Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 13, No. 1, pp. 1 – 15, January 2017

A NONSMOOTH GLOBAL QUASI-NEWTON METHOD FOR NONLINEAR COMPLEMENTARITY PROBLEMS*

Carlos Andrés Arias, Héctor Jairo Martínez and Rosana Pérez

Abstract: In this paper, we present a new quasi Newton algorithm for the solution of nonlinear complementarity problems. The algorithm is based on a nonsmooth equations reformulation of the complementarity problem and the fact that the natural merit function associated to the reformulation is continuously differentiable. We proved global and local convergence results for the algorithm. Some numerical experiments show a good performance of this algorithm.

Key words: *nonsmooth systems, nonlinear complementarity problems, merit function, quasi-Newton methods, global convergence, superlinear convergence*

Mathematics Subject Classification: *90C30, 90C33*

1 Introduction

Let $F\colon \mathbb{R}^n \to \mathbb{R}^n$, $F(\boldsymbol{x}) = (F_1(\boldsymbol{x}), \ldots, F_n(\boldsymbol{x}))$ be a continuously differentiable mapping. The Nonlinear Complementarity Problem, NCP for short, consists of finding a vector $\boldsymbol{x} \in \mathbb{R}^n$ such that,

$$\boldsymbol{x} \geq 0, \quad F(\boldsymbol{x}) \geq 0, \quad \boldsymbol{x}^T F(\boldsymbol{x}) = 0. \tag{1.1}$$

Here, the inequality $\boldsymbol{y} \geq 0$ for $\boldsymbol{y} \in \mathbb{R}^n$ means $y_i \geq 0$ for all $i = 1, \ldots, n$. The third condition in (1.1) requires that the vectors $\boldsymbol{x}$ and $F(\boldsymbol{x})$ are ortogonal; for this reason, it is called *complementarity condition*.

The NCP arises in many applications such as Friction Mechanical Contact problems [9], Structural Mechanics Design problems, Lubrication Elasto-hydrodynamic problems [12], Traffic Equilibrium problems [17], as well as problems related to Economic Equilibrium Models [9]. The importance of NCP in the areas of Physics, Engineering and Economics is due to the fact that the concept of complementarity is synonymous with the notion of system in equilibrium. In recent years, various techniques have been studied to solve the NCP, one of which is to reformulate it as a nonsmooth system of nonlinear equations by using special functions called *complementarity functions* [18]. That is, a function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\varphi(a, b) = 0 \Longleftrightarrow a \geq 0, \quad b \geq 0, \quad ab = 0. \tag{1.2}$$

*The authors would like to thank the Universidad del Cauca for providing time for this work through research project VRI ID 3908.

In 1998, *Kanzow* and *Kleinmichel* [11] introduced an one-parametric class of complementarity functions φ_λ defined by

$$\varphi_\lambda(a,b) = \sqrt{(a-b)^2 + \lambda ab} - a - b, \tag{1.3}$$

where $\lambda \in (0,4)$. We will refer to (1.3) as the *Kanzow function*. Note that, this function is nonsmooth at $(0,0)$. Moreover, in the special case $\lambda = 2$, φ_λ reduces to the *Fischer function* [10], whereas in the limiting case $\lambda \to 0$, the function φ_λ becomes a multiple of the *minimum function* [16] [15].

Now, if we define the function $\Phi_\lambda \colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by

$$\Phi_\lambda(\boldsymbol{x}) = \begin{pmatrix} \varphi_\lambda(x_1, F_1(\boldsymbol{x})) \\ \vdots \\ \varphi_\lambda(x_n, F_n(\boldsymbol{x})) \end{pmatrix} \tag{1.4}$$

then it follows immediately from the definition of an NCP-function that $\boldsymbol{x}^*$ solves NCP if and only if $\boldsymbol{x}^*$ solves $\Phi_\lambda(\boldsymbol{x}) = 0$ [11]. Alternatively, if $\Psi_\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}$ denotes the corresponding merit function

$$\Psi_\lambda(\boldsymbol{x}) = \frac{1}{2}\Phi_\lambda(\boldsymbol{x})^T\Phi_\lambda(\boldsymbol{x}) = \frac{1}{2}\|\Phi_\lambda(\boldsymbol{x})\|_2^2, \tag{1.5}$$

then we may rewrite the NCP as the unconstrained minimization problem

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{Minimize} \quad \Psi_\lambda(\boldsymbol{x}). \tag{1.6}$$

This *merit function* is continuously differentiable [11].

In [11], the authors present a global semismooth Newton-type method for solving the minimization problem (6) and recently, in [1], the author proposes a *local quasi-Newton* method for solving $\Phi_\lambda(\boldsymbol{x}) = 0$ therefore, solving the NCP. Moreover, this author leaves open the possibility of introducing globalization strategies to improve the performance of his algorithm.

Considering the limitations that may have a local algorithm, in this paper, we propose a global quasi-Newton method for the NCP solving alternatively the unconstrained minimization problem (1.5). For this, we propose a globalization of the algorithm proposed in [1] and we established global and local convergence properties. Moreover, we include numerical test, that show a good performance of the algorithm proposed.

We organized this paper as follows. In Section 2, we present some theoretical results that will be useful in the development of the convergence results of our algorithm. In Section 3, we introduce a global nonsmooth quasi-Newton method for solving nonlinear complementarity problems through the minimization of the *merit function* (1.5) and we prove, under suitable assumptions, global and local convergence results for this method. In Section 4, we analyze numerically, the local performance of the algorithm introduced in the last section, using some *test problems* proposed in [15] and we compare these results with those obtained by solving the same problems with the algorithm type Newton proposed in [11]. Finally, Section 5 contains some remarks on what we have done in this paper and present possibilities for future works.

2 Preliminaries

In this section, we present some definitions and theoretical results that have been previously demonstrated by some authors which will be useful in the development of our proposal.

From (1.1), a solution $\boldsymbol{x}^*$ of NCP is characterized by $\boldsymbol{x}^* \geq 0$, $F(\boldsymbol{x}^*) \geq 0$ and $F(\boldsymbol{x}^*)^T \boldsymbol{x}^* = 0$. Associated with $\boldsymbol{x}^*$, we have the sets of indices $\alpha = \{i \in I\colon x_i^* > 0,\ F_i(\boldsymbol{x}^*) = 0\}$, $\beta = \{i \in I\colon x_i^* = 0 = F_i(\boldsymbol{x}^*)\}$ and $\gamma = \{i \in I\colon , x_i^* = 0,\ F_i(\boldsymbol{x}^*) > 0\}$.

Since the function Φ_λ defined by (1.4) is nondifferentiable and *Lipschitz continuous*, its *generalized jacobian* exists and it is defined by

$$\partial\Phi_\lambda(\boldsymbol{x}) \;=\; conv\left\{ \lim_{k\to\infty} \Phi'_\lambda(\boldsymbol{x}_k) \in \mathbb{R}^{n\times n} : \lim_{k\to\infty} \boldsymbol{x}_k \;=\; \boldsymbol{x},\ \boldsymbol{x}_k \in D_{\Phi_\lambda} \right\} \;=\; conv\{\partial_B \Phi_\lambda(\boldsymbol{x})\} \tag{2.1}$$

where D_{Φ_λ} is the set of all points where Φ_λ is differentiable and *conv* denotes the convex hull of a set.

In [1], the author builds matrices $H \in \partial\Phi_\lambda(\boldsymbol{x})$ whose i th row is given by

$$[H]_i = \begin{cases} (\chi(x_i, F_i(\boldsymbol{x})) - 1)\, \mathbf{e}_i^T + (\psi(x_i, F_i(\boldsymbol{x})) - 1)\, \nabla F_i(\boldsymbol{x})^T, & i \notin \beta \\ \big(\chi(z_i, \nabla F_i(\boldsymbol{x})^T \boldsymbol{z}) - 1\big)\, \mathbf{e}_i^T + \big(\psi(z_i, \nabla F_i(\boldsymbol{x})^T \boldsymbol{z}) - 1\big)\, \nabla F_i(\boldsymbol{x})^T, & i \in \beta. \end{cases} \tag{2.2}$$

where $\boldsymbol{z}$ is a vector such that $z_i \neq 0$ when $i \in \beta$, $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ is a canonical basis of $\mathbb{R}^n$ and the functions χ and ψ are the the partial derivatives of the function

$$G_\lambda(a, b) = \sqrt{(a-b)^2 + \lambda a b}\,, \tag{2.3}$$

that is,

$$\chi(a, b) = \frac{2(a-b) + \lambda b}{2G_\lambda(a, b)} \qquad \textit{and} \qquad \psi(a, b) = \frac{-2(a-b) + \lambda a}{2G_\lambda(a, b)}\,. \tag{2.4}$$

For any nonzero vector $(a,\, b)$, the partial derivatives (2.4) are bounded as follows [1] [11]:

$$|\chi(a, b)| \leq 1 \qquad \textit{and} \qquad |\psi(a, b)| \leq \sqrt{2}. \tag{2.5}$$

Also, using the structure of $H \in \partial\Phi_\lambda(\boldsymbol{x})$ given by (2.2), the author in [1] proposed *quasi-Newton approximations* B to these matrices given by

$$[B]_i = \begin{cases} (\chi(x_i, F_i(\boldsymbol{x})) - 1)\boldsymbol{e}_i^T + (\psi(x_i, F_i(\boldsymbol{x})) - 1)[A]_i, & i \notin \beta \\ (\chi(z_i, [A]_i \boldsymbol{z}) - 1)\boldsymbol{e}_i^T \;+\; (\psi(z_i, [A]_i \boldsymbol{z}) \;- 1)[A]_i, & i \in \beta \end{cases} \tag{2.6}$$

where $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a canonical basis of $\mathbb{R}^n$, and the matrix A is an approximation of the Jacobian matrix of F at $\boldsymbol{x}$.

With the above information, in [1], the author proposes a *local quasi-Newton* algorithm for solving NCP based on its reformulation as the nonsmooth system of equations. The generic form of this algorithm is as follows.

Algorithm 2.1 (quasi-Newton for $\Phi_\lambda(\mathbf{x}) = 0$)**.** Given $\mathbf{x}_0$ and $\lambda \in (0,\, 4)$, for $k = 1,\, 2, \ldots,\ x_{k+1}$ compute

$$x_{k+1} = x_k - B_k^{-1} \Phi_\lambda(x_k)$$

where B_k is given by (2.6).

Moreover, in [1], it is shown that the distance between a matrix H and its approximation B is bounded by a constant value, and under certain conditions, the matrix B given by (2.6) is nonsingular. We recall this results.

Theorem 2.2 ([1]). *Let* $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $F \in C^1$, *such that its Jacobian matrix is Lipschitz continuous;* $\boldsymbol{x}^*$ *a solution of* NCP*;* H *and* B *defined by* (2.2) *and* (2.6)*, respectively, and given positive constants* ϵ *and* δ*. Then, for each* $\boldsymbol{x} \in \mathcal{B}(\boldsymbol{x}^*, \epsilon)$ *and* $A \in \mathcal{B}(F'(\boldsymbol{x}^*), \delta)$*, there exists a positive constant* θ *such that*

$$\|H - B\|_\infty \leq \theta.$$

Theorem 2.3 ([1]). *Let* $\boldsymbol{x}^*$ *be a solution of* NCP *and* B *defined by* (2.6)*. There exists positive constants* ϵ_0 *and* δ_0*, such that if* $\|\boldsymbol{x} - \boldsymbol{x}^*\|_\infty \leq \epsilon_0$ *and* $\|A - F'(\boldsymbol{x}^*)\|_\infty \leq \delta_0$ *then the function* $\mathcal{Q}$ *defined by*

$$\mathcal{Q}(\boldsymbol{x}, A) = \boldsymbol{x} - B^{-1}\Phi_\lambda(\boldsymbol{x})$$

is well defined.

In particular, this theorem guarantees that, if a sequence $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}^*$, then there exists a positive integer $\overline{k}$ such that, for all $k > \overline{k}$, the matrix B_k^{-1} exists.

For the development of convergence results for algorithms that solve the NCP, it is important to establish sufficient conditions for the nonsingularity of matrices in the generalized Jacobian at a solution of the problem. For the this, we recalling the concept of *BD regularity.*

Definition 2.4. Let $\mathbf{x}^*$ be a solution of *NCP.*

1. If all matrices $H \in \partial_B \Phi_\lambda(\mathbf{x}^*)$ are nonsingular, then $\mathbf{x}^*$ is a BD regular solution.
2. If the submatrix $F'(\mathbf{x}^*)_{\alpha\alpha}$ is nonsingular and the Schur-complement

$$F'(\mathbf{x}^*)_{\beta\beta} - F'(\mathbf{x}^*)_{\beta\alpha}F'(\mathbf{x}^*)^{-1}_{\alpha\alpha}F'(\mathbf{x}^*)_{\alpha\beta}$$

is a P-matriz, $\mathbf{x}^*$ is called a solution R-regular.

The following theorem gives a sufficient condition to ensure nonsingularity of the matrices of *generalized Jacobian* $\partial\Phi_\lambda(\boldsymbol{x}^*)$ [11].

Theorem 2.5 ([11]). *Assume that* $\boldsymbol{x}^* \in \mathbb{R}^n$ *is a R-regular solution of* NCP*. Then all elements in the generalized Jacobian* $\partial\Phi_\lambda(\boldsymbol{x}^*)$ *are nonsingular.*

An immediate consequence of this result is given in the following corollary.

Corollary 2.6. *If* $\boldsymbol{x}^*$ *is a solution R-regular of* NCP *then* $\boldsymbol{x}^*$ *is a BD-regular solution.*

Proof. Let $\boldsymbol{x}^*$ be a solution *R-regular* of NCP. by hypothesis and Theorem 2.5, all elements in the generalized Jacobian $\partial\Phi_\lambda(\boldsymbol{x}^*)$ are nonsingular. Moreover, from (2.1)

$$\partial\Phi_\lambda(\boldsymbol{x}^*) = \mathit{conv}\{\partial_B\Phi_\lambda(\boldsymbol{x}^*)\},$$

and $\partial_B\Phi_\lambda(\boldsymbol{x}^*) \subseteq \partial\Phi_\lambda(\boldsymbol{x}^*)$. Thus, if $H \in \partial_B\Phi_\lambda(\boldsymbol{x}^*)$, then $H \in \partial\Phi_\lambda(\boldsymbol{x}^*)$ and therefore, is nonsingular. Accordingly, $\boldsymbol{x}^*$ is a BD-regular solution of NCP. □

The following two definitions are useful for the presentation of Theorem 2.9 in [8].

Given a matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ and index sets η y τ, the $A_{\eta\tau}$ is a matrix where a_{ij} with $i \in \eta$ and $j \in \tau$.

A matrix $M \in \mathbb{R}^{n\times n}$ is a *P-matrix,* if for any vector nonzero $\boldsymbol{y} \in \mathbb{R}^n$, exists a index $i_0 = i_0(\boldsymbol{y}) \in \{1, 2, \ldots, n\}$, such that $y_{i_0}[M\boldsymbol{y}]_{i_0} > 0$.

Definition 2.7 ([11]). Let $G\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be locally Lipschitzian and directionally differentiable at $\mathbf{x} \in \mathbb{R}^n$. We say that , G is semismooth at $\mathbf{x}$, if for any $H \in \partial G(\mathbf{x}+\mathbf{d})$ and for any $\mathbf{d} \to \mathbf{0}$,

$$H\mathbf{d} - G'(\mathbf{x}\,;\mathbf{d}) = o(\|\mathbf{d}\|).$$

Definition 2.8 ([7]). A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is an SC^1 function if f is continuously differentiable and its gradient is semismooth.

For the statement of the following theorem, we suppose that $\boldsymbol{x}^*$ is an accumulation point of the $\{\boldsymbol{x}_k\}$.

Theorem 2.9 ([8]). *Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ an SC^1 function and $\delta > 0$. If $\overline{Df}(\boldsymbol{x};\,\boldsymbol{d})$ is an approximation to the directional derivative of f at $\boldsymbol{x}$ in the direction $\boldsymbol{d}$, such that, for all $\boldsymbol{x}_k \in \mathcal{B}(\boldsymbol{x}^*;\delta)$,*

$$\nabla f(\boldsymbol{x}_k)^T \boldsymbol{d}_k = \overline{Df}(\boldsymbol{x}_k;\boldsymbol{d}_k) + o(\|\boldsymbol{d}_k\|^2),$$

then. for any $\sigma \in \left(0,\,\frac{1}{2}\right)$, there exists a $\overline{k}$ such that, for all $k > \overline{k}$

$$f(\boldsymbol{x}_k + \boldsymbol{d}) \le f(\boldsymbol{x}_k) + \sigma\overline{Df}(\boldsymbol{x}_k;\boldsymbol{d}_k).$$

The following result in [7] presents a particular feature of the *merit function* (1.5).

Theorem 2.10 ([7]). *Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $F(\boldsymbol{x}) = (F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), \ldots, F_n(\boldsymbol{x}))$ be a nonlinear and continuously differentiable function, Φ_λ given by* (1.4) *associated to F and Ψ_λ the merit function given by* (1.5) *associated to Φ_λ. If for all $i = 1, 2, \ldots, n$, the function F_i is SC^1 function, then Ψ_λ is also a SC^1 function.*

Finally, we recall that the merit function (1.5) is *differentiable* and its gradient can be calculated using matrices $H \in \partial\Phi_\lambda(\boldsymbol{x})$ as follows [11].

Theorem 2.11 ([11]). *For any matrix $H \in \partial\Phi_\lambda(\boldsymbol{x})$, the merit function Ψ_λ given by* (1.5) *is continuously differentiable and*

$$\nabla\Psi_\lambda(\boldsymbol{x}) = H^T\Phi_\lambda(\boldsymbol{x}). \tag{2.7}$$

3 Algorithm and Convergence

In this section, we propose a *global nonsmooth quasi Newton* algorithm for solving NCP by solving the minimization problem (1.6). In fact, for this algorithm, we prove global and local convergence results.

The algorithm applies a nonsmooth *quasi Newton* method as introduced and investigated by [1] for solving the nonlinear system of equations $\Phi_\lambda(\boldsymbol{x}) = 0$ and then, the method is globalized used a *merit function* Ψ_λ. For this, we use a linear search strategy as used in [11]. The following is a precise statement of our algorithm.

Algorithm 3.1 (Global quasi Newton method $\Phi_\lambda(\mathbf{x}) = 0$).

P.0. Initialization
Choose $\lambda \in (0,4)$, $\mathbf{x}_0 \in \mathbb{R}^n$, $\rho > 0$, $\mu \in (0,1)$, $\sigma \in (0,\frac{1}{2})$, $p > 2$, $\varepsilon \ge 0$, $N > 0$ and set $k := 0$ and $A_0 = F'(x_0)$.

P.1. Stopping criterion
If $\|B_k^T\Phi_\lambda(\mathbf{x}_k)\| \le \varepsilon$ or $k > N$ stop.

P.2. Search direction

Calculate B_k as in *(2.6)* and find $\mathbf{d}_k \in \mathbb{R}^n$ such that

$$B_k \mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k). \tag{3.1}$$

If this system is not solvable or if $\Phi_\lambda(\mathbf{x}_k)^T B_k \mathbf{d}_k > -\rho\|\mathbf{d}_k\|^p$ set $\mathbf{d}_k = -B_k^T \Phi_\lambda(\mathbf{x}_k)$.

P.3. Linear search

Compute $t_k := \max\{\mu^l \,|\, l = 0, 1, 2, ...\}$ such that

$$\Psi_\lambda(\mathbf{x}_k + t_k \mathbf{d}_k) \leq \Psi_\lambda(\mathbf{x}_k) + \sigma t_k B_k^T \Phi_\lambda(\mathbf{x}_k)\mathbf{d}_k \,. \tag{3.2}$$

P.4. Update

Set $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_k$, $k = k + 1$, update A_k and go to **P.1**.

We observe that the approximation matrix B_k given by (2.6) depends on the approximation A_k to the jacobian matrix $F'(\boldsymbol{x}_k)$. Thus, a natural question is "how to update the matrix A_k" that is, "what is A_{k+1}" The way we update A_k results in a diverse range of *quasi-Newton* methods, among which there are the so-called *Least Change Secant Update methods*, in which the matrix A_{k+1}, must satisfy the *secant equation* [13]

$$A_{k+1}(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) = F(\boldsymbol{x}_{k+1}) - F(\boldsymbol{x}_k) \tag{3.3}$$

and its change, measured in some norm, respect to the A_k, must be minimum. Examples of this type of updates are the *"Good" Broyden update* [5], *"Bad" Broyden update* [4] and *Schubert update* [6]. We will use these three updates in the Section 4, when we analyze the numerical performance of our algorithm. Moreover, in our code we chose the vector $\boldsymbol{z}$ in (2.6) such that $z_i = 0$ if $i \notin \beta$ and $z_i = 1$ if $i \in \beta$.

Given that we do not want to calculate matrices $H_k \in \partial\Phi_\lambda(\boldsymbol{x}_k)$, we proposed some variations in the classical *search direction calculation* and in the *linear search* (*Armijo* condition [13]). By Theorem 2.11, to calculate the gradient of Ψ_λ, in each iteration, the matrices $H_k \in \partial\,\Phi_\lambda(\boldsymbol{x}_k)$ must be calculated. As an alternative, we propose the following approximation to the $\nabla\Psi_\lambda(\boldsymbol{x}_k)$,

$$\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k) = B_k^T \Phi_\lambda(\boldsymbol{x}_k). \tag{3.4}$$

This approximation requires the matrices B_k defined by (2.6) and to evaluate Φ_λ in $\boldsymbol{x}_k$.

Under the following assumptions, we will prove global and local convergence properties of Algorithm 3.1.

3.1 Hypothesis

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

A2. The functions F_i are SC^1 functions.

A3. There exist positive constants γ and ϵ, such that $\|F'(\boldsymbol{x}) - F'(\boldsymbol{x}^*)\| \leq \gamma\|\boldsymbol{x} - \boldsymbol{x}^*\|$, for all $\boldsymbol{x} \in \mathcal{B}\,(\boldsymbol{x}^*, \epsilon)$.

3.2 Convergence

The following two lemmas will be useful in demonstrating the convergence theorems of Algorithm 3.1. In the first one, we prove that under certain conditions, the norm of the matrices that approximate the respective matrices in $\partial\Phi_\lambda\,(\boldsymbol{x})$ is bounded. In the second lemma, we prove an equivalence between the gradient of the *merit function* Ψ_λ and the approximation that we propose in (3.4).

Lemma 3.2. *Let B given by (2.6) and $\delta > 0$. For any matrix $A \in \mathcal{B}(F'(\boldsymbol{x}^*); \delta)$, there exists a positive constant η such that*

$$\|B\| \leq \eta. \tag{3.5}$$

Proof. We consider the matrix B defined by (2.6). By definition of *infinite matrix norm*, we have

$$\|B\|_\infty = \max_{1\leq i\leq n}\{\|[B]_i\|_1\}. \tag{3.6}$$

We assume that the maximum is reached in the row j. If $j \notin \beta$, then $\boldsymbol{x}_j \neq 0$ or $F_j(\boldsymbol{x}) \neq 0$, from (2.5)

$$|\chi(x_j,\, F_j(\boldsymbol{x}))| \leq 1 \qquad \textit{and} \qquad |\psi(x_j,\, F_j(\boldsymbol{x}))| \leq \sqrt{2},$$

moreover,

$$\begin{aligned}
\|B\|_\infty &= \|(\chi(x_j,\, F_j(\boldsymbol{x})) - 1)e_j^T + (\psi(x_j,\, F_j(\boldsymbol{x})) - 1)[A]_j\|_\infty \\
&\leq |\chi(x_j,\, F_j(\boldsymbol{x})) - 1| + |\psi(x_j,\, F_j(\boldsymbol{x})) - 1|\|[A]_j\|_\infty \\
&\leq 2 + \left(1+\sqrt{2}\right)\|[A]_j\|_\infty \\
&\leq 2 + \left(1+\sqrt{2}\right)(\|F'(\boldsymbol{x}^*)\|_\infty + \delta).
\end{aligned}$$

If $j \in \beta$, following the same procedure, we obtain the same upper bound for $\|B\|_\infty$. Therefore, exists

$$\eta = 2 + \left(1+\sqrt{2}\right)(\|F'(\boldsymbol{x}^*)\|_\infty + \delta)$$

such that $\|B\|_\infty \leq \eta$. □

Lemma 3.3. *If $\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)$ is the approximation to $\nabla\Psi_\lambda(\boldsymbol{x}_k)$ given by (3.4), then*

$$\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k = \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k + o(\|\boldsymbol{d}_k\|_\infty^2). \tag{3.7}$$

Proof. From (2.7) and (3.4),

$$|\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k - \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k| \;=\; |\Phi_\lambda(\boldsymbol{x}_k)^T(H_k - B_k)\boldsymbol{d}_k|. \tag{3.8}$$

Using the *Cauchy-Schwarz* inequality and the relation between the *Euclidean* and *infinity* vector norms in (3.8), we obtain the inequality

$$|\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k - \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k| \leq\; n\,\|\Phi_\lambda(\boldsymbol{x}_k)\|_\infty\|H_k - B_k\|_\infty\|\boldsymbol{d}_k\|_\infty.$$

By Theorem 2.2, for k sufficiently large, there exists a constant θ such that $\|H_k - B_k\|_\infty \leq \theta$, then

$$|\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k - \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k| \;\leq\; n\,\theta\,\|\Phi_\lambda(\boldsymbol{x}_k)\|_\infty\|\boldsymbol{d}_k\|_\infty$$

thus,

$$0 \leq \;\frac{|\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k - \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k|}{\|\boldsymbol{d}_k\|_\infty^2} \leq\; \frac{n\,\theta\,\|\Phi_\lambda(\boldsymbol{x}_k)\|_\infty}{\|\boldsymbol{d}_k\|_\infty},$$

therefore, given that with the Algorithm 3.1 we want to minimize the *merit function* Ψ_λ and that $\boldsymbol{d}_k$ is a decent direction, we conclude that although the sequence $\{\|\boldsymbol{d}_k\|\}$ converges to zero, the sequence $\{\|\Phi_\lambda(\boldsymbol{x}_k)\|\}$ will make it faster, thus

$$\lim_{k\to\infty}\frac{|\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k - \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k|}{\|\boldsymbol{d}_k\|_\infty^2} = 0$$

proving (3.7). □

The following convergence theorems are analogous to those presented in [11] for a semismooth Newton-type method to solve the NCP.

Theorem 3.4. *Every accumulation point of a sequence* $\{\boldsymbol{x}_k\}$ *generated by Algorithm 3.1 is a stationary point of* Ψ_λ.

Proof. If, for an infinite set of indices J, the direction given by the second step of Algorithm 3.1 is always $\boldsymbol{d}_k = -\nabla\Psi_\lambda(\boldsymbol{x}_k)$, then any limit point of subsequence $\{\boldsymbol{x}_k\}_J$ is a stationary point of Ψ_λ [2]. Thus, we can assume, without loss of generality, that the direction is always given by

$$B_k\,\boldsymbol{d}_k = -\Phi_\lambda(\boldsymbol{x}_k). \tag{3.9}$$

Suppose now, that $\boldsymbol{x}^*$ is not an stationary point of Ψ_λ, that is, $\nabla\Psi_\lambda(\boldsymbol{x}^*) \neq \boldsymbol{0}$. We will prove that there are positive constant m and M such that

$$m \leq \|\boldsymbol{d}_k\| \leq M. \tag{3.10}$$

Indeed, from (3.9)

$$\|\Phi_\lambda(\boldsymbol{x}_k)\| \leq \|B_k\|\|\boldsymbol{d}_k\|$$

then

$$\frac{\|\Phi_\lambda(\boldsymbol{x}_k)\|}{\|B_k\|} \leq \|\boldsymbol{d}_k\|$$

thus, if the sequence $\{\boldsymbol{d}_k\}$ converges to the zero vector, then the sequence $\{\|\Phi_\lambda(\boldsymbol{x}_k)\|\}$ converges to cero since $\|B_k\|$ is bounded for all k (Lemma 3.2), which implies that the $\{\Phi_\lambda(\boldsymbol{x}_k)\}$ converges to the vector zero, whereby $\nabla\Psi_\lambda(\boldsymbol{x}_k)$ also converges to vector zero and given that the function Φ_λ is continuous, then

$$\nabla\Psi_\lambda(\boldsymbol{x}^*) = \boldsymbol{0},$$

contradicting our assumption. Thus, there exists a positive constant m such that $m \leq \|\boldsymbol{d}_k\|$.

On the other hand, if $\|\boldsymbol{d}_k\|$ had no upper bound, but $\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)$ is bounded and $p \geq 2$, then

$$\lim_{k\to\infty} \frac{\|\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)\|\cos\theta}{\|\boldsymbol{d}_k\|^{p-1}} = 0$$

where θ is the angle between $\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)$ and $\boldsymbol{d}_k$. equivalently,

$$\lim_{k\to\infty} \frac{\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|^{p}} = 0$$

contradicting the fact

$$\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k \leq -\rho\|\boldsymbol{d}_k\|^p. \tag{3.11}$$

Thus, there exists a positive constant M such that $\|\boldsymbol{d}_k\| \leq M$.

Now, since the sequence $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}^*$ and the next iteration of Algorithm 3.1 is generated by $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + t_k\,\boldsymbol{d}_k$, then

$$t_k\,\boldsymbol{d}_k \to \boldsymbol{0}. \tag{3.12}$$

By (3.10), we can assume that, if $k \to \infty$, $\boldsymbol{d}_k \to \overline{\boldsymbol{d}} \neq \boldsymbol{0}$, then $t_k = \mu^{l_k} \to 0$. Moreover, from (3.2)

$$\frac{\Psi_\lambda(\boldsymbol{x}_k + \mu^{l_k-1}\boldsymbol{d}_k) - \Psi_\lambda(\boldsymbol{x}_k)}{\mu^{l_k-1}} > \sigma\overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k \tag{3.13}$$

Thus, from (3.13), if $k \to \infty$

$$\nabla\Psi_\lambda(\boldsymbol{x}^*)^T\,\overline{\boldsymbol{d}} \geq \sigma\,\overline{\nabla\Psi_\lambda}(\boldsymbol{x}^*)^T\,\overline{\boldsymbol{d}}$$

then, by Lemma 3.3, we can conclude that

$$\sigma \geq 1,$$

contradicting that $\sigma \in \left(0, \frac{1}{2}\right)$. Thus, from (3.12), we have that necessarily the sequence $\{\boldsymbol{d}_k\}$ converges to zero vector and therefore, $\{\|\boldsymbol{d}_k\|\}$ converges to zero, contradicting the fact $\nabla\Psi_\lambda(\boldsymbol{x}^*) \neq \mathbf{0}$. In conclusion, $\nabla\Psi_\lambda(\boldsymbol{x}^*) = \mathbf{0}$. □

In the Theorem 3.6, we refer to the apparently contradictory term, in order to avoid ambiguity, we include de following definition.

Definition 3.5. Let Ω be the accumulation points set of the sequence $\{\mathbf{x}_k\}$. We say that $\mathbf{x}^* \in \Omega$ is an isolated accumulation point, if $\mathbf{x}^*$ is an isolated point in Ω, i.e., If there exist $\delta > 0$ such that

$$\mathcal{B}(\mathbf{x}^*; \delta) \cap \Omega = \{\mathbf{x}^*\}.$$

Theorem 3.6. *Suppose that $\boldsymbol{x}^*$ is isolated accumulation point of a sequence generated by Algorithm 3.1. Then the entire sequence converges to $\boldsymbol{x}^*$.*

Proof. Let $\{\boldsymbol{x}_k\}$ be the sequence generated by Algorithm 3.1 and $\boldsymbol{x}^*$ an isolated accumulation point of this sequence. As $\Psi(\boldsymbol{x}^*) = 0$ and Ψ is a convex function, then $\boldsymbol{x}^*$ is a global minimizer of Ψ.

Let Ω be the accumulation points set of $\{\boldsymbol{x}_k\}$, then $\Omega \neq \emptyset$, since $\boldsymbol{x}^* \in \Omega$. We define

$$\delta = \begin{cases} dist(\boldsymbol{x}^*;\, \Omega \backslash \{\boldsymbol{x}^*\}) & \text{if} \quad \Omega \backslash \{\boldsymbol{x}^*\} \neq \emptyset \\ 1 & \text{if} \quad \Omega = \{\boldsymbol{x}^*\} \end{cases}$$

where $dist(\boldsymbol{a},\, A) = \inf_{x \in A} \|\boldsymbol{a} - \boldsymbol{x}\|$ is the distance of point $\boldsymbol{a}$ to the set A.

Given that, locally, $\boldsymbol{x}^*$ is unique, then $\delta > 0$. If

$$\Omega_1 = \left\{ \boldsymbol{y} \in \mathbb{R}^n : \; dist(\boldsymbol{y};\, \Omega) \leq \frac{\delta}{4} \right\},$$

then there exists $\overline{k}$ such that $\boldsymbol{x}_k \in \Omega_1$ for all $k \geq \overline{k}$.

Now, if

$$\Lambda = \left\{ k \in \mathbb{N} : \; \|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leq \frac{\delta}{4} \right\},$$

then $\{\boldsymbol{x}_k\}_\Lambda \subset \overline{\mathcal{B}}\left(x^*;\, \frac{\delta}{4}\right)$. Since $\boldsymbol{x}^*$ is the only accumulation point of $\{\boldsymbol{x}_k\}$ in $\overline{\mathcal{B}}\left(\boldsymbol{x}^*;\, \frac{\delta}{4}\right)$, we can conclude that $\{\boldsymbol{x}_k\}_\Lambda$ converges to $\boldsymbol{x}^*$ and since $\Psi(\boldsymbol{x}^*) = 0$, then $\{\|\Psi(\boldsymbol{x}_k)\|\}_\Lambda$ converges to zero, which in turn, by (3.11) and (3.7) implies that the sequence $\{\boldsymbol{d}_k\}$ converges to the zero vector. Thus, there exists $\widetilde{k}$ such that, if $k \in \Lambda$ and $k \geq \widetilde{k} \geq \overline{k}$, then

$$\|\boldsymbol{d}_k\| \leq \frac{\delta}{4}.$$

Let $\widehat{k} \in \Lambda$ be such that $\widehat{k} \geq \widetilde{k}$, we observe that

$$\boldsymbol{x}_{\widehat{k}+1} = \boldsymbol{x}_{\widehat{k}} + t_{\widehat{k}}\boldsymbol{d}_{\widehat{k}}, \tag{3.14}$$

with $t_{\widehat{k}} \in (0, 1]$. Thus

$$\|\boldsymbol{x}_{\widehat{k}+1} - \boldsymbol{x}_{\widehat{k}}\| \leq \|\boldsymbol{d}_{\widehat{k}}\| \leq \frac{\delta}{4},$$

therefore,

$$\begin{aligned} dist(\boldsymbol{x}^*; \Omega\backslash\{\boldsymbol{x}^*\}) &\leq dist(\boldsymbol{x}_{\widehat{k}+1}; \Omega\backslash\{\boldsymbol{x}^*\}) + \|\boldsymbol{x}^* - \boldsymbol{x}_{\widehat{k}+1}\| \\ &\leq dist(\boldsymbol{x}_{\widehat{k}+1}; \Omega\backslash\{\boldsymbol{x}^*\}) + \|\boldsymbol{x}^* - \boldsymbol{x}_{\widehat{k}}\| + \|\boldsymbol{x}_{\widehat{k}} - \boldsymbol{x}_{\widehat{k}+1}\| \\ &\leq dist(\boldsymbol{x}_{\widehat{k}+1}; \Omega\backslash\{\boldsymbol{x}^*\}) + \frac{\delta}{4} + \frac{\delta}{4} \end{aligned}$$

that is,

$$dist(\boldsymbol{x}_{\widehat{k}+1}; \Omega\backslash\{\boldsymbol{x}^*\}) \geq dist(\boldsymbol{x}^*; \Omega\backslash\{\boldsymbol{x}^*\}) - \frac{\delta}{2} \geq \delta - \frac{\delta}{2} = \frac{\delta}{2}$$

thus, $\boldsymbol{x}_{\widehat{k}+1} \notin \Omega_1\backslash\overline{\mathcal{B}}\left(\boldsymbol{x}^*; \frac{\delta}{4}\right)$ and, since $\boldsymbol{x}_{\widehat{k}+1} \in \Omega_1$, then $\boldsymbol{x}_{\widehat{k}+1} \in \overline{\mathcal{B}}\left(x^*; \frac{\delta}{4}\right)$, i.e., $\widehat{k}+1 \in \Lambda$ and given that $\widehat{k}+1 > \widetilde{k}$, then, for induction, $\widehat{k} \in \Lambda$ for all $\widehat{k} > \widetilde{k}$, thus, $\boldsymbol{x}_k \in \overline{\mathcal{B}}\left(\boldsymbol{x}^*; \frac{\delta}{4}\right)$ for all $k \geq \widetilde{k}$, i.e., $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}^*$. □

The following theorem shows that there exist a positive integer $\overline{k}$ such that, for all $k > \overline{k}$, the *local* behavior of the Algorithm 3.1 is identical to the Algorithm 4.1 proposed in [1] allowing us to guarantee convergence rate similar to that obtained for the latter algorithm. It is important to observe that, if $\boldsymbol{x}^*$ is *R-regular*, the assumption **H3** in [1] is immediately satisfied.

Theorem 3.7. *Assume that $\boldsymbol{x}^*$ is an accumulation point of a sequence $\{\boldsymbol{x}_k\}$ generated by Algorithm 3.1 such that $\boldsymbol{x}^*$ is an R-regular solution of NCP. Then $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}^*$, the search direction $\boldsymbol{d}_k$ is eventually given by the solution of the linear system $B_k\boldsymbol{d}_k = -\Phi_\lambda(\boldsymbol{x}_k)$ and the full stepsize $t_k = 1$ is accepted for all k sufficiently large.*

Proof. If $\boldsymbol{x}^*$ is a *R-regular* solution of NCP, x^* is also a *BD-regular* solution (Corollary 2.6), then, by Proposition 3 in [14], $\boldsymbol{x}^*$ is an isolated accumulation point of the sequence $\{\boldsymbol{x}_k\}$ generated by the Algorithm 3.1 and, from the Theorem 3.6, we have that the sequence $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}^*$.

On the other hand, by the Theorem 2.3, for some k sufficiently large, the matrix B_k^{-1} exists and therefore, the linear system of equations $B_k\boldsymbol{d}_k = -\Phi_\lambda(\boldsymbol{x}_k)$ has a solution. Thus,

$$\boldsymbol{d}_k = -B_k^{-1}\Phi_\lambda(\boldsymbol{x}_k), \tag{3.15}$$

For any vector norm and its associated matricial norm, we have

$$\|\boldsymbol{d}_k\| \leq \|\Phi_\lambda(\boldsymbol{x}_k)\|\|B_k^{-1}\|. \tag{3.16}$$

From (3.4), (3.15), (3.16)

$$\begin{aligned} \overline{\nabla\Psi_\lambda}(\boldsymbol{x}_k)^T\boldsymbol{d}_k &= \Phi_\lambda(\boldsymbol{x}_k)^T B_k\boldsymbol{d}_k \\ &= -\Phi_\lambda(\boldsymbol{x}_k)^T B_k B_k^{-1}\Phi_\lambda(\boldsymbol{x}_k) \\ &= -\|\Phi_\lambda(\boldsymbol{x}_k)\|^2 \\ &\leq -\frac{\|d_k\|^2}{\|B_k^{-1}\|^2} \end{aligned}$$

then, we can define the constant $\rho = \eta^2$, where η is the upper bound of $\|B_k\|$ given by the Theorem 2.3. That is, since the sequence $\{\|\boldsymbol{d}_k\|\}$ converges to zero then

$$\overline{\nabla\Psi_\lambda(\boldsymbol{x}_k)\boldsymbol{d}_k} \leq -\rho\|\boldsymbol{d}_k\|^p$$

for all $p \geq 2$. This proves that for a k sufficiently large, the search direction is always given by (3.10).

Now, by the Lemma 3.3, the Theorem 2.9 and since Ψ_λ is a SC^1 function, we have that, for any $\sigma \in \left(0, \frac{1}{2}\right)$, there exist $\overline{k}$ such that, for all $k > \overline{k}$

$$\Psi_\lambda(\boldsymbol{x}_k + \boldsymbol{d}_k) \leq \Psi_\lambda(\boldsymbol{x}_k) + \sigma B_k^T \Phi_\lambda(\boldsymbol{x}_k)\boldsymbol{d}_k$$

i.e., the stepsize $t_k = 1$ is accepted for k sufficiently large. □

4 Numerical Results

In this section, we analyzed numerically the global behavior of the algorithm proposed in the last section (Algorithm 3.1). For this, we compare our algorithm with the *semismooth Newton-type method* proposed in [11]. This latter algorithm uses at each iteration, a matrix $H_k \in \partial\Phi_\lambda(\boldsymbol{x})$. This is basically the difference with our algorithm.

For the numerical tests, we consider four nonlinear complementarity problems associated respectively with the functions: Kojima Shindo (Koj-Shi), Kojima Josephy (Koj-Jo), Modify Mathiesen (Mathiesen), Billups (Billups) [3, 14]. The latter problem was constructed by *Billups* [3] in order to make almost all state-of-the-art methods to fail on this problem [11]. We observe that, our method has not problem to solve this example.

We implemented the algorithms and tests functions in MATLAB® and for each problem, we use all the starting points used in [11] for their tests. Moreover, we use the following values for the parameters required in the algorithm $\rho = 10^{-8}$, $\sigma = 10^{-4}$, $p = 2.1$, $\varepsilon = 10^{-12}$, $N = 200$, $\mu = 0.5$. Similarly, considering the results presented in [11], we use the same dynamical choice of λ that they used; i.e., in each iteration, we update λ as follows

1. Set $\lambda = 2$ at the beginning of each iteration.

2. if $\Psi_\lambda(\boldsymbol{x}_k) \leq \gamma_1$, then set $\lambda := \Psi_\lambda(\boldsymbol{x}_k)$, else set $\lambda := \min\{c_1\Psi(\boldsymbol{x}_k), \lambda\}$

3. if $\Psi(\boldsymbol{x}_k) \leq \gamma_2$, then set $\lambda := \min\{c_2, \lambda\}$

where $\gamma_1 = 10^{-2}$, $\gamma_2 = 10^{-4}$, $c_1 = 10$, $c_2 = 10^{-8}$. With the dynamical choice of the parameter λ, the algorithms perform their first iterations using *Fisher function* [10], which gives the best global convergence properties, and completes its cycle with complementary functions that tend to be multiples of the *minimum function* [16] [15], which usually results in a faster local convergence [16].

Table 1 presents the results of our numerical tests. Its columns contains the following information: *Problem* means the problem name, $\boldsymbol{x}_0$ is the starting point, t is the runtime (in seconds) of the algorithms, k is the number of iterations and $\boldsymbol{x}^*$ is the solution obtained. We also include a column with the algorithm and the secant update used. Thus, *GN* means *Generalized Newton*; *GB*, *BB* and *SC* means Algorithm 3.1 with the *"Good" Broyden update*, *"Bad" Broyden update* and the *Schubert update*, respectively. A **-** sign means divergence.

Problem	*Algorithm*	$\boldsymbol{x}_0$	$t\,(sec)$	k	$\boldsymbol{x}^*$
Koj-Jo	GN	$(0\ 0\ 0\ 0)^T$	0.152	21	$(1\ 0\ 3\ 0)^T$
Koj-Jo	GB	$(0\ 0\ 0\ 0)^T$	0.141	21	$(1\ 0\ 3\ 0)^T$
Koj-Jo	BB	$(0\ 0\ 0\ 0)^T$	-	-	-
Koj-Jo	SC	$(0\ 0\ 0\ 0)^T$	-	-	-
Koj-Jo	GN	$(1\ 1\ 1\ 1)^T$	0.060	8	$(1\ 0\ 3\ 0)^T$
Koj-Jo	GB	$(1\ 1\ 1\ 1)^T$	0.086	13	$(1\ 0\ 3\ 0)^T$
Koj-Jo	BB	$(1\ 1\ 1\ 1)^T$	0.074	12	$(1\ 0\ 3\ 0)^T$
Koj-Jo	SC	$(1\ 1\ 1\ 1)^T$	-	-	-
Koj-Jo	NG	$(0\ 0\ 0\ 100)^T$	-	-	-
Koj-Jo	GB	$(0\ 0\ 0\ 100)^T$	0.167	24	$(1\ 0\ 3\ 0)^T$
Koj-Jo	BB	$(0\ 0\ 0\ 100)^T$	-	-	-
Koj-Jo	SC	$(0\ 0\ 0\ 100)^T$	-	-	-
Koj-Jo	GN	$(1\ 0\ 1\ 0)^T$	0.038	6	$(1\ 0\ 3\ 0)^T$
Koj-Jo	GB	$(1\ 0\ 1\ 0)^T$	0.051	8	$(1\ 0\ 3\ 0)^T$
Koj-Jo	BB	$(1\ 0\ 1\ 0)^T$	0.056	8	$(1\ 0\ 3\ 0)^T$
Koj-Jo	SC	$(1\ 0\ 1\ 0)^T$	-	-	-
Koj-Jo	GN	$(1\ 0\ 0\ 0)^T$	0.082	8	$(1\ 0\ 3\ 0)^T$
Koj-Jo	GB	$(1\ 0\ 0\ 0)^T$	0.067	10	$(1\ 0\ 3\ 0)^T$
Koj-Jo	BB	$(1\ 0\ 0\ 0)^T$	0.087	12	$(1\ 0\ 3\ 0)^T$
Koj-Jo	SC	$(1\ 0\ 0\ 0)^T$	-	-	-
Koj-Jo	GN	$(0\ 1\ 1\ 0)^T$	0.066	9	$(1\ 0\ 3\ 0)^T$
Koj-Jo	GB	$(0\ 1\ 1\ 0)^T$	0.089	14	$(1\ 0\ 3\ 0)^T$
Koj-Jo	BB	$(0\ 1\ 1\ 0)^T$	0.229	17	$(1\ 0\ 3\ 0)^T$
Koj-Jo	SC	$(0\ 1\ 1\ 0)^T$	-	-	-
Koj-Shi	GN	$(0\ 0\ 0\ 0)^T$	0.212	15	$(1\ 0\ 3\ 0)^T$
Koj-Shi	GB	$(0\ 0\ 0\ 0)^T$	0.131	21	$(1\ 0\ 3\ 0)^T$
Koj-Shi	BB	$(0\ 0\ 0\ 0)^T$	-	-	-
Koj-Shi	SC	$(0\ 0\ 0\ 0)^T$	-	-	-
Koj-Shi	GN	$(1\ 1\ 1\ 1)^T$	0.067	10	$(1\ 0\ 3\ 0)^T$
Koj-Shi	GB	$(1\ 1\ 1\ 1)^T$	0.141	16	$(1\ 0\ 3\ 0)^T$
Koj-Shi	BB	$(1\ 1\ 1\ 1)^T$	0.102	16	$(1\ 0\ 3\ 0)^T$
Koj-Shi	SC	$(1\ 1\ 1\ 1)^T$	-	-	-
Koj-Shi	GC	$(0\ 0\ 0\ 100)^T$	0.129	17	$(1\ 0\ 3\ 0)^T$
Koj-Shi	GB	$(0\ 0\ 0\ 100)^T$	0.168	22	$(1\ 0\ 3\ 0)^T$
Koj-Shi	BB	$(0\ 0\ 0\ 100)^T$	-	-	-
Koj-Shi	SC	$(0\ 0\ 0\ 100)^T$	-	-	-
Koj-Shi	GN	$(1\ 0\ 1\ 0)^T$	0.039	6	$(1\ 0\ 3\ 0)^T$
Koj-Shi	GB	$(1\ 0\ 1\ 0)^T$	0.046	7	$(1\ 0\ 3\ 0)^T$

Problem	*Algorithm*	$\boldsymbol{x}_0$	$t\,(sec)$	k	$\boldsymbol{x}^*$
Koj-Shi	BB	$(1\ 0\ 1\ 0)^T$	0.044	7	$(1\ 0\ 3\ 0)^T$
Koj-Shi	GN	$(1\ 0\ 0\ 0)^T$	0.051	8	$\left(\frac{\sqrt{6}}{2}\ 0\ 0\ \frac{1}{2}\right)^T$
Koj-Shi	GB	$(1\ 0\ 0\ 0)^T$	0.060	9	$\left(\frac{\sqrt{6}}{2}\ 0\ 0\ \frac{1}{2}\right)^T$
Koj-Shi	BB	$(1\ 0\ 0\ 0)^T$	0.057	9	$\left(\frac{\sqrt{6}}{2}\ 0\ 0\ \frac{1}{2}\right)^T$
Koj-Shi	SC	$(1\ 0\ 0\ 0)^T$	-	-	-
Koj-Shi	GN	$(0\ 1\ 1\ 0)^T$	0.070	10	$(1\ 0\ 3\ 0)^T$
Koj-Shi	GB	$(0\ 1\ 1\ 0)^T$	0.093	15	$(1\ 0\ 3\ 0)^T$
Koj-Shi	BB	$(0\ 1\ 1\ 0)^T$	0.125	15	$(1\ 0\ 3\ 0)^T$
Koj-Shi	SC	$(0\ 1\ 1\ 0)^T$	-	-	-
Mathiesen	GN	$(1\ 1\ 1\ 1)^T$	0.186	11	$(0.0302\ 0\ 0\ 0)^T$
Mathiesen	GB	$(1\ 1\ 1\ 1)^T$	0.178	11	$(0.0302\ 0\ 0\ 0)^T$
Mathiesen	BB	$(1\ 1\ 1\ 1)^T$	0.091	11	$(0.0302\ 0\ 0\ 0)^T$
Mathiesen	SC	$(1\ 1\ 1\ 1)^T$	0.070	5	$(1.497\ 0\ 0\ 0)^T$
Mathiesen	GN	$(100\ 100\ 100\ 100)^T$	0.102	13	$(2.999\ 0\ 0\ 0)^T$
Mathiesen	GB	$(100\ 100\ 100\ 100)^T$	0.132	7	$(3\ 0\ 0\ 0)^T$
Mathiesen	BB	$(100\ 100\ 100\ 100)^T$	0.054	7	$(3\ 0\ 0\ 0)^T$
Mathiesen	SC	$(100\ 100\ 100\ 100)^T$	0.094	9	$(2.320\ 0\ 0\ 0)^T$
Mathiesen	GN	$(1\ 0\ 1\ 0)^T$	0.091	4	$(0\ 0\ 0\ 0)^T$
Mathiesen	GB	$(1\ 0\ 1\ 0)^T$	0.046	6	$(0\ 0\ 0\ 0)^T$
Mathiesen	BB	$(1\ 0\ 1\ 0)^T$	0.043	6	$(0\ 0\ 0\ 0)^T$
Mathiesen	SC	$(1\ 0\ 1\ 0)^T$	0.071	6	$(0.018\ 0\ 0\ 0)^T$
Mathiesen	GN	$(0\ 1\ 1\ 0)^T$	0.037	5	$(0.752\ 0\ 0\ 0)^T$
Mathiesen	GB	$(0\ 1\ 1\ 0)^T$	0.036	5	$(0.608\ 0\ 0\ 0)^T$
Mathiesen	BB	$(0\ 1\ 1\ 0)^T$	0.036	5	$(0.647\ 0\ 0\ 0)^T$
Mathiesen	SC	$(0\ 1\ 1\ 0)^T$	0.037	5	$(0.778\ 0\ 0\ 0)^T$
Billups	GN	0	-	-	-
Billups	GB	0	2.843	16	2.0488
Billups	BB	0	2.827	16	2.0488
Billups	SC	0	-	-	-

Table 1: Numerical results for the algorithms.

These results show the importance of taking into account the running time of the algorithms because, comparing only the number of iterations performed by each of them is not enough because, although in general the number of iterations performed by the Algorithm 3.1 is greater than those made by the *Generalized Newton* algorithm, the number of operations performed at each iteration by the Algorithm 3.1 is generally smaller than the number of operations performed by de *Generalize Newton* algorithm. In particular, for the *Mathiesen* problem, the Algorithm 3.1 was faster than the other one.

The results show the good performance of Algorithm 3.1 when it uses the *"Good" Broyden update*, it converges for a 93% of the problems and starting points. Moreover, the *Billups* problem is solved with the Algorithm 3.1 while the other algorithm diverges.

5 Final Remarks

In this paper, we propose a global nonsmooth quasi-Newton method for solving the NCP solving alternatively the unconstrained minimization problem (1.6) associated with the reformulation of NCP as a nonlinear system of equations. Moreover, we established global and local convergence results for the algorithm proposed.

Following the philosophy of quasi Newton methods, we were not interested in calculating matrices of generalized Jacobian of Φ_λ, therefore we proposed an approximation to the gradient of *merit function* (1.6) and we show how good it is when we compared it to the exact gradient.

Preliminary numerical experiments shows a good performance of our algorithm, but it is necessary more numerical tests. We believe that other dynamical choice of parameter λ could improve the performance of our algorithm.

Acknowledgment

The authors are indebted to the anonymous referees for their comments and suggestions which improve the paper.

References

[1] F. E. Arenas, *Metodos secantes de cambio minimo para el problema de complementariedad no lineal*, Tesis de Maestria, Universidad del Cauca, 2013.

[2] D. Bertzekas, *Constrained Optimization and Lagrange Multiplier Methods*, Athena Scientific, 1996.

[3] S.C. Billups, *Algorithms for complementarity problems and Generalized Equations*, Ph. D. Thesis, University of Wisconsin, 1995.

[4] P.N. Brown and Y. Saad, Convergence Theory of Nonlinear Newton-Krylov Algorithms, *SIAM J. Optim.* 4 (1994) 297–330.

[5] C.G. Broyden, A Class of Methods for Solving Nonlinear Simultaneous Equations, *Math. Comp.* 19 (1965) 577–593.

[6] T.F. Coleman, B.S. Garbow and J.J. More, Software for estimating sparse Jacobian matrices, *ACM Trans. Math. Software.* 10 (1984) 329–345.

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

[8] F. Facchinei, Minimization of SC^1 Functions and the Maratos Effect, *Oper. Res. Lett.* 23 (1994) 131–137.

[9] M. C. Ferris and J. Pang, Engineering and economic applications of complementarity problems, *SIAM Rev.* 39 (1997) 669–713.

[10] A. Fischer and C. Kanzow, On finite termination of an iterative method for linear complementarity problems, *Math. Program.* 74 (1996) 279–292.

[11] C, Kanzow and H. Kleinmichel, A New Class of Semismooth Newton-Type Methods for Nonlinear Complementarity Problems, *Comput. Optim. Appl.* 11 (1993) 227–251.

[12] M. Kostreva, Elasto-hydrodynamic lubrication: A non-linear complementarity problem, *Internat. J. Numer. Methods Fluids.* 4 (1984) 377–397.

[13] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research, 2000.

[14] J. Pang and S. A. Gabriel, A robust algorithm for the nonlinear complementarity problem, *Math. Program.* 60 (1993) 295–337.

[15] J. Pang and L. Qi, Nonsmooth Equations: Motivation and Algorithms, *SIAM J. Optim.* 3 (1993) 443–465.

[16] V. L. Rocha, J. M. Martinez and R. Perez, On the local convergence of quasi-Newton methods for nonlinear complementary problems, *Appl. Numer. Math.* 30 (1999) 3–22.

[17] J. G. Wardrop, Some theoretical aspects of road traffic research, *OR Oper. Res. Q.* 4 (1953) 72–73.

[18] L. Yong, Nonlinear complementarity problem and solution methods, in *Artificial Intelligence and Computational Intelligence: Proceedings-ICAICI Part I*, F. L. Wang, H. Deng and L. Jingshen (eds) Springer-Verlag, 2010, pp. 461–469.

Manuscript received 16 May 2014
revised 1 June 2015
accepted for publication 5 June 2015

Carlos Andrés Arias
Deparment of Mathematics, Universidad del Cauca, Popayán Colombia
E-mail address: carlosarias@unicauca.edu.co

Héctor Jairo Martínez
Deparment of Mathematics, Universidad del Valle, Cali, Colombia
E-mail address: hector.martinez@correounivalle.edu.co

Rosana Pérez
Deparment of Mathematics, Universidad del Cauca, Popayán Colombia
E-mail address: rosana@unicauca.edu.co