

# A NEW TRUST REGION METHOD FOR NONSMOOTH EQUATIONS

Y. F. YANG<sup>1</sup>

(Received 10 March, 2000; revised 3 December, 2000)

## Abstract

We propose a new trust region algorithm for solving the system of nonsmooth equations  $F(x) = 0$  by using a smooth function satisfying the Jacobian consistency property to approximate the nonsmooth function  $F(x)$ . Compared with existing trust region methods for systems of nonsmooth equations, the proposed algorithm possesses some nice convergence properties. Global convergence is established and, in particular, locally superlinear or quadratical convergence is obtained if  $F$  is semismooth or strongly semismooth at the solution.

## 1. Introduction

In this paper, we consider the system of nonsmooth equations

$$F(x) = 0, \tag{1.1}$$

where  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is locally Lipschitz continuous but not differentiable.

Many methods have been developed for solving the above nonsmooth system, see for example [1, 6, 13–17]. Some of these methods have established locally superlinear or quadratic convergence, see for example [1, 6, 14, 17].

If  $F$  is smooth, then damped Newton methods are a class of important iterative methods with global convergence for solving system (1.1). In general, given an iterate  $x^k$ , one computes a search direction  $d_k$  by solving the Newton equation

$$F(x^k) + F'(x^k)d_k = 0, \tag{1.2}$$

then let  $x^{k+1} = x^k + \alpha_k d_k$ , where step size  $\alpha_k \in [0, 1]$  can be obtained by some line search. The line search method is a class of important globalisation strategies.

<sup>1</sup>College of Mathematics and Econometrics, Hunan University, Changsha 410082, China; e-mail: yyangf@hotmail.com.

© Australian Mathematical Society 2003, Serial-fee code 1446-1811/03

In the case when  $F$  is nonsmooth, Qi and Chen [16] decomposed  $F(x)$  into  $F(x) = f_k(x) + g_k(x)$ , where  $f_k(x)$  is a smooth function,  $\|g_k\| = \sup_{x \in \mathbb{R}^n} \|g_k(x)\| \leq \alpha \|F(x^k)\|$  and  $\alpha \in (0, 1)$  is a constant; then  $f'_k(x^k)$  was used to replace  $F'(x^k)$  in (1.2), thus the successive approximation method was proposed and global convergence was established. Recently, by applying a smooth function satisfying the Jacobian consistency property to approximate a nonsmooth function, Chen, Qi and Sun [1] proposed a smoothing Newton method and obtained global and locally superlinear/quadratic convergence.

The trust region method is another class of well-known globalisation strategies and is often said to be more reliable than the corresponding line search method for smooth problems. In the past twenty years, many trust region methods for solving nonsmooth optimisation problems have been proposed, see for example [2, 10, 18–21]. Globally and superlinearly convergent trust region methods for systems of smooth equations have also been constructed, see [12] for a general survey on this development. However, there are relatively few studies on trust region methods for solving nonsmooth systems. Based on the work in [16], Qi [15] proposed two trust region methods for solving nonsmooth systems and established their global convergence, but did not obtain any rate of convergence result.

Since the function  $F$  is nonsmooth, it is not possible to use a standard trust region method for systems of smooth equations. In this paper, based on the Jacobian consistency property, we propose a trust region method for solving the nonsmooth system (1.1). The proposed trust region method is different from the classical trust region method in that after a successful iteration we use some fixed positive constant as a lower bound for the new radius. This type of updating rule for the trust region radius has also been used in recent works on nonsmooth trust region methods [3, 5, 7, 9, 11]. Compared with existing trust region methods for systems of nonsmooth equations, this modified updating rule enables us not only to show global convergence but also to establish locally superlinear or even quadratic convergence under suitable conditions.

Throughout this paper, the symbol  $\|\cdot\|$  will refer to the Euclidean vector norm or its associated matrix norm and  $\mathcal{N} := \{0, 1, \dots\}$ . For the sake of convenience, we denote

$$\mathbb{R}_+ = \{\varepsilon \mid \varepsilon \geq 0, \varepsilon \in \mathbb{R}^n\} \quad \text{and} \quad \mathbb{R}_{++} = \{\varepsilon \mid \varepsilon > 0, \varepsilon \in \mathbb{R}^n\}.$$

This paper is organised as follows. In the next section, we introduce the concept of the Jacobian consistency property of a smoothing approximation function and recall some important properties of semismooth functions. In Section 3, we present a trust region algorithm for systems of nonsmooth equations and prove that it is well-defined.

In Section 4, we show that the proposed algorithm is globally convergent and locally superlinearly/quadratically convergent under mild conditions. Some conclusive remarks are given in the last section.

## 2. Preliminaries

In this section we introduce several fundamental concepts. First we give several definitions of the generalised Jacobian on nonsmooth functions (see [1, 14] for details).

Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be locally Lipschitz continuous. According to Rademacher's theorem,  $F$  is differentiable almost everywhere. Let  $D_F$  be the set where  $F$  is differentiable. The  $B$ -differential of  $F$  at  $x$  is defined by

$$\partial_B F(x) = \lim_{\substack{x^k \rightarrow x \\ x^k \in D_F}} F'(x^k).$$

The generalised Jacobian of  $F$  at  $x$  in the sense of Clarke is

$$\partial F(x) = \text{conv } \partial_B F(x).$$

In this paper we use the following definition of generalised Jacobians:

$$\partial_C F(x) = \partial F_1(x) \times \partial F_2(x) \times \cdots \times \partial F_n(x).$$

As was defined in [1], we introduce the Jacobian consistency property of a smoothing approximation function as follows.

**DEFINITION 2.1.** Let  $F$  be a locally Lipschitz continuous function in  $\mathbb{R}^n$ . We call  $f : \mathbb{R}^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}^n$  a smoothing approximation function of  $F$  if  $f$  is continuously differentiable with respect to the first variable and there is a constant  $\mu > 0$  such that for any  $x \in \mathbb{R}^n$  and  $\varepsilon \in \mathbb{R}_{++}$ ,

$$\|f(x, \varepsilon) - F(x)\| \leq \mu\varepsilon. \quad (2.1)$$

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

$$\lim_{\varepsilon \downarrow 0} \text{dist}((\nabla_x f(x, \varepsilon))^T, \partial_C F(x)) = 0, \quad (2.2)$$

then we say  $f$  satisfies the Jacobian consistency property.

For simplicity, we denote  $f_x(x, \varepsilon) = (\nabla_x f(x, \varepsilon))^T$ ,  $\Phi(x, \varepsilon) = \|f(x, \varepsilon)\|^2/2$ . Qi and Chen [16] proved that for any continuous function  $F$ , a smoothing approximation function  $f$  of  $F$  can be constructed using convolution. Chen, Qi and Sun [1] investigated the case in which  $f$  has the Jacobian consistency property, and proved that the smoothing approximation functions introduced in [4, 16] possess this property.

To discuss the convergence rate of the proposed algorithm, we introduce the concept of the semismooth operator and give some of its properties. For more details, we refer the reader to [14, 17].

DEFINITION 2.2. We say that  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is semismooth at  $x$  if  $F$  is locally Lipschitzian at  $x$  and

$$\lim_{\substack{V \in \partial F(x+th') \\ h' \rightarrow h, t \downarrow 0}} \{Vh'\} \tag{2.3}$$

exists for any  $h \in \mathbb{R}^n$ .

Obviously, if  $F$  is semismooth at  $x$ , then for any  $h \in \mathbb{R}^n$ , the directional derivative  $F'(x; h)$  of  $F$  at  $x$  along the direction  $h$  exists and is equal to the limit in (2.3). Moreover, we have the following lemma.

LEMMA 2.1. *Suppose that  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a locally Lipschitzian function. Then  $F$  is semismooth at  $x$  if and only if for any  $V \in \partial F(x + h)$  and  $h \rightarrow 0$ ,*

$$Vh - F'(x; h) = o(\|h\|). \tag{2.4}$$

The above property motivates the following definition.

DEFINITION 2.3. If  $F$  is locally Lipschitzian at  $x$  and for any  $V \in \partial F(x + h)$  and  $h \rightarrow 0$ ,

$$Vh - F'(x; h) = O(\|h\|^2), \tag{2.5}$$

then we say that  $F$  is strongly semismooth at  $x$ .

The following results play an important role in the analysis of the convergence rate of the proposed algorithm.

LEMMA 2.2. (i) *If  $F$  is semismooth at  $x$ , then for any  $h \rightarrow 0$ ,*

$$F(x + h) - F(x) - F'(x; h) = O(\|h\|). \tag{2.6}$$

(ii) *If  $F$  is strongly semismooth at  $x$ , then for any  $h \rightarrow 0$ ,*

$$F(x + h) - F(x) - F'(x; h) = O(\|h\|^2). \tag{2.7}$$

### 3. Algorithm

In this section, we shall propose a trust region algorithm for solving the nonsmooth system (1.1) and prove that the proposed algorithm is well-defined.

Throughout this and the following section, we assume that  $f$  has the Jacobian consistency property.

Let  $x^0$  be the initial point of the algorithm,  $\alpha \in (0, 1)$  and  $M_0 := (1 + \alpha)\|F(x^0)\| \neq 0$ . We make the following assumptions.

ASSUMPTION A. (i) The level set  $L_0 := \{x \in \mathbb{R}^n \mid \|F(x)\| \leq M_0\}$  is bounded;  
 (ii)  $\forall \varepsilon \in \mathbb{R}_{++}$  and  $x \in L_0$ ,  $f_x(x, \varepsilon)$  is nonsingular.

We now state our trust region algorithm for solving the nonsmooth system (1.1) and then give some remarks.

ALGORITHM 3.1. Step 0. Choose positive constants  $c_1, c_2, c_3, c_4, \gamma, \eta, \Delta_{\min}, \Delta_0$  satisfying  $c_2 < c_1 < 1, c_3 < 1 < c_4, \eta < 1$ . Let  $\beta_0 := \|F(x^0)\|, \varepsilon_0 := \alpha\beta_0^2/(2M_0\mu)$  and set  $k := 0$ .

Step 1. Solve the subproblem

$$\min \{ Q_k(d) := \|f(x^k, \varepsilon_k) + f_x(x^k, \varepsilon_k)d\|^2/2 \mid \|d\| \leq \Delta_k \}. \tag{3.1}$$

Let  $d_k$  be the solution of (3.1).

Step 2. Set

$$r_k := \frac{\Phi(x^k, \varepsilon_k) - \Phi(x^k + d_k, \varepsilon_k)}{\Phi(x^k, \varepsilon_k) - Q_k(d_k)}, \tag{3.2}$$

$$x^{k+1} := \begin{cases} x^k + d_k, & \text{if } r_k > c_2, \\ x^k, & \text{otherwise,} \end{cases} \tag{3.3}$$

$$\Delta_{k+1} := \begin{cases} c_3\Delta_k, & \text{if } r_k \leq c_2, \\ \max\{\Delta_{\min}, \Delta_k\}, & \text{if } c_2 < r_k \leq c_1, \\ \max\{\Delta_{\min}, c_4\Delta_k\}, & \text{otherwise.} \end{cases} \tag{3.4}$$

Step 3. If  $F(x^{k+1}) = 0$ , stop; otherwise, go to Step 4.

Step 4. If  $\|F(x^{k+1})\| \leq \max\{\eta\beta_k, \alpha^{-1}\|F(x^{k+1}) - f(x^{k+1}, \varepsilon_k)\|\}$ , then set

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

and choose  $\varepsilon_{k+1}$  satisfying

$$0 < \varepsilon_{k+1} \leq \min \left\{ \frac{\alpha}{2M_0\mu}\beta_{k+1}^2, \frac{\varepsilon_k}{2} \right\} \tag{3.5}$$

and

$$\text{dist}(f_x(x^{k+1}, \varepsilon_{k+1}), \partial_C F(x^{k+1})) \leq \gamma\beta_{k+1}; \tag{3.6}$$

otherwise, let  $\beta_{k+1} := \beta_k$  and  $\varepsilon_{k+1} := \varepsilon_k$ .

Step 5. Set  $k := k + 1$ , go to Step 1.

REMARK 1. (i) When a successful iteration arises, the updating rule of the trust region radius in the above trust region method is different from one in the classical trust region method. More precisely, when  $r_k > c_2$ , Algorithm 3.1 uses a fixed positive constant  $\Delta_{\min}$  as a lower bound for the new radius. This idea has also been used in recent literature [3, 5, 7, 9, 11].

(ii) It is not difficult to prove that if Algorithm 1 in [15] uses (3.4) as the updating rule of the trust region radius, its convergence result is still valid.

(iii) Notice that the choice of the parameter  $\varepsilon_{k+1}$  in the above algorithm might be difficult for general nonsmooth systems. However, it has been shown in [1] how to choose an  $\varepsilon_{k+1}$  satisfying (3.5) and (3.6) for general box-constrained variational inequality problems and order complementarity problems.

Without loss of generality, we assume that  $\|F(x^k)\| \neq 0$  for all  $k \geq 0$ . Set

$$\begin{aligned}
 K &:= \{0\} \cup \{k \mid \|F(x^k)\| \leq \max\{\eta\beta_{k-1}, \alpha^{-1}\|f(x^k, \varepsilon_{k-1}) - F(x^k)\|\}, k \in \mathcal{N}\} \\
 &= \{k_0 = 0 < k_1 < k_2 \dots\}.
 \end{aligned}
 \tag{3.7}$$

The following proposition shows that Algorithm 3.1 is well defined.

**PROPOSITION 3.1.** *Under Assumption A (ii), Algorithm 3.1 is well defined and the generated sequence  $\{x^k\}$  remains in  $L_0$ .*

**PROOF.** For all  $k \geq 0$ , let  $k_j$  be the largest number in  $K$  such that  $k_j \leq k$ , then  $\varepsilon_k = \varepsilon_{k_j}$  and  $\beta_k = \beta_{k_j}$ . Thus iterations  $k_j$  to  $k$  of Algorithm 3.1 may be regarded as some iterations of Algorithm 1 in [15] with  $F(x) = p(x) = f(x, \varepsilon_{k_j})$ ,  $q(x) = 0$  and the updating rule (3.4).

Let  $U_j := \{x \in \mathbb{R}^n \mid \|f(x, \varepsilon_{k_j})\| \leq \|f(x^{k_j}, \varepsilon_{k_j})\|\}$ . In the following, we will prove by induction that

$$U_j \subseteq L_0, \quad \forall j \geq 0.
 \tag{3.8}$$

In fact,  $\forall x \in U_j$ ,

$$\begin{aligned}
 \|F(x)\| &\leq \|f(x, \varepsilon_{k_j})\| + \mu\varepsilon_{k_j} \leq \|f(x^{k_j}, \varepsilon_{k_j})\| + \mu\varepsilon_{k_j} \\
 &\leq \|F(x^{k_j})\| + 2\mu\varepsilon_{k_j} = \beta_{k_j} + 2\mu\varepsilon_{k_j}.
 \end{aligned}
 \tag{3.9}$$

For  $j = 0$ ,  $\forall x \in U_0$ ,  $\|F(x)\| \leq \|F(x^0)\| + 2\mu\varepsilon_0 \leq (1 + \alpha)\|F(x^0)\|$ , which implies  $U_0 \subseteq L_0$ . Suppose that  $U_{j-1} \subseteq L_0$  for some  $j \geq 1$ . Set

$$\begin{aligned}
 K_1 &:= \{k \in K \mid \eta\beta_{k-1} \geq \alpha^{-1}\|f(x^k, \varepsilon_{k-1}) - F(x^k)\|\}, \\
 K_2 &:= \{k \in K \mid \eta\beta_{k-1} < \alpha^{-1}\|f(x^k, \varepsilon_{k-1}) - F(x^k)\|\}.
 \end{aligned}$$

By Step 4 of Algorithm 3.1, we have  $\beta_{k_j} \leq \eta\beta_{k_j-1} = \eta\beta_{k_{j-1}}$ , if  $k_j \in K_1$ , or

$$\beta_{k_j} \leq \frac{\mu}{\alpha}\varepsilon_{k_j-1} = \frac{\mu}{\alpha}\varepsilon_{k_{j-1}} \leq \frac{1}{2M_0}\beta_{k_{j-1}}^2 \leq \frac{1}{2}\beta_{k_{j-1}}, \quad \text{if } k_j \in K_2.$$

Let  $r = \max\{1/2, \eta\}$ . Then  $\beta_{k_j} \leq r\beta_{k_{j-1}}$  and  $\varepsilon_{k_j} \leq \varepsilon_{k_{j-1}}/2 = \varepsilon_{k_{j-1}}/2, \forall k_j \in K$ . Thus, for  $j \geq 1$ , we get

$$\varepsilon_{k_j} \leq \frac{1}{2^j} \varepsilon_0 \leq \frac{\alpha}{2^{j+1} M_0 \mu} \|F(x^0)\|^2 \leq \frac{\alpha}{2^{j+1} \mu} \|F(x^0)\|$$

and  $\beta_{k_j} \leq r^j \beta_0 = r^j \|F(x^0)\|$ .

It follows from (3.9) and the above analysis that  $\forall x \in U_j$ ,

$$\|F(x)\| \leq r^j \|F(x^0)\| + \frac{\alpha}{2^j} \|F(x^0)\| \leq r^j M_0, \tag{3.10}$$

which implies that  $U_j \subseteq L_0$  and hence (3.8) holds. This shows

$$x^k \in U_j \subseteq L_0, \quad \text{as } k_j \leq k < k_{j+1}. \tag{3.11}$$

Moreover, by (2.1) and Step 4 of Algorithm 3.1, we deduce

$$\|f(x^k, \varepsilon_k) - F(x^k)\| \leq \alpha \|F(x^k)\|, \quad \forall k \geq 0. \tag{3.12}$$

This implies  $f(x^k, \varepsilon_k) \neq 0$  for all  $k \geq 0$  and hence by Assumption A (ii),  $d_k \neq 0$ . The proof is complete.

#### 4. Global and superlinear convergence

In this section, we shall prove that Algorithm 3.1 converges to the solution of the nonsmooth system (1.1) globally and superlinearly/quadratically. To this end, we shall first show that the index set  $K$  is infinite so that the global convergence follows from (3.10) and (3.11). Then we shall prove that there exists infinitely many  $k \in K$  such that for each  $k$ , the  $(k - 1)$ th iteration is successful and under mild conditions, eventually the  $k$ th iteration is also successful and the trust region radius in (3.1) is inactive, thus the desired result is obtained.

**PROPOSITION 4.1.** *Under Assumption A, the index set  $K$  defined by (3.7) is infinite.*

**PROOF.** The proof is by contradiction. Assume that  $K$  is finite and let  $\hat{k}$  be the largest number in  $K$ . Then for all  $k \geq \hat{k}$ ,  $\varepsilon_k = \varepsilon_{\hat{k}}$  and  $\beta_k = \beta_{\hat{k}}$ . Denote  $\tilde{\varepsilon} := \varepsilon_{\hat{k}}$ ,  $\tilde{\beta} := \beta_{\hat{k}}$  and  $q(x) := F(x) - f(x, \tilde{\varepsilon})$ . Then  $\forall k > \hat{k}$ ,

$$\|F(x^k)\| > \max\{\eta\tilde{\beta}, \alpha^{-1}\|q(x^k)\|\} \tag{4.1}$$

and

$$F(x^k) = f(x^k, \tilde{\varepsilon}) + q(x^k). \tag{4.2}$$

Therefore Algorithm 3.1 may be regarded as the case in [15, Algorithm 1] with  $F(x) = p(x) = f(x, \bar{\varepsilon})$  and the updating rule (3.4). From [15, Theorem 2] and Remark 1 (ii), we deduce that if  $\bar{x}$  is an accumulation point of  $\{x^k\}$ , then  $f(\bar{x}, \bar{\varepsilon}) = 0$ .

On the other hand, Assumption A (i) implies that there exist an accumulation point  $\bar{x} \in L_0$  and a subsequence  $\{x^k\}_{k \in K_3}$  such that as  $k \in K_3$  and  $k \rightarrow \infty, x^k \rightarrow \bar{x}$ . Thus

$$\lim_{\substack{k \rightarrow \infty \\ k \in K_3}} f(x^k, \bar{\varepsilon}) = f(\bar{x}, \bar{\varepsilon}) = 0,$$

which shows that there exists  $\tilde{k} \geq \hat{k}$  such that for all  $k \in K_3$  with  $k \geq \tilde{k}$ ,

$$\|f(x^k, \bar{\varepsilon})\| \leq (1 - \alpha)\eta\tilde{\beta}.$$

By (4.1), (4.2) and the above expression, we deduce that for all  $k \in K_3$  with  $k \geq \tilde{k}$ ,

$$\|f(x^k, \bar{\varepsilon})\| < (1 - \alpha)\|F(x^k)\| \leq (1 - \alpha)(\|q(x^k)\| + \|f(x^k, \bar{\varepsilon})\|),$$

which implies that  $\|f(x^k, \bar{\varepsilon})\| < (\alpha^{-1} - 1)\|q(x^k)\|$  and hence

$$\|F(x^k)\| \leq \|f(x^k, \bar{\varepsilon})\| + \|q(x^k)\| < \alpha^{-1}\|q(x^k)\|.$$

This contradicts (4.1). The proof is complete.

It follows from (3.10) and (3.11) that for  $k_j \leq k < k_{j+1}, \|F(x^k)\| \leq r^j M_0$ . This, combined with Proposition 3.1 and Proposition 4.1, shows that the following global convergence result holds.

**THEOREM 4.1.** *Under Assumption A, we have  $\lim_{k \rightarrow \infty} F(x^k) = 0$ ; that is, every accumulation point of  $\{x^k\}$  is a solution of the nonsmooth system (1.1).*

From the proof of Proposition 4.1, we see that Proposition 4.1 is still true if Assumption A (i) is replaced by the assumption that there exists at least one accumulation point in the sequence  $\{x^k\}$ . Hence the following result is deduced.

**COROLLARY 4.1.** *Under Assumption A (ii), every accumulation point of  $\{x^k\}$  is a solution of the nonsmooth system (1.1).*

The next proposition shows that  $K$  includes infinitely many indices obtained by successful iterations.

**PROPOSITION 4.2.** *Under Assumption A, there exist infinitely many  $k \in K$  such that for each  $k$ , the  $(k - 1)$ th iteration is successful.*

PROOF. We proceed by contradiction. Suppose that the proposition is false. Then there exists an index  $\bar{k} \in K$  such that for all  $k \in K$  with  $k \geq \bar{k}$ ,

$$x^k = x^{k-1} \tag{4.3}$$

and

$$\|F(x^k)\| \leq \max \{ \eta\beta_{k-1}, \alpha^{-1}\|f(x^k, \varepsilon_{k-1}) - F(x^k)\| \}. \tag{4.4}$$

We claim that  $k - 1 \in K$ . In fact, if  $k - 1 \notin K$ , then  $\beta_{k-1} = \beta_{k-2}$  and  $\varepsilon_{k-1} = \varepsilon_{k-2}$ . This, together with (4.3) and (4.4), shows that  $k - 1 \in K$  and hence we get a contradiction.

By Proposition 4.1 and repeating the above process, we may prove that  $k \in K$ , for all  $k \geq \bar{k}$ , which implies that for all  $k > \bar{k}$ ,  $x^k = x^{\bar{k}}$ ,  $\beta_k = \|F(x^k)\|$  and  $\varepsilon_k \leq \varepsilon_{\bar{k}}/2^{k-\bar{k}}$ . By (2.1) and (4.4), we get  $\|F(x^{\bar{k}})\| \leq \max \{ \eta\|F(x^{\bar{k}})\|, \mu\varepsilon_{k-1}/\alpha \}$ , which contradicts  $F(x^{\bar{k}}) \neq 0$ ,  $\varepsilon_k \rightarrow 0$  and  $0 < \eta < 1$ . The assertion is proved.

In order to analyse the convergence rate of Algorithm 3.1, we first give an important proposition.

PROPOSITION 4.3. *Suppose that Assumption A holds and that  $K_0$  is an infinite subset of  $K$  such that for all  $k \in K_0$ , the  $(k - 1)$ th iteration is successful and  $\{x^k\}_{k \in K_0}$  converges to  $x^*$ . Suppose that all  $V \in \partial_C F(x^*)$  are nonsingular. Then there exists  $\bar{k} \in K_0$  such that for all  $k \in K_0$  with  $k \geq \bar{k}$ , the  $k$ th iteration is successful and the trust region radius in the subproblem (3.1) is inactive; that is,*

$$x^{k+1} = x^k + d_k \quad \text{and} \quad d_k = -f_x(x^k, \varepsilon_k)^{-1}f(x^k, \varepsilon_k). \tag{4.5}$$

PROOF. From (3.12) and Theorem 4.1, we get  $F(x^*) = 0$  and

$$\lim_{k \rightarrow \infty} f(x^k, \varepsilon_k) = 0. \tag{4.6}$$

we note that for any  $x \in \mathbb{R}^n$ ,  $\partial_C F(x)$  is a compact set. Let  $V_k \in \partial_C F(x^k)$  such that

$$\text{dist}(f_x(x^k, \varepsilon_k), \partial_C F(x^k)) = \|f_x(x^k, \varepsilon_k) - V_k\|.$$

By (3.6), we have

$$\|f_x(x^k, \varepsilon_k) - V_k\| \leq \gamma\beta_k, \quad \forall k \in K_0. \tag{4.7}$$

It follows from Theorem 4.1 that  $\beta_k \rightarrow 0$  as  $k \rightarrow \infty$ . By the compactness of  $\partial_C F(x^*)$ , the nonsingularity of all  $V \in \partial_C F(x^*)$  and the upper semicontinuity of  $\partial_C F(\cdot)$  at  $x^*$ , we deduce from (4.7) that there exist  $M > 0$  and  $\hat{k} > 0$  such that for all  $k \in K_0$  with  $k \geq \hat{k}$ ,  $\|f_x(x^k, \varepsilon_k)^{-1}\| \leq M$ .

We now consider the subproblem (3.1). By (4.6), we deduce that there exists  $\bar{k} \geq \hat{k}$  such that for all  $k \in K_0$  with  $k \geq \bar{k}$ ,

$$\|f_x(x^k, \varepsilon_k)^{-1}f(x^k, \varepsilon_k)\| \leq \Delta_{\min} \leq \Delta_k,$$

which implies that

$$d_k = -f_x(x^k, \varepsilon_k)^{-1} f(x^k, \varepsilon_k) \quad \text{and} \quad Q_k(d_k) = 0. \tag{4.8}$$

This shows  $\Phi(x^k, \varepsilon_k) - Q_k(d_k) = \|f(x^k, \varepsilon_k)\|^2/2$  and

$$\begin{aligned} \Phi(x^k, \varepsilon_k) - \Phi(x^k + d_k, \varepsilon_k) &= \frac{1}{2} \|f(x^k, \varepsilon_k)\|^2 - o(\|d_k\|^2) \\ &= \frac{1}{2} \|f(x^k, \varepsilon_k)\|^2 - o(\|f(x^k, \varepsilon_k)\|^2). \end{aligned}$$

Hence we have

$$\lim_{\substack{k \rightarrow \infty \\ k \in K_0}} r_k = \lim_{\substack{k \rightarrow \infty \\ k \in K_0}} \frac{\Phi(x^k, \varepsilon_k) - \Phi(x^k + d_k, \varepsilon_k)}{\Phi(x^k, \varepsilon_k) - Q_k(d_k)} = 1,$$

which implies that there exists  $\tilde{k} \geq \bar{k}$  such that for all  $k \in K_0$  with  $k \geq \tilde{k}$ ,  $r_k > c_2$ ; that is,

$$x^{k+1} = x^k + d_k. \tag{4.9}$$

The assertion then follows from (4.8) and (4.9).

We are now ready to present a rate of convergence result for Algorithm 3.1.

**THEOREM 4.2.** *Suppose that Assumption A holds. Suppose that for any accumulation point  $\bar{x}$  of the sequence  $\{x^k\}$  generated by Algorithm 3.1, all  $V \in \partial_C F(\bar{x})$  are nonsingular and  $F$  is semismooth at  $\bar{x}$ . Then the whole sequence  $\{x^k\}$  converges to a solution  $x^*$  of  $F(x) = 0$  superlinearly. Moreover, if  $F$  is strongly semismooth at  $x^*$ , then the convergence rate is quadratic.*

**PROOF.** Let  $\bar{x}$  be an accumulation point of  $\{x^k\}$ . By Theorem 4.1, we have  $F(\bar{x})=0$ . Since  $\partial_B F(\bar{x}) \subseteq \partial_C F(\bar{x})$ , by [14, Proposition 2.5], there exists a neighbourhood of  $\bar{x}$  such that  $\bar{x}$  is the unique solution in this neighbourhood. Therefore the sequence  $\{x^k\}$  only has finitely many accumulation points.

It follows from Proposition 4.2 and Assumption A (i) that for at least one accumulation point  $x^*$ , there exists an infinite subset  $K_0$  of  $K$  such that for all  $k \in K_0$ , the  $(k - 1)$ th iteration is successful and  $\{x^k\}_{k \in K_0}$  converges to  $x^*$ . By the proof of Proposition 4.3, there exist  $M > 0$  and  $\tilde{k} \in K_0$  such that for all  $k \in K_0$  with  $k \geq \tilde{k}$ , (4.5) holds and  $\|f_x(x^k, \varepsilon_k)^{-1}\| \leq M$ .

If  $F$  is semismooth at  $x^*$ , from (2.4) and (2.6), we have

$$\|V_k(x^k - x^*) - F(x^k) + F(x^*)\| = o(\|x^k - x^*\|) \quad \text{as } k \rightarrow \infty, k \in K_0. \tag{4.10}$$

Therefore, by (2.1), (4.5), (4.7) and (4.10), we deduce that for all  $k \in K_0$  with  $k \geq \tilde{k}$ ,

$$\begin{aligned} \|x^{k+1} - x^*\| &= \|x^k + d^k - x^*\| \\ &\leq \|f_x(x^k, \varepsilon_k)^{-1}\| (\|(f_x(x^k, \varepsilon_k) - V_k)(x^k - x^*)\| \\ &\quad + \|V_k(x^k - x^*) - F(x^k) + F(x^*)\| + \|F(x^k) + f_x(x^k, \varepsilon_k)d_k\|) \\ &\leq M(\gamma\beta_k\|x^k - x^*\| + o(\|x^k - x^*\|) + \mu\varepsilon_k). \end{aligned} \tag{4.11}$$

By the local Lipschitz continuity of  $F$ , for all  $k \in K_0$ ,

$$\beta_k = \|F(x^k)\| = O(\|x^k - x^*\|)$$

and

$$\varepsilon_k = O(\|x^k - x^*\|^2). \tag{4.12}$$

Thus, it follows from (4.11) that

$$\|x^{k+1} - x^*\| = o(\|x^k - x^*\|) \quad \text{as } k \rightarrow \infty, k \in K_0. \tag{4.13}$$

From the proof of Theorem 3.1 in [14], we deduce

$$\|F(x^{k+1})\| = o(\|F(x^k)\|), \quad \text{as } k \rightarrow \infty, k \in K_0.$$

So there exists  $\hat{k} \geq \tilde{k}$  such that for all  $k \in K_0$  with  $k \geq \hat{k}$ ,

$$\|F(x^{k+1})\| \leq \eta\|F(x^k)\| = \eta\beta_k,$$

which implies that  $k + 1 \in K$  and  $\{x^{k+1}\}_{k \in K_0}$  converges to  $x^*$ . Hence, by Proposition 4.3,  $k + 1 \in K_0$ .

Repeating the above process, we may prove that  $k \in K_0$ , for all  $k \geq \hat{k}$ . This, together with (4.13), implies that  $\{x^k\}$  converges to  $x^*$  superlinearly.

If  $F$  is strongly semismooth at  $x^*$ , from (2.5) and (2.7), we have

$$\|V_k(x^k - x^*) - F(x^k) + F(x^*)\| = O(\|x^k - x^*\|^2) \quad \text{as } k \rightarrow \infty. \tag{4.14}$$

By (4.11), (4.12) and (4.14), we deduce that as  $k \rightarrow \infty$ ,

$$\|x^{k+1} - x^*\| \leq M(\gamma\beta_k\|x^k - x^*\| + O(\|x^k - x^*\|^2) + \mu\varepsilon_k) = O(\|x^k - x^*\|^2),$$

which shows that  $\{x^k\}$  quadratically converges to  $x^*$ .

## 5. Conclusion

In this paper we developed a new trust region algorithm for solving systems of nonsmooth equations. The proposed algorithm is based on the Jacobian consistency property of a smoothing approximation function and has some nice convergence properties. We not only establish global convergence but also recover locally superlinear or even quadratic convergence under mild conditions in spite of the fact that the function involved in the system of equations is nonsmooth.

We observe that Kanzow and Pieper [8] considered a line search method without a condition like Assumption A (ii) and established the desired convergence properties. How to develop a corresponding method within the trust region framework might be an interesting topic for future research.

## Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant 10171030.

The author would like to thank the referee for his helpful comments on this paper and drawing his attention to reference [8].

## References

- [1] X. Chen, L. Qi and D. Sun, "Global and superlinear convergence of the smoothing Newton method and its applications to general box constrained variational inequalities", *Math. Comp.* **222** (1998) 519–540.
- [2] J. E. Dennis, S. B. Li and R. A. Tapia, "A unified approach to global convergence of trust region methods for nonsmooth optimization", *Math. Program.* **68** (1995) 319–346.
- [3] A. Friedlander, J. M. Martínez and S. A. Santos, "A new trust region algorithm for bound constrained minimization", *Appl. Math. Optim.* **30** (1994) 235–266.
- [4] S. A. Gabriel and J. J. Moré, "Smoothing of mixed complementarity problems", in *Complementarity and variational problems: State of the Art* (eds. M. C. Ferris and J. S. Pang), (SIAM, Philadelphia, 1997) 105–116.
- [5] S. A. Gabriel and J. S. Pang, "A trust region method for constrained nonsmooth equations", in *Large scale optimization: State of the Art* (eds. W. W. Hager, D. W. Hearn and P. M. Pardalos), (Kluwer Academic Publishers, Boston, MA, 1994) 159–186.
- [6] S. P. Han, J. S. Pang and N. Rangaraj, "Globally convergent Newton methods for nonsmooth equations", *Math. Oper. Res.* **17** (1992) 586–607.
- [7] H. Jiang, M. Fukushima, L. Qi and D. Sun, "A trust region method for solving generalized complementarity problems", *SIAM J. Optim.* **8** (1998) 140–157.
- [8] C. Kanzow and H. Pieper, "Jacobian smoothing methods for nonlinear complementarity problems", *SIAM J. Optim.* **9** (1999) 342–373.

- [9] C. Kanzow and M. Zupke, "Inexact trust region methods for nonlinear complementarity problems", in *Reformulation: Nonsmooth, piecewise smooth, semismooth and smoothing methods* (eds. M. Fukushima and L. Qi), (Kluwer Academic Publishers, Norwell, MA, 1998) 211–233.
- [10] J. M. Martínez and A. C. Moretti, "A trust region method for minimization of nonsmooth functions with linear constraints", *Math. Program.* **76** (1997) 431–449.
- [11] J. M. Martínez and S. A. Santos, "A trust region strategy for minimization on arbitrary domains", *Math. Program.* **68** (1995) 267–302.
- [12] J. J. Moré, "Recent developments in algorithms and software for trust region methods", in *Mathematical programming: State of the Art* (eds. A. Bachem, M. Grottschel and B. Korte), (Springer, Berlin, 1983) 258–287.
- [13] J. S. Pang and L. Qi, "Nonsmooth equations: motivation and algorithms", *SIAM J. Optim.* **3** (1993) 443–465.
- [14] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations", *Math. Oper. Res.* **18** (1993) 227–244.
- [15] L. Qi, "Trust region algorithms for solving nonsmooth equation", *SIAM J. Optim.* **5** (1995) 219–230.
- [16] L. Qi and X. Chen, "A globally convergent successive approximation method for nonsmooth equations", *SIAM J. Control Optim.* **38** (1995) 402–418.
- [17] L. Qi and J. Sun, "A nonsmooth version of Newton's method", *Math. Program.* **58** (1993) 353–367.
- [18] L. Qi and J. Sun, "A trust region algorithm for minimization of locally Lipschitzian functions", *Math. Program.* **66** (1994) 25–43.
- [19] Y. Yuan, "An example of only linearly convergence of trust region algorithms for nonsmooth optimization", *IMA J. Numer. Anal.* **4** (1984) 327–335.
- [20] Y. Yuan, "Conditions for convergence of trust region algorithms for nonsmooth optimization", *Math. Program.* **31** (1985) 220–228.
- [21] Y. Yuan, "On the superlinear convergence of a trust region algorithm for nonsmooth optimization", *Math. Program.* **31** (1985) 269–285.