MATHEMATICS OF COMPUTATION
Volume 67, Number 222, April 1998, Pages 519–540
S 0025-5718(98)00932-6

GLOBAL AND SUPERLINEAR CONVERGENCE OF THE SMOOTHING NEWTON METHOD AND ITS APPLICATION TO GENERAL BOX CONSTRAINED VARIATIONAL INEQUALITIES

X. CHEN, L. QI, AND D. SUN

Abstract. The smoothing Newton method for solving a system of nonsmooth equations $F(x) = 0$, which may arise from the nonlinear complementarity problem, the variational inequality problem or other problems, can be regarded as a variant of the smoothing method. At the kth step, the nonsmooth function F is approximated by a smooth function $f(\cdot, \varepsilon_k)$, and the derivative of $f(\cdot, \varepsilon_k)$ at x^k is used as the Newton iterative matrix. The merits of smoothing methods and smoothing Newton methods are global convergence and convenience in handling. In this paper, we show that the smoothing Newton method is also superlinearly convergent if F is semismooth at the solution and f satisfies a Jacobian consistency property. We show that most common smooth functions, such as the Gabriel-Moré function, have this property. As an application, we show that for box constrained variational inequalities if the involved function is P–uniform, the iteration sequence generated by the smoothing Newton method will converge to the unique solution of the problem globally and superlinearly (quadratically).

1. Introduction

Let $p, q : \Re^n \to \Re^n$ be two smooth (continuously differentiable) mappings and X be a closed convex set in $\Re^n$. The general variational inequality problem, $\mathrm{GVI}(X, p, q)$ for short, is to find a vector $x \in \Re^n$ such that

$$q(x) \in X, \quad (y - q(x))^T p(x) \geq 0 \quad \text{for all } y \in X. \tag{1.1}$$

It is known (for example, see [10] for a proof in the case that $q(x) = x$) that $\mathrm{GVI}(X, p, q)$ is equivalent to finding a zero of the following nonsmooth equation

$$q(x) - \Pi_X[q(x) - p(x)] = 0, \tag{1.2}$$

where Π_X is the projection operator onto X under the Euclidean norm. Equation (1.2) is called the generalized normal equation in [39]. $\mathrm{GVI}(X, p, q)$ is a generalization of variational inequalities and general complementarity problems. The variational inequality problem is to find an $x \in X$ such that

$$(y - x)^T p(x) \geq 0 \quad \text{for all } y \in X. \tag{1.3}$$

Received by the editor June 17, 1996 and, in revised form, January 9, 1997.
1991 *Mathematics Subject Classification.* Primary 90C33, 90C30, 65H10.
Key words and phrases. Variational inequalities, nonsmooth equations, smoothing approximation, smoothing Newton method, convergence.
This work is supported by the Australian Research Council.

The general complementarity problem, GCP(p,q) for short, is to find an $x \in \Re^n$ such that

$$q(x) \geq 0, \quad p(x) \geq 0, \quad q(x)^T p(x) = 0. \tag{1.4}$$

When $q(x) = x$, GCP(p,q) reduces to the nonlinear complementarity problem of finding an $x \in \Re^n$ such that

$$x \geq 0, \quad p(x) \geq 0, \quad x^T p(x) = 0. \tag{1.5}$$

A lot of effort has been spent on complementarity problems and variational inequalities, for a comprehensive survey see [20], [37].

In this paper we focus on general box constrained variational inequalities, i.e., we assume that in (1.1) X has the following box form:

$$X = \{x \in \Re^n |\ l \leq x \leq u\}, \tag{1.6}$$

where $l \in \{\Re \cup \{-\infty\}\}^n$, $u \in \{\Re \cup \{+\infty\}\}^n$ and $l < u$. In this case GVI(X,p,q) will be denoted by GVI(l,u,p,q). GVI(l,u,p,q) includes two very useful models: general complementarity problems and box constrained variational inequalities while the latter is actually equivalent to what is called mixed complementarity problems in some papers [2], [4], [9], [18]. Furthermore, GVI(l,u,p,q) models many important problems in engineering, management and economics [20], [37].

When X is of the structure (1.6), problem (1.2) is equivalent to

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

Here mid$(\cdot)$ is the median operator, i.e., for three vectors $a,b,c \in \{\Re \cup \{\pm\infty\}\}^n$ and $a \leq b$,

$$(\mathrm{mid}(a,b,c))_i = \mathrm{mid}(a_i,b_i,c_i) = \begin{cases} a_i & \text{if } c_i < a_i, \\ c_i & \text{if } a_i \leq \ c_i \leq b_i, \\ b_i & \text{if } b_i < c_i, \end{cases} \qquad i = 1,...,n.$$

If $q(x) = x$, (1.7) reduces to

$$x - \mathrm{mid}(l,u,x-p(x)) = 0. \tag{1.8}$$

Since the median operator is piecewise smooth, (1.7) and (1.8) are systems of nonsmooth equations.

A considerable number of generalizations of Newton-type methods [15], [21], [23], [29], [35], [36], [44], [47] have been developed for solving nonsmooth equation

$$F(x) = 0, \tag{1.9}$$

where $F : \Re^n \to \Re^n$ is locally Lipschitz continuous but not differentiable. Some of these methods solve a linear complementarity problem or a linear variational inequality problem at each step. A natural extension of the classical Newton method for solving (1.9) is

$$x^{k+1} = x^k - V_k^{-1} F(x^k), \tag{1.10}$$

where V_k is an $n \times n$ matrix in a generalized Jacobian of F at x^k. For this method, only a system of linear equations needs to be solved at each step. There are several possible definitions of generalized Jacobians. We will discuss them in the next section. It was proved in [40], [43] for the generalized Jacobians in the sense of [7] and [40] that the sequence generated by (1.10) superlinearly (quadratically) converges in a neighbourhood of a solution x^* of (1.9) if F is (strongly) semismooth at x^* and all matrices in the generalized Jacobian of F at x^* are nonsingular. Since most common nonsmooth functions, such as convex functions, piecewise smooth functions, the Burmeister-Fischer function which is useful for nonsmooth equations arising from the nonlinear complementarity problem [16], and their compositions are semismooth functions, this extends the superlinear convergence theory of the classical Newton method to the nonsmooth case. The function

$$F(x) = q(x) - \mathrm{mid}(l, u, q(x) - p(x)) \tag{1.11}$$

is (strongly) semismooth if p and q are (twice) smooth. Globally and superlinearly convergent methods for solving (1.9) can be constructed by combining this theory with some global convergence techniques. Applications of this theory to the nonlinear complementarity problem and the variational inequality problem can be found in [8], [11], [12], [13], [14], [25], [27], [30], [31], [33], [34], [38], [48], [51]. For a general survey on this development, see [26].

Another approach for solving (1.9) is the smoothing method [2], [3], [4], [18], [28]. The feature of smoothing methods is to construct a smoothing approximation function $f : \Re^n \times \Re_{++} \to \Re^n$ of F such that for any $\varepsilon > 0$, $f(\cdot, \varepsilon)$ is continuously differentiable and

$$\|F(x) - f(x, \varepsilon)\| \to 0 \quad \text{as} \quad \varepsilon \to 0_+ \qquad \text{for all } x \in \Re^n,$$

and then to find a solution of (1.9) by (inexactly) solving the following problems for a given positive sequence $\{\varepsilon_k\}$, $k = 0, 1, 2...$,

$$f(x, \varepsilon_k) = 0. \tag{1.12}$$

In [4], Chen and Mangasarian introduced a class of smoothing approximation functions for nonlinear complementarity problems. Gabriel and Moré [18] extended Chen-Mangasarian's smoothing approach to box constrained variational inequalities (1.8). Another class of smoothing approximation functions for general complementarity problems (1.4) was given in [5]. Problems (1.8) and (1.4) are special cases of general box constrained variational inequalities GVI(l, u, p, q). The merits of the smoothing method are global convergence and convenience in handling smooth functions instead of nonsmooth functions. However, (1.12), which needs to be solved at each step, is nonlinear in general.

The smoothing Newton method can be regarded as a variant of the smoothing method. It uses the derivative of f with respect to the first variable in the Newton method, namely

$$x^{k+1} = x^k - t_k f_x(x^k, \varepsilon_k)^{-1} F(x^k), \tag{1.13}$$

where $\varepsilon_k > 0$, $f_x(x^k, \varepsilon_k)$ denotes the derivative of f with respect to the first variable at (x^k, ε_k) and $t_k > 0$ is the stepsize. The smoothing Newton method (1.13) for solving nonsmooth equation (1.9) has been studied for decades in different areas [1], [5], [6], [22], [32], [42], [46], [50]. In some previous papers, method (1.13) is called a splitting method because $F(\cdot)$ is split into a smooth part $f(\cdot, \varepsilon)$ and a nonsmooth part $F(\cdot) - f(\cdot, \varepsilon)$. The global and linear convergence of (1.13) has been discussed in [42], but so far no superlinear convergence result has been obtained. In this paper we will address this problem by investigating the relation between the derivative $f_x(x, \varepsilon)$ and the generalized Jacobian of F at x. We define a Jacobian consistency property and show that the smoothing approximation functions in [4], [5], [18], [42] have this property. Under mild conditions, we prove that the sequence $\{x^k\}$ generated by the smoothing Newton method is bounded and each accumulation point is a solution of (1.9). Furthermore, the convergence rate is superlinear if F is semismooth at the solution and the smoothing approximation function satisfies this Jacobian consistency property. Moreover, the convergence rate is quadratic if F is strongly semismooth at the solution. In particular, for box constrained variational inequalities if p is a uniform P-function, then the smoothing Newton method has three advantages:

- Solving a linear system of equations at each step;
- Guaranteeing that $\{x^k\}$ is bounded and converges to the unique solution;
- Having superlinear convergence rate.

There are a wide variety of algorithms for the solution of variational inequalities with box constraints [4], [9], [13], [18], [27], [45], [48]. As said before, some of them have nonlinear subproblems. The algorithm proposed in [8] for nonlinear complementarity problem (1.5), which is a special case of the box constrained variational inequality problem, has the above three properties. The algorithms proposed in [27], [48] based on a differentiable merit function (for a survey on merit functions, see [17]) for the box constrained variational inequality problem have the above three properties if p is a strongly monotone function, which is a stronger condition than that of a uniform P-function. Hence, as an application, we present a method for solving the general box constrained variational inequality problem with better convergence properties.

This paper is organized as follows. In section 2, we define the Jacobian consistency property. In section 3, we present the smoothing Newton method in detail and prove that the method is globally and superlinearly convergent. In section 4, we discuss the application of the smoothing Newton method to GVI(l, u, p, q) and verify various assumptions. In section 5, we give some final remarks and point out the possible availability of the smoothing Newton method to the order complementarity problem and the variational inequality problem (1.3).

We let $\|\cdot\|$ denote the Euclidean norm of $\Re^n$ and let

$$\Re_+ = \{\varepsilon \mid \varepsilon \geq 0, \varepsilon \in \Re\}$$

and

$$\Re_{++} = \{\varepsilon \mid \varepsilon > 0, \varepsilon \in \Re\}.$$

We denote the set of all nonnegative integers by $N = \{0, 1, ...\}$.

2. Jacobian consistency property

Let $H : \Re^n \to \Re^m$ be locally Lipschitz continuous. According to Rademacher's theorem, H is differentiable almost everywhere. Let D_H be the set where H is differentiable. There are several definitions of generalized Jacobians of H, which can be used in the generalized Newton method. The B-differential of H [40] is defined by

$$\partial_B H(x) = \{ \lim_{\substack{x^k \to x \\ x^k \in D_H}} H'(x^k)\}.$$

The generalized Jacobian of H at x in the sense of Clarke [7] is

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

The superlinear convergence of (1.10) is established for these two kinds of generalized Jacobians in [40], [43]. Some other variants of Jacobians and their perturbations are used in the literature [48], [49], [51]. A general range of different kinds of generalized Jacobians, which are associated with superlinear convergence of (1.10), is discussed in [41]. In this paper, for the function F, we use a kind of generalized Jacobian, denoted by $\partial_C F$ and defined as

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

This definition can be seen as a special case of the C-differential operator discussed in [41] and is more suitable to the discussion in this paper.

We are now able to define the Jacobian consistency property.

Definition 2.1. Let F be a Lipschitz continuous function in $\Re^n$. We call $f : \Re^n \times \Re_{++} \to \Re^n$ 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 $x \in \Re^n$ and $\varepsilon \in \Re_{++}$,

$$\|f(x, \varepsilon) - F(x)\| \le \mu\varepsilon. \tag{2.1}$$

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

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

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

Remark 2.1. In condition (2.1), $\mu\varepsilon$ may be replaced by any nondecreasing function $\sigma : \Re_+ \to \Re_+$ such that $\sigma(0) = 0$ and $\sigma(t) > 0$ for $t > 0$. In this paper, however, we will restrict our discussion to (2.1), because it makes the analysis significantly simple.

For simplicity, in the remainder of this paper we denote

$$f_x(x, \varepsilon) \equiv (\nabla_x f(x, \varepsilon))^T.$$

It was proved in [42] that for any continuous function F by using convolution we can construct a smoothing approximation function f of F. We now investigate the cases in which f has the Jacobian consistency property.

Chen and Mangasarian [4] introduced a class of smoothing functions for the nonlinear complementarity problem (1.5). Gabriel and Moré [18] extended Chen-Mangasarian's smoothing approach to the box constrained variational inequality problem (1.8). The result in [18] may be easily generalized to the function F defined in (1.11). Let $\rho : \Re \to \Re_+$ be a density function with a bounded absolute mean, that is

$$\kappa := \int_{-\infty}^{\infty} |s|\rho(s)ds < \infty. \tag{2.3}$$

Define the smoothing approximation $h(x, \varepsilon) = (h_i(x, \varepsilon))$ to the mid function in (1.11) by

$$h_i(x, \varepsilon) = \int_{-\infty}^{\infty} \operatorname{mid}(l_i, u_i, q_i(x) - p_i(x) - \varepsilon s)\rho(s)ds.$$

Let

$$f(x, \varepsilon) = q(x) - h(x, \varepsilon). \tag{2.4}$$

Following Lemma 2.3 and Theorem 3.3 in [18], we can show that f_i is continuously differentiable with respect to x and satisfies (2.1) with $\mu = \sqrt{n}\kappa$. Hence f is a smoothing approximation function of F.

Now we show that f has the Jacobian consistency property. For $i = 1, 2, ..., n$, it is easy to verify

(2.5)
$$\partial_B F_i(x) = \begin{cases} \{p_i'(x)\} & \text{if } q_i(x) - p_i(x) \in (l_i, u_i), \\ \{q_i'(x)\} & \text{if } q_i(x) - p_i(x) \notin [l_i, u_i], \\ \{q_i'(x), p_i'(x)\} & \text{if } q_i(x) - p_i(x) = l_i \text{ or } q_i(x) - p_i(x) = u_i. \end{cases}$$

Following Lemma 2.3 in [18] again, we have

$$(f_x(x, \varepsilon))_i = q_i'(x) - \left(\int_{(q_i(x)-p_i(x)-u_i)/\varepsilon}^{(q_i(x)-p_i(x)-l_i)/\varepsilon} \rho(s)ds \right) (q_i'(x) - p_i'(x)). \tag{2.6}$$

Hence for any fixed x,

$$\begin{aligned} &\lim_{\varepsilon \downarrow 0} (f_x(x, \varepsilon))_i \\ &= \begin{cases} p_i'(x) & \text{if } q_i(x) - p_i(x) \in (l_i, u_i), \\ q_i'(x) & \text{if } q_i(x) - p_i(x) \notin [l_i, u_i], \\ q_i'(x) - \left(\int_{-\infty}^0 \rho(s)ds\right) (q_i'(x) - p_i'(x)) & \text{if } q_i(x) - p_i(x) = l_i, \\ q_i'(x) - \left(\int_0^\infty \rho(s)ds\right) (q_i'(x) - p_i'(x)) & \text{if } q_i(x) - p_i(x) = u_i. \end{cases} \end{aligned}$$

Since $\int_{-\infty}^0 \rho(s)ds$ and $\int_0^\infty \rho(s)ds$ are in $[0, 1]$, we obtain (2.2). Hence f has the Jacobian consistency property.

The limit (2.2) implies that for any $\delta > 0$ there is an $\varepsilon(x, \delta) > 0$ such that for any $\varepsilon \in (0, \varepsilon(x, \delta)]$

$$\mathrm{dist}(f_x(x, \varepsilon), \partial_C F(x)) \le \delta.$$

Based on the Gabriel-Moré function, such $\varepsilon(x, \delta)$ for GVI(l, u, p, q) can be chosen as follows.

Let

$$\begin{aligned} \gamma(x) \quad = \quad & \min_{1 \le i,j \le n} \{|q_i(x) - p_i(x) - l_i|, |q_j(x) - p_j(x) - u_j| : \\ & q_i(x) - p_i(x) \ne l_i, q_j(x) - p_j(x) \ne u_j\}. \end{aligned} \tag{2.7}$$

Since $\int_0^\infty s^{-1}ds = \infty$, (2.3) implies that the density function ρ satisfies

$$\lim_{s \to \infty} |s|^2 \rho(s) = 0.$$

Thus we may choose a positive constant $\tau \le 1$ such that for any ν with $|\nu| \in (0, \tau]$,

$$\rho(\frac{1}{|\nu|}) < \nu^2.$$

This implies that for any $\varepsilon \in (0, \tau]$

$$\{\int_{-\infty}^{-1/\varepsilon} \rho(s)ds, \int_{1/\varepsilon}^\infty \rho(s)ds\} \le \int_{1/\varepsilon}^\infty \frac{1}{s^2} ds = \varepsilon. \tag{2.8}$$

If $\|q'(x) - p'(x)\| = 0$, then by (2.6) from any $\varepsilon > 0$, $f_x(x, \varepsilon) = q'(x) \in \partial_C F(x)$. Suppose that $\|q'(x) - p'(x)\| \ne 0$. Let

$$\varepsilon(x, \delta) = \min\{\tau\gamma(x), \frac{\gamma(x)\delta}{2\sqrt{n}\|q'(x) - p'(x)\|}\}.$$

If $q_i(x) - p_i(x) \notin [l_i, u_i]$, then F_i is differentiable at x and $F_i'(x) = q_i'(x)$. Hence in this case for any $\varepsilon \in (0, \varepsilon(x, \delta)]$,

$$\begin{aligned}
&\|(f_x(x,\varepsilon) - F'(x))_i\| \\
&= \left(\int_{(q_i(x)-p_i(x)-u_i)/\varepsilon}^{(q_i(x)-p_i(x)-l_i)/\varepsilon} \rho(s)ds \right) \|q_i'(x) - p_i'(x)\| \qquad (2.9)\\
&\le \left(\int_{\gamma(x)/\varepsilon}^{\infty} \frac{1}{s^2} ds \right) \|q_i'(x) - p_i'(x)\| \\
&\le \frac{\varepsilon}{\gamma(x)} \|q_i'(x) - p_i'(x)\| \\
&\le \frac{\delta}{2\sqrt{n}},
\end{aligned}$$

where the first inequality follows from (2.8) and $\tau \ge \varepsilon/\gamma(x)$.

If $q_i(x) - p_i(x) \in (l_i, u_i)$, then F_i is differentiable at x and $F_i'(x) = p_i'(x)$. Hence in this case for any $\varepsilon \in (0, \varepsilon(x, \delta)]$

$$\begin{aligned}
&\|(f_x(x,\varepsilon) - F'(x))_i\| \\
&= \left(1 - \int_{(q_i(x)-p_i(x)-u_i)/\varepsilon}^{(q_i(x)-p_i(x)-l_i)/\varepsilon} \rho(s)ds \right) \|q_i'(x) - p_i'(x)\| \qquad (2.10)\\
&= \left(\int_{-\infty}^{(q_i(x)-p_i(x)-u_i)/\varepsilon} \rho(s)ds + \int_{(q_i(x)-p_i(x)-l_i)/\varepsilon}^{\infty} \rho(s)ds \right) \|q_i'(x) - p_i'(x)\| \\
&\le \frac{2\varepsilon}{\gamma(x)} \|q_i'(x) - p_i'(x)\| \\
&\le \delta/\sqrt{n}.
\end{aligned}$$

If $q_i(x) - p_i(x) = l_i$ or $q_i(x) - p_i(x) = u_i$, then

$$(f_x(x,\varepsilon))_i = q_i'(x) - \lambda_\varepsilon(q_i'(x) - p_i'(x)),$$

where $\lambda_\varepsilon \in [0, 1]$. Thus in this case $(f_x(x,\varepsilon))_i \in \partial F_i(x)$ for any $\varepsilon > 0$.

Hence, in any case, we have

$$\operatorname{dist}((f_x(x,\varepsilon))_i, \partial F_i(x)) \le \delta/\sqrt{n}.$$

Thus

$$\operatorname{dist}(f_x(x,\varepsilon), \partial_C F(x)) \le \delta.$$

By choosing a special density function, for any fixed $x \in \Re^n$, we even can ask that

$$f_x(x,\varepsilon) \in \partial_C F(x) \qquad (2.11)$$

for any $\varepsilon > 0$ sufficiently small. For instance, we consider the uniform density function

$$\rho(s) = \begin{cases} \frac{1}{b-a} & \text{if } s \in [a,b], \\ 0 & \text{otherwise}, \end{cases}$$

where a and b are two finite real numbers and $a < b$. Notice that the uniform density function has

$$\int_{-\infty}^{c} \rho(s)ds = 0, \quad \int_{d}^{\infty} \rho(s)ds = 0 \quad \text{and} \quad \int_{c}^{d} \rho(s)ds = 1$$

for any $d \geq b$ and $c \leq a$. Then it is easy to verify that for any $\varepsilon \in (0, \frac{\gamma(x)}{\max\{|a|,|b|\}}]$,

$$\int_{(q_i(x)-p_i(x)-u_i)/\varepsilon}^{(q_i(x)-p_i(x)-l_i)/\varepsilon} \rho(s)ds = \begin{cases} 0 & \text{if } q_i(x)-p_i(x) \notin [l_i, u_i], \\ 1 & \text{if } q_i(x)-p_i(x) \in (l_i, u_i). \end{cases}$$

By (2.9) and (2.10), we obtain (2.11).

To see a smoothing approximation function with a concrete density function, we consider

$$\rho(s) = \begin{cases} 1 & \text{if } -\frac{1}{2} \leq s \leq \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

By a straightforward calculation we have that

$$\begin{aligned} f_i(x,\varepsilon) &= q_i(x) - \int_{-1/2}^{1/2} \operatorname{mid}(l_i, u_i, q_i(x)-p_i(x)-\varepsilon s)ds \\ &= \begin{cases} q_i(x) - \frac{1}{\varepsilon}(q_i(x)-p_i(x))(u_i-l_i) - \frac{1}{2}(l_i+u_i) - \frac{1}{2\varepsilon}(l_i^2-u_i^2) \\ \qquad \text{if } |q_i(x)-p_i(x)-l_i| \leq \frac{\varepsilon}{2}, |q_i(x)-p_i(x)-u_i| \leq \frac{\varepsilon}{2}, \\ \frac{1}{2}(q_i(x)+p_i(x)) - \frac{1}{2\varepsilon}(q_i(x)-p_i(x)-l_i)^2 - \frac{\varepsilon}{8} - \frac{l_i}{2} \\ \qquad \text{if } |q_i(x)-p_i(x)-l_i| \leq \frac{\varepsilon}{2}, q_i(x)-p_i(x)-u_i < -\frac{\varepsilon}{2}, \\ \frac{1}{2}(q_i(x)+p_i(x)) + \frac{1}{2\varepsilon}(q_i(x)-p_i(x)-u_i)^2 - \frac{u_i}{2} + \frac{\varepsilon}{8} \\ \qquad \text{if } |q_i(x)-p_i(x)-u_i| \leq \frac{\varepsilon}{2}, q_i(x)-p_i(x)-l_i > \frac{\varepsilon}{2}, \\ F_i(x) \qquad \text{otherwise.} \end{cases} \end{aligned}$$

We can simplify the definition of $f(x,\varepsilon)$ for special ε. For example, let

$$\bar{\varepsilon} = \min_{1 \leq i \leq n} \{u_i - l_i\}. \tag{2.13}$$

For $i = 1, 2, ..., n$ and $\varepsilon \in (0, \bar{\varepsilon}]$, the smoothing approximation function (2.4) with density function (2.12) reduces to

$$f_i(x,\varepsilon) = \begin{cases} \frac{1}{2}(p_i(x)+q_i(x)) + \frac{1}{2\varepsilon}(p_i(x)-q_i(x)+u_i)^2 + \frac{\varepsilon}{8} - \frac{u_i}{2} \\ \qquad \text{if } |q_i(x)-p_i(x)-u_i| \leq \frac{\varepsilon}{2}, \\ \frac{1}{2}(p_i(x)+q_i(x)) - \frac{1}{2\varepsilon}(p_i(x)-q_i(x)+l_i)^2 - \frac{\varepsilon}{8} - \frac{l_i}{2} \\ \qquad \text{if } |q_i(x)-p_i(x)-l_i| \leq \frac{\varepsilon}{2}, \\ F_i(x) \qquad \text{otherwise.} \end{cases} \tag{2.14}$$

Let

$$\varepsilon(x) = \min\{\bar{\varepsilon}, \gamma(x)\},$$

where $\bar{\varepsilon}$ is defined by (2.13) and $\gamma(x)$ is defined by (2.7). Then for any $\varepsilon \in (0, \varepsilon(x)]$,

$$(f_x(x,\varepsilon))_i = \begin{cases} \frac{1}{2}(p_i'(x)+q_i'(x)) & \text{if } q_i(x)-p_i(x) = u_i \text{ or } q_i(x)-p_i(x) = l_i, \\ F_i'(x) & \text{otherwise.} \end{cases}$$

By (2.5), this implies that for any $\varepsilon \in (0, \varepsilon(x)]$,

$$\operatorname{dist}(f_x(x,\varepsilon), \partial_C F(x)) = 0.$$

For the general complementarity problem GCP(p,q), $f(x,\varepsilon)$ defined by (2.14) reduces to

$$f_i(x,\varepsilon) = \begin{cases} \frac{1}{2}(p_i(x)+q_i(x)) - \frac{1}{2\varepsilon}(p_i(x)-q_i(x))^2 - \frac{\varepsilon}{8} & \text{if } |q_i(x)-p_i(x)| \leq \frac{\varepsilon}{2}, \\ F_i(x) & \text{otherwise.} \end{cases} \tag{2.15}$$

The function defined by (2.15) has been studied in [5], [42].

Some smoothing approximation functions are not derivable from the integration of probability densities, for example, the Auto-scaling interior point smooth function [3], but also satisfy the Jacobian consistency property.

3. A smoothing Newton method

In this section we assume that f has the Jacobian consistency property, and present a smoothing Newton method with a line search based on f. We prove that this method converges globally and superlinearly.

We denote

$$\Theta(x) = \frac{1}{2}\|F(x)\|^2$$

and

$$\theta_k(x) = \frac{1}{2}\|f(x, \varepsilon_k)\|^2.$$

Algorithm 3.1. *Given $\rho, \alpha, \eta \in (0,1)$, $\gamma \in (0, +\infty)$ and a starting point $x^0 \in \Re^n$. Choose $\sigma \in (0, \frac{1}{2}(1-\alpha))$, and $\mu > 0$ satisfying* (2.1).

Initial step. *Let $\beta_0 = \|F(x^0)\|$ and $\varepsilon_0 = \frac{\alpha}{2\mu}\beta_0$. For $k \geq 0$,*

1. *Solve*

$$F(x^k) + f_x(x^k, \varepsilon_k)d^k = 0. \tag{3.1}$$

Let d^k be the solution of (3.1).

2. *Let m_k be the smallest nonnegative integer m such that*

$$\theta_k(x^k + \rho^m d^k) - \theta_k(x^k) \leq -2\sigma\rho^m\Theta(x^k). \tag{3.2}$$

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

3. 3.1 *If $\|F(x^{k+1})\| = 0$, terminate.*

   3.2 *If $\|F(x^{k+1})\| > 0$ and*

$$\|F(x^{k+1})\| \leq \max\{\eta\beta_k, \alpha^{-1}\|F(x^{k+1}) - f(x^{k+1}, \varepsilon_k)\|\}, \tag{3.3}$$

we let

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

and choose an ε_{k+1} satisfying

$$0 < \varepsilon_{k+1} \leq \min\{\frac{\alpha}{2\mu}\beta_{k+1}, \frac{\varepsilon_k}{2}\} \tag{3.4}$$

and

$$dist(f_x(x^{k+1}, \varepsilon_{k+1}), \partial_C F(x^{k+1})) \leq \gamma\beta_{k+1}. \tag{3.5}$$

   3.3 *If $\|F(x^{k+1})\| > 0$ but* (3.3) *does not hold, we let $\beta_{k+1} = \beta_k$ and $\varepsilon_{k+1} = \varepsilon_k$.*

Without loss of generality, we assume that $\|F(x^k)\| \neq 0$ for all k in the following convergence analysis.

Remark 3.1. Condition (3.5) is the crucial condition for superlinear convergence of Algorithm 3.1. To guarantee global convergence, one only needs to choose a smoothing approximation function f and ignore (3.5). If f has the Jacobian consistency property, we can find an $\varepsilon_{k+1} > 0$ such that (3.4) and (3.5) hold by Definition 2.1. Moreover, we have shown, in section 2, how to choose an ε_{k+1} satisfying (3.4) and (3.5) for GVI(l,u,p,q); we also show that it is possible to choose an ε_{k+1} satisfying (3.4) and (3.5) for order complementarity problems in section 5. From the construction of Algorithm 3.1, we have

$$\|F(x^k) - f(x^k,\varepsilon_k)\| \le \alpha\|F(x^k)\| \qquad \text{for } k \ge 0 \tag{3.6}$$

and

$$\mathrm{dist}(f_x(x^k,\varepsilon_k), \partial_C F(x^k)) \le \gamma\|F(x^k)\| \qquad \text{for } k \in K,\ \ k \ge 1, \tag{3.7}$$

where

$$K = \{0\} \cup \{k \mid \|F(x^k)\| \le \max\{\eta\beta_{k-1}, \alpha^{-1}\|f(x^k,\varepsilon_{k-1}) - F(x^k)\|\},\ k \in N\}. \tag{3.8}$$

Notice that (3.7) may not hold for those k such that (3.3) fail to hold.

Lemma 3.1. *Suppose that $f_x(x^k,\varepsilon_k)$ is nonsingular. Then there exists a finite nonnegative integer m_k such that* (3.2) *holds.*

Proof. The continuous differentiability of $f(\cdot,\varepsilon_k)$ implies that θ_k is continuously differentiable and $\theta_k'(x^k) = f(x^k,\varepsilon_k)^T f_x(x^k,\varepsilon_k)$. By the construction of Algorithm 3.1, $f_x(x^k,\varepsilon_k)d^k = -F(x^k)$. Then, from (3.6) we have

$$\begin{aligned}
&\theta_k(x^k + td^k) - \theta_k(x^k)\\
&= t\theta_k'(x^k)d^k + o(t)\\
&= -tF(x^k)^T f(x^k,\varepsilon_k) + o(t)\\
&= -2t\Theta(x^k) + tF(x^k)^T(F(x^k) - f(x^k,\varepsilon_k)) + o(t)\\
&\le -2t\Theta(x^k) + 2t\alpha\Theta(x^k) + o(t)\\
&= -2t(1-\alpha)\Theta(x^k) + o(t).
\end{aligned}$$

Since $\sigma < \frac{1}{2}(1-\alpha) < 1-\alpha$, there exists a finite nonnegative integer m_k such that (3.2) holds. $\square$

Assumption 1. The level set

$$D_0 = \{x \in \Re^n : \Theta(x) \le (1+\alpha)^2\Theta(x^0)\}$$

is bounded.

Assumption 2. For any $\varepsilon \in \Re_{++}$ and $x \in D_0$, $f_x(x,\varepsilon)$ is nonsingular.

Theorem 3.1. *Suppose that Assumptions 1 and 2 hold. Then Algorithm 3.1 is well defined and the generated sequence $\{x^k\}$ remains in D_0 and satisfies*

$$\lim_{k\to 0} F(x^k) = 0. \tag{3.9}$$

Proof. Let us denote

$$K_1 = \{k \in K \mid \eta\beta_{k-1} \ge \alpha^{-1}\|f(x^k,\varepsilon_{k-1}) - F(x^k)\|\}$$

and

$$K_2 = \{k \in K \mid \eta\beta_{k-1} < \alpha^{-1}\|f(x^k,\varepsilon_{k-1}) - F(x^k)\|\}.$$

Then $K_1 \cup K_2 \cup \{0\} = K$, which is defined in (3.8). Assume that K consists of $k_0 = 0 < k_1 < k_2 < \dots$. Let k be an arbitrary nonnegative integer. Let k_j be the largest number in K such that $k_j \le k$. Then

$$\varepsilon_k = \varepsilon_{k_j} \quad \text{and} \quad \beta_k = \beta_{k_j}.$$

Notice that f is a smoothing approximation function. By the line search rule (3.2),

$$\|f(x^k,\varepsilon_{k_j})\| \le \|f(x^{k_j},\varepsilon_{k_j})\|.$$

Then by (2.1), for $j \ge 0$,

$$\begin{aligned}
\|F(x^k)\| &\le \|f(x^k,\varepsilon_k)\| + \|F(x^k) - f(x^k,\varepsilon_k)\| \\
&= \|f(x^k,\varepsilon_{k_j})\| + \|F(x^k) - f(x^k,\varepsilon_{k_j})\| \\
&\le \|f(x^{k_j},\varepsilon_{k_j})\| + \mu\varepsilon_{k_j} \\
&\le \|F(x^{k_j})\| + \mu\varepsilon_{k_j} + \mu\varepsilon_{k_j} \\
&= \beta_{k_j} + 2\mu\varepsilon_{k_j}.
\end{aligned} \tag{3.10}$$

If $j = 0$, $\beta_{k_j} = \beta_0, \varepsilon_{k_j} = \varepsilon_0$ and

$$\|F(x^k)\| \le \beta_0 + 2\mu\varepsilon_0 \le (1+\alpha)\|F(x^0)\|.$$

If $j \ge 1$, by step 3 of Algorithm 3.1,

$$\varepsilon_{k_j} \le \frac{1}{2}\varepsilon_{k_j-1} = \frac{1}{2}\varepsilon_{k_{j-1}}$$

and

$$\beta_{k_j} \le \eta\beta_{k_j-1} = \eta\beta_{k_{j-1}}, \qquad \text{if } k_j \in K_1,$$

or

$$\beta_{k_j} \le \alpha^{-1}\|f(x^{k_j},\varepsilon_{k_j-1}) - F(x^{k_j})\| \le \frac{\mu}{\alpha}\varepsilon_{k_j-1} = \frac{\mu}{\alpha}\varepsilon_{k_{j-1}} \le \frac{1}{2}\beta_{k_{j-1}}, \qquad \text{if } k_j \in K_2.$$

Let

$$r = \max\{\frac{1}{2},\eta\}.$$

Then by the definitions of ε_0 and β_0, for $j \ge 1$,

$$\varepsilon_{k_j} \le \frac{1}{2^{(j-1)}}\varepsilon_0 = \frac{1}{2^j}\frac{\alpha}{\mu}\|F(x^0)\| \tag{3.11}$$

and

$$\beta_{k_j} \le r^{j-1}\beta_0 = r^{j-1}\|F(x^0)\|. \tag{3.12}$$

Hence by (3.10), for $j \ge 1$

$$\begin{aligned}
\|F(x^k)\| &\le (r^{j-1} + \frac{\alpha}{2^{j-1}})\|F(x^0)\| \\
&\le r^{j-1}(1+\alpha)\|F(x^0)\|,
\end{aligned} \tag{3.13}$$

where the last inequality follows from the fact that $\frac{1}{2} \le r$.

Therefore in any case

$$\|F(x^k)\| \le (1+\alpha)\|F(x^0)\|.$$

This implies that the sequence $\{x^k\}$ remains in the level set D_0.

Now we prove (3.9). If K is infinite, by (3.13),

$$\lim_{k\to\infty}\|F(x^k)\| \le \lim_{j\to\infty} r^{j-1}(1+\alpha)\|F(x^0)\| = 0.$$

Hence to prove (3.9), it suffices to prove that K is infinite. Suppose that K is finite. This means that both K_1 and K_2 are finite. Let $\hat{k}$ be the largest number in K. Then for all $k > \hat{k}$,

$$\varepsilon_k = \varepsilon_{\hat{k}}, \quad \beta_k = \beta_{\hat{k}} = \|F(x^{\hat{k}})\|, \tag{3.14}$$

$$\|F(x^k)\| > \eta\beta_k = \eta\|F(x^{\hat{k}})\| > 0 \tag{3.15}$$

and

$$\alpha\|F(x^k)\| > \|f(x^k, \varepsilon_{\hat{k}}) - F(x^k)\|. \tag{3.16}$$

By (3.15), for all $k > \hat{k}$,

$$\Theta(x^k) \geq \eta^2\Theta(x^{\hat{k}}). \tag{3.17}$$

Let

$$\hat{\varepsilon} = \varepsilon_{\hat{k}},$$
$$\hat{f}(x) = f(x, \hat{\varepsilon})$$

and

$$\hat{\theta}(x) = \frac{1}{2}\|\hat{f}(x)\|^2.$$

Notice that for all $k > \hat{k}$,

$$f(x^k, \varepsilon_k) = \hat{f}(x^k) \quad \text{and} \quad \theta_k(x^k) = \hat{\theta}(x^k).$$

By Assumptions 1 and 2 there is an $M > 0$ such that for all $x \in D_0$, $\|f_x(x, \hat{\varepsilon})^{-1}\| \leq M$. Then for all $k > \hat{k}$,

$$\begin{array}{rcl} \|d^k\| & = & \|f_x(x^k, \hat{\varepsilon})^{-1}F(x^k)\| \\ & \leq & M\|F(x^k)\| \\ & \leq & M(1+\alpha)\|F(x^0)\| \\ & =: & L. \end{array}$$

If $\inf_k t_k = t^* > 0$, then from (3.17) and the line search rule (3.2), for all $k \geq 0$,

$$\hat{\theta}(x^{k+1}) - \hat{\theta}(x^k) \leq -2\sigma t_k\Theta(x^k) \leq -2\sigma t^*\eta^2\Theta(x^{\hat{k}}) < 0.$$

This, together with the monotonicity of $\{\hat{\theta}(x^k)\}_{k\geq\hat{k}}$, implies that $\hat{\theta}(x^k) \to -\infty$ as $k \to \infty$. This contradicts the fact that $\hat{\theta}(x^k) \geq 0$ for all $k \geq 0$. Hence K cannot be finite. Thus (3.9) holds.

Now we consider the case that $\inf_k t_k = 0$. Let K_0 be a subsequence of N such that $\{t_k\}_{k\in K_0}$ converges to zero. Since $\{x^k\}$ is bounded, without loss of generality, we assume that $\{x^k\}_{k\in K_0}$ converges to x^*.

By the line search rule (3.2) for all $k \geq \hat{k}$,

$$-2\sigma\rho^{m_k-1}\Theta(x^k) < \hat{\theta}(x^k + \rho^{m_k-1}d^k) - \hat{\theta}(x^k). \tag{3.18}$$

Dividing both sides by ρ^{m_k-1}, we obtain

$$\begin{array}{rcl} -2\sigma\Theta(x^k) & < & \dfrac{\hat{\theta}(x^k + \rho^{m_k-1}d^k) - \hat{\theta}(x^k)}{\rho^{m_k-1}} \\ & = & \hat{\theta}'(x^k)d^k + \displaystyle\int_0^1((\hat{\theta}'(x^k + \lambda\rho^{m_k-1}d^k) - \hat{\theta}'(x^k))d^k)d\lambda. \end{array}$$

Notice that

$$\begin{array}{rcl} \hat{\theta}'(x^k)d^k & = & -F(x^k)^T\hat{f}(x^k) \\ & = & -2\Theta(x^k) + F(x^k)(\hat{f}(x^k) - F(x^k)) \\ & \leq & -2\Theta(x^k) + 2\alpha\Theta(x^k), \end{array}$$

where the last inequality follows from (3.16).

By the continuity of $\hat{\theta}'$, the boundedness of $\{d^k\}$ and $\lim_{\substack{k\in K_0\\k\to\infty}} m_k = \infty$, we have

$$\lim_{\substack{k\in K_0\\k\to\infty}} \int_0^1 ((\hat{\theta}'(x^k + \lambda\rho^{m_k-1}d^k) - \hat{\theta}'(x^k))d^k)d\lambda = 0.$$

By taking the limit in (3.18) on the subsequence $k \in K_0$, we obtain

$$-2\sigma\Theta(x^*) \le -2(1-\alpha)\Theta(x^*) < 0.$$

This implies $\sigma \ge (1-\alpha)$, which contradicts the fact that $\sigma < (1-\alpha)/2$. Hence K cannot be finite. Thus (3.9) holds. ☐

To show the superlinear convergence rate, we give the following lemma.

Lemma 3.2. *If there exists a scalar*

$$\lambda \in [\frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}, \frac{1}{2}] \tag{3.19}$$

such that for some $k \in K$,

$$\Theta(y) - \Theta(x^k) \le -2\lambda\Theta(x^k), \tag{3.20}$$

then it holds that

$$\theta_k(y) - \theta_k(x^k) \le -2\sigma\Theta(x^k). \tag{3.21}$$

Proof. By the definition of K, we have

$$0 < \varepsilon_k \le \frac{\alpha}{2\mu}\|F(x^k)\|, \quad k \in K.$$

Hence, from (2.1), for any $y \in \Re^n$, $k \in K$,

$$\|f(y,\varepsilon_k)\| \le \|F(y)\| + \frac{\alpha}{2}\|F(x^k)\|$$

and

$$\|f(x^k,\varepsilon_k)\| \ge \|F(x^k)\| - \frac{\alpha}{2}\|F(x^k)\|.$$

Using these two inequalities and (3.20), we obtain

$$\begin{aligned}
\theta_k(y) - \theta_k(x^k) &= \frac{1}{2}\|f(y,\varepsilon_k)\|^2 - \frac{1}{2}\|f(x^k,\varepsilon_k)\|^2 \\
&\le \frac{1}{2}(\|F(y)\| + \frac{\alpha}{2}\|F(x^k)\|)^2 - \frac{1}{2}(1-\frac{\alpha}{2})^2\|F(x^k)\|^2 \\
&= \Theta(y) + \frac{1}{2}\alpha\|F(y)\|\|F(x^k)\| + \frac{\alpha^2}{4}\Theta(x^k) - (1-\alpha+\frac{\alpha^2}{4})\Theta(x^k) \\
&= \Theta(y) + \frac{1}{2}\alpha\|F(y)\|\|F(x^k)\| - (1-\alpha)\Theta(x^k) \\
&\le \Theta(y) + \alpha\sqrt{1-2\lambda}\Theta(x^k) - (1-\alpha)\Theta(x^k) \\
&= \Theta(y) - \Theta(x^k) + \alpha(1+\sqrt{1-2\lambda})\Theta(x^k) \\
&\le -(2\lambda - \alpha(1+\sqrt{1-2\lambda}))\Theta(x^k),
\end{aligned}$$

where the second and last inequalities follow from (3.20).

Let us denote

$$\phi(\lambda) = \lambda - \frac{1}{2}\alpha(1+\sqrt{1-2\lambda}).$$

To prove (3.21), it suffices to show

$$\phi(\lambda) \geq \sigma \qquad \text{for} \quad \lambda \in [\frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}, \frac{1}{2}]. \tag{3.22}$$

Since $0 < (1-\alpha-2\sigma)/(2+\alpha) < 1$,

$$\frac{(1-\alpha-2\sigma)^2}{(2+\alpha)^2} \leq \frac{1-\alpha-2\sigma}{2+\alpha}. \tag{3.23}$$

Notice that ϕ is monotone increasing in $[0, \frac{1}{2}]$. We only need to show (3.22) at $\bar{\lambda} = \frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}$. By the definition of ϕ and (3.23),

$$\begin{aligned}
\phi(\bar{\lambda}) &= \frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2} - \frac{1}{2}\alpha(1 + \frac{1-\alpha-2\sigma}{2+\alpha}) \\
&\geq \frac{1}{2} - \frac{1-\alpha-2\sigma}{2+\alpha} - \frac{\alpha}{2}(1 + \frac{1-\alpha-2\sigma}{2+\alpha}) \\
&= \frac{1}{2} - \frac{\alpha}{2} \\
&\geq \sigma.
\end{aligned}$$

This completes the proof. $\square$

Theorem 3.2. *Suppose that Assumptions 1 and 2 hold. Suppose that for an accumulation point x^* of the sequence $\{x^k\}$, all $V \in \partial_C F(x^*)$ are nonsingular and that F is semismooth at x^*. Then x^* is a solution of $F(x) = 0$ and the sequence $\{x^k\}$ generated by Algorithm 3.1 converges to x^* superlinearly. Moreover, if F is strongly semismooth at x^*, then $\{x^k\}$ converges to x^* quadratically.*

Proof. By Theorem 3.1, x^* is a solution of $F(x) = 0$. Notice that $\partial_B F(x^*) \subseteq \partial_C F(x^*)$. By Proposition 2.5 in [40] there is a neighbourhood of x^* such that x^* is the unique solution in this neighbourhood.

By Theorem 3.1, the set K defined by (3.8) is infinite, and there is a subsequence K_0 of K such that $\{x^k\}_{k \in K_0}$ converges to x^*. Now we consider the convergence behaviour of the subsequence $\{x^k\}_{k \in K_0}$.

Notice that for any $x \in \Re^n$, $\partial_C F(x)$ is a compact set. Let $V_k \in \partial_C F(x^k)$ be such that

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

By construction of Algorithm 3.1,

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

By Theorem 3.1, $\beta_k \to 0$ as $k \to \infty$. This, together with 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^*, implies that there exist $M > 0$ and $\hat{k} \geq 0$ such that for all $k \geq \hat{k}$ and $k \in K_0$, $\|f_x(x^k, \varepsilon_k)^{-1}\| \leq M$. Therefore, by the construction of Algorithm 3.1, for all $k \geq \hat{k}$ and $k \in K_0$,

$$\begin{aligned}
&\|x^k + d^k - x^*\| = \|x^k - x^* - f_x(x^k, \varepsilon_k)^{-1} F(x^k)\| \\
&= \|f_x(x^k, \varepsilon_k)^{-1}(f_x(x^k, \varepsilon_k)(x^k - x^*) - F(x^k) + F(x^*))\| \\
&\leq \|f_x(x^k, \varepsilon_k)^{-1}\| \left(\|(f_x(x^k, \varepsilon_k) - V_k)(x^k - x^*)\| \right. \\
&\qquad\qquad \left. + \|V_k(x^k - x^*) - F(x^k) + F(x^*)\|\right) \\
&\leq M(\gamma\beta_k\|x^k - x^*\| + \|V_k(x^k - x^*) - F(x^k) + F(x^*)\|).
\end{aligned} \tag{3.24}$$

Since F is semismooth at x^* if and only if each F_i is semismooth at x^* [43], by Theorem 3.2 in [43],

$$\begin{aligned}
&\|V_k(x^k - x^*) - F(x^k) + F(x^*)\| \\
&\leq \sqrt{\sum_{i=1}^{n} \|V_k^i(x^k - x^*) - F_i(x^k) + F_i(x^*)\|^2} \\
&= o(\|x^k - x^*\|) \quad \text{as} \quad k \to \infty,\ k \in K_0,
\end{aligned} \tag{3.25}$$

where V_k^i denotes the ith row of V_k. Hence

$$\|x^k + d^k - x^*\| = o(\|x^k - x^*\|) \quad \text{as} \quad k \to \infty,\ k \in K_0. \tag{3.26}$$

Furthermore, following the proof of Theorem 3.1 in [40]

$$\|F(x^k + d^k)\| = o(\|F(x^k)\|) \quad \text{as} \quad k \to \infty,\ k \in K_0. \tag{3.27}$$

Let $\lambda = \max\left\{\frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}, \frac{1-\eta^2}{2}\right\}$. Then (3.27) implies that there is $\bar{k} \geq \hat{k}$ such that $\bar{k} \in K_0$ and for any $k \geq \bar{k}$ and $k \in K_0$,

$$\Theta(x^k + d^k) - \Theta(x^k) \leq -2\lambda\Theta(x^k). \tag{3.28}$$

By Lemma 3.2, for any $k \geq \bar{k}$ and $k \in K_0$,

$$\theta_k(x^k + d^k) - \theta_k(x^k) \leq -2\sigma\Theta(x^k),$$

that is, $t_k \equiv 1$ and $x^{k+1} = x^k + d^k$ for all $k \geq \bar{k}$ and $k \in K_0$. In particular, $x^{\bar{k}+1} = x^{\bar{k}} + d^{\bar{k}}$ and from (3.28),

$$\|F(x^{\bar{k}+1})\| \leq \sqrt{1 - 2\lambda}\|F(x^{\bar{k}})\| \leq \eta\|F(x^{\bar{k}})\| = \eta\beta_{\bar{k}},$$

which implies that $x^{\bar{k}+1} \in K_0$. Repeating the above process we may prove that for all $k \geq \bar{k}$,

$$k \in K_0$$

and

$$x^{k+1} = x^k + d^k.$$

Then by using (3.26) we have proved that $\{x^k\}$ converges to x^* superlinearly.

If F is strongly semismooth at x^*, then each F_i is also strongly semismooth at x^*. By (3.25) and Lemma 2.3 in [40],

$$\|V_k(x^k - x^*) - F(x^k) + F(x^*)\| = O(\|x^k - x^*\|^2).$$

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

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

Hence the quadratic convergence follows easily from (3.24) and the above proof. □

4. Application

In this section we discuss an application of Algorithm 3.1 to general box constrained variational inequalities GVI(l, u, p, q). Conditions used in the convergence analysis of Algorithm 3.1 are Assumptions 1 and 2, and the condition that F is semismooth at a solution point x^* and all $V \in \partial_C F(x^*)$ are nonsingular. Now we consider when these conditions hold for the smoothing approximation function f defined by (2.4) and the nonsmooth function F defined by (1.11).

Let A be an $n \times n$ matrix. A is called a P_0-matrix, if its principal minors are all nonnegative, and A is called a P-matrix, if its principal minors are all positive.

We call $p : \Re^n \to \Re^n$ a uniform P-function with respect to q if there is a positive constant κ such that

$$\max_{1\le i\le n} (q_i(x) - q_i(y))(p_i(x) - p_i(y)) \ge \kappa \|x - y\|^2. \tag{4.1}$$

If q is the identity map and (4.1) holds, then we call p a uniform P-function directly.

The concepts of a P-matrix and a uniform P-function have been frequently used in complementarity and variational inequalities areas.

Proposition 4.1. *Suppose that q is norm coercive, i.e., $\|q(x)\| \to \infty$ if and only if $\|x\| \to \infty$. Then the level sets*

$$D(\Gamma) = \{x \in \Re^n \ : \ \|F(x)\| \le \Gamma\}$$

are bounded for all positive numbers Γ if one of the following two conditions is satisfied:

(i) l and u are both bounded;

(ii) p is a uniform P-function with respect to q and q is surjective and Lipschitz continuous.

Proof. Since q is norm coercive, the boundedness of $D(\Gamma)$ under assumption (i) follows easily.

Next we prove the boundedness of $D(\Gamma)$ under assumption (ii). It is not difficult to verify that for $a, b \in \Re^n$,

$$|a_i - \mathrm{mid}(l_i, u_i, a_i - b_i)| \to \infty \quad \text{as} \quad |a_i|, |b_i| \to \infty. \tag{4.2}$$

Suppose that there exists one $\Gamma > 0$ such that $D(\Gamma)$ is unbounded, i.e., there exists a sequence $\{x^k\} \subseteq D(\Gamma)$ such that $\|x^k\| \to \infty$. Since q is norm coercive, $\{\|q(x^k)\|\}$ is unbounded. Define the index set J by $J := \{i| \ \{q_i(x^k)\}$ is unbounded, $i = 1, 2, ..., n\}$. Then $J \ne \emptyset$. Since q is surjective, we can choose $y^k \in \Re^n$ such that

$$q_i(y^k) = \begin{cases} q_i(x^k) & \text{if } i \notin J, \\ 0 & \text{if } \ i \in J. \end{cases}$$

Then $\{\|q(y^k)\|\}$, and so $\{\|y^k\|\}$, is bounded. Since p is a uniform P-function with respect to q, there is a positive number κ such that

$$\begin{aligned} \kappa\|x^k - y^k\|^2 &\le \max_{1\le i\le n} (q_i(x^k) - q_i(y^k))(p_i(x^k) - p_i(y^k)) \\ &\le \max_{1\le i\le n} |q_i(x^k) - q_i(y^k)||p_i(x^k) - p_i(y^k)|, \end{aligned}$$

which, together with the Lipschitz continuity of q, implies that

$$\begin{aligned} L^{-2}\kappa \sum_{i\in J}(q_i(x^k))^2 &= L^{-2}\kappa\|q(y^k)-q(x^k)\|^2 \\ &\le \kappa\|x^k-y^k\|^2 \\ &\le \max_{1\le i\le n}|q_i(y^k)-q_i(x^k)||p_i(x^k)-p_i(y^k)| \\ &= \max_{i\in J}|q_i(x^k)||p_i(x^k)-p_i(y^k)| \\ &\le \sqrt{\sum_{i\in J}(q_i(x^k))^2}\max_{i\in J}|p_i(x^k)-p_i(y^k)|, \end{aligned}$$

where L is the Lipschitz constant of q. Then $\max_{i\in J}|p_i(x^k)-p_i(y^k)|\to\infty$ as $k\to\infty$. Since $\{\|p(y^k)\|\}$ is bounded, for each k there exists at least one $i_k\in J$ such that

$$|p_{i_k}(x^k)|\to\infty.$$

Since J has only a finite number of elements, by taking a subsequence if necessary, we may assume that there exists an $i\in J$ such that

$$|p_i(x^k)|\to\infty.$$

Then we have proved that there exists at least one $i\in J$ such that

$$|q_i(x^k)|,|p_i(x^k)|\to\infty,$$

which, together with (4.2), implies that for such i, $\{|F_i(x^k)|\}$ is unbounded. This is a contradiction. So for every $\Gamma>0$, the level set $D(\Gamma)$ is bounded. □

Remark 4.1. Let $I=\{i|\ l_i=-\infty \text{ or } u_i=\infty,\ i=1,...,n\}$. Notice that the index set J in the proof of Proposition 4.1 is a subset of I. Conditions (i) and (ii) in Proposition 4.1 can be replaced by the following condition

$$\kappa\sum_{i\in I}(q_i(x)-q_i(y))^2\le\max_{1\le i\le n}(q_i(x)-q_i(y))(p_i(x)-p_i(y)). \tag{4.3}$$

Moreover, if condition (i) holds, then $I=\emptyset$, so (4.3) holds directly. If condition (ii) holds, then from

$$\sum_{i\in I}(q_i(x)-q_i(y))^2\le\|q(x)-q(y)\|^2,$$

we have (4.3).

Proposition 4.2. *Suppose that $q'(x)$ is nonsingular at $x\in\Re^n$.*

(i) If $p'(x)q'(x)^{-1}$ is a P-matrix, then for any $\varepsilon>0$, $f_x(x,\varepsilon)$ and all elements in $\partial_C F(x)$ are nonsingular;

(ii) If $p'(x)q'(x)^{-1}$ is a P_0-matrix and the support of the density function ρ,

$$\operatorname{supp}(\rho)=\{s\ :\ \rho(s)>0\}$$

is the whole real line. Then for any $\varepsilon>0$, $f_x(x,\varepsilon)$ is nonsingular.

Proof. (i) By (2.6),

$$f_x(x,\varepsilon) = q'(x) - D(x)(q'(x) - p'(x)), \tag{4.4}$$

where $D(x) = \mathrm{diag}(d_i(x))$ has $d_i(x) \in [0,1]$. Let

$$\mathcal{P}(x) := q'(x) - \begin{pmatrix} [0,1] & & \\ & \ddots & \\ & & [0,1] \end{pmatrix} (q'(x) - p'(x)).$$

By Lemma 5.1 in [39], all elements in $\mathcal{P}(x)$ are nonsingular. Obviously, $f_x(x,\varepsilon) \in \mathcal{P}(x)$ and $\partial_C F(x) \subseteq \mathcal{P}(x)$. Hence we obtain (i).

(ii) $\mathrm{Supp}(\rho) = \Re$ implies that the diagonal matrix $D(x)$ in (4.4) has $d_i(x) \in [0,1)$. By Theorem 4.2 in [18],

$$I - D(x)(I - p'(x)q'(x)^{-1})$$

is nonsingular. Thus

$$f_x(x,\varepsilon) = (I - D(x)(I - p'(x)q'(x)^{-1}))q'(x)$$

is nonsingular. □

The condition in Proposition 4.1 that $p'(x)q'(x)^{-1}$ is a P-matrix can be weakened to the condition that $[p'(x)q'(x)^{-1}]_{\mathcal{I}}$ is nonsingular and its Schur complement in $[p'(x)q'(x)^{-1}]_{\mathcal{I}\cup\mathcal{B}}$,

$$[p'(x)q'(x)^{-1}]_{\mathcal{I}\cup\mathcal{B}} / [p'(x)q'(x)^{-1}]_{\mathcal{I}},$$

is a P-matrix, where

$$\mathcal{I} = \{i \ : \ l_i < q_i(x) - p_i(x) < u_i\},$$
$$\mathcal{B} = \{i \ : \ q_i(x) - p_i(x) = l_i\} \cup \{i \ : \ q_i(x) - p_i(x) = u_i\}.$$

Corollary 4.1. *Suppose that $p, q : \Re^n \to \Re^n$ are continuously differentiable, p is a uniform P-function with respect to q, and q is norm coercive, surjective and Lipschitz continuous. Then the iteration sequence $\{x^k\}$ generated by Algorithm 3.1 for F given by* (1.11) *and f defined by* (2.4) *is well defined and converges to the unique solution x^* of $F(x) = 0$ superlinearly. Furthermore, if p' and q' are locally Lipschitz continuous around x^*, the convergence is quadratic.*

Proof. By the assumption that p is a uniform P-function with respect to q, there exists a $\kappa > 0$ such that (4.1) holds. Since p, q are continuously differentiable, from (4.1) for any $x \in \Re^n$ we have

$$\max_{1\le i\le n} (q'(x)z)_i (p'(x)z)_i \ge \kappa z^T z \quad \text{for all} \quad z \in \Re^n.$$

This, from Lemma 5.1 of [39], implies that both $q'(x)$ and $p'(x)$ are nonsingular and $p'(x)q'(x)^{-1}$ is a P-matrix. Then from Proposition 4.2, for any $x \in \Re^n$ and $\varepsilon > 0$, $f_x(x,\varepsilon)$ is nonsingular and all elements in $\partial_C F(x)$ are nonsingular. The boundedness of D_0 follows from Proposition 4.1. The uniqueness of the solution follows from the nonsingularity of all elements in $\partial_C F(x)$. Finally, F is piecewise smooth, hence semismooth everywhere, from Theorem 3.2 we may conclude that $\{x^k\}$ is well defined and converges to the unique solution x^* superlinearly.

The quadratic convergence follows from the fact that when p' and q' are Lipschitz continuous around x^*, F is strongly semismooth at x^*. □

Remark 4.2. When q is the identity map, Corollary 4.1 says that if p is a continuously differentiable uniform P-function, then the sequence $\{x^k\}$ generated by Algorithm 3.1 is well defined and converges to the unique solution of box constrained variational inequalities superlinearly. Such a result was only obtained for the nonlinear complementarity problem in [8]. In [27], [48], a similar result for box constrained variational inequalities was obtained by assuming that p is a strongly monotone function, which is a stronger condition than that of a uniform P-function.

5. Final remarks

In this paper we have shown that the smoothing Newton method for solving nonsmooth equations can have a convergence rate better than the linear rate. We have established global and superlinear convergence of the smoothing Newton method based on the Jacobian consistency property. Furthermore, we have investigated our conditions used in the convergence analysis for general box constrained variational inequalities GVI(l, u, p, q). We have shown that all assumptions hold if p is a uniform P-function with respect to q and q is norm coercive, surjective and Lipschitz continuous. In contrast with other methods for GVI(l, u, p, q), the smoothing Newton method has three advantages: solution of a linear system at each iteration, the boundedness of the iterates and superlinear convergence rate.

Algorithm 3.1 and the convergence analysis in section 3 are not restricted to GVI(l, u, p, q). More applications of the smoothing Newton method are possible. For example, we can use the smoothing Newton method to solve the order complementarity problem. The order complementarity problem is to find an $x \in \Re^n$ such that

$$q^i(x) \geq 0,\ i = 1, ..., m, \quad \prod_{i=1}^{m} q^i_j(x) = 0, \quad j = 1, ..., n, \tag{5.1}$$

where all $q^i : \Re^n \to \Re^n$ are continuously differentiable. Problem (5.1) has received as increasing amount of interest recently [19], [24]. Let "min" be the component minimum operator. Problem (5.1) is equivalent to finding a zero of the following nonsmooth equation

$$\min(q^i(x) : i = 1, ..., m) = 0.$$

For simplicity, let us consider

$$F(x) = \min(r(x), q(x), p(x)). \tag{5.2}$$

By using the results of Zang [52], (5.2) is equivalent to

$$F(x) = r(x) - \max(r(x) - q(x) + \max(q(x) - p(x), 0), 0).$$

A smoothing approximation function f, which has the Jacobian consistency property, can be given as

$$\begin{aligned} &f_i(x, \varepsilon) = r_i(x) \\ &- \int_{-\infty}^{\infty} \max\left(r_i(x) - q_i(x) - \varepsilon t + \int_{-\infty}^{\infty} \max(q_i(x) - p_i(x) - s\varepsilon, 0)\rho(s)ds, 0 \right) \rho(t)dt, \end{aligned}$$

where $\rho : \Re \to \Re_+$ is a density function with a bounded absolute mean.

Another possible application of the smoothing Newton method is for the variational inequality problem (1.3) with X given by

$$X = \{x \in R^n |\ g(x) \leq 0,\ h(x) = 0,\ l \leq x \leq u\},$$

where $g : R^n \to R^{m_1}$ and $h : R^n \to R^{m_2}$ are assumed to be twice continuously differentiable. The Karush-Kuhn-Tucker conditions of problem (1.3) can be written as [49]

$$\begin{pmatrix} x - \Pi_{[l,u]}[x - L(x,\lambda,\mu)] \\ \\ \lambda - \Pi_{R^n_+}[\lambda - (-g(x))] \\ \\ -h(x) \end{pmatrix} = 0, \tag{5.3}$$

where

$$L(x,\lambda,\mu) = p(x) + \sum_{i=1}^{m_1} \nabla g_i(x)\lambda_i + \sum_{j=1}^{m_2} \nabla h_j(x)\mu_j.$$

Problem (5.3) is a special box constrained variational inequality problem, so the smoothing Newton method can be used directly to find a solution of (5.3), and in turn to find a solution of (1.3) under some constraint qualification conditions [20]. However for problem (5.3), the boundedness assumption on the level sets may not hold even if p is strongly monotone. We will leave this and the comparison of different smoothing approximation functions in the smoothing Newton methods as further research topics.

Acknowledgment

The authors would like to thank H. Jiang, C. Kanzow, S.M. Robinson and two referees for their helpful comments.

References

[1] O. Axelsson and I.E. Kaporin, *On the solution of nonlinear equations for nondifferentiable mappings*, Technical Report, Department of Mathematics, University of Nijmegen, The Netherlands, 1994.

[2] S.C. Billups, S.P. Dirkse and M.C. Ferris, *A comparison of algorithms for large-scale mixed complementarity problems*, Comp. Optim. Appl., 7 (1997), pp. 3-25. CMP 97:09

[3] B. Chen and P. Harker, *Smooth approximations to nonlinear complementarity problems*, SIAM J. Optim., 7 (1997), pp. 403-420. CMP 97:11

[4] C. Chen and O.L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems,* Comp. Optim. Appl., 5 (1996), pp. 97-138. MR **96m:**90102

[5] X. Chen and L. Qi, *A parameterized Newton method and a quasi-Newton method for solving nonsmooth equations*, Comp. Optim. Appl., 3 (1994), pp. 157-179. MR **95e:**90090

[6] X. Chen and T. Yamamoto, *Convergence domains of certain iterative methods for solving nonlinear equations,* Numer. Funct. Anal. Optim., 10 (1989), pp. 37-48. MR **90a:**65137

[7] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983. MR **85m:**49002

[8] T. De Luca, F. Facchinei and C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems,* Math. Programming, 75 (1996), pp. 407-439. CMP 97:05

[9] S.P. Dirkse and M.C. Ferris, *The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems*, Optim. Method. Softw., 5 (1995), pp. 123-156.

[10] B.C. Eaves, *On the basic theorem of complementarity,* Math. Programming, 1 (1971), pp. 68-75. MR **44:**5103

[11] F. Facchinei, A. Fischer and C. Kanzow, *Inexact Newton methods for semismooth equations with applications to variational inequality problems,* in: G. Di Pillo and F. Giannessi, eds., "Nonlinear Optimization and Applications", Plenum Press, New York, 1996, pp. 155-170.

[12] F. Facchinei, A. Fischer and C. Kanzow, *A semismooth Newton method for variational inequalities: Theoretical results and preliminary numerical experience,* Preprint 102, Institute of Applied Mathematics, University of Hamburg, Hamburg, December 1995.

[13] F. Facchinei, A. Fischer and C. Kanzow, *A semismooth Newton method for variational inequalities: The case of box constraints*, in: M.C. Ferris and J.S. Pang, eds., "Complementarity and Variational Problems: State of the Art," SIAM, Philadelphia, Pennsylvania, 1997, pp. 76-90. CMP 97:11

[14] F. Facchinei and C. Kanzow, *A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems,* Math. Programming, 76 (1997), pp. 493-512. CMP 97:08

[15] A. Fischer, *A special Newton-type optimization method,* Optim., 24 (1992), pp. 269-284. MR **94g:**90109

[16] A. Fischer, *An NCP-function and its use for the solution of complementarity problems,* in: D. Du, L. Qi and R. Womersley, eds., "Recent Advances in Nonsmooth Optimization", World Scientific Publishers, New Jersey, 1995, pp. 88-105.

[17] M. Fukushima, *Merit functions for variational inequality and complementarity problems,* in: G. Di Pillo and F. Giannessi eds., "Nonlinear Optimization and Applications", Plenum Publishing Corporation, New York, 1996, pp. 155-170.

[18] S.A. Gabriel and J.J. Moré, *Smoothing of mixed complementarity problems*, in: M.C. Ferris and J.S. Pang, eds., "Complementarity and Variational Problems: State of the Art," SIAM, Philadelphia, Pennsylvania, 1997, pp. 105-116. CMP 97:11

[19] M.S. Gowda and R. Sznajder, *The generalized order linear complementarity problem*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 779-795. MR **95e:**90108

[20] 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), pp. 161-220. MR **91g:**90166

[21] P.T. Harker and B. Xiao, *Newton's method for the nonlinear complementarity problem: A B-differentiable equation approach,* Math. Programming, 48 (1990), pp. 339-357. MR **91m:**90176

[22] M. Heinkenschloß, C.T. Kelley and H.T. Tran, *Fast algorithms for nonsmooth compact fixed point problems*, SIAM J. Numer. Anal., 29 (1992), pp. 1769-1792. MR **94c:**65073

[23] C.M. Ip and J. Kyparisis, *Local convergence of quasi-Newton methods for B-differentiable equations,* Math. Programming, 56 (1992), pp. 71-89. MR **93c:**65072

[24] G. Isac and M. M. Kostreva, *The generalized order complementarity problem, models and iterative methods*, Ann. Oper. Res., 44 (1993), pp. 63-92.

[25] H. Jiang and L. Qi, *A new nonsmooth equations approach to nonlinear complementarity problems,* SIAM J. Control Optim., 35 (1997), pp. 178-193. CMP 97:07

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

[27] C. Kanzow and M. Fukushima, *Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities,* Math. Programming, to appear.

[28] C. Kanzow and H. Jiang, *A continuation method for (strongly) monotone variational inequalities,* Math. Programming, to appear.

[29] B. Kummer, *Newton's method for non-differentiable functions,* in: J. Guddat, B. Bank, H. Hollatz, P. Kall, D. Klatte, B. Kummer, K. Lommatzsch, L. Tammer, M. Vlach and K. Zimmerman, eds., "Advances in Mathematical Optimization", Akademi-Verlag, Berlin, 1988, pp. 114-125. MR **89h:**90201

[30] Z.Q. Luo and P. Tseng, *A new class of merit functions for the nonlinear complementarity problem*, in: M.C. Ferris and J.S. Pang, eds., "Complementarity and Variational Problems: State of the Art," SIAM, Philadelphia, Pennsylvania, 1997, pp. 204-225. CMP 97:11

[31] J.M. Martínez, and L. Qi, *Inexact Newton methods for solving nonsmooth equations,* J. Comp. Appl. Math., 60 (1995), pp. 127-145. MR **96h:**65076

[32] J.M. Ortega and W.C. Rheinboldt, Iterative Solutions of Nonlinear Equations in Several Variables, Academic Press, New York, 1970. MR **42:**8686

[33] J.V. Outrata, *On optimization problems with variational inequality constraints,* SIAM J. Optim., 4 (1994), pp. 340-357. MR **95g:**90075

[34] J.V. Outrata and J. Zowe, *A Newton method for a class of quasi-variational inequalities,* Comp. Optim. Appl., 4 (1995), pp. 5-21. MR **95k:**49023

[35] J.S. Pang, *Newton's method for B-differentiable equations,* Math. Oper. Res., 15 (1990), pp. 311-341. MR **91m:**49029

[36] J.S. Pang, *A B-differentiable equation based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems,* Math. Programming, 51 (1991), pp. 101-131. MR **92f:**90062

[37] J.S. Pang, *Complementarity problems,* in: R. Horst and P. Pardalos, eds., "Handbook of Global Optimization", Kluwer Academic Publishers, Boston, pp. 271-338, 1995. MR **97a:**90095

[38] J.S. Pang and L. Qi, *Nonsmooth equations: Motivation and algorithms,* SIAM J. Optim., 3 (1993), pp. 443–465. MR **94g:**90145

[39] J.S. Pang and J.C. Yao, *On a generalization of a normal map and equation*, SIAM J. Control Optim., 33 (1995), pp. 168-184. MR **95m:**90131

[40] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res., 18 (1993), pp. 227-244. MR **95f:**65109

[41] L. Qi, *C-differentiability, C-Differential operators and generalized Newton methods*, AMR 96/5, Applied Mathematics Report, University of New South Wales, Sydney, 1996.

[42] L. Qi and X. Chen, *A globally convergent successive approximation method for severely nonsmooth equations*, SIAM J. Control Optim., 33 (1995), pp. 402-418. MR **96a:**90035

[43] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Math. Programming, 58 (1993), pp. 353-367. MR **94b:**90077

[44] D. Ralph, *Global convergence of damped Newton's method for nonsmooth equations via the path search,* Math. Oper. Res., 19 (1994), pp. 352-389. MR **95e:**49012

[45] D. Ralph and S. Wright, *Superlinear convergence of an interior-point method for monotone variational inequalities*, Technical Report MCS-P556-0196, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, 1996.

[46] W.C. Rheinboldt, *A unified convergence theory for a class of iterative processes*, SIAM J. Numer. Anal., 5 (1968), pp. 42-63. MR **37:**1061

[47] S. M. Robinson, *Newton's method for a class of nonsmooth functions,* Set-Valued Anal., 2 (1994), pp. 291-305. MR **95c:**65093

[48] D. Sun, M. Fukushima and L. Qi, *A computable generalized Hessian of the D-gap function and Newton-type methods for variational inequality problem*, in: M.C. Ferris and J.S. Pang, eds., "Complementarity and Variational Problems: State of the Art," SIAM, Philadelphia, Pennsylvania, 1997, pp. 452-473. CMP 97:11

[49] D. Sun and J. Han, *Newton and quasi-Newton methods for a class of nonsmooth equations and related problems*, SIAM J. Optim., 7 (1997), pp. 463-480. CMP 97:11

[50] T. Yamamoto, *Split nonsmooth equations and verification of solution*, in: "Numerical Analysis, Scientific Computing, Computer Science", the Zeitschrift fuer Angewandte Mathematik und Mechanik (ZAMM) with the Akademie Verlag, Berlin, 1996, pp. 199-202. CMP 97:11

[51] N. Yamashita and M. Fukushima, *Modified Newton methods for solving semismooth reformulations of monotone complementarity problems,* Math. Programming, 76 (1997), pp. 469-491. CMP 97:08

[52] I. Zang, *A smoothing-out technique for min–max optimization*, Math. Programming, 19 (1980), pp. 61-77. MR **82c:**90085

School of Mathematics, The University of New South Wales, Sydney 2052, Australia
E-mail address: `X.Chen@unsw.edu.au`

E-mail address: `L.Qi@unsw.edu.au`

E-mail address: `sun@alpha.maths.unsw.edu.au`