Novi Sad J. Math.
Vol. 40, No. 2, 2010, 47-60

JACOBIAN SMOOTHING INEXACT NEWTON METHOD FOR NCP WITH A SPECIAL CHOICE OF FORCING PARAMETERS

Sanja Rapajić[1]

Abstract. The inexact Newton method with a special choice of forcing parameters is proposed for solving nonlinear complementarity problems. This method belongs to the class of Jacobian smoothing methods. Linear system is solved approximately in every iteration. The sequence of forcing terms controls the accuracy level of the approximate solution and influences the behavior of the method. Globalization strategy is based on nonmonotone rule.

AMS Mathematics Subject Classification (2000): 65H10 ; 90C33

Key words and phrases: Nonlinear complementarity problems, inexact Newton method, semismooth systems, forcing terms

1. Introduction

Nonlinear complementarity problems (NCP) are mathematical models of many engineering, economical, technological and science phenomena. These problems have also important applications in operations research and equilibrium models. For this reason, there is a growing interest in finding efficient and robust algorithms for solving them.

Most algorithms for NCP are based on suitable reformulation to equivalent nonsmooth system of nonlinear equations. Nonsmoothness complicates the application of classical Newton-type methods, so semismooth systems arising from NCP can be solved by applying generalized derivatives or various smoothing techniques.

Generalizations of classical Newton method for smooth to nonsmooth case require solving linear system in each iteration, which can be expensive if the dimension of the system is large. Therefore, it seems reasonable to use an iterative method which solves system approximately. The theory for semismooth methods can be carried over to the inexact case, where iterative solver finds the approximate solution of a linear system. Hence, the semismooth inexact Newton methods seem to be reliable algorithms for NCP (see De Luca et al. [6], Dingguo, Weiwen [7], Facchinei, Kanzow [9], Kanzow [14], Krejić, Martínez [16], Martínez, Qi [21]). The sequence of forcing parameters controls the accuracy degree and determines the behavior of the method, in the sense of convergence rate. It also has a strong influence on the robustness and efficiency of the method. Because of

[1]Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, Serbia. E-mail: sanja@dmi.uns.ac.rs

that, these parameters have a very important role in inexact Newton methods. The forcing term connects the norm of residual vector with the norm of function in the current iteration. An et al. [1] proposed a new way to choose forcing parameters, in order to improve the efficiency and robustness of the inexact Newton method in smooth case. That new choice is determined by the ratio of actual reduction and predicted reduction of the function, so it can reflect both the agreement between the local linear and nonlinear model and also the function reduction in some degree.

In this paper we consider globally convergent inexact Newton method for NCP proposed in Krejić, Rapajić [17], but with special choice of forcing terms defined in An et al. [1]. This method belongs to the class of Jacobian smoothing methods, whose main characteristic is solving mixed Newton equation, which links the semismooth function with the Jacobian of its smooth operator. Jacobian smoothing methods are introduced by many authors (see Chen et al. [3], Kanzow, Peiper [15], Krejić, Rapajić [17], Krejić et al. [18]).

The paper is organized as follows. In Section 2 we review some definitions and preliminary results to be used. Convergence analysis is presented in Section 3. Some numerical results are reported in Section 4.

2. Preliminaries

Let $F : R^n \to R^n$ be a continuously differentiable function. We consider the nonlinear complementarity problem, which is to find a vector $x \in R^n$ satisfying conditions

$$x \geq 0,\ F(x) \geq 0,\ x^\top F(x) = 0.$$

The semismooth equation reformulation of NCP given in Fischer [11] is

$$\Phi(x) = 0, \tag{2.1}$$

where $\Phi : R^n \to R^n$ is defined componentwise by

$$\Phi_i(x) := \phi(x_i, F_i(x)), \quad i = 1, 2, ..., n,$$

for Fischer-Burmeister mapping $\phi : R^2 \to R$ given with

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

One way to deal with the nonsmoothness given in Kanzow [13] is to approximate the function Φ by the smooth operator $\Phi_\mu : R^n \to R^n$, defined componentwise by

$$\Phi_i(x, \mu) := \phi_\mu(x_i, F_i(x)), \quad i = 1, 2, ..., n,$$

where $\mu > 0$ is a smoothing parameter and $\phi_\mu : R^2 \to R$ given with

$$\phi_\mu(a, b) = \sqrt{a^2 + b^2 + 2\mu} - a - b$$

is Kanzow's smooth approximation of the Fischer-Burmeister function. Function Φ_μ is smooth for any $\mu > 0$.

Some words about notation are necessary. The distance between the given matrix $A \in R^{n,n}$ and nonempty set of matrices $\mathcal{A} \subset R^{n,n}$ is denoted by $dist(A, \mathcal{A}) = \inf_{B \in \mathcal{A}} \|A - B\|$. Landau symbol $o(\cdot)$ is defined in usual way. For smooth function $F : R^n \to R^n$ we denote by F_i its ith component function, $F_i : R^n \to R$. The Jacobian of F at x is denoted by $F'(x)$.

Let us denote by $\partial_B \Phi(x)$ the B-subdifferential of Φ at x

$$\partial_B \Phi(x) = \{ \lim_{x^k \to x} \Phi'(x^k) : x^k \in D_\Phi \},$$

where D_Φ is the set where Φ is differentiable. The convex hull of B-subdifferential

$$\partial \Phi(x) = conv \partial_B \Phi(x)$$

is called the generalized Jacobian of Φ at x in the sense of Clarke [4]. Usually, computing $\partial \Phi(x)$ is complicate, so it is common to define the C-subdifferential of Φ at x by

$$\partial_C \Phi(x) = \partial \Phi_1 (x) \times \partial \Phi_2 (x) \times \ldots \times \partial \Phi_n (x),$$

where $\partial \Phi_i (x)$, $i = 1, 2, ..., n$ is the generalized gradient of Φ_i at x.

It is well known that all elements of the set $\partial_C \Phi(x)$ have the form

$$\partial_C \Phi(x) = D_a(x) + D_b(x) F'(x),$$

where $D_a(x) = diag(a_1(x), \ldots, a_n(x))$, $D_b(x) = diag(b_1(x), \ldots, b_n(x))$ are diagonal matrices with elements

$$a_i(x) = \frac{x_i}{\sqrt{x_i^2 + F_i^2(x)}} - 1, \quad b_i(x) = \frac{F_i(x)}{\sqrt{x_i^2 + F_i^2(x)}} - 1,$$

when $(x_i, F_i(x)) \neq (0, 0)$ and

$$a_i(x) = \xi_i - 1, \quad b_i(x) = \rho_i - 1, \quad (\xi_i, \rho_i) \in R^2, \quad \|(\xi_i, \rho_i)\| \leq 1,$$

for $(x_i, F_i (x)) = (0, 0)$.

We recall some background concepts which are necessary in the subsequent analysis.

Definition 2.1. *A function $F : R^n \to R^n$ is a*

- *P_0-function if for every $x, y \in R^n$, with $x \neq y$, there is an index i such that*
$$x_i \neq y_i, \quad (x_i - y_i)(F_i(x) - F_i(y)) \geq 0,$$
- *uniform P-function if there exists a positive constant c such that for every $x, y \in R^n$, there is an index i such that*
$$(x_i - y_i)(F_i(x) - F_i(y)) \geq c\|y - x\|^2.$$

Lemma 2.1. *[19] If F from NCP is a P_0-function then the Jacobian $\Phi'_\mu(x)$ is a nonsingular matrix for every $\mu > 0$ and any $x \in R^n$.*

Lemma 2.2. *[15] Assume that $\{x^k\} \subseteq R^n$ is any convergent sequence with limit point $x^* \in R^n$. If the function Φ is semismooth, then*

$$\|\Phi(x^k) - \Phi(x^*) - V_k(x^k - x^*)\| = o(\|x^k - x^*\|)$$

holds for any $V_k \in \partial_C \Phi(x^k)$.

Lemma 2.3. *[10] Let $\Phi : R^n \to R^n$ be a semismooth function and let $x^* \in R^n$ be a solution of $\Phi(x) = 0$ such that all elements of $\partial\Phi(x^*)$ are nonsingular. Suppose that two sequences $\{x^k\}$ and $\{s^k\}$ are given in such a way that $\{x^k\} \to x^*$ and $\|x^k + s^k - x^*\| = o(\|x^k - x^*\|)$. Then*

$$\|\Phi(x^k + s^k)\| = o(\|\Phi(x^k)\|).$$

Lemma 2.4. *[22] Suppose that x^* is the solution of (2.1). If Φ is semismooth at x^* and all elements of $\partial_B \Phi(x^*)$ are nonsingular, then there exists a neighborhood of x^* such that x^* is the unique solution of (2.1) in it.*

Some basic properties of the smoothing procedure are introduced in Kanzow, Peiper [15], where for any $x \in R^n$ it is proved that

$$\lim_{\mu \to 0} dist\left(\Phi'_\mu(x), \partial_C \Phi(x)\right) = 0,$$

which means that the function Φ_μ has the Jacobian consistency property. The sequence of smoothing parameters leads to the smoothing procedure and the precise definition of the threshold value for the smoothing parameters is also given in [15].

3. The algorithm and convergence results

Jacobian smoothing inexact Newton method with nonmonotone line-search rule is introduced in Krejić, Rapajić [17], where global convergence and local q-superlinear convergence is analyzed. Since inexact directions are not descent directions in general, a nonmonotone technique is used for globalization.

In this paper we will consider the same algorithm, but with the special choice of forcing parameters, proposed in An et al. [1].

Globalization procedure for solving NCP is based on unconstrained minimization of the merit function. Merit function is a function whose global minima are coincident with the solutions of the original NCP. The major advantage of the Fischer-Burmeister reformulation is that the merit function

$$\Psi(x) = \frac{1}{2}\|\Phi(x)\|^2, \quad \Psi : R^n \to R \tag{3.1}$$

is smooth although the original operator Φ is nonsmooth. The globalization can be done in a standard way, using the related merit function

$$\Psi_\mu(x) = \frac{1}{2}\|\Phi_\mu(x)\|^2, \quad \Psi_\mu : R^n \to R.$$

It was shown in Facchinei, Soares [10] that if F is a P_0-function, then every stationary point of Ψ is a global minima of the unconstrained problem $\min_{x\in R^n}\Psi(x)$ and hence solves the NCP. Moreover, the level sets of Ψ, defined by $\mathcal{L}(\alpha)=\{x\in R^n,\quad \Psi(x)\le\alpha\}$ are bounded, under the assumption that F is an uniform P-function.

The algorithm is described as follows.

Algorithm JSIN: Jacobian Smoothing Inexact Newton Method

S0: Choose $\sigma,\alpha,\bar{\xi}\in(0,1)$, $0<\tau_{\min}<\tau_{\max}<1$, $\gamma>0$, $t,\theta\in[0,1)$, such that $t<\frac{1-\alpha}{1+\alpha}-\sigma(1-\theta)(1+\alpha)$, $\varepsilon\ge 0$ and $x^0\in R^n$. Let $\{\eta_k\}>0$ be a sequence such that $\sum_{k=0}^{\infty}\eta_k\le\eta<\infty$ and $\{t_k\}$, $0\le t_k\le t$. Set $\beta_0=\|\Phi(x^0)\|$, $\mu_0=(\frac{\alpha\beta_0}{2\sqrt{2n}})^2$ and $k=0$.

S1: If $\|\Phi(x^k)\|\le\varepsilon$ STOP.

S2: Compute s^k from the mixed Newton equation

$$\Phi'_{\mu_k}(x^k)s^k=-\Phi(x^k)+\bar{r}^k, \tag{3.2}$$

where

$$\|\bar{r}^k\|\le t_k\|\Phi(x^k)\|. \tag{3.3}$$

S3: Set $\tilde{\alpha}=1$. If

$$\Psi_{\mu_k}(x^k+\tilde{\alpha}s^k)\le(1+\tilde{\alpha}\sigma(\theta-1))^2\Psi_{\mu_k}(x^k)+\eta_k, \tag{3.4}$$

set $\alpha_k=\tilde{\alpha}$ and $x^{k+1}=x^k+\alpha_k s^k$.

If (3.4) is not satisfied, choose $\alpha_{new}\in[\tilde{\alpha}\tau_{\min},\tilde{\alpha}\tau_{\max}]$, set $\tilde{\alpha}=\alpha_{new}$ and repeat (3.4).

S4: If

$$\|\Phi(x^{k+1})\|\le\max\{\bar{\xi}\beta_k,\frac{1}{\alpha}\|\Phi(x^{k+1})-\Phi_{\mu_k}(x^{k+1})\|\} \tag{3.5}$$

then

$$\beta_{k+1}=\|\Phi(x^{k+1})\|$$

and choose μ_{k+1} such that

$$0<\mu_{k+1}\le\min\{(\frac{\alpha\beta_{k+1}}{2\sqrt{2n}})^2,\frac{\mu_k}{4},\frac{\mu_k^2}{\|\Phi_{\mu_k}(x^{k+1})\|^2},\bar{\mu}(x^{k+1},\gamma\beta_{k+1})\}. \tag{3.6}$$

If (3.5) does not hold then

$$\beta_{k+1}=\beta_k,\quad \mu_{k+1}=\mu_k.$$

S5: Set $k:=k+1$ and return to step S1.

The above algorithm is globally convergent.

Theorem 3.1. *[17] Assume that F is a uniform P-function and $\{x^k\}$ is a sequence generated by Algorithm JSIN. Then every accumulation point of the sequence $\{x^k\}$ is a stationary point of Ψ.*

Since solving the NCP is equivalent to unconstrained minimization of Ψ, Theorem 3.1 and the fact that any stationary point of Ψ is its global minima, imply that every accumulation point of the sequence generated by Algorithm JSIN is a solution of the NCP.

Step S3 of the Algorithm indicates that the forcing parameter t_k is used to control the accuracy of approximate solution obtained by some linear solver in the kth iteration. As we mentioned before, the sequence of forcing terms has an important role. It determines the convergence rate and also reflects the efficiency and robustness of the algorithm. If the current iteration x_k is away from the solution x_*, then choosing t_k too small may lead to "oversolving" mixed Newton equation. This may result in little or no reduction of the original function, but requires additional costs. Thus, the main purpose is to choose forcing parameters that achieve a desirably fast local convergence and also tend to avoid "oversolving". A good choice of forcing terms is of great importance and it should be related to specific problems. Many researchers have proposed some promising strategies for practical computing forcing parameters (see Dembo, Steihaug [5], Eisenstat, Walker [8]) and most of them use some information about the original function. A new way of choosing forcing terms, proposed in An et al. [1], reflects not only the agreement between the local linear model and nonlinear model, but also the reduction rate of $\|\Phi(x)\|$ in some degree. It is determined by the ratio of the actual reduction to the predicted reduction. Let

$$r_k = \frac{Ared_k(s^k)}{Pred_k(s^k)},$$

where actual reduction $Ared_k(s^k)$ and predicted reduction $Pred_k(s^k)$ of the function $\|\Phi(x)\|$ at x^k with step s^k are given with

$$Ared_k(s^k) = \|\Phi(x^k)\| - \|\Phi(x^k + s^k)\|,$$

$$Pred_k(s^k) = \|\Phi(x^k)\| - \|\Phi(x^k) + \Phi'_{\mu_k}(x^k)s^k\|.$$

In this choice, forcing parameters t_k are defined by r_k. If $r_k \approx 0$, then the local linear model and nonlinear model disagree and the function $\|\Phi(x)\|$ will be reduced very little or will be enlarged, so this case is the worst. If $r_k \approx 1$, then the local linear model and nonlinear model will agree well and the function $\|\Phi(x)\|$ will be reduced obviously, so the case $r_k \approx 1$ is the best one. According to this, we choose forcing terms in the following way.

Let t_0 be given such that

$$0 < t_0 < \frac{1-\alpha}{1+\alpha} - \sigma(1-\theta)(1+\alpha), \tag{3.7}$$

where the parameters α, σ and θ are chosen as in step S0 of Algorithm JSIN. For $k = 1, 2, ...$ define

$$t_k = \begin{cases} 1 - 2p_1, & r_{k-1} < p_1 \\ t_{k-1}, & p_1 \le r_{k-1} < p_2 \\ 0.8t_{k-1}, & p_2 \le r_{k-1} < p_3 \\ 0.5t_{k-1}, & p_3 \le r_{k-1}, \end{cases} \tag{3.8}$$

where

$$r_k = \frac{Ared_k(s^k)}{Pred_k(s^k)} = \frac{\|\Phi(x^k)\| - \|\Phi(x^k + s^k)\|}{\|\Phi(x^k)\| - \|\Phi(x^k) + \Phi'_{\mu_k}(x^k)s^k\|},$$

$$0 < p_1 < p_2 < p_3 < \frac{1-\alpha}{1+\alpha} - \sigma(1-\theta)(1+\alpha),$$

and

$$p_1 \in (\frac{1}{2} - \frac{1-\alpha}{2(1+\alpha)} + \frac{\sigma(1-\theta)(1+\alpha)}{2}, \frac{1}{2}). \tag{3.9}$$

This strategy of choosing forcing parameters determines t_k by using r_{k-1}. If $r_{k-1} < p_1$, i.e. r_{k-1} is near 0, then the local linear model and nonlinear model can not agree well, which means the worst case, and because of that the accuracy of approximate solution should be relaxed, i.e. $t_k = 1-2p_1$ is relatively large. If $r_{k-1} \ge p_2$, i.e. r_{k-1} is near 1, then the local linear model and nonlinear model agree well and $\|\Phi(x)\|$ will be reduced, which means the best case, and because of that the mixed Newton equation should be solved more accurately, i.e. t_k should be smaller. Otherwise, t_k remains the same.

The previous value r_{k-1} determines the current forcing parameter t_k, while t_k determines the current value r_k through approximate solving mixed Newton equation. In this way, the sequences $\{t_k\}$ and $\{r_k\}$ are interrelated.

Now, we are going to prove that JSIN method with this choice of forcing terms is globally convergent.

Theorem 3.2. *Assume that F is a uniform P-function and $\{x^k\}$ is a sequence generated by Algorithm JSIN with the forcing terms defined by (3.7) and (3.8). Then every accumulation point of the sequence $\{x^k\}$ is a stationary point of Ψ.*

Proof. It suffices to show that the sequence t_k defined by (3.7) and (3.8) satisfies the conditions given in step S0 of Algorithm JSIN, i.e.

$$0 \le t_k \le t < \frac{1-\alpha}{1+\alpha} - \sigma(1-\theta)(1+\alpha), \quad k = 0, 1, 2, ... \tag{3.10}$$

Choose $t = \max\{t_0, 1 - 2p_1\}$. First, we show that

$$t < \frac{1-\alpha}{1+\alpha} - \sigma(1-\theta)(1+\alpha). \tag{3.11}$$

If $t = t_0$, then (3.11) is true because of (3.7). If $t = 1 - 2p_1$, then (3.11) follows from (3.9). We are going to prove by the induction

$$0 \leq t_k \leq t, \quad k = 0, 1, 2, ... \tag{3.12}$$

Obviously, it is true for $k = 0$. As an inductive hypothesis, suppose that

$$0 \leq t_{k-1} \leq t$$

for some $k \geq 1$. If $t_k = 1 - 2p_1$, then (3.12) follows trivially from (3.7) and (3.9). If $t_k = t_{k-1}$ or $t_k = 0.8t_{k-1}$ or $t_k = 0.5t_{k-1}$, then induction hypothesis implies (3.12). Thus, (3.10) is satisfied, which means that the Algorithm JSIN can be applied with this choice of t_k and Theorem 3.1 holds for it. □

To obtain superlinear convergence of Jacobian smoothing inexact Newton method with the given choice of forcing terms and the special choice of sequence $\{\eta_k\}$ we need some results from [15].

Let us define the set

$$K = \{0\} \cup \left\{ k, k \in N;\ \|\Phi(x^k)\| \leq \max\{\bar{\xi}\beta_{k-1}, \frac{1}{\alpha}\|\Phi(x^k) - \Phi_{\mu_{k-1}}(x^k)\|\} \right\}.$$

Lemma 3.1. *[15] Let $\{x^k\}$ be a sequence generated by Algorithm JSIN. Then the following statements hold*

$$\|\Phi(x^k) - \Phi_{\mu_k}(x^k)\| \leq \alpha\|\Phi(x^k)\|, \quad \textit{for} \quad k \geq 0,$$

$$dist\left(\Phi'_{\mu_k}(x^k), \partial_C\Phi(x^k)\right) \leq \gamma\|\Phi(x^k)\|, \quad \textit{for} \quad k \geq 1,\ k \in K.$$

Lemma 3.2. *[15] Let $\{x^k\}$ be a sequence generated by Algorithm JSIN. Assume that $\{x^k\}$ has an accumulation point x^* which is a solution of NCP. Then the index set K is infinite and $\{\mu_k\} \to 0$.*

Using the same idea as in Krejić, Rapajić [17] the next theorem can be proved.

Theorem 3.3. *Assume that F is a uniform P-function. Let x^* be an accumulation point of a sequence $\{x^k\}$ generated by Algorithm JSIN with a special choice of forcing terms defined by (3.7), (3.8) and with the sequence $\{\eta_k\}$,*

$$\eta_k = (2 + \sigma(\theta - 1))^2 n\mu_k + (2 + \sigma(\theta - 1))\sqrt{2n\mu_k}(1 + \sigma(\theta - 1))\|\Phi_{\mu_k}(x^k)\|.$$

Assume that all elements of $\partial_C\Phi(x^*)$ are nonsingular. Then x^* is a solution of NCP and the sequence $\{x^k\}$ converges q-superlinearly to x^*.

Proof. From the statement of Theorem 3.2, it follows that every accumulation point of the sequence $\{x^k\}$ generated by JSIN with the forcing terms t_k given with (3.7) and (3.8) is a stationary point of Ψ, so it is clear that x^* is a solution of NCP. The fact that $\partial_B\Phi(x^*) \subseteq \partial_C\Phi(x^*)$ and Lemma 2.4 imply

that there is a neighborhood of the solution such that x^* is a unique solution in it. Since x^* is an accumulation point and also a solution of NCP, there follows from Lemma 3.2 that K is an infinite set. Thus, there is a subsequence K_0 of K such that $\{x^k\}_{K_0} \to x^*$.

The sequence $\{\eta_k\}$ satisfies $\sum_{k=0}^{\infty} \eta_k \leq \eta < \infty$, which is proved in Krejić, Rapajić [17]. Using the same technique as in Chen et al. [3] with the fact that $\Phi'_{\mu_k}(x^k)$ is nonsingular matrix because of Lemma 2.1, it is easy to prove that

$$\|\Phi'_{\mu_k}(x^k)^{-1}\| \leq M, \tag{3.13}$$

for all $k \in K_0$ large enough and some positive constant M. Since $\partial_C \Phi(x^k)$ is nonempty and compact set, there exists $V_k \in \partial_C \Phi(x^k)$ such that

$$dist(\Phi'_{\mu_k}(x^k), \partial_C \Phi(x^k)) = \|\Phi'_{\mu_k}(x^k) - V_k\|. \tag{3.14}$$

Second statement of Lemma 3.1 implies

$$\|\Phi'_{\mu_k}(x^k) - V_k\| \leq \gamma\beta_k, \quad k \in K_0. \tag{3.15}$$

By inexact Newton conditions (3.2) and (3.3), there follows

$$\|s^k\| \leq (1 + t_k)\|\Phi'_{\mu_k}(x^k)^{-1}\|\|\Phi(x^k)\|, \tag{3.16}$$

$$\|\Phi(x^k)\| - \|\Phi'_{\mu_k}(x^k)s^k + \Phi(x^k)\| \geq (1 - t_k)\|\Phi(x^k)\|, \tag{3.17}$$

and Lemma 2.2 implies

$$\|\Phi(x^k) - \Phi(x^k + s^k) + V_k s^k\| = o(\|s^k\|), \;\; V_k \in \partial_C \Phi(x^k). \tag{3.18}$$

We are going to prove that $\lim_{k\to\infty} t_k = 0$. It suffices to show that $r_k \geq p_3$ for all sufficiently large $k > \bar{k} \in K_0$.

Since $\beta_k \to 0$, it follows

$$\beta_k \leq \frac{1}{2\gamma M} \tag{3.19}$$

for $k \in K_0$ large enough, because of the algorithm construction and the fact that K is an infinite set. Choosing $\delta < 1$ such that

$$\delta + \frac{1}{2M} = \frac{(1 - p_3)(1 - t)}{M(1 + t)}, \tag{3.20}$$

$V_k \in \partial_C \Phi(x^k)$ such that (3.14) holds and using (3.16)-(3.20) there follows

$$\begin{aligned}
r_k &= \frac{\|\Phi(x^k)\| - \|\Phi(x^k + s^k)\|}{\|\Phi(x^k)\| - \|\Phi(x^k) + \Phi'_{\mu_k}(x^k)s^k\|} \\
&= \frac{\|\Phi(x^k)\| - \| - \Phi(x^k + s^k) \pm \Phi'_{\mu_k}(x^k)s^k \pm \Phi(x^k) \pm V_k s^k\|}{\|\Phi(x^k)\| - \|\Phi(x^k) + \Phi'_{\mu_k}(x^k)s^k\|} \\
&\geq 1 - \frac{\|\Phi(x^k) - \Phi(x^k + s^k) + V_k s^k\| + \|(V_k - \Phi'_{\mu_k}(x^k))s^k\|}{\|\Phi(x^k)\| - \|\Phi(x^k) + \Phi'_{\mu_k}(x^k)s^k\|} \\
&\geq 1 - \frac{\|\Phi(x^k) - \Phi(x^k + s^k) + V_k s^k\| + \|V_k - \Phi'_{\mu_k}(x^k)\|\|s^k\|}{(1 - t_k)\|\Phi(x^k)\|} \\
&\geq 1 - \frac{(\delta + \gamma\beta_k)\|s^k\|}{(1 - t_k)\|\Phi(x^k)\|} \\
&\geq 1 - \frac{(\delta + \gamma\beta_k)M(1 + t)}{1 - t} \\
&\geq p_3
\end{aligned}$$

for every $k > \bar{k} \in K_0$ sufficiently large. Thus, the definition of forcing terms shows that $\lim_{k\to\infty} t_k = 0$. This fact with Lipschitz continuity of Φ, Lemma 2.2, (3.13), (3.15) and the construction of Algorithm JSIN imply

$$\begin{aligned}
\|x^k + s^k - x^*\| &= \|x^k - x^* - \Phi'_{\mu_k}(x^k)^{-1}(\Phi(x^k) - \bar{r}^k)\| \\
&\leq \|\Phi'_{\mu_k}(x^k)^{-1}\|(\|\Phi(x^k) - \Phi(x^*) - V_k(x^k - x^*)\| \\
&\quad + \|V_k - \Phi'_{\mu_k}(x^k)\|\|x^k - x^*\| + \|\bar{r}^k\|) \\
&\leq \|\Phi'_{\mu_k}(x^k)^{-1}\|(\|\Phi(x^k) - \Phi(x^*) - V_k(x^k - x^*)\| \\
&\quad + \|V_k - \Phi'_{\mu_k}(x^k)\|\|x^k - x^*\| + t_k\|\Phi(x^k)\|) \\
&\leq M(o(\|x^k - x^*\|) + \gamma\beta_k\|x^k - x^*\| + t_k\|\Phi(x^k) - \Phi(x^*)\|) \\
&= o(\|x^k - x^*\|) \qquad (3.21)
\end{aligned}$$

for $k \in K_0$ sufficiently large, since $\beta_k \to 0$, where $V_k \in \partial_C \Phi(x^k)$ is chosen such that (3.14) holds. Hence by (3.21) and Lemma 2.3 there follows

$$\|\Phi(x^k + s^k)\| = o(\|\Phi(x^k)\|) \qquad (3.22)$$

for $k \in K_0$, $k \to \infty$. In the same way as in [17], it can be proved that there exists an index $\breve{k} \in K_0$ such that for every $k \geq \breve{k}$, $k \in K_0$ the index $k + 1$ belongs to the set K_0 and $x^{k+1} = x^k + s^k$. This fact and (3.21) imply q-superlinear convergence. □

4. Numerical results

Some numerical experiments are reported in this section. We compared Jacobian Smoothing Inexact Newton methods with four different choices of forcing

terms. The following notation is used.

- JSIN1 - Jacobian Smoothing Inexact Newton method with $t_k = 0.5$
- JSIN2 - Jacobian Smoothing Inexact Newton method with $t_k = 2^{-k}$
- JSIN3 - Jacobian Smoothing Inexact Newton method with special choice of t_k defined by (3.7) and (3.8)
- JSIN4 - Jacobian Smoothing Inexact Newton method with $t_k = \|\Phi(x^k)\|$.

The main stopping criterion was

$$\|\Phi(x^k)\| \leq 10^{-5}\sqrt{n}.$$

Algorithms were stopped after $k_{max} = 200$ iterations if the stopping criterion was not satisfied. Test problems are generated in the usual way, proposed by Gomes-Ruggiero et al. [12].

Let $x^* = (1, 0, 1, 0, ...)^\top \in R^n$. For $i = 1, 2, ..., n$ set

$$F_i(x) = \begin{cases} f_i(x) - f_i(x^*), & \text{if } \ i \ \text{ odd or } \ i > r \\ f_i(x) - f_i(x^*) + 1, & \text{otherwise} \end{cases}$$

where $r \geq 0$ is an integer and $f(x) = (f_1(x), f_2(x), ..., f_n(x))^\top$ is a smooth nonlinear mapping from R^n to R^n, given with all test problems proposed in Lukšan [20], problems 2, 4, 6, 7, 12, 13, 25 and 27 from Spedicato, Huang [23] and problems 1.1, 1.2, 1.3 and 1.5 from Bus [2]. Vector x^* is a solution of the NCP defined with this function F, but not necessarily its unique solution. If $r < n$, x^* is a degenerate solution of NCP, while for $r = n$ it is a nondegenerate one. All examples are tested in three dimensions $n = 10$, $n = 100$, $n = 1000$, with two different starting points. The first one, denoted by x^0, is suggested in Bus [2], Lukšan [20], Spedicato, Huang [23] and the second one denoted by $\tilde{x}^0$ is defined by

$$\tilde{x}_i^0 = \begin{cases} 10x_i^0, & \text{if } \ x_i^0 \neq 0 \\ 10, & \text{otherwise.} \end{cases}$$

We consider the degenerate solution ($r = n/2$) and the nondegenerate one ($r = n$) for each dimension of the problem. In our implementation we used GMRES as linear solver and the following parameters $\sigma = 10^{-4}$, $\alpha = 0.1$, $\bar{\xi} = 0.5$, $\gamma = 20$, $\theta = 0.8$, $\tau_{\min} = 0.3$, $\tau_{\max} = 0.8$ in all methods and $p_1 = 0.1$, $p_2 = 0.4$, $p_3 = 0.7$ in JSIN3 method. The obtained results are compared using three indices: the index of robustness, the efficiency index and the combined robustness and efficiency index. The robustness index is defined by

$$R_j = \frac{t_j}{n_j},$$

the efficiency index is

$$E_j = \sum_{i=1, r_{ij} \neq 0}^{m} (\frac{r_{ib}}{r_{ij}})/t_j,$$

and the combined index is

$$E_j \times R_j = \sum_{i=1, r_{ij} \neq 0}^{m} (\frac{r_{ib}}{r_{ij}})/n_j,$$

where r_{ij} is the number of iterations required to solve the problem i by the method j, $r_{ib} = \min_j r_{ij}$, t_j is the number of successes by method j and n_j is the number of problems attempted by method j.

The results of all tested methods are given in the following tables.

	JSIN1	JSIN2	JSIN3	JSIN4
R	0.6956	0.8116	0.7826	0.6956
E	0.7895	0.8860	0.8550	0.8009
E × R	0.5492	0.7191	0.6691	0.5572

Table 1. NONDEGENERATE CASE ($r = n$) WITH STARTING POINT x^0

	JSIN1	JSIN2	JSIN3	JSIN4
R	0.7027	0.8243	0.7838	0.6351
E	0.7143	0.9032	0.8388	0.5704
E × R	0.5020	0.7445	0.6574	0.3623

Table 2. NONDEGENERATE CASE ($r = n$) WITH STARTING POINT $\bar{x}^0$

	JSIN1	JSIN2	JSIN3	JSIN4
R	0.6805	0.7639	0.7639	0.6389
E	0.5870	0.8299	0.7851	0.7646
E × R	0.3994	0.6340	0.5997	0.4885

Table 3. DEGENERATE CASE ($r = n/2$) WITH STARTING POINT x^0

	JSIN1	JSIN2	JSIN3	JSIN4
R	0.6479	0.7746	0.7183	0.5211
E	0.5620	0.8570	0.7459	0.6493
E × R	0.3641	0.6638	0.5358	0.3384

Table 4. DEGENERATE CASE ($r = n/2$) WITH STARTING POINT $\bar{x}^0$

Jacobian smoothing inexact Newton method with a special choice of forcing parameters and the sequence $\{\eta_k\}$ is q-superlinearly convergent. Numerical results show that the JSIN3 has better performance than JSIN1 and JSIN4 on this collection of test problems. The JSIN3 is close to JSIN2, which is the best method among all tested inexact methods. These results confirm theoretical expectations in the sense of convergence rate. In the case of degenerate solution, the JSIN2 and JSIN3 have the same robustness index, which can be seen from Table 3.

The special choice of forcing parameters results in a desirably fast local convergence and tends to avoid "oversolving". Choosing a very small forcing term may risk needless expenses in obtaining an unnecessarily accurate approximate solution of mixed Newton equation. The disagreement between the local linear model and nonlinear model when "oversolving" occurs, may require significant work from globalization process and even cause it to fail. In addition, a very small forcing term may require more residual reduction than the linear iterative solver such as GMRES can accurately provide. This is the reason for many failures of JSIN4, which should be the best method from the theoretical point of view. On the other hand, choosing a larger forcing term may reduce "oversolving" and avoid inaccuracy in the linear iterative solver, but increase the number of inexact Newton steps necessary for convergence.

References

[1] An, H.B., Mo, Z.Y., Liu, X.P., A choice of forcing terms in inexact Newton method. Journal of Computational and Applied Mathematics, 200 (2007), 47-60.

[2] Bus, J.C.P., Numerical Solution of Systems of Nonlinear Equations. Mathematical Centre Tracts 122, Amsterdam, 1980.

[3] Chen, X., Qi, L., Sun, D., Global and superlinear convergence of the smoothing Newton method and its applications to general box constrained variational inequalities. Mathematics of Computation, 67 (1998), 519-540.

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

[5] Dembo, R., Steihaug, T., Truncated Newton algorithms for large-scale optimization. Mathematical Programming, 26 (1983), 190-212.

[6] De Luca, T., Facchinei, F., Kanzow, C., A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. Computational Optimization and Applications, 16 (2000), 173-205.

[7] Dingguo, P., Weiwen, T., Globally convergent inexact genaralized Newton's methods for nonsmooth equations. Journal of Computational and Applied Mathematics, 138 (2002), 37-49.

[8] Eisenstat, S.C., Walker, H.F., Choosing the forcing term in an inexact Newtin method. SIAM J. Sci. Comput., 17 (1996), 16-32.

[9] Facchinei, F., Kanzow, C., A nonsmooth inexact Newton method for the solution of large scale complementarity problems. Mathematical Programming, 76 (1997), 493-512.

[10] Facchinei, F., Soares, J., A new merit function for nonlinear complementarity problems and a related algorithm. SIAM Journal on Optimization, 7 (1997), 225-247.

[11] Fischer, A., A special Newton-type optimization method. Optimization, 24 (1992), 269-284.

[12] Gomes-Ruggiero, M.A., Martínez, J.M., Santos, S.A., Solving nonsmooth equations by means of Quasi-Newton methods with globalization. Recent Advances in Noonsmooth Optimization, World Scientific Publisher, Singapore, 1995, 121-140.

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

[14] Kanzow, C., Inexact semismooth Newton methods for large-scale complementarity problems. Optimization Methods and Software, 19 (2004), 309-325.

[15] Kanzow, C., Peiper, H., Jacobian smoothing methods for general nonlinear complementarity problems. SIAM Journal on Optimization, 9 (1999), 342-373.

[16] Krejić, N., Martínez, J.M., Inexact-Newton methods for semismooth systems of equations with block-angular structure. Journal of Computational and Applied Mathematics, 103 (1999), 239-249.

[17] Krejić, N., Rapajić, S., Globally convergent Jacobian smoothing inexact Newton methods for NCP. Computational Optomization and Applications, 41 (2008), 243-261.

[18] Krejić,N., Lužanin, Z., Rapajić, S., Jacobian smoothing Brown's method for NCP", Nonlinear Analysis, 70 (2009), 642-657.

[19] Li, D., Fukushima, M., Globally convergent Broyden-like methods for semismooth equations and applications to VIP, NCP and MCP. Annals of Operations Research, 103 (2001), 71-97.

[20] Lukšan, L., Inexact trust region method for large sparse systems of nonlinear equations. Journal of Optimization Theory and Applications, 81 (1994), 569-590.

[21] Martínez, J.M., Qi, L., Inexact Newton's method for solving nonsmooth equations. Journal of Computational and Applied Mathematics, 60 (1995), 127-145.

[22] Qi, L., Convergens analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res., 18 (1993), 227-244.

[23] Spedicato, E., Huang, Z., Numerical experience with Newton-like methods for nonlinear algebraic systems. Computing, 58 (1997), 69-89.

Received by the editors February 17, 2010