Numerical Algorithms (2024) 95:551–574
https://doi.org/10.1007/s11075-023-01581-2

ORIGINAL PAPER

A nonsmooth Newton method for solving the generalized complementarity problem

Hevert Vivas[1] · **Rosana Pérez**[1] · **Carlos A. Arias**[1]

Received: 10 August 2022 / Accepted: 16 May 2023 / Published online: 1 July 2023

Abstract
In this paper, we present a new nonsmooth Newton-type algorithm for solving the generalized complementarity problem based on its reformulation as a system of nonlinear equations using a one-parametric family of complementarity functions. We demonstrate, under suitable hypotheses, that this algorithm converges locally and q-quadratically. In addition, we show numerical experiments that allow us to see the good performance of the proposed algorithm.

Keywords Nonsmooth Newton · Generalized complementarity · q-quadratically · Algorithm · Complementarity functions

Mathematics Subject Classification (2010) 90C33 · 90C30 · 65K10 · 65H10

1 Introduction

Let $F\colon \mathbb{R}^n \to \mathbb{R}^n$, $(F_1(\boldsymbol{x}), \ldots, F_n(\boldsymbol{x}))$ and $G\colon \mathbb{R}^n \to \mathbb{R}^n$, $(G_1(\boldsymbol{x}), \ldots, G_n(\boldsymbol{x}))$ be continuously differentiable functions. The generalized complementarity problem (GCP(F, G)) is to find a solution of the following system of equations and inequalities

$$F_i(\boldsymbol{x}) \geq 0, \qquad G_i(\boldsymbol{x}) \geq 0, \qquad F_i(\boldsymbol{x})G_i(\boldsymbol{x}) = 0, \quad \forall i = 1, \ldots, n. \tag{1}$$

Hevert Vivas, Rosana Pérez, and Carlos A. Arias contributed equally to this work.

✉ Hevert Vivas
hevivas@unicauca.edu.co

Rosana Pérez
rosana@unicauca.edu.co

Carlos A. Arias
carlosarias@unicauca.edu.co

[1] Department of Mathematics, Universidad del Cauca, Calle 5 No. 4-70, 190003, Cauca Popayán, Colombia

Particular cases of GCP(F, G), of great interest due to its numerous applications, are the *linear complementarity problem* ($F(\boldsymbol{x}) = M\boldsymbol{x} + \boldsymbol{q}$, $M \in \mathbb{R}^{n\times n}$ and $G(\boldsymbol{x}) = \boldsymbol{x}$) [1–5] *nonlinear complementarity problem* ($G(\boldsymbol{x}) = \boldsymbol{x}$) [2–4] and *implicit complementarity problem* ($G(\boldsymbol{x}) = \boldsymbol{x} - E(\boldsymbol{x})$, with E a continuously differentiable function [5–8].

As in the particular cases mentioned previously, the GCP(F, G) can be reformulated as the following system of nonlinear equations,

$$\Phi(\boldsymbol{x}) = (\varphi(F_1(\boldsymbol{x}), G_1(\boldsymbol{x})),\ \ldots,\ \varphi(F_n(\boldsymbol{x}), G_n(\boldsymbol{x}))^T = 0, \tag{2}$$

where $\Phi\colon \mathbb{R}^n \to \mathbb{R}^n$ and $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$. The last one is called *complementarity function* and satisfies that $\varphi(a, b) = 0 \Leftrightarrow a \geq 0,\ b \geq 0,\ ab = 0$. From this equivalence, we conclude two facts. First, the trace of φ by the intersection with the xy-plane is the curve formed by the non negative x and y semi-axes which is nonsmooth, and therefore, φ and Φ are nonsmooth functions [9]. Second, $\boldsymbol{x}^*$ solves GCP(F, G) if only if $\Phi(\boldsymbol{x}^*) = 0$ which guarantees that solving the generalized complementarity problem is equivalent to solve its reformulation.

Alternatively, if $\Psi\colon \mathbb{R}^n \to \mathbb{R}$ denotes the natural merit function defined by $\Psi(\boldsymbol{x}) = \frac{1}{2}\Phi(\boldsymbol{x})^T\Phi(\boldsymbol{x})$, then we may rewrite the generalized complementarity problem as the following unconstrained minimization problem

$$\begin{aligned} &Minimize \quad \Psi(\boldsymbol{x}). \\ &\boldsymbol{x} \in \mathbb{R}^n \end{aligned} \tag{3}$$

Observe that any solution to (2) is a global minimizer of Ψ in $\mathbb{R}^n$. Reciprocally, any local solution $\boldsymbol{x}$ of (3) such that $\Psi(\boldsymbol{x}) = 0$ is a solution of (2).

Nonsmooth Newton-type methods are popular ones to solve nonlinear complementarity problems [2, 10–12]. They use the concept of *generalized Jacobian* [13]. Some of them have been extended to generalized complementarity problems [14, 15].

In this paper, we consider the reformulation (2) with $\varphi = \varphi_\lambda$, the one-parametric family of complementarity functions introduced in [12], defined by

$$\varphi_\lambda(a, b) = \sqrt{(a-b)^2 + \lambda ab} - a - b\,, \qquad \lambda \in (0, 4)\,. \tag{4}$$

Thus, we may rewrite the nonlinear system of equations (2) as $\Phi_\lambda(\boldsymbol{x}) = 0$ and the minimization problem (3) as

$$\begin{aligned} &Minimize \quad \Psi_\lambda(\boldsymbol{x}) = \tfrac{1}{2}\,\|\Phi_\lambda(\boldsymbol{x})\|_2^2\,, \\ &\boldsymbol{x} \in \mathbb{R}^n \end{aligned} \tag{5}$$

whose objective function is continuously differentiable [12].

We chose the family φ_λ for two reasons. First, because its relationship with the two probably most prominent complementarity functions: *Fischer* function [2, 16] and the *Minimum* function [17] defined by $\varphi(a, b) = \sqrt{a^2 + b^2} - a - b$ and $\varphi(a, b) =$

$\min\{a, b\}$, respectively. Observe that, for $\lambda = 2$, φ_λ reduces to the *Fischer* function and, when λ tends to 0, φ_λ converges to a multiple of the *Minimum* function.

The second reason is that, as far as we know, the family has not been used in connection with the generalized complementarity problem, only the particular case $\lambda = 2$ (Fischer function) was used in [15] to propose a generalized Newton-type method to solve the GCP (F, G), and the minimum function was used in [14] to analyze a local convergence of Levenberg-Marquardt type-method for the GCP(F, G).

In this paper, we analyze and extend to generalized complementarity problems, the properties of the operator Φ_λ investigated in [12] for the particular case of nonlinear complementarity, and we present a new nonsmooth Newton-type algorithm for solving the GCP (F, G) based on its reformulation as a system of nonlinear equations using a one-parametric family of complementarity functions. We demonstrate, under suitable hypotheses, that this algorithm converges locally and q-quadratically. We show numerical tests that allow us to see the good performance of the new algorithm.

We organize this article as follows. In Section 2, we recall some definitions and results from nonsmooth analysis which we use in the document. We also extend some definitions of regularity to generalized complementarity. In Section 3, we analyze properties of the operator Φ_λ as extensions of those investigated in [12] for nonlinear complementarity. In Section 4, we present the algorithm and convergence results. In Section 5, we present the analysis of the numerical performance of the algorithm proposed. Finally, In Section 6, we present some conclusions.

2 Preliminaries

We recall some definitions and results from nonsmooth analysis which we use in this work.

Definition 1 [18]. Let $K : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitzian function and let D_K be the set where K is differentiable. For all $\boldsymbol{x} \in \mathbb{R}^n$, the set given by

$$\partial_B K(\boldsymbol{x}) = \left\{ \lim_{k\to\infty} K'(\boldsymbol{x}_k) \in \mathbb{R}^{n\times n} : \lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x},\ \boldsymbol{x}_k \in D_K \right\} \tag{6}$$

is known as the generalized B-Jacobian of K at $\boldsymbol{x}$.

Definition 2 The convex hull of $\partial_B K(\boldsymbol{x})$ is called the generalized Jacobian of K at $\boldsymbol{x}$, denoted $\partial K(\boldsymbol{x})$.

Usually, the set $\partial K(\boldsymbol{x})$ is difficult to compute. An alternative is the following overestimation [13],

$$\partial K(\boldsymbol{x})^T \subseteq \partial K_1(\boldsymbol{x}) \times \cdots \times \partial K_n(\boldsymbol{x}), \tag{7}$$

where the right side, for short $\partial_C K(\boldsymbol{x})$ is called *C-subdifferential* of K at $\boldsymbol{x}$ [19], and it is the set of matrices in $\mathbb{R}^{n\times n}$ whose ith column is the generalized gradient of the ith component of the function K.

Definition 3 [20]. A locally Lipschitz continuous and directionally differentiable function $K : \mathbb{R}^n \to \mathbb{R}^n$ is called semismooth at a point $\boldsymbol{x} \in \mathbb{R}^n$ if

$$H\boldsymbol{d} - K'(\boldsymbol{x}; \boldsymbol{d}) = o(\|\boldsymbol{d}\|)$$

holds for every $\boldsymbol{d} \to \boldsymbol{0}$ and every $H \in \partial K(\boldsymbol{x} + \boldsymbol{d})$ and, in addition

$$\lim_{H \in \partial K(\boldsymbol{x}+t\boldsymbol{h}'),\, \boldsymbol{h}' \to \boldsymbol{h},\, t \to 0^+} \left\{ H\boldsymbol{h}' \right\} \tag{8}$$

exists for any $\boldsymbol{h} \in \mathbb{R}^n$, and strongly semismooth at $\boldsymbol{x} \in \mathbb{R}^n$ if

$$H\boldsymbol{d} - K'(\boldsymbol{x}; \boldsymbol{d}) = O(\|\boldsymbol{d}\|^2)$$

holds for every $\boldsymbol{d} \to \boldsymbol{0}$, and every $H \in \partial K(\boldsymbol{x} + \boldsymbol{d})$. Here, $K'(\boldsymbol{x}; \boldsymbol{d})$ denotes the usual directional derivative of K at $\boldsymbol{x}$ in the direction $\boldsymbol{d}$.

Definition 4 [12]. The function $K : \mathbb{R}^n \to \mathbb{R}^n$ is class LC^1, if it is differentiable and its derivative is locally Lipschitz continuous.

Definition 5 [10] The function $K : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is B-differentiable at $\boldsymbol{x}$, if it is directionally differentiable at $\boldsymbol{x}$ and

$$K(\boldsymbol{x} + \boldsymbol{h}) - K(\boldsymbol{x}) - K'(\boldsymbol{x}, \boldsymbol{h}) = o(\|\boldsymbol{h}\|). \tag{9}$$

Moreover, K is B-differentiable of degree 2 en $\boldsymbol{x}$ if and only if

$$K(\boldsymbol{x} + \boldsymbol{h}) = K(\boldsymbol{x}) + K'(\boldsymbol{x}; \boldsymbol{h}) + O(\|\boldsymbol{h}\|^2). \tag{10}$$

Furthermore, the directional derivative $K'(\cdot, \cdot)$ is semicontinuous of degree 2 in $\boldsymbol{x}$ [10] if there exists a constant L and a neighborhood N of $\boldsymbol{x}$ such that for all $\boldsymbol{x} + \boldsymbol{h} \in N$,

$$\left\| K'(\boldsymbol{x} + \boldsymbol{h}, \boldsymbol{h}) - K'(\boldsymbol{x}, \boldsymbol{h}) \right\| \leq L \|\boldsymbol{h}\|^2. \tag{11}$$

The following two lemmas will be useful in determining the rate of convergence of our algorithm.

Lemma 1 (**Lemma 2.2** [10]) *Suppose that $K : \mathbb{R}^n \to \mathbb{R}^m$ is directionally differentiable at a neighborhood of $\boldsymbol{x}$. The following statements are equivalent:*

(1) *K is semismooth at $\boldsymbol{x}$.*
(2) *$K'(\cdot, \cdot)$ is semicontinuous at $\boldsymbol{x}$.*
(3) *for any $H \in \partial K(\boldsymbol{x} + \boldsymbol{h})$, $\boldsymbol{h} \longrightarrow \boldsymbol{0}$, $H\boldsymbol{h} - K'(\boldsymbol{x}; \boldsymbol{h}) = o(\|\boldsymbol{h}\|)$.*

Lemma 2 (**Lemma 2.3** [10]) *Suppose that $K : \mathbb{R}^n \to \mathbb{R}^m$ is directionally differentiable at a neighborhood of $\boldsymbol{x}$. The following statements are equivalent:*

(1) *K is semicontinuous of degree 2 at $\boldsymbol{x}$.*

(2) *for any* $H \in \partial K(\boldsymbol{x}+\boldsymbol{h})$, $\boldsymbol{h} \longrightarrow \boldsymbol{0}$, $H\boldsymbol{h} - K'(\boldsymbol{x};\boldsymbol{h}) = O(\|\boldsymbol{h}\|^2)$.

If (1) or (2) holds, then K is B-differentiable of degree 2 *at* $\boldsymbol{x}$.

Next, we define the concepts of BD-regularity and R-regularity in generalized complementarity context. For this, we consider a solution $\boldsymbol{x}_*$ of GCP(F, G) and the following index sets, $\alpha_* = \{i \in I : F_i(\boldsymbol{x}_*) > 0 = G_i(\boldsymbol{x}_*)\}$, $\beta_* = \{i \in I : F_i(\boldsymbol{x}_*) = 0 = G_i(\boldsymbol{x}_*)\}$ and $\gamma_* = \{i \in I : F_i(\boldsymbol{x}_*) = 0 < G_i(\boldsymbol{x}_*)\}$.

Definition 6 [4]. Let $\boldsymbol{x}_*$ be a solution of GCP(F, G).

1. If all matrices $H \in \partial_B \Phi(\boldsymbol{x}_*)$ are nonsingular, $\boldsymbol{x}_*$ is called a BD-regular solution.
2. Let $F'(\boldsymbol{x}_*)$ be nonsingular and $K = G'(\boldsymbol{x}_*)F'(\boldsymbol{x}_*)^{-1}$. If the submatrix[1] $K_{\alpha\alpha}$ is nonsingular and the Schur-complement of K,

$$K_{\beta\beta} - K_{\beta\alpha}K_{\alpha\alpha}^{-1}K_{\alpha\beta} \in \mathbb{R}^{|\beta|\times|\beta|}$$

is a P-matrix,[2] $\boldsymbol{x}_*$ is called an R-regular solution.

Observe that for the particular case $G(\boldsymbol{x}) = \boldsymbol{x}$, the **Definition** 6 is reduced to the one of BD-regularity and R-regularity for nonlinear complementarity.

Following [12], we will use the notation $f_\lambda(a, b)$ for the first term on the right side of (4).

Lemma 3 [12]. *Let* $f_\lambda : \mathbb{R}^2 \to \mathbb{R}$ *defined by*

$$f_\lambda(a, b) = \sqrt{(a-b)^2 + \lambda ab}, \qquad \lambda \in (0, 4). \tag{12}$$

There exists a constant $c_\lambda \in (0, 2)$ *such that* $\|\nabla f_\lambda(a, b)\|^2 \leq c_\lambda$, *for all nonzero vector* $(a, b) \in \mathbb{R}^2$.

Finally, we next present a characterization of the class of P-matrices for nonlinear complementarity, which will be useful later.

Proposition 1 (**Proposition 2.7** [12]) *A matrix of the form* $D_a + D_b N$ *is nonsingular for all positive (negative) semidefinite diagonal matrices* $D_a, D_b \in \mathbb{R}^{n\times n}$ *such that* $D_a + D_b$ *is positive (negative) definite if and only if* $N \in \mathbb{R}^{n\times n}$ *is a P-matrix.*

3 The operator Φ_λ

As we mentioned earlier, properties of the operator Φ_λ were investigated in detail in [12] for the reformulation of nonlinear complementarity problem (i.e., when $G(\boldsymbol{x}) = \boldsymbol{x}$.). In this section, we verify that these properties are easily extend to generalized complementarity problem.

The following lemma and corollary give upper bounds for the partial derivatives of the function f_λ, defined by (12), which will be used in later results.

[1] Let $A = (a_{ij}) \in \mathbb{R}^{m\times n}$. $A_{\alpha\beta}$ is the one with elements a_{ij} such that $i \in \alpha$ and $j \in \beta$.

[2] A matrix $M \in \mathbb{R}^{n\times n}$ is a P-matrix if for every nonzero vector z there is an index $j \in \{1, \ldots, n\}$ such that $z_j[Mz]_j > 0$.

Lemma 4 *The partial derivatives of the function* f_λ *are bounded from above by* 1, *for all nonzero vector* $(a, b) \in \mathbb{R}^2$.

Proof We must prove that for all nonzero vector $(a, b) \in \mathbb{R}^2$, the following inequalities are satisfied:

$$\frac{\partial f_\lambda(a,b)}{\partial a} \leq 1 \qquad \text{and} \qquad \frac{\partial f_\lambda(a,b)}{\partial b} \leq 1. \tag{13}$$

The first inequality was proved in [9]. Analogously, we prove the second below. From the inequality $\lambda a^2(\lambda - 4) \leq 0$, adding $4(a-b)^2 + 4\lambda ab$ and after some algebraic manipulations, we obtain the inequalities

$$4(a-b)^2 + 4\lambda ab + \lambda^2 a^2 - 4\lambda a^2 \leq 4 f_\lambda^2(a,b)$$

$$4(a-b)^2 - 4\lambda a(a-b) + \lambda^2 a^2 \leq 4 f_\lambda^2(a,b)$$

$$[-2(a-b) + \lambda a]^2 \leq 4 f_\lambda^2(a,b)$$

which implies that $|-2(a-b) + \lambda a| \leq 2 f_\lambda(a,b) > 0$, then

$$\left| \frac{-2(a-b) + \lambda a}{2 f_\lambda(a,b)} \right| \leq 1,$$

thus

$$-1 \leq \frac{-2(a-b) + \lambda a}{2 f_\lambda(a,b)} \leq 1.$$

Therefore, we obtain the second inequality in (13). □

Corollary 1 *The partial derivatives of the function* f_λ *satisfy the inequality*

$$\frac{\partial f_\lambda(a,b)}{\partial a} + \frac{\partial f_\lambda(a,b)}{\partial b} < 2,$$

for all nonzero vector $(a, b) \in \mathbb{R}^2$.

Proof From **Lemma** 4, $\frac{\partial f_\lambda(a,b)}{\partial a} + \frac{\partial f_\lambda(a,b)}{\partial b} \leq 2$. We assume that the equality holds. By **Lemma** 4, we have that $\frac{\partial f_\lambda(a,b)}{\partial a} = 1$ and $\frac{\partial f_\lambda(a,b)}{\partial b} = 1$, then $\|\nabla f_\lambda(a,b)\|^2 = 2$ which contradicts **Lemma** 3. Therefore, $\frac{\partial f_\lambda(a,b)}{\partial a} + \frac{\partial f_\lambda(a,b)}{\partial b} < 2$. □

The following lemma gives a compact expression for the gradient of the component functions of Φ_λ, where φ_λ is differentiable. This expression will be useful to characterize the generalized Jacobian matrices of Φ_λ in $\boldsymbol{x}$.

Lemma 5 *Let* $\Phi_{\lambda,i}(\boldsymbol{x}) = \varphi_\lambda(F_i(\boldsymbol{x}), G_i(\boldsymbol{x}))$ *for* $i \in \{1, 2 \ldots, n\}$, *such that* $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) \neq (0, 0)$. *Then the gradient of* $\Phi_{\lambda,i}(\boldsymbol{x})$ *at* $\boldsymbol{x}$ *is given by*

$$\nabla \Phi_{\lambda,i}(\boldsymbol{x}) = (a_i(\boldsymbol{x}) - 1)\nabla F_i(\boldsymbol{x}) + (b_i(\boldsymbol{x}) - 1)\nabla G_i(\boldsymbol{x}), \tag{14}$$

where

$$a_i(\boldsymbol{x}) = \frac{2\left(F_i(\boldsymbol{x}) - G_i(\boldsymbol{x})\right) + \lambda G_i(\boldsymbol{x})}{2\sqrt{\left(F_i(\boldsymbol{x}) - G_i(\boldsymbol{x})\right)^2 + \lambda F_i(\boldsymbol{x}) G_i(\boldsymbol{x})}} \tag{15}$$

$$b_i(\boldsymbol{x}) = \frac{-2\left(F_i(\boldsymbol{x}) - G_i(\boldsymbol{x})\right) + \lambda F_i(\boldsymbol{x})}{2\sqrt{\left(F_i(\boldsymbol{x}) - G_i(\boldsymbol{x})\right)^2 + \lambda F_i(\boldsymbol{x}) G_i(\boldsymbol{x})}}. \tag{16}$$

Proof Let $\boldsymbol{x} \in \mathbb{R}^n$ such that $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) \neq (0,0)$. Therefore, Φ_i and φ_λ are differentiable at $\boldsymbol{x}$. Moreover, $\nabla \Phi_{\lambda,i}(\boldsymbol{x}) = \nabla \varphi_\lambda(F_i(\boldsymbol{x}), G_i(\boldsymbol{x}))$. After some algebraic calculations using the Chain rule, we have that

$$\nabla \Phi_{\lambda,i}(\boldsymbol{x}) = \nabla \varphi_\lambda(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) = \begin{bmatrix} (a_i(\boldsymbol{x}) - 1)\dfrac{\partial F_i(\boldsymbol{x})}{\partial x_1} + (b_i(\boldsymbol{x}) - 1)\dfrac{\partial G_i(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ (a_i(\boldsymbol{x}) - 1)\dfrac{\partial F_i(\boldsymbol{x})}{\partial x_n} + (b_i(\boldsymbol{x}) - 1)\dfrac{\partial G_i(\boldsymbol{x})}{\partial x_n} \end{bmatrix},$$

from which it follows immediately that

$$\nabla \Phi_{\lambda,i}(\boldsymbol{x}) = (a_i(\boldsymbol{x}) - 1)\nabla F_i(\boldsymbol{x}) + (b_i(\boldsymbol{x}) - 1)\nabla G_i(\boldsymbol{x}),$$

where $a_i(\boldsymbol{x})$ and $b_i(\boldsymbol{x})$ are given by (15) and (16), respectively. □

The following result extends the **Proposition 2.5** in [12] to generalized complementarity. It gives an overestimation of the generalized Jacobian of Φ_λ at $\boldsymbol{x}$.

Proposition 2 *For all $\boldsymbol{x} \in \mathbb{R}^n$, we have*

$$\partial \Phi_\lambda(\boldsymbol{x}) \subseteq D_a(\boldsymbol{x}) F'(\boldsymbol{x}) + D_b(\boldsymbol{x}) G'(\boldsymbol{x}),$$

where $D_a(\boldsymbol{x}) = diag\,(a_1(\boldsymbol{x}) - 1, \ldots, a_n(\boldsymbol{x}) - 1)$ and $D_b(\boldsymbol{x}) = diag(b_1(\boldsymbol{x}) - 1, \ldots, b_n(\boldsymbol{x}) - 1)$ are diagonal matrices, with $a_i(\boldsymbol{x})$ and $b_i(\boldsymbol{x})$ given by (15) and (16), if $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) \neq (0,0)$, and by

$$a_i(\boldsymbol{x}) = \xi_i, \quad b_i(\boldsymbol{x}) = \chi_i, \quad \text{for all } (\xi_i, \chi_i) \in \mathbb{R}^2 \text{ such that } \|(\xi_i, \chi_i)\| \leq \sqrt{c_\lambda},$$

if $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) = (0,0)$, where c_λ is the constant from Lemma 3.

Proof From (7) and by Lipschitz continuity of Φ_λ, we have

$$\partial \Phi_\lambda(\boldsymbol{x})^T \subseteq \partial \Phi_{\lambda,1}(\boldsymbol{x}) \times \cdots \times \partial \Phi_{\lambda,n}(\boldsymbol{x}),$$

where $\Phi_{\lambda,i}$ is the ith component function of Φ_λ and $\partial \Phi_{\lambda,i}(\boldsymbol{x})$ denotes the generalized gradient of $\Phi_{\lambda,i}$ at $\boldsymbol{x}$.

If $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) \neq (0,0)$, the function $\Phi_{\lambda,i}$ is continuously differentiable. Therefore the generalized gradient $\partial \Phi_{\lambda,i}(\boldsymbol{x})$ is the set whose only element is the gradient of $\Phi_{\lambda,i}$ at $\boldsymbol{x}$, which by **Lemma** 5 is given by

$$\partial \Phi_{\lambda,i}(\boldsymbol{x}) = \{(a_i(\boldsymbol{x}) - 1)\nabla F_i(\boldsymbol{x}) + (b_i(\boldsymbol{x}) - 1)\nabla G_i(\boldsymbol{x}) \nabla G_i(\boldsymbol{x})\}, \tag{17}$$

where $a_i(\boldsymbol{x})$ and $b_i(\boldsymbol{x})$ are given by (15) and (16), respectively.

If $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) = (0, 0)$ then $\Phi_{\lambda,i}$ is not differentiable at $\boldsymbol{x}$ and therefore, for $H \in \partial_B \Phi_{\lambda,i}(\boldsymbol{x})$, we have

$$H = \lim_{\boldsymbol{x}_k \to \boldsymbol{x}} \nabla \Phi_{\lambda,i}(\boldsymbol{x}_k), \tag{18}$$

where $\{\boldsymbol{x}_k\}$ is a sequence of points in $D_{\Phi_{\lambda,i}}$ such that $\boldsymbol{x}_k \to \boldsymbol{x}$, when $k \to \infty$. Thus, by **Lemma** 5, for each $\boldsymbol{x}_k \in D_\Phi$, the ith row of gradient from $\Phi_{\lambda,i}$ at $\boldsymbol{x}_k$ is given by

$$\nabla \Phi_{\lambda,i}(\boldsymbol{x}_k) = (a_i(\boldsymbol{x}_k) - 1)\nabla F_i(\boldsymbol{x}_k)^T + (b_i(\boldsymbol{x}_k) - 1)\nabla G_i(\boldsymbol{x}_k)^T. \tag{19}$$

Since the limit in (19) exists, $\nabla \Phi_{\lambda,i}$, F and G are continuously differentiable at $\boldsymbol{x}_k$ and from (18), we have

$$\begin{aligned} H &= \lim_{\boldsymbol{x}_k \to \boldsymbol{x}} \left[(a_i(\boldsymbol{x}_k) - 1)\, \nabla F_i(\boldsymbol{x}_k) + (b_i(\boldsymbol{x}_k) - 1)\, \nabla G_i(\boldsymbol{x}_k)\right] \\ &= (\xi_i - 1)\nabla F_i(\boldsymbol{x}) + (\chi_i - 1)\nabla G_i(\boldsymbol{x}), \end{aligned}$$

where $\xi_i = \lim_{\boldsymbol{x}_k \to \boldsymbol{x}} a_i(\boldsymbol{x}_k)$ and $\chi_i = \lim_{\boldsymbol{x}_k \to \boldsymbol{x}} a_i(\boldsymbol{x}_k)$. On the other hand, for each k,

$$(a_i(\boldsymbol{x}_k), b_i(\boldsymbol{x}_k))^T = \nabla f_\lambda((F_i(\boldsymbol{x}_k), (G_i(\boldsymbol{x}_k)), \tag{20}$$

where f_λ, is defined by (12). By **Lemma** 4, for each $i = 1, \ldots, n$ and for each $k \in N$, $\|\nabla f_\lambda((F_i(\boldsymbol{x}_k), (G_i(\boldsymbol{x}_k))\|^2 \leq c_\lambda$; so

$$\lim_{k \to \infty} \|\nabla f_\lambda((F_i(\boldsymbol{x}_k), (G_i(\boldsymbol{x}_k))\|^2 \leq c_\lambda.$$

Therefore, $\|(\xi_i, \chi_i)\| \leq \sqrt{c_\lambda}$. For the above, the statement of the lemma follows easily. □

The following result generalizes the **Proposition 2.6** in [12] to the case where G is any continuously differentiable function. Furthermore, it guarantees that the diagonal elements $a_i(\boldsymbol{x}) - 1$ and $b_i(\boldsymbol{x}) - 1$ defined in **Proposition** 2 are nonpositive (namely for those indices for which $(F_i(\boldsymbol{x}), G_i(\boldsymbol{x})) = (0, 0)$).

Proposition 3 *Any $H \in \partial \Phi_\lambda(\boldsymbol{x})$ can be written in the form*

$$H = (D_a - I)F'(\boldsymbol{x}) + (D_b - I)G'(\boldsymbol{x}), \tag{21}$$

where $(D_a - I)$ and $(D_b - I)$ are negative semidefinite diagonal matrices such that $D_a + D_b - 2I$ is negative definite.

Proof It is analogous to the proof of **Proposition 2.6** in [12], taking into account that we must use **Proposition** 2. □

The next proposition gives a new characterization of the class of P-matrices in the context of generalized complementarity. It reduces to the one presented in [12] for nonlinear complementarity when G is the identity function.

Proposition 4 *Let D_a and D_b be positive (negative) semidefinite diagonal matrices such that $D_a + D_b$ is positive (negative) definite, and M, $N \in \mathbb{R}^{n\times n}$, with M nonsingular. The matrix $D_aM + D_bN$ is nonsingular if and only if $NM^{-1} \in \mathbb{R}^{n\times n}$ is a P-matrix.*

Proof Clearly,

$$D_aM + D_bN = \left(D_a + D_bNM^{-1}\right) M. \tag{22}$$

$\Rightarrow$) We assume that D_aM+D_bN is nonsingular. Since M is nonsingular (hypotheses), then from (22), we have that $D_a + D_bNM^{-1}$ is also nonsingular and by **proposition** 1, NM^{-1} is a P-matrix.
$\Leftarrow$) If NM^{-1} is a P-matrix, **proposition** 1 guarantees that the matrix $D_a + D_bNM^{-1}$ is nonsingular. Using this fact in (22) and recalling that M es nonsingular, we conclude that $D_aM + D_bN$ is also nonsingular.

□

The following theorem gives a sufficient condition to guarantee the nonsingularity of the elements of the generalized Jacobian of Φ_λ in a solution $\boldsymbol{x}_*$ of the GCP(F, G).

Theorem 1 *If $\boldsymbol{x}_*$ is an R-regular solution of the GCP(F, G) then all the matrices in the generalized Jacobian $\partial\Phi_\lambda(\boldsymbol{x}_*)$ are nonsingular.*

Proof Let $H \in \partial\Phi_\lambda(\boldsymbol{x}_*)$ be fixed but arbitrary. By **proposition** 2, there are diagonal matrices $D_a = D_a(\boldsymbol{x}_*)$ and $D_b = D_b(\boldsymbol{x}_*)$ such that

$$H = D_aF'(\boldsymbol{x}_*) + D_bG'(\boldsymbol{x}_*). \tag{23}$$

Since $\boldsymbol{x}_*$ is an R-regular solution of the GCP(F, G), we have that $F'(\boldsymbol{x}_*)$ is nonsingular. Thus, from (23), we have

$$\overline{H} = HF'(\boldsymbol{x}_*)^{-1} = D_a + D_bG'(\boldsymbol{x}_*)F'(\boldsymbol{x}_*)^{-1} = D_a + D_bK,$$

where $K = G'(\boldsymbol{x}_*)F'(\boldsymbol{x}_*)^{-1}$. From here, the proof can be carried out in essentially the same way as the one for **Theorem 2.8** in [12]. □

Corollary 2 *If $\boldsymbol{x}_*$ is an R-regular solution of the GCP(F, G) then $\boldsymbol{x}_*$ is also a BD-regular solution.*

Proof Let $\boldsymbol{x}_*$ be an R-regular solution of the GCP(F, G). Then all matrices in $\partial\Phi_\lambda(\boldsymbol{x}_*)$ are nonsingular. Since $\partial_B\Phi_\lambda(\boldsymbol{x}_*) \subseteq \partial\Phi_\lambda(\boldsymbol{x}_*)$, we have that if $H \in \partial_B\Phi_\lambda(\boldsymbol{x}_*)$ then H is nonsingular. Therefore $\boldsymbol{x}_*$ is a BD-regular solution of the GCP(F, G). □

The following theorem guarantees the semismoothness of Φ_λ and gives a sufficient condition for its strong semismoothness.

Theorem 2 *The following propositions are fulfilled:*

1. *Φ_λ is semismooth.*
2. *If F and G are class LC^1 then Φ_λ is strongly semismooth.*

Proof The proof is analogous to one given in [12] in the nonlinear complementarity context, and it is an immediate consequence of the fact that φ_λ is strongly semismooth. □

Finally, we present the following result which will be very useful in the next section.

Theorem 3 *The function Ψ_λ is continuously differentiable with $\nabla\Psi_\lambda(\boldsymbol{x}) = H^T\Phi_\lambda(\boldsymbol{x})$, for any $H \in \partial\Phi_\lambda(\boldsymbol{x})$.*

Proof Using **proposition** 3 and properties of the family φ_λ [12], the proof is practically the same as the one for **proposition 3.4** in [21]. □

4 Algorithm and convergence results

In this section, we present a new global nonsmooth Newton-type algorithm to solve the GCP(F, G) and some convergence results of this algorithm.

Algorithm 1

Require: $\boldsymbol{x}_0$, $\lambda \in (0, 4)$, $\rho > 0$, $\sigma \in (0, 1/2)$, $p > 2$, $\epsilon \geq 0$.
Ensure: Approach to the solution of the system (2).
1: **while** $\|\nabla\Psi_\lambda(\boldsymbol{x}_k)\| \leq \epsilon$ **do**
2: Select an element H_k in $\partial_B\Phi_\lambda(\boldsymbol{x}_k)$ and find $\boldsymbol{d}_k$ such that

$$H_k\boldsymbol{d}_k = -\Phi_\lambda(\boldsymbol{x}_k). \tag{24}$$

3: **if** (24) is not solvable or

$$\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k > -\rho\|d_k\|^p \tag{25}$$

then
4: $\boldsymbol{d}_k = -\nabla\Psi_\lambda(\boldsymbol{x}_k)$.
5: **end if**
6: Compute $t_k = \max\{2^{-i_k} : i_k = 0, 1, \ldots\}$ such that

$$\Psi_\lambda(\boldsymbol{x}_k + t_k\boldsymbol{d}_k) \leq \Psi_\lambda(\boldsymbol{x}_k) + \sigma t_k\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k \tag{26}$$

7: $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + t_k\boldsymbol{d}_k$.
8: **end while**

Remark 1 The algorithm proposed uses in each iteration the matrix H defined in (21) which in turn uses the Jacobian matrices of F and G. It can be seen as an extension of the algorithm proposed in [15] to all members of the family of complementary functions φ_λ defined in (12).

Remark 2 The **Algorithm 1** differs from the proposed in [15], not only in the complementarity function but also in the linear search. Our algorithm uses the one-parametric family of complementarity functions defined in (4) and its linear search guarantees that the step satisfies the Armijo condition [22] while the other algorithm uses the

Fischer function and its linear search guarantees that the step satisfies the *Goldstein* conditions [22].

Remark 3 From the previous *Remark*, **Algorithm 1** can be seen as a generalization of algorithm proposed in [15] to the family of complementarity functions (4), and comparing their linear search, **Algorithm 1** is less restrictive than the other.

For **Algorithm 1**, we develop its global convergence theory. For this purpose, we present two lemmas that will be useful in proving the convergence theorems of **Algorithm 1** and in which a characterization of the accumulation points of the sequence $\{\boldsymbol{x}_k\}$ generated by **Algorithm 1** is given.

Theorem 4 *Each of the accumulation points of the sequence* $\{\boldsymbol{x}_k\}$ *generated by* ***Algorithm 1*** *is a stationary point of* Ψ_λ.

Proof Let $\boldsymbol{x}_*$ be an accumulation point of the sequence $\{\boldsymbol{x}_k\}$ generated by the **Algorithm 1**. If for an infinite set of indices J, the direction given by Step 4 of the algorithm ($\boldsymbol{d}_k = -\nabla\Psi_\lambda(\boldsymbol{x}_k)$, for all $k \in J$) then any limit point of the subsequence $\{\boldsymbol{x}_k\}_J$ is a stationary point of Ψ_λ [23].

We assume, without losing generality, that $\boldsymbol{d}_k$ is given by (24). We will prove that for all k, there are positive constants m and M such that

$$m \leq \|\boldsymbol{d}_k\| \leq M. \tag{27}$$

Let k be any index of the sequence generated by **Algorithm 1**. Since $\boldsymbol{d}_k$ is given by (24), we have $H_k\boldsymbol{d}_k = -\Phi_\lambda(\boldsymbol{x}_k)$ then

$$\|\Phi_\lambda(\boldsymbol{x}_k)\| \leq \|H_k\|\,\|\boldsymbol{d}_k\|, \tag{28}$$

for a vector norm $\|\cdot\|$. Clearly $\|H_k\| \neq 0$, otherwise, by **Theorem** 3 we would have that $\nabla\Psi_\lambda(\boldsymbol{x}_k) = 0$; that is, $\boldsymbol{x}_k$ would be a stationary point of Ψ_λ and **Algorithm 1** would stop. Thus, from (28), we have the inequality

$$\frac{\|\Phi_\lambda(\boldsymbol{x}_k)\|}{\|H_k\|} \leq \|\boldsymbol{d}_k\|. \tag{29}$$

We suppose that $\boldsymbol{x}_*$ is not a stationary point of $\Psi_\lambda(\nabla\Psi_\lambda(\boldsymbol{x}_*) \neq 0)$. If for an infinite set of indices J, $\{\|\boldsymbol{d}_k\|\}_J \to 0$, when $k \to \infty$ then since the generalized Jacobian is compact, H_k is bounded and from (28), we get, $\{\|\Phi_\lambda(\boldsymbol{x}_k)\|\}_J \to 0$

Now, by the continuity of Φ_λ and by taking the limit $k \to \infty$ in (29), we obtain $\Phi_\lambda(\boldsymbol{x}_*) = 0$ so $\nabla\Psi_\lambda(\boldsymbol{x}_*) = 0$ which contradicts that $\boldsymbol{x}_*$ is not a stationary point. Then $\{\|\boldsymbol{d}_k\|\}$ cannot converge to 0. Therefore there exists $m > 0$ such that $m \leq \|\boldsymbol{d}_k\|$, for all k.

On the other hand, if $\|\boldsymbol{d}_k\|$ were not bounded above then, since $p > 2$ and $\nabla\Psi_\lambda(\boldsymbol{x}_k)$ is bounded, we have

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

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

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

which contradicts (25). Then there exists $M > 0$ such that $\|\boldsymbol{d}_k\| \leq M$.

Finally, since (25) is satisfied at each iteration and Ψ_λ is continuously differentiable, $\{\Psi_\lambda(\boldsymbol{x}_{k+1}) - \Psi_\lambda(\boldsymbol{x}_x)\} \to 0$ when $k \to \infty$, which implies by (25), that $\left\{\sigma 2^{-i_k}\|\nabla\Psi_\lambda(\boldsymbol{x}_k)\|\boldsymbol{d}_k\right\} \to 0$, therefore

$$\left\{2^{-i_k}\|\nabla\Psi_\lambda(\boldsymbol{x}_k)\|\boldsymbol{d}_k\right\} \to 0. \tag{30}$$

Now, we want to show that 2^{-i_k} is bounded away from 0. Suppose the contrary. Then, subsequencing if necessary, we have that $\{2^{-i_k}\} \to 0$, so that at each iteration the stepsize is reduced at least once and (26) gives

$$\frac{\Psi_\lambda(\boldsymbol{x}_k + 2^{-(i_k-1)}\boldsymbol{d}_k) - \Psi_\lambda(\boldsymbol{x}_k)}{2^{-(i_k-1)}} > \sigma\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k. \tag{31}$$

From (27), we can assume that $\{\boldsymbol{d}_k\} \to \overline{\boldsymbol{d}} \neq 0$, and passing to the limit at (31), we obtain

$$\nabla\Psi_\lambda(\boldsymbol{x}_*)^T\overline{\boldsymbol{d}} \geq \sigma\nabla\Psi_\lambda(\boldsymbol{x}_*)^T\overline{\boldsymbol{d}} \tag{32}$$

in addition, by (25), we have

$$\nabla\Psi_\lambda(\boldsymbol{x}_*)^T\overline{\boldsymbol{d}} \leq -\rho\|\boldsymbol{d}_k\|^p < 0, \tag{33}$$

which contradicts (32) since $\sigma \in (0, 1/2)$. Thus, we can assume that the sequence $\{2^{-i_k}\}$ is bounded from 0. On the other hand, (25) and (30) imply that $\{\boldsymbol{d}_k\} \to 0$, which contradicts (27). Therefore, $\nabla\Psi_\lambda(\boldsymbol{x}_*) = 0$. That is, $\boldsymbol{x}_*$ is a stationary point of Ψ_λ. □

Theorem 5 *If $\boldsymbol{x}_*$ is an isolated accumulation point of the sequence $\{\boldsymbol{x}_k\}$ generated by **Algorithm 1**. Then the sequence converges to $\boldsymbol{x}_*$.*

Proof Let $\{\boldsymbol{x}_k\}$ be the sequence generated by **Algorithm 1** and $\boldsymbol{x}_*$ an isolated accumulation point of the sequence. By **Theorem** 4, $\boldsymbol{x}_*$ is a stationary point of the convex function Ψ_λ then $\boldsymbol{x}_*$ is an isolated global minimizer of Ψ_λ.

Let Ω be the set of accumulation points of $\{\boldsymbol{x}_k\}$. Then $\Omega \neq \emptyset$ since $\boldsymbol{x}_* \in \Omega$. we define

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

$\delta > 0$ since $\boldsymbol{x}_*$ is an isolated accumulation point. Now, if

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

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

$$K = \{k \in \mathbb{N} : dist(\boldsymbol{x}_k; \boldsymbol{x}_*) \leq \delta/4\}$$

(K is obviously not empty because $\boldsymbol{x}_*$ is a limit point of the sequence) then $\{\boldsymbol{x}_k\}_K \subset \overline{B}(\boldsymbol{x}_*; \delta/4)$.

Since all points of the subsequence $\{\boldsymbol{x}_k\}_K$ belong to the compact set $\overline{B}(\boldsymbol{x}_*; \delta/4)$ and all its limit points are also limit points of $\{\boldsymbol{x}_k\}$, we conclude that $\{\boldsymbol{x}_k\}_K$ converges to $\boldsymbol{x}_*$. Furthermore, because $\nabla\Psi_\lambda(\boldsymbol{x}_*) = 0$, the **Theorem** 4 guarantees that $\{\|\nabla\Psi_\lambda(\boldsymbol{x}_k)\|\}_K$ converges to zero, which by (25) implies that the sequence $\{\boldsymbol{d}_k\}$ converges to the zero vector. Thus, for all ϵ, there exists $k_1 \in \mathbb{N}$ such that if $k \in K$ and $k \geq k_1 \geq \overline{k}$, then $\|\boldsymbol{d}_k\| \leq \epsilon$. Particularly, for $\epsilon = \delta/4$ we have $\|\boldsymbol{d}_k\| \leq \delta/4$.

Let $k_2 \in K$ such that $k_2 \geq k_1$. By **Algorithm 1**, we have that

$$\boldsymbol{x}_{k_2+1} = \boldsymbol{x}_{k_2} + t_{k_2}\boldsymbol{d}_{k_2}$$

with $t_{k_2} \in (0, 1]$, whereby $\|\boldsymbol{x}_{k_2+1} - \boldsymbol{x}_{k_2}\| \leq \|\boldsymbol{d}_{k_2}\| \leq \delta/4$. Thus,

$$\begin{aligned} dist(\boldsymbol{x}_*; \Omega\backslash\{\boldsymbol{x}_*\}) &\leq dist(\boldsymbol{x}_{k_2+1}; \Omega\backslash\{\boldsymbol{x}_*\}) + \|\boldsymbol{x}_* - \boldsymbol{x}_{k_2+1}\| \\ &\leq dist(\boldsymbol{x}_{k_2+1}; \Omega\backslash\{\boldsymbol{x}_*\}) + \|\boldsymbol{x}_* - \boldsymbol{x}_{k_2}\| + \|\boldsymbol{x}_{k_2} - \boldsymbol{x}_{k_2+1}\| \\ &\leq dist(\boldsymbol{x}_{k_2+1}; \Omega\backslash\{\boldsymbol{x}_*\}) + \delta/4 + \delta/4. \end{aligned}$$

That is,

$$dist(\boldsymbol{x}_{k_2+1}; \Omega\backslash\{\boldsymbol{x}_*\}) \geq dist(\boldsymbol{x}_*; \Omega\backslash\{\boldsymbol{x}_*\}) - \delta/2 \geq \delta - \delta/2 = \delta/2,$$

consequently, $\boldsymbol{x}_{k_2+1} \notin \Omega_1\backslash\overline{B}(\boldsymbol{x}_*; \delta/4)$ and since $\boldsymbol{x}_{k_2+1} \in \Omega_1$ then $\boldsymbol{x}_{k_2+1} \in \overline{B}(\boldsymbol{x}_*; \delta/4)$; that is, $k_2 + 1 \in K$ and since $(k_2 + 1) > k_1$, then, using induction $k_2 \in K$, for all $k_2 > k_1$, which implies that $\boldsymbol{x}_k \in \overline{B}(\boldsymbol{x}_*; \delta/4)$, for all $k_2 > k_1$. Thus, $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}_*$. □

Theorem 6 *Let $\boldsymbol{x}_*$ be an accumulation point of the sequence $\{\boldsymbol{x}_k\}$ generated by **Algorithm 1** such that $\boldsymbol{x}_*$ is an R-regular solution of GCP(F, G). Then, the sequence $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}_*$, the search direction $\boldsymbol{d}_k$ is given by the solution of the linear system (24) eventually, and the full stepsize $t_k = 1$ is accepted for all k sufficiently large.*

Proof By hypothesis, $\boldsymbol{x}_*$ is an R-regular solution of the GCP(F, G) and by **Corollary** 2, $\boldsymbol{x}_*$ is also a BD-regular solution then the **Proposition 3** in [8] guarantees that $\boldsymbol{x}_*$ is an isolated accumulation point Therefore, the sequence $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}_*$ by **Theorem** 5.

Since $\{\boldsymbol{x}_k\}$ converges to a BD-regular solution of the system $\Phi_\lambda(\boldsymbol{x}) = 0$ then for k large enough, H_k is nonsingular, therefore the system (24) always has solution [10]. Let's see that it satisfies the inequality

$$\nabla\Psi_\lambda(\boldsymbol{x}_k)^T\boldsymbol{d}_k \leq -\rho_1\|\boldsymbol{d}_k\|^2, \tag{34}$$

for some $\rho_1 > 0$.

From (24),

$$\|\boldsymbol{d}_k\| \leq \left\| H_k^{-1} \right\| \|\Phi_\lambda(\boldsymbol{x}_k)\| . \tag{35}$$

Furthermore, Since $\nabla\Psi_\lambda(\boldsymbol{x}_k) = H_k^T \Phi_\lambda(\boldsymbol{x}_k)$ and $H_k \boldsymbol{d}_k = -\Phi_\lambda(\boldsymbol{x}_k)$, we get

$$\nabla\Psi_\lambda(\boldsymbol{x}_k)^T \boldsymbol{d}_k = -\|\Phi_\lambda(\boldsymbol{x}_k)\|^2. \tag{36}$$

Combining (35) and (36), we have

$$\nabla\Psi_\lambda(\boldsymbol{x}_k)^T \boldsymbol{d}_k \leq -\frac{\|\boldsymbol{d}_k\|^2}{N^2}, \tag{37}$$

where N is an upper bound of $\left\| H_k^{-1} \right\|$ (which exists by **Lemma 2.6** of [10]). Taking $\rho_1 = 1/N^2$ in (37), we get (34). Since the sequence $\{\|\boldsymbol{d}_k\|\}$ converges to 0. From (34) it follows that (25), holds for any $p > 2$ and any positive constant ρ.

Finally, let's see that the full step size is achieved; that is, $i_k = 0$. First, Ψ_λ is class LC^1 by **Theorem 2.3** in [12]. With this hypothesis and **Theorem 3.2** in [3] guarantees that, starting from a certain k, $t_k = 1$. □

Finally, we present a result that guarantees the rate of convergence of algorithm.

Theorem 7 *If $\boldsymbol{x}_*$ is an accumulation point of the sequence $\{\boldsymbol{x}_k\}$ generated by **Algorithm 1**, such that $\boldsymbol{x}_*$ is an R-regular solution of the GCP(F,G) then the sequence $\{\boldsymbol{x}_k\}$ generated by **Algorithm 1** converges q-superlinearly to $\boldsymbol{x}_*$. In addition, if Φ_λ is an LC^1 mapping then the convergence rate is q-quadratic.*

Proof Since $\boldsymbol{x}_*$ is R-regular and by **Lemma 2.6** of [10], H_k is nonsingular and $\left\| H_k^{-1} \right\| \leq N$, for some positive constant N; so, $t_k = 1$ is accepted from a certain value of k, the sequence $\{\boldsymbol{x}_k\}$ is well defined and

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - H_k^{-1}\Phi_\lambda(\boldsymbol{x}_k).$$

Subtracting $\boldsymbol{x}_*$ and using a norm $\|\cdot\|$, we have

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_*\| = \left\| \boldsymbol{x}_k - \boldsymbol{x}_* - H_k^{-1}\Phi_\lambda(\boldsymbol{x}_k) \right\| .$$

After some algebraic manipulations, we obtain

$$\begin{aligned} \|\boldsymbol{x}_{k+1} - \boldsymbol{x}_*\| &= \left\| \boldsymbol{x}_k - \boldsymbol{x}_* - H_k^{-1}\Phi_\lambda(\boldsymbol{x}_k) \right\| \\ &\leq \left\| H_k^{-1} \left[\Phi_\lambda(\boldsymbol{x}_k) - \Phi_\lambda(\boldsymbol{x}_*) - \Phi'_\lambda(\boldsymbol{x}_*; \boldsymbol{x}_k - \boldsymbol{x}_*) \right] \right\| \\ &\quad + \left\| H_k^{-1} \left[H_k(\boldsymbol{x}_k - \boldsymbol{x}_*) - \Phi'_\lambda(\boldsymbol{x}_*; \boldsymbol{x}_k - \boldsymbol{x}_*) \right] \right\| . \end{aligned} \tag{38}$$

To bound the two terms on the right side in (38), we must take into account that Φ_λ is semismooth (**Theorem** 2) and, therefore it is directionally differentiable. Then, from (9),

$$\left\| \Phi_\lambda(\boldsymbol{x}_k) - \Phi_\lambda(\boldsymbol{x}_*) - \Phi'_\lambda(\boldsymbol{x}_*; \boldsymbol{x}_k - \boldsymbol{x}_*) \right\| = o\left\| \boldsymbol{x}_k - \boldsymbol{x}_* \right\|,$$

and by **Lemma** 1,

$$\left\| H_k(\boldsymbol{x}_k - \boldsymbol{x}_*) - \Phi'_\lambda(\boldsymbol{x}_*; \boldsymbol{x}_k - \boldsymbol{x}_*) \right\| = o\left\| \boldsymbol{x}_k - \boldsymbol{x}_* \right\|,$$

using the last two equalities and since $\left\| H_k^{-1} \right\| \leq N$ in (38), we conclude

$$\left\| \boldsymbol{x}_{k+1} - \boldsymbol{x}_* \right\| = o\left\| \boldsymbol{x}_k - \boldsymbol{x}_* \right\|, \tag{39}$$

which shows that the sequence $\{\boldsymbol{x}_k\}$ converges q-superlinearly to $\boldsymbol{x}_*$.

Now, we assume that Φ_λ is an LC^1 mapping. Then, by **Theorem** 2, it is also *strongly semismooth* at $\boldsymbol{x}_*$. Moreover, the directional derivative $\Phi'_\lambda(\cdot, \cdot)$ is *semi-continuous of degree* 2 [2]. Then, from (11) and by **Lemma** 2, we have

$$\left\| H_k(\boldsymbol{x}_k - \boldsymbol{x}_*) - \Phi'_\lambda(\boldsymbol{x}_*; \boldsymbol{x}_k - \boldsymbol{x}_*) \right\| = O\left\| \boldsymbol{x}_k - \boldsymbol{x}_* \right\|^2$$

and $\left\| \Phi_\lambda(\boldsymbol{x}_k) - \Phi_\lambda(\boldsymbol{x}_*) - \Phi'_\lambda(\boldsymbol{x}_*; \boldsymbol{x}_k - \boldsymbol{x}_*) \right\| = O\left\| \boldsymbol{x}_k - \boldsymbol{x}_* \right\|^2$. Hence, from (38),

$$\left\| \boldsymbol{x}_{k+1} - \boldsymbol{x}_* \right\| = O\left\| \boldsymbol{x}_k - \boldsymbol{x}_* \right\|^2,$$

which shows that the sequence $\{\boldsymbol{x}_k\}$ converges q-quadratically to $\boldsymbol{x}_*$. □

5 Numerical experiments

In this section, we analyze the numerical performance of the **Algorithm 1** proposed in **Section** 4, which we call **Algorithm NG**, and compare it with the algorithm proposed in [15] that we call **Algorithm NGS**. Moreover, we incorporate to **Algorithm NG** the dynamic procedure to update, at each iteration, the parameter λ proposed in [12], obtaining a new algorithm that we call **Algorithm NGD**, which we also compare with the two previous ones.

The algorithms were implemented in MATLAB® R2019a and tested on a computer with an AMD Sempron (tm) processor of 2.21 GHz. The parameters are the following: $\epsilon = 10^{-4}$, $N = 100$, $\rho = 10^{-8}$, $p = 2.1$ and $\sigma = 10^{-4}$.

We did four types of experiments. First, we compare the performance of the **Algorithms NGD** and **NGS** in terms of iterations number and CPU time. Second, we analyze the global performance of **Algorithms NGD** and **NGS**. Third, we analyze the behavior of **Algorithm NG** in terms of iterations number for different values of λ. Finally, in the fourth experiment, we compare the **Algorithms NG** and **NGD**.

The experiments were carried out with seven problems. Below, we describe each of them, its solutions and the starting points used. For simplicity, we denote $\boldsymbol{e}_n$ as the vector of ones of order n.

Problem 1 [24]. F is the *Kojima-Shindo* function, and G is the identity function. Thus, the GCP(F, G) reduces to a nonlinear complementarity problem. Explicitly $F, G : \mathbb{R}^4 \to \mathbb{R}^4$ are defined by

$$F(\boldsymbol{x}) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_2^2 + x_1 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix} \quad \text{and} \quad G(\boldsymbol{x}) = \boldsymbol{x}\,.$$

The solutions are $\boldsymbol{x}_* = (1,\ 0,\ 3,\ 0)$ and $\boldsymbol{x}_* = \left(\sqrt{6}/2,\ 0,\ 0,\ 1/2\right)$ and the starting points are $\boldsymbol{x}_1 = (0,\ 0,\ 0,\ 0)\,,\ \boldsymbol{x}_2 = (1,\ 0,\ 1,\ 0)\,,\ \boldsymbol{x}_3 = (1,\ 0,\ 0,\ 0)\,,$ $\boldsymbol{x}_4 = (0,\ 1,\ 1,\ 0)\,.$

Problem 2 [15]. The functions $F, G : \mathbb{R}^2 \to \mathbb{R}^2$ are defined by

$$F(\boldsymbol{x}) = \begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix} \quad \text{and} \quad G(\boldsymbol{x}) = \begin{pmatrix} x_1^2 + 10 \\ x_2^2 + 1 \end{pmatrix}.$$

The unique solution is $\boldsymbol{x}_* = (0,\ 0)$ and the starting points are $\boldsymbol{x}_1 = (10,\ 1)\,,\ \boldsymbol{x}_2 = 100\boldsymbol{e}_2,\ \boldsymbol{x}_3 = 1000\boldsymbol{e}_2$ and $\boldsymbol{x}_4 = 10000\boldsymbol{e}_2$.

Problem 3 [5]. The functions $F, G : \mathbb{R}^2 \to \mathbb{R}^2$ are defined by

$$F(\boldsymbol{x}) = \begin{pmatrix} -\frac{100}{3} + 2x_1 + \frac{8}{3}x_2 \\ -22.5 + 2x_2 + \frac{5}{4}x_1 \end{pmatrix} \quad \text{and} \quad G(\boldsymbol{x}) = \begin{pmatrix} 15 - x_2 \\ 20 - x_1 \end{pmatrix}.$$

The solutions are $\boldsymbol{x}_* = (10,\ 5)$ and $\boldsymbol{x}_* = (20,\ 15)\,.$ The starting points are $\boldsymbol{x}_1 = (0,\ 0)\,,\ \boldsymbol{x}_2 = (5,\ 0)$ and $\boldsymbol{x}_3 = (11,\ 0)\,.$

Problem 4 [5]. This is an implicit complementarity problem where the functions $F, G : \mathbb{R}^4 \to \mathbb{R}^4$ are defined by $F(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ and $G(\boldsymbol{x}) = \boldsymbol{x} - h(\boldsymbol{x})$, where $A = tridiag(-1, 2, -1)$, $\boldsymbol{b} = \boldsymbol{e}_4$ and $h_i(\boldsymbol{x}) = -0.5 - x_i$.

The solution is $\boldsymbol{x}_* = (-0.9,\ -1.2,\ -1.2,\ -0.9)$ and the starting points are $\boldsymbol{x}_1 = (0,\ 0,\ 0,\ 0)\,,\ \boldsymbol{x}_2 = -0.5\boldsymbol{e}_4$ and $\boldsymbol{x}_3 = -\boldsymbol{e}_4$.

Problem 5 [8]. The functions $F, G : \mathbb{R}^5_+ \to \mathbb{R}^5$ are defined by

$$G(\boldsymbol{x}) = F(\boldsymbol{x}) = \boldsymbol{c} + L^{\frac{1}{b}}\boldsymbol{x}^{\frac{1}{b}} - \left(\frac{5000}{\sum_{i=1}^{5} x_i}\right)^{\frac{1}{\gamma}} \left(\boldsymbol{e}_5 - \frac{1}{\gamma\left(\sum_{i=1}^{5} x_i\right)}\boldsymbol{x}\right),$$

where $\boldsymbol{c} = (10,\ 8,\ 6,\ 4,\ 2)^T$, $\boldsymbol{b} = (1.2,\ 1.1,\ 1,\ 0.9,\ 0.8)^T$, $L = 5\boldsymbol{e}_5$, and $\gamma = 1.1$. In this problem all operations are component-wise.

The solution is $\boldsymbol{x}_* = (15.4293,\ 12.4986,\ 9.6635,\ 7.1651,\ 5.1326)^T$ and the starting points are $\boldsymbol{x}_1 = \boldsymbol{e}_5$, $\boldsymbol{x}_2 = 10\boldsymbol{e}_5$ and $\boldsymbol{x}_3 = 20\boldsymbol{e}_5$.

Problems 6 and **7** [25]. The functions $F, G: \mathbb{R}^n \to \mathbb{R}^n$ are defined by

$$F(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{q} + \Gamma(\boldsymbol{x}) \quad \text{and} \quad G(\boldsymbol{x}) = \boldsymbol{x} - \Upsilon(\boldsymbol{x}),$$

where $A \in \mathbb{R}^{n\times n}$, $\boldsymbol{q} = (-1, 1, \ldots, (-1)^n) \in \mathbb{R}^n$, $\Upsilon(\boldsymbol{x}) = \left(x_1^3, x_2^3, \ldots, x_n^3\right)$ and $\Gamma(\boldsymbol{x}) = \left(x_1^2, x_2^2, \ldots, x_n^2\right)$.

The solution for both problems is $\boldsymbol{x}_* = \boldsymbol{e}_n$ and the starting points are $\boldsymbol{x}_1 = (1, 0.6, 1, 0.6, \ldots)$, $\boldsymbol{x}_2 = 5\boldsymbol{e}_n$ and $\boldsymbol{x}_3 = 15\boldsymbol{e}_n$.

For **Problem 6**, we define as in [25]:

$$A = tridiag(-I, S, -I) = \begin{pmatrix} S & -I & 0 \ldots & 0 & 0 \\ -I & S & -I \ldots & 0 & 0 \\ 0 & -I & S \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \ldots & S & -I \\ 0 & 0 & 0 \ldots & -I & S \end{pmatrix},$$

with

$$S = tridiag(-1, 4, -1) = \begin{pmatrix} 4 & -1 & 0 \ldots & 0 & 0 \\ -1 & 4 & -1 \ldots & 0 & 0 \\ 0 & -1 & 4 \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \ldots & 4 & -1 \\ 0 & 0 & 0 \ldots & -1 & 4 \end{pmatrix} \in \mathbb{R}^{m\times m}.$$

For **Problem 7**, the matrix A is defined by $A = tridiag(-1.5I, S, -0.5I)$, with $S = tridiag(-1.5, 4, -0.5)$.

The results of the experiments are presented in Fig. 1 and Tables 1, 2, and 3 with the following information: number of the test problem (**Prob**), starting point (*SP*), number of iterations (k), *CPU* time (*CPU*), value of the objective function at the final iteration (Ψ), success rate (S), divergence ($-$).

5.1 Experiment 1

In this experiment, we compared the **Algorithms NGD** and **NGS** in terms of number of iterations, the CPU time and the accuracy at the solution. In Table 1, we report the results obtained when executing the algorithms to solve each of the problems described above.

In this table, the notation **6** (n) and **7** (n), in the first column, means that the algorithms were executed to solve the **Problems 6** and **7**, for $n = m^2$, with $m = 8$ and $m = 10$.

From the Table 1, we see that for problems 1 to 4, the algorithms have a similar behavior in terms of number of iterations and CPU time, with a slight CPU time advantage of the **Algorithm NGD**. However, when we compare the value of the merit function, $\Psi(\boldsymbol{x})$, in the final iteration, **Algorithm NGD** reached, in 90% of the cases,

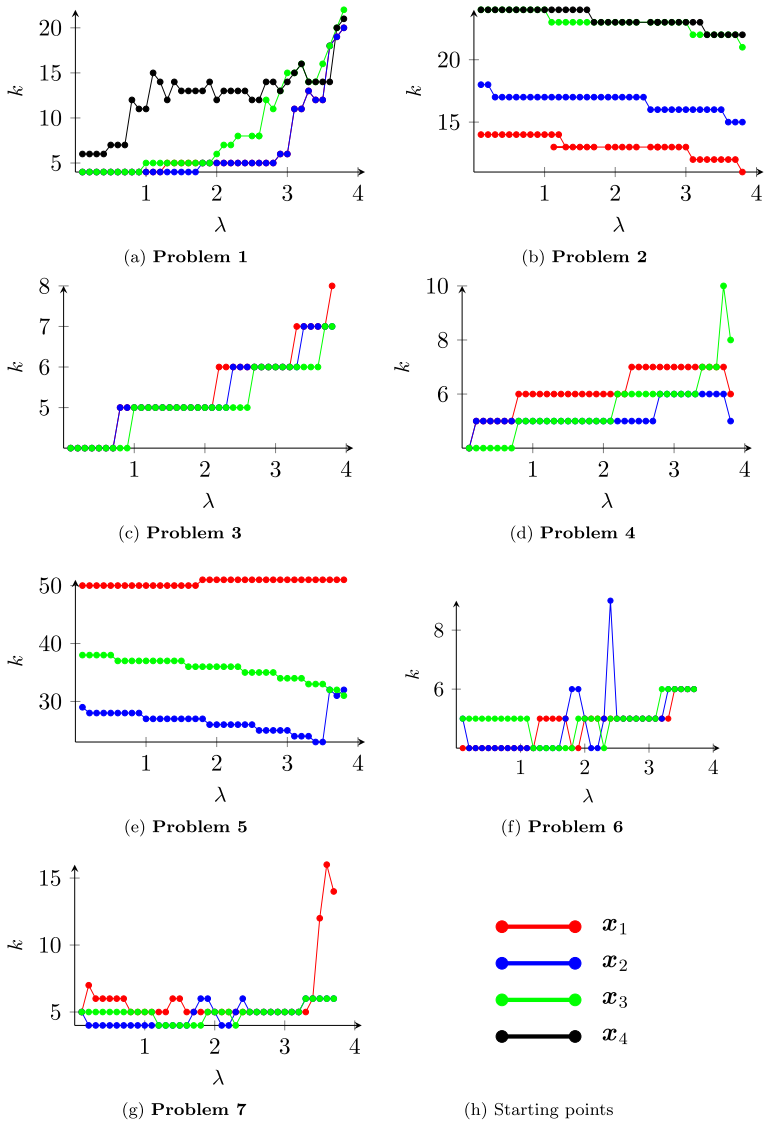

Fig. 1 Performance of the **Algorithm NG** for different values of λ

a value lower than that reached by **Algorithm NGS**. In some cases, the difference in these values was of the order of 10^{-5}; this suggests that **Algorithm NGD** probably gives a better approximation to the solution of GCP(F, G) by about same number of iterations as **Algorithm NGS**. For the remaining **Problems** (**5**, **6** and **7**), the **NGS Algorithm** did not converge, while the **NGD Algorithm** did. Furthermore, it converges in few iterations, with an acceptable CPU time.

On the other hand, this experiment shows us that the λ-dynamic strategy implemented in **Algorithm NGD** takes advantage when the iterations are close of the

Table 1 Results for the algorithms NGD and NGS

Prob	*SP*	NGD			NGS		
		k	*CPU*	Ψ	k	*CPU*	Ψ
1	$\boldsymbol{x}_1$	12	1.1	1.2×10^{-10}	12	1.5	3.8×10^{-8}
	x_2	5	0.4	1.6×10^{-11}	5	0.7	5.8×10^{-10}
	x_3	6	0.5	1.4×10^{-8}	6	0.6	1.8×10^{-7}
	sx_4	12	1.1	1.9×10^{-12}	–	–	–
2	$\boldsymbol{x}_1$	14	0.5	1.4×10^{-6}	13	0.5	2.8×10^{-6}
	$\boldsymbol{x}_2$	17	0.6	3.1×10^{-6}	17	0.6	1.5×10^{-6}
	$\boldsymbol{x}_3$	24	1.3	1.3×10^{-6}	23	0.8	2.6×10^{-6}
	$\boldsymbol{x}_4$	24	0.9	1.8×10^{-6}	23	0.9	3.6×10^{-6}
3	$\boldsymbol{x}_1$	5	0.2	1.1×10^{-6}	5	0.3	1.6×10^{-6}
	$\boldsymbol{x}_2$	5	0.3	1.5×10^{-6}	5	0.2	3.9×10^{-6}
	$\boldsymbol{x}_3$	5	0.3	2.1×10^{-10}	5	0.2	2.3×10^{-8}
4	$\boldsymbol{x}_1$	6	0.3	2.4×10^{-12}	6	0.3	7.5×10^{-7}
	$\boldsymbol{x}_2$	5	0.2	2.7×10^{-10}	5	0.2	1.1×10^{-6}
	$\boldsymbol{x}_3$	5	0.2	3.8×10^{-10}	5	0.2	1.7×10^{-6}
5	$\boldsymbol{x}_1$	8	0.8	0.0	–	–	–
	$\boldsymbol{x}_2$	4	0.3	0.0	–	–	–
	$\boldsymbol{x}_3$	5	0.4	1.0×10^{-15}	–	–	–
6 (64)	$\boldsymbol{x}_1$	6	13.3	3.6×10^{-13}	–	–	–
	$\boldsymbol{x}_2$	10	11.9	1.3×10^{-8}	–	–	–
	$\boldsymbol{x}_3$	12	14.3	9.1×10^{-10}	–	–	–
6 (100)	$\boldsymbol{x}_1$	6	37.26	4.7×10^{-11}	–	–	–
	$\boldsymbol{x}_2$	14	41.7	3.2×10^{-12}	–	–	–
	$\boldsymbol{x}_3$	13	73.2	4.7×10^{-9}	–	–	–
7 (64)	$\boldsymbol{x}_1$	6	17.1	5.5×10^{-13}	–	–	–
	$\boldsymbol{x}_2$	10	20.3	3.51×10^{-9}	–	–	–
	$\boldsymbol{x}_3$	13	22.4	1.1×10^{-10}	–	–	–
7 (100)	$\boldsymbol{x}_1$	6	71.04	5.4×10^{-11}	–	–	–
	$\boldsymbol{x}_2$	11	92.4	4.7×10^{-8}	–	–	–
	$\boldsymbol{x}_3$	13	101.2	4.3×10^{-10}	–	–	–

solution $\boldsymbol{x}_*$. Proof of this is the fast decrease in the value of $\Psi(\boldsymbol{x})$ in the last iterations.

5.2 Experiment 2

In this experiment, we study the global performance of the **Algorithms NGD** and **NGS**. For this, we tested the algorithms taking one hundred random starting points

Table 2 Success rate for the **Algorithms NGD** and **NGS**

Prob	NGD		NGS	
	$\overline{k}$	S (%)	$\overline{k}$	S (%)
1	10	96	9	87
2	14	98	14	98
3	6	100	6	97
4	6	100	6	100
5	34	100	35	100
6 (64)	41	98	40	96
6 (100)	23.1	84	24	79
7 (64)	42	94	42	94
7 (100)	25	97	27	95

for each of the five problems. For **Problems 1** to **4**, we take each component of the starting points in the interval $[-30, 30]$ and for the remaining problems, the interval $[1, 50]$. The results of this experiment are summarized in Table 2, in which $\overline{k}$ is the mean of the iterations of the hundred of experiments for each problem.

As we expected, the **Algorithm NGD** seems to be a robust algorithm. We want to highlight the performance of our algorithm for the first problem since this is considered a hard one [12]. In this problem, **Algorithm NGD** had a success rate of 96% while its counterpart, **Algorithm NGS**, had a success rate of 87%, so we can conclude that the λ-dynamic strategy combined with a linear search less restrictive, such as an Armijo-type linear search, works very well and can lead to a robust and efficient algorithm.

5.3 Experiment 3

The above experiments show that the λ-dynamic strategy is a good option to complement the **Algorithm NG**, but a reasonable question is the following: What is the performance of this algorithm without λ-dynamic strategy? What is the behavior of the algorithm when it works with the same value of λ throughout the execution? And what is the better choice of λ to run the algorithm? In this experiment we investigate these situations. For this, we tested the algorithm with different values for λ.

In Fig. 1, we show the number of iterations required by the **Algorithm NG** to solve each of the five problems when it was executed with 38 values of λ equally spaced in the interval $(0,\ 4)$. We observe that **Algorithm NG** converges in all experiments and for all values of λ which reflects in a sense, the robustness of algorithm. On the other hand, we can see that the **Algorithm NG** had his best performance taking values of λ at the beginning or at the end of interval $(0,\ 4)$.

Comparing the results of Fig. 1 with those on Table 1, we can conclude that in all the experiments, it is possible to find a value of λ for which **Algorithm NG** is faster than **Algorithm NGS**; however, there is no a single value of λ that works well for all

Table 3 Performance of the **Algorithms NG** and **NGD**

Prob	SP	NG λ_B	λ_W	k_B	k_W	$\bar{k}$	NGD k
1	x_1	(0, 0.4)	(3.5, 3.8)	11	19	15	12
	x_2	(0, 1.7)	(3.7, 3.8)	4	20	12	5
	x_3	(0, 0.9)	(3.7, 3.8)	4	22	13	6
	x_4	(0, 0.4)	(3.7, 3.8)	6	21	13.5	12
2	x_1	(3.7, 3.8)	(0, 1.2)	11	14	12.5	9
	x_2	(3.5 3.8)	(0, 0.2)	15	17	16	12
	x_3	(3.7, 3.8)	(0, 1.0)	21	24	22.5	19
	x_4	(3.2, 3.8)	(0, 1.6)	22	24	23	19
3	x_1	(0, 0.7)	(3.7, 3.8)	4	8	6	5
	x_2	(0, 0.7)	(3.3, 3.8)	4	7	5.5	5
	x_3	(0, 0.9)	(3.6, 3.7)	4	7	5.5	5
4	x_1	(0, 0.1)	(2.3, 3.7)	4	7	5.5	6
	x_2	(0, 0.1)	(2.7, 3.7)	4	6	5	5
	x_3	(0, 0.7)	(3.6, 3.7)	4	10	7	5
5	x_1	(0, 1.7)	(1.7, 3.8)	50	51	50.5	49
	x_2	(3.3, 3.5)	(0, 0.9)	23	28	25.5	29
	x_3	(3.6, 3.8)	(0, 0.5)	32	38	35	39
6 (100)	x_1	(0, 1.2)	(3.4, 3.8)	4	6	5	6
	x_2	(0.2, 1.6)	(3.3, 3.8)	4	6	5	14
	x_3	(1.2, 1.8)	(3.2, 3.8)	4	6	5	13
7 (100)	x_1	(0.8, 1.3)	(3.5, 3.8)	5	14	9.5	6
	x_2	(0.2, 1, 6)	(3.3, 0.9)	4	6	5	11
	x_3	(1.2, 1.8)	(3.3, 0.5)	4	6	5	13

problems; therefore, the λ-dynamic strategy seems to be the best option to complement the **Algorithm NG**.

5.4 Experiment 4

In this experiment, we summarize the results obtained in **Experiments 1** and **3** for **Algorithms NGD** and **NG** with the aim to compare the best and worst performance of the **Algorithm NG** with the one of the **Algorithm NGD**.

The results are shown in Table 3, where λ_B indicates the subintervals for which, taking values of λ into them, the **Algorithm NG** shows his best behavior, and λ_W indicates the subintervals for which the **Algorithm NG** had his worst behavior. Analogously, k_B and k_W are the number of iterations needed for the **Algorithm NG** in his best and worst behavior, respectively.

In Table 3, we observe that, for all problems except **Problem 2**, it is possible to find a value of λ for which **Algorithm NG** solved the corresponding GCP(F, G) in fewer iterations than **Algorithm NGD**. However, there is no single value of λ that makes **Algorithm NG** perform better than **Algorithm NGD** in all experiments.

On the other hand, we observe that in general, in the worst case, the **Algorithm NG** needed many more iterations to solve the GCP(F, G) than **Algorithm NGD** which increases significantly its average of iterations.

From above observations, we can conclude that when we do not have the best value of λ for **Algorithm NG** available, the λ-dynamic strategy is a good option. However, we made some preliminary tests varying the initial λ in the strategy proposed in [12], and the results showed us that by making some changes in the initial value of λ, the **Algorithm NGD** can have a better performance, so we think that this is an open problem in this area.

6 Conclusions

In this work, we present a new nonsmooth Newton-type algorithm for solving the generalized complementarity problem indirectly, reformulating it as a system of nonlinear equations using a one-parametric family of complementarity functions.

We demonstrate that this algorithm converges locally and q-quadratically. In addition, we verify that properties of operator Φ_λ, investigated in [12] for the nonlinear complementarity case, can easily be extended to the generalized complementarity problem. For this, it was necessary to introduce some new definitions such as the concepts of BD-regularity and R-regularity in generalized complementarity.

We also present some numerical experiments that show the good performance of the new algorithm. In particular, these experiments showed that for each problem, there exists a subinterval of values of λ that give us better performance of **Algorithm NG**. For this, we consider important to modify the variation procedure of parameter λ changing its initial value by one that belongs to some of the subintervals previously mentioned. We leave this as a future research topic.

Acknowledgements We are grateful to the University of Cauca for the time allowed for this research.

Author Contributions All authors contributed equally to this work.

Funding Open Access funding provided by Colombia Consortium

Data Availability The data obtained in this research are available.

Code Availability This work introduces an algorithm completely available for the readers.

Declarations

Ethical approval The authors declare that they followed all the rules of a good scientific practice.

Consent to participate All authors approve their participation in this work.

Consent for publication the authors approve the publication of this research.

Conflict of interest The authors declare no competing interests.

Human and animal ethics Not applicable

References

1. Cottle, R.W., Pang, J. S., Stone, R.E. The linear complementarity problem, Philadelphia (2009)
2. De Luca, T., Facchinei, F., Kanzow, C.: A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 407–439 (1996)
3. Facchinei, F.: Minimization of SC1 functions and the Maratos effect. Operation Res. Lett. **17**(3), 131–137 (1995)
4. Arias, C., Martínez, H.J., Pérez, R.: A global quasi-Newton algorithms for nonlinear complementarity problems. Pacific J Optim. **13**(1), 1–15 (2017)
5. Outrata, J.V., Zowe, J.: A Newton method for a class of quasi-variational inequalities. Comput. Optim. Appl. **4**(1), 5–21 (1995)
6. Pang, J.S.: On the convergence of a basic iterative method for the implicit complementarity problem. J Optim. Theory Appl. **37**(2), 149–162 (1982)
7. Dolcetta, I.C., Mosco, U. Implicit complementarity problems and quasivariational inequalities, in variational inequalities and complementarity problems: theory and applications. R. W. Cottle, F. Giannessi, and J. L. Lions, 75–87 (1980)
8. Pang, J.S.: On the convergence of a basic iterative method for the implicit complementarity problem. J Optim. Theory y Appl. **37**(2), 149–162 (1982)
9. Arenas, F., Mart ínez, H.J., P érez, R.: Least change secant update methods for nonlinear complementarity problem. Ingenier ía y Ciencia 11(21), 11–36 (2015)
10. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**(1), 227–244 (1993)
11. Jiang, H., Fukushima, M., Qi, L., Sun, D.: A trust region method for solving generalized complementarity problems. SIAM J Optim. **8**(1), 140–157 (1998)
12. Kanzow, C., Kleinmichel, H.: A new class of semismooth Newton-type methods for nonlinear complementarity problems. Comput. Optim. Appl. **11**(3), 227–251 (1998)
13. Clarke, F.H. Necessary conditions for nonsmooth problems in optimal control and the calculus of variations. PhD thesis, University of Washington (1973)
14. Song, L., Gao, Y.: On the local convergence of a Levenberg-Marquardt method for nonsmooth nonlinear complementarity problem. ScienceAsia **43**(1), 377–382 (2017)
15. Du, S.-Q.: Generalized Newton method for a kind of complementarity problem. Abstract Appl. Anal. **2014**, 1–5 (2014)
16. Fischer, A., Kanzow, C.: On finite termination of an iterative method for linear complementarity problems. Math. Program. **74**(3), 279–292 (1996)
17. Pang, J.S.: Newton's method for B-differentiable equations. Math. Oper. Res. **15**(2), 311–347 (1990)
18. Lopes, V.L.R., Martínez, J.M., Pérez, R.: On the local convergence of quasi-Newton methods for nonlinear complementarity problems. Appl. Numer. Math. **30**(1), 3–22 (1999)
19. Qi, L. C-differentiability, C-differential operators and generalized Newton methods. Technical report, Applied Mathematics Report AMR96/5, University of New South Wales, Sydney, Australia (1996)

20. Sun, D., Han, J.: Newton and quasi-Newton methods for a class of nonsmooth equations and related problems. SIAM J Optim. **7**(2), 463–480 (1997)
21. Facchinei, F., Soares, J.: A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J Optim. **7**(1), 225–247 (1997)
22. Nocedal, J., Wright, S.: Numerical Optimization. Springer, New York (2006)
23. Bertsekas, D.P.: Constrained Optimization and Lagrange Multipliers Methods. Academic Press, INC, New York (2014)
24. Kojima, M., Shindo, S.: Extension of Newton and quasi-Newton methods to systems of PC1 equations. J Oper. Res. Soc. Japan **29**(4), 352–375 (1986)
25. Wu, S.L., Guo, P.: Modulus-based matrix splitting algorithms for the quasi-complementarity problems. Appl. Numer. Math. **132**, 127–137 (2018)