Journal of Computational Mathematics, Vol.21, No.6, 2003, 747–758.

A NEW SMOOTHING EQUATIONS APPROACH TO THE NONLINEAR COMPLEMENTARITY PROBLEMS [*1)]

Chang-feng Ma
(*Department of Computational Science and Applied Physics, Guilin Institute of Electronic Technology, Guangxi 541004, China*)
(*Institute of Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Pu-yan Nie
(*Department of Mathematics, Institute of Information Science and Technology, Jinan University, Guangzhou 510623, China*)

Guo-ping Liang
(*Institute of Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract

The nonlinear complementarity problem can be reformulated as a nonsmooth equation. In this paper we propose a new smoothing Newton algorithm for the solution of the nonlinear complementarity problem by constructing a new smoothing approximation function. Global and local superlinear convergence results of the algorithm are obtained under suitable conditions. Numerical experiments confirm the good theoretical properties of the algorithm.

Key words: Nonlinear complementarity problem, Smoothing Newton method, Global convergence, Superlinear convergence.

1. Introduction

Let $F : R^n \to R^n$ be a continuously differentiable mapping and X be a nonempty closed convex set in R^n. The variational inequality problems, denoted by $\mathrm{VIP}(F, X)$, is to find a vector $x^* \in X$ such that

$$F(x^*)^T(x - x^*) \geq 0 \quad \text{for all } x \in X \tag{1.1}$$

If $X = R^n_+$, $\mathrm{VIP}(F, X)$ reduces to the nonlinear complementarity problem, denoted $\mathrm{NCP}(F)$, which is to find $x \in R^n$ such that

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

Two comprehensive surveys of variational inequality problems and nonlinear complementarity problems are [1] and [3]. The study on iterative methods for solving $\mathrm{VIP}(F, X)$ and $\mathrm{NCP}(F)$ has been rapidly developed in the last decade. One of the most popular approaches is to reformulate $\mathrm{NCP}(F)$ as an equivalent nonsmooth equation so that generalized Newton-type methods can be applied in a way similar to those for smooth equations.

Much effort has been made to construct smoothing approximation functions for approach to the solution of $\mathrm{NCP}(F)$ in recent years [2, 4, 5, 6, 7, 18, 19]. This class of algorithms, called smoothing Newton method, is due to Chen, Qi, and Sun [2]. In [2], the locally superlinear

---

* Received July 4, 2001; final revised March 3, 2002.
1) This work is supported by the National Natural Science Foundation of China.

convergence of a smoothing Newton method is established. In this paper, we will construct a new smoothing approximation function and present a new smoothing Newton method. The proposed smoothing Newton method meets the demands used in the Chen et al. in [2] and is easy to implement. We will show global and superlinear convergence of the proposed method under the same assumptions as used by Chen et al. [2] and by Qi et al. [19].

Next we introduce some words about our notation: Let $G : R^n \to R^m$ be continuously differentiable. The $\nabla G(x) \in R^{m\times n}$ denotes the Jacobian of G at a point $x \in R^n$. If $m = 1$, $\nabla G(x)$ denotes the gradient of G at a point $x \in R^n$. If is $G : R^n \to R^m$ only local Lipschitzian, we can define Clarke's [12] generalized Jacobian as follows:

$$\partial G(x) := \mathrm{conv}\{H \in R^{m\times n} | \exists \{x^k\} \subseteq D_G : x^k \to x \text{ and } G'(x^k) \to H\};$$

here D_G denotes the set of differentiable points of G and conv$S$ is the convex hull of a set S. If $m = 1$, we call $\partial G(x)$ the generalized gradient of G at x for obvious reasons.

Usually, $\partial G(x)$ is not easy to compute, especially for $m > 1$. Based on this reason, we use in this paper a kind of generalized Jacobian for the function G, denoted by $\partial_C G$ and defined as(see [13])

$$\partial_C G = \partial G_1(x) \times \partial G_2(x) \times \cdots \times \partial G_n(x),$$

where $G_i(x)$ is ith component function of G.

Furthermore, we denote by $\|x\|$ the Euclidian norm of x if $x \in R^n$ and by $\|A\|$ the spectral norm of a matrix $A \in R^{n\times n}$ which is the induced matrix norm of the Euclidian vector norm. If $A \in R^{n\times n}$ is any given matrix and $\mathcal{M} \subseteq \mathcal{R}^{\setminus\times\setminus}$ is a nonempty set of matrices, we demote by $dist(A, \mathcal{M}) := \inf_{\mathcal{B}\in\mathcal{M}} \|\mathcal{A} - \mathcal{B}\|$ the distance between A and $\mathcal{M}$.

The remainder of the paper is organized as follows: In the next section, the mathematical background and some preliminary results are summarized. The algorithm is proposed in detail in section 3. Section 4 is devoted to proving global local superlinear convergence of the algorithm. Numerical results are reported in section 5.

2. Preliminaries

In this section, we first introduce the conception of NCP-function. A function $\phi : R^2 \to R$ is called an NCP-function if $\phi(a, b) = 0$ is equivalent to $a \geq 0,\ b \geq 0,\ ab = 0$. Let us define the function $H(x) = (h_1(x), h_2(x), \cdots, h_n(x))^T$, where for each $i = 1, 2, \cdots, n$,

$$h_i(x) = \min\{x_i, F_i(x)\}. \tag{2.1}$$

Then NCP (F) can be reformulated as the following nonsmooth equation:

$$H(x) = 0. \tag{2.2}$$

Function h_i and hence H are not differentiable everywhere but semismooth in the sense of Mifflin [17] and Qi [11] if F is continuously differentiable. Denote

$$\alpha(x) \;=\; \{i : F_i(x) < x_i\}, \beta(x) = \{i : F_i(x) = x_i\}, \gamma(x) = \{i : F_i(x) > x_i\}.$$

Then we have

$$h_i(x) = \begin{cases} F_i(x), & \text{if } i \in \alpha(x) \\ \min\{x_i, F_i(x)\}, & \text{if } i \in \beta(x) \\ x_i, & \text{if } i \in \gamma(x) \end{cases}$$

By using the chain rule for generalized derivatives of Lipschitz functions(see [12]), we have the following expression of $\partial_C\Phi(x) = \partial h_1(x) \times \partial h_2(x) \times \cdots \times \partial h_n(x)$ for each $i = 1, 2, \cdots, n$,

$$\partial h_i(x) = \begin{cases} \{\nabla F_i(x)\}, & \text{if } i \in \alpha(x) \\ \{\frac{1}{2}(1+\rho)e_i, \frac{1}{2}(1-\rho)\nabla F_i(x)\}, & \text{if } i \in \beta(x) \\ \{e_i\}, & \text{if } i \in \gamma(x) \end{cases} \tag{2.3}$$

where e_i denotes the ith unit vector in R^n and parameter $\rho \in [-1, 1]$.

Smoothing techniques have attracted much attention in recent years [2,4-7,10-11,21-23]. Now let us construct a new smoothing approximation function of NCP(F) and discuss its some useful properties. we define the function $H_\mu(x) = (h_1(x,\mu), h_2(x,\mu), \cdots, h_n(x,\mu))^T$ as follows:

$$h_i(x,\mu) := \begin{cases} F_i(x), \quad i \in \alpha(x,\mu), \\ F_i(x) + \dfrac{(x_i - F_i(x) - \mu)^3}{6\mu^2}, \ i \in \beta_2(x,\mu), \\ x_i + \dfrac{(F_i(x) - x_i - \mu)^3}{6\mu^2}, \ i \in \beta_1(x,\mu), \\ x_i, \quad i \in \gamma(x,\mu) \end{cases} \tag{2.4}$$

where

$$\begin{aligned} \alpha(x,\mu) &= \{i : F_i(x) < x_i - \mu\}, \\ \beta_1(x,\mu) &= \{i : x_i - \mu \le F_i(x) \le x_i\}, \\ \beta_2(x,\mu) &= \{i : x_i < F_i(x) \le x_i + \mu\}, \\ \gamma(x,\mu) &= \{i : F_i(x) > x_i + \mu\}, \end{aligned}$$

It is not difficult to see that $h_i(x,\mu)$, for every $i = 1, 2, \cdots, n$, is continuously differentiable if $\mu > 0$ whenever F is differentiable. And it is easy to find the fact that $\alpha(x,\mu) \to \alpha(x), \beta_1(x,\mu) \cup \beta_2(x,\mu) \to \beta(x), \gamma(x,\mu) \to \gamma(x)$ as $\mu \to 0$.

It is not difficult to show the following result.

Lemma 2.1. *For any $\mu' > \mu \ge 0$, we have for all $i = 1, 2, \cdots, n$,*

$$|h_i(x,\mu') - h_i(x,\mu)| \le \frac{4}{3}(\mu' - \mu), \ \forall x \in R^n.$$

Particularly, we have for every $\mu > 0$,

$$|h_i(x,\mu) - h_i(x)| \le \frac{4}{3}\mu, \ \forall x \in R^n.$$

Lemma 2.1 shows that $h_i(x,\mu) \to h(x)$ uniformly as $\mu \to 0$. Furthermore, we have

Lemma 2.2. *For any $\mu' > \mu \ge 0$, we have,*

$$\|H_{\mu'}(x) - H_\mu(x)\| \le \frac{4\sqrt{n}}{3}(\mu' - \mu), \ \forall x \in R^n. \tag{2.5}$$

Particularly, we have for every $\mu > 0$,

$$\|H_\mu(x) - H(x)\| \le \frac{4\sqrt{n}}{3}\mu, \ \forall x \in R^n. \tag{2.6}$$

Note that $h_i(x,\mu)$, for all $i = 1, 2, \cdots, n$, is differentiable everywhere if $\mu > 0$ whenever F is differentiable. By direct deduction we obtain for each $i = 1, 2, \cdots, n$,

$$\nabla h_i(x,\mu) = \begin{cases} \nabla F_i(x), & i \in \alpha(x,\mu), \\ \nabla F_i(x) + \dfrac{(x_i - F_i(x) - \mu)^2}{2\mu^2}(e_i - \nabla F_i(x)), & i \in \beta_1(x,\mu), \\ e_i + \dfrac{(F_i(x) - x_i - \mu)^2}{2\mu^2}(\nabla F_i(x) - e_i), & i \in \beta_2(x,\mu), \\ e_i, & i \in \gamma(x,\mu). \end{cases} \tag{2.7}$$

Denote

$$a_i(x,\mu) = \begin{cases} 0, & i \in \alpha(x,\mu), \\ \dfrac{(x_i - F_i(x) - \mu)^2}{2\mu^2}, & i \in \beta_1(x,\mu), \\ 1 - \dfrac{(F_i(x) - x_i - \mu)^2}{2\mu^2}, & i \in \beta_2(x,\mu), \\ 1, & i \in \gamma(x,\mu), \end{cases}$$

$$b_i(x,\mu)=\begin{cases}1, & i\in\alpha(x,\mu),\\ 1-\dfrac{(x_i-F_i(x)-\mu)^2}{2\mu^2}, & i\in\beta_1(x,\mu),\\ \dfrac{(F_i(x)-x_i-\mu)^2}{2\mu^2}, & i\in\beta_2(x,\mu),\\ 0, & i\in\gamma(x,\mu).\end{cases}$$

Then we get

$$\nabla h_i(x,\mu)=a_i(x,\mu)e_i+b_i(x,\mu)\nabla F_i(x). \tag{2.8}$$

Let us introduce the definition of the Jacobian consistency property [2, 10].

Definition 2.3. *Let $H=(h_1,h_2,\cdots,h_n)^T: R^n\to R^n$ be a Lipschitz function in R^n. We call $H_\mu: R^n\times R_{++}\to R^n$ a smoothing approximation function of H, if H_μ is continuously differentiable and there exists a constant $c>0$ such that for any $x\in R^n$ and $\mu>0$,*

$$\|H_\mu(x)-H(x)\|\le c\mu. \tag{2.9}$$

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

$$\lim_{\mu\downarrow 0} dist(\nabla H_\mu(x),\ \partial_C H(x))=0, \tag{2.10}$$

then we say H and H_μ satisfy the Jacobian consistency property, where $\partial_C H(x)$ defined by

$$\partial_C H(x)=\partial h_1(x)\times\partial h_2(x)\times\cdots\times\partial h_n(x).$$

The inequality (2.6) implies that $H_\mu(x)$ approximates $H(x)$ uniformly. Moreover, we can show that $H(x)$ and $H_\mu(x)$ satisfy the Jacobian consistency property if F is continuously differentiable.

Lemma 2.4. *Let $x\in R^n$ be arbitrary but fixed. Then functions $H(x)$ and $H_\mu(x)$ satisfy the Jacobian consistency property, i.e.,*

$$\lim_{\mu\downarrow 0} dist(\nabla H_\mu(x),\ \partial_C H(x))=0. \tag{2.11}$$

Proof. By (2.7),

$$\nabla h_i(x,\mu)=\begin{cases}\nabla F_i(x)+\dfrac{(x_i-F_i(x)-\mu)^2}{2\mu^2}(e_i-\nabla F_i(x)), & i\in\beta_1(x,\mu),\\ e_i+\dfrac{(F_i(x)-x_i-\mu)^2}{2\mu^2}(\nabla F_i(x)-e_i), & i\in\beta_2(x,\mu).\end{cases}$$

Hence, we get

$$\nabla h_i(x,\mu)=\frac{1}{2}(e_i+\nabla F_i(x)),\quad i\in\beta(x). \tag{2.12}$$

Thus, the assertion follows from (2.3) with $\rho=0$ if $i\in\beta(x)$.

3. Algorithm

In this section, we give a detailed description of our smoothing Newton method for the nonlinear complementarity problem. In the algorithm, the subproblem is the following linear equation

$$\nabla H_{\mu_k}(x_k)d+H(x_k)=0. \tag{3.1}$$

Let $\{\eta_k\}$ be a positive sequence satisfying

$$\sum_{k=1}^{\infty}\eta_k\le\eta<\infty, \tag{3.2}$$

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

$$\|H(x_k + \lambda d_k, \mu_k)\| \leq \|H_{\mu_k}(x_k)\| - \sigma\|\lambda d_k\|^2 + \eta_k, \tag{3.3}$$

where $\sigma > 0$ is a constant and d_k is the solution of (3.1). Clearly, at each iteration k, (3.3) holds for all $\lambda > 0$ sufficiently small, since η_k is positive and independent of λ. The proposed smoothing Newton method is stated as follows.

Algorithm 3.1. *(Smoothing Newton Method)*

(S.0) Choose constants $\rho_1, \rho_2 \in (0, 1)$, $0 < \gamma < \min\{\frac{1}{3\sqrt{n}}, \frac{\rho_2}{\sqrt{n}}\}$, $\sigma_1, \sigma_2 > 0$. Select a positive sequence $\{\eta_k\}$ satisfying (3.2). Choose an initial point $x_0 \in R^n$ and a positive constant $\mu_0 \leq \frac{\gamma}{2}\|H(x_0)\|$. Set $k := 0$.

(S.1) If $\|H(x_k)\| = 0$, stop. Otherwise, solve the linear equation (3.1) to get d_k.

(S.2) If

$$\|H(x_k + d_k, \mu_k)\| \leq \rho_2\|H_{\mu_k}(x_k)\| - \sigma_1\|d_k\|^2, \tag{3.4}$$

then let $\lambda_k := 1$ and go to (S.4).

(S.3) Let λ_k be the maximum number in the set $\{1, \rho_1, \rho_1^2, \cdots,\}$ such that $\lambda = \rho_1^i$ satisfies the line search condition (3.3) with $\sigma = \sigma_2$.

(S.4) Let $x_{k+1} := x_k + \lambda_k d_k$.

(S.5) If (3.4) holds or $\gamma\|H(x_{k+1})\| \leq \mu_k$, let

$$\mu_{k+1} := \min\{\frac{\gamma}{2}\|H(x_{k+1})\|, \frac{1}{2}\mu_k\}. \tag{3.5}$$

Otherwise, let $\mu_{k+1} := \mu_k$.

(S.6) Let $k := k + 1$. Go to (S.1).

Remark 3.2. It is easy to see that if Algorithm 3.1 generates an infinite sequence $\{x_k\}$, then $H_{\mu_k}(x_k) \neq 0$ for k, and the positive sequence $\{\mu_k\}$ is not increasing and satisfies

$$\mu_k \leq \gamma\|H(x_k)\|, \ \forall k. \tag{3.6}$$

It then follows from (2.6) that

$$\mu_k \leq \gamma(\frac{4\sqrt{n}}{3}\mu_k + \|H_{\mu_k}(x_k)\|).$$

Since $\sqrt{n}\gamma < \frac{1}{3}$ by $(S.0)$, we obtain

$$\mu_k \leq \upsilon\|H_{\mu_k}(x_k)\|, \tag{3.7}$$

where $\upsilon = \gamma/(1 - \frac{4\sqrt{n}}{3}\gamma) \in (0, \frac{1}{\sqrt{n}})$.

Remark 3.3. Define index set $K = \{0\} \cup K_1 \cup K_2$, where

$$K_1 = \{k|\gamma\|H(x_{k+1})\| \leq \mu_k\}, \tag{3.8}$$

and

$$K_2 = \{k|\|H(x_{k+1}, \mu_k)\| \leq \rho_2\|H(x_k, \mu_k)\| - \sigma_1\|d_k\|^2\}, \tag{3.9}$$

Then, we have

$$\mu_{k+1} = \begin{cases} \leq \frac{1}{2}\mu_k, \text{ if } k \in K, \\ = \mu_k, \text{ if } k \notin K. \end{cases} \tag{3.10}$$

Define matrix $G_\mu(x) = (g_1(x,\mu), g_2(x,\mu), \cdots, g_n(x,\mu))^T$, where $g_i(x,\mu) \in R^n$, $i = 1, 2, \cdots, n$, are defined by

$$g_i(x,\mu) = \begin{cases} a_i(x,\mu)e_i + b_i(x,\mu)\nabla F_i(x), & i \in \alpha(x) \cup \gamma(x), \\ (\frac{1}{2} - \frac{\mu}{1+\mu^2})e_i + (\frac{1}{2} + \frac{\mu}{1+\mu^2})\nabla F_i(x), & i \in \beta(x,). \end{cases} \tag{3.11}$$

It is clear from (2.3), (2.8) and (3.11) that $g_i(x,\mu) \in \partial h_i(x)$ for each $i = 1, 2, \cdots, n$ and any $\mu \neq 0$, and $G_\mu(x) \in \partial_C H(x)$. By (2.8) and (2.12) , we get

$$\nabla h_i(x,\mu) - g_i(x,\mu) = \begin{cases} 0, & i \in \alpha(x) \cup \gamma(x), \\ \frac{\mu}{1+\mu^2}(e_i - \nabla F_i(x)), & i \in \beta(x). \end{cases} \tag{3.12}$$

Lemma 3.4. *Let $G(x,\mu)$ be defined by* (3.11). *If $\mu > 0$, then we have*

$$\|\nabla H_\mu(x) - G(x,\mu)\| \le \sqrt{n}\mu(1 + \|\nabla F(x)\|). \tag{3.13}$$

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

$$\|\nabla h_i(x,\mu) - g_i(x,\mu)\| \le \mu(1 + \|\nabla F(x)\|). \tag{3.14}$$

For $i \in \alpha(x) \cup \gamma(x)$, (3.14) holds clearly. If $i \in \beta(x)$, then we get $\frac{\mu}{1+\mu^2} \le \mu$. This together with (3.12) implies that (3.14) holds for every $i \in \beta(x)$.

It is not difficult to see from Lemma 3.4 and (3.6) that

$$dist(\nabla H_{\mu_k}(x_k), \partial_C H(x_k)) \le \sqrt{n}\gamma(1 + \|\nabla F(x_k)\|)\|H(x_k)\|. \tag{3.15}$$

4. Convergence Analysis

In this section, we discuss the global convergence of Algorithm 3.1. Firstly, we give the following assumption.

Assumption A.

(i) The level set

$$\Omega = \{x | \|H(x)\| \le 2\|H(x_0)\| + \eta\}$$

is bounded, where the positive constant η is given in (3.2);

(ii) For each $\mu > 0$, $\nabla H_\mu(x)$ is nonsingular for any $x \in \Omega$.

Lemma 4.1. *Let $\{x_k\}$ be generated by Algorithm* 3.1. *Then the following two statements are equivalent:*

$$(i)\ \lim_{k\to\infty} \inf \|H(x_k)\| = 0.$$

$$(ii)\ \lim_{k\to\infty} \inf \|H_{\mu_k}(x_k)\| = 0.$$

Proof. From (2.6) and (3.6), we get

$$\|H(x_k)\| \le \|H_{\mu_k}(x_k)\| + \frac{4\sqrt{n}}{3}\mu_k \le \|H_{\mu_k}(x_k)\| + \frac{4}{3}\sqrt{n}\gamma\|H(x_k)\|.$$

This implies,

$$\|H(x_k)\| \le \frac{3}{3 - 4\sqrt{n}\gamma}\|H_{\mu_k}(x_k)\|. \tag{4.1}$$

Noticing that, by (S.0) of Algorithm 3.1, $\sqrt{n}\gamma < 1/3$, so $3 - 4\sqrt{n}\gamma > 0$. Similarly we have

$$H_{\mu_k}(x_k)\| \le \|H(x_k)\| + \frac{4\sqrt{n}}{3}\mu_k \le \frac{3 + 4\sqrt{n}\gamma}{3}\|H(x_k)\|. \tag{4.2}$$

This together with (4.2) shows the equivalence between (i) and (ii).

Lemma 4.1 reveals that if NCP(F) has an accumulation point that is a solution of VIP(F, X), then every accumulation point of $\{x_k\}$ is a solution of NCP(F).

Lemma 4.2. *Let $\{x_k\}$ be generated by Algorithm 3.1. Then for every k, we have $\{x_k\} \subset \Omega$.*

Proof. By (2.6), we have for every k,

$$\|H_{\mu_k}(x_k)\| \leq \|H_{\mu_{k-1}}(x_k)\| + \frac{4\sqrt{n}}{3}(\mu_{k-1} - \mu_k). \tag{4.3}$$

If λ_k is determined by (S.2), then we have

$$\|H_{\mu_k}(x_k)\| \leq \rho_2 \|H_{\mu_{k-1}}(x_{k-1})\| - \sigma_1 \|x_k - x_{k-1}\|^2 + \frac{4\sqrt{n}}{3}(\mu_{k-1} - \mu_k). \tag{4.4}$$

If λ_k is determined by (S.3), then we get

$$\|H_{\mu_k}(x_k)\| \leq \|H_{\mu_{k-1}}(x_{k-1})\| - \sigma_2 \|x_k - x_{k-1}\|^2 + \frac{4\sqrt{n}}{3}(\mu_{k-1} - \mu_k) + \eta_{k-1}. \tag{4.5}$$

Notice that $\rho_2 \in (0, 1)$, for every k we have from (4.4) and (4.5),

$$\|H_{\mu_k}(x_k)\| \leq \|H_{\mu_{k-1}}(x_{k-1})\| + \frac{4\sqrt{n}}{3}(\mu_{k-1} - \mu_k) + \eta_{k-1}. \tag{4.6}$$

Summing two sides of (4.6) for k, we obtain

$$\begin{aligned} \|H_{\mu_k}(x_k)\| &\leq \|H_{\mu_0}(x_0)\| + \frac{4\sqrt{n}}{3}(\mu_0 - \mu_k) + \textstyle\sum_{i=0}^{k-1} \eta_i \\ &\leq \|H(x_0)\| + \frac{4\sqrt{n}}{3}\mu_0 + \frac{4\sqrt{n}}{3}(\mu_0 - \mu_k) + \eta \end{aligned}$$

Thus,

$$\|H(x_k)\| \leq \|H_{\mu_k}(x_k)\| + \frac{4\sqrt{n}}{3}\mu_k \leq \|H(x_0)\| + \frac{2}{3}\sqrt{n}\gamma\|H(x_0)\| + \eta \leq 2\|H(x_0)\| + \eta.$$

This completes proof of the lemma.

Lemma 4.3. *Let $\{x_k\}$ be generated by Algorithm 3.1. If the index set K defined in Remark 3.3 is infinite, then we have*

$$\lim_{k\to\infty} \|H(x_k)\| = 0, \tag{4.7}$$

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

Proof. Since $K = \{0\} \cup K_1 \cup K_2$. 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, it means (4.4) holds infinitely often. Since $\rho_2 \in (0, 1)$, this implies $\lim_{k\to\infty} \|H_{\mu_k}(x_k)\| = 0$. Then (4.7) follows from Lemma 4.1.

For the latter case, without loss of generality, let $K_1 = K = \{0 = k_0 < k_1 < k_2 < \cdots\}$. By (S.5) of Algorithm 3.1, we get

$$\|H(x_{k_j})\| \leq \gamma^{-1}\mu_{k_j - 1} = \gamma^{-1}\mu_{k_{j-1}} \leq \frac{1}{2}\|H(x_{k_{j-1}})\| \leq \cdots \leq (\frac{1}{2})^j \|H(x_0)\|$$

This shows that $\lim_{k\to\infty} \|H(x_k)\| = 0$. Hence every accumulation point of $\{x_k\}$ is a zero point of H, or equivalently, a solution of NCP(F).

Theorem 4.4. *Suppose that Assumption A holds and the sequence $\{x_k\}$ is generated by Algorithm 3.1 . Then the index set K must be infinite and hence*

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

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

Proof. Suppose to the contradiction that K is finite. Then there exists an integer $\bar{k}$ such that $\mu_k = \mu_{\bar{k}}$ for all $k \geq \bar{k}$. Denote $\mu_{\bar{k}} \equiv \bar{\mu}$. Notice that the finiteness of K implies that K_2 defined (3.9) is also finite, and hence there is an index $\tilde{k}$ such that when $k \geq \tilde{k}$, the stepsize λ_k is determined by (S.3) of Algorithm 3.1. That is, we have for every $k \geq \max\{\bar{k}, \tilde{k}\}$,

$$\|H_{\bar{\mu}}(x_k + \lambda_k d_k)\| \leq \|H_{\bar{\mu}}(x_k)\| - \sigma_2 \|\lambda_k d_k\|^2 + \eta_k. \tag{4.8}$$

By Assumption A(ii), there exists a constant $M_1 > 0$ such that for all $x \in \Omega$, $\|\nabla H_{\bar{\mu}}(x)^{-1}\| \leq M$. Then for all $k \geq \max\{\bar{k}, \tilde{k}\}$ sufficiently large ,

$$\|d_k\| = \|\nabla H_{\bar{\mu}}(x_k)^{-1} \cdot H(x_k)\| \leq M_1 \|H(x_k)\|, \tag{4.9}$$

and there is a constant $M_2 > 0$ such that for all $k \geq \max\{\bar{k}, \tilde{k}\}$ sufficiently large ,

$$\|H(x_k)\| = \|\nabla H_{\bar{\mu}}(x_k) d_k\| \leq M_2 \|d_k\|. \tag{4.10}$$

Without loss of generality, we can assume that $\{d_k\}_{k \in \bar{K}} \to \bar{d}(k \to \infty)$ and let $\bar{x}$ be an accumulation point of the corresponding sequence $\{x_k\}_{k \in \bar{K}}$, that is, $\{x_k\}_{k \in \bar{K}} \to \bar{x}(k \to \infty)$. Let $\bar{\lambda} = \limsup_{k(\in \bar{K}) \to \infty} \lambda_k$. Then $\bar{\lambda} \in [0, 1]$. If $\bar{\lambda} > 0$, then $\bar{d} = 0$. So it follows from (3.1) that $H(\bar{x}) = 0$ since (4.8) particular implies $\lambda_k d_k \to 0$ as $k \to \infty$. This is a contradiction.

We now consider the case of $\bar{\lambda} = 0$, or equivalently, $\limsup_{k(\in \bar{K}) \to \infty} \lambda_k = 0$. By (S.3) of Algorithm 3.1, when $k \in \bar{K}$ is large enough, $\lambda'_k = \lambda_k / \rho_1$ does not satisfy (4.8). Therefore, we have

$$\|H_{\bar{\mu}}(x_k + \lambda'_k d_k)\| - \|H_{\bar{\mu}}(x_k)\| > -\sigma_2 \|\lambda_k d_k\|^2.$$

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

$$2H_{\bar{\mu}}(\bar{x})^T \nabla H_{\bar{\mu}}(\bar{x}) \bar{d} \geq 0. \tag{4.11}$$

Since $H(x_k) = -\nabla H_{\mu_k}(x_k) d_k$, by (4.1), taking the limit as $k(\in \bar{K}) \to \infty$ yield $H(\bar{x}) = -\nabla H_{\bar{\mu}_k}(\bar{x})\bar{d}$. It then follows from (4.11) that

$$H_{\bar{\mu}}(\bar{x})^T H(\bar{x}) \leq 0.$$

This together with (3.6) implies,

$$\begin{aligned}(1 - \frac{16}{9} n\gamma^2)\|H(\bar{x})\|^2 &\leq \|H_{\bar{\mu}}(\bar{x})\|^2 + (1 - \frac{16}{9} n\gamma^2)\|H(\bar{x})\|^2 \\ &\leq \|H_{\bar{\mu}}(\bar{x})\|^2 + \|H(\bar{x})\|^2 - \frac{16}{9} n\bar{\mu}^2 \\ &= \|H_{\bar{\mu}}(\bar{x}) - H(\bar{x})\|^2 - \frac{16}{9} n\bar{\mu}^2 + 2H_{\bar{\mu}}(\bar{x})^T H(\bar{x}) \\ &\leq (\frac{4}{3}\sqrt{n}\bar{\mu})^2 - \frac{16}{9} n\bar{\mu}^2 + 2H_{\bar{\mu}}(\bar{x})^T H(\bar{x}) \leq 0\end{aligned} \tag{4.12}$$

where the third inequality follows from (2.6). Notice that $0 < \gamma < \frac{1}{3\sqrt{n}}$, that is, $\frac{65}{81} < 1 - \frac{16}{9} n\gamma^2 < 1$, hence (4.12) is also a contradiction. So, K must be infinite, and thus the assertion follows from Lemma 4.3.

We now turn to analyze the convergence rate of $\{x_k\}$ generated by Algorithm 3.1. Throughout this section, we assume that the sequence $\{x_k\}$ converges to a solution x^* of NCP(F), or equivalently, $H(x^*) = 0$.

Theorem 4.5. *Let Assumption A hold and $\{x_k\}$ be generated by Algorithm 3.1. Suppose that there is an accumulation point x^* of $\{x_k\}$ such that all matrices in $\partial_C H(x^*)$ are nonsingular. then $\{x_k\}$ converges to x^* superlinearly. Moreover, if ∇F is Lipschitzian at x^*, then the convergence rate is quadratic.*

Proof. Theorem 4.4 implies that the index set K defined in Remark 3.3 is infinite. By Assumption A(ii), there exists a constant M_1 such that $\|\nabla H_{\mu_k}(x_k)^{-1}\| \le M_1$ for $k \in K$ sufficiently large. Thus, we get

$$\begin{aligned}&\|x_k + d_k - x^*\| = \|x_k - x^* - \nabla H_{\mu_k}(x_k)^{-1}H(x_k)\| \\ &\le \|\nabla H_{\mu_k}(x_k)^{-1}\|\|\nabla H_{\mu_k}(x_k)(x_k - x^*) - H(x_k) + H(x^*)\| \\ &\le M_1(\|(\nabla H_{\mu_k}(x_k) - G_{\mu_k}(x_k))(x_k - x^*)\| \\ &\qquad + \|G_{\mu_k}(x_k)(x_k - x^*) - H(x_k) + H(x^*)\|) \\ &\le M_1(\sqrt{n}\gamma(1 + \|\nabla F(x)\|)\|H(x_k)\|\|x_k - x^*\| \\ &\qquad + \|H(x_k) - H(x^*) - G_{\mu_k}(x_k)(x_k - x^*)\|)\end{aligned} \tag{4.13}$$

where $G_{\mu_k}(x_k) \equiv G(x_k, \mu_k) \in \partial_C H(x_k)$ is defined by (3.11) and the third inequality follows from (3.15). Since H is semismooth at x^*, we have for $k \in K$ sufficiently large,

$$\|H(x_k) - H(x^*) - G_{\mu_k}(x_k)(x_k - x^*)\| = o(\|x_k - x^*\|). \tag{4.14}$$

Notice that $\{\|H(x_k)\|\} \to 0$, it hence together with (4.14) implies that we have

$$\|x_k + d_k - x^*\| = o(\|x_k - x^*\|) \tag{4.15}$$

for $k \in K$ sufficiently large. Furthermore, we obtain (see [11])

$$\|H(x_k + d_k)\| = o(\|H(x_k)\|) \tag{4.16}$$

when $k \in K$ is large enough. This together with (4.1) and (4.2) implies

$$\|H_{\mu_k}(x_k)\| \ge (1 - \frac{4}{3}\sqrt{n}\gamma)\|H(x_k)\| \tag{4.17}$$

and

$$\|H_{\mu_k}(x_k + d_k)\| \le (1 + \frac{4}{3}\sqrt{n}\gamma)\|H(x_k + d_k)\| = o(\|H(x_k)\|) \tag{4.18}$$

for $k \in K$ is sufficiently large. Therefore, we get that when $k \in K$ is large enough,

$$\begin{aligned}&\|H_{\mu_k}(x_k + d_k)\| - \rho_2\|H_{\mu_k}(x_k)\| + \sigma_1\|d_k\|^2 \\ &\le -\rho_2(1 - \tfrac{4}{3}\sqrt{n}\gamma)\|H(x_k)\| + o(\|H(x_k)\|) + \sigma_1 M_1^2\|H(x_k)\|^2 \\ &= -\rho_2(1 - \tfrac{4}{3}\sqrt{n}\gamma)\|H(x_k)\| + o(\|H(x_k)\|).\end{aligned} \tag{4.19}$$

It is then not difficult to see that there is an integer $\hat{k} > 0$ such that when $k \ge \hat{k}$ and $k \in K$,

$$\|H_{\mu_k}(x_k + d_k)\| - \rho_2\|H_{\mu_k}(x_k)\| + \sigma_1\|d_k\|^2 \le 0. \tag{4.20}$$

In particular, for $x_{\hat{k}+1} = x_{\hat{k}} + d_{\hat{k}}$, we obtain from (4.20)

$$\|H_{\mu_{\hat{k}}}(x_{\hat{k}+1})\| \le \rho_2\|H_{\mu_{\hat{k}}}(x_{\hat{k}})\| - \sigma_1\|d_{\hat{k}}\|^2, \tag{4.21}$$

which implies that $\hat{k} \in K$. Repeating the above process, we may prove that for all $k \ge \hat{k}$,

$$k \in K \quad \text{and} \quad x_{k+1} = x_k + d_k.$$

Thus, the superlinear convergence follows immediately from (4.15).

Moreover, if ∇F is Lipschitzian at x^*, then H is strongly semismooth at x^* so that

$$\|H(x_k) - H(x^*) - G_{\mu_k}(x_k)(x_k - x^*)\| = O(\|x_k - x^*\|^2).$$

Since H is obviously locally Lipschitzian, we further have

$$\|H(x_k)\| = \|H(x_k) - H(x^*)\| \le L\|x_k - x^*\|.$$

Hence the quadratic rate of convergence of $\{x_k\}$ to x^* follows from (4.13) by using similar argument as for the proof the local superlinear convergence.

5. Numerical Experiments

In this section we present some numerical experiments for the algorithm proposed in section 3. Throughout the computational experiments, the parameters used in Algorithm 3.1 were $\sigma_1 = \sigma_2 = 0.25,\ \rho_1 = \rho_2 = 0.9,\ \eta_k = 2^{-k}$. The stopping criterion is: $\|H(x_k)\| \le 10^{-6}$.

Example 1. *We first consider the following nonlinear complementarity problem. Test functions are of following forms:*

$$F(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}$$

This problem has two solutions on interval [0, $+\infty$):

$$x^* = (1,0,3,0)^T \quad \text{and} \quad x^{**} = (\frac{\sqrt{6}}{2}, 0, 0, \frac{1}{2})^T.$$

The numerical results are given in table 1 using different starting points.

Table 1. Numerical results of Example 1

Starting points	Num. of iter.	Starting points	Num. of iter.
$(0,\ 0,\ 0,\ 0)^T$	7^{**}	$(1,\ 1,\ 1,\ 1)^T$	4^*
$(0,1,1,1)^T$	5^{**}	$(10^2,10^2,10^2,10^2)^T$	7^{**}
$(0,1,0,1)^T$	6^{**}	$(10^5,10^5,10^5,10^5)^T$	7^*
$(1,0,1,0)^T$	5^*	$(-10^5,-10^5,-10^5,-10^5)^T$	7^*

Example 2. *We consider following linear complementarity problem(see [15, 16]):*

$$F(x) = Mx + q,$$

where the matrix M and vector q are of following forms respectively,

$$M = \begin{pmatrix} 4 & -2 & & & & \\ 1 & 4 & -2 & & & \\ & & \cdots & \cdots & \cdots & \\ & & & 1 & 4 & -2 \\ & & & & 1 & 4 \end{pmatrix}, \qquad q = \begin{pmatrix} -1 \\ -1 \\ \vdots \\ -1 \\ -1 \end{pmatrix}.$$

For different dimensions n, the numerical results are given in the following table 2 using starting point $x_0 = (0.5, 0.5, \cdots, 0.5)^T$.

Table 2. Numerical results of Example 2

Dimensions	10	40	80	160	240	320	400	480
Num. of iter.	4	4	4	4	4	4	4	4

Example 3. *Finally we consider another nonlinear complementarity problem(see [14]). This problem was tested by Kanzow with five variables defined by*

$$f_i(x) = 2(x_i - i + 2)\exp\{\sum_{i=1}^{5}(x_i - i + 2)^2\}, \quad 1 \le i \le 5.$$

This example has one degenerate solution $x^* = (0,0,1,2,3)^T$. The numerical results are given in Table 3 using different starting points.

Table 3. Numerical results of Example 3

Starting points	Num. of iter.	Final θ_value
$(1,\ 1,\ 1,\ 1, 1)^T$	7	5.2310e-14
$(-1,-1,-1,-1,-1)^T$	10	2.3864e-13
$(2,2,2,2,2)^T$	6	6.3321e-13
$(-2,-2,-2,-2,-2)^T$	25	9.4276e-14
$(3,2,1,2,3)^T$	3	2.3406e-15
$(1,0,1,3,5)^T$	5	7.8962e-14
$(0,\ 0,\ 0,\ 0,\ 0)^T$	14	1.2794e-13

Acknowledgements. The authors express their heartfelt thanks to the anonymous referees for many constructive proposals that improved the paper greatly.

References

[1] P. T. Harker and J. S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problem; A survey of theory algorithms and applications, *Math. Programming,* **48** (1990), 161-220.

[2] 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. of Computation,* **67** (1998), 519-540.

[3] J. S. Pang, Complementarity problems, in Handbook, eds., Kluwer Academic Publishers, Boston, MA, 1995, 271-338.

[4] B. Chen and P.T. Harker, Smooth approximations to nonlinear complementarity problems, *SIAM J. Optim.,* **7** (1997), 403-420.

[5] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear complementarity problems, *Comp. Appl. MAth.,* **80** (1997), 105-126.

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

[7] C. Kanzow and H. Jiang, A continuation method for (strong) monotone variational inequalities, *Math. Programming,* **81** (1998), 103-135.

[8] B. Chen and P.T. Harker, A non-interior-point continuation method for linear complementarity problems, *SIAM J. Matrix Anal. Appl.,* **14** (1993), 1168-1190.

[9] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, *SIAM J. Matrix Anal.,* **17** (1996), 178-193.

[10] C. Kanzow and H. Pieper, Jacobian Smoothing Methods for Nonlinear Complementarity Problems, *SIAM J. Optim.,* **9** (1999), 342-373.

[11] L. Qi, Convergence analysis of some algorithms for sovling nonsmooth equations, *Math. Oper. Res.,* **18** (1993), 227-244.

[12] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[13] L. Qi, C-differentiablity, C-differential Operators and Generailized Newton Methods, Tech. Report, School of Mathematics, The University of New South Wales, Sydney, Australia, Jan., 1996.

[14] H. Y. Jiang and L. Qi, A new nonsmooth equations approach to nonlinear complementarity problems, *SIAM J. Control Optim.,* **35** (1997), 178-193.

[15] K. G. Murty, Linear complementartiy, linear and nonlinear programming, Helderman, Berlin, 1988.

[16] D. Sun, A projection and contraction method for the nonlinear complementarity problem and its extensions, *Math. Numer. Sinica,* **16** (1994), 183-194.

[17] R. Mafflin, Semismooth and semiconvex functions in constrained optimization, *SIAM J. Control Optim.,* **15** (1977), 957-972.

[18] 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.

[19] L. Qi, D. Sun and G. Zhou, A new look at Smoothing Newton method for complementarity problems and box constrained variational inequality problems, AMR 97/13, AppliedMathematics Report, University of New South Wales, Sydney, June 1997.

[20] S. Smale, Algorithms for solving the equatins, in Proceeding of the International Congress of Mathematicians, Berkeley, California, 172-195, 1986.