Smoothing Newton and Quasi-Newton Methods for Mixed Complementarity Problems

Donghui Li [1]
Department of Applied Mathematics
Hunan University
Changsha, China 410082
e-mail: dhli@mail.hunu.edu.cn

Masao Fukushima
Department of Applied Mathematics and Physics
Graduate School of Informatics
Kyoto University
Kyoto 606-8501, Japan
e-mail: fuku@kuamp.kyoto-u.ac.jp

February 3, 1999

Abstract

The mixed complementarity problem can be reformulated as a nonsmooth equation by using the median operator. In this paper, we first study some useful properties of this reformulation and then derive the Chen-Harker-Kanzow-Smale smoothing function for the mixed complementarity problem. On the basis of this smoothing function, we present a smoothing Newton method for solving the mixed complementarity problem. The smoothing Newton method converges globally if the problem involves a differentiable P_0 function. Under suitable conditions, the method exhibits a quadratic convergence property. We also present a smoothing Broyden-like method based on the same smoothing function. Under appropriate conditions, the method converges globally and superlinearly.

Key words: Mixed complementarity problem, smoothing function, Newton's method, quasi-Newton method

[1] Present address (available until October, 1999): Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan, e-mail: lidh@kuamp.kyoto-u.ac.jp

1. Introduction

Let $F = (F_1, F_2, \ldots, F_n)$ be a mapping from R^n into itself and $X = \prod_{i=1}^{n} [l_i, u_i]$ be a box in R^n, where $-\infty \leq l_i < u_i \leq \infty$, $i = 1,2,\ldots,n$. The variational inequality problem of finding an $x \in X$ such that

$$F(x)^T(y - x) \geq 0, \qquad \forall y \in X \tag{1.1}$$

is particularly called the mixed complementarity problem [4] and is simply denoted MCP(F). The mixed complementarity problem is also called the box constrained variational inequality problem (BVIP) [1, 25, 29, 31]. MCP(F) contains, as a special case, the nonlinear complementarity problem, NCP(F), of finding an $x \in R^n$ such that

$$x \geq 0, \;\; F(x) \geq 0, \;\; x^T F(x) = 0, \tag{1.2}$$

which corresponds to the case where $l_i = 0$ and $u_i = +\infty$, $i = 1,2,\ldots,n$. Another extreme case of MCP(F) is the problem with $l_i = -\infty$, $i = 1,2,\ldots,n$ and $u_i = +\infty$, $i = 1,2,\ldots,n$. In this case, MCP(F) reduces to the nonlinear equation

$$F(x) = 0. \tag{1.3}$$

The study on iterative methods for solving NCP(F) and MCP(F) has been rapidly developed in the last decade. One of the most popular approaches is to reformulate NCP(F) or MCP(F) as an equivalent nonsmooth equation so that generalized Newton-type methods can be applied in a way similar to those for smooth equations. For example, consider the KKT system for the MCP(F):

$$\left\{ \begin{array}{l} F(x) - \lambda + \mu = 0, \\ \lambda \geq 0, \; x - l \geq 0, \; \lambda^T(x - l) = 0, \\ \mu \geq 0, \; x - u \leq 0, \; \mu^T(x - u) = 0. \end{array} \right. \tag{1.4}$$

It is easy to see that the KKT system (1.4) can be rewritten as the equation

$$\left(\begin{array}{c} F(x) - \lambda + \mu \\ \min(\lambda, x - l) \\ \min(\mu, -(x - u)) \end{array} \right) = 0 \tag{1.5}$$

or equivalently

$$\left(\begin{array}{c} \min(F(x) + \mu, x - l) \\ \min(\mu, -(x - u)) \end{array} \right) = 0. \tag{1.6}$$

Another nonsmooth equation reformulation of MCP(F) is the equation

$$\mathrm{mid}\{x - l, F(x), x - u\} = 0. \tag{1.7}$$

Here and throughout the paper, the operator mid$\{a,b,c\}$ stands for the median of three scalars $a,b,c \in R\cup\{\pm\infty\}$. For vectors $x,y,z \in (R\cup\{\pm\infty\})^n$, the mid operation is done elementwise. We will focus on the reformulation (1.7) because it does not increase the dimension of the problem. For NCP(F), (1.7) reduces to the following well known nonsmooth equation:

$$\min\{x, F(x)\} = 0. \tag{1.8}$$

There are a variety of nonsmooth Newton methods based on (1.5) – (1.8) and other reformulations for NCP and MCP. The interested readers may find details in recent papers such as [8, 9, 11, 15, 22, 23, 27] and the references therein.

In this paper, we pay our attention to another approach to NCP and MCP, called smoothing methods. A smoothing method tries to find a solution of the nonsmooth equation (1.7) by solving a sequence of smooth equations that tend to the original equation (1.7). Smoothing methods for NCP and MCP have also received much attention in recent years [3, 4, 6, 14, 18, 28, 29]. In particular, Chen et al. [6] studied the Jacobian smoothing Newton method and established superlinear convergence of a smoothing Newton method. However, the merit function used in [6] may have an unbounded level set even if F is a uniform P function. Recently, Chen, Chen and Kanzow [1] proposed a new NCP-function by penalizing the Fisher-Burmeister function. The natural merit function of this new NCP-function has all nice features of Kanzow-Yamashita-Fukushima merit function [19] for global convergence. In particular, it provides a bounded level set. By means of this merit function, a smoothing Newton method for solving NCP with global and superlinear convergence was proposed in [1]. Merit functions that possess similar properties have also been proposed recently in [32, 33]. Another interesting smoothing Newton method was proposed by Qi, Sun and Zhou [29] for solving MCP(F) based on a $2n$-dimensional semismooth equation reformulation. An advantage of this method is that it only requires the function F to be defined on the box X. Moreover, though the reformulation is $2n$-dimensional, subproblems in the smoothing Newton method reduce to n-dimensional linear equations. However, it is not clear whether it is possible to develop a related quasi-Newton method in which subproblems still remain to be n-dimensional linear equations. In this paper, we present a new smoothing Newton method based on Chen-Harker-Kanzow-Smale smoothing function. The proposed smoothing Newton method retains global and superlinear convergence under the same assumptions as used by Chen et al. [6] and by Qi et al. [29], and the subproblems are n-dimensional linear equations. Moreover, we develop a smoothing quasi-Newton method where subproblems are also n-dimensional linear equations. We show that if F is a differentiable P_0 function with Lipschitz continuous Jacobian, then the smoothing quasi-Newton method converges globally. If in addition, the strict complementarity holds at a limit point, then the convergence rate is superlinear.

The paper is organized as follows. In the next section, we introduce the notion of MCP-

function and show some useful properties of it. In Section 3, we present a smoothing function for an MCP function and investigate its properties. In Sections 4 and 5, we propose a smoothing Newton method and a smoothing quasi-Newton method, respectively, for solving MCP, and establish their global and superlinear convergence under appropriate conditions.

Throughout the paper, we adopt the following convention: For any real number x, $\infty + x - \infty = x$, $x + \infty - \infty = x$, $0 \cdot \infty = \infty \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0 = 0$, and $\infty / \infty = 1$.

2. An MCP-Function and Its Properties

In this section, we first introduce a natural extension of NCP-function. Recall that a function $\phi : R^2 \to R$ is called an NCP-function if $\phi(a,b) = 0$ holds if and only if $a \geq 0$, $b \geq 0$ and $ab = 0$. In other words, a function $\phi : R^2 \to R$ is called an NCP-function if $\phi(a,b) = 0$ is equivalent to $\min\{a,b\} = 0$. Comparing (1.7) with (1.8), we see that the role of the mid function in MCP is very similar to that of the min function in NCP. Since $x - l \geq x - u$ for all x, we define an MCP-function as follows.

Definition 1 *Denote $\bar{R}^3 = (R \cup \{-\infty\}) \times R \times (R \cup \{+\infty\})$. A function $\phi : \bar{R}^3 \to R$ is called an MCP-function if the following relation holds:*

$$\phi(a,b,c) = 0, \ \ a > c \qquad \Longleftrightarrow \qquad \text{mid } \{a,b,c\} = 0, \ \ a > c. \tag{2.1}$$

Note that we allow a and c to be $+\infty$ and $-\infty$, respectively. In particular, when $c = -\infty$ the relation (2.1) is rewritten as

$$\phi(a,b,-\infty) = 0 \qquad \Longleftrightarrow \qquad \min\{a,b\} = 0.$$

In other words, $\phi(a,b,-\infty)$ is an MCP-function if and only if the function $\hat{\phi}(a,b) \stackrel{\triangle}{=} \phi(a,b,-\infty)$ is an NCP function. In this sense, the class of MCP-functions contains NCP-functions.

We notice that MCP-function is different from the so-called BVIP-function in [25]. For any $\alpha < \beta$, a function $\psi_{(\alpha,\beta)} : R^2 \to R$ is said to be a BVIP-function if $\psi_{(\alpha,\beta)}(a,b) = 0$ if and only if either $a = \alpha$ and $b \geq 0$, or $a = \beta$ and $b \leq 0$, or $\alpha \leq a \leq \beta$ and $b = 0$. However, both MCP-function and BVIP-function are closely related to the MCP(F). Indeed, it is easy to prove the following proposition.

Proposition 2.1 *Let ϕ be an MCP-function and $\psi_i \stackrel{\triangle}{=} \psi_{(l_i,u_i)}$ be BVIP-functions for all $i = 1,2,\ldots,n$. Then, the following statements are equivalent:*
(i) x is a solution of MCP(F);
(ii) $\phi(x_i - l_i, F_i(x), x_i - u_i) = 0$ for each $i = 1,2,\ldots,n$;
(iii) $\psi_i(x_i, F_i(x)) = 0$ for each $i = 1,2,\ldots,n$.

When $a > c$, the relation

$$\mathrm{mid}\{a,b,c\} = 0$$

is equivalent to

$$\begin{aligned} a+b+c &= \max\{a,b,c\} + \min\{a,b,c\} \\ &= \max\{a,b\} + \min\{b,c\} \\ &= \frac{(a+b)+|a-b|}{2} + \frac{(b+c)-|b-c|}{2} \\ &= \frac{(a+b)+\sqrt{(a-b)^2}}{2} + \frac{(b+c)-\sqrt{(b-c)^2}}{2}, \end{aligned}$$

or equivalently,

$$a + c = \sqrt{(a-b)^2} - \sqrt{(b-c)^2}\,.$$

Hence if we define the function $\phi : \bar{R}^3 \to R$ by

$$\phi(a,b,c) = a + c - \sqrt{(a-b)^2} + \sqrt{(b-c)^2}\,, \tag{2.2}$$

then ϕ is an MCP-function. Indeed it is not difficult to see that

$$\phi(a,b,c) = 2\ \mathrm{mid}\{a,b,c\},$$

whenever $a > c$.

The MCP-function ϕ defined by (2.2) is not new but it will play an important role in this paper. We now reformulate the MCP(F) by means of ϕ. In view of (1.7), let us define the function $\Phi(x) = (\phi_1(x), \phi_2(x), \ldots, \phi_n(x))^T$, where for each $i = 1, 2, \ldots, n$

$$\begin{aligned} \phi_i(x) &= \phi(x_i - l_i, F_i(x), x_i - u_i) \\ &= 2x_i - (l_i + u_i) - \sqrt{(x_i - l_i - F_i(x))^2} + \sqrt{(x_i - u_i - F_i(x))^2}. \end{aligned} \tag{2.3}$$

Then the MCP(F) is reformulated as the following nonsmooth equation:

$$\Phi(x) = 0. \tag{2.4}$$

For the NCP(F), the function ϕ_i becomes

$$\psi_i(x) = x_i + F_i(x) - \sqrt{(x_i - F_i(x))^2}. \tag{2.5}$$

when $l_i = -\infty$ and $u_i = +\infty$, the function ϕ_i simply reduces to $2F_i$.

We now show some useful properties of ϕ. We first consider the differentiability of ϕ. It is obvious that ϕ and hence Φ are not differentiable everywhere. Nevertheless, we can show that it has a property called semismoothness. The concept of semismoothness was introduced for scalar

functions by Mifflin [21] and was extended to vector functions by Qi [24]. It has been widely used in the context of nonlinear complementarity problems and variational inequality problems [16, 23, 24]. A locally Lipschitz function $\Psi : R^n \to R^n$ is said to be semismooth at a point x if the limit

$$\lim_{\substack{V \in \partial\Psi(x+h') \\ h' \to h}} \{Vh'\}$$

exists for any $h \in R^n$, where $\partial\Psi(x)$ stands for the Clarke generalized Jacobian of Ψ at x [7]. If Ψ is semismooth everywhere, we say that Ψ is a semismooth function. It is clear that a semismooth function is directionally differentiable [27]. Let $\Psi'(x;p)$ denote the directional derivative of Ψ at x along direction p. Then the following lemma holds [27].

Lemma 2.1 *Let Ψ be semismooth at x. Then it holds that for any $V \in \partial\Psi(x+h)$ and any $h \to 0$,*

$$\Psi(x+h) - \Psi(x) - Vh = o(\|h\|) \tag{2.6}$$

and

$$Vh - \Psi'(x;p) = o(\|h\|). \tag{2.7}$$

A locally Lipschitz function Ψ is said to be strongly semismooth at x if

$$Vh - \Psi'(x;p) = O(\|h\|^2). \tag{2.8}$$

It is not difficult to show that if Ψ is strongly semismooth at x, then we have for any $V \in \partial\Psi(x)$ and any $h \to 0$,

$$\Psi(x+h) - \Psi(x) - Vh = O(\|h\|^2). \tag{2.9}$$

It is obvious that differentiable functions and strongly semismooth functions are semismooth functions. If Ψ has Lipschitz continuous Jacobian, then it is a strongly semismooth function.

The following proposition shows the semismoothness of the function Φ defined by (2.3).

Proposition 2.2 *Let $\phi : \bar{R}^3 \to R$ be defined by (2.2) and $\Phi(x) = (\phi_1(x), \phi_2(x), \ldots, \phi_n(x))^T$ be defined by (2.3). Then*

(i) ϕ is strongly semismooth everywhere;

(ii) if F is differentiable, then Φ is semismooth;

(iii) if F has Lipschitz continuous Jacobian, then Φ is strongly semismooth.

Proof It is not difficult to show that the function $g : x \mapsto |x|$ is strongly semismooth. From the definition of ϕ, we have $\phi(a,b,c) = a + c - g(a-b) + g(b-c)$. By Theorem 19 in [12], it follows that ϕ is strongly semismooth. This concludes (i). Since Φ is (strongly) semismooth if and only if its elements ϕ_i, $i = 1, 2, \ldots, n$ are (strongly) semismooth, (ii) and (iii) follow again from Theorem 19 in [12]. □

Denote

$$\begin{aligned}
\alpha(x) &= \{i \,|\, x_i - u_i < F_i(x) < x_i - l_i\}, \\
\beta_1(x) &= \{i \,|\, F_i(x) = x_i - l_i\}, \quad \beta_2(x) = \{i \,|\, F_i(x) = x_i - u_i\}, \\
\gamma_1(x) &= \{i \,|\, F_i(x) > x_i - l_i\}, \quad \gamma_2(x) = \{i \,|\, F_i(x) < x_i - u_i\}, \\
\beta(x) &= \beta_1(x) \cup \beta_2(x), \quad \gamma(x) = \gamma_1(x) \cup \gamma_2(x).
\end{aligned}$$

Then we have

$$\phi_i(x) = \begin{cases} 2F_i(x), & \text{if} \quad i \in \alpha(x), \\ 2(x_i - l_i), & \text{if} \quad i \in \beta_1(x) \cup \gamma_1(x), \\ 2(x_i - u_i), & \text{if} \quad i \in \beta_2(x) \cup \gamma_2(x). \end{cases}$$

By using the chain rule for generalized derivatives of Lipschitz functions (see Theorem 2.3.9 (iii) in [7]), we have the following expression of $\partial\phi_i(x)$.

Proposition 2.3 *Let $F : R^n \to R^n$ be differentiable. Then we have*

$$\partial\phi_i(x) = \begin{cases} \{2\nabla F_i(x)\}, & \text{if } \ i \in \alpha(x), \\ \{(1+\rho)e_i + (1-\rho)\nabla F_i(x) \,|\, \rho \in [-1,1]\}, & \text{if } \ i \in \beta_1(x) \cup \beta_2(x), \\ \{2e_i\}, & \text{if } \ i \in \gamma_1(x) \cup \gamma_2(x), \end{cases} \tag{2.10}$$

where e_i denotes the ith unit vector in R^n.

It is clear that $\partial\Phi(x) \subseteq \partial_C\Phi(x) \stackrel{\triangle}{=} \partial\phi_1(x) \times \cdots \times \partial\phi_n(x)$ for any $x \in R^n$, though the relation $\partial\Phi(x) = \partial_C\Phi(x)$ does not necessarily hold.

The following proposition shows that Φ provides a global error bound for the MCP(F) when F is a Lipschitz continuous uniform P function.

Proposition 2.4 *Let F be Lipschitz continuous. If it is a uniform P function, then there are positive constants $m \le M$ such that*

$$m\|\Phi(x) - \Phi(y)\| \le \|x - y\| \ \le M\|\Phi(x) - \Phi(y)\|, \quad \forall x, y \in R^n. \tag{2.11}$$

In particular, we have

$$\|x - \bar{x}\| \ \le M\|\Phi(x)\|, \quad \forall x \in R^n, \tag{2.12}$$

where $\bar{x}$ stands for the unique solution of the MCP(F).

Proof First we show that the left-hand inequality in (2.11), i.e., the Lipschitz continuity of Φ. It is not difficult to show that for any $x \in R^n$, $\Phi(x) = 2\{x - \mathrm{Proj}_X(x - F(x))\}$, where $\mathrm{Proj}_X(x)$ denotes the projection of x onto the box X. The left-hand inequality of (2.11) then follows from the Lipschitz continuity of a projection operator.

Now we prove the right-hand inequality of (2.11). We first show that there exists a constant $c > 0$ such that for every $i = 1, 2, \ldots, n$, and all $x, y \in R^n$

$$c(x_i - y_i)(F_i(x) - F_i(y)) \leq \|\Phi(x) - \Phi(y)\|\|x - y\|. \tag{2.13}$$

First consider the case $i \in \alpha(x) = \Big(\alpha(x) \cap \alpha(y)\Big) \cup \Big(\alpha(x) \cap \beta(y)\Big) \cup \Big(\alpha(x) \cap \gamma(y)\Big)$.

When $i \in \alpha(x) \cap \Big(\alpha(y) \cup \beta(y)\Big)$, we have $\phi_i(x) = 2F_i(x)$ and $\phi_i(y) = 2F_i(y)$. It then follows that

$$(F_i(x) - F_i(y))(x_i - y_i) = \frac{1}{2}(\phi_i(x) - \phi_i(y))(x_i - y_i) \leq \frac{1}{2}\|\Phi(x) - \Phi(y)\| \, \|x - y\|.$$

When $i \in \alpha(x) \cap \gamma_1(y)$, we have $\phi_i(x) = 2F_i(x)$ and $\phi_i(y) = 2(y_i - l_i)$. If $x_i \geq y_i$, then we get

$$\begin{aligned}(F_i(x) - F_i(y))(x_i - y_i) &\leq \Big(F_i(x) - (y_i - l_i)\Big)(x_i - y_i) \\ &= \frac{1}{2}(\phi_i(x) - \phi_i(y))(x_i - y_i) \\ &\leq \frac{1}{2}\|\Phi(x) - \Phi(y)\| \, \|x - y\|.\end{aligned}$$

If $x_i < y_i$, then $F_i(x) < x_i - l_i < y_i - l_i < F_i(y)$. It follows that

$$\begin{aligned}(F_i(x) - F_i(y))(x_i - y_i) &= (F_i(x) - F_i(y))\Big((x_i - l_i) - \frac{1}{2}\phi_i(y)\Big) \\ &\leq \frac{1}{2}(F_i(x) - F_i(y))(\phi_i(x) - \phi_i(y)) \\ &= \frac{1}{2}\|F(x) - F(y)\| \, \|\Phi(x) - \Phi(y)\| \\ &\leq \frac{1}{2}L\|\Phi(x) - \Phi(y)\| \, \|x - y\|,\end{aligned}$$

where L stands for the Lipschitz constant of F. Similarly, we can show that (2.13) holds for every $i \in \alpha(x) \cap \gamma_2(y)$.

By similar arguments, it is not difficult to show that (2.13) holds for every $i \in \beta(x) = \Big(\beta(x) \cap \alpha(y)\Big) \cup \Big(\beta(x) \cap \beta(y)\Big) \cup \Big(\beta(x) \cap \gamma(y)\Big)$ and $i \in \gamma(x) = \Big(\gamma(x) \cap \alpha(y)\Big) \cup \Big(\gamma(x) \cap \beta(y)\Big) \cup \Big(\gamma(x) \cap \gamma(y)\Big)$.

Since F is a uniform P function, there is a constant $c' > 0$ such that

$$\max_{1 \leq i \leq n} (F_i(x) - F_i(y))(x_i - y_i) \geq c'\|x - y\|^2, \quad \forall x, y \in R^n.$$

This together with (2.13) implies

$$\|\Phi(x) - \Phi(y)\| \, \|x - y\| \geq cc'\|x - y\|^2, \quad \forall x, y \in R^n.$$

Letting $M = cc'$, we get the right-hand inequality of (2.11). □

Proposition 2.4, particularly, (2.12) shows that if F is a Lipschitz continuous uniform P function, then the level set $\{x \mid \|\Phi(x)\| \leq C\}$ is bounded for every $C > 0$.

3. Chen-Harker-Kanzow-Smale Smoothing Function for MCP

In this section, we consider a smooth function that approximates Φ and show its useful properties. For a scalar ϵ, let $\phi^\epsilon : \bar{R}^3 \to R$ be defined by

$$\phi^\epsilon(a,b,c) = a + c - \sqrt{(a-b)^2 + \epsilon^2} + \sqrt{(b-c)^2 + \epsilon^2}\,. \tag{3.1}$$

It is obvious that for any $\epsilon \neq 0$, ϕ^ϵ is smooth on R^3, and $\phi^0(a,b,c) = \phi(a,b,c)$ holds for any $(a,b,c) \in R^3$. It is interesting to notice that though ϕ^ϵ is obtained by smoothing the nonsmooth terms $\sqrt{(a-b)^2}$ and $\sqrt{(b-c)^2}$ in ϕ, the function $\frac{1}{2}\phi^\epsilon$ is just the Chen-Harker-Kanzow-Smale smoothing function [2, 17, 30], which is derived by letting

$$\rho(s) = \frac{2}{(s^2+4)^{3/2}}$$

in the Gabriel-Moré smoothing function [29]

$$\phi(\mu,c,d,w) \stackrel{\triangle}{=} \int_{\infty}^{\infty} \ \mathrm{mid}\{c,d,w-\mu s\}\rho(s)ds$$

with $w = b$, $d = a$ and $\mu = \frac{1}{2}\epsilon$. Moreover, in the case where $c = -\infty$, we have

$$\phi^\epsilon(a,b,-\infty) = a + b - \sqrt{(a-b)^2 + \epsilon^2}.$$

Notice that this function is essentially Chen-Harker-Kanzow-Smale smoothing function for NCP. In the following, we show that ϕ^ϵ possesses similar properties to those of Chen-Harker-Kanzow-Smale smoothing function for NCP. First the following result shows that ϕ^ϵ provides a uniform approximation to ϕ.

Lemma 3.1 *For any $\epsilon' > \epsilon \geq 0$, we have*

$$|\phi^{\epsilon'}(a,b,c) - \phi^\epsilon(a,b,c)| \leq \epsilon' - \epsilon, \quad \forall (a,b,c) \in \bar{R}^3.$$

In particular, we have for every $\epsilon > 0$

$$|\phi^\epsilon(a,b,c) - \phi(a,b,c)| \leq \epsilon. \quad \forall (a,b,c) \in \bar{R}^3.$$

Proof From the definition of ϕ^ϵ and ϕ, we get

$$\begin{aligned}
&|\phi^{\epsilon'}(a,b,c) - \phi^\epsilon(a,b,c)| \\
&= \left|\sqrt{(b-c)^2 + (\epsilon')^2} - \sqrt{(b-c)^2 + \epsilon^2} - \sqrt{(a-b)^2 + (\epsilon')^2} + \sqrt{(a-b)^2 + \epsilon^2}\right| \\
&= \left|\frac{[(b-c)^2 + (\epsilon')^2] - [(b-c)^2 + \epsilon^2]}{\sqrt{(b-c)^2 + (\epsilon')^2} + \sqrt{(b-c)^2 + \epsilon^2}} - \frac{[(a-b)^2 + (\epsilon')^2] - [(a-b)^2 + \epsilon^2]}{\sqrt{(a-b)^2 + (\epsilon')^2} + \sqrt{(a-b)^2 + \epsilon^2}}\right|
\end{aligned}$$

$$
\begin{aligned}
&= \Big|\frac{(\epsilon')^2-\epsilon^2}{\sqrt{(b-c)^2+(\epsilon')^2}+\sqrt{(b-c)^2+\epsilon^2}} - \frac{(\epsilon')^2-\epsilon^2}{\sqrt{(a-b)^2+(\epsilon')^2}+\sqrt{(a-b)^2+\epsilon^2}}\Big| \\
&\le \max\Big\{\frac{(\epsilon')^2-\epsilon^2}{\sqrt{(b-c)^2+(\epsilon')^2}+\sqrt{(b-c)^2+\epsilon^2}}, \frac{(\epsilon')^2-\epsilon^2}{\sqrt{(a-b)^2+(\epsilon')^2}+\sqrt{(a-b)^2+\epsilon^2}}\Big\} \qquad (3.2)\\
&\le \frac{(\epsilon')^2-\epsilon^2}{\epsilon'+\epsilon} = \epsilon'-\epsilon.
\end{aligned}
$$

□

Lemma 3.1 particularly shows that $\phi^\epsilon \to \phi$ uniformly as $\epsilon \to 0$. It is then natural to construct a smoothing function for Φ based on ϕ^ϵ. For each $i = 1, 2, \ldots, n$, define $\phi_i^\epsilon : R^n \to R$ by

$$
\phi_i^\epsilon(x) = 2x_i - (l_i + u_i) - \sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2} + \sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}\,. \qquad (3.3)
$$

Let $\Phi^\epsilon(x) = (\phi_1^\epsilon(x), \phi_2^\epsilon(x), \ldots, \phi_n^\epsilon(x))^T$. It is obvious that $\Phi^0(x) \equiv \Phi(x)$. Moreover, it follows from Lemma 3.1 that for any $\epsilon' > \epsilon \ge 0$

$$
\|\Phi^{\epsilon'}(x) - \Phi^\epsilon(x)\| \le \sqrt{n}(\epsilon' - \epsilon), \qquad \forall x \in R^n, \qquad (3.4)
$$

and

$$
\|\Phi^\epsilon(x) - \Phi(x)\| \le \sqrt{n}\epsilon, \qquad \forall x \in R^n. \qquad (3.5)
$$

Note that, for every $\epsilon \ne 0$, ϕ^ϵ is differentiable everywhere whenever F is differentiable. Specifically, by direct deduction we get for any $\epsilon \ne 0$ and $i = 1, 2, \ldots, n$

$$
\begin{aligned}
\nabla\phi_i^\epsilon(x) \;=\; & \Big(2 - \frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + \frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}}\Big)e_i \\
& + \Big(\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} - \frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}}\Big)\nabla F_i(x).
\end{aligned} \qquad (3.6)
$$

Denote for each $i = 1, 2, \ldots, n$

$$
b_i^\epsilon(x) = \frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} - \frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} \qquad (3.7)
$$

and

$$
a_i^\epsilon(x) = 2 - b_i^\epsilon(x). \qquad (3.8)
$$

Then we have

$$
\nabla\phi_i^\epsilon(x) = a_i^\epsilon(x)e_i + b_i^\epsilon(x)\nabla F_i(x), \qquad (3.9)
$$

or, in a matrix form,

$$
\nabla\Phi^\epsilon(x) = \ \mathrm{diag}\,(a_i^\epsilon(x)) + \nabla F(x)\ \mathrm{diag}\,(b_i^\epsilon(x)). \qquad (3.10)
$$

Lemma 3.2 *We have for every $\epsilon \neq 0$ and any $x \in R^n$*

$$0 \leq a_i^\epsilon(x) < 2, \quad 0 < b_i^\epsilon(x) \leq 2, \qquad \forall i \notin \beta_1(x) \cup \beta_2(x) \tag{3.11}$$

and

$$1 \leq a_i^\epsilon(x) \leq 1 + \frac{\epsilon^2}{2(u_i - l_i)^2}, \quad 1 - \frac{\epsilon^2}{2(u_i - l_i)^2} \leq b_i^\epsilon(x) \leq 1, \qquad \forall i \in \beta_1(x) \cup \beta_2(x). \tag{3.12}$$

If either l_i or u_i is finite, then all inequalities in (3.11) hold strictly.

Proof If $l_i = -\infty$ and $u_i = \infty$, then $\beta_1(x) = \beta_2(x) = \emptyset$. Moreover, we have for each $i \notin \beta_1(x) \cup \beta_2(x)$, $b_i^\epsilon(x) = 2$ and $a_i^\epsilon(x) = 0$, and hence (3.11) holds. Now, assume that either l_i or u_i is finite. It is not difficult to see that for every fixed $\epsilon \neq 0$, the function $t \mapsto \frac{t}{\sqrt{t^2+\epsilon^2}}$ is increasing with t. It then follows that

$$\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} > \frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}},$$

which implies that $b_i^\epsilon(x) > 0$ and hence $a_i^\epsilon(x) < 2$ for each i. On the other hand, it is obvious that for any $\epsilon \neq 0$

$$\Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}}\Big| \leq 1 \quad \text{and} \quad \Big|\frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}}\Big| \leq 1. \tag{3.13}$$

Moreover, at least one of the inequalities in (3.13) must hold strictly because at least one of l_i and u_i is finite. This yields $b_i^\epsilon(x) < 2$ and hence $a_i^\epsilon(x) > 0$. Thus we have (3.11). Note that, in this case, the inequalities in (3.11) hold strictly.

Next, we prove (3.12). If l_i or u_i is infinite, then it follows from (3.7) that $a_i^\epsilon(x) = b_i^\epsilon(x) = 1$ for every $i \in \beta(x)$. Suppose that both l_i and u_i are finite. Then, we have for each $i \in \beta(x)$

$$b_i^\epsilon(x) = \frac{u_i - l_i}{\sqrt{(u_i - l_i)^2 + \epsilon^2}} \leq 1.$$

Moreover, we have

$$\begin{aligned} b_i^\epsilon(x) &= 1 - 1 + \frac{u_i - l_i}{\sqrt{(u_i - l_i)^2 + \epsilon^2}} \\ &= 1 - \frac{\epsilon^2}{\Big(\sqrt{(u_i - l_i)^2 + \epsilon^2} + (u_i - l_i)\Big)\sqrt{(u_i - l_i)^2 + \epsilon^2}} \\ &\geq 1 - \frac{1}{2(u_i - l_i)^2}\epsilon^2. \end{aligned}$$

By the definition of $a_i^\epsilon(x)$, it is not difficult to complete the proof. □

The inequality (3.5) implies that Φ^ϵ approximates Φ uniformly. In the following, we will see that $\nabla\Phi^\epsilon(x)$ serves as a good approximation to an element of $\partial_C\Phi(x)$. That is, $\nabla\Phi^\epsilon$ and $\nabla\Phi$ satisfy the Jacobian consistence property introduced by Chen, Qi and Sun [6].

Definition 2 *Let $\Psi = (\psi_1, \psi_2, \ldots, \psi_n)^T : R^n \to R^n$ be a Lipschitz continuous function. We call $\Psi^\epsilon : R^n \to R^n$ a smoothing approximation function of Ψ if Ψ^ϵ is continuously differentiable and there is a constant $\mu > 0$ such that for any $x \in R^n$ and $\epsilon > 0$*

$$\|\Psi^\epsilon(x) - \Psi(x)\| \leq \mu\epsilon. \tag{3.14}$$

Furthermore, if for any $x \in R^n$

$$\lim_{\epsilon \downarrow 0} \mathrm{dist}(\nabla\Psi^\epsilon(x)^T, \partial_C \Psi(x)) = 0, \tag{3.15}$$

then we say Ψ and Ψ^ϵ satisfy the Jacobian consistence property, where $\mathrm{dist}(\nabla\Psi^\epsilon(x)^T, \partial_C\Psi(x))$ denotes the distance from the matrix $\nabla\Psi^\epsilon(x)^T$ to the set of matrices $\partial_C\Psi(x)$ defined by

$$\partial_C \Psi(x) = \partial\psi_1(x) \times \partial\psi_2(x) \times \cdots \times \partial\psi_n(x).$$

From (3.5), it is obvious that the function Φ^ϵ defined by (3.3) is a smoothing function of Φ defined by (2.3). We now show that Φ and Φ^ϵ satisfy the Jacobian consistence property. To this end, we construct the matrix $V^\epsilon(x) = (v_1^\epsilon(x), v_2^\epsilon(x), \ldots, v_n^\epsilon(x))$, where vectors $v_i^\epsilon(x) \in R^n$, $i = 1, 2, \ldots, n$, are defined by

$$v_i^\epsilon(x) = \begin{cases} 2\nabla F_i(x), & i \in \alpha(x), \\ a_i^\epsilon(x)e_i + b_i^\epsilon(x)\nabla F_i(x), & i \in \beta_1(x) \cup \beta_2(x), \\ 2e_i, & i \in \gamma_1(x) \cup \gamma_2(x). \end{cases} \tag{3.16}$$

Since $a_i^\epsilon(x)$ and $b_i^\epsilon(x)$ satisfy (3.8) and (3.11), it is clear from (2.10) and (3.16) that $v_i^\epsilon(x) \in \partial\phi_i(x)$ for each $i = 1, 2, \ldots, n$ and any $\epsilon \neq 0$, and hence $V^\epsilon(x) \in \partial_C\Phi(x)$. By (3.8), we get

$$\nabla\phi_i^\epsilon(x) - v_i^\epsilon(x) = \begin{cases} a_i^\epsilon(x)(e_i - \nabla F_i(x)), & i \in \alpha(x), \\ 0, & i \in \beta_1(x) \cup \beta_2(x), \\ b_i^\epsilon(x)(\nabla F_i(x) - e_i), & i \in \gamma_1(x) \cup \gamma_2(x). \end{cases} \tag{3.17}$$

It is easy to see by the definition of (3.7) and (3.8) of $a_i^\epsilon(x)$ and $b_i^\epsilon(x)$ that

$$\lim_{\epsilon \downarrow 0} b_i^\epsilon(x) = \lim_{\epsilon \downarrow 0}(2 - a_i^\epsilon(x)) = 0, \quad \forall i \in \gamma_1(x) \cup \gamma_2(x),$$

and

$$\lim_{\epsilon \downarrow 0} a_i^\epsilon(x) = \lim_{\epsilon \downarrow 0}(2 - b_i^\epsilon(x)) = 0, \quad \forall i \in \alpha(x).$$

We then deduce from (3.17) that for every $i = 1, 2, \ldots, n$

$$\lim_{\epsilon \downarrow 0}(\nabla\phi_i^\epsilon(x) - v_i^\epsilon(x)) = 0. \tag{3.18}$$

Let $V(x) = (v_1(x), v_2(x), \ldots, v_n(x)) \stackrel{\triangle}{=} \lim_{\epsilon \downarrow 0} V^\epsilon(x)$. Then we get from (3.12) and (3.16)

$$v_i(x) = \begin{cases} 2\nabla F_i(x), & i \in \alpha(x), \\ e_i + \nabla F_i(x), & i \in \beta_1(x) \cup \beta_2(x), \\ 2e_i, & i \in \gamma_1(x) \cup \gamma_2(x). \end{cases}$$

It is not difficult to see that $V(x) \in \partial_C \Phi(x)$. Therefore, (3.18) implies

$$\lim_{\epsilon \downarrow 0} \operatorname{dist}(\nabla \Phi^\epsilon(x), \partial_C \Phi(x)) \leq \lim_{\epsilon \downarrow 0} \|\nabla \Phi^\epsilon(x) - V(x)\| = 0.$$

The above discussion has shown the following proposition.

Proposition 3.1 *Let F be continuously differentiable. Then Φ and Φ^ϵ satisfy the Jacobian consistence property.*

In the next proposition, we give sufficient conditions to ensure the nonsingularity of $\nabla \Phi^\epsilon(x)$.

Proposition 3.2 *Let F be continuously differentiable. Then $\nabla \Phi^\epsilon(x)$ is nonsingular for every $\epsilon \neq 0$ if one of the following conditions holds.*
(i) $\nabla F(x)$ is a P matrix;
(ii) $\nabla F(x)$ is a P_0 matrix and, for each $i \notin \beta_1(x) \cup \beta_2(x)$, either l_i or u_i is finite.

Proof (i) From Lemma 3.2, $a_i^\epsilon(x) \geq 0$ and $b_i^\epsilon(x) > 0$ hold for every i and any x. Therefore, in view of (3.10), the nonsingularity of $\nabla \Phi^\epsilon(x)$ is equivalent to the nonsingularity of the matrix $\operatorname{diag}\left(a_i^\epsilon(x) b_i^\epsilon(x)^{-1}\right) + \nabla F(x)$. If $\nabla F(x)$ is a P matrix, then so is $\operatorname{diag}\left(a_i^\epsilon(x) b_i^\epsilon(x)^{-1}\right) + \nabla F(x)$ and hence nonsingular. If $\nabla F(x)$ is a P_0 matrix and for each $i \notin \beta_1(x) \cup \beta_2(x)$, either l_i or u_i is finite, then by Lemma 3.2 again, the matrix $\operatorname{diag}\left(a_i^\epsilon(x) b_i^\epsilon(x)^{-1}\right)$ is a diagonal matrix with positive diagonal elements, and hence the nonsingularity of the matrix $\operatorname{diag}\left(a_i^\epsilon(x) b_i^\epsilon(x)^{-1}\right) + \nabla F(x)$ follows from Lemma 5.1 in De Luca et al. [8]. □

To conclude this section, we show that if F is a continuously differentiable uniform P function, then the smooth equation

$$\Phi^\epsilon(x) = 0 \tag{3.19}$$

has a unique solution x^ϵ for each $\epsilon > 0$, and x^ϵ tends to the unique solution of MCP(F) as $\epsilon \to 0$.

Lemma 3.3 *Let F be continuously differentiable. If it is a uniform P function. then there is a constant $M > 0$ such that for any $\epsilon \geq 0$*

$$\|x^\epsilon - \bar{x}\| \leq M\sqrt{n}\epsilon, \tag{3.20}$$

where $\bar{x}$ is the unique solution of MCP(F). In particular, we have $\lim_{\epsilon \downarrow 0} x^\epsilon = \bar{x}$.

Proof We first prove the existence of a unique solution of (3.19). By Proposition 2.4, the function $\|\Phi\|$ has bounded level sets. Since Φ^ϵ satisfies (3.5), it follows that the level set $\Omega^\epsilon = \{x \,|\, \|\Phi^\epsilon(x)\| \leq C\}$ is nonempty and bounded for an arbitrary constant C such that $C > \inf_x \|\Phi^\epsilon(x)\|$. Moreover, it is clear that $\|\Phi^\epsilon\|^2$ attains its infimum at some point $x^\epsilon \in \text{ int } \Omega^\epsilon$. Since x^ϵ satisfies $\nabla\Phi^\epsilon(x^\epsilon)\Phi^\epsilon(x^\epsilon) = 0$ and $\nabla\Phi^\epsilon(x)$ is nonsingular by Proposition 3.2, we must have $\Phi^\epsilon(x^\epsilon) = 0$. This shows that (3.19) has a solution. The uniqueness of the solution of (3.19) follows from the nonsingularity of $\nabla\Phi^\epsilon(x)$ for every x.

Next, we show that x^ϵ converges to $\bar{x}$. Since $\bar{x}$ is the unique solution of MCP(F), it is the unique zero point of Φ. So, it follows from (2.12) in Proposition 2.4 that

$$\|x^\epsilon - \bar{x}\| \leq M\|\Phi(x^\epsilon)\| = M\|\Phi(x^\epsilon) - \Phi^\epsilon(x^\epsilon)\| \leq M\sqrt{n}\epsilon,$$

where the last inequality follows from . □

4. A Smoothing Newton Method and Its Convergence

In this section, we propose a smoothing Newton method based on Φ^ϵ for solving the MCP(F). For simplicity, we abbreviate Φ^{ϵ_k}, $\phi_i^{\epsilon_k}$, etc. as Φ^k, ϕ_i^k, etc., respectively. In the rest of the paper, we assume that F is continuously differentiable.

In the algorithm, the subproblem is the following linear equation

$$\nabla\Phi^k(x_k)p + \Phi(x_k) = 0. \tag{4.1}$$

This type of subproblems has been used by many authors (see e.g., [6, 26, 29]). For line search, it is natural to use the squared residual function

$$\varphi^k(x) \stackrel{\triangle}{=} \frac{1}{2}\|\Phi^k(x)\|^2. \tag{4.2}$$

as the merit function. Generally, in Newton's method, the steplength λ_k is determined by

$$\varphi^k(x_k + \lambda_k p_k) \leq \varphi^k(x_k) + \sigma\lambda_k\nabla\varphi^k(x_k)^T p_k$$

or

$$\varphi^k(x_k + \lambda_k p_k) \leq \varphi^k(x_k) + 2\sigma\lambda_k\varphi(x_k),$$

where $\varphi(x) = \varphi^0(x)$, p_k is the solution of (4.1), and $\sigma > 0$ is a suitable constant. Although it is possible to construct a smoothing Newton method with those line searches, we here adopt a different line search that is also effective in the smoothing quasi-Newton method to be presented in the next section. Let $\{\eta_k\}$ be a positive sequence satisfying

$$\sum_{k=1}^{\infty} \eta_k \leq \eta < \infty, \tag{4.3}$$

where η is a given positive constant. At iteration k, we determine a steplength $\lambda_k > 0$ so that the following inequality holds for $\lambda = \lambda_k$:

$$\|\Phi^k(x_k + \lambda p_k)\| - \|\Phi^k(x_k)\| \leq -\sigma\|\lambda p_k\|^2 + \eta_k, \tag{4.4}$$

where $\sigma > 0$ is a constant. Clearly, at each iteration k, (4.4) holds for all $\lambda > 0$ sufficiently small since η_k is positive and independent of λ.

Let $\Delta_i(x) = \min\{(x_i - l_i - F_i(x))^2, (x_i - u_i - F_i(x))^2\}$, $i = 1, 2, \ldots, n$, and

$$\Delta(x) = \min\{\Delta_i(x) \,|\, i \in \alpha(x) \cup \gamma_1(x) \cup \gamma_2(x)\}. \tag{4.5}$$

The proposed smoothing Newton method is stated as follows.

Algorithm 1 (Smoothing Newton Method):

Step 0. Choose constants $\rho_1, \rho_2 \in (0,1)$, $0 < \gamma < \min\{\frac{1}{2\sqrt{n}}, \frac{\rho_2}{\sqrt{n}}\}$, $\sigma_1, \sigma_2 > 0$. Select a positive sequence $\{\eta_k\}$ satisfying (4.3). Choose an initial point $x_1 \in R^n$ and a positive constant $\epsilon_1 \leq \frac{\gamma}{2}\|\Phi(x_1)\|$. Let $k := 1$.

Step 1. Stop, if $\Phi(x_k) = 0$. Otherwise, solve the linear equation (4.1) to get p_k.

Step 2. If

$$\|\Phi^k(x_k + p_k)\| \leq \rho_2\|\Phi^k(x_k)\| - \sigma_1\|p_k\|^2, \tag{4.6}$$

then let $\lambda_k := 1$ and go to Step 4.

Step 3. Let λ_k be the maximum number in the set $\{1, \rho_1, \rho_1^2, \ldots, \}$ such that $\lambda = \rho_1^i$ satisfies the line search condition (4.4) with $\sigma = \sigma_2$.

Step 4. Let $x_{k+1} := x_k + \lambda_k p_k$.

Step 5. Update ϵ_k by the following rule: If (4.6) holds or $\gamma\|\Phi(x_{k+1})\| \leq \epsilon_k$, let

$$\epsilon_{k+1} := \min\{\frac{\gamma}{2}\|\Phi(x_{k+1})\|, \frac{1}{2}\epsilon_k, \Delta(x_{k+1})\}, \tag{4.7}$$

where $\Delta(x)$ is defined by (4.5). Otherwise, let $\epsilon_{k+1} := \epsilon_k$.

Step 6. Let $k := k + 1$. Go to Step 1.

Remark (i) It is not difficult to see that if Algorithm 1 generates an infinite sequence $\{x_k\}$, then $\Phi^k(x_k) \neq 0$ for all k, and the positive sequence $\{\epsilon_k\}$ is nonincreasing and satisfies

$$\epsilon_k \leq \gamma\|\Phi(x_k)\|, \qquad \forall k. \tag{4.8}$$

It then follows from (3.5) that

$$\epsilon_k \leq \gamma\|\Phi(x_k)\| \leq \gamma(\|\Phi^k(x_k)\| + \sqrt{n}\epsilon_k).$$

Since $\sqrt{n}\gamma < \frac{1}{2}$ by Step 0, we get

$$\epsilon_k \leq \nu\|\Phi^k(x_k)\|, \tag{4.9}$$

where $\nu = \gamma/(1-\sqrt{n}\gamma) \in (0, \frac{1}{\sqrt{n}})$.
(ii) Define index set $K = K_1 \cup K_2$, where

$$K_1 = \{k \,|\, \gamma\|\Phi(x_{k+1})\| \le \epsilon_k\},\ K_2 = \{k \,|\, \|\Phi^k(x_k + p_k)\| \le \rho_2\|\Phi^k(x_k)\| - \sigma_1\|p_k\|^2\}. \tag{4.10}$$

Then, $\epsilon_k \le \frac{1}{2}\epsilon_{k-1}$ if $k-1 \in K$, and $\epsilon_k = \epsilon_{k-1}$ if $k-1 \notin K$.

Lemma 4.1 *Let $\{x_k\}$ be generated by Algorithm 1. Then for every k, we have*

$$\|\Phi^{k+1}(x_{k+1})\| \le \rho_2\|\Phi^k(x_k)\| - \sigma_1\|x_{k+1} - x_k\|^2 + \sqrt{n}(\epsilon_k - \epsilon_{k+1}) \tag{4.11}$$

if λ_k is determined by Step 2, and

$$\|\Phi^{k+1}(x_{k+1})\| \le \|\Phi^k(x_k)\| - \sigma_2\|x_{k+1} - x_k\|^2 + \sqrt{n}(\epsilon_k - \epsilon_{k+1}) + \eta_k \tag{4.12}$$

if λ_k is determined by Step 3. In particular, we have for every k

$$\|\Phi^{k+1}(x_{k+1})\| \le \|\Phi^k(x_k)\| - \sigma\|x_{k+1} - x_k\|^2 + \sqrt{n}(\epsilon_k - \epsilon_{k+1}) + \eta_k, \tag{4.13}$$

where $\sigma = \min\{\sigma_1, \sigma_2\}$. Consequently, we have

$$\sum_{k=1}^{\infty} \|x_{k+1} - x_k\|^2 < \infty. \tag{4.14}$$

Moreover, the sequence $\{\|\Phi^k(x_k)\|\}$ converges.

Proof By (3.4), we clearly have for every k

$$\|\Phi^{k+1}(x_{k+1})\| \le \|\Phi^k(x_{k+1})\| + \sqrt{n}(\epsilon_k - \epsilon_{k+1}). \tag{4.15}$$

If λ_k is determined by Step 2, then (4.11) follows from (4.6) and (4.15). Since $\rho_2 \in (0,1)$, (4.11) implies (4.13). If λ_k is determined by Step 3, then (4.12) follows from (4.15) and (4.4) with $\sigma = \sigma_2$, and hence (4.13) holds. Let us denote $\bar{\eta}_k = \sqrt{n}(\epsilon_k - \epsilon_{k+1}) + \eta_k > 0$. Since η_k satisfies (4.3), we have $\sum_{k=1}^{\infty} \bar{\eta}_k < \infty$. Therefore, (4.14) follows from (4.13) and, by Lemma 3.3 in [10], $\{\|\Phi^k(x_k)\|\}$ converges. □

Lemma 4.2 *Let $\{x_k\}$ be generated by Algorithm 1. Then the following four statements are equivalent.*

(i)

$$\liminf_{k\to\infty} \|\Phi(x_k)\| = 0.$$

(ii)

$$\liminf_{k\to\infty} \|\Phi^k(x_k)\| = 0.$$

(iii)

$$\lim_{k\to\infty} \|\Phi(x_k)\| = 0.$$

(iv)

$$\lim_{k\to\infty} \|\Phi^k(x_k)\| = 0.$$

Proof Since $\{\|\Phi^k(x_k)\|\}$ converges, by Lemma 4.1, (ii) and (iv) are equivalent. From (3.5) and (4.8), we have

$$\|\Phi(x_k)\| \le \|\Phi^k(x_k)\| + \sqrt{n}\epsilon_k \le \|\Phi^k(x_k)\| + \gamma\sqrt{n}\|\Phi(x_k)\|.$$

This implies

$$\|\Phi(x_k)\| \le \frac{1}{1-\sqrt{n}\gamma}\|\Phi^k(x_k)\|. \tag{4.16}$$

Similarly we have

$$\|\Phi^k(x_k)\| \le \|\Phi(x_k)\| + \sqrt{n}\epsilon_k \le (1+\gamma\sqrt{n})\|\Phi(x_k)\|.$$

This together with (4.16) shows the equivalence between (i) and (ii) as well as the equivalence between (iii) and (iv). □

Lemma 4.2 reveals that $\{x_k\}$ has an accumulation point that is a solution of MCP(F), and every accumulation point of $\{x_k\}$ is a solution of MCP(F).

Theorem 4.1 *Let $\{x_k\}$ be generated by Algorithm 1. If the index set K defined by (4.10) is infinite, then we have*

$$\lim_{k\to\infty} \|\Phi(x_k)\| = 0, \tag{4.17}$$

and every accumulation point of $\{x_k\}$ is a solution of MCP(F).

Proof The assumption that K is infinite means that either K_2 is infinite, or K_2 is finite but K_1 is infinite. In the first case, by Lemma 4.1, (4.11) holds infinitely often. Since $\rho_2 \in (0,1)$, this implies $\lim_{k\to\infty} \|\Phi^k(x_k)\| = 0$. Then (4.17) follows from Lemma 4.2.

Now consider the latter case. Without loss of generality, we assume that $K_1 = K = \{k_1 < k_2 < \cdots\}$. By Step 5, we have

$$\begin{aligned} \|\Phi(x_{k_i+1})\| &\le \gamma^{-1}\epsilon_{k_i} = \gamma^{-1}\epsilon_{k_{i-1}+1} \le \frac{1}{2}\|\Phi(x_{k_{i-1}+1})\| \\ &\le \cdots \le \left(\frac{1}{2}\right)^{i-1}\|\Phi(x_{k_1+1})\|. \end{aligned}$$

This shows that $\liminf_{k\to\infty} \|\Phi(x_k)\| = 0$. Consequently, by Lemma 4.2 again, we get (4.17). Hence every accumulation point of $\{x_k\}$ is a zero point of Φ, or equivalently, a solution of MCP(F). □

Theorem 4.1 shows that to prove global convergence of Algorithm 1, it suffices to verify that the index set K defined by (4.10) is infinite. To this end, we need the following assumption.

Assumption A For each $\epsilon > 0$, $\nabla\Phi^\epsilon(x)$ is nonsingular for any x.

Sufficient conditions for Assumption A to hold are given in Proposition 3.2. The following theorem shows that under Assumption A, the index set K is infinite.

Theorem 4.2 *Let Assumption A hold and $\{x_k\}$ be generated by Algorithm 1. Then the index set K defined by (4.10) is infinite and hence (4.17) holds. Consequently, every accumulation point of $\{x_k\}$ is a solution of MCP(F).*

Proof We assume to the contrary that K is finite. Then by Step 5 of Algorithm 1, there is a positive integer $\bar{k}$ such that $\epsilon_k = \epsilon_{\bar{k}}$ holds for all $k \geq \bar{k}$, which in turn implies that $\Phi^k = \Phi^{\bar{k}}$ for all $k \geq \bar{k}$. Moreover, we have by (4.8), $\|\Phi(x_k)\| \geq \gamma^{-1}\epsilon_{\bar{k}} > 0$ for all $k \geq \bar{k}$ and hence by Lemma 4.2, $\|\Phi^{\bar{k}}(x_k)\| \geq \epsilon'$ for all k sufficiently large and some constant $\epsilon' > 0$. We denote $\Phi^{\bar{k}}$ by $\bar{\Phi}$. Since K_2 is finite, λ_k is determined by Step 3 for all k sufficiently large. So, we have for all k sufficiently large

$$\|\bar{\Phi}(x_k + \lambda_k p_k)\| - \|\bar{\Phi}(x_k)\| \leq -\sigma_2\|\lambda_k p_k\|^2 + \eta_k. \tag{4.18}$$

This particularly implies that $\|\lambda_k p_k\| \to 0$ as $k \to \infty$. Let $\bar{x}$ be an accumulation point of $\{x_k\}$ and $\{x_k\}_{k\in\bar{K}}$ be a subsequence converging to $\bar{x}$. Since $\bar{\Phi}$ is continuously differentiable and $\nabla\bar{\Phi}(\bar{x})$ is nonsingular, it follows that when $k \in \bar{K}$ is sufficiently large, $\nabla\bar{\Phi}(x_k)$ is uniformly nonsingular. Therefore, (4.1) implies that $\{p_k\}_{k\in\bar{K}}$ is bounded. Without loss of generality, we assume that $\{p_k\}_{k\in\bar{K}}$ converges to some vector $\bar{p}$. Let $\bar{\lambda} = \limsup_{k\in\bar{K},k\to\infty} \lambda_k$. Then $\bar{\lambda} \in [0, 1]$. If $\bar{\lambda} > 0$, then $\bar{p} = 0$. So it follows from (4.1) that $\Phi(\bar{x}) = 0$. This is a contradiction. Now, suppose $\bar{\lambda} = 0$, or equivalently, $\lim_{k\in\bar{K},k\to\infty} \lambda_k = 0$. By Step 3 of Algorithm 1, when $k \in \bar{K}$ is sufficiently large, $\lambda'_k \triangleq \lambda_k/\rho_1$ does not satisfy (4.18). So, we have

$$\|\bar{\Phi}(x_k + \lambda'_k p_k)\| - \|\bar{\Phi}(x_k)\| \geq -\sigma_2\|\lambda'_k p_k\|^2.$$

Multiplying both sides by $(\lambda'_k)^{-1}\left(\|\bar{\Phi}(x_k+\lambda'_k p_k)\| + \|\bar{\Phi}(x_k)\|\right)$ and then taking the limit as $k \to \infty$ with $k \in \bar{K}$ yield

$$2\bar{\Phi}(\bar{x})^T\nabla\bar{\Phi}(\bar{x})\bar{p} \geq 0.$$

It then follows from (4.1) that

$$\bar{\Phi}(\bar{x})^T\Phi(\bar{x}) \leq 0.$$

This together with (4.8) implies

$$\begin{aligned}(1-n\gamma^2)\|\Phi(\bar{x})\|^2 &\leq \|\bar{\Phi}(\bar{x})\|^2 + (1-n\gamma^2)\|\Phi(\bar{x})\|^2 \leq \|\bar{\Phi}(\bar{x})\|^2 + \|\Phi(\bar{x})\|^2 - n\bar{\epsilon}^2 \\ &= \|\bar{\Phi}(\bar{x}) - \Phi(\bar{x})\|^2 - n\bar{\epsilon}^2 + 2\bar{\Phi}(\bar{x})\Phi(\bar{x}) \leq 2\bar{\Phi}(\bar{x})^T\Phi(\bar{x}) \leq 0, \qquad (4.19)\end{aligned}$$

where the third inequality follows from (3.5). Since $\gamma < \frac{1}{\sqrt{n}}$, (4.19) is also a contradiction. Therefore, K must be infinite, and hence the assertion follows from Theorem 4.1. □

Define the level set

$$\Omega = \{x \,|\, \|\Phi(x)\| \leq \frac{7}{4}\|\Phi(x_1)\| + \eta\}, \qquad (4.20)$$

where the positive constant η is given in (4.3). The following theorem shows that if Ω is bounded, then there is at least an accumulation point of $\{x_k\}$ that solves MCP(F).

Theorem 4.3 *Let Assumption A hold and the level set Ω be contained in a bounded convex set $\bar{\Omega}$. Then the sequence $\{x_k\}$ generated by Algorithm 1 has at least one accumulation point. Moreover, every accumulation point of $\{x_k\}$ is a solution of MCP(F).*

Proof From Theorem 4.2, it suffices to show that $\{x_k\}$ is bounded. By the boundedness of Ω, it then suffices to verify that $\{x_k\} \subset \Omega$. Indeed, from (4.13), we have for every k

$$\begin{aligned}\|\Phi^{k+1}(x_{k+1})\| &\leq \|\Phi^k(x_k)\| + \sqrt{n}(\epsilon_k - \epsilon_{k+1}) + \eta_k \\ &\leq \cdots \leq \|\Phi^1(x_1)\| + \sqrt{n}\epsilon_1 + \eta \\ &\leq \|\Phi(x_1)\| + \|\Phi^1(x_1) - \Phi(x_1)\| + \sqrt{n}\epsilon_1 + \eta \\ &\leq \|\Phi(x_1)\| + 2\sqrt{n}\epsilon_1 + \eta,\end{aligned}$$

where the last inequality follows from (3.5). This implies

$$\begin{aligned}\|\Phi(x_{k+1})\| &\leq \|\Phi^{k+1}(x_{k+1})\| + \|\Phi^{k+1}(x_{k+1}) - \Phi(x_{k+1})\| \\ &\leq \|\Phi(x_1)\| + 2\sqrt{n}\epsilon_1 + \eta + \sqrt{n}\epsilon_{k+1} \leq \|\Phi(x_1)\| + 3\sqrt{n}\epsilon_1 + \eta \\ &\leq \|\Phi(x_1)\| + \frac{3}{2}\gamma\sqrt{n}\|\Phi(x_1)\| + \eta \leq \frac{7}{4}\|\Phi(x_1)\| + \eta, \qquad (4.21)\end{aligned}$$

where the last two inequalities follow from Step 0 of Algorithm 1. Since (4.21) holds for all k, we get $\{x_k\} \subset \Omega$ as desired. □

By Proposition 2.4, if F is a uniform P function, then the level set Ω is bounded. So, we have the following corollary.

Corollary 4.1 *Let F be a continuously differentiable uniform P function. Then $\{x_k\}$ generated by Algorithm 1 converges to the unique solution of MCP(F).*

We say that a semismooth function $\Psi : R^n \to R^n$ is BD-regular at a point x if every matrix $V \in \partial\Psi_B(x)$ is nonsingular, where $\partial_B\Psi(x)$ is the B-differential of Ψ at x (see [24]).

By using BD-regularity, we can prove a strong convergence property for Algorithm 1.

Theorem 4.4 *Let Assumption A hold and $\{x_k\}$ be generated by Algorithm 1. Suppose that there is an accumulation point x^* of $\{x_k\}$ at which Φ is BD-regular. Then the whole sequence $\{x_k\}$ converges to x^*.*

Proof The BD-regularity implies that x^* is an isolated solution of MCP(F) and hence an isolated accumulation point of $\{x_k\}$ ([24], Proposition 2.5). Since $x_{k+1} - x_k \to 0$ by Lemma 4.1, the whole sequence $\{x_k\}$ converges. □

We now turn to proving superlinear/quadratic convergence of Algorithm 1. We first prove a useful lemma.

Lemma 4.3 *Let $V^\epsilon(x)$ be defined by (3.16) and $\Delta(x)$ be defined by (4.5). If $0 < \epsilon \le \Delta(x)$, then we have*

$$\|\nabla\Phi^\epsilon(x) - V^\epsilon(x)\| \le \sqrt{n}\epsilon(1 + \|\nabla F(x)\|). \tag{4.22}$$

Proof It suffices to show that for every $i = 1, 2, \ldots, n$

$$\|\nabla\phi_i^\epsilon(x) - v_i^\epsilon(x)\| \le \epsilon(1 + \|\nabla F(x)\|). \tag{4.23}$$

First, we have for any scalar t

$$\Big|1 - \frac{1}{\sqrt{1+t^2}}\Big| = \frac{|1 - \sqrt{1+t^2}|}{\sqrt{1+t^2}} \le |1 - \sqrt{1+t^2}| = \frac{t^2}{1+\sqrt{1+t^2}} \le \frac{1}{2}t^2. \tag{4.24}$$

For $i \in \gamma_1(x)$, we have $x_i - l_i - F_i(x) < 0$ and $x_i - u_i - F_i(x) < 0$. Letting $t = -\dfrac{\epsilon}{x_i - l_i - F_i(x)}$ in (4.24), we get

$$\Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + 1\Big| \le \frac{\epsilon^2}{2(x_i - l_i - F_i(x))^2} \le \frac{\epsilon\Delta(x)}{2(x_i - l_i - F_i(x))^2} \le \frac{1}{2}\epsilon.$$

Similarly, letting $t = -\dfrac{\epsilon}{x_i - u_i - F_i(x)}$ in (4.24), we get

$$\Big|\frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} + 1\Big| \le \frac{1}{2}\epsilon.$$

Therefore, we get for every $i \in \gamma_1(x)$

$$\begin{aligned} |b_i^\epsilon(x)| &= \Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + 1 - \frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} - 1\Big| \\ &\le \Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + 1\Big| + \Big|\frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} + 1\Big| \\ &\le \epsilon. \end{aligned}$$

This together with (3.17) implies that (4.23) holds for every $i \in \gamma_1(x)$. By a similar discussion, we can show that (4.23) holds for every other i. The proof is then complete. □

Let K be defined by (4.10). Then, by Step 5 of Algorithm 1, for every k with $k-1 \in K$, we have $\epsilon_k \leq \Delta(x_k)$. Since $\epsilon_k \leq \gamma\|\Phi(x_k)\|$ holds for all k, Theorem 4.2 implies $\lim_{k\to\infty} \epsilon_k = 0$. So, we get the following corollary from Theorem 4.2 and Lemma 4.3.

Corollary 4.2 *Let the conditions of Theorem 4.2 hold. Then*

$$\lim_{\substack{k \to \infty \\ k-1 \in K}} \|\nabla\Phi^k(x_k) - V_k\| = 0, \tag{4.25}$$

where $V_k = V^k(x_k)$ is defined by (3.16).

The following theorem shows another good property of the smoothing function Φ^ϵ.

Theorem 4.5 *Let Assumption A hold and $\{x_k\}$ be generated by Algorithm 1. Suppose that $\{x_k\}$ converges to x^*. Then we have*

$$\lim_{\substack{k \to \infty \\ k-1 \in K}} \operatorname{dist}(\nabla\Phi^k(x_k), \partial_C\Phi(x^*)) = \lim_{k\to\infty} \operatorname{dist}(V_k, \partial_C\Phi(x^*)) = 0, \tag{4.26}$$

where $V_k = V^k(x_k)$ is defined by (3.16).

Proof Since (4.25) holds, we only need to verify the last equality. Let $I_1 \cup I_2 \cup I_3$ be an arbitrary partition of the index set $\beta(x^*) = \beta_1(x^*) \cup \beta_2(x^*)$. That is, $I_i \cap I_j = \emptyset$, $i \neq j$, $i,j = 1,2,3$ and $I_1 \cup I_2 \cup I_3 = \beta(x^*)$. Since, by Proposition 2.3, $\{2e_i, 2\nabla F_i(x^*), e_i + \nabla F_i(x^*)\} \subset \partial\phi_i(x^*)$ for each $i \in \beta(x^*)$, $\partial_C\Phi(x^*)$ contains matrix $V \stackrel{\triangle}{=} (v_1, v_2, \ldots, v_n)$ with

$$v_i = \begin{cases} 2\nabla F_i(x^*), & i \in \alpha(x^*) \cup I_1, \\ 2e_i, & i \in \gamma_1(x^*) \cup \gamma_2(x^*) \cup I_2, \\ e_i + \nabla F_i(x^*), & i \in I_3. \end{cases} \tag{4.27}$$

We verify the last equality of (4.26) by showing that every accumulation point of $\{V_k\}$ is in $\partial_C\Phi(x^*)$. It suffices to show that for an arbitrary accumulation point $V^* = (v_1^*, v_2^*, \ldots, v_n^*)$ of $\{V_k\}$, there is a partition $I_1 \cup I_2 \cup I_3$ of $\beta(x^*)$ such that, for each i, v_i^* is expressed in the form of (4.27). Let V^* be an accumulation point of $\{V_k\}$, and $\{V_k\}_{k\in\bar{K}}$ be a subsequence converging to V^*. For every $i \in \alpha(x^*)$, we have $i \in \alpha(x_k)$ for all $k \in \bar{K}$ sufficiently large. Therefore, from (3.16), we get $v_i^k = 2\nabla F_i(x_k)$ for every $i \in \alpha(x^*)$ and all k sufficiently large, and hence $v_i^* = 2\nabla F_i(x^*)$ for every $i \in \alpha(x^*)$. Similarly, we have $v_i^* = 2e_i$ for each $i \in \gamma(x^*)$. Now, consider the indices $i \in \beta(x^*)$. By taking a further subsequence of $\{x_k\}_{k\in\bar{K}}$ if necessary, we can partition the index $\beta(x^*)$ as I_1, I_2, I_3 such that $I_1 = \{i \in \beta(x^*) \,|\, i \in \alpha(x_k) \text{ for all } k \in \bar{K}\}$,

$I_2 = \{i \in \beta(x^*) \,|\, i \in \gamma(x_k) \text{ for all } k \in \bar{K}\}$, $I_3 = \{i \in \beta(x^*) \,|\, i \in \beta(x_k) \text{ for all } k \in \bar{K}\}$. Then, taking into account (3.12), we see from (3.16) that $v_i^* = 2\nabla F_i(x^*)$ for each $i \in I_1$, $v_i^* = 2e_i(x^*)$ for each $i \in I_2$, and $v_i^* = e_i + \nabla F_i(x^*)$ for each $i \in I_3$. Hence $V^* = (v_1^*, v_2^*, \ldots, v_n^*)$ is represented in the form of (4.27). □

Now, we show superlinear/quadratic convergence of Algorithm 1.

Theorem 4.6 *Let Assumption A hold and $\{x_k\}$ be generated by Algorithm 1. Suppose that there is an accumulation point x^* of $\{x_k\}$ such that all matrices in $\partial_C \Phi(x^*)$ are nonsingular. Then $\{x_k\}$ converges to x^* superlinearly. If in addition, ∇F is Lipschitz continuous at x^*, then the convergence rate is quadratic.*

Proof It is clear that the given conditions imply that Φ is BD-regular at x^*. Therefore, by Theorem 4.4, $\{x_k\} \to x^*$.

Let K be defined by (4.10). By Theorem 4.5, there is a constant $M_1 > 0$ such that $\|\nabla \Phi^k(x_k)^{-1}\| \le M_1$ for all k sufficiently large with $k-1 \in K$. So, it follows from (4.1) that for every k with $k-1 \in K$

$$\begin{aligned}
\|x_k + p_k - x^*\| &= \| - \nabla \Phi^k(x_k)^{-1} \Big(\Phi(x_k) - \Phi(x^*) - \nabla \Phi^k(x_k)(x_k - x^*) \Big) \| \\
&\le \|\nabla \Phi^k(x_k)^{-1}\| \Big(\|\Phi(x_k) - \Phi(x^*) - V_k(x_k - x^*)\| \\
&\quad + \|(\nabla \Phi^k(x_k) - V_k)(x_k - x^*)\| \Big) \\
&\le M_1 \Big(\sum_{i=1}^{n} \|\phi_i(x_k) - \phi_i(x^*) - (v_i^k)^T (x_k - x^*)\| + \|\nabla \Phi^k(x_k) - V_k\| \, \|x_k - x^*\| \Big) \\
&= o(\|x_k - x^*\|), \qquad (4.28)
\end{aligned}$$

where $V_k = V^{\epsilon_k}(x_k)$ is given by (3.16), and the last equality follows from (4.25) and the fact that $v_i^k \in \partial \phi_i(x_k)$ together with the semismoothness of ϕ_i. Then we have for all k sufficiently large with $k-1 \in K$

$$\begin{aligned}
\|\Phi^k(x_k + p_k)\| &\le \|\Phi^k(x_k + p_k) - \Phi(x_k + p_k)\| + \|\Phi(x_k + p_k) - \Phi(x^*)\| \\
&\le \sqrt{n}\epsilon_k + O(\|x_k + p_k - x^*\|) \\
&\le \sqrt{n}\nu \|\Phi^k(x_k)\| + o(\|x_k - x^*\|) \\
&= \rho_2 \|\Phi^k(x_k)\| - (\rho_2 - \sqrt{n}\nu)\|\Phi^k(x_k)\| + o(\|x_k - x^*\|), \qquad (4.29)
\end{aligned}$$

where the second inequality follows from (3.5) and the local Lipschitzian property of Φ, and the third inequality follows from (4.9) and (4.28).

On the other hand, since Φ is BD-regular at x^*, there is a constant $m > 0$ such that $\|\Phi(x)\| = \|\Phi(x) - \Phi(x^*)\| \ge m\|x - x^*\|$ for all x near x^*. So, it follows that for all k sufficiently large with

$k-1 \in K$

$$\begin{aligned} \|\Phi^k(x_k)\| &\geq \|\Phi(x_k)\| - \|\Phi^k(x_k) - \Phi(x_k)\| \\ &\geq \|\Phi(x_k)\| - \sqrt{n}\epsilon_k \\ &\geq (1 - \sqrt{n}\gamma)\|\Phi(x_k) - \Phi(x^*)\| \\ &\geq (1 - \sqrt{n}\gamma)m\|x_k - x^*\| = \mu\|x_k - x^*\|, \end{aligned} \tag{4.30}$$

where $\mu = (1 - \sqrt{n}\gamma)m > 0$ and the third inequality follows from (4.8). Also it follows from (4.1) that for all k sufficiently large with $k-1 \in K$

$$\|p_k\| \leq \|\nabla\Phi^k(x_k)^{-1}\|\,\|\Phi(x_k)\| \leq M_1\|\Phi(x_k) - \Phi(x^*)\| \leq M_1 L\|x_k - x^*\|, \tag{4.31}$$

where $L > 0$ stands for the Lipschitz constant of Φ at x^*. Therefore, it follows from (4.29), (4.30), and (4.31) that for all k sufficiently large with $k-1 \in K$

$$\|\Phi^k(x_k + p_k)\| - \rho_2\|\Phi^k(x_k)\| + \sigma_1\|p_k\|^2 \leq -(\rho_2 - \sqrt{n}\nu)\|\Phi^k(x_k)\| + o(\|x_k - x^*\|) \leq 0. \tag{4.32}$$

This shows that (4.6) holds, and hence $k \in K$. By induction, we claim that $k \in K$ for all k sufficiently large. Moreover, by Step 2 in Algorithm 1, (4.32) means that $\lambda_k = 1$ is accepted. Consequently, (4.28) shows the superlinear convergence of $\{x_k\}$.

If ∇F is Lipschitz continuous at x^*, then Φ is strongly semismooth at x^*. By means of (4.22) and (4.8), we have $\|\nabla\Phi^k(x_k) - V_k\| = O(\|x_k - x^*\|^2)$. Therefore, in a similar way to (4.28), we can deduce $\|x_k + p_k - x^*\| = O(\|x_k - x^*\|^2)$ and hence, $\|x_{k+1} - x^*\| = O(\|x_k - x^*\|^2)$. That is, $\{x_k\}$ converges to x^* quadratically. □

The proof of Theorem 4.6 shows that the index set K defined by (4.10) contains all k sufficiently large. Therefore, we get the following corollary from Theorem 4.5.

Corollary 4.3 *Let the conditions of Theorem 4.4 hold. Then we have*

$$\lim_{k\to\infty} \operatorname{dist}(\nabla\Phi^k(x_k),\, \partial_C\Phi(x^*)) = 0.$$

Since $\partial_C\Phi(x^*)$ is closed, Corollary 4.3 shows that any accumulation point of $\{\nabla\Phi^k(x_k)\}$ is an element of $\partial_C\Phi(x^*)$. In particular, if Φ is differentiable at x^*, then $\partial_C\Phi(x^*)$ reduces to a singleton, and hence

$$\lim_{k\to\infty} \nabla\Phi^k(x_k) = \nabla\Phi(x^*).$$

5. A Smoothing Quasi-Newton Method

In this section, based on the smoothing Newton method presented in Section 4, we propose a smoothing Broyden-like method for solving MCP(F). The basic idea is to use a nonsingular matrix B_k instead of $\nabla\Phi^k(x_k)$ in Algorithm 1. The algorithm is stated as follows.

Algorithm 2 (Smoothing Broyden-like Method)

Step 0. Choose constants $\rho_1, \rho_2 \in (0,1)$, $0 < \gamma < \min\{\frac{1}{2\sqrt{n}}, \frac{\rho_2}{\sqrt{n}}\}$, $\sigma_1, \sigma_2 > 0$. Choose an initial point $x_1 \in R^n$, an $n \times n$ nonsingular matrix B_1, and positive constants $\epsilon_1 \leq \frac{\gamma}{2}\|\Phi(x_1)\|$ and $\eta_1 \leq \|\Phi(x_1)\|$. Let $k := 1$.

Step 1. Stop, if $\Phi(x_k) = 0$. Otherwise, solve the linear equation

$$B_k p + \Phi(x_k) = 0 \tag{5.1}$$

to get p_k.

Step 2. If

$$\|\Phi^k(x_k + p_k)\| \leq \rho_2\|\Phi^k(x_k)\| - \sigma_1\|p_k\|^2, \tag{5.2}$$

then let $\lambda_k := 1$ and go to Step 4.

Step 3. Let λ_k be the maximum number in the set $\{1, \rho_1, \rho_1^2, \ldots, \}$ such that $\lambda = \rho_1^i$ satisfies the line search condition (4.4) with $\sigma = \sigma_2$.

Step 4. Let $x_{k+1} := x_k + \lambda_k p_k$.

Step 5. Update B_k by the Broyden-like update formula

$$B_{k+1} = B_k + \theta_k \frac{(y_k - B_k s_k)s_k^T}{\|s_k\|^2},$$

where $s_k := x_{k+1} - x_k$ and $y_k := \Phi^k(x_{k+1}) - \Phi^k(x_k)$. The parameter θ_k is chosen so that $|\theta_k - 1| \leq \bar{\theta}$ for some constant $\bar{\theta} \in (0,1)$ and B_{k+1} is nonsingular.

Step 6. Update ϵ_k by the following rule: If (5.2) holds or $\gamma\|\Phi(x_{k+1})\| \leq \epsilon_k$, let

$$\epsilon_{k+1} := \min\{\frac{\gamma}{2}\|\Phi(x_{k+1})\|, \frac{1}{2}\epsilon_k\}.$$

Otherwise, let $\epsilon_{k+1} := \epsilon_k$.

Step 7. Let $\eta_{k+1} := \min\{\frac{1}{2}\eta_k, \|\Phi(x_{k+1})\|\}$ and $k := k + 1$. Go to Step 1.

Remark (i) Since $\eta_k \leq \frac{1}{2}\eta_{k-1}$ for all k, it is obvious that (4.3) holds.

(ii) Similar to Algorithm 1, we have for every k

$$\epsilon_k \leq \gamma\|\Phi(x_k)\|, \quad \epsilon_k \leq \nu\|\Phi^k(x_k)\| \quad \text{and} \quad \eta_k \leq \|\Phi(x_k)\|, \tag{5.3}$$

where $\nu = \frac{\gamma}{1 - \sqrt{n}\gamma} \in (0, \frac{1}{\sqrt{n}})$ (see (4.16)).

(iii) We still denote $K = K_1 \cup K_2$ with

$$K_1 = \{k \,|\, \gamma\|\Phi(x_{k+1})\| \leq \epsilon_k\},\ K_2 = \{k \,|\, \|\Phi^k(x_k + p_k)\| \leq \rho_2\|\Phi^k(x_k)\| - \sigma_1\|p_k\|^2\}. \tag{5.4}$$

Then, $\epsilon_k \leq \frac{1}{2}\epsilon_{k-1}$ holds for every k with $k - 1 \in K$, and $\epsilon_k = \epsilon_{k-1}$ holds for every k with $k - 1 \notin K$.

(iv) It is not difficult to see that the sequence $\{x_k\}$ generated by Algorithm 2 is contained in the level set Ω defined by (4.20). Moreover, Lemmas 4.1, 4.2 and Theorem 4.1 remain true for Algorithm 2. This means that to show global convergence of Algorithm 2, it suffices to verify that K defined by (5.4) is infinite.

In order to establish global convergence for Algorithm 2, we need the following assumption.

Assumption B (i) The level set Ω defined by (4.20) is contained in a bounded convex set $\bar{\Omega}$. (ii) The function F is continuously differentiable and ∇F is Lipschitz continuous on $\bar{\Omega}$.

For given $\epsilon > 0$ and $\theta > 0$, define the set

$$\Pi^{\epsilon}(\theta) = \{x \,|\, (x_i - l_i - F_i(x))^2 + \epsilon^2 \geq \theta,\ (x_i - u_i - F_i(x))^2 + \epsilon^2 \geq \theta, i = 1, 2, \ldots, n\} \cap \bar{\Omega}. \quad (5.5)$$

The following lemma shows that under Assumption B, $\nabla \Phi^{\epsilon}$ is Lipschitz continuous on $\Pi^{\epsilon}(\theta)$.

Lemma 5.1 *Let Assumption B hold. Then for any $\epsilon > 0$ and $\theta > 0$, there exists a positive constant L_θ independent of ϵ, such that*

$$\|\nabla \Phi^{\epsilon}(x) - \nabla \Phi^{\epsilon}(y)\| \leq L_\theta \|x - y\|, \quad \forall x, y \in \Pi^{\epsilon}(\theta). \quad (5.6)$$

Proof We prove (5.6) by showing that there is a positive constant l_θ such that for each $i = 1, 2, \ldots, n$

$$\|\nabla \phi_i^{\epsilon}(x) - \nabla \phi_i^{\epsilon}(y)\| \leq l_\theta \|x - y\|, \quad \forall x, y \in \Pi^{\epsilon}(\theta). \quad (5.7)$$

We first show that there is a constant $l'_\theta > 0$ such that for each $i = 1, 2, \ldots, n$

$$|a_i^{\epsilon}(x) - a_i^{\epsilon}(y)| = |b_i^{\epsilon}(x) - b_i^{\epsilon}(y)| \leq l'_\theta \|x - y\|, \quad \forall x, y \in \Pi^{\epsilon}(\theta). \quad (5.8)$$

By an elementary deduction, we get for any scalars t_1, t_2 and $\epsilon > 0$

$$\begin{aligned}
\Big|\frac{t_1}{\sqrt{t_1^2 + \epsilon^2}} - \frac{t_2}{\sqrt{t_2^2 + \epsilon^2}}\Big| &\leq \frac{|t_1 - t_2|}{\sqrt{t_1^2 + \epsilon^2}} + |t_2|\Big|\frac{1}{\sqrt{t_1^2 + \epsilon^2}} - \frac{1}{\sqrt{t_2^2 + \epsilon^2}}\Big| \\
&= \frac{|t_1 - t_2|}{\sqrt{t_1^2 + \epsilon^2}} + |t_2|\frac{|\sqrt{t_1^2 + \epsilon^2} - \sqrt{t_2^2 + \epsilon^2}|}{\sqrt{t_1^2 + \epsilon^2}\sqrt{t_2^2 + \epsilon^2}} \\
&\leq \frac{1}{\sqrt{t_1^2 + \epsilon^2}}\Big(|t_1 - t_2| + |\sqrt{t_1^2 + \epsilon^2} - \sqrt{t_2^2 + \epsilon^2}|\Big) \\
&= \frac{1}{\sqrt{t_1^2 + \epsilon^2}}\Big(|t_1 - t_2| + \frac{|t_1^2 - t_2^2|}{\sqrt{t_1^2 + \epsilon^2} + \sqrt{t_2^2 + \epsilon^2}}\Big) \\
&\leq \frac{1}{\sqrt{t_1^2 + \epsilon^2}}\Big(|t_1 - t_2| + \frac{(|t_1| + |t_2|)|t_1 - t_2|}{|t_1| + |t_2|}\Big) \\
&= \frac{2|t_1 - t_2|}{\sqrt{t_1^2 + \epsilon^2}}.
\end{aligned}$$

So, we get by the definition (3.7) of b_i^ϵ

$$
\begin{aligned}
&|b_i^\epsilon(x) - b_i^\epsilon(y)| \\
&= \Big|\Big(\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} - \frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}}\Big) \\
&\quad - \Big(\frac{y_i - l_i - F_i(y)}{\sqrt{(y_i - l_i - F_i(y))^2 + \epsilon^2}} - \frac{y_i - u_i - F_i(y)}{\sqrt{(y_i - u_i - F_i(y))^2 + \epsilon^2}}\Big)\Big| \\
&\le \Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} - \frac{y_i - l_i - F_i(y)}{\sqrt{(y_i - l_i - F_i(y))^2 + \epsilon^2}}\Big| \\
&\quad + \Big|\frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} - \frac{y_i - u_i - F_i(y)}{\sqrt{(y_i - u_i - F_i(y))^2 + \epsilon^2}}\Big| \\
&\le \frac{2|(x_i - l_i - F_i(x)) - (y_i - l_i - F_i(y))|}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + \frac{2|(x_i - u_i - F_i(x)) - (y_i - u_i - F_i(y))|}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} \\
&\le \frac{4}{\theta}\Big(|x_i - y_i| + |F_i(x) - F_i(y)|\Big).
\end{aligned}
$$

Since F is continuously differentiable, it is Lipschitz continuous on the bounded set $\Pi^\epsilon(\theta)$. The last inequality then implies that (5.8) holds for each i with a suitable constant l'_θ independent of ϵ. Therefore, by (3.9), we get for each i

$$
\begin{aligned}
\|\nabla\phi_i^\epsilon(x) - \nabla\phi_i^\epsilon(y)\| &\le \Big\|(a_i^\epsilon(x) - a_i^\epsilon(y))e_i\Big\| + \Big\|b_i^\epsilon(x)\nabla F_i(x) - b_i^\epsilon(y)\nabla F_i(y)\Big\| \\
&\le |a_i^\epsilon(x) - a_i^\epsilon(y)| + \|(b_i^\epsilon(x) - b_i^\epsilon(y))\nabla F_i(x)\| + \|b_i^\epsilon(y)(\nabla F_i(x) - \nabla F_i(y))\|.
\end{aligned}
$$

Since ∇F is Lipschitz continuous and $|b_i^\epsilon(x)| \le 2$, the last inequality and (5.8) imply (5.7), which in turn implies (5.6) with some constant $L_\theta > 0$. □

Define

$$
\delta_k = \frac{\|y_k - B_k s_k\|}{\|s_k\|} \quad \text{and} \quad A_k = \int_0^1 \nabla\Phi^k(x_k + \tau s_k)d\tau. \tag{5.9}
$$

Then, we have $y_k = A_k s_k$ and

$$
\delta_k = \frac{\|(A_k - B_k)s_k\|}{\|s_k\|} = \frac{\|(A_k - B_k)p_k\|}{\|p_k\|}. \tag{5.10}
$$

So, it follows that

$$
\begin{aligned}
\|(B_k - \nabla\Phi^k(x_k))p_k\| &\le \|(A_k - B_k)p_k\| + \|(\nabla\Phi^k(x_k) - A_k)p_k\| \\
&\le (\delta_k + \|\nabla\Phi^k(x_k) - A_k\|)\|p_k\|. \qquad (5.11)
\end{aligned}
$$

In a similar way to Lemma 4.1 in [20], it is not difficult to prove the following useful lemma.

Lemma 5.2 *Let Assumption B hold. If the index set K defined by (5.4) is finite, then we have*

$$
\lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} \delta_i^2 = 0. \tag{5.12}
$$

Now, we are ready to prove global convergence for Algorithm 2.

Theorem 5.1 *Let Assumptions A and B hold and $\{x_k\}$ be generated by Algorithm 2. Then the index set K must be infinite and hence*

$$\lim_{k\to\infty} \|\Phi(x_k)\| = 0.$$

Moreover, every accumulation of $\{x_k\}$ is a solution of MCP(F).

Proof Assume to the contrary that K is finite. Then there is an integer $\bar{k}$ such that $\epsilon_k = \epsilon_{\bar{k}}$ for all $k \geq \bar{k}$. Denote $\bar{\Phi} = \Phi^{\bar{k}}$. Notice that the finiteness of K implies that K_2 defined by (5.4) is also finite, and hence there is an index $\tilde{k}$ such that when $k \geq \tilde{k}$, the steplength λ_k is determined by Step 3. That is, we have for every $k \geq \max\{\bar{k}, \tilde{k}\}$

$$\|\bar{\Phi}(x_k + \lambda_k p_k)\| - \|\bar{\Phi}(x_k)\| \leq -\sigma_2 \|\lambda_k p_k\|^2 + \eta_k.$$

By (5.12), there is a subsequence $\{\delta_k\}_{k\in\bar{K}}$ converging to zero. Let $\bar{x}$ be an arbitrary accumulation point of the corresponding sequence $\{x_k\}_{k\in\bar{K}}$. Without loss of generality, we assume $\lim_{k\in\bar{K},k\to\infty} x_k = \bar{x}$. By (5.1) and (5.11), we have

$$\|\Phi(x_k)\| \leq \|(B_k - \nabla\bar{\Phi}(x_k))p_k\| + \|\nabla\bar{\Phi}(x_k)p_k\| \leq (\delta_k + \|\nabla\Phi^k(x_k) - A_k\| + \|\nabla\bar{\Phi}(x_k)\|)\|p_k\|$$

and

$$\begin{aligned} \|p_k\| &= \|\nabla\bar{\Phi}(x_k)^{-1}(\Phi(x_k) + (B_k - \nabla\bar{\Phi}(x_k))p_k\| \\ &\leq \|\nabla\bar{\Phi}(x_k)^{-1}\|\Big(\|\Phi(x_k)\| + \|(B_k - \nabla\bar{\Phi}(x_k))p_k\|\Big) \\ &\leq \|\bar{\Phi}(x_k)^{-1}\|\Big(\|\Phi(x_k)\| + (\delta_k + \|\nabla\bar{\Phi}(x_k) - A_k\|)\,\|p_k\|\Big). \end{aligned} \tag{5.13}$$

So, there exists a constant $C > 0$ such that for all $k \in \bar{K}$ sufficiently large

$$\|\Phi(x_k)\| \leq C\|p_k\|$$

and

$$\|p_k\| \leq C\|\Phi(x_k)\|. \tag{5.14}$$

In a similar way to the proof of Theorem 4.2, we can deduce a contradiction. □

Following the same arguments as in the proof of Theorem 4.4, it is not difficult to prove the following theorem.

Theorem 5.2 *Let Assumptions A and B hold. Suppose that there is an accumulation point x^* of $\{x_k\}$ at which Φ is BD-regular. Then the whole sequence $\{x_k\}$ converges to x^*.*

Now we turn to proving superlinear convergence of Algorithm 2. For this purpose, we make further assumptions.

Assumption C (i) The sequence $\{x_k\}$ generated by Algorithm 2 converges to a solution x^* of MCP(F) at which Φ is BD-regular.

(ii) The strict complementarity holds at x^*, i.e., $\beta_1(x^*) = \beta_2(x^*) = \emptyset$; in other words,

$$(x_i^* - l_i)^2 + F_i(x^*)^2 > 0 \quad \text{and} \quad (x_i^* - u_i)^2 + F_i(x^*)^2 > 0, \quad i = 1, 2, \ldots, n. \tag{5.15}$$

Let

$$\bar{\tau} = \frac{1}{2} \min\{|x_i^* - l_i - F_i(x^*)|,\ |x_i^* - u_i - F_i(x^*)| \,|\, i = 1, 2, \ldots, n\}.$$

and

$$U^* = \{x \,|\, |x_i - l_i - F_i(x)| \geq \bar{\tau},\ |x_i - u_i - F_i(x)| \geq \bar{\tau}, i = 1, 2, \ldots, n\}.$$

It is clear that, under Assumption C, $\bar{\tau} > 0$ and $U^* \neq \emptyset$. It is also obvious that $\beta_1(x) = \beta_2(x) = \emptyset$ for every $x \in U^*$, and hence, Φ is continuously differentiable on U^*. Specifically, for every $x \in U^*$, we have $\nabla \Phi(x) = (\nabla \phi_1(x), \nabla \phi_2(x), \ldots, \nabla \phi_n(x))$ with

$$\nabla \phi_i(x) = \begin{cases} 2\nabla F_i(x), & \text{if } i \in \alpha(x), \\ 2e_i, & \text{if } i \in \gamma(x). \end{cases} \tag{5.16}$$

Moreover, there exists a constant $M_2 > 0$ such that

$$\|\Phi(x) - \Phi(x^*)\| \leq M_2 \|x - x^*\|, \quad \forall x \in U^*. \tag{5.17}$$

Let $\Pi^\epsilon(\theta)$ be defined by (5.5). Then we have $U^* \subset \Pi^\epsilon(\bar{\tau}^2)$ for any ϵ. Therefore, in a similar way to Lemma 5.1, we can prove the following result.

Lemma 5.3 *Let Assumption B hold. Then there is a positive constant L_0 independent of ϵ such that*

$$\|\nabla \Phi(x) - \nabla \Phi(y)\| \leq L_0 \|x - y\|, \quad \forall x, y \in U^*$$

and

$$\|\nabla \Phi^\epsilon(x) - \nabla \Phi^\epsilon(y)\| \leq L_0 \|x - y\|, \quad \forall \epsilon > 0,\ \forall x, y \in U^*.$$

Lemma 5.4 *Let a solution x^* of MCP(F) satisfy (5.15). Then there are constants $C_1 > 0$ and $C_2 > 0$ such that for every $\epsilon > 0$*

$$\|\Phi^\epsilon(x) - \Phi(x)\| \leq C_1 \epsilon^2, \qquad \forall x \in U^* \tag{5.18}$$

and

$$\|\nabla \Phi^\epsilon(x) - \nabla \Phi(x)\| \leq C_2 \epsilon^2, \qquad \forall x \in U^*. \tag{5.19}$$

Proof We deduce from (3.2) that for every $(a,b,c) \in \bar{R}^3$

$$|\phi^\epsilon(a,b,c) - \phi(a,b,c)| \le \max\Big\{\frac{\epsilon^2}{2|b-c|}, \frac{\epsilon^2}{2|a-b|}\Big\}.$$

Substituting $x_i - l_i$ for a, $F_i(x)$ for b, and $x_i - u_i$ for c in the last inequality yields

$$|\phi_i^\epsilon(x) - \phi_i(x)| \le \max\Big\{\frac{\epsilon^2}{2|x_i - u_i - F_i(x)|}, \frac{\epsilon^2}{2|x_i - l_i - F_i(x)|}\Big\}.$$

This implies that for all $x \in U^*$

$$|\phi_i^\epsilon(x) - \phi_i(x)| \le \frac{\epsilon^2}{2\bar{\tau}}.$$

Letting $C_1 = \frac{\sqrt{n}}{2\bar{\tau}}$, we get (5.18).

Now we prove (5.19). From (3.9) and (5.16), we have for every $\epsilon > 0$ and any $x \in U^*$

$$\nabla\phi_i^\epsilon(x) - \nabla\phi_i(x) = \begin{cases} a_i^\epsilon(x)(e_i - \nabla F_i(x)), & i \in \alpha(x), \\ b_i^\epsilon(x)(-e_i + \nabla F_i(x)), & i \in \gamma(x). \end{cases} \tag{5.20}$$

For each $i \in \gamma_1(x)$, we have by (3.7)

$$|b_i^\epsilon(x)| \le \Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + 1\Big| + \Big|\frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} + 1\Big|. \tag{5.21}$$

Moreover, we have

$$\begin{aligned} &\Big|\frac{x_i - l_i - F_i(x)}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}} + 1\Big| \qquad (5.22)\\ &= \Big|\frac{(F_i(x) - x_i + l_i) - \sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}}\Big| \\ &= \frac{\epsilon^2}{\sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}\Big((F_i(x) - x_i + l_i) + \sqrt{(x_i - l_i - F_i(x))^2 + \epsilon^2}\Big)} \\ &\le \frac{1}{2\bar{\tau}^2}\epsilon^2. \qquad (5.23) \end{aligned}$$

Similarly, we have

$$\Big|\frac{x_i - u_i - F_i(x)}{\sqrt{(x_i - u_i - F_i(x))^2 + \epsilon^2}} + 1\Big| \le \frac{1}{2\bar{\tau}^2}\epsilon^2. \tag{5.24}$$

Combining (5.23) and (5.24) with (5.21), we get

$$|b_i^\epsilon(x)| \le \frac{1}{\bar{\tau}^2}\epsilon^2. \tag{5.25}$$

By a similar argument, we can show that (5.25) holds for each $i \in \gamma_2(x)$, every $\epsilon > 0$ and any $x \in U^*$. Moreover, we have for each $i \in \alpha(x)$, every $\epsilon > 0$ and any $x \in U^*$

$$|a_i^\epsilon(x)| \le \frac{1}{\bar{\tau}^2}\epsilon^2.$$

Then by using (5.20), we can deduce (5.19) with some constant $C_2 > 0$. □

Lemma 5.5 *Let Assumptions B and C hold, $\{x_k\}$ be generated by Algorithm 2, and δ_k be defined by (5.9). Then there exist a constant $\delta > 0$ and an integer $k' > 0$ such that $\lambda_k = 1$ whenever $\delta_k \le \delta$ and $k \ge k'$.*

Proof Since $x_k \to x^*$, there is an index k_1' such that $x_k \in U^*$ for all $k \ge k_1'$. Let A_k be defined by (5.9). Then we have for all $k \ge k_1'$

$$\begin{aligned}
\|A_k - \nabla\Phi(x^*)\| &\le \int_0^1 \|\nabla\Phi^k(x_k + \tau s_k) - \nabla\Phi(x_k + \tau s_k)\| d\tau + \int_0^1 \|\nabla\Phi(x_k + \tau s_k) - \nabla\Phi(x^*)\| d\tau \\
&\le C_2\epsilon_k^2 + L_0 \int_0^1 \|\tau(x_{k+1} - x^*) + (1-\tau)(x_k - x^*)\| d\tau \\
&\le C_2\gamma^2\|\Phi(x_k) - \Phi(x^*)\|^2 + L_0(\|x_{k+1} - x^*\| + \|x_k - x^*\|) \\
&\le C_2\gamma^2 M_2^2\|x_k - x^*\|^2 + L_0(\|x_{k+1} - x^*\| + \|x_k - x^*\|) \\
&= (L_0 + C_2\gamma^2 M_2^2\|x_k - x^*\|)\|x_k - x^*\| + L_0\|x_{k+1} - x^*\| \\
&\triangleq \bar{\sigma}_k, \qquad (5.26)
\end{aligned}$$

where the second inequality follows from (5.19) and Lemma 5.3, and the third and the fourth inequalities follow from (5.3) and (5.17), respectively. It is obvious that $\bar{\sigma}_k \to 0$ as $k \to \infty$.

By using (5.1), we have

$$\nabla\Phi(x^*)p_k = (\nabla\Phi(x^*) - A_k)p_k + (A_k - B_k)p_k - F(x_k).$$

This together with (5.10) yields

$$\begin{aligned}
\|p_k\| &\le \|\nabla\Phi(x^*)\|\Big(\|\nabla\Phi(x^*) - A_k\|\,\|p_k\| + \|(A_k - B_k)p_k\| + \|F(x_k)\|\Big) \\
&\le \|\nabla\Phi(x^*)\|(\bar{\sigma}_k + \delta_k)\|p_k\| + \|\nabla\Phi(x^*)\|\,\|F(x_k)\|.
\end{aligned}$$

Since $\bar{\sigma}_k \to 0$ as $k \to \infty$, the last inequality shows that there are constants $\delta' > 0$ and $C_3 > 0$ such that when $\delta_k \le \delta'$ and k is sufficiently large

$$\|p_k\| \le C_3\|\Phi(x_k)\| = \|\Phi(x_k) - \Phi(x^*)\| \le C_3 M_2\|x_k - x^*\|. \qquad (5.27)$$

From (5.1) again, we have

$$\nabla\Phi(x^*)(x_k + p_k - x^*) = \nabla\Phi(x^*)(x_k - x^*) - \Phi(x_k) - (B_k - \nabla\Phi(x^*))p_k.$$

Let $M_3 = \|\nabla\Phi(x^*)^{-1}\|$. It then follows that

$$\begin{aligned}
&\|x_k + p_k - x^*\| \\
&\le \|\nabla\Phi(x^*)^{-1}\|\,\Big\|\Phi(x_k) - \nabla\Phi(x^*)(x_k - x^*) + (B_k - \nabla\Phi(x^*))p_k\Big\| \\
&\le M_3\Big(\|\Phi(x_k) - \Phi(x^*) - \nabla\Phi(x^*)(x_k - x^*)\| + \|(B_k - A_k)p_k\| + \|(A_k - \nabla\Phi(x^*))p_k\|\Big) \\
&\le \bar{\sigma}_k'\|x_k - x^*\| + M_3(\delta_k + \bar{\sigma}_k)\|p_k\| \\
&\le (\bar{\sigma}_k' + C_3 M_2 M_3(\bar{\sigma}_k + \delta_k))\|x_k - x^*\| \\
&\triangleq (\tilde{\sigma}_k + C_3 M_2 M_3\delta_k)\|x_k - x^*\|, \qquad (5.28)
\end{aligned}$$

where $\bar{\sigma}'_k = M_3\|\Phi(x_k) - \Phi(x^*) - \nabla\Phi(x^*)(x_k - x^*)\|/\|x_k - x^*\|$, $\tilde{\sigma}_k = \bar{\sigma}'_k + C_3M_2M_3\bar{\sigma}_k)$, the third inequality follows from (5.10) and (5.26), and the last inequality follows from (5.27). So, we deduce that when $\delta_k \leq \delta'$ and k is sufficiently large

$$\begin{aligned}
\|\Phi^k(x_k + p_k)\| &\leq \|\Phi^k(x_k + p_k) - \Phi(x_k + p_k)\| + \|\Phi(x_k + p_k) - \Phi(x^*)\| \\
&\leq C_1\epsilon_k^2 + M_2\|x_k + p_k - x^*\| \\
&\leq C_1\gamma^2\|\Phi(x_k) - \Phi(x^*)\|^2 + M_2(\tilde{\sigma}_k + C_3M_2M_3\delta_k)\|x_k - x^*\| \\
&\leq (C_1\gamma^2M_2^2\|x_k - x^*\| + M_2(\tilde{\sigma}_k + C_3M_2M_3\delta_k))\|x_k - x^*\| \\
&\stackrel{\Delta}{=} (\hat{\sigma}_k + C_4\delta_k)\|x_k - x^*\|, \qquad (5.29)
\end{aligned}$$

where $\hat{\sigma}_k = C_1\gamma^2M_2^2\|x_k - x^*\| + M_2\tilde{\sigma}_k$ and $C_4 = C_3M_2^2M_3$, the second inequality follows from (5.18) and the third inequality follows from (5.3) and (5.28).

On the other hand, by the given assumptions, it is not difficult to see that (4.30) still holds. Therefore, we get that when $\delta_k \leq \delta'$ and k is sufficiently large

$$\begin{aligned}
&\|\Phi^k(x_k + p_k)\| - \rho_2\|\Phi^k(x_k)\| + \sigma_1\|p_k\|^2 \\
&\quad\leq (\hat{\sigma}_k + C_4\delta_k)\|x_k - x^*\| - \rho_2\mu\|x_k - x^*\| + \sigma_1C_3^2M_2^2\|x_k - x^*\|^2 \\
&\quad= -\rho_2\mu\|x_k - x^*\| + C_4\delta_k\|x_k - x^*\| + o(\|x_k - x^*\|),
\end{aligned}$$

where the inequality follows from (5.29), (4.30) and (5.27). It is then not difficult to see that there exist a constant $\delta \leq \delta'$ such that when $\delta_k \leq \delta$ and k is sufficiently large

$$\|\Phi^k(x_k + p_k)\| - \rho_2\|\Phi^k(x_k)\| + \sigma_1\|p_k\|^2 \leq 0.$$

This means $\lambda_k = 1$. The proof is complete. □

The following lemma shows that under Assumptions B and C, (5.12) holds even if K is infinite.

Lemma 5.6 *Let Assumptions B and C hold. Then*

$$\lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\delta_i^2 = 0. \qquad (5.30)$$

Proof It is clear that when k is sufficiently large, $x_k + \tau s_k \in U^*$ for any $\tau \in [0,1]$. It follows from Lemma 5.3 that there is a constant $L_0 > 0$ such that for any $\tau \in [0,1]$

$$\begin{aligned}
\|\nabla\Phi^{k-1}(x_k + \tau s_k) - \nabla\Phi^{k-1}(x_{k-1} + \tau s_{k-1})\| &\leq L_0\|(x_k + \tau s_k) - (x_{k-1} + \tau s_{k-1})\| \\
&\leq L_0\Big(\tau\|s_k\| + (1-\tau)\|s_{k-1}\|\Big) \\
&\leq L_0\Big(\|s_k\| + \|s_{k-1}\|\Big). \qquad (5.31)
\end{aligned}$$

Let A_k be defined by (5.9). Then we have

$$\begin{aligned}
\|A_k - A_{k-1}\| &= \Big\| \int_0^1 \nabla\Phi^k(x_k + \tau s_k)d\tau - \int_0^1 \nabla\Phi^{k-1}(x_{k-1} + \tau s_{k-1})d\tau \Big\| \\
&\le \int_0^1 \|(\nabla\Phi^k(x_k + \tau s_k) - \nabla\Phi^{k-1}(x_k + \tau s_k))d\tau\| \\
&\quad + \int_0^1 \|(\nabla\Phi^{k-1}(x_k + \tau s_k) - \nabla\Phi^{k-1}(x_{k-1} + \tau s_{k-1}))d\tau\| \\
&\le \sqrt{n}(\epsilon_{k-1} - \epsilon_k) + L_0\Big(\|s_k\| + \|s_{k-1}\|\Big) \\
&\triangleq \nu_k,
\end{aligned}$$

where the last inequality follows from (3.4) and (5.31). It is easy to deduce that

$$\sum_{k=1}^{\infty} \nu_k^2 \le 2n \sum_{k=1}^{\infty} (\epsilon_{k-1} - \epsilon_k)^2 + 4L_0^2 \sum_{k=1}^{\infty} \Big(\|s_k\|^2 + \|s_{k-1}\|^2\Big).$$

Since $\sum_{k=1}^{\infty} \|s_k\|^2 < \infty$ (see Remark (iv) after Algorithm 2) and $\{\epsilon_k\}$ is nonincreasing, it is clear that $\sum_{k=1}^{\infty} \nu_k^2 < \infty$. This implies $\sum_{k=1}^{\infty} \|A_k - A_{k-1}\|^2 < \infty$. In a way similar to Lemma 4.1 in [20], we can then prove (5.30). □

The following theorem shows superlinear convergence of Algorithm 2.

Theorem 5.3 *Let Assumptions B and C hold. Then $\{x_k\}$ converges superlinearly.*

Proof We first show that there is a constant $c > 0$ such that for all k large enough

$$\|\Phi(x_{k+1})\| \le (\omega_k + c\delta_k)\|\Phi(x_k)\|, \tag{5.32}$$

where $\{\omega_k\}$ is a positive sequence converging to zero.

Define the index set $\tilde{K} = \{k \,|\, \delta_k \le \delta \quad \text{and} \quad k \ge k'\}$, where δ and k' are constants given in Lemma 5.5. Then we get for each $k \in \tilde{K}$

$$\begin{aligned}
\|\Phi(x_{k+1})\| &\le \|\Phi(x_{k+1}) - \Phi^k(x_{k+1})\| + \|\Phi^k(x_{k+1})\| \\
&\le C_1\epsilon_k^2 + \|\Phi^k(x_{k+1})\| \\
&\le C_1\gamma^2\|\Phi(x_k) - \Phi(x^*)\|^2 + (\hat{\sigma}_k + C_4\delta_k)\|x_k - x^*\| \\
&\le (C_1\gamma^2 M_2^2\|x_k - x^*\| + \hat{\sigma}_k + C_4\delta_k)\|x_k - x^*\|,
\end{aligned} \tag{5.33}$$

where the second inequality follows from (5.18), the third inequality follows from (5.3) and (5.29), and the fourth inequality follows from (5.17). By the nonsingularity of $\nabla\Phi(x^*)$, there exists a constant $\bar{m} > 0$ such that $\|\Phi(x_k) - \Phi(x^*)\| \ge \bar{m}\|x_k - x^*\|$ for all k sufficiently large. Letting $\bar{\omega}_k = \bar{m}^{-1}(C_1\gamma^2 M_2^2\|x_k - x^*\| + \hat{\sigma}_k)$ and $C_5 = \bar{m}^{-1}C_4$, it then follows from (5.33) that

$$\|\Phi(x_{k+1})\| \le (\bar{\omega}_k + C_5\delta_k)\|\Phi(x_k)\|. \tag{5.34}$$

Clearly, $\bar{\omega}_k \to 0$ as $k \to \infty$.

By Steps 2 and 3, it is clear that

$$\|\Phi^k(x_{k+1})\| \leq \|\Phi^k(x_k)\| + \eta_k \leq \|\Phi^k(x_k)\| + \|\Phi(x_k)\|,$$

where the last inequality follows from (5.3). Also we have $\delta_k > \delta$ for any $k \notin \tilde{K}$ such that $k > k'$. Therefore, we get for all $k \notin \tilde{K}$ sufficiently large

$$\begin{aligned}
\|\Phi(x_{k+1})\| &\leq \|\Phi(x_{k+1}) - \Phi^k(x_{k+1})\| + \|\Phi^k(x_{k+1})\| \\
&\leq C_1\epsilon_k^2 + \|\Phi^k(x_k)\| + \|\Phi(x_k)\| \\
&\leq C_1\epsilon_k^2 + \|\Phi^k(x_k) - \Phi(x_k)\| + 2\|\Phi(x_k)\| \\
&\leq 2C_1\epsilon_k^2 + 2\|\Phi(x_k)\| \\
&\leq 2C_1\gamma^2\|\Phi(x_k)\|^2 + xm2\|\Phi(x_k)\| \\
&= 2(C_1\gamma^2\|\Phi(x_k)\| + 1)\|\Phi(x_k)\| \\
&\leq 2(C_1\gamma^2\|\Phi(x_k)\| + \frac{\delta_k}{\delta})\|\Phi(x_k)\|, \qquad (5.35)
\end{aligned}$$

where the second and the fourth inequalities follow from (5.18), and the fifth inequality follows from (5.3). Letting $\omega_k = \max\{\bar{\omega}_k, 2C_1\gamma^2\|\Phi(x_k)\|\}$ and $c = \max\{C_5, 2\delta^{-1}\}$, we obtain (5.32) from (5.34) and (5.35).

Next, we show the inequality

$$\sum_{k=1}^{\infty} \|x_k - x^*\| < \infty. \qquad (5.36)$$

Without loss of generality, we assume that (5.32) holds for all k. Multiplying those inequalities from $k = 1$ to t yields

$$\begin{aligned}
\|\Phi(x_{t+1})\| &\leq \|\Phi(x_1)\| \prod_{k=1}^{t} (\omega_k + c\delta_k) \\
&= \|\Phi(x_1)\| \Big\{ \prod_{k=1}^{t} (\omega_k + c\delta_k)^2 \Big\}^{\frac{1}{2}} \\
&\leq \|\Phi(x_1)\| \Big\{ \prod_{k=1}^{t} (2\omega_k^2 + 2c^2\delta_k^2) \Big\}^{\frac{1}{2}} \\
&\leq \|\Phi(x_1)\| \Big\{ \frac{1}{t} \sum_{k=1}^{t} (2\omega_k^2 + 2c^2\delta_k^2) \Big\}^{\frac{t}{2}} \\
&= \|\Phi(x_1)\| \Big\{ \frac{2}{t} \sum_{k=1}^{t} \omega_k^2 + \frac{2c^2}{t} \sum_{k=1}^{t} \delta_k^2 \Big\}^{\frac{t}{2}}.
\end{aligned}$$

Note that both terms in the last braces tend to zero as $t \to \infty$, since ω_k converges to zero and (5.30) holds. Therefore we have $\lim_{k\to\infty} \|\Phi(x_{k+1})\|^{\frac{1}{k}} = 0$ and hence $\sum_{k=1}^{\infty} \|\Phi(x_{k+1})\| <$

∞. However, $\|\Phi(x_{k+1})\| = \|\Phi(x_{k+1}) - \Phi(x^*)\| \geq \bar{m}\|x_{k+1} - x^*\|$ for all k sufficiently large. Consequently, we get (5.36).

Now, we show superlinear convergence of Algorithm 2. By virtue of (5.36), in a similar way to the proof of Lemma 5.4 in [20], we can verify that $\delta_k \to 0$ as $k \to \infty$. Therefore, by Lemma 5.5, we have $\lambda_k \equiv 1$ for all k sufficiently large. Consequently, (5.28) shows that $\{x_k\}$ converges to x^* superlinearly. □

6. Conclusion

In this paper, we proposed a smoothing Newton method and a smoothing quasi-Newton method for solving MCP. Under appropriate conditions, the proposed methods converge globally and superlinearly/quadratically. An advantage of these methods is that the dimension of the subproblems is the same as that of MCP. Moreover, the level set is bounded if the MCP involves a uniform P function. We showed an interesting property for the smoothing function. That is, every accumulation point of $\{\nabla\Phi^k(x_k)\}$ is an element of $\partial_C\Phi(x^*)$. In particular, if Φ is differentiable at x^*, which is particularly the case when the strict complementarity holds at x^*, then $\nabla\Phi^k(x_k) \to \nabla\Phi(x^*)$. A drawback of the proposed quasi-Newton method is that it is not a descent method even if F is a strongly monotone function, though it converges globally. It is interesting to construct a descent quasi-Newton method for MCP with global and superlinear convergence. We leave it as a possible further research topic.

References

[1] B. Chen, X. Chen and C. Kanzow, "A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results", Technical Report AMR 97/28, School of Mathematics, University of New South Wales, Sydney, September 1997.

[2] B. Chen and P.T. Harker, "A non-interior-point continuation method for linear complementarity problems", SIAM J. Matrix Anal. Appl., vol. 14, pp. 1168-1190, 1993.

[3] B. Chen and P.T. Harker, "Smooth approximations to nonlinear complementarity problems", SIAM J. Optim., vol. 7, pp. 403-420, 1997.

[4] C. Chen and O.L. Mangasarian, "A class of smoothing functions for nonlinear and mixed complementarity problems", Comput. Optim. Appl., vol. 5, pp. 97-138, 1996.

[5] X. Chen, Z. Nashed and L. Qi, "Convergence of Newton's method for singular smooth and nonsmooth equations using adaptive outer inverses", SIAM J. Optim., vol. 7, pp. 445-462, 1997.

[6] 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", Math. Comp., vol. 67, pp. 519-540, 1998.

[7] F.H. Clarke, Optimization and Nonsmooth Analysis, John Wiley & Sons: New York, 1983.

[8] T. De Luca, F. Facchinei and C. Kanzow, "A semismooth equation approach to the solution of nonlinear complementarity problems", Math. Programming, vol. 75, pp. 407-439, 1996.

[9] T. De Luca, F. Facchinei and C. Kanzow, "A theoretical and numerical comparison of some semismooth algorithms for complementarity problems", Mathematical Programming Technical Report 97-15, Computer Sciences Department, University of Wisconsin - Madison, Madison, WI, December 1997.

[10] J.E. Dennis, Jr. and J.J. Moré, "A characterization of superlinear convergence and its applications to quasi-Newton methods", Math. Comp., vol. 28, pp. 549-560, 1974.

[11] F. Facchinei, A. Fischer and C. Kanzow, "A semismooth Newton method for variational inequalities: The case of box constraints", in Complementarity and Variational Problems: State of the Art, M.C. Ferris and J-S. Pang (Eds.), SIAM, Philadelphia, pp. 76-90, 1997.

[12] A. Fischer, "Solution of monotone complementarity problems with locally Lipschitzian functions", Math. Programming, vol. 76, pp. 513-532, 1997.

[13] M. Fukushima and L. Qi, (Eds.), Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, Boston, 1998.

[14] S.A. Gabriel and J.J. More, "Smoothing of mixed complementarity problems", in Complementarity and Variational Problems: State of the Art, M.C. Ferris and J.S. Pang (Eds.), SIAM, Philadelphia, pp. 105-116, 1997.

[15] S.A. Gabriel and J.S. Pang, "An inexact NE/SQP method for solving the nonlinear complementarity problem", Comput. Optim. Appl., vol. 1, pp. 67-91, 1992.

[16] H. Jiang, L. Qi, X. Chen and D. Sun, "Semismoothness and superlinear convergence in nonsmooth optimization and nonsmooth equations", in Nonlinear Optimization and Applications, G. Di Pillo and F. Giannessi (Eds.), Plenum Press: New York, pp. 197-212, 1996.

[17] C. Kanzow, "Some noninterior continuation methods for linear complementarity problems", SIAM J. Matrix Anal. Appl., vol. 17, pp. 851-868, 1996.

[18] C. Kanzow and H. Jiang, "A continuation method for (strongly) monotone variational inequalities", Math. Programming, vol. 81, pp. 103-125, 1998.

[19] C. Kanzow, N. Yamashita and M. Fukushima, "New NCP-functions and their properties", J. Optim. Theory Appl., vol. 94, pp. 115-135, 1997.

[20] D.H. Li and M. Fukushima, "Broyden-like methods for variational inequality and nonlinear complementarity problems with global and superlinear convergence", Technical Report 98016, Department of Applied Mathematics and Physics, Kyoto University, Kyoto, September 1998.

[21] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization", SIAM J. Control Optim., vol. 15, pp. 957-972, 1977.

[22] J.S. Pang, "A B-differentiable equation based, globally, and locally quadratically convergent algorithm for nonlinear complementarity and variational inequality problems", Math. Programming, vol. 51, pp. 101-131, 1991.

[23] J.S. Pang and L. Qi, "Nonsmooth equations: Motivation and algorithms", SIAM J. Optim., vol. 3, pp. 443-465, 1993.

[24] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations", Math. Oper. Res., vol. 18, pp. 227-244, 1993.

[25] L. Qi, "Regular pseudo-smooth NCP and BVIP functions and globally and quadratically convergent generalized Newton methods for complementarity and variational inequality problems", Technical Report AMR 97/14, School of Mathematics, University of New South Wales, Sydney, July, 1997.

[26] L. Qi and X. Chen, "A globally convergent successive approximation method for severely nonsmooth equations", SIAM J. Control Optim., vol. 33, pp. 402-418, 1995.

[27] L. Qi and J. Sun, "A nonsmooth version of Newton's method", Math. Programming, vol. 58, pp. 353-367, 1993.

[28] L. Qi and D. Sun, "Smoothing functions and a smoothing Newton method for complementarity problems and variational inequality problems", School of Mathematics, University of New South Wales, Sydney, October 1998.

[29] L. Qi, D. Sun and G. Zhou, "A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities", AMR 97/13, Applied Mathematics Report, University of New South Wales, Sydney, June 1997.

[30] S. Smale, "Algorithms for solving the equations", in Proceedings of the International Congress of Mathematicians, Berkeley, California, pp. 172-195, 1986.

[31] D. Sun and R.S. Womersley, "A new unconstrained differentiable merit function for box constrained variational inequality problems and a damped Gauss-Newton method", to appear in SIAM J. Optim..

[32] K. Yamada, N. Yamashita and M. Fukushima, "A new derivative-free descent method for the nonlinear complementarity problem", to appear in Nonlinear Optimization and Applications 2, G. Di Pillo and F. Giannessi (Eds.), Kluwer Academic Publishers: Boston.

[33] N. Yamashita and M. Fukushima, "Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems", Math. Programming, vol. 76, pp. 469-491, 1997.

[34] N. Yamashita and M. Fukushima, "Equivalent unconstrained minimization and global error bounds for variational inequality problems", SIAM J. Control Optim., vol. 35, pp. 273-284, 1997.