $F_i(x) = 0$. Hence, $c_i(x) = 0$ and $d_i(x) = 1$. From Proposition 3.3(c), it follows that

$$\partial_C \Phi_{p,i}(x) = \{a_i(x)e_i^T + b_i(x)F_i'(x)\} \tag{19}$$

with

$$(a_i(x), b_i(x)) \in \left\{(\xi - 1, \zeta - 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1\right\}.$$

On the other hand, since $a_i(x,\varepsilon) = -1$, $d_i(x,\varepsilon) = 1$ and

$$b_i(x,\varepsilon) = \frac{|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}}{\left(|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^p + \varepsilon^p\right)^{\frac{p-1}{p}}} - 1,$$

$$c_i(x,\varepsilon) = 1 - \frac{|u_i - x_i|^{p-1}}{(|u_i - x_i|^p + \varepsilon^p)^{(p-1)/p}},$$

from Proposition 3.2(a3) it follows that

$$\nabla_x \Psi_{p,i}(x,\varepsilon)^T = (-1 + b_i(x,\varepsilon)c_i(x,\varepsilon))e_i^T + b_i(x,\varepsilon)F_i'(x). \tag{20}$$

Taking

$$\xi = 0 \quad \text{and} \quad \zeta = \frac{|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}}{\left(|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^p + \varepsilon^p\right)^{\frac{p-1}{p}}},$$

it is not hard to verify that $|\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1$, and consequently

$$-e_i^T + b_i(x,\varepsilon)F_i'(x) \in \partial_C \Phi_{p,i}(x).$$

Noting that

$$\lim_{\varepsilon\downarrow 0} \left\| \nabla_x \Psi_{p,i}(x,\varepsilon)^T - \left(-e_i^T + b_i(x,\varepsilon)F_i'(x)\right) \right\| = \lim_{\varepsilon\downarrow 0} \|b_i(x,\varepsilon)c_i(x,\varepsilon)e_i^T\| = 0,$$

it then follows that

$$\lim_{\varepsilon\downarrow 0} \operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) = 0, \quad i \in \beta_3(x).$$

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,

$$c_i(x,\varepsilon) = 1, \qquad d_i(x,\varepsilon) = 1 \quad \text{for } i \in \beta_4(x)$$

and

$$\lim_{\varepsilon\downarrow 0} c_i(x,\varepsilon) = c_i(x), \qquad \lim_{\varepsilon\downarrow 0} d_i(x,\varepsilon) = d_i(x) \quad \text{for } i \in \bar{\beta}_4(x).$$

Using Proposition 3.3(c) and noting that

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

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,

$$\lim_{\varepsilon\downarrow 0} \nabla_x \Psi_{p,i}(x,\varepsilon)^T \in \partial_C \Phi_{p,i}(x) \quad \text{for } i \in I_{lu}. \tag{21}$$

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

$$\lim_{\varepsilon\downarrow 0} \nabla_x \Psi_{p,i}(x,\varepsilon)^T \in \partial_C \Phi_{p,i}(x) \quad \text{for } i \in I_f. \tag{22}$$

Now the desired result follows from (18)–(22) and $\{1, 2, \ldots, n\} = I_f \cup I_l \cup I_u \cup I_{lu}$. $\square$

Proposition 3.4 implies that for any $\delta > 0$, there exists an $\varepsilon(x, \delta) > 0$ such that

$$\operatorname{dist}\left(\nabla_x \Psi_p(x, \varepsilon)^T, \partial_C \Phi_p(x)\right) \leq \delta \quad \text{for all } 0 < \varepsilon \leq \varepsilon(x, \delta).$$

The following lemma gives a way to choose such $\varepsilon(x, \delta)$, whose proof is seen in Appendix.

Lemma 3.2. *Let Ψ_p be defined by (10). Suppose that x is not a solution of (1). Let*

$$\alpha(x) := \min\{\alpha_1(x), \alpha_2(x), \alpha_3(x)\} > 0, \qquad \gamma(x) := \max\{\gamma_1(x), \gamma_2(x), \gamma_3(x)\} \geq 0$$

with

$$\begin{aligned}
\alpha_1(x) &:= \min_{i \in \bar{\beta}_1(x)} |x_i - l_i|^p + |F_i(x)|^p, \\
\alpha_2(x) &:= \min_{i \in \bar{\beta}_2(x) \cup \bar{\beta}_4(x)} |u_i - x_i|^p + |F_i(x)|^p, \\
\alpha_3(x) &:= \min_{i \in \bar{\beta}_4(x) \cup \{i \mid |x_i - l_i| \neq 0\}} |x_i - l_i|^p + |\phi_p(u_i - x_i, -F_i(x))|^p \\
\gamma_1(x) &:= \max_{i \in \bar{\beta}_1(x)} \left\| \operatorname{sign}(x_i - l_i)|x_i - l_i|^{p-1} e_i + \operatorname{sign}(F_i(x))|F_i(x)|^{p-1} \nabla F_i(x) \right\| \\
\gamma_2(x) &:= \max_{i \in \bar{\beta}_2(x)} \left\| \operatorname{sign}(F_i(x))|F_i(x)|^{p-1} \nabla F_i(x) - \operatorname{sign}(u_i - x_i)|u_i - x_i|^{p-1} e_i \right\| \\
\gamma_3(x) &:= \max_{i \in \bar{\beta}_4(x)} |u_i - x_i|^{p-1} + |F_i(x)|^{p-1}.
\end{aligned}$$

Then, for any $\delta > 0$, there exists an $\varepsilon(x, \delta) > 0$ such that

$$\operatorname{dist}\left(\nabla_x \Psi_p(x, \varepsilon)^T, \partial_C \Phi_p(x)\right) \leq \delta \quad \textit{for all } 0 < \varepsilon \leq \varepsilon(x, \delta), \tag{23}$$

where

$$\varepsilon(x, \delta) := \min\left\{ \varepsilon_0(x, \delta), \varepsilon_1(x, \delta), \varepsilon_2(x, \delta), \varepsilon_3(x, \delta), \left(\frac{\delta}{\sqrt{n} M(x)}\right)^{\frac{p-1}{p}} \right\}$$

with

$$\varepsilon_0(x, \delta) := \min_{i \in \beta_3(x)} \left[\frac{|u_i - x_i|^{p-1}}{(1 - \delta/\sqrt{n})^{\frac{p}{p-1}}} - |u_i - x_i|^p \right]^{1/p}, \qquad \varepsilon_2(x, \delta) := \min_{i \in \beta_4(x)} \frac{1}{2} |x_i - l_i|,$$

$$\varepsilon_1(x, \delta) := \begin{cases} 1 & \textit{if } \left(\dfrac{\sqrt{n}\gamma(x)}{\delta}\right)^{\frac{p}{p-1}} - \alpha(x) \leq 0, \\[2ex] \alpha(x)^{2/p} \left(\dfrac{\sqrt{n}\gamma(x)}{\delta}\right)^{(p/(p-1) - \alpha(x))^{-1/p}} & \text{otherwise,} \end{cases}$$

$$\varepsilon_3(x, \delta) := \begin{cases} 1 & \text{if } \phi_p(u_i - x_i, -F_i(x)) \geq 0, \\ \dfrac{1}{2}\left[(u_i - x_i - F_i(x))^p - |u_i - x_i|^p - |F_i(x)|^p\right]^{1/p} & \text{otherwise.} \end{cases}$$

4. Smoothing algorithm and convergence results

In this section, we describe the iteration steps of the smoothing algorithm based on the smooth approximation $\Psi_p(x, \varepsilon) = 0$ of $\Phi_p(x) = 0$, and then present the global and local convergence results of the algorithm. To this end, we need the following merit functions:

$$\Theta_p(x) := \frac{1}{2} \|\Phi_p(x)\|^2$$

and

$$H_p(x, \varepsilon) := \frac{1}{2} \|\Psi_p(x, \varepsilon)\|^2.$$

The algorithm follows the same line as the one proposed by Chen et al. [19].

Algorithm 4.1 (*Smoothing Algorithm*).

(S.0) Given a starting point $x^0 \in \mathbb{R}^n$, the parameters $\rho, \alpha, \eta \in (0, 1)$ and $\nu \in (0, +\infty)$. Choose $\sigma \in (0, (1 - \alpha)/2)$. Let $\beta_0 = \|\Phi_p(x^0)\|$ and $\varepsilon_0 := \frac{\alpha}{2\sqrt{n}}$. Set $k := 0$.

(S.1) Solve the following linear system of equations

$$\Phi_p(x^k) + \Psi_p(x^k, \varepsilon^k)d = 0,$$

and denote its solution by d^k.

(S.2) Let m_k be the smallest nonnegative integer m such that

$$H_p(x^k + \rho^m d^k, \varepsilon^k) - H_p(x^k, \varepsilon^k) \le -2\sigma\rho^m \Theta_p(x^k).$$

Set $t_k := \rho^{m_k}$ and $x^{k+1} := x^k + t_k d^k$.

(S.3) If $\|\Phi_p(x^{k+1})\| = 0$, then terminate. If

$$0 < \|\Phi_p(x^{k+1})\| \le \max\left\{\eta\beta_k, \alpha^{-1}\|\Phi_p(x^{k+1}) - \Psi(x^{k+1}, \varepsilon^k)\|\right\}, \tag{24}$$

let $\beta_{k+1} = \|\Phi_p(x^{k+1})\|$ and choose an ε_{k+1} satisfying

$$0 < \varepsilon_{k+1} \le \min\left\{\frac{\alpha\beta_{k+1}}{2\sqrt{n}}, \frac{\varepsilon_k}{2}\right\} \tag{25}$$

and

$$\operatorname{dist}\left(\nabla_x \Psi_p(x^{k+1}, \varepsilon^{k+1}), \partial_C \Phi_p(x^{k+1})\right) \le \beta_{k+1}\nu. \tag{26}$$

If $\|\Phi_p(x^{k+1})\| > 0$ but (24) does not hold, then let $\beta_{k+1} = \beta_k$ and $\varepsilon_{k+1} = \varepsilon_k$.

(S.4) Set $k := k + 1$, and go to (S.1).

In Algorithm 4.1, the parameter σ chosen from $(0, (1-\alpha)/2)$ has twofold purposes: one is to guarantee the existence of m_k in (S.2), and the other is to lend itself to the superlinear convergence analysis of the algorithm, the initial β_0 and ε_0 are chosen as $\|\Phi_p(x^0)\|$ and $\frac{\alpha}{2\sqrt{n}}$, respectively, just for the global convergence analysis of the algorithm. Such choices for these parameters are also used in the numerical experiments of Section 5. Algorithm 4.1 has inexpensive computation work and only a system of linear equations is solved at each iteration. Since the operator Ψ_p has the Jacobian consistency property, we can find an $\varepsilon_{k+1} > 0$ such that (25) and (26) hold by the definition, and moreover, Lemma 3.2 shows how to choose an $\varepsilon_{k+1} > 0$ satisfying (25) and (26) for the MCP (1).

Lemma 4.1. *For any fixed $\varepsilon > 0$, the Jacobian matrix of Ψ_p at any $x \in \mathbb{R}^n$ is nonsingular if F is a P_0-function and the submatrix $[F'(x)]_{I_f I_f}$ is nonsingular. Particularly, if $I_f = \emptyset$, the Jacobian matrix of Ψ_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 Ψ_p at any $x \in \mathbb{R}^n$ is given by

$$\nabla_x \Psi_p(x, \varepsilon)^T = D_a(x, \varepsilon) + D_b(x, \varepsilon)F'(x)$$

where $D_a(x, \varepsilon)$ and $D_b(x, \varepsilon)$ are $n \times n$ diagonal matrices whose diagonal elements $(D_a)_{ii}(x, \varepsilon)$ and $(D_b)_{ii}(x, \varepsilon)$ are negative for $i \in I_l \cup I_u \cup I_{lu}$, and $(D_a)_{ii}(x, \varepsilon) = 0$, $(D_b)_{ii}(x, \varepsilon) = -1$ for $i \in I_f$. Now suppose that $\nabla_x \Psi_p(x, \varepsilon)^T z = 0$. Then,

$$z_i = -\frac{(D_b)_{ii}(x, \varepsilon)}{(D_a)_{ii}(x, \varepsilon)}\left(F'(x)z\right)_i, \quad \text{for } i \in I_l \cup I_u \cup I_{lu} \tag{27}$$

and

$$\left(F'(x)z\right)_i = 0, \quad \text{for } i \in I_f. \tag{28}$$

Since F is a continuously differentiable P_0-function, $F'(x)$ is a P_0-matrix. From Lemma 2.1, we get $z_i = 0$ for $i \in I_l \cup I_u \cup I_{lu}$. Substituting this into (28), we obtain

$$[F'(x)_{I_f I_f}]z_{I_f} = 0,$$

where z_{I_f} is a vector consisting of z_i with $i \in I_f$. This along with the nonsingularity of $[F'(x)]_{I_f I_f}$ 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 4.1. We want to point out when $p \to +\infty$, the diagonal elements $(D_a)_{ii}(x, \varepsilon)$ and $(D_b)_{ii}(x, \varepsilon)$ for $i \in I_l \cup I_u \cup I_{lu}$ will tend to 0, though $(D_a)_{ii}(x, \varepsilon) + (D_b)_{ii}(x, \varepsilon) < 0$. This implies that for a larger p the nonsingularity of $\nabla\Psi_p(x, \varepsilon)$ actually requires stronger conditions than those given by Lemma 4.1.

By Lemma 4.1 and Lemma 3.1 of [19], Algorithm 4.1 is well-defined under the conditions that F is a P_0 function and $[F'(x)]_{I_f I_f}$ is nonsingular. The following lemma provides a condition to guarantee that the merit function $\Theta_p(x)$ has bounded level sets.

Lemma 4.2. *The level sets* $\mathcal{L}(\gamma) := \left\{x \in \mathbb{R}^n \mid \|\Phi_p(x)\| \le \gamma\right\}$ *are bounded for all* $\gamma > 0$ *if one of the following two conditions is satisfied:*

(a) *l and u are both bounded.*
(b) *F is a uniform P-function.*

Proof. Under the condition (a), we have $\{1, 2, \ldots, n\} = I_{lu}$. The result is clear by the definition of Φ_p and Lemma 2.2(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 $\{x^k\} \subseteq \mathcal{L}(\gamma)$ such that $\|x^k\| \to \infty$. Define the index set

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

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

$$\begin{aligned} \mu\|x^k - y^k\|^2 &\le \max_{1\le i\le n}(x_i^k - y_i^k)(F_i(x^k) - F_i(y^k)) \\ &= \max_{i\in J}(x_i^k)(F_i(x^k) - F_i(y^k)) \\ &\le |x_{j_0}^k||F_{j_0}(x^k) - F_{j_0}(y^k)| \end{aligned}$$

where j_0 is an index from $\{1, 2, \ldots, n\}$ for which the maximum is attained, and without loss of generality it is assumed to be independent of k. Clearly, $j_0 \in J$, which means that $\{x_{j_0}^k\}$ is unbounded. Consequently, there exists a subsequence, assumed to be $\{x_{j_0}^k\}$ without loss of generality, such that $|x_{j_0}^k| \to \infty$. Notice that

$$\|x^k - y^k\|^2 \ge |x_{j_0}^k - y_{j_0}^k|^2 = |x_{j_0}^k|^2 \quad \text{for each } k.$$

Therefore, $\mu|x_{j_0}^k|^2 \le |x_{j_0}^k||F_{j_0}(x^k) - F_{j_0}(y^k)|$ and

$$\mu|x_{j_0}^k| \le |F_{j_0}(x^k) - F_{j_0}(y^k)| \le |F_{j_0}(x^k)| + |F_{j_0}(y^k)|,$$

which in turn implies $|F_{j_0}(x^k)| \to \infty$ as $|x_{j_0}^k| \to \infty$. Thus, we prove that

$$|x_{j_0}^k| \to +\infty \quad \text{and} \quad |F_{j_0}(x^k)| \to +\infty. \tag{29}$$

On the other hand, we notice that (29) implies that

$$|x_{j_0}^k - l_i| \to +\infty \quad \text{and} \quad |F_{j_0}(x^k)| \to +\infty.$$

Combining the last two equations with Lemma 2.2(d), we have $|\Phi_{p,j_0}(x^k)| \to +\infty$ from the definition of Φ_p. This contradicts the fact that $\{x^k\} \subseteq \mathcal{L}(\gamma)$. $\square$

Using Lemmas 4.1 and 4.2 and following the same arguments as in [19], we have the following global and local convergence results.

Theorem 4.1. *Suppose that F is a uniform P-function. Then the iteration sequence $\{x^k\}$ generated by Algorithm 4.1 is well defined and converges to the unique solution x^* of the MCP (1) superlinearly. Furthermore, if F' is locally Lipschitz continuous around x^*, then the convergence rate is quadratic.*

5. Numerical experiments

We implemented Algorithm 4.1 in MATLAB 7.0 for solving the MCPLIB test problem collection [11]. The actual implementation is same as the description of Algorithm 4.1 except that in Step 3 we choose ε_{k+1} satisfying (25) only whenever the inequality (24) holds. Although the condition in (26) is crucial for the superlinear convergence analysis of Algorithm 4.1, numerical results reported in Table 1indicate that Algorithm 4.1 without (26) seems to possess the superlinear convergence.

All experiments were done with a PC of Intel Pentium Dual CPU E2200 and 2047MB memory. The parameters of Algorithm 4.1 were chosen as follows:

$$\rho = 0.5, \qquad \sigma = 10^{-2}, \qquad \alpha = 0.5, \qquad \eta = 0.01. \tag{30}$$

Table 1
Numerical results for the MCPLIB problems with different p.

Problem	$p = 1.001$			$p = 1.1$			$p = 2$			$p = 1000$		
	It	NF	$\Vert\Phi_p(x^f)\Vert$	It	NF	$\Vert\Phi_p(x^f)\Vert$	It	NF	$\Vert\Phi_p(x^f)\Vert$	It	NF	$\Vert\Phi_p(x^f)\Vert$
badfree	8	39	6.23e−11	41	607	1.57e−11	56	626	2.87e−12	53	934	1.22e−12
bertsekas(1)	35	101	8.30e−11	22	58	6.20e−12	27	122	2.21e−11	–	–	–
bertsekas(2)	45	154	8.53e−11	29	97	3.05e−11	30	137	2.33e−11	–	–	–
bertsekas(3)	60	212	8.95e−11	29	87	3.56e−11	28	146	1.07e−14	–	–	–
bert-oc	61	273	9.49e−11	15	20	4.44e−11	18	40	2.40e−13	8	20	3.08e−14
bratu	61	283	8.23e−12	16	21	1.94e−12	10	15	1.69e−14	–	–	–
billups	–	–	–	–	–	–	241	5759	0	33	715	1.00e−12
choi	14	38	4.51e−11	12	13	5.43e−13	6	7	8.30e−11	6	7	2.14e−14
colvdual(1)	–	–	–	–	–	–	–	–	–	–	–	–
colvdual(2)	51	371	3.07e−12	–	–	–	–	–	–	–	–	–
colvnlp(1)	123	2305	7.49e−11	23	53	2.02e−11	19	41	3.91e−11	–	–	–
colvnlp(2)	14	23	5.54e−12	16	27	4.38e−11	19	42	6.86e−12	–	–	–
cycle	24	50	7.68e−11	8	12	6.44e−11	4	6	0	4	6	0
degen	5	6	1.77e−13	8	9	4.35e−12	6	7	1.11e−16	3	4	0
duopoly	–	–	–	–	–	–	–	–	–	–	–	–
ehl-k40	–	–	–	–	–	–	–	–	–	–	–	–
ehl-k60	22	45	2.36e−12	25	54	1.28e−12	21	43	1.96e−12	–	–	–
ehl-k80	29	76	3.41e−12	25	49	8.83e−12	27	71	5.21e−11	56	334	3.12e−12
ehl-kost	26	59	1.85e−11	29	61	1.03e−11	34	97	4.47e−12	–	–	–
electric	–	–	–	–	–	–	–	–	–	–	–	–
explcp	31	117	1.28e−13	31	96	3.68e−12	20	47	0	19	54	0
forcebsm	–	–	-	53	281	8.35e−11	–	–	–	–	–	–
forceda	–	–	–	–	–	–	–	–	–	–	–	–
freebert(1)	61	418	6.86e−11	27	105	1.07e−12	24	115	5.67e−12	–	–	–
freebert(2)	63	418	5.77e−11	68	336	9.04e−11	24	108	8.70e−12	–	–	–
freebert(3)	16	35	1.55e−13	17	33	1.19e−12	19	69	1.28e−14	–	–	–
freebert(4)	18	75	5.01e−11	26	97	1.24e−12	23	126	1.12e−14	–	–	–
freebert(5)	62	386	5.81e−11	70	346	3.19e−12	11	21	1.01e−14	–	–	–
freebert(6)	21	58	2.56e−11	20	52	1.90e−11	18	72	1.23e−14	–	–	–
gafni(1)	24	75	6.75e−11	16	30	2.54e−13	11	22	2.10e−15	17	64	3.83e−15
gafni(2)	28	70	5.66e−11	13	26	6.51e−11	14	35	5.55e−15	13	45	1.42e−15
gafni(3)	28	67	6.71e−11	15	32	3.79e−11	15	38	3.76e−15	15	53	3.56e−15
games	11	17	8.87e−12	13	23	2.76e−11	–	–	–	–	–	–
hanskoop(1)	33	60	5.15e−13	30	57	4.96e−11	28	55	4.19e−16	–	–	–
hanskoop(2)	32	59	3.47e−12	14	21	4.32e−12	11	16	1.51e−14	–	–	–
hanskoop(3)	34	61	8.52e−13	32	59	2.31e−12	47	266	1.89e−14	2	4	6.80e−17
hanskoop(4)	34	61	7.80e−12	32	59	2.14e−12	30	57	9.90e−14	–	–	–
hanskoop(5)	44	116	6.83e−11	36	74	9.42e−11	96	743	2.36e−16	–	–	–
hydroc06	9	15	4.06e−12	9	13	2.93e−12	5	7	4.15e−12	5	7	1.06e−12
hydroc20	16	31	1.34e−12	–	–	–	9	12	5.04e−11	9	11	9.77e−14
jel	7	9	9.58e−11	11	13	8.52e−13	23	88	4.56e−12	10	19	1.59e−14
josephy(1)	24	47	6.88e−11	10	12	5.93e−11	9	13	1.02e−12	12	24	1.78e−15
josephy(2)	25	51	6.27e−11	10	15	3.41e−11	8	13	0	8	16	1.07e−13
josephy(3)	28	55	7.01e−11	15	25	8.10e−12	–	–	–	13	25	1.78e−15
josephy(4)	6	7	4.86e−11	8	9	4.42e−12	5	6	1.48e−11	4	5	1.83e−15
josephy(5)	7	8	9.31e−14	9	10	9.19e−13	6	7	8.22e−14	4	5	3.24e−12
josephy(6)	30	85	9.19e−11	14	24	4.14e−12	7	10	0	7	11	0
kojshin(1)	26	52	5.37e−11	11	14	9.23e−12	11	20	5.28e−14	10	21	5.35e−13
kojshin(2)	28	64	7.59e−11	16	30	1.74e−11	11	24	2.17e−11	8	15	1.78e−15
kojshin(3)	31	71	6.17e−11	19	35	3.38e−13	12	16	5.32e−13	11	22	5.35e−13
kojshin(4)	29	61	9.97e−11	20	37	7.73e−12	11	16	7.09e−13	10	16	8.88e−16
kojshin(5)	7	8	4.74e−13	8	9	1.26e−11	11	17	2.44e−13	6	8	0
kojshin(6)	–	–	–	14	22	2.86e−12	9	15	1.98e−11	7	12	8.88e−16
lincont	56	447	2.01e−10	33	119	2.37e−10	–	–	–	–	–	–
mathinum(1)	5	6	2.75e−12	7	8	2.77e−11	5	6	4.35e−14	6	12	0
mathinum(2)	7	8	1.41e−13	9	10	4.43e−12	6	7	1.83e−12	6	7	0
mathinum(3)	7	9	6.36e−12	10	12	1.31e−12	7	9	0	7	12	0
mathinum(4)	7	8	1.80e−11	10	11	1.12e−12	8	9	4.44e−16	7	8	8.88e−16
mathisum(1)	7	9	1.19e−12	11	16	7.39e−12	9	13	0	–	–	–
mathisum(2)	8	12	1.31e−11	12	18	1.51e−12	7	9	4.45e−13	5	8	0
mathisum(3)	7	9	2.54e−11	9	10	1.15e−12	5	6	1.48e−11	4	6	0
mathisum(4)	10	20	1.34e−11	13	22	7.80e−13	8	10	4.44e−16	6	9	0
methan08	7	8	5.80e−13	8	9	3.56e−12	4	5	1.17e−11	4	5	7.87e−12
nash(1)	8	9	5.00e−13	10	11	9.73e−11	8	9	4.50e−14	9	16	6.31e−14
nash(2)	14	32	6.14e−12	15	28	2.65e−12	13	28	5.54e−14	14	42	8.20e−14
ne-hard	35	92	4.10e−11	35	92	4.10e−11	35	92	4.10e−11	35	92	4.10e−11
obstacle	47	226	2.63e−11	14	18	7.24e−12	9	12	3.37e−15	8	16	3.25e−15
opt-cont	52	214	2.64e−11	14	18	3.91e−11	9	10	3.93e−15	11	16	2.63e−15
opt-cont31	49	168	3.18e−11	15	20	4.11e−12	10	13	6.09e−15	12	29	5.87e−15

Table 1 (*continued*)

Problem	$p = 1.001$			$p = 1.1$			$p = 2$			$p = 1000$		
	It	NF	$\|\Phi_p(x^f)\|$	It	NF	$\|\Phi_p(x^f)\|$	It	NF	$\|\Phi_p(x^f)\|$	It	NF	$\|\Phi_p(x^f)\|$
opt-cont127	61	243	1.16e−11	21	31	1.27e−12	16	36	1.52e−11	30	136	1.08e−14
opt-cont255	73	277	1.95e−11	16	29	1.79e−11	15	32	1.54e−14	39	241	1.62e−14
pgvon106(1)	100	584	5.40e−11	–	–	–	–	–	–	–	–	–
pgvon106(2)	80	364	9.81e−11	42	107	7.66e−11	–	–	–	–	–	–
pgvon106(3)	36	70	1.61e−13	287	2334	8.04e−11	–	–	–	–	–	–
pies	59	676	1.24e−11	17	31	4.60e−12	16	46	7.29e−12	–	–	–
powell(1)	–	–	–	13	30	9.43e−11	–	–	–	–	–	–
powell(2)	120	1285	8.51e−11	17	48	8.67e−11	–	–	–	–	–	–
powell(3)	17	86	9.36e−11	20	46	6.03e−11	12	33	8.99e−12	–	–	–
powell(4)	63	883	2.65e−13	–	–	–	–	–	–	–	–	–
powell-mcp(1)	6	7	6.46e−12	6	7	6.46e−12	6	7	6.46e−12	6	7	6.46e−12
powell-mcp(2)	7	8	2.17e−12	7	8	2.17e−12	7	8	2.17e−12	7	8	2.17e−12
powell-mcp(3)	9	10	5.43e−15	9	10	5.43e−15	9	10	5.43e−15	9	10	5.43e−15
powell-mcp(4)	8	9	1.98e−14	8	9	1.98e−14	8	9	1.98e−14	8	9	1.98e−14
qp	7	9	6.51e−14	8	9	3.66e−12	6	7	8.88e−16	3	4	0
scarfanum(1)	37	90	7.69e−14	39	92	8.97e−12	40	76	5.13e−11	–	–	–
scarfanum(2)	36	89	2.73e−13	38	91	8.35e−12	38	71	3.20e−15	–	–	–
scarfanum(3)	–	–	–	–	–	–	118	676	2.28e−15	39	124	1.67e−12
scarfasum(1)	52	79	2.28e−13	58	85	1.43e−12	14	21	9.49e−15	–	–	–
scarfasum(2)	51	78	2.77e−11	60	88	9.08e−11	14	20	3.16e−15	–	–	–
scarfasum(3)	34	107	6.32e−11	13	17	4.18e−12	40	70	5.71e−13	–	–	–
scarfbsum(1)	63	251	7.90e−12	50	174	5.52e−12	34	168	5.45e−11	–	–	–
scarfbsum(2)	28	72	1.35e−12	89	370	2.28e−12	24	142	5.52e−11	–	–	–
simple-red	12	13	7.94e−13	14	15	3.29e−12	12	13	1.87e−15	12	13	1.15e−15
simple-ex	196	3321	3.50e−13	–	–	–	–	–	–	–	–	–
sppe(1)	13	18	5.92e−12	11	13	5.99e−11	6	7	1.11e−12	–	–	–
sppe(2)	7	8	3.71e−12	9	10	1.49e−12	7	9	9.98e−14	–	–	–
shubik	49	207	1.87e−11	64	225	9.12e−11	–	–	–	–	–	–
tinloi	23	83	2.19e−11	17	33	3.43e−11	14	26	3.77e−15	19	93	2.32e−15
tobin(1)	16	70	1.91e−12	12	21	1.14e−11	9	12	3.48e−14	16	28	8.27e−13
tobin(2)	13	23	2.47e−11	13	18	1.49e−11	9	12	2.04e−14	11	16	1.25e−12
trafelas	50	233	9.96e−11	27	62	6.22e−11	–	–	–	–	–	–

We started Algorithm 4.1 with the standard starting point provided by the MCPLIB collection, and terminated the iteration if one of the following conditions is satisfied

$$\|\Phi_p(x)\| \le 10^{-10} \quad \text{or} \quad k > 300.$$

The numerical results corresponding to $p = 1.001$, $p = 1.1$, $p = 2$ and $p = 1000$, respectively, are summarized in Table 1. In these tables, the first column gives the names of problems, and the number after each problem specifies which starting point from the library is used; **Iter** denotes the number of iterations; **NF** means the number of function evaluations for the mapping F, and $\|\Phi_p(x^f)\|$ column denotes the values of $\|\Phi_p(x)\|$ at the final iterate $x = x^f$.

Table 1 show that Algorithm 4.1 based on the smoothing approximation $\Psi_p(x, \varepsilon)$ with $p \in (1, 2]$ was able to solve almost all MCPLIB test problems, including a number of examples known to be very bad. Among 55 test problems, there are 7 problems failure for $p = 1.001$, which are **billups, duopoly, ehl-k40, electric, forcebsm, forceda, simple-ex**; there are 8 problems failure for $p = 1.1$, which are **billups, colvdual, duopoly, ehl-k40, electric, forceda, hydroc20, simple-ex**; and there 12 problems failure for $p = 2$, which are **billups, colvdual, duopoly, ehl-k40, electric, forcebsm, forceda, lincont, simple-ex, games, shubik, trafelas**. It is known that the problems such as "duopoly, forcebsm, electric, shubik" are also very difficult for other Newton-type methods in the literature. Unlike the least-square semismooth Newton method based on ϕ_p (see [23]), Algorithm 4.1 with $p = 1000$ fails for most of test problems due to the singularity of $\nabla_x \Psi_p(x^k, \varepsilon^k)$. This also coincides with the observations in Remark 4.1.

From Table 1, we see that Algorithm 4.1 with a smaller p, to say $p < 2$, has better robustness than a larger p (>2), but when p is closer to 1, Algorithm 4.1 generally requires more iterations. Therefore, we conclude that Algorithm 4.1 with p chosen from $[1.1, 2]$ should be desirable. In addition, we want to point out that the value of α will give an influence on numerical performance of Algorithm 4.1, and the favorable α should be chosen from the interval $[0.3, 0.7]$.

6. Conclusions

In this paper, we have studied the smoothing Newton method [19] based on the smooth approximation ψ_p of the generalized FB function. The smooth operator Ψ_p is shown to possess the Jacobian consistency property, which implies the fast convergence of this smoothing algorithm. Numerical experiments indicate that the algorithm with $p \in (1, 2]$ has better numerical performance than the one with $p > 2$, and it has better robustness when p is closer to 1. Further numerical experiments are needed to check whether imposing the condition (26) on ε_{k+1} may improve the performance of Algorithm 4.1.

Acknowledgements

The first author's work is partially supported by National Science Council of Taiwan. The second author's work is supported by National Young Natural Science Foundation (No. 10901058) and Guangdong Natural Science Foundation (No. 9251802902000001).

Appendix. The proof of Lemma 3.2

From Eq. (15), clearly, the index set $\{1, 2, \ldots, n\}$ can be partitioned as

$$I_f \cup \beta_1(x) \cup \bar{\beta}_1(x) \cup \beta_2(x) \cup \bar{\beta}_2(x) \cup \beta_3(x) \cup \beta_4(x) \cup \bar{\beta}_4(x).$$

In view of this, we proceed the proof by the following several cases.

Case 1: $i \in I_f$. From Proposition 3.2(a4) and Proposition 3.3(d), we have

$$\nabla_x \Psi_{p,i}(x,\varepsilon)^T = -F_i'(x) \quad \text{and} \quad \partial_C \Phi_{p,i}(x) = -F_i'(x),$$

which implies that

$$\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) = 0 \quad \text{for all } \varepsilon > 0. \tag{31}$$

Case 2: $i \in \beta_1(x) \cup \beta_2(x)$. From Proposition 3.2(a1) and (a2), it follows that

$$\nabla_x \Psi_{p,i}(x,\varepsilon)^T = -e_i^T - F_i'(x).$$

In addition, by Proposition 3.3(a) and (b), we have $\nabla_x \Psi_{p,i}(x,\varepsilon)^T \in \partial \Phi_{p,i}(x)$ since

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

Therefore,

$$\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) = 0 \quad \text{for all } \varepsilon > 0. \tag{32}$$

Case 3: $i \in \beta_3(x)$. Under this case, $x_i - l_i = 0$, $\phi_p(u_i - x_i, -F_i(x)) = 0$, $u_i - x_i > 0$ and $F_i(x) = 0$. Hence, $c_i(x) = 0$ and $d_i(x) = 1$. From Proposition 3.3(c), it follows that

$$\partial_C \Phi_{p,i}(x) = \{a_i(x) e_i^T + b_i(x) F_i'(x)\} \tag{33}$$

with

$$(a_i(x), b_i(x)) \in \left\{(\xi - 1, \zeta - 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1\right\}.$$

On the other hand, since $a_i(x,\varepsilon) = -1$, $d_i(x,\varepsilon) = 1$ and

$$b_i(x,\varepsilon) = \frac{|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}}{\left(|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^p + \varepsilon^p\right)^{\frac{p-1}{p}}} - 1,$$

$$c_i(x,\varepsilon) = 1 - \frac{|u_i - x_i|^{p-1}}{(|u_i - x_i|^p + \varepsilon^p)^{(p-1)/p}},$$

from Proposition 3.2(a3) it follows that

$$\nabla_x \Psi_{p,i}(x,\varepsilon)^T = (-1 + b_i(x,\varepsilon) c_i(x,\varepsilon)) e_i^T + b_i(x,\varepsilon) F_i'(x). \tag{34}$$

Taking

$$\xi = 0 \quad \text{and} \quad \zeta = \frac{|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}}{\left(|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^p + \varepsilon^p\right)^{\frac{p-1}{p}}},$$

we can verify that $|\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1$, and consequently $-e_i^T + b_i(x,\varepsilon) F_i'(x) \in \partial_C \Phi_{p,i}(x)$. Using the definition of $\varepsilon_0(x,\delta)$, it is easy to verify that, for all $\varepsilon \le \varepsilon_0(x,\delta)$,

$$\left\| \nabla_x \Psi_{p,i}(x,\varepsilon)^T - \left(-e_i^T + b_i(x,\varepsilon) F_i'(x)\right) \right\| = \|b_i(x,\varepsilon) c_i(x,\varepsilon) e_i^T\| \le |c_i(x,\varepsilon)| \le \frac{\delta}{\sqrt{n}}.$$

Therefore, for all $0 < \varepsilon \le \varepsilon_0(x,\delta)$,

$$\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) \le \frac{\delta}{\sqrt{n}}. \tag{35}$$

Case 4: $i \in \bar{\beta}_1(x)$. From Proposition 3.3(a) and Proposition 3.2(a1), it follows that

$$\begin{aligned}
\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) &= \left\| \nabla_x \Psi_{p,i}(x,\varepsilon)^T - \nabla \Phi_{p,i}(x)^T \right\| \\
&= \left(\frac{1}{\|(x_i - l_i, F_i(x))\|_p^{p-1}} - \frac{1}{\|(x_i - l_i, F_i(x), \varepsilon)\|_p^{p-1}} \right) \\
&\quad \left\| \operatorname{sign}(x_i - l_i)|x_i - l_i|^{p-1} e_i + \operatorname{sign}(F_i(x))|F_i(x)|^{p-1} \nabla F_i(x) \right\| \\
&\le \left(\alpha_1(x)^{\frac{1-p}{p}} - (\alpha_1(x) + \varepsilon^p)^{\frac{1-p}{p}} \right) \gamma_1(x) \\
&\le \left(\alpha(x)^{\frac{1-p}{p}} - (\alpha(x) + \varepsilon^p)^{\frac{1-p}{p}} \right) \gamma(x) \\
&= \frac{(\alpha(x) + \varepsilon^p)^{\frac{p-1}{p}} - \alpha(x)^{\frac{p-1}{p}}}{[\alpha(x)(\alpha(x) + \varepsilon^p)]^{\frac{p-1}{p}}} \gamma(x) \\
&\le \frac{\varepsilon^{p-1}}{[\alpha(x)(\alpha(x) + \varepsilon^p)]^{\frac{p-1}{p}}} \gamma(x) \qquad (36)
\end{aligned}$$

where the inequalities are using Lemma 2.4 and the definition of $\alpha(x)$ and $\gamma(x)$. Now using Eq. (36), it is not hard to verify that for all $0 < \varepsilon \le \varepsilon_1(x, \delta)$

$$\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) \le \frac{\delta}{\sqrt{n}}. \qquad (37)$$

Indeed, if $\gamma(x) = 0$, this inequality obviously holds for all $\varepsilon > 0$. Suppose that $\gamma(x) > 0$. Then, a simple calculation shows that

$$\frac{\varepsilon^{p-1}\gamma(x)}{[\alpha(x)(\alpha(x) + \varepsilon^p)]^{\frac{p-1}{p}}} \le \frac{\delta}{\sqrt{n}} \iff \alpha(x)^2 \ge \varepsilon^p \left(\left(\frac{\sqrt{n}\gamma(x)}{\delta} \right)^{p/(p-1)} - \alpha(x) \right).$$

Clearly, the inequality on the right hand side holds for all $0 < \varepsilon \le \varepsilon_1(x, \delta)$. Consequently, the result in (37) follows from the above equivalence and (36).

Case 5: $i \in \bar{\beta}_2(x)$. From Proposition 3.3(b) and Proposition 3.2(a2), it follows that

$$\begin{aligned}
&\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) \\
&= \left(\frac{1}{\|(u_i - x_i, F_i(x))\|_p^{p-1}} - \frac{1}{\|(u_i - x_i, F_i(x), \varepsilon)\|_p^{p-1}} \right) \left\| \operatorname{sign}(F_i(x))|F_i(x)|^{p-1} \nabla F_i(x) - \operatorname{sign}(u_i - x_i)|u_i - x_i|^{p-1} e_i \right\| \\
&\le \left(\alpha_2(x)^{\frac{1-p}{p}} - (\alpha_2(x) + \varepsilon^p)^{\frac{1-p}{p}} \right) \gamma_2(x) \\
&\le \left(\alpha(x)^{\frac{1-p}{p}} - (\alpha(x) + \varepsilon^p)^{\frac{1-p}{p}} \right) \gamma(x).
\end{aligned}$$

where the inequalities are using Lemma 2.4 and the definition of $\alpha(x)$ and $\gamma(x)$. Using the same arguments as Case 4, we can prove that for all $0 < \varepsilon \le \varepsilon_1(x, \delta)$,

$$\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) \le \frac{\delta}{\sqrt{n}}. \qquad (38)$$

Case 6: $i \in \beta_4(x)$. Since $(u_i - x_i, -F_i(x)) = (0, 0)$, we necessarily have

$$x_i - l_i > 0, \qquad \phi_p(u_i - x_i, -F_i(x)) = 0 \quad \text{and} \quad \psi_p(u_i - x_i, -F_i(x), \varepsilon) = \varepsilon, \qquad (39)$$

which in turn implies $a_i(x) = 0$ and $b_i(x) = -1$. By Proposition 3.3(c),

$$\partial_C \Phi_{p,i}(x) = \left\{ -c_i(x) e_i^T - d_i(x) F_i'(x) \right\}$$

with

$$(c_i(x), d_i(x)) \in \left\{ (\xi + 1, \zeta + 1) \in \mathbb{R}^2 \mid |\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \le 1 \right\}.$$

In addition, we notice that under this case $c_i(x, \varepsilon) = 1$, $d_i(x, \varepsilon) = 1$ and

$$a_i(x, \varepsilon) = \frac{|x_i - l_i|^{p-1}}{\left(\sqrt[p]{|x_i - l_i|^p + 2\varepsilon^p}\right)^{p-1}} - 1, \qquad b_i(x, \varepsilon) = \frac{\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i - l_i|^p + 2\varepsilon^p}\right)^{p-1}} - 1.$$

Therefore, from Proposition 3.2(a3), it follows that

$$\begin{aligned}\nabla_x \Psi_{p,i}(x,\varepsilon)^T &= (a_i(x,\varepsilon)+b_i(x,\varepsilon))e_i^T + b_i(x,\varepsilon)d_i(x,\varepsilon)F_i'(x)\\ &= -\left(1-\frac{|x_i-l_i|^{p-1}+\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}+1\right)e_i^T-\left(-\frac{\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}+1\right)F_i'(x).\end{aligned}$$

We next want to prove that for any $0<\varepsilon\le\varepsilon_2(x,\delta)$,

$$\left|1-\frac{|x_i-l_i|^{p-1}+\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}\right|^{\frac{p}{p-1}}+\left|\frac{\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}\right|^{\frac{p}{p-1}}\le 1, \tag{40}$$

and consequently $\nabla_x\Psi_{p,i}(x,\varepsilon)^T\in\partial_C\Phi_{p,i}(x)$. It is easily verified that the function

$$h_1(\varepsilon)=\frac{|x_i-l_i|^{p-1}+\varepsilon^{p-1}}{(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\,)^{p-1}}$$

is increasing on $[0,\varepsilon_2(x,\delta)]$. Since $h_1(0)=1$, we have $h_1(\varepsilon)\ge 1$ on $[0,\varepsilon_2(x,\delta)]$. Therefore,

$$\begin{aligned}&\left|1-\frac{|x_i-l_i|^{p-1}+\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}\right|^{\frac{p}{p-1}}+\left|\frac{\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}\right|^{\frac{p}{p-1}}\\ &=\left(\frac{|x_i-l_i|^{p-1}+\varepsilon^{p-1}}{\left(\sqrt[p]{|x_i-l_i|^p+2\varepsilon^p}\right)^{p-1}}-1\right)^{\frac{p}{p-1}}+\frac{\varepsilon^p}{|x_i-l_i|^p+2\varepsilon^p}\\ &:= h_2(\varepsilon).\end{aligned}$$

We can verify that $h_2(\varepsilon)$ is strictly increasing on $[0,\varepsilon_2(x,\delta)]$ and

$$\begin{aligned}h_2(\varepsilon_2(x,\delta))=h_2(|x_i-l_i|/2)&\le\left[\left(1+\frac{1}{2^{p-1}}\right)^{1/p}-1\right]^{\frac{p}{p-1}}+1/2\\ &\le\left[1+(1/2)^{\frac{p-1}{p}}-1\right]^{\frac{p}{p-1}}+1/2\le 1,\end{aligned}$$

where the second inequality is since $(1+t)^{1/p}\le 1+t^{1/p}$ for $t\ge 0$. The last two equations imply that the inequality (40) holds. Consequently,

$$\operatorname{dist}\left(\nabla_x\Psi_{p,i}(x,\varepsilon)^T,\partial_C\Phi_{p,i}(x)\right)\le\frac{\delta}{\sqrt{n}}\quad\text{for all } 0<\varepsilon\le\varepsilon_2(x,\delta). \tag{41}$$

Case 7: $i\in\bar{\beta}_4(x)$. Since $(u_i-x_i,-F_i(x))\ne(0,0)$, by Proposition 3.3(c) and Proposition 3.2(a3),

$$\begin{aligned}&\operatorname{dist}\left(\nabla_x\Psi_{p,i}(x,\varepsilon)^T,\partial_C\Phi_{p,i}(x)\right)\\ &=\|(a_i(x,\varepsilon)-a_i(x))\,e_i+(b_i(x,\varepsilon)c_i(x,\varepsilon)-b_i(x)c_i(x))\,e_i+(b_i(x,\varepsilon)d_i(x,\varepsilon)-b_i(x)c_i(x))\,\nabla F_i(x)\|\\ &=\|(a_i(x,\varepsilon)-a_i(x))\,e_i+(b_i(x,\varepsilon)-b_i(x))\,c_i(x)e_i+(b_i(x,\varepsilon)-b_i(x))\,d_i(x)\nabla F_i(x)\\ &\quad+\,b_i(x,\varepsilon)\,(c_i(x,\varepsilon)-c_i(x))\,e_i+b_i(x,\varepsilon)\,(d_i(x,\varepsilon)-d_i(x))\,\nabla F_i(x)\|\,.\end{aligned} \tag{42}$$

In what follows, we will successively estimate the value of $|a_i(x,\varepsilon)-a_i(x)|$, $|b_i(x,\varepsilon)-b_i(x)|$, $|c_i(x,\varepsilon)-c_i(x)|$ and $|d_i(x,\varepsilon)-d_i(x)|$ for $0<\varepsilon<\varepsilon_3(x,\delta)$. Note that $\psi_p(u_i-x_i,-F_i(x),\varepsilon)$ and $\phi_p(u_i-x_i,-F_i(x))$ have the same sign for all $0<\varepsilon\le\varepsilon_3(x,\delta)$. Indeed, if $\phi_p(u_i-x_i,-F_i(x))\ge 0$, then $\psi_p(u_i-x_i,-F_i(x),\varepsilon)>0$ clearly holds. Otherwise, since

$$\begin{aligned}\psi_p(u_i-x_i,-F_i(x),\varepsilon)<0 &\iff |u_i-x_i|^p+|F_i(x)|^p+\varepsilon^p<(u_i-x_i-F_i(x))^p,\\ &\iff \varepsilon<\left((u_i-x_i-F_i(x))^p-|u_i-x_i|^p-|F_i(x)|^p\right)^{1/p},\end{aligned}$$

the definition of $\varepsilon_3(x,\delta)$ implies that $\psi_p(u_i-x_i,-F_i(x),\varepsilon)<0$ for all $0<\varepsilon\le\varepsilon_3(x,\delta)$.

Step 1: To estimate $|a_i(x,\varepsilon)-a_i(x)|$. For $0<\varepsilon\le\varepsilon_3(x,\delta)$, we first estimate

$$r(x,\varepsilon):=\left|\frac{1}{\left\|(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon),\varepsilon)\right\|_p^{p-1}}-\frac{1}{\left\|(x_i-l_i,\phi_p(u_i-x_i,-F_i(x)))\right\|_p^{p-1}}\right|.$$

Let

$$g_1(\varepsilon) := |\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^p - |\phi_p(u_i - x_i, -F_i(x))|^p$$

and

$$\Delta(\varepsilon) := \psi_p(u_i - x_i, -F_i(x), \varepsilon) - \phi_p(u_i - x_i, -F_i(x))$$

for $0 < \varepsilon \le \varepsilon_3(x, \delta)$. If $\phi_p(u_i - x_i, -F_i(x)) \ge 0$, then $\psi_p(u_i - x_i, -F_i(x), \varepsilon) > 0$, and hence $g_1(\varepsilon) > 0$. In addition, applying the mean-value theorem and Lemma 3.1(c), we have,

$$\begin{aligned} g_1(\varepsilon) &= \psi_p(u_i - x_i, -F_i(x), \varepsilon)^p - \phi_p(u_i - x_i, -F_i(x))^p \\ &= p\left[\phi_p(u_i - x_i, -F_i(x)) + t_1\Delta(\varepsilon)\right]^{p-1}\Delta(\varepsilon) \quad \text{for some } t_1 \in (0,1) \\ &\le p\left[\phi_p(u_i - x_i, -F_i(x)) + \varepsilon_3(x,\delta)\right]^{p-1}\varepsilon. \end{aligned} \tag{43}$$

Under this case, taking into account the definition of $\alpha_3(x)$ and $a(x)$, we have

$$\begin{aligned} r(x,\varepsilon) &= \frac{1}{\left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)))\right\|_p^{p-1}} - \frac{1}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{p-1}} \\ &\le \left[\alpha_3(x)^{-\frac{p-1}{p}} - (\alpha_3(x) + g_1(\varepsilon) + \varepsilon^p)^{-\frac{p-1}{p}}\right] \\ &\le \left[\alpha(x)^{-\frac{p-1}{p}} - (\alpha(x) + g_1(\varepsilon) + \varepsilon^p)^{-\frac{p-1}{p}}\right] \\ &= \frac{(\alpha(x) + g_1(\varepsilon) + \varepsilon^p)^{\frac{p-1}{p}} - \alpha(x)^{\frac{p-1}{p}}}{[\alpha(x)(\alpha(x) + g_1(\varepsilon) + \varepsilon^p)]^{\frac{p-1}{p}}} \le \frac{(g_1(\varepsilon) + \varepsilon^p)^{\frac{p-1}{p}}}{\alpha(x)^{\frac{2(p-1)}{p}}} \le M_1(x)\varepsilon^{\frac{p-1}{p}} \end{aligned}$$

where the first three inequalities are due to Lemma 2.4, the last one is by (43), and

$$M_1(x) := \left[\frac{p\left[\phi_p(u_i - x_i, -F_i(x)) + \varepsilon_3(x,\delta)\right]^{p-1} + \varepsilon_3(x,\delta)^{p-1}}{\alpha(x)^{2/p}}\right]^{p-1}. \tag{44}$$

If $\phi_p(u_i - x_i, -F_i(x)) < 0$, then $\psi_p(u_i - x_i, -F_i(x), \varepsilon) < 0$, and hence $g_1(\varepsilon) < 0$. Now,

$$\begin{aligned} r(x,\varepsilon) \le\; & \frac{1}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{p-1}} - \frac{1}{\left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)), \varepsilon)\right\|_p^{p-1}} \\ &+ \left|\frac{1}{\left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)), \varepsilon)\right\|_p^{p-1}} - \frac{1}{\left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)))\right\|_p^{p-1}}\right| \\ \le\; & \frac{\left[|\phi_p(u_i - x_i, -F_i(x))|^{p-1} - |\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}\right]}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{p-1}\left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)), \varepsilon)\right\|_p^{p-1}} \\ &+ \left[\alpha_3(x)^{-\frac{p-1}{p}} - (\alpha_3(x) + \varepsilon^p)^{-\frac{p-1}{p}}\right] \\ \le\; & \frac{\left[|\phi_p(u_i - x_i, -F_i(x))|^{p-1} - |\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1}\right]}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{2p-2}} + \left[\alpha_3(x)^{-\frac{p-1}{p}} - (\alpha_3(x) + \varepsilon^p)^{-\frac{p-1}{p}}\right]. \end{aligned} \tag{45}$$

Notice that

$$\begin{aligned} &|\phi_p(u_i - x_i, -F_i(x))|^{p-1} - |\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1} \\ &= \left[-\phi_p(u_i - x_i, -F_i(x))\right]^{p-1} - \left[-\psi_p(u_i - x_i, -F_i(x), \varepsilon)\right]^{p-1} \\ &= (p-1)\left[-\psi_p(u_i - x_i, -F_i(x), \varepsilon) + t_2\Delta(\varepsilon)\right]^{p-2}\Delta(\varepsilon) \quad \text{for some } t_2 \in (0,1) \\ &\le \begin{cases} (p-1)\left[-\phi_p(u_i - x_i, -F_i(x))\right]^{p-2}\varepsilon & \text{if } p \ge 2; \\ (p-1)\left[-\psi_p(u_i - x_i, -F_i(x), \varepsilon_3(x,\delta))\right]^{p-2}\varepsilon & \text{if } 1 < p < 2, \end{cases} \end{aligned}$$

and

$$\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p \ge \left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon_3(x,\delta)))\right\|_p.$$

Therefore,

$$\begin{aligned}
&\frac{\left[|\phi_p(u_i-x_i,-F_i(x))|^{p-1}-|\psi_p(u_i-x_i,-F_i(x),\varepsilon)|^{p-1}\right]}{\left\|(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon),\varepsilon)\right\|_p^{2p-2}}\\
&\quad\le (p-1)\frac{\left[-\phi_p(u_i-x_i,-F_i(x))\right]^{p-2}\varepsilon+\left[-\psi_p(u_i-x_i,-F_i(x),\varepsilon_3(x,\delta))\right]^{p-2}\varepsilon}{\left\|(x_i-l_i,\psi_p(u_i-x_i,-F_i(x)),\varepsilon_3(x,\delta))\right\|_p^{2p-2}}\\
&\quad\le (p-1)\left(\frac{\left[-\phi_p(u_i-x_i,-F_i(x))\right]^{p-2}}{\left\|(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon_3(x,\delta)))\right\|_p^{p-2}}+1\right)\varepsilon\\
&\quad:= M_2(x)\varepsilon.
\end{aligned}$$

This together with (45) and Lemma 2.4 yields that

$$\begin{aligned}
r(x,\varepsilon) &\le M_2(x)\varepsilon+\frac{(\alpha(x)+\varepsilon^p)^{\frac{p-1}{p}}-\alpha(x)^{\frac{p-1}{p}}}{[\alpha(x)(\alpha(x)+\varepsilon^p)]^{\frac{p-1}{p}}}\\
&\le M_2(x)\varepsilon+\frac{\varepsilon^{p-1}}{\alpha(x)^{\frac{2(p-1)}{p}}}\\
&\le \begin{cases} M_3(x)\varepsilon & \text{if } p\ge 2\\ M_3(x)\varepsilon^{p-1} & \text{if } 1<p<2\end{cases}
\end{aligned}$$

where

$$M_3(x):=\begin{cases} M_2(x)+\dfrac{\varepsilon_3(x,\delta)^{p-2}}{\alpha(x)^{\frac{2p-2}{p}}} & \text{if } p\ge 2;\\ M_2(x)\varepsilon_3(x,\delta)^{2-p}+\alpha(x)^{\frac{2-2p}{p}} & \text{if } 1<p<2.\end{cases} \tag{46}$$

Summing up the above discussions, it then follows that

$$\begin{aligned}
r(x,\varepsilon) &\le \begin{cases} \max\left\{M_1(x),M_3(x)\varepsilon_3(x,\delta)^{1/p}\right\}\varepsilon^{\frac{p-1}{p}} & \text{if } p\ge 2;\\ \max\left\{M_1(x),M_3(x)\varepsilon_3(x,\delta)^{(p+\frac{1}{p}-2)}\right\}\varepsilon^{\frac{p-1}{p}} & \text{if } 1<p<2,\end{cases}\\
&\le M_4(x)\varepsilon^{\frac{p-1}{p}},
\end{aligned}$$

where

$$M_4(x):=\max\left\{M_1(x),M_3(x)\varepsilon_3(x,\delta)^{1/p},M_3(x)\varepsilon_3(x,\delta)^{(p+1/p-2)}\right\}. \tag{47}$$

Consequently,

$$|a_i(x,\varepsilon)-a_i(x)|=r(x,\varepsilon)|x_i-l_i|^{p-1}\le M_4(x)|x_i-l_i|^{p-1}\varepsilon^{\frac{p-1}{p}}.$$

Step 2: To estimate $|b_i(x,\varepsilon)-b_i(x)|$. From the expressions of $b_i(x,\varepsilon)$ and $b_i(x)$,

$$\begin{aligned}
|b_i(x,\varepsilon)-b_i(x)| &= \left|\frac{\mathrm{sign}(\psi_p(u_i-x_i,-F_i(x),\varepsilon))|\psi_p(u_i-x_i,-F_i(x),\varepsilon)|^{p-1}}{\left\|\left(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon),\varepsilon\right)\right\|_p^{p-1}}\right.\\
&\quad-\frac{\mathrm{sign}(\phi_p(u_i-x_i,-F_i(x)))|\phi_p(u_i-x_i,-F_i(x))|^{p-1}}{\left\|\left(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon),\varepsilon\right)\right\|_p^{p-1}}\\
&\quad+\frac{\mathrm{sign}(\phi_p(u_i-x_i,-F_i(x)))|\phi_p(u_i-x_i,-F_i(x))|^{p-1}}{\left\|\left(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon),\varepsilon\right)\right\|_p^{p-1}}\\
&\quad\left.-\frac{\mathrm{sign}(\phi_p(u_i-x_i,-F_i(x)))|\phi_p(u_i-x_i,-F_i(x))|^{p-1}}{\left\|\left(x_i-l_i,\phi_p(u_i-x_i,-F_i(x))\right)\right\|_p^{p-1}}\right|\\
&\le \frac{g_2(\varepsilon)}{\left\|(x_i-l_i,\psi_p(u_i-x_i,-F_i(x),\varepsilon),\varepsilon)\right\|_p^{p-1}}+r(x,\varepsilon)|\phi_p(u_i-x_i,-F_i(x))|^{p-1},
\end{aligned} \tag{48}$$

where $r(x,\varepsilon)$ is same as above, and $g_2(\varepsilon)$ is defined by

$$\begin{aligned}
g_2(\varepsilon):=&\left|\mathrm{sign}(\psi_p(u_i-x_i,-F_i(x),\varepsilon))|\psi_p(u_i-x_i,-F_i(x),\varepsilon)|^{p-1}\right.\\
&\left.-\ \mathrm{sign}(\phi_p(u_i-x_i,-F_i(x)))|\phi_p(u_i-x_i,-F_i(x))|^{p-1}\right|.
\end{aligned}$$

If $\phi_p(u_i - x_i, -F_i(x)) \geq 0$, then $\psi_p(u_i - x_i, -F_i(x), \varepsilon) > 0$, and therefore

$$\begin{aligned} g_2(\varepsilon) &= \psi_p(u_i - x_i, -F_i(x), \varepsilon)^{p-1} - \phi_p(u_i - x_i, -F_i(x))^{p-1} \\ &= (p-1)\left[\phi_p(u_i - x_i, -F_i(x)) + t_3\Delta(\varepsilon)\right]^{p-2}\Delta(\varepsilon) \quad \text{for some } t_3 \in (0,1) \\ &\leq \begin{cases} (p-1)\left[\phi_p(u_i - x_i, -F_i(x)) + \varepsilon_3(x,\delta)\right]^{p-2}\varepsilon & \text{if } p \geq 2; \\ (p-1)\left[\phi_p(u_i - x_i, -F_i(x))\right]^{p-2}\varepsilon & \text{if } 1 < p < 2. \end{cases} \end{aligned}$$

If $\phi_p(u_i - x_i, -F_i(x)) < 0$, then $\psi_p(u_i - x_i, -F_i(x), \varepsilon) < 0$ and

$$|\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1} < |\phi_p(u_i - x_i, -F_i(x))|^{p-1} \quad \text{for } 0 < \varepsilon \leq \varepsilon_3(x,\delta).$$

Consequently,

$$\begin{aligned} g_2(\varepsilon) &= \left[-\phi_p(u_i - x_i, -F_i(x))\right]^{p-1} - \left[-\psi_p(u_i - x_i, -F_i(x), \varepsilon)\right]^{p-1} \\ &= (p-1)\left[-\psi_p(u_i - x_i, -F_i(x), \varepsilon) + t_4\Delta(\varepsilon)\right]^{p-2}\Delta(\varepsilon) \quad \text{for some } t_4 \in (0,1) \\ &\leq \begin{cases} (p-1)\left[-\phi_p(u_i - x_i, -F_i(x))\right]^{p-2}\varepsilon & \text{if } p \geq 2; \\ (p-1)\left[-\psi_p(u_i - x_i, -F_i(x), \varepsilon_3(x,\delta))\right]^{p-2}\varepsilon & \text{if } 1 < p < 2. \end{cases} \end{aligned}$$

In addition, if $\phi_p(u_i - x_i, -F_i(x)) \geq 0$, then

$$\left\|\left(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon\right)\right\|_p^{p-1} > \left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)))\right\|_p^{p-1},$$

whereas if $\phi_p(u_i - x_i, -F_i(x)) < 0$, then for all $0 < \varepsilon \leq \varepsilon_3(x,\delta)$,

$$\begin{aligned} \left\|\left(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon\right)\right\|_p^{p-1} &\geq |\psi_p(u_i - x_i, -F_i(x), \varepsilon)|^{p-1} \\ &\geq |\psi_p(u_i - x_i, -F_i(x), \varepsilon_3(x,\delta))|^{p-1}. \end{aligned}$$

The above discussions show that for all $0 < \varepsilon \leq \varepsilon_3(x,\delta)$, we have

$$\frac{g_2(\varepsilon)}{\left\|(x_i - l_i, \psi_p(u_i - x_i, -F_i(x), \varepsilon), \varepsilon)\right\|_p^{p-1}} \leq (p-1)M_5(x)\varepsilon,$$

where

$$M_5(x) := \begin{cases} \dfrac{\left[|\phi_p(u_i - x_i, -F_i(x))| + \varepsilon_3(x,\delta)\right]^{p-2}}{\left\|(x_i - l_i, \phi_p(u_i - x_i, -F_i(x)))\right\|_p^{p-1}} & \text{if } p \geq 2, \\ \dfrac{\max\{|\phi_p(u_i - x_i, -F_i(x))|^{p-2}, |\psi_p(u_i - x_i, -F_i(x), \varepsilon_3(x,\delta))|^{p-2}\}}{|\psi_p(u_i - x_i, -F_i(x), \varepsilon_3(x,\delta))|^{p-1}} & \text{if } 1 < p < 2. \end{cases}$$

This along with (48) and the result of Step 1 gives $|b_i(x,\varepsilon) - b_i(x)| \leq M_6(x)\varepsilon^{\frac{p-1}{p}}$ where

$$M_6(x) := (p-1)M_5(x)\varepsilon_3(x,\delta)^{1/p} + M_4(x)|\phi_p(u_i - x_i, -F_i(x))|^{p-1}. \tag{49}$$

Step 3: To estimate $|c_i(x,\varepsilon) - c_i(x)|$ and $|d_i(x,\varepsilon) - d_i(x)|$. Using Lemma 2.4,

$$\begin{aligned} |c_i(x,\varepsilon) - c_i(x)| &= \left|\frac{\text{sign}(u_i - x_i)|u_i - x_i|^{p-1}}{\|(u_i - x_i, -F_i(x), \varepsilon)\|_p^{p-1}} - \frac{\text{sign}(u_i - x_i)|u_i - x_i|^{p-1}}{\|(u_i - x_i, -F_i(x))\|_p^{p-1}}\right| \\ &= \frac{|u_i - x_i|^{p-1}}{\|(u_i - x_i, -F_i(x))\|_p^{p-1}} - \frac{|u_i - x_i|^{p-1}}{\|(u_i - x_i, -F_i(x), \varepsilon)\|_p^{p-1}} \\ &\leq \left[\alpha_2(x)^{\frac{1-p}{p}} - \left(\alpha_2(x) + \varepsilon^p\right)^{\frac{1-p}{p}}\right]|u_i - x_i|^{p-1} \\ &\leq \left[\alpha(x)^{\frac{1-p}{p}} - \left(\alpha(x) + \varepsilon^p\right)^{\frac{1-p}{p}}\right]|u_i - x_i|^{p-1} \\ &= \frac{\left(\alpha(x) + \varepsilon^p\right)^{\frac{p-1}{p}} - \alpha(x)^{\frac{p-1}{p}}}{\left[\alpha(x)(\alpha(x) + \varepsilon)\right]^{\frac{p-1}{p}}}|u_i - x_i|^{p-1} \\ &\leq \frac{|u_i - x_i|^{p-1}\varepsilon^{p-1}}{\alpha(x)^{\frac{2p-2}{p}}}. \end{aligned}$$

Using the similar arguments, we also have $|d_i(x,\varepsilon) - d_i(x)| \leq \frac{|F_i(x)|^{p-1}\varepsilon^{p-1}}{\alpha(x)^{\frac{2p-2}{p}}}$.

Now using (42) and the results of the above three steps, and noting that $|b_i(x,\varepsilon)| \le 1$, $|d_i(x)| \le 1$ and $|c_i(x)| \le 1$, it follows that for all $0 < \varepsilon \le \varepsilon_3(x,\delta)$,

$$\begin{aligned}
&\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) \\
&\quad \le M_4(x)|x_i - l_i|\varepsilon^{\frac{p-1}{p}} + M_6(x)\left(1 + \|\nabla F_i(x)\|\right)\varepsilon^{\frac{p-1}{p}} + \frac{|u_i - x_i|^{p-1}\varepsilon^{p-1}}{\alpha(x)^{\frac{2p-2}{p}}} + \frac{|F_i(x)|^{p-1}\varepsilon^{p-1}}{\alpha(x)^{\frac{2p-2}{p}}} \\
&\quad \le M(x)\varepsilon^{\frac{p-1}{p}}
\end{aligned}$$

where

$$M(x) := M_4(x)|x_i - l_i| + M_6(x)\left(1 + \|\nabla F_i(x)\|\right) + \frac{\gamma_3(x)}{\alpha(x)^{\frac{2p-2}{p}}}. \tag{50}$$

Therefore, when $i \in \bar{\beta}_4(x)$, we have

$$\operatorname{dist}\left(\nabla_x \Psi_{p,i}(x,\varepsilon)^T, \partial_C \Phi_{p,i}(x)\right) \le \frac{\delta}{\sqrt{n}} \quad \text{for all } 0 < \varepsilon \le \left(\frac{\delta}{\sqrt{n}M(x)}\right)^{\frac{p-1}{p}}. \tag{51}$$

From the discussions in the above seven cases and the definition of $\varepsilon(x,\delta)$, we obtain the desired result.

References

[1] R.W. Cottle, J.-S. Pang, R.-E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.
[2] M. Ferris, J-S. Pang, Engineering and economic applications of complementarity problems, SIAM Review 39 (1997) 669–713.
[3] M. Ferris, K. Sinapiromsaran, Formulating and solving nonlinear programs as mixed complementarity problems, in: V. Nguyen, J. Strodiot, P. Tossings (Eds.), Optimization, in: Lecture Notes in Economics and Mathematical Systems, vol. 481, 2000.
[4] P.T. Harker, J.-S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications, Mathematical Programming 48 (1990) 161–220.
[5] S.C. Billups, S.P. Dirkse, M.C. Soares, A comparison of algorithms for large scale mixed complementarity problems, Computational Optimization and Applications 7 (1997) 3–25.
[6] S.C. Billups, M.C. Soares, QPCOMP: A quadratic programming based solver for mixed complementarity problems, Mathematical Programming 76 (1997) 533–562.
[7] M. Ferris, C. Kanzow, T. Munson, Feasible descent algorithms for mixed complementarity problems, Mathematical Programming 86 (1999) 475–497.
[8] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, SIAM Journal on Matrix Analysis and Applications 7 (1996) 851–868.
[9] C. Kanzow, S. Petra, On a semismooth least squares formulation of mixed complementarity problems with gap reduction, Optimization Methods and Softwares 19 (2004) 507–525.
[10] C. Kanzow, S. Petra, Projected filter trust region methods for a semismooth least squares formulation of mixed complementarity problems, Optimization Methods and Softwares 22 (2007) 713–735.
[11] S. Dirkse, M. Ferris, MCPLIB: a collection of nonlinear mixed complementarity problems, Optimization Methods and Softwares 5 (1995) 319–345.
[12] F. Facchinei, J-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I and II, Springer, New York, 2003.
[13] S. Dirkse, Robusr solution of mixed complementarity problems, Ph.D. Thesis, Computer Sciences Department, University of Wisconsin, 1994.
[14] B. Chen, P.T. Harker, A non-intertior-point continuation method for linear complementarity problem, SIAM Journal on Matrix Analysis and Applications 14 (1993) 1168–1190.
[15] C. Chen, O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, Computational Optimization and Applications 5 (1995) 97–138.
[16] X. Chen, L. Qi, A parameterized Newton method and a Broyden-like method for solving nonsmooth equations, Computational Optimization and Applications 3 (1994) 157–179.
[17] A. Fischer, A special Newton-type optimization method, Optimization 24 (1992) 269–284.
[18] S.A. Gabriel, J.J. Moré, Smoothing of mixed complementarity probblems, in: M.C. Ferris, J.S. Pang (Eds.), Complementarity and Variational Problems: State of the Art, SIAM, Pliladelphia, Pennsylvania, 1997, pp. 105–116.
[19] X. Chen, L. Qi, D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, Mathematics of Computation 67 (1998) 519–540.
[20] J.-S. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, Journal of Global Optimization 36 (2006) 565–580.
[21] J.-S. Chen, On some NCP-functions based on the generalized Fischer–Burmeister function, Asia-Pacific Journal of Opertional Research 24 (2007) 401–420.
[22] J.-S. Chen, S. Pan, A family of NCP-functions and a descent method for the nonlinear complementarity problem, Computational Optimization and Applications 40 (2008) 389–404.
[23] J.-S. Chen, S. Pan, C.Y. Yang, Numerical comparisons of two effective methods for mixed complementarity problems, Journal of Computational and Applied Mathematics (submitted for publication).
[24] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[25] L. Qi, C-differentiability, C-Differential operators and generalized Newton methods, AMR 96/5, Applied Mathematics Report, University of New South Wales, Sydney, 1996.
[26] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Mathematics of Operations Research 18 (1993) 227–244.
[27] L. Qi, J. Sun, A nonsmooth version of Newton's method, Mathematical Programming 58 (1993) 353–367.
[28] X. Chen, Smoothing methods for complementarity problems and their applications: A survery, Journal of the Operations Research 43 (2000) 32–47.