Applied Numerical Mathematics 30 (1999) 3–22

On the local convergence of quasi-Newton methods for nonlinear complementarity problems

Vera Lúcia Rocha Lopes *,1, José Mario Martínez 2, Rosana Pérez 3

Departamento de Matemática Aplicada, IMECC-UNICAMP, Caixa Postal 6065, 13083-970 Campinas, SP, Brazil

Abstract

A family of Least-Change Secant-Update methods for solving nonlinear complementarity problems based on nonsmooth systems of equations is introduced. Local and superlinear convergence results for the algorithms are proved. Two different reformulations of the nonlinear complementarity problem as a nonsmooth system are compared, both from the theoretical and the practical point of view. A global algorithm for solving the nonlinear complementarity problem which uses the algorithms introduced here is also presented. Some numerical experiments show a good performance of this algorithm. © 1999 Elsevier Science B.V. and IMACS. All rights reserved.

Keywords: Nonsmooth systems; Nonlinear complementarity problems; Quasi-Newton methods

1. Introduction

Let $F:\mathbb{R}^n \to \mathbb{R}^n$, $F(x) = (f_1(x), f_2(x), \dots, f_n(x))$ be a continuously differentiable mapping. The nonlinear complementarity problem (NCP) consists of finding a vector $x \in \mathbb{R}^n$ such that $x \geqslant 0$, $F(x) \geqslant 0$, $\langle x, F(x)\rangle = 0$. Variational inequalities problems, linear complementarity problems, mixed complementarity and horizontal complementarity problems are related with the NCP. The NCP appears in many problems of Physics and Economy (see [4,6,15]). In the last few years, much work has been done with the aim of finding efficient Newton-type algorithms to solve the NCP, trying to find merit functions whose minimizers agree with the solutions of the NCP and also seeking methods with good local convergence rate (see [8]). A well-known way to deal with the NCP is to reformulate it as a nonsmooth nonlinear system of equations. See [16] and references therein. In [19] a quasi-Newton approach for nonsmooth nonlinear equations is proposed:

$$G(x) = \min\{x, F(x)\}, \tag{1}$$

* Corresponding author. E-mail: vlopes@ime.unicamp.br.

[1] This author was supported by FINEP, FAPESP and FAEP-UNICAMP.

[2] This author was supported by CNPq, FAPESP and FAEP-UNICAMP.

[3] This author was supported by FAPESP, Grant 96/0607-5.

PII: S0168-9274(98)00080-4

where min is taken componentwise (see [16]). It is easy to verify that x is a zero of G if and only if x solves the NCP. If there exists i such that $x_i = f_i(x)$, the function G may be nonsmooth at x. The second function is the Burmeister–Fischer function $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\Phi(x) = \big(\varphi\big(x_1, f_1(x)\big), \varphi\big(x_2, f_2(x)\big), \ldots, \varphi\big(x_n, f_n(x)\big)\big), \tag{2}$$

where $\varphi : \mathbb{R}^2 \to \mathbb{R}$ is given by

$$\varphi(x, y) = \big\|(x, y)\big\|_2 - x - y.$$

See [5].

This function is such that $\varphi(x, y) = 0$ if and only if $x \geqslant 0$, $y \geqslant 0$ and $xy = 0$ and so it is obvious that the NCP is equivalent to solving the nonlinear system $\Phi(x) = 0$. If there exists i such that $x_i = f_i(x) = 0$, we also have that the function Φ is nonsmooth at x. The function G will be called here the "Min function" while Φ will be called the "Fischer function". In this work we develop and analyze methods to solve the NCP using systems (1) and (2). In both cases we develop a family of Least-Change Secant-Update (LCSU) methods, following the lines of [12]. For these families we prove local and superlinear convergence under suitable assumptions.

This work is organized as follows. In Section 2 we state the main assumptions, we prove some consequences and we develop the LCSU theory for the Min function. In Section 3 we state similar hypotheses under which the same results hold for the Fischer function. Section 4 is dedicated to show that the assumption of BD-regularity of G at x^* made in Section 2 is not equivalent to the assumption of Φ having all the elements nonsingular in a subset Z_* of $\partial_B \Phi(x^*)$ made in Section 3. In Section 5 we present some numerical results which show the sensitivity of the Fischer function to degeneracy. We also present the numerical performance of both formulations when applied to 16 test problems given in [7]. In Section 6 we present a globalizing strategy to solve the NCP which take the ideas of the hybrid algorithm given in [7], using the global algorithm given in [2]. We present the results of some numerical experiments which show a good performance of our algorithm. Section 7 contains some remarks on what we have done in this paper and presents some possibilities for future work.

A few words about notation. Given a matrix $M \in \mathbb{R}^{n \times n}$ we denote by $[M]_i$ its ith row. We will denote the Jacobian matrix of F at x by $F'(x)$.

$$\left(\frac{\partial f_i}{\partial x_1}(x), \ldots, \frac{\partial f_i}{\partial x_n}(x)\right)$$

is denoted by $f_i'(x)$ and $\mathcal{B}(x, \varepsilon)$ means the open ball centered at x with radius ε.

We finish this section recalling some concepts which we use in the text (see [1,16,18]).

Definition 1.1. Let $H : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitzian function and let D_H denotes the set where H is differentiable. For all $x \in \mathbb{R}^n$, the set given by

$$\partial_B H(x) = \Big\{ \lim_{x^k \to x} H'\big(x^k\big) \colon x^k \in D_H \Big\},$$

is called the generalized B-Jacobian of H at x.

Definition 1.2. The convex hull of $\partial_B H(x)$ is called $\partial H(x)$.

Definition 1.3. Let $H:\mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitzian at $x \in \mathbb{R}^n$. We say that G is semismooth at x if

$$\lim_{\substack{V \in \partial H(x+ty') \\ y' \to y,\ t \downarrow 0}} V y'$$

exists for every $y \in \mathbb{R}^n$.

Definition 1.4. We say that a semismooth function $H:\mathbb{R}^n \to \mathbb{R}^n$ is BD-regular at x if all elements in $\partial_B H(x)$ are nonsingular.

2. A local convergence theory for $G(x) = 0$

2.1. Introduction

In this section we present a family of LCSU methods for solving $G(x) = 0$, where G is given by (1), and we prove that the algorithms are locally and superlinearly convergent. Given $x^0 \in \mathbb{R}^n$ an initial approximation to the solution of the problem, the basic quasi-Newton algorithm applied to $G(x) = 0$ will be given by

$$x^{k+1} = x^k - B_k^{-1} G(x^k),$$

where each row of B_k is given by

$$[B_k]_i = \begin{cases} \boldsymbol{e}_i & \text{if } x_i^k < f_i(x^k), \\ [A_k]_i & \text{if } x_i^k > f_i(x^k), \\ \boldsymbol{e}_i \ \text{ or } [A_k]_i & \text{if } x_i^k = f_i(x^k). \end{cases} \tag{3}$$

Here $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ is the canonical basis of $\mathbb{R}^n$ and

$$A_k = \begin{pmatrix} [A_k]_1 \\ \vdots \\ [A_k]_n \end{pmatrix}$$

is an approximation of the Jacobian matrix of F at x^k. Most times, the matrix A_{k+1} is obtained from A_k using secant updates.

2.2. Local assumptions and convergence results

Under the following assumptions we will prove that the sequences generated by the basic quasi-Newton algorithm of Section 2.1 are well defined and converge linearly to a solution of $G(x) = 0$.

(A1) $x^* \in \mathbb{R}^n$ is such that $G(x^*) = 0$.

(A2) There exist $\gamma > 0$, $\tilde{\varepsilon} > 0$ such that

$$\left\| F'(x) - F'(x^*) \right\| \leqslant \gamma \|x - x^*\|$$

for all $x \in \mathcal{B}(x^*, \tilde{\varepsilon})$, where $\|\cdot\|$ denotes an arbitrary norm on $\mathbb{R}^n$ and its associated matricial norm.

(A3) If

$$B_* = \begin{pmatrix} [B_*]_1 \\ \vdots \\ [B_*]_n \end{pmatrix}$$

is such that, for all $i = 1, \ldots, n$,

$$[B_*]_i \in \{ f_i'(x^*), \boldsymbol{e}_i \},$$

and

$$[B_*]_i = \begin{cases} \boldsymbol{e}_i & \text{if } x_i^* < f_i(x^*), \\ f_i'(x^*) & \text{if } x_i^* > f_i(x^*), \end{cases} \tag{4}$$

then B_* is nonsingular.

We observe that $\partial_B G(x^*) = \{G'(x^*)\}$ if x^* is nondegenerate and that associated to each component x_i^* of x^*, for which $x_i^* = f_i(x^*) = 0$, there are two matrices whose ith rows are given by (4) in $\partial_B G(x^*)$. So, assumption (3) means that we are assuming that the function G is BD-regular at x^*. Since there are at most 2^m of such matrices, where m is the number of degenerate components of x^*, we can define $\theta > 0$, a bound to $\|B_*^{-1}\|$ for all of them.

The next lemma prepares the "theorem of the two neighborhoods". Its proof follows well-known arguments of quasi-Newton theories but is included here for the sake of completeness.

For each $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, define

$$\Gamma(x, A) = x - B^{-1}G(x), \tag{5}$$

where

$$B = \begin{pmatrix} [B]_1 \\ \vdots \\ [B]_n \end{pmatrix}, \quad \text{with } [B]_i \in \{[A]_i, \boldsymbol{e}_i\}, \tag{6}$$

and

$$[B]_i = \begin{cases} \boldsymbol{e}_i & \text{if } x_i < f_i(x), \\ [A]_i & \text{if } x_i > f_i(x). \end{cases} \tag{7}$$

As observed after (4), if $x_i = f_i(x)$, then $[B_*]_i$ will be either $\boldsymbol{e}_i$ or $[A]_i$.

Until the end of this section, $\|\cdot\|$ means $\|\cdot\|_\infty$.

Lemma 2.1. *Let all the local assumptions be verified and let $r \in (0, 1)$. Then there exist ε_1 and δ_1 such that, if*

$$\|x - x^*\| \leqslant \varepsilon_1 \quad \textit{and} \quad \|A - F'(x^*)\| \leqslant \delta_1,$$

then the function $\Gamma(x, A)$ is well defined, and satisfies

$$\|\Gamma(x, A) - x^*\| \leqslant r\|x - x^*\|.$$

Proof. Let $\varepsilon_1 > 0$ be such that, for all $i = 1, \ldots, n$,

$$\text{if } f_i(x^*) > x_i^* \quad \text{then } f_i(x) > x_i,$$
$$\text{if } f_i(x^*) < x_i^* \quad \text{then } f_i(x) < x_i$$

for all $x \in \mathcal{B}(x^*, \varepsilon_1)$. It is obvious that ε_1 exists by the continuity of F. Let $\delta_1 \leqslant r/(4\theta)$, where θ is given by assumption (3) for $\|\cdot\|_\infty$.

For each $x \in \mathcal{B}(x^*, \varepsilon_1)$, we take $A \in \mathcal{B}(F'(x^*), \delta_1)$ and B associated to A by (6) and (7) and we consider in $\partial_B G(x^*)$ the matrix B_* that corresponds to the matrix B, that is to say, B_* is the matrix in $\partial_B G(x^*)$ that has for the ith row either $\boldsymbol{e}_i$ or ${f_i}'(x^*)$ according to the ith row of the matrix B being $\boldsymbol{e}_i$ or $[A]_i$. Then it is easily seen that

$$\|B - B_*\| \leqslant \delta_1. \tag{8}$$

So, by Banach lemma, B^{-1} exists and

$$\|B^{-1}\| \leqslant 2\|B_*^{-1}\| \leqslant 2\theta. \tag{9}$$

From (5), (8) and (9):

$$\begin{aligned}\|\varGamma(x, A) - x^*\| &= \|(x - x^*) - B^{-1}G(x) + B^{-1}B_*(x - x^*) - B^{-1}B_*(x - x^*)\| \\ &= \|(I - B^{-1}B^*)(x - x^*) - B^{-1}[G(x) - G(x^*) - B_*(x - x^*)]\| \\ &\leqslant \|B^{-1}\|[\|B - B^*\|\|x - x^*\| + \|G(x) - G(x^*) - B_*(x - x^*)\|] \\ &\leqslant 2\theta\left[\delta_1 + \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|}\right]\|x - x^*\|.\end{aligned} \tag{10}$$

But $g_i(x) - g_i(x^*) \in \{x_i - x_i^*, f_i(x) - f_i(x^*)\}$, and, by the continuity of f_i,

$$g_i(x) - g_i(x^*) = \begin{cases} x_i - x_i^* & \text{if } x_i^* < f_i(x^*), \\ f_i(x) - f_i(x^*) & \text{if } x_i^* > f_i(x^*), \end{cases} \tag{11}$$

$$[B_*(x - x^*)]_i = \begin{cases} x_i - x_i^* & \text{if } x_i^* < f_i(x^*), \\ f_i'(x^*)(x - x^*) & \text{if } x_i^* > f_i(x^*). \end{cases} \tag{12}$$

If $x_i^* = f_i(x^*)$ we can make any of the choices either in (11) or in (12), since, in this case,

$$\min\{x_i^*, f_i(x^*)\} = x_i^* = f_i(x^*) = 0.$$

So, from (11) and (12):

$$\frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} = \frac{\|\overline{F}(x)\|}{\|x - x^*\|},$$

where

$$\overline{f}_i(x) = \begin{cases} 0 & \text{if } x_i^* \leqslant f_i(x^*), \\ f_i(x) - f_i(x^*) - f_i'(x^*)(x - x^*) & \text{if } x_i^* > f_i(x^*). \end{cases}$$

Now,

$$\frac{\|\overline{F}(x)\|}{\|x - x^*\|} = \max_{1 \leqslant i \leqslant n} \frac{|f_i(x) - f_i(x^*) - f_i'(x^*)(x - x^*)|}{\|x - x^*\|},$$

and so, by the differentiability of F, given $\rho = r/(4\theta)$, there exists $\varepsilon_r > 0$ such that

$$\|x - x^*\| < \varepsilon_r \Longrightarrow \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leqslant \rho.$$

Thus,

$$\left\|\Gamma(x, A) - x^*\right\| \leqslant 2\theta\left[\frac{r}{4\theta} + \frac{r}{4\theta}\right]\|x - x^*\| = r\|x - x^*\|.$$

Therefore, the desired result is proved. □

The following is the so-called theorem of the two neighborhoods. As in Lemma 2.1, $\|\cdot\|$ will mean $\|\cdot\|_\infty$.

Theorem 2.1. *Let all the local assumptions be verified and let* $r \in (0, 1)$. *Then there exist* $\varepsilon, \delta > 0$ *such that, if*

$$\left\|x^0 - x^*\right\| \leqslant \varepsilon \quad \text{and} \quad \left\|A_k - F'(x^*)\right\| \leqslant \delta, \quad \text{for all } k,$$

the sequence $\{x^k\}$ *generated by*

$$x^{k+1} = \Gamma\left(x^k, A^k\right) = x^k - B_k^{-1}G\left(x^k\right),$$

is well defined, converges to x^* *and satisfies*

$$\left\|x^{k+1} - x^*\right\| \leqslant r\left\|x^k - x^*\right\|, \quad \text{for all } k = 0, 1, 2, \ldots.$$

Proof. To prove this we just need to use an inductive argument associated to Lemma 2.1. □

We observe that Theorem 2.1 proves the q-linear convergence of the sequence $\{x^k\}$ in the infinity norm. Therefore, if $e_k = \|x^k - x_*\|$ is the error related to any other norm, then $e_k \leqslant Cr^k e_0$ where C is a positive constant that does not depend on k and r is as in Theorem 2.1. So, r-linear convergence holds in any other norm.

Using similar arguments as in Lemma 2.1, it is easy to prove the following lemma:

Lemma 2.2. *Assume that the assumptions* (A1)–(A3) *are verified. Then there exist* $\varepsilon, \beta > 0$ *such that, if* $\|x - x^*\| < \varepsilon$ *then*

$$\left\|G(x)\right\| \geqslant \beta\|x - x^*\|.$$

Even though we have used the infinity norm to prove Lemma 2.2, the above result remains valid for any other norm, with a suitable change in the constant β.

In the next theorem we prove that, for the reformulation of the NCP by means of $G(x) = 0$, a Dennis–Moré–Walker type condition ensures superlinear convergence. We use the infinity norm in this proof but we recall that superlinear convergence results are norm-independent.

Theorem 2.2. *Assume that the assumptions* (A1)–(A3) *are verified and that for some* x^0 *the sequence generated by*

$$x^{k+1} = x^k - B_k^{-1}G\left(x^k\right), \tag{13}$$

where B_k *is given in* (3), *satisfies*

$$\lim_{k\to\infty} x^k = x^*.$$

Define A_k *as in* (3) *and* $s^k = x^{k+1} - x^k$. *If*

$$\lim_{k\to\infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} = 0,$$

then the sequence $\{x^k\}$ *converges superlinearly to* x^*.

Proof. Let us assume that $x^k \in \mathcal{B}(x^*, \varepsilon_1)$ for all k. By the considerations made in (3),

$$\|(A_k - F'(x^*))s^k\| = \left\| \begin{pmatrix} [A_k]_1 - f_1'(x^*) \\ \vdots \\ [A_k]_n - f_n'(x^*) \end{pmatrix} s^k \right\|. \tag{14}$$

Since B_* is the matrix in $\partial_B G(x^*)$ that has for the ith row $\boldsymbol{e}_i$ or ${f'}_i(x^*)$, we have that

$$\|(B_k - B_*)s^k\| = \left\| \begin{pmatrix} [B_k]_1 - [B_*]_1 \\ \vdots \\ [B_k]_n - [B_*]_n \end{pmatrix} s^k \right\|, \tag{15}$$

where

$$[B_k]_i - [B_*]_i = \begin{cases} 0 & \text{if } x_i^* \leqslant f_i(x^*), \\ [A_k]_i - f_i'(x^*) & \text{if } x_i^* > f_i(x^*), \end{cases}$$

Thus, from (14) and (15),

$$\|(B_k - B_*)s^k\| \leqslant \|(A_k - F'(x^*))s^k\|.$$

Now by (13), we have

$$\begin{aligned} 0 &= B_k s^k + G(x^k), \\ -G(x^{k+1}) &= (B_k - B_*)s^k - G(x^{k+1}) + G(x^k) + B_* s^k. \end{aligned} \tag{16}$$

Observe that to compute $\|-G(x^{k+1}) + G(x^k) + B_* s^k\|$ we work componentwise and so we can use the fact that $F'(x)$ is Lipschitz continuous with constant γ. Using this observation and (16) we get

$$\begin{aligned} \frac{\|G(x^{k+1})\|}{\|s^k\|} &\leqslant \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} + \frac{\|-G(x^{k+1}) + G(x^k) + B_* s^k\|}{\|s^k\|} \\ &\leqslant \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} + \gamma \max\{\|x^{k+1} - x^*\|, \|x^k - x^*\|\}. \end{aligned}$$

Since

$$\lim_{k\to\infty} \frac{\|(A_k - F'(x_*))s^k\|}{\|s^k\|} = 0 \quad \text{and} \quad \lim_{k\to\infty} (x^k - x^*) = 0,$$

we have

$$\lim_{k\to\infty} \frac{\|G(x^{k+1})\|}{\|s^k\|} = 0.$$

But, by Lemma 2.2, there exists a positive constant β such that

$$\lim_{k\to\infty} \frac{\|G(x^{k+1})\|}{\|s^k\|} \geqslant \lim_{k\to\infty} \beta \frac{\|x^{k+1} - x^*\|}{\|s^k\|}.$$

So,

$$0 \geqslant \lim_{k\to\infty} \beta \frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|+\|x^{k+1}-x^*\|} = \lim_{k\to\infty} \beta \frac{\|x^{k+1}-x^*\|/\|x^k-x^*\|}{1+\|x^{k+1}-x^*\|/\|x^k-x^*\|},$$

which implies

$$\lim_{k\to\infty} \frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|} = 0,$$

and thus, x^k converges to x^* superlinearly. □

2.3. LCSU family for $G(x)=0$

The least-change theory, which is well established for smooth and also for some nonsmooth problems [10,13,14], states sufficient conditions under which the hypotheses of the Theorem 2.1 hold.

In this section we define a least-change algorithm for problems like (1) satisfying assumptions (A1)–(A3) and we prove the corresponding local convergence theorems.

$|\cdot|$ will denote an arbitrary norm on $\mathbb{R}^n$ and its associated matrix norm.

Assume that for each pair $x, y \in \mathbb{R}^n$, $x \neq y$, $V(x, y)$ is a linear subspace of $\mathbb{E} = \mathbb{R}^{n\times n}$ and $\|\cdot\|_{xy}$ is the norm on $\mathbb{E}$ related to the scalar product $\langle\, ,\rangle_{xy}$. Moreover, assume that $\|\cdot\|$ is a norm on $\mathbb{E}$ associated to a scalar product $\langle\, ,\rangle$ and let P_{xy} be the orthogonal projection on $V(x, y)$ with respect to the norm $\|\cdot\|_{xy}$.

Algorithm 2.1. Assume that x^0 and A_0 are arbitrary. For $k = 0, 1, 2, \ldots, x^{k+1}$, A_{k+1} are generated as follows:

$$x^{k+1} = x^k - B_k^{-1} G(x^k), \tag{17}$$

$$A_{k+1} = P_{x^k x^{k+1}}(A_k), \tag{18}$$

where B_k is taken following (3) and (4).

In addition to (A1)–(A3), we will assume, as in [12], that:

(A4) There exists $\alpha_1 > 0$ such that, for all $x, y \in \mathbb{R}^n$, there exists a matrix $\overline{A} \in V(x, y)$ satisfying

$$\left\|\overline{A} - F'(x^*)\right\| \leqslant \alpha_1 \sigma(x, y), \tag{19}$$

where $\sigma(x, y) = \max\{|x - x^*|, |y - x^*|\}$.

(A5) There exists $\alpha_2 > 0$ such that, for all $x, y \in \mathbb{R}^n$, $A \in \mathbb{E}$,

$$\|A\|_{xy} \leqslant \left[1 + \alpha_2 \sigma(x, y)\right] \|A\|, \tag{20}$$

$$\|A\| \leqslant \left[1 + \alpha_2 \sigma(x, y)\right] \|A\|_{xy}. \tag{21}$$

In the next lemma we prove a result, known as a Bounded Deterioration Principle, which ensures that the distance between $P_{xy}(A)$ and $F'(x^*)$ cannot be much larger than that between A and $F'(x^*)$.

Lemma 2.3. *Let the assumptions* (A1)–(A5) *be verified. There exist* $\alpha_3, \alpha_4 > 0$ *such that for all* $x, y \in \mathcal{B}(x^*, \varepsilon_1)$, $A \in \mathbb{E}$,

$$\left\|P_{xy}(A) - F'(x^*)\right\| \leqslant \left[1 + \alpha_4 \sigma(x, y)\right] \left\|A - F'(x^*)\right\| + \alpha_3 \sigma(x, y).$$

Proof. The proof is analogous to the proof of [12, Lemma 3.1]. □

Corollary 2.1. *There exists $\alpha_5 > 0$ such that*

$$\|P_{xy}(A) - F'(x^*)\| \leqslant \|A - F'(x^*)\| + \alpha_5|x - x^*|, \tag{22}$$

when $x, y \in \mathcal{B}(x^*, \varepsilon_1)$, $A \in \mathcal{B}(F'(x^*), \delta_1)$ and $|y - x^*| \leqslant |x - x^*|$.

Proof. The proof is analogous to the proof of [12, Corollary 3.1]. □

These two results and the assumptions (A1)–(A3) are of fundamental importance to prove the next linear convergence result for Algorithm 2.1.

Theorem 2.3. *Let assumptions* (A1)–(A5) *be verified and consider the sequence $\{A_k\}$ defined by* (18)*. Given $r \in (0, 1)$ there exist $\overline{\varepsilon}$ and $\overline{\delta}$ such that, if*

$$\|x^0 - x^*\|_\infty \leqslant \overline{\varepsilon} \quad \text{and} \quad \|A_0 - F'(x^*)\| \leqslant \overline{\delta},$$

the sequence x_k generated by

$$x^{k+1} = x^k - B_k^{-1}G(x^k),$$

is well defined, converges to x^* and, for all $k = 0, 1, 2, \ldots$

$$\|x^{k+1} - x^*\|_\infty \leqslant r\|x^k - x^*\|_\infty.$$

Moreover, for all $k, j = 0, 1, 2, \ldots$ there exist positive numbers α_6 and α_7 such that:

$$\|A_{k+j} - F'(x^*)\| \leqslant \|A_k - F'(x^*)\| + \alpha_6|x^k - x^*|, \tag{23}$$

$$\|A_{k+j} - F'(x^*)\|^2 \leqslant \|A_k - F'(x^*)\|^2 + \alpha_7|x^k - x^*|. \tag{24}$$

Proof. It is very similar to that of Theorem 3.3 and Corollary 3.2 in [12]. We observe though that in this proof we use Lemma 2.1 which was proved before using both norms $|\cdot|$ and $\|\cdot\|$ as $\|\cdot\|_\infty$. So, we must be careful in the choice of $\overline{\varepsilon}$ and $\overline{\delta}$ here. For instance, we need $\|A_0 - F'(x^*)\| \leqslant \overline{\delta}$ implying in $\|A_0 - F'(x^*)\|_\infty \leqslant \delta_1$. After making the right choices, the proof follows in a straightforward way. □

Theorem 2.4. *Assume the same hypotheses of Theorem* 2.3*. Then,*

$$\lim_{k\to\infty} \|A_{k+1} - A_k\| = 0. \tag{25}$$

Proof. It follows the lines of that of [12, Theorem 3.2], with the same remarks that we made in Theorem 2.3. □

With this result we can derive a necessary and sufficient condition to have superlinear convergence as shown by the next theorem.

Theorem 2.5. *Assume that the assumptions* (A1)–(A5) *are verified and let the sequences $\{A_k\}$ and $\{x^k\}$ be generated by Algorithm* 2.1*, with B_k generated as in* (3)*. Assume that*

$$\lim_{k\to\infty} x^k = x^*.$$

If

$$\lim_{k\to\infty} \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} = 0, \tag{26}$$

then the sequence $\{x^k\}$ *converges superlinearly to* x^*.

Proof. The proof follows in a straightforward way from Theorems 2.2 and 2.4:

$$\lim_{k\to\infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} \leqslant \lim_{k\to\infty} \left[\frac{\|(A_k - A_{k+1})s^k\|}{\|s^k\|} + \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|}\right]. \tag{27}$$

From Theorem 2.4, we have that the right-hand side expression of the last inequality is equal to zero. So,

$$\lim_{k\to\infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} = 0$$

and by Theorem 2.2 the convergence to x^* is superlinear. □

All the results obtained in this section can be incorporated in the following theorem:

Theorem 2.6. *Assume that the assumptions* (A1)–(A5) *are verified and let the sequences* $\{A_k\}$ *and* $\{x^k\}$ *be generated by Algorithm* 2.1, *with* B_k *generated as in* (3). *Given* $r \in (0,1)$ *there exist* $\overline{\varepsilon}$ *and* $\overline{\delta}$ *such that, if*

$$\|x^0 - x^*\|_\infty \leqslant \overline{\varepsilon} \quad \text{and} \quad \left\|A_0 - F'(x^*)\right\| \leqslant \overline{\delta},$$

the sequence x_k *is well defined and converges linearly to* x^*.

Moreover, if

$$\lim_{k\to\infty} \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} = 0,$$

then the convergence is superlinear.

Thus, we have seen that the family of LCSU methods generates sequences that are locally and superlinearly convergent.

3. The theory for $\Phi(x) = 0$

For the reformulation of the NCP given by (2) it is generated a family of Least-Change Secant-Update (LCSU) methods, as in Section 2.

Given $x^0 \in \mathbb{R}^n$ an initial approximation to x^*, the basic algorithm for this formulation is given by

$$x^{k+1} = x^k - B_k^{-1}\Phi\left(x^k\right),$$

where

$$B_k = \begin{pmatrix} ([B_k]_1) \\ \vdots \\ ([B_k]_n) \end{pmatrix}, \tag{28}$$

with

$$[B_k]_i = \begin{cases} \left(\dfrac{x_i^k}{\|(x_i^k, f_i(x^k))\|_2} - 1\right)\boldsymbol{e}_i + \left(\dfrac{f_i(x^k)}{\|(x_i^k, f_i(x^k))\|_2} - 1\right)[A_k]_i, \\ \qquad\qquad x_i^k \neq 0 \text{ or } f_i(x^k) \neq 0, \\ \left(\dfrac{z_i^k}{\|(z_i^k, \langle [A_k]_i, z^k\rangle)\|_2} - 1\right)\boldsymbol{e}_i + \left(\dfrac{\langle [A_k]_i, z^k\rangle}{\|(z_i^k, \langle [A_k]_i), z^k\rangle)\|_2} - 1\right)[A_k]_i, \\ \qquad\qquad x_i^k = f_i(x^k) = 0. \end{cases} \tag{29}$$

Here $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ is the canonical basis of $\mathbb{R}^n$,

$$A_k = \begin{pmatrix} ([A_k]_1) \\ \vdots \\ ([A_k]_n) \end{pmatrix}$$

is an approximation of the Jacobian matrix of F at x^k and $z^k \in \mathbb{R}^n$ is such that $z_i^k \neq 0$, if $x_i^k = f_i(x^k) = 0$.

As pointed out in Section 1, the function Φ is nondifferentiable at x, if for some i, $1 \leqslant i \leqslant n$, $x_i = f_i(x) = 0$. Facchinei and Kanzow [3] give a procedure to calculate elements of $\partial_B \Phi(x)$ in these cases. They construct a sequence of points where Φ is differentiable and such that the sequence of the Jacobian matrices at these points converges to a matrix belonging to $\partial_B \Phi(x)$. The sequence that they propose is

$$y^k = x + \varepsilon^k z,$$

where $\{\varepsilon^k\}$ is a sequence of positive numbers that converges to zero and z is the vector such that $z_i \neq 0$ if $x_i = 0$ like z^k in (28).

It follows from the results of [3] that defining

$$y^k = x^* + \varepsilon^k z,$$

where x^* is a solution of the NCP, then the matrix $B_*(z)$ belongs to $\partial_B \Phi(x^*)$, where for $i = 1, \ldots, n$,

$$[B_*(z)]_i = \begin{cases} -\boldsymbol{e}_i & \text{if } 0 = x_i^* < f_i(x^*), \\ -f_i'(x^*) & \text{if } x_i^* > f_i(x^*) = 0, \\ (\alpha_i^* - 1)\boldsymbol{e}_i + (\beta_i^* - 1) f_i'(x^*) & \text{if } x_i^* = f_i(x^*) = 0, \end{cases} \tag{30}$$

with

$$\alpha_i^* = \frac{z_i}{\|(z_i, \langle f_i'(x^*), z\rangle)\|_2} \quad \text{and} \quad \beta_i^* = \frac{\langle f_i'(x^*), z\rangle}{\|(z_i, \langle f_i'(x^*), z\rangle)\|_2}.$$

Actually, these matrices $B_*(z)$ form a (generally infinite) compact subset Z_* of $\partial_B \Phi(x^*)$, since there are infinite many ways to choose the vector $z \in \mathbb{R}^n$. This set is given by

$$Z_* = \left\{B_*(z)\colon\ z \in \mathbb{R}^n \text{ is such that } z_i \neq 0, \text{ if } x_i^* = f_i(x^*) = 0\right\}. \tag{31}$$

The algorithm for the LCSU family for $\Phi(x) = 0$ is given by:

Algorithm 3.1. Assume that $x^0 \in \mathbb{R}^n$ and A_0 is arbitrary. For $k = 0, 1, \ldots, n$, let the sequences x^k and $\{A_k\}$ be generated by

$$x^{k+1} = x^k - B_k^{-1}\Phi\bigl(x^k\bigr), \tag{32}$$

$$A_{k+1} = P_{x^k,x^{k+1}}(A_k), \tag{33}$$

and B_k is computed following (27) and (28).

As in Section 2, $|\cdot|$ will denote an arbitrary norm on $\mathbb{R}^n$ and its associated matrix norm and we will assume that for each pair $x, y \in \mathbb{R}^n$, $x \neq y$, $V(x, y)$ is a linear subspace of $\mathbb{E} = \mathbb{R}^{n\times n}$ and $\|\cdot\|_{xy}$ is the norm on $\mathbb{E}$ related to the scalar product $\cdot, \cdot_{xy}$. Moreover, assume that $\|\cdot\|$ is a norm on $\mathbb{E}$ associated to a scalar product $\cdot, \cdot$ and let P_{xy} be the orthogonal projection on $V(x, y)$ with respect to the norm $\|\cdot\|_{xy}$.

The five analogous local assumptions for this formulation are given by:

(H1) $x^* \in \mathbb{R}^n$ is such that $\Phi(x^*) = 0$.

(H2) There exist $\gamma > 0$, $\varepsilon > 0$, such that

$$\bigl\|F'(x) - F'(x^*)\bigr\| \leqslant \gamma|x - x^*|,$$

for all $x \in \mathcal{B}(x^*, \widetilde{\varepsilon})$.

(H3) All the matrices in Z_* are nonsingular.

(H4) There exists $\alpha_1 > 0$ such that for all $x, y \in \mathbb{R}^n$ there exists a matrix $\overline{A} \in V(x, y)$ satisfying

$$\bigl\|\overline{A} - F'(x^*)\bigr\| \leqslant \alpha_1\sigma(x, y), \tag{34}$$

where $\sigma(x, y) = \max\{|x - x^*|, |y - x^*|\}$.

(H5) There exists $\alpha_2 > 0$ such that for all $x, y \in \mathbb{R}^n$, $A \in \mathbb{E}$,

$$\|A\|_{xy} \leqslant \bigl[1 + \alpha_2\sigma(x, y)\bigr]\,\|A\|, \tag{35}$$

$$\|A\| \leqslant \bigl[1 + \alpha_2\sigma(x, y)\bigr]\,\|A\|_{xy} \tag{36}$$

Under these local assumptions, the same convergence results as those obtained in Section 2 are proved. The details of the proofs are given in [17].

4. The nonsingularity assumption

One of the assumptions under which we developed the LCSU theory is the BD-regularity of G at x^*, that is, the assumption that all the elements in $\partial_B G(x^*)$ are nonsingular, and the nonsingularity of the matrices in $Z_* \subset \partial_B\Phi(x^*)$. Since any $B \in \partial_B G(x^*)$ can be written as $-\overline{B} \in Z_*$ if we take $\alpha = 1$, $\beta = 0$; then we have $\partial_B G(x^*) \subset Z_*$ and so, if all the elements in Z_* are nonsingular, then G is BD-regular at x_*. But it is not true that if all the elements in $\partial_B G(x^*)$ are nonsingular, then all the elements in $\partial_B\Phi(x^*)$ are nonsingular. This is shown in this very simple example:

Example 4.1. Define

$$F : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto F(x) = (x_1 + 3x_2 - 1, x_1 + x_2 - 1).$$

$x^* = (1, 0)$ is a degenerate solution of the NCP, since

$$x_1^* > f_1(x^*) = 0, \qquad x_2^* = f_2(x^*) = 0.$$

Thus,

$$\partial_B G(x^*) = \left\{ \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \right\},$$

with both elements nonsingular. As was seen in Section 3, in this case the elements of $Z_* \subset \partial_B \Phi(x^*)$ are given by

$$\overline{B} = \begin{pmatrix} -1 & -3 \\ (\alpha - 1)\boldsymbol{e}_2 + & (\beta - 1)(1,1) \end{pmatrix},$$

with $\alpha^2 + \beta^2 = 1$, where

$$\alpha = \frac{z_2}{\|(z_2, \langle f_2'(x^*), z\rangle)\|_2} = \frac{z_2}{\|(z_2, z_1 + z_2)\|_2},$$

$$\beta = \frac{\langle f_2'(x^*), z\rangle}{\|(z_2, \langle f_2'(x^*), z\rangle)\|_2} = \frac{z_1 + z_2}{\|(z_2, z_1 + z_2)\|_2},$$

and $z \in \mathbb{R}^2$ is such that $z_2 \neq 0$, since $x_2^* = f_2(x^*) = 0$.

For $z = (1,3)$, $\|(z_2, z_1 + z_2)\|_2 = 5$ and

$$\overline{B} = \begin{pmatrix} -1 & -3 \\ (\frac{3}{5} - 1)\boldsymbol{e}_2 + & (\frac{4}{5} - 1)(1,1) \end{pmatrix} = \begin{pmatrix} -1 & -3 \\ -\frac{1}{5} & -\frac{3}{5} \end{pmatrix},$$

which is a singular matrix. So, there exists at least one singular element in $Z_* \subset \partial_B \Phi(x^*)$.

Theorem 4.1 gives necessary and sufficient conditions for the case $n = 2$ to have singular elements in $Z_* \subset \partial_B \Phi(x^*)$ under BD-regularity of G at x^*.

We recall that, for $n = 2$, the set $\partial_B G(x^*)$ is given, for $p \neq 0$, by

$$\partial_B G(x^*) = \left\{ \begin{pmatrix} \frac{\partial f_1(x^*)}{\partial x_1} & \frac{\partial f_1(x^*)}{\partial x_2} \\ 0 & p \end{pmatrix}, \begin{pmatrix} \frac{\partial f_1(x^*)}{\partial x_1} & \frac{\partial f_1(x^*)}{\partial x_2} \\ \frac{\partial f_2(x^*)}{\partial x_1} & \frac{\partial f_2(x^*)}{\partial x_2} \end{pmatrix} \right\}. \tag{37}$$

Theorem 4.1. *Let $n = 2$, $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F \in C^1(\mathbb{R}^2)$, G and Φ be defined as in* (1) *and* (2)*. If G is BD-regular at x^*, then there will be singular elements in $Z_* \subset \partial_B \Phi(x^*)$ if and only if*
- $\partial f_1(x^*)/\partial x_1$ *and determinant of the second matrix in* (37) *have opposite signs,*
- $p < 0$ *in* (37).

Proof. See [17]. □

For the generic case, i.e., for $F : \mathbb{R}^n \to \mathbb{R}^n$, $n > 2$, $F \in C^1$, let x^* be a degenerate solution of the NCP. The analysis of this case gives us similar conditions to those of the case $n = 2$.

Based on these results we can expect a better local numerical performance of the Min function reformulation of problem (1). In the next section we discuss this fact with more detail.

5. Some numerical experiments

In this section we analyze the sensitivity of the functions G and Φ to degenerate solutions of the NCP. We do this by taking the functions defined in [7], i.e.,

$$F:\mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto F(x) = \big(f_1(x), \ldots, f_n(x)\big),$$

where

$$f_i(x) = \begin{cases} h_i(x) - h_i(x^*) & \text{if } i \text{ is odd or } i > n/2, \\ h_i(x) - h_i(x^*) + 1 & \text{otherwise}. \end{cases}$$

For all these functions the vector $x^* = (1, 0, 1, 0, \ldots) \in \mathbb{R}^n$ is a solution. For $i = 1, \ldots, n$, h_i are the functions given by Lukšan [11]. In these cases F is nonsmooth at x^*, since, if i is even and $i > n/2$ we have

$$f_i(x^*) = x_i^* = 0.$$

Thus, x^* is a degenerate solution of the NCP and it is also a solution of the nonlinear systems $G(x) = 0$ and $\Phi(x) = 0$. To compare the sensitivity we worked with the problems of [11], fixing a value for n. For each problem we calculated the maximum condition number of the matrices in $\partial_B G(x^*)$ for the Min function and we maximized the condition number of the elements of Z_* as a function of z, for the Fischer function.

To run the test problems we used MATLAB and worked on a Sun Sparc station 2. The first column of Table 1 shows which problem has been tested and the second one, shows the n we fixed.

With the results given in the Table 1 we conclude that the Fischer function is much more sensitive to degeneracies at a solution x^* than the Min function and this will affect the local convergence of the method that uses the Fischer reformulation of problem (1). In other words, in degenerated problems, if for both reformulations convergence takes place, the convergence of the Fischer reformulation is expected to be slower than that of the Min reformulation.

In what follows we analyze the local behavior of the algorithms proposed in Sections 2 and 3. All the tests were done using MATLAB. We used the 17 test problems proposed in [7] for both cases and we tested the generalized Newton method and the generalized Schubert method for all of them.

We recall that in the Schubert method the matrices are updated in the following way:

For $y_k = F(x^{k+1}) - F(x^k)$, let $v_k = y_k - A_k s_k$ and $w_k = \widehat{s}_k$, where $\widehat{s}_k$ means the vector derived from s_k by setting s_j^k to zero whenever the corresponding element of $[A_k]_i$ is a known constant. Then

$$A_{k+1} = A_k + \frac{C_k}{\langle \widehat{s}_k, \widehat{s}_k \rangle},$$

where C_k is the matrix with elements $c_{ij} = v_i w_j$.

The stopping criteria used were:

- $\|G(x^k)\|_2 < \sqrt{n}10^{-5}$ for Algorithm 2.1,
- $\|\Phi(x^k)\|_2 < \sqrt{n}10^{-5}$ for Algorithm 3.1,
- $k > 100$,
- $\|G(x^k)\|_\infty > 10^{20}$ for Algorithm 2.1,
- $\|\Phi(x^k)\|_\infty > 10^{20}$ for Algorithm 3.1.

The results for these numerical experiments are shown in Table 2. In all the problems the initial approximation vector was the vector $(0.9, 0.1, \ldots,)$. Prob means the number of the problem from [11] that was tested, Dim is the dimension of the problem and each one of the other columns tells what happened in terms of convergence: a number means how many iterations were performed to converge to the solution that we were looking for; a – sign means divergence and k^* means that in k iterations the

Table 1
The maximum condition number of the matrices in $\partial_B G(x^*)$ for the Min function and a maximization of the condition number for the Fischer function

Prob	Dim	Min	Fischer
1	6	28.37	∞
2	6	38.45	∞
3	10	5.50	∞
4	6	15.63	22.34
5	7	∞	∞
6	6	∞	∞
7	6	33.00	∞
8	6	35.01	∞
9	6	34.18	1.83×10^{18}
10	6	35.37	48.55
11	6	50.08	∞
12	8	256.20	∞
13	8	76.90	253.86
14	6	38.47	2.19×10^{19}
15	6	25.58	31.89
16	6	7.23	13.23
17	6	10.93	2.98×10^{18}

process converged to another solution. Since problem 6 from [11] does not satisfy the assumption (3) of the theories developed in Sections 2 and 3, we did not consider it.

The results in Table 2 show that, for this set of experiments, the local behavior of the method that uses the Min function is slightly better than that of the the method that uses the Fischer function. This was observed also in [9] from numerical experiments and the authors use this observation to introduce a globalizing strategy.

Table 2
The performance of the generalized Newton's and Schubert's methods

Prob	Dim	Newton Min	Newton Fischer	Schubert Min	Schubert Fischer
1	100	3	4	3	6
2	100	4	5	3	6
3	10	–	–	–	–
4	100	3	4	4	4
5	101	–	–	–	–
7	100	4	4	4	6
8	100	3	4	6	–
9	100	5	4	6	6
10	100	4	4	7	7
11	100	1*	–	1*	–
12	100	1*	6*	–	–
13	100	–	–	–	–
14	100	–	–	–	–
15	100	15	17	–	–
16	100	2	4	2	4
17	100	5	5	6	6

6. A globalizing strategy

In this section we present a global algorithm to solve the NCP. This is an hybrid algorithm like the one proposed in [7] that combines the good local behavior of the Min function with the global behavior of the Fischer function.

We start the iterations with the local method which uses the Min function and continue with it while the value of $\|G(x)\|$ is decreasing. If it does not decrease we use the global minimization algorithm proposed in [2,9] for $\lambda = 2$, to solve the NCP.

We also present some numerical results of our algorithm and compare them with the results that we obtained using the algorithm proposed in [2].

We will call the local iteration $x^{k+1} = x^k - {B_k}^{-1} G(x^k)$ an *ordinary iteration*, and an iteration generated by the global minimization algorithm, will be called a *special iteration*.

Ordinary and special iterations are combined following [7] in this hybrid algorithm.

For each $k \in \mathbb{N}$, let

$$w^k = \operatorname{argmin}\{\|G(x^0)\|, \ldots, \|G(x^k)\|\}.$$

For the sake of completeness we define $\|G(w^j)\| = \|G(x^0)\|$ if $k < j$.

Algorithm 6.1. Initialize $k \leftarrow 0$, $FLAG \leftarrow 1$. Let $q \geqslant 0$ be an integer, $\gamma \in (0,1)$ and the initial approximation x^0 be given.

Step 0. $k \leftarrow 0$, $FLAG \leftarrow 1$.

Step 1. If $FLAG = 1$, obtain x^{k+1} using an ordinary iteration.
Otherwise, obtain x^{k+1} using a special iteration.

Step 2. If $\|G(x^{k+1})\| \leqslant \gamma \|G(w^{k-q})\|$, set $FLAG \leftarrow 1$, $k \leftarrow k+1$ and go to Step 1.
Otherwise, re-define $x^{k+1} \leftarrow w^{k+1}$, $FLAG \leftarrow -1$, $k \leftarrow k+1$ and go to Step 1.

6.1. Numerical performance

We tested Algorithm 6.1 with the problems suggested in [11] with the same initial approximations. The parameters used were: $\gamma = 0.9$, $q = 5$, and, for the special iterations, $\rho = 10^{-8}$, $\beta = 0.5$, $\sigma = 10^{-4}$, $p = 2.1$ and $t_{\min} = 10^{-12}$.

These are the stopping criteria used:

- $\|G(x^k)\|_2 < \sqrt{n}10^{-5}$,
- $k > 100$, and
- $t_k < t_{\min}$ in the special iterations.

Table 3 presents the results when we applied to the problems in [11], Algorithm 6.1 (Min-Fischer) and the Global Algorithm from [2] that uses only the Fischer function (Fischer). Prob means the number of the problem from [11] that was tested, Dim is the dimension that we used for it. The columns Min-Fischer and Fischer contain the total number of iterations performed. A – sign means divergence and k^* means that in k iterations the process converged to another solution. Since problem 6 from [11] does not satisfy the assumption (3) of the theories developed in Sections 2 and 3, we did not consider it.

We observe in Table 3 that, in most cases of convergence of both algorithms, Algorithm 6.1 takes less iterations than the other one, and we notice that, for problem 14 our algorithm attained convergence in 12 iterations while the other failed. In fact these experiments show that the globalizing strategy that uses the hybrid algorithm is more effective.

7. Final remarks

The technique of reducing nonlinear complementarity problems to nonlinear systems of equations is very important for solving this type of problems because, in this way, the main work of most iterations is the resolution of a single linear system. In the Newtonian approach, the matrix of this system is a Jacobian

Table 3
Comparison between the Global Algorithm 6.1 and a Global Algorithm that uses only the Fischer function

Prob	Dim	Min-Fischer	Fischer
1	100	4 (4, 0)	6
2	100	6 (1, 5)	29
3	100	–	–
4	100	10* (7, 3)	13*
5	101	–	–
7	100	–	–
8	100	–	–
9	100	–	–
10	100	6 (6, 0)	7
11	100	1 (1, 0)	4
12	100	23 (9, 14)	13
13	100	11 (6, 5)	11
14	100	12 (7, 5)	–
15	100	–	–
16	100	4* (4, 0)	6*
17	100	7 (7, 0)	7

and the exact solution is required, while in the inexact-Newton framework, only an approximate solution is necessary. In this paper, we considered the quasi-Newton approach, that can be very useful when the derivatives of the system are very expensive or difficult to obtain.

The fact that, ultimately, an iteration consists on the resolution of a linear system together with only one functional evaluation is associated to the possibility of obtaining high convergence rates (generally, superlinear convergence) of pure local methods. Globalization procedures are usually devised in such a way that global iterations coincide with local iterations in a neighborhood of the solution, so that fast local convergence is maintained.

However, fast local convergence usually depends on characteristics of the problem, the main of which is the nonsingularity of the (generalized) Jacobians of the nonlinear system at the solution considered. De Luca et al. [2] showed that, when one uses the nonlinear system induced by the componentwise application of the Fischer function, the nonsingularity of the generalized Jacobians is not directly associated to degeneracy of the solution. In other words, nonsingular generalized Jacobians can be encountered at degenerate solutions, so that fast convergence can be expected even in these cases. It is easy to show that the same result holds for the classical Min function considered in this paper and for the generalizations of the Fischer function introduced in [9]. The sum of squares of the Fischer-related system is smooth, therefore globally convergent methods related to its minimization can be developed. Unhappily, this is not the case of the Min function.

On the other hand, it has been shown in this paper that singular generalized Jacobians can appear at degenerate solutions of the nonlinear complementarity problem when one uses the Fischer system, as a result of the algebraic form of this two-variable function. To understand geometrically why this happens, consider the two-dimensional complementarity problem defined by $f_1(x_1, x_2) = x_1 + 2x_2 - 1$ and $f_2(x_1, x_2) = x_1 + x_2 - 1$, which has the nondegenerate solution $(0, 1)$ and the degenerate solution $(1, 0)$. In a neighborhood of $(1, 0)$ the Min system is formed by the line $x_1 + 2x_2 - 1 = 0$ and the piecewise linear "curve" $\min\{x_2, x_1 + x_2 - 1\} = 0$. Therefore, the level sets of the two Fischer functions involved are smooth approximations of the level sets of $-(x_1 + 2x_2 - 1)$ and $-\min\{x_2, x_1 + x_2 - 1\}$. It is easy to see, geometrically, that the set of points at which the level sets of the first Fischer function are tangent to the level sets of the second one form a continuous curve that emanates from $(1, 0)$. Obviously, the Jacobian of the Fischer system is singular at all the points of this curve and, so, there is a singular generalized Jacobian at $(1, 0)$. Clearly, this phenomenon does not occur in the case of the Min function. (In the Min system the set of generalized Jacobians is formed by two nonsingular matrices.)

The observations above seem to indicate that the development of Newton-like local theories for the Min system can be a useful tool to understand the behavior of practical methods. In this paper we developed the Least-Change Secant-Update theory for quasi-Newton methods based on secant-like projections. (Newton and inexact-Newton theories are mere applications of existing theories for semismooth systems.)

Finally, the association between locally convergent methods and globally convergent ones should be considered. Here we suggested to combine local strategies based on the Min function with global strategies based on the Fischer function like the ones developed by [7]. Preliminary computational results seem to show that this combination is worthwhile. However, much research is necessary along these lines both from the theoretical and the practical point of view.

References

[1] F.H. Clarke, *Optimization and Nonsmooth Analysis* (Wiley, New York, 1983). Reprinted by SIAM, Philadelphia, PA, 1990.

[2] T. De Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, *Math. Programming* 75 (1996) 407–439.

[3] F. Facchinei and C. Kanzow, A nonsmooth inexact Newton method for the solution of large scale nonlinear complementarity problems, *Math. Programming* 76 (1997) 493–512.

[4] M.C. Ferris and J.-S. Pang, Engineering and economic applications of complementarity problems, Technical Report 95-07, Computer Science Department, University of Wisconsin, Madison, WI (1995).

[5] A. Fischer, An NCP-function and its use for the solution of complementarity problems, in: D.-Z. Du, L. Qi and R. Wormersley, eds., *Recent Advances in Nonsmooth Optimization* (World Scientific, Singapore, 1995) 88–105.

[6] T.L. Friesz, R.L. Tobin, T.E. Smith and P.T. Harker, A nonlinear complementarity formulation and solution procedure for the general derived demand network equilibrium problem, *J. Reg. Sci.* 23 (1983) 337–359.

[7] M.A. Gomes-Ruggiero, J.M. Martínez and S.A. Santos, Solving nonsmooth equations by means of quasi-Newton methods with globalization, in: D.-Z. Du, L. Qi and R. Wormersley, eds., *Recent Advances in Nonsmooth Optimization* (World Scientific, Singapore, 1995) 121–140.

[8] P.T. Harker and J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, *Math. Programming* 48 (1990) 161–220.

[9] C. Kanzow and H.A. Kleimichel, A new class of semismooth Newton-type methods for nonlinear complementarity problems, Reihe A, Preprint 118, Hamburger Beiträge zur Angewandten Mathematik (1997).

[10] V.L.R. Lopes and J.M. Martínez, On the convergence of quasi-Newton methods for nonsmooth problems, *Numer. Funct. Anal. Optim.* 16 (1995) 1193–1209.

[11] L. Lukšan, Inexact trust region method for large sparse systems of nonlinear equations, *J. Optim. Theory Appl.* 81 (1994) 569–590.

[12] J.M. Martínez, Local convergence theory of inexact Newton methods based on structured least change updates, *Math. Comput.* 55 (1990) 143–168.

[13] J.M. Martínez, On the relation between two local convergence theories of least-change secant update methods, *Math. Comput.* 59 (1992) 457–481.

[14] J.M. Martínez and M.C. Zambaldi, Least-change update methods for nonlinear systems with nondifferentiable terms, *Numer. Funct. Anal. Optim.* 14 (1993) 405–415.

[15] K.P. Oh, The formulation of the mixed lubrication problem as a generalized nonlinear complementarity problem, *Trans. ASME* 108 (1986) 598–604.

[16] J.-S. Pang and L. Qi, Nonsmooth equations: motivation and algorithms, *SIAM J. Optim.* 3 (3) (1993) 443–465.

[17] R. Pérez, Algoritmos para complementaridade não linear e problemas relacionados, Ph.D. Thesis, IMECC, UNICAMP, Campinas, Brazil (1997).

[18] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.* 18 (1) (1993) 227–244.

[19] D. Sun and J. Han, Newton and quasi-Newton methods for a class of nonsmooth equations and related problems, *SIAM J. Optim.* 7 (2) (1997) 463–480.