Journal of Global Optimization **24:** 261–283, 2002.

NCP Functions Applied to Lagrangian Globalization for the Nonlinear Complementarity Problem

LIQUIN QI[1] and YU-FEI YANG[2]
[1]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: maqilq@polyu.edu.hk);* [2]*College of Mathematics and Econometrics, Hunan University, Chang-sha, P.R. China*

Abstract. Based on NCP functions, we present a Lagrangian globalization (LG) algorithm model for solving the nonlinear complementarity problem. In particular, this algorithm model does not depend on some specific NCP function. Under several theoretical assumptions on NCP functions we prove that the algorithm model is well-defined and globally convergent. Several NCP functions applicable to the LG-method are analyzed in details and shown to satisfy these assumptions. Furthermore, we identify not only the properties of NCP functions which enable them to be used in the LG method but also their properties which enable the strict complementarity condition to be removed from the convergence conditions of the LG method. Moreover, we construct a new NCP function which possesses some favourable properties.

Key words: NCP function; Nonlinear complementarity problem; Lagrangian globalization; Strict complementarity condition; Global convergence

1. Introduction

In this paper we consider the nonlinear complementarity problem (NCP for abbreviation): Find a vector $x \in \mathbb{R}^n$ satisfying the conditions

$$x \geqslant 0\,, \qquad F(x) \geqslant 0\,, \qquad x^T F(x) = 0\,, \tag{1}$$

where $F: \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be continuously differentiable.

The NCP has many applications in economics, engineering and various equilibrium models [6, 9, 17]. Many numerical methods for solving the NCP have been developed, e.g., see [2, 5, 8, 10, 14, 23]. Among them, an important class of iterative methods are based on reformulating the NCP as a system of nonlinear equations.

By means of a so-called NCP function, i.e., a mapping $\psi: \mathbb{R}^2 \to \mathbb{R}$ satisfying

$$\psi(a, b) = 0 \Leftrightarrow a \geqslant 0\,, \quad b \geqslant 0 \text{ and } ab = 0\,,$$

the NCP is reformulated as an equivalent system of nonsmooth equations:

$$\Phi(x) = 0\,, \tag{2}$$

where $\Phi: \mathbb{R}^n \to \mathbb{R}^n$ is defined by

$$\Phi_j(x) := \psi(x_j, F_j(x))\,, \quad j = 1, \ldots, n\,.$$

In order to globalize such methods, a line search technique is used to achieve a sufficient decrease of the natural merit function

$$\Psi(x) = \frac{1}{2}\|\Phi(x)\|^2 , \quad (3)$$

which is often, if not always, continuously differentiable.

However, such methods are not always successful in finding a solution of the NCP. In general, such methods can only find a stationary point of Ψ and typically require strong assumptions in order to guarantee that every stationary point of Ψ is a solution of the original NCP. Therefore, when one obtains a stationary point $\bar{x}$ of Ψ, which is not a solution of the NCP, it is very important and meaningful to have methods to find another point at which the value of Ψ is less than $\Psi(\bar{x})$. In order to resolve this problem, Chen et al. [3] proposed a Lagrangian globalization (LG) method for solving the NCP, which can be regarded as a generalization of the LG method for solving smooth systems [15, 16]. The success of Chen-Qi-Yang method [3] relies on the use of an NCP function which possesses some favourable properties.

The NCP function used in [3] is not a unique choice, i.e., there exist many other NCP functions which may be used to obtain results similar to those in [3]. This observation motivates us to study further the properties of NCP functions. One question is: What properties should NCP functions possess to be used in the LG method? Moreover, the LG method in [3] suffers a drawback in that the strict complementarity condition is needed for global convergence. Therefore, its applications are limited. Hence, another question arises: Is it possible to remove the strict complementarity condition from the LG method? If possible, what properties should NCP functions have to guarantee to remove the strict complementarity condition in the LG method?

In this paper we will analyze in details the properties of the following NCP functions:

$$\phi_1(a, b) = |\phi(a, b)| ,$$
$$\phi_2(a, b) = \sqrt{\{[-\phi(a, b)]_+\}^2 + [(-a)_+]^2 + [(-b)_+]^2} ,$$
$$\phi_3(a, b) = \sqrt{[\phi(a, b)]^2 + \alpha[(ab)_+]^2} , \quad \alpha > 0 ,$$
$$\phi_4(a, b) = \sqrt{[\phi(a, b)]^2 + \alpha[(ab)_+]^4} , \quad \alpha > 0 ,$$
$$\phi_5(a, b) = \sqrt{[\phi(a, b)]^2 + \alpha[(a_+ b_+)_+]^2} , \quad \alpha > 0 ,$$
$$\phi_6(a, b) = \sqrt{\{[\phi(a, b)]_+\}^2 + \alpha[(ab)_+]^2} , \quad \alpha > 0 ,$$

where $\phi(a, b) := \sqrt{a^2 + b^2} - a - b$ is the Fischer–Burmeister function [7] and for any $u \in \mathbb{R}$, $u_+ := \max\{0, u\}$. The function ϕ_1 is the absolute value function of ϕ and has been studied in [3] while the function ϕ_2 was discussed originally in [22] and then in [20]. Other functions ϕ_i for $i = 3, 4, 5, 6$, have been studied in a recently review on NCP functions [21], also see [1, 11, 12, 24].

As in [3], the aim of the LG method is not to present a new method for solving the NCP, but to find $\hat{x}$ such that $\|\Phi(\hat{x})\| < \|\Phi(\bar{x})\|$ when the NCP has a solution and $\bar{x}$ is a stationary point of Ψ but not a solution of the NCP. In this paper we first present an LG-type algorithm model for solving the NCP and make several theoretical assumptions on NCP functions to ensure global convergence of the algorithm. Then we analyze in details several NCP functions and answer the two questions presented above.

The organization of this paper is as follows. In the next section, we will briefly review several basic concepts and some related properties. In Section 3, based on NCP functions, we first reformulate the nonlinear complementarity problem as an unconstrained optimization problem and present some preliminary results. Then we describe an LG-type algorithm model for solving the NCP. In Section 4, we state several basic assumptions and prove that under these assumptions, the LG method is well-defined and globally convergent. In Section 5 we give a detailed analysis of properties of several NCP functions applicable to the LG method. In Section 6, we show that NCP functions analyzed in the previous section satisfy those basic assumptions under mild limitations and answer the two questions proposed above positively. Some conclusive remarks are given in the last section. Meanwhile, we construct a new NCP function, which is applicable to the LG method and can remove the strict complementarity condition from the convergence conditions for the LG method.

Notation: For a given function $G: \mathbb{R}^n \to \mathbb{R}^m$, we denote by G_i its ith component function. If G is continuously differentiable, $\nabla G(x)$ is the transposed Jacobian of G at $x \in \mathbb{R}^n$. If G is directional differentiable, we denote the directional derivative of G at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ by $G'(x; d)$. The set of all nonnegative real numbers is denoted by $\mathbb{R}_+$. The symbol $\|\cdot\|$ indicates the Euclidean norm, E indicates the $n \times n$ unit matrix and H_j indicates the jth column of the matrix $H \in \mathbb{R}^{n \times n}$.

2. Preliminaries

In this section we state several basic definitions and summarize some related results.

Let $\Theta: \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitzian. Then Θ is differentiable almost everywhere on $\mathbb{R}^n$. Denote the set of points at which Θ is differentiable by D_Θ. The generalized Jacobian of Θ at x in the sense of Clarke [4] is defined by

$$\partial\Theta(x) = \text{co}\{\lim_{\substack{x^k \to x \\ x^k \in D_\Theta}} \nabla\Theta(x^k)^T\}.$$

Let $f: \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near a given point $x \in \mathbb{R}^n$ and let v be a vector in $\mathbb{R}^n$.

The generalized directional derivative of f at x in the direction v, denoted $f^\circ(x; v)$, is defined by

$$f^\circ(x; v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y + tv) - f(y)}{t} .$$

The next lemma displays the relationship between generalized gradients and generalized directional derivatives of f, see Proposition 2.1.2 in [4] for details.

LEMMA 2.1. *Let $f: \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x. Then for any $v \in \mathbb{R}^n$, one has*

$$f^\circ(x; v) = \max\{\xi^T v \mid \xi \in \partial f(x)\} .$$

We now introduce the regularity concept. The function $f: \mathbb{R}^n \to \mathbb{R}$ is said to be regular at $x \in \mathbb{R}^n$ provided
(i) for all $v \in \mathbb{R}^n$, the directional derivative $f'(x; v)$ exists;
(ii) for all $v \in \mathbb{R}^n$, $f'(x; v) = f^\circ(x; v)$.

The next lemma provides some conditions under which the function f is regular, see Proposition 2.3.6 in [4].

LEMMA 2.2. *Let f be Lipschitz near x.*
(i) If f is continuously differentiable at x, then f is regular at x.
(ii) If f is convex, then f is regular at x.

Semismoothness was originally introduced by Mifflin [13] for functionals and shown to be very important in the global convergence theory of nonsmooth optimization. Convex functions, smooth functions and piecewise linear functions are examples of semismooth functions. Qi and Sun [19] extended the definition of semismooth functions to vector valued functions. It has been proven that $\Theta: \mathbb{R}^n \to \mathbb{R}^m$ is semismooth at $x \in \mathbb{R}^n$ if and only if all its component function are and that the composition of (strong) semismooth functions is still a (strong) semismooth function. The following results can be found in [4, 8, 19]. For more details about semismooth functions, see Qi and Sun [19].

LEMMA 2.3. *Let $\theta: \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian near $x \in \mathbb{R}^n$. Then*
(i) $0 \in \partial\theta(x)$ if θ attains a (local) minimum at x.
(ii) $\partial\theta(\cdot)$ is upper semicontinuous in the sense that $\{x^k \to \hat{x},\ y^k \in \partial\theta(x^k)$ and $y^k \to \hat{y}\} \Rightarrow \{\hat{y} \in \partial\theta(\hat{x})\}$.

LEMMA 2.4. *Suppose that $p: \mathbb{R}^n \to \mathbb{R}^n$ and $g: \mathbb{R}^m \to \mathbb{R}$ are Lipschitzian near $x \in \mathbb{R}^n$ and near $p(x)$, respectively. Then the following statements hold.*
(i) The composite function $\vartheta = g \circ p$ is Lipschitzian near x and

$$\partial\vartheta(x) \subseteq co\{\partial g(p(x))\, \partial p(x)\} .$$

In particular, if g is regular at p(x) and p is continuously differentiable at x, then

$$\partial \vartheta(x) = \partial g(p(x)) \nabla p(x)^T .$$

(ii) If p and g are semismooth at x and p(x), respectively, then the composite function ϑ *is also semismooth at x and*

$$\vartheta'(x; d) = g'(p(x); p'(x; d)) , \quad d \in \mathbb{R}^n .$$

(iii) If p and g are strongly semismooth at x and p(x), respectively, then the composite function ϑ *is also strongly semismooth at x.*

3. Algorithm model

In this section we first reformulate the NCP (1) as an unconstrained optimization problem and deduce some relationships between the two problems. Then we present an LG-type algorithm model which does not depend on some specific NCP function.

The LG method for solving nonsmooth system (2) is usually used in an algorithmic framework for solving (2), which is composed of two phases. At the first phase, a standard algorithm such as the generalized Newton method or one of its variants, is used to solve nonsmooth system (2). If it is stuck at a point, say $\bar{x}$, which is not a solution to (2), then go into the next phase. At the second phase, a more expensive algorithm (LG method) is employed to find a new point, say $\hat{x}$ so that $\|\Phi(\hat{x})\| < \|\Phi(\bar{x})\|$. Then the first phase can be reinitiated at $\hat{x}$.

In this paper we will focus on the second phase of the above algorithmic framework, i.e., the LG method for solving (2). As to the whole algorithmic framework above, see [3] for details. The LG method associates an objective function f with system (2) to produce an equality constrained optimization problem of the form:

$$\begin{aligned} &\min \quad f(x) , \\ &\text{s.t.} \quad \Phi(x) = 0 , \end{aligned} \tag{4}$$

where $f: \mathbb{R}^n \to \mathbb{R}$ is specifically chosen and is assumed to be continuously differentiable. For example, $f(x) = \delta e^T x$ where e is the vector of all ones and $\delta \neq 0$ is a constant. See [16] for more details about choices of f.

The Lagrangian function associated with (4) is defined by

$$L(x, \lambda) := f(x) + \lambda^T \Phi(x) \tag{5}$$

and the associated augmented Lagrangian function is defined by

$$P_c(z) := P_c(x, \lambda) := L(x, \lambda) + \frac{1}{2} c \|\Phi(x)\|^2 , \tag{6}$$

where $\lambda \in \mathbb{R}^n$ is the Lagrange multiply vector, c is a nonnegative real parameter and

$z := (x, \lambda)$. Therefore, the NCP (1) is reformulated as the following unconstrained optimization problem:

$$\min_{z \in \mathbb{R}^{2n}} P_c(z) .$$

A point $(x, \lambda) \in \mathbb{R}^{2n}$ is called a critical point of P_c if

$$\begin{aligned} 0 &\in \frac{\partial}{\partial x} P_c(x, \lambda) , \\ 0 &= \frac{\partial}{\partial \lambda} P_c(x, \lambda) = \Phi(x) , \end{aligned} \tag{7}$$

where the generalized derivative of P_c at (x, λ) can be written as

$$\begin{aligned} \frac{\partial}{\partial x} P_c(x, \lambda) &= \{\nabla f(x) + H\lambda + cH\Phi(x) \mid H \in \partial\Phi(x)^T\} , \\ \frac{\partial}{\partial \lambda} P_c(x, \lambda) &= \Phi(x) . \end{aligned} \tag{8}$$

It is easy to deduce the following results, also see [3].

LEMMA 3.1. *(i) If x^* is not a solution of the NCP, then for any $\lambda \in \mathbb{R}^n$, (x^*, λ) is not a critical point of P_c.*

(ii) If (x^*, λ^*) is a critical point of P_c, then x^* is a solution of the NCP.

(iii) If x^* is a solution of the NCP and $\nabla f(x^*)$ belongs to the range space of an element $H^* \in \partial\Phi(x^*)^T$. Then there exists a $\lambda^* \in \mathbb{R}^n$ such that (x^*, λ^*) is a critical point of P_c for any $c \in \mathbb{R}_+$.

We now state an LG-type algorithm model for solving the NCP (1) as follows:

ALGORITHM 3.1.

Step 0. Choose $\sigma, \beta \in (0, 1)$ and initial vector $z^0 = (\bar{x}, \lambda^0) \in \mathbb{R}^{2n}$ with $\lambda^0 \leqslant 0$. Set $k := 0$.

Step 1. If the termination criterion is satisfied, then let $\hat{x} := x^k$, stop.

Step 2. Choose $H^k \in \partial\Phi(x^k)^T$. Denote $V^k := \nabla f(x^k) + H^k(\lambda^k + c\Phi(x^k))$ and

$$q^k := \begin{pmatrix} -V^k \\ -\Phi(x^k) \end{pmatrix} .$$

Step 3. Determine $t^k = \beta^{m_k}$, where m_k is the smallest nonnegative integer m such that

$$P_c(z^k + \beta^m q^k) - P_c(z^k) \leqslant -\sigma\beta^m \|q^k\|^2 . \tag{9}$$

Step 4. Set $z^{k+1} := z^k + t^k q^k$, $k := k + 1$. Go to Step 1

The above algorithm is similar to that in [3]. However, the above algorithm model does not depend on some specific NCP function.

At Step 1 any reasonable termination criterion can be used. Since the LG method aims at looking for another point at which the value of Ψ is less than $\Psi(\bar{x})$ and we will prove in Theorem 4.1 that x-part of any accumulation point of the sequence generated by the above algorithm is a solution of the NCP, in our case we can use the termination criterion of the form: $\|\Phi(x^k)\| \leqslant \gamma \|\Phi(\bar{x})\|$ with some fixed constant $\gamma \in (0, 1)$. In order to analyze the behaviour of the above algorithm, we assume that the algorithm produces an infinite sequence of points $\{z^k\}$.

4. Basic assumptions and convergence analysis

In this section we first state several basic assumptions. Based on these assumptions, we analyze the descentness of the generalized gradients of P_c and show that Algorithm 3.1 is well-defined and globally convergent.

Note that Algorithm 3.1 does not depend on some specific NCP function. In order to show the well-definedness and global convergence of the algorithm, we give some assumptions. In the sequel we will prove that under mild limitations, all NCP functions listed in Section 1 satisfy these assumptions. Denote

$$I(x) := \{j \mid \Phi_j(x) = 0\} \quad \text{and} \quad \bar{I}(x) := \{1, \ldots, n\} \backslash I(x) .$$

ASSUMPTION A1. $\psi(\cdot, \cdot)$ is nonnegative, semismooth and regular on $\mathbb{R}^2$.

ASSUMPTION A2. If the sequence $\{(a^k, b^k)\}$ converges and $(\xi^k, \eta^k) \in \partial\psi(a^k, b^k)$ for every k, then the sequence $\{(\xi^k, \eta^k)\}$ is bounded.

ASSUMPTION A3. For any $(a, b) \in \mathbb{R}^2$, $(\xi, \eta) \in \partial\psi(a, b)$ is always chosen to satisfy the condition:

$$\forall v \in \mathbb{R}^2 , \qquad \psi'((a, b); v) - (\xi, \eta)v \begin{cases} = 0 , & \text{if } \psi(a, b) \neq 0 , \\ \geqslant 0 , & \text{if } \psi(a, b) = 0 . \end{cases} \tag{10}$$

In Assumption A3, we do not assume that any $(\xi, \eta) \in \partial\psi(a, b)$ satisfies the condition (10). As was done in [3], we have to give mild limitations so as to ensure the algorithm to choose generalized gradients satisfying (10). In the sequel we will find that some NCP functions do not require any limitation because any generalized gradient of these NCP functions satisfies (10). This suggests us to give a stronger assumption in place of Assumption A3:

ASSUMPTION A3′. For any $(a, b) \in \mathbb{R}^2$ and any $(\xi, \eta) \in \partial\psi(a, b)$, the condition (10) always holds.

From Proposition 2.6.2(e) in [4], we have

$$\partial\Phi(x)^T \subseteq \Phi_1(x) \times \partial\Phi_2(x) \times \cdots \times \partial\Phi_n(x) .$$

By Assumption A1, it follows from Lemma 2.4(i) that for any $j \in \{1, \dots, n\}$ and any $x \in \mathbb{R}^n$, $H_j \in \partial\Phi_j(x)^T$ can be written as the form

$$H_j = \mu_j(x)E_j + \nu_j(x)\nabla F_j(x) \tag{11}$$

with

$$(\mu_j(x), \nu_j(x)) \in \partial\psi(x_j, F_j(x)) .$$

Furthermore, it follows from Lemma 2.4(ii) that for any $d \in \mathbb{R}^n$,

$$\Phi_j'(x; d) = \psi'((x_j, F_j(x)); (d_j, \nabla F_j(x)^T d)^T = \psi'((x_j, F_j(x)); p_j) ,$$

where $p_j := (d_j, \nabla F_j(x)^T d)^T$. Hence, we get

$$\Phi_j'(x; d) - H_j^T d = \psi'((x_j, F_j(x)); p_j) - (\mu_j(x), \nu_j(x))p_j , \tag{12}$$

The following result follows immediately from (10) and (12).

PROPOSITION 4.1. *Suppose that $H \in \partial\Phi(x)^T$ and that for every $j \in \{1, \dots, n\}$, the generalized gradient of ψ in the computation of H_j satisfies the condition (10). Then H satisfies*

$$\forall d \in \mathbb{R}^n , \qquad \Phi_j'(x; d) - H_j^T d \begin{cases} = 0 , & \forall j \in \bar{I}(x) , \\ \geq 0 , & \forall j \in I(x) . \end{cases} \tag{13}$$

The next proposition shows that at any noncritical point of P_c, some negative generalized gradient direction of P_c is its descent direction if suitable conditions hold.

PROPOSITION 4.2. *Suppose that $z := (x, \lambda)$ is not a critical point of P_c and that λ is nonpositive. Let $q := ((-V)^T, (-\Phi(x))^T)^T$ with $V := \nabla f(x) + H(\lambda + c\Phi(x))$ and $H \in \partial\Phi(x)^T$. If H satisfies (13), then q is a descent of P_c at z, i.e., $P_c'(z; q) < 0$.*

Proof. It is not difficult to deduce

$$\begin{aligned} P_c'(z; q) &= -\|\Phi(x)\|^2 - \nabla f(x)^T V + (\lambda + c\Phi(x))^T \Phi'(x; -V) \\ &= -\|\Phi(x)\|^2 - \nabla f(x)^T V + \sum_{j \in I(x)} \lambda_j \Phi_j'(x; -V) \\ &\quad + \sum_{j \in \bar{I}(x)} (\lambda_j + c\Phi_j(x))\Phi_j'(x; -V) \end{aligned}$$

and

$$\begin{aligned} \|V\|^2 &= \nabla f(x)^T V + (\lambda + c\Phi(x))^T H^T V \\ &= \nabla f(x)^T V + \sum_{j \in I(x)} \lambda_j H_j^T V + \sum_{j \in \bar{I}(x)} (\lambda_j + c\Phi_j(x))H_j^T V , \end{aligned}$$

which together with (13) and $\lambda \leq 0$, show

$$
\begin{aligned}
P_c'(z;q) &+ \|V\|^2 + \|\Phi(x)\|^2 \\
&= \sum_{j\in I(x)} \lambda_j(\Phi_j'(x;-V) + H_j^T V) + \sum_{j\in \bar{I}(x)} (\lambda_j + c\Phi_j(x))(\Phi_j'(x;-V) + H_j^T V) \\
&= \sum_{j\in I(x)} \lambda_j(\Phi_j'(x;-V) + H_j^T V) \\
&\leqslant 0\,.
\end{aligned}
$$

This implies

$$
\begin{aligned}
P_c'(z;q) &\leqslant -\|V\|^2 - \|\Phi(x)\|^2 \\
&= -\|q\|^2 \\
&< 0\,.
\end{aligned}
$$

□

By Proposition 4.2 and similar to the proof of Proposition 4.1 in [3], it is easy to deduce the following result.

PROPOSITION 4.3. *Suppose that Assumptions A1 and A3 hold. If Algorithm 3.1 does not stop at Step 1, then the algorithm is well-defined, i.e., there exiss a finite nonnegative integer m such that (9) holds.*

To prove the global convergence of Algorithm 3.1, we need another assumption which is associated with the NCP (1).

ASSUMPTION A4. Assume that the sequence $\{x^k\}$ converges and that for every $j\in\{1,\dots,n\}$, the sequence $\{(\mu_j^k,\nu_j^k)\}$ with $(\mu_j^k,\nu_j^k)\in\partial\psi(x_j^k,F_j(x^k))$ converges to (μ_j^*,ν_j^*). If for every j and any k, (μ_j^k,ν_j^k) satisfies the condition (10), then (μ_j^*,ν_j^*) satisfies (10).

It is obvious that Assumption A3′ implies Assumption A4 and hence Assumption A3′ can replace both Assumtions A3 and A4.

We are now ready to state the global convergence result for Algirithm 3.1 whose proof is motivated by that of Theorem 4.1 in [3], also see Theorem 2.5 in [18]. However, the convergence result here is based on the above assumptions and does not exactly depend on the strict complementarity condition.

THEOREM 4.1. *Suppose that Assumptions A1–A4 hold. Then any accumulation point $z^*=(x^*,\lambda)$ of the sequence $\{z^k\}$ generated by Algorithm 3.1 is a critical point of P_c, i.e., $0\in\partial P_c(z^*)$. Furthermore, x^* is a solution of the NCP (1).*

Proof. Assume that $\{z^k\}_{k\in K}\to z^*$. Let $H_j^k\in\partial\Phi_j(x^k)^T$. As (11), we may assume that, for every $j\in\{1,\dots,n\}$ and any k,

$$
H_j^k = \mu_j^k E_j + \nu_j^k \nabla F_j(x^k) \quad \text{with} \quad (\mu_j^k,\nu_j^k)\in\partial\psi(x_j^k,F_j(x^k))\,.
$$

It follows from Assumption A2 that $\{(\mu_j^k, \nu_j^k)\}_{k\in K}$ is bounded. Without loss of generality, we assume that $\{(\mu_j^k, \nu_j^k)\}_{k\in K} \to (\mu_j^*, \nu_j^*)$.

For every $j \in \{1, \dots, n\}$, denote

$$H_j^* := \mu_j^* E_j + \nu_j^* \nabla F_j(x^*) .$$

By Lemma 2.3(ii), we have that $(\mu_j^*, \nu_j^*) \in \partial\psi(x_j^*, F_j(x^*))$ and hence $H^* := (H_1^*, \dots, H_n^*) \in \partial\Phi(x^*)^T$. Proposition 4.1, combined with Assumptions A3 and A4, shows that H^* satisfies (13) at x^*. Moreover, from Assumption A1 and the structure of Algorithm 3.1, we deduce $\lambda^k \leqslant 0$ for all k and hence $\lambda^* \leqslant 0$.

Let

$$q^* = \begin{pmatrix} -V^* \\ -\Phi(x^*) \end{pmatrix}$$

with $V^* = \nabla f(x^*) + H^*(\lambda^* + c\Phi(x^*))$. If $0 \notin \partial P_c(z^*)$, it follows from Proposition 4.2 that q^* is a descent direction of P_c at z^*. Choose $\sigma_1 \in (\sigma, 1)$. By Proposition 4.3, there exists some constant $t^* > 0$ such that

$$P_c(z^* + t^* q^*) - P_c(z^*) \leqslant -t^* \sigma_1 \|q^*\|^2 .$$

Since $\{z^k\}_{k\in K} \to z^*$, $\{q^k\}_{k\in K} \to q^*$, $P_c(\cdot)$ is continuous and $\sigma_1 > \sigma$, there exists some positive integer k_0 such that for all $k \in K$ and $k \geqslant k_0$,

$$P_c(z^k + t^* q^k) - P_c(z^k) \leqslant -t^* \sigma \|q^k\|^2 ,$$

which implies $t^k \geqslant \beta t^*$ for all $k \in K$ and $k \geqslant k_0$. Therefore,

$$\begin{aligned} P_c(z^{k+1}) - P_c(z^k) &\leqslant -t^k \sigma \|q^k\|^2 \\ &\leqslant -\beta t^* \sigma \|q^k\|^2 . \end{aligned} \tag{14}$$

However, $\{P_c(z^k)\}$ is monotonicially decreasing and hence it is convergent. By (14), we get a contradiction: $0 \leqslant -\beta t^* \sigma \|q^*\|^2 < 0$. This shows $0 \in \partial P_c(z^*)$. It follows from Lemma 3.1(ii) that x^* is a solution of the NCP (1). □

In this theorem, we assume that an accumulation point of the sequence generated by Algorithm 3.1 exists. The existence of such an accumultion point can be guaranteed if the level sets of P_c are bounded.

The following is a direct corollary of Theorem 4.1, which does not depend on a particular nonlinear complementarity problem. Hence it implies that under Assumptions A1, A2 and A3′, the global convergence analysis does not require the strict complementarity condition.

COROLLARY 4.1. *Under Assumptions A1, A2 and A3′, the statements in Theorem 4.1 hold.*

5. Properties of several NCP functions

In this section we will give the detailed materials about the generalized gradients and the directional derivatives of function ϕ_i for every $i \in \{1, 2, 3, 5, 6\}$, which are very useful in the analysis of the next section. The next lemma shows the semismoothness of function ϕ_i for every $i \in \{1, 2, 3, 4, 5, 6\}$.

LEMMA 5.1. *For every* $i \in \{1, 2, 3, 4, 5, 6\}$, ϕ_i *is strongly semismooth and* ϕ_i^2 *is continuously differentiable.*

Proof. The results on ϕ_1 and ϕ_i for $i \in \{3, 4, 5, 6\}$ have been proved in [3] and [21], respectively. Since ϕ, $\sqrt{(\cdot)^2 + (\cdot)^2}$ and $(\cdot)_+$ are strongly semismooth functions, it follows from Lemma 2.4(iii) that ϕ_2 is strongly semismooth. It has been proved in [22] that ϕ_2^2 is continuously differentiable. □

We are now ready to deduce the generalized gradients and the directional derivatives of function ϕ_i for every $i \in \{1, 2, 3, 4, 5, 6\}$. We first rewrite function ϕ_i for every $i \in \{1, 2, 3, 4, 5, 6\}$ as follows:

$$\phi_1(a, b) = \begin{cases} a + b - \sqrt{a^2 + b^2}\,, & \text{if } a > 0,\ b > 0\,, \\ \sqrt{a^2 + b^2} - a - b\,, & \text{otherwise}\,, \end{cases}$$

$$\phi_2(a, b) = \begin{cases} a + b - \sqrt{a^2 + b^2}\,, & \text{if } a \geqslant 0,\ b \geqslant 0\,, \\ -b\,, & \text{if } a \geqslant 0,\ b < 0\,, \\ -a\,, & \text{if } a < 0,\ b \geqslant 0\,, \\ \sqrt{a^2 + b^2}\,, & \text{if } a < 0,\ b < 0\,, \end{cases}$$

$$\phi_3(a, b) = \begin{cases} \sqrt{[\phi(a, b)]^2 + \alpha[ab]^2}\,, & \text{if } ab > 0\,, \\ \phi(a, b)\,, & \text{otherwise}\,, \end{cases}$$

$$\phi_4(a, b) = \begin{cases} \sqrt{[\phi(a, b)]^2 + \alpha[ab]^4}\,, & \text{if } ab > 0\,, \\ \phi(a, b)\,, & \text{otherwise}\,, \end{cases}$$

$$\phi_5(a, b) = \begin{cases} \sqrt{[\phi(a, b)]^2 + \alpha[ab]^2}\,, & \text{if } a > 0,\ b > 0\,, \\ \phi(a, b)\,, & \text{otherwise}\,, \end{cases}$$

$$\phi_6(a, b) = \begin{cases} \sqrt{\alpha}ab\,, & \text{if } a > 0,\ b > 0\,, \\ \sqrt{[\phi(a, b)]^2 + \alpha[ab]^2}\,, & \text{if } a < 0,\ b < 0\,, \\ \phi(a, b)\,, & \text{otherwise}\,. \end{cases}$$

The following results on function ϕ_1 are due to Proposition 3.1 in [3].

PROPOSITION 5.1. *(i) If* $\phi_1(a, b) \neq 0$, *then* ϕ_1 *is continuously differeniable at* (a, b) *and*

$$\nabla\phi_1(a,b)=\begin{cases}\left(1-\dfrac{a}{\sqrt{a^2+b^2}},1-\dfrac{b}{\sqrt{a^2+b^2}}\right)^T, & \textit{if } a>0,\ b>0\,,\\ \left(\dfrac{a}{\sqrt{a^2+b^2}}-1,\dfrac{b}{\sqrt{a^2+b^2}}\right)^T, & \textit{if } a<0,\ b<0\,.\end{cases} \tag{15}$$

If $\phi_1(a,b)=0$, *then the generalized gradient of* ϕ_1 *at* (a,b) *is*

$$\partial\phi_1(a,b)=\begin{cases}\{(\rho,0)\mid\rho\in[-1,1]\}, & \textit{if } a=0,\ b>0\,,\\ \{(0,\rho)\mid\rho\in[-1,1]\}, & \textit{if } a>0,\ b=0\,,\\ \Omega_1\,, & \textit{if } a=0,\ b=0\,,\end{cases} \tag{16}$$

where $\Omega_1=co\{\Omega_{11}\cup\Omega_{12}\}$, *here*

$$\Omega_{11}=\{(1-\xi,1-\eta)\mid\xi\geqslant 0,\ \eta\geqslant 0,\ \xi^2+\eta^2=1\}\,,$$
$$\Omega_{12}=\{(\xi-1,\eta-1)\mid\xi\leqslant 0 \textit{ or } \eta\leqslant 0,\ \xi^2+\eta^2=1\}\,.$$

(ii) The directional derivative of ϕ_1 *at* (a,b) *in the direction* $v=(v_1,v_2)^T$ *is*

$$\phi_1'((a,b);v)=\left\{\begin{array}{ll}\dfrac{\partial\phi_1}{\partial a}v_1+\dfrac{\partial\phi_1}{\partial b}v_2\,, & \textit{if } \phi_1(a,b)\neq 0\,,\\ |v_1|\,, & \textit{if } a=0,\ b>0\,,\\ |v_2|\,, & \textit{if } a>0,\ b=0\,,\\ |\phi(v_1,v_2)|\,, & \textit{if } a=0,\ b=0\,.\end{array}\right\} \tag{17}$$

PROPOSITION 5.2. *(i) If* $\delta_2(a,b)\neq 0$, *then* ϕ_2 *is continuously differentiable at* (a,b) *and*

$$\nabla\phi_2(a,b)=\begin{cases}\left(1-\dfrac{a}{\sqrt{a^2+b^2}},1-\dfrac{b}{\sqrt{a^2+b^2}}\right)^T, & \textit{if } a>0,\ b>0\,,\\ (0,-1)^T\,, & \textit{if } a\geqslant 0,\ b<0\,,\\ (-1,0)^T\,, & \textit{if } a>0,\ b\geqslant 0\,,\\ \left(\dfrac{a}{\sqrt{a^2+b^2}},\dfrac{b}{\sqrt{a^2+b^2}}\right)^T, & \textit{if } a<0,\ b<0\,.\end{cases} \tag{18}$$

If $\phi_2(a,b)=0$, *then the generalized gradient of* ϕ_2 *at* (a,b) *is*

$$\partial\phi_2(a,b)=\begin{cases}\{(\rho,0)\mid\rho\in[-1,1]\}, & \textit{if } a=0,\ b>0\,,\\ \{(0,\rho)\mid\rho\in[-1,1]\}, & \textit{if } a>0,\ b=0\,,\\ \Omega_2\,, & \textit{if } a=0,\ b=0\,,\end{cases} \tag{19}$$

where $\Omega_2=co\{\Omega_{21}\cup\Omega_{22}\}$, *here*

$$\Omega_{21}=\{(1-\xi,1-\eta)\mid\xi\geqslant 0,\ \eta\geqslant 0,\ \xi^2+\eta^2=1\}\,,$$
$$\Omega_{22}=\{(\xi,\eta)\mid\xi\leqslant 0,\ \eta\leqslant 0,\ \xi^2+\eta^2=1\}\,.$$

(ii) The directional derivative of ϕ_2 *at* (a,b) *in the direction* $v=(v_1,v_2)^T$ *is*

$$\phi_2'((a,b);v)=\begin{cases}\dfrac{\partial\phi_2}{\partial a}v_1+\dfrac{\partial\phi_2}{\partial b}v_2\,, & \text{if } \phi_2(a,b)\neq 0\,,\\ |v_1|\,, & \text{if } a=0,\ b>0\,,\\ |v_2|\,, & \text{if } a>0,\ b=0\,,\\ |\phi(v_1,v_2)|\,, & \text{if } a=0,\ b=0\,.\end{cases} \tag{20}$$

Proof. (i) (18) is obvious while (19) follows directly from the definition of the generalized gradient and the expression of ϕ_2.

(ii) If $\phi_2(a,b)\neq 0$, then ϕ_2 is differentiable at (a,b) and hence the assertion obviously holds.

If $a=0$ and $b>0$, then

$$\begin{aligned}\phi_2'((0,b);v)&=\lim_{t\downarrow 0}\frac{\phi_2(tv_1,b+tv_2)-\phi_2(0,b)}{t}\\ &=\begin{cases}\lim\limits_{t\downarrow 0}\dfrac{tv_1+(b+tv_2)-\sqrt{(tv_1)^2-(b+tv_2)^2}}{t}\,, & \text{if } v_1\geq 0\,,\\ \lim\limits_{t\downarrow 0}\dfrac{-tv_1}{t}\,, & \text{if } v_1<0\,,\end{cases}\\ &=|v_1|\,.\end{aligned}$$

If $a>0$ and $b=0$, the assertion can be similarly deduced. Moreover, since $\phi_2(tv_1,tv_2)=t\phi_2(v_1,v_2)$ for any $t>0$, we get

$$\phi_2'((0,0);v)=\lim_{t\downarrow 0}\frac{\phi_2(tv_1,tv_2)}{t}=\phi_2(v_1,v_2)\,.$$ □

PROPOSITION 5.3. *(i) If $\phi_3(a,b)\neq 0$, then ϕ_3 is continuously differentiable at (a,b) and*

$$\nabla\phi_3(a,b)=\begin{cases}\left(\dfrac{1}{\phi_3}\left[\phi\dfrac{\partial\phi}{\partial a}+\alpha ab^2\right],\dfrac{1}{\phi_3}\left[\phi\dfrac{\partial\phi}{\partial b}+\alpha a^2b\right]\right)^T, & \text{if } ab>0\,,\\ \left(\dfrac{\partial\phi}{\partial a},\dfrac{\partial\phi}{\partial b}\right)^T=\left(\dfrac{a}{\sqrt{a^2+b^2}}-1,\dfrac{b}{\sqrt{a^2+b^2}}-1\right)^T, & \\ \quad \text{if } a\geq 0 \text{ and } b<0, \text{ or } a<0 \text{ and } b\geq 0\,. & \end{cases} \tag{21}$$

If $\phi_3(a,b)=0$, then the generalized gradient of ϕ_3 at (a,b) is

$$\partial\phi_3(a,b)=\begin{cases}\{(\rho,0)\,|\,\rho\in[-1,\sqrt{1+\alpha b^2}]\}\,, & \text{if } a=0,\ b>0\,,\\ \{(0,\rho)\,|\,\rho\in[-1,\sqrt{1+\alpha a^2}]\}\,, & \text{if } a>0,\ b=0\,,\\ \Omega_3\,, & \text{if } a=0,\ b=0\,,\end{cases} \tag{22}$$

where $\Omega_3=\Omega_1$.

(ii) The directional derivative of ϕ_3 at (a,b) in the direction $v=(v_1,v_2)^T$ is

$$\phi_3'((a,b);v) = \begin{cases} \frac{\partial\phi_3}{\partial a} v_1 + \frac{\partial\phi_3}{\partial b} v_2\,, & \text{if } \phi_3(a,b) \neq 0\,, \\ \sqrt{1+\alpha b^2}\, v_1\,, & \text{if } a=0,\ b>0 \text{ and } v_1>0\,, \\ -v_1\,, & \text{if } a=0,\ b>0 \text{ and } v_1 \leqslant 0\,, \\ \sqrt{1+\alpha a^2}\, v_2\,, & \text{if } a>0,\ b=0 \text{ and } v_2>0\,, \\ -v_2\,, & \text{if } a>0,\ b=0 \text{ and } v_2 \leqslant 0\,, \\ |\phi(v_1,v_2)|\,, & \text{if } a=0,\ b=0\,. \end{cases} \tag{23}$$

Proof. If $ab \neq 0$, the assertions are obvious. For $a=0$ and $b<0$, we obtain

$$\lim_{\substack{s\to 0^-\\ t\to b}} \frac{\phi(s,t)}{\phi_3(s,t)} = 1 \quad \text{and} \quad \lim_{\substack{s\to 0^-\\ t\to b}} \frac{1}{\phi_3(s,t)} = \frac{1}{2b}\,,$$

which imply

$$\lim_{\substack{s\to 0^-\\ t\to b}} \nabla\phi_3(s,t) = (-1,2)^T = \lim_{\substack{s\to 0^+\\ t\to b}} \nabla\phi_3(s,t)\,.$$

So, ϕ_3 is continuously differentiable at $(0,b)$. Similarly, we can deduce that ϕ_3 is also continuously differentiable at $(a,0)$ if $a<0$.

For $a=0$ and $b>0$, it is easy to deduce

$$\lim_{\substack{s\to 0^+\\ t\to b}} \frac{s}{\phi_3(s,t)} = \frac{1}{\sqrt{1+\alpha b^2}} \quad \text{and} \quad \lim_{\substack{s\to 0^+\\ t\to b}} \frac{\phi(s,t)}{s} = -1\,.$$

Hence, we get

$$\lim_{\substack{s\to 0^+\\ t\to b}} \nabla\phi_3(s,t) = (\sqrt{1+\alpha b^2},0)^T \quad \text{and} \quad \lim_{\substack{s\to 0^-\\ t\to b}} \nabla\phi_3(s,t) = (-1,0)^T\,,$$

which shows

$$\partial\phi_3(0,b) = \{(\rho,0)\,|\,\rho\in[-1,\sqrt{1+\alpha b^2}]\}\,.$$

Furthermore, we have

$$\begin{aligned} \phi_3'((0,b);v) &= \lim_{t\downarrow 0} \frac{\phi_3(tv_1, b+tv_2) - \phi_3(0,b)}{t} \\ &= \begin{cases} \lim_{t\downarrow 0} \frac{\sqrt{[\phi(tv_1,b+tv_2)]^2 + \alpha[tv_1(b+tv_2)]^2}}{t}\,, & \text{if } v_1>0\,, \\ \lim_{t\downarrow 0} \frac{\phi(tv_1,b+tv_2)}{t}\,, & \text{if } v_1\leqslant 0\,, \end{cases} \\ &= \begin{cases} \sqrt{1+\alpha b^2}\, v_1\,, & \text{if } v_1>0\,, \\ -v_1\,, & \text{if } v_1 \leqslant 0\,. \end{cases} \end{aligned}$$

If $a>0$ and $b=0$, the assertions can be similarly deduced.

Moreover, we deduce

$$\begin{aligned}\phi_3'((0,0);v) &= \lim_{t\downarrow 0}\frac{\sqrt{[\phi(tv_1,tv_2)]^2+\alpha[(t^2v_1v_2)_+]^2}}{t}\\ &= \lim_{t\downarrow 0}\frac{\sqrt{[t\phi(v_1,v_2)]^2+\alpha t^4[(v_1v_2)_+]^2}}{t}\\ &= |\phi(v_1,v_2)|\,.\end{aligned}$$

By the definition of the generalized gradient and noting that

$$\lim_{\substack{a\to 0^+\\ b\to 0^+}}\frac{\phi(a,b)}{\phi_3(a,b)}=-1\,,\qquad \lim_{\substack{a\to 0^-\\ b\to 0^-}}\frac{\phi(a,b)}{\phi_3(a,b)}=1\quad\text{and}\quad \lim_{\substack{a\to 0\\ b\to 0}}\frac{ab}{\phi_3(a,b)}=0\,,$$

it follows from (21) that $\Omega_3=\mathrm{co}\{\Omega_{31}\cup\Omega_{32}\cup\Omega_{33}\}$, where

$$\begin{aligned}\Omega_{31}&=\{(1-\xi,1-\eta)\,|\,\xi\geqslant 0,\ \eta\geqslant 0,\ \xi^2+\eta^2=1\}\,,\\ \Omega_{32}&=\{(\xi-1,\eta-1)\,|\,\xi\leqslant 0,\ \eta\leqslant 0,\ \xi^2+\eta^2=1\}\,,\\ \Omega_{33}&=\{(\xi-1,\eta-1)\,|\,\xi\eta\leqslant 0,\ \xi^2+\eta^2=1\}\,.\end{aligned}$$

The proof is complete. □

Note that for every $i\in\{4,5,6\}$, function ϕ_i is only a variant of ϕ_3 and $\phi_i(a,b)=\phi(a,b)$ for any $i\in\{3,4,5,6\}$ if $ab\leqslant 0$; i.e., function ϕ_i for every $i\in\{3,4,5,6\}$ is different only in the first and third quadrant. According to the previous analysis, it is not difficult to deduce the generalized gradients and the directional derivatives of function ϕ_i for every $i\in\{4,5,6\}$, we omit the process.

PROPOSITION 5.4. *(i) If $\phi_4(a,b)\neq 0$, then ϕ_4 is continuously differentiable at (a,b) and*

$$\nabla\phi_4(a,b)=\begin{cases}\left(\dfrac{1}{\phi_4}\left[\phi\dfrac{\partial\phi}{\partial a}+2\alpha ab^3b^4\right],\dfrac{1}{\phi_4}\left[\phi\dfrac{\partial\phi}{\partial b}+2\alpha a^4b^3\right]\right)^T, & \\ \quad \text{if } ab>0\,, & \\ \left(\dfrac{a}{\sqrt{a^2+b^2}}-1,\dfrac{b}{\sqrt{a^2+b^2}}-1\right)^T, & \\ \quad \text{if } a\geqslant 0 \text{ and } b<0, \text{ or } a<0 \text{ and } b\geqslant 0\,. & \end{cases}\tag{24}$$

If $\phi_4(a,b)=0$, then the generalized gradient of ϕ_4 at (a,b) is

$$\partial\phi_4(a,b)=\begin{cases}\{(\rho,0)\,|\,\rho\in[-1,1]\}\,, & \text{if } a=0,\ b>0\,,\\ \{(0,\rho)\,|\,\rho\in[-1,1]\}\,, & \text{if } a>0,\ b=0\,,\\ \Omega_4\,, & \text{if } a=0,\ b=0\,,\end{cases}\tag{25}$$

where $\Omega_4=\Omega_1$.

(ii) The directional derivative of ϕ_4 at (a,b) in the direction $v=(v_1,v_2)^T$ is

$$\phi_4'((a,b);v)=\begin{cases}\dfrac{\partial\phi_4}{\partial a}v_1+\dfrac{\partial\phi_4}{\partial b}v_2\,, & \text{if } \phi_4(a,b)\neq 0\,,\\ |v_1|\,, & \text{if } a=0,\ b>0\,,\\ |v_2|\,, & \text{if } a>0,\ b=0\,,\\ |\phi(v_1,v_2)|\,, & \text{if } a=0,\ b=0\,.\end{cases} \tag{26}$$

PROPOSITION 5.5. *(i) If* $\phi_5(a,b)\neq 0$, *then* ϕ_5 *is continuously differentiable at* (a,b) *and*

$$\nabla\phi_5(a,b)=\begin{cases}\left(\dfrac{1}{\phi_5}\left[\phi\dfrac{\partial\phi}{\partial a}+\alpha ab^2\right],\dfrac{1}{\phi_5}\left[\phi\dfrac{\partial\phi}{\partial b}+\alpha a^2 b\right]\right)^T, \\ \quad \text{if } a>0,\ b>0\,,\\ \left(\dfrac{a}{\sqrt{a^2+b^2}}-1,\dfrac{b}{\sqrt{a^2+b^2}}-1\right)^T, \\ \quad \text{if } a<0 \text{ or } b<0\,.\end{cases} \tag{27}$$

If $\phi_5(a,b)=0$, *then the generalized gradient of* ϕ_5 *at* (a,b) *is*

$$\partial\phi_5(a,b)=\begin{cases}\{(\rho,0)\,|\,\rho\in[-1,\sqrt{1+\alpha b^2}]\}\,, & \text{if } a=0,\ b>0\,,\\ \{(0,\rho)\,|\,\rho\in[-1,\sqrt{1+\alpha a^2}]\}\,, & \text{if } a>0,\ b=0\,,\\ \Omega_5\,, & \text{if } a=0,\ b=0\,,\end{cases} \tag{28}$$

where $\Omega_5=\Omega_1$.
(ii) The directional derivative of ϕ_5 *at* (a,b) *in the direction* $v=(v_1,v_2)^T$ *is*

$$\phi_5'((a,b);v)=\begin{cases}\dfrac{\partial\phi_5}{\partial a}v_1+\dfrac{\partial\phi_5}{\partial b}v_2\,, & \text{if } \phi_5(a,b)\neq 0\,,\\ \sqrt{1+\alpha b^2}\,v_1\,, & \text{if } a=0,\ b>0 \text{ and } v_1>0\,,\\ -v_1\,, & \text{if } a=0,\ b>0 \text{ and } v_1\leqslant 0\,,\\ \sqrt{1+\alpha a^2}\,v_2\,, & \text{if } a>0,\ b=0 \text{ and } v_2>0\,,\\ -v_2\,, & \text{if } a>0,\ b=0 \text{ and } v_2\leqslant 0\,,\\ |\phi(v_1,v_2)|\,, & \text{if } a=0,\ b=0\,.\end{cases} \tag{29}$$

PROPOSITION 5.6. *(i) If* $\phi_6(a,b)\neq 0$, *then* ϕ_6 *is continuously differentiable at* (a,b) *and*

$$\nabla\phi_6(a,b)=\begin{cases}(\sqrt{\alpha}b,\sqrt{\alpha}a)^T\,, \quad \text{if } a>0,\ b>0\,,\\ \left(\dfrac{1}{\phi_6}\left[\phi\dfrac{\partial\phi}{\partial a}+\alpha ab^2b^2\right],\dfrac{1}{\phi_6}\left[\phi\dfrac{\partial\phi}{\partial b}+\alpha a^2 b\right]\right)^T, \\ \quad \text{if } a<0,\ b<0\,,\\ \left(\dfrac{a}{\sqrt{a^2+b^2}}-1,\dfrac{b}{\sqrt{a^2+b^2}}-1\right)^T, \\ \quad \text{if } a\geqslant 0 \text{ and } b<0, \text{ or } a<0 \text{ and } b\geqslant 0\,.\end{cases} \tag{30}$$

If $\phi_6(a,b)=0$, *then the generalized gradient of* ϕ_6 *at* (a,b) *is*

$$\partial\phi_6(a,b) = \begin{cases} \{(\rho,0) \mid \rho \in [-1, \sqrt{\alpha}b]\}, & \text{if } a=0,\ b>0, \\ \{(0,\rho) \mid \rho \in [-1, \sqrt{\alpha}a]\}, & \text{if } a>0,\ b=0, \\ \Omega_6, & \text{if } a=0,\ b=0, \end{cases} \tag{31}$$

where $\Omega_6 = \text{co}\{\{(0,0)\} \cup \Omega_{61} \cup \Omega_{62}\}$, *here*

$$\Omega_{61} = \{(\xi-1, \eta-1) \mid \xi \leqslant 0,\ \eta \leqslant 0,\ \xi^2+\eta^2=1\},$$
$$\Omega_{62} = \{(\xi-1, \eta-1) \mid \xi\eta \leqslant 0,\ \xi^2+\eta^2=1\}.$$

(ii) The directional derivative of ϕ_6 *at* (a,b) *in the direction* $v=(v_1,v_2)^T$ *is*

$$\phi_6'((a,b);v) = \begin{cases} \frac{\partial\phi_6}{\partial a}v_1 + \frac{\partial\phi_6}{\partial b}v_2, & \text{if } \phi_6(a,b)\neq 0, \\ \sqrt{\alpha}bv_1, & \text{if } a=0,\ b>0 \text{ and } v_1>0, \\ -v_1, & \text{if } a=0,\ b>0 \text{ and } v_1\leqslant 0, \\ \sqrt{\alpha}av_2, & \text{if } a>0,\ b=0 \text{ and } v_2>0, \\ -v_2, & \text{if } a>0,\ b=0 \text{ and } v_2\leqslant 0, \\ [\phi(v_1,v_2)]_+, & \text{if } a=0,\ b=0. \end{cases} \tag{32}$$

6. Verification of assumptions

In Section 4 we have proven that Algorithm 3.1 converges globally if suitable assumptions hold. In this section we will show that function ϕ_i for every $i \in \{1,\dots,6\}$ satisfies Assumptions A1 and A2. For every $i \in \{1,\dots,5\}$, function ϕ_i satisfies Assumptions A3 and A4 if restricting the choice of the generalized gradients of ϕ_i at the origin and assuming that the strict complementarity condition holds at the limit point, while function ϕ_6 satisfies Assumption A3′ without any condition.

The next theorem follows immediately from Lemma 5.1 and Propositions 5.1–5.6.

THEOREM 6.1. *For every* $i \in \{1,\dots,6\}$, $\psi = \phi_i$ *satisfies Assumptions A1 and A2.*

Now we turn to deduce the conditions under which Assumption A3 holds. To this end, we first show the next proposition.

PROPOSITION 6.1. *For every* $i \in \{1,\dots,6\}$, $\forall (a,b) \in \mathbb{R}^2$ *and* $\forall(\xi,\eta) \in \partial\phi_i(a,b)$, *one has*

$$\phi_i'((a,b);v) - (\xi,\eta)v \begin{cases} =0, & \text{if } \phi_i(a,b)\neq 0, \\ \geqslant 0, & \text{if } \phi_i(a,b)=0 \text{ but } (a,b)\neq(0,0). \end{cases}$$

Proof. The proof is based on Propositions 5.1–5.6.

If $\phi_i(a,b)\neq 0$. Then ϕ_i is continuously differentiable at (a,b) and

$$\phi_i'((a,b);v)=\frac{\partial\phi_i}{\partial a}v_1+\frac{\partial\phi_i}{\partial b}v_2 \quad \text{and} \quad \partial\phi_i(a,b)=\left\{\left(\frac{\partial\phi_i}{\partial a},\frac{\partial\phi_i}{\partial b}\right)\right\},$$

which imply $\phi_i'((a,b);v)-(\xi,\eta)v=0$.

If $a=0$ and $b>0$. For $i=1, 2, 4$, we have

$$\phi_i'((a,b);v)-(\xi,\eta)v=|v_1|-\xi v_1\geqslant 0, \quad \forall\xi\in[-1,1].$$

For $i=3, 5$, we have

$$\phi_i'((a,b);v)-(\xi,\eta)v=\begin{cases}\sqrt{1+\alpha b^2}v_1-\xi v_1, & \text{if } v_1>0,\\ -v_1-\xi v_1, & \text{if } v_1\leqslant 0,\end{cases}$$
$$\geqslant 0, \quad \forall\xi\in[-1,\sqrt{1+\alpha b^2}].$$

For $i=6$, we have

$$\phi_i'((a,b);v)-(\xi,\eta)v=\begin{cases}\sqrt{\alpha}bv_1-\xi v_1, & \text{if } v_1>0,\\ -v_1-\xi v_1, & \text{if } v_1\leqslant 0,\end{cases}$$
$$\geqslant 0, \quad \forall\xi\in[-1,\sqrt{\alpha}b].$$

If $a>0$ and $b=0$. In this case, the proof is similar to that of the previous case. The assertion is proved. □

From the above proposition, in order to verify Assumption A3, we only need to consider the case: $(a,b)=(0,0)$. The following basic lemma is due to Proposition 3.5 in [3] and will play a very important role in the analysis of this section.

LEMMA 6.1. *For any* $(s,t)\in\Omega_0$, *we have*

$$|\phi(a,b)|-as-bt\geqslant 0, \tag{33}$$

where

$$\Omega_0:=\{(\xi,\eta)\,|\,(\xi+1)^2+(\eta+1)^2\leqslant 1\}\cup\{(\xi,\eta)\,|\,-1\leqslant\xi\leqslant 0,\ -1\leqslant\eta\leqslant 0\}.$$

Denote

$$\tilde{\Omega}_2:=\Omega_2\cap\{(\xi,\eta)\,|\,\xi\leqslant 0,\ \eta\leqslant 0\} \quad \text{and} \quad \tilde{\Omega}_i:=\Omega_0 \quad \text{for} \quad i=1,3,4,5,6.$$

It is obvious that $\tilde{\Omega}_i\subseteq\Omega_i$ for every $i\in\{1,\ldots,5\}$ and $\tilde{\Omega}_6=\Omega_6$. From the above basic lemma and Propositions 5.1–5.6, we deduce the next proposition.

PROPOSITION 6.2. *For every* $i\in\{1,\ldots,6\}$ *and* $\forall v\in\mathbb{R}^2$, *one has*

$$\phi_i'((0,0);v)-(\xi,\eta)v\geqslant 0, \tag{34}$$

whenever $(\xi,\eta)\in\tilde{\Omega}_i$.

Proof. For every $i \in \{1, 3, 4, 5\}$, since $\tilde{\Omega}_i = \Omega_0$ and $\phi_i'((0,0); v) = |\phi(v_1, v_2)|$, it follows directly from Lemma 6.1 that

$$\phi_i'((0,0); v) - (\xi, \eta)v = |\phi(v_1, v_2)| - \xi v_1 - \eta v_2 \geqslant 0\,, \quad \forall(\xi, \eta) \in \tilde{\Omega}_i\,.$$

Moreover, it is not difficult to deduce

$$\begin{aligned}\phi_6'((0,0); v) - (\xi, \eta)v &= \begin{cases} -\xi v_1 - \eta v_2\,, & \text{if } v_1 \geqslant 0,\ v_2 \geqslant 0\,,\\ |\phi(v_1, v_2)| - \xi v_1 - \eta v_2\,, & \text{otherwise}\,,\end{cases}\\ &\geqslant 0\,, \quad \forall(\xi, \eta) \in \tilde{\Omega}_6\,;\end{aligned}$$

and

$$\begin{aligned}\phi_2'((0,0); v) - (\xi, \eta)v &= \begin{cases} |\phi(v_1, v_2)| - \xi v_1 - \eta v_2\,, & \text{if } v_1 \geqslant 0,\ v_2 \geqslant 0\,,\\ -\xi v_1 - (1+\eta)v_2\,, & \text{if } v_1 \geqslant 0,\ v_2 < 0\,,\\ -(1+\xi)v_1 - \eta v_2\,, & \text{if } v_1 < 0,\ v_2 \geqslant 0\,,\\ \sqrt{v_1^2 + v_2^2} - \xi v_1 - \eta v_2\,, & \text{if } v_1 < 0,\ v_2 < 0\,,\end{cases}\\ &\geqslant 0\,, \quad \forall(\xi, \eta) \in \tilde{\Omega}_2\,,\end{aligned}$$

where the last item above is due to

$$\begin{aligned}v_1^2 + v_2^2 - (\xi v_1 + \eta v_2)^2 &= (1 - \xi^2)v_1^2 + (1 - \eta^2)v_2^2 - 2\xi\eta v_1 v_2\\ &\geqslant \eta^2 v_1^2 + \xi^2 v_2^2 - 2\xi\eta v_1 v_2\\ &\geqslant 0\,, \quad \in (\xi, \eta) \in \tilde{\Omega}_2\,.\end{aligned}$$

The proof is complete. □

The next theorem is a direct consequence of Propositions 6.1 and 6.2.

THEOREM 6.2. *For every $i \in \{1, \ldots, 5\}$, function $\psi = \phi_i$ satisfies Assumption A3 whenever the generalized gradients of ϕ_i at the origin are chosen within $\tilde{\Omega}_i$.*

Assume that $\{x^k\} \to x^*$. If for some $j \in \{1, \ldots, n\}$, the sequence $\{(\mu_j^k, \nu_j^k)\}$ with $(\mu_j^k, \nu_j^k) \in \partial\psi(x_j^k, F_j(x^k))$ converges to (μ_j^*, ν_j^*) and for any k, (μ_j^k, ν_j^k) satisfies (10). It follows from Lemma 2.3(ii) that $(\mu_j^*, \nu_j^*) \in \partial\psi(x_j^*, F_j(x^*))$. However, Proposition 6.2 implies that (μ_j^*, ν_j^*) does not necessarily satisfy (10). If we suppose that the strict complementarity condition holds at x^*, i.e.,

$$I_0(x^*) = \emptyset\,,$$

where $I_0(x) := \{j \mid x_j = 0,\ F_j(x) = 0\}$, then (μ_j^*, ν_j^*) satisfies (10). Hence, we deduce the next theorem.

THEOREM 6.3. *For every $i \in \{1, 5\}$, function $\psi = \phi_i$ satisfies Assumption A4 whenever the generalized gradients of ϕ_i at the origin are chosen within $\tilde{\Omega}_i$ and the strict complementarity condition holds at the limit point of $\{x^k\}$.*

Since $\tilde{\Omega}_6 = \Omega_6$, in the proof of Proposition 6.2, we have actually proved that (34) holds for any $(\xi, \eta) \in \Omega_6$. This means that we have deduced the next theorem.

THEOREM 6.4. *Function $\psi = \phi_6$ satisfies Assumption A3′.*

From the analysis above, all NCP functions listed in Section 1 satisfy the needed assumptions so as to ensure the global convergence of the LG method whenever the generalized gradients of the corresponding ϕ_i at the origin are chosen within $\tilde{\Omega}_i$ and the strict complementarity condition holds at the limit point of $\{x^k\}$. However, due to $\tilde{\Omega}_6 = \Omega_6$, Theorem 6.4 shows that if $\psi = \phi_6$ is used in the LG method, then the strict complementarity condition and the limitation in the choice of the generalized Jacobians of ϕ_6 are unnecessary.

On the other hand, the above analysis also answers positively the two questions presented in the introduction. If some NCP function satisfies Assumptions A1–A4, then it can be used in the LG method. Moreover, if it satisfies Assumption A3′ instead of Assumptions A3 and A4, then the strict complementarity condition can be removed from the LG method.

7. Final remarks

In this paper we proposed an LG-type algorithm model for solving the nonlinear complementarity problem. In particular our algorithm model does not depend on some specific NCP function. We find that under our algorithm model, in order to guarantee the global convergence of the algorithm, some NCP functions require mild limitations while other NCP functions do not require any limitation. For example, if we use function ϕ_6 in our algorithm model, then the global convergence of the algorithm does not need the strict complementarity condition whereas if we use other functions., even its variant ϕ_i for $i = 3, 4, 5$, we cannot remove this condition. Actually, we have deduced the assertions:

- NCP functions satisfying Assumptions A1–A4 are applicable to the LG method.
- NCP functions satisfying Assumptions A1, A2 and A3′ are applicable to the LG method without requiring the strict complementarity condition at the solution.

Meanwhile, we studied in details six NCP functions applicable to the LG method. Based on the observation to these NCP functions, we can construct a new NCP function possibly with the same properties as ϕ_6. Set

$$\phi_7(a, b) := \sqrt{[(-a)_+]^2 + [(-b)_+]^2 + \alpha[(ab)_+]^2}\,, \quad \alpha > 0\,,$$
$$= \begin{cases} \sqrt{\alpha}ab\,, & \text{if } a \geqslant 0,\ b \geqslant 0\,, \\ -b, & \text{if } a \geqslant 0,\ b < 0\,, \\ -a\,, & \text{if } a < 0,\ b \geqslant 0\,, \\ \sqrt{a^2 + b^2 + \alpha(ab)^2}\,, & \text{if } a < 0,\ b < 0\,. \end{cases}$$

We observe that the function ϕ_7 with $\alpha = 1$ can also be reformulated from a merit function studied in [11]. It is not difficult to deduce the following results:

- ϕ_7 is strongly semismooth and ϕ_7^2 is continuously differentiable.
- If $\phi_7(a, b) \neq 0$, then ϕ_7 is continuously differentiable at (a, b).
- If $\phi_7(a, b) = 0$, then the generalized gradient of ϕ_7 at (a, b) is

$$\partial\phi_7(a, b) = \begin{cases} \{(\rho, 0) \mid \rho \in [-1, \sqrt{\alpha}b]\}, & \text{if } a = 0,\ b > 0, \\ \{(0, \rho) \mid \rho \in [-1, \sqrt{\alpha}a]\}, & \text{if } a > 0,\ b = 0, \\ \Omega_7, & \text{if } a = 0,\ b = 0, \end{cases}$$

  where $\Omega_7 := \{(\xi, \eta) \mid \xi \leqslant 0,\ \eta \leqslant 0,\ \xi^2 + \eta^2 \leqslant 1\}$.
- The directional derivative of ϕ_7 at (a, b) in the direction $v = (v_1, v_2)^T$ is

$$\phi_7'((a, b); v) = \begin{cases} \dfrac{\partial\phi_7}{\partial a} v_1 + \dfrac{\partial\phi_7}{\partial b} v_2, & \text{if } \phi_7(a, b) \neq 0, \\ \sqrt{\alpha} b v_1, & \text{if } a = 0,\ b > 0 \text{ and } v_1 > 0, \\ -v_1, & \text{if } a = 0,\ b > 0 \text{ and } v_1 \leqslant 0, \\ \sqrt{\alpha} a v_2, & \text{if } a > 0,\ b = 0 \text{ and } v_2 > 0, \\ -v_2, & \text{if } a > 0,\ b = 0 \text{ and } v_2 \leqslant 0, \\ \sqrt{[(-v_1)_+]^2 + [(-v_2)_+]^2}, & \text{if } a = 0,\ b = 0. \end{cases}$$

- Function $\psi = \phi_7$ satisfies Assumptions A1, A2 and A3′.

The last item above shows that: if we use $\psi = \phi_7$ in our algorithm model, then the global convergence of the algorithm does not require the strict complementarity condition.

An interesting question is whether the LG-type algorithm model presented in this paper can be generalized to the mixed complementarity problem, the box constrained variational inequality problem even the general variational inequality problem, or whether it is possible to avoid the unboundedness of the multiplier sequence in the implementation of the algorithm.

Moreover, in this paper we do not discuss the conditions under which the iterates generated by the LG method are bounded. This is also a very important and interesting question. We will leave these questions as further research topics.

Acknowledgments

This research was partially supported by the Research Grant Council of Hong Kong (Grant 491101) and the National Natural Science Foundation of China (Grant 10171030). The authors would like to thank the anonymous referees and Professor A.M. Rubinov for their helpful comments and suggestions.

References

[1] Chen, B., Chen, X. and Kanzow, C. (2000), A Penalized Fischer-Burmeister NCP-Function, *Math. Prog.* 88, 211–216.
[2] Chen, X., Qi, L. and Sun, D. (1998), Global and Superlinear Convergence of the Smoothing Newton Method and its Application to General Box Constrained Variational Inequalities, *Math. Comp.* 67, 519–540.
[3] Chen, X., Qi, L. and Yang, Y.F. (2002), Lagrangian Globalization Methods for Nonlinear Complementarity Problems, *J. Optim. Theory Appl.* 112, 77–95.
[4] Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.
[5] De Luca, T., Facchinei, F. and Kanzow, C. (1996), A Semismooth Equation Approach to the Solution of Nonlinear Complementarity Problem, *Math. Prog.* 75, 407–439.
[6] Ferris, M.C. and Pang, J.S. (1997), Engineering and Economic Applications of Complementarity Problems, *SIAM Review* 39, 669–713.
[7] Fischer, A. (1992), A Special Newton-Type Optimization Method, *Optimization* 24, 269–284.
[8] Fischer, A. (1997), Solution of Monotone Complementarity Problems with Locally Lipschitzian Functions, *Math. Prog.* 76, 513–532.
[9] Harker, P.T. and Pang, S.J. (1990), Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications, *Math. Prog.* 48, 161–220.
[10] Kanzow, C. and Pieper, H. (1999), Jacobian Smoothing Methods for Nonlinear Complementarity Problems, *SIAM J. Optim.* 9, 342–373.
[11] Kanzow, C., Yamashita, N. and Fukushima, M. 91997), New NCP-Functions and their Properties, *J. Optim. Theory Appl.* 94, 115–135.
[12] Luo, Z.-Q. and Tseng, P. (1997), A New Class of Merit Functions for the Nonlinear Complementarity Problem, in Ferris, M.C. and Pang, J.S. (eds.), *Complementarity and Variational Problems: State of the Arts*, SIAM, Philadelphia, pp. 204–225.
[13] Mifflin, R. (1977), Semismooth and Semiconvex Functions in Constrained Optimization, *SIAM J. Control Optim.* 15, 957–972.
[14] Moré, J.J. (1996), Global Methods for Nonlinear Complementarity Problems, *Math. Oper. Res.* 21, 589–614.
[15] Nazareth, J.L. (1996), Lagrangian Globalization: Solving Nonlinear Equations via Constrained Optimization, in Renegar, J., Shub, M. and Smale, S. (eds.), *The Mathematics of Numerical Analysis*, the American Mathematical Society, Providence, RI, *Lectures in Applied Mathematics* 32, pp. 533–542.
[16] Nazareth, J.L. and Qi, L. (1996), Globalization of Newton's Methods for Solving Nonlinear Equations, J. Numer. Algebra App;. 3, 239–249.
[17] Pang, J.S. (1995), Complementarity problems, in Horst, R. and Pardalos, P. (eds.), *Handbook of Global Optimization*, Kluwer Academic Publishers, Boston, MA, pp. 271–338.
[18] Polak, E., Mayne, D.Q. and Wardi, Y. (1983), On the Extension of Constrained Optimization Algorithms from Differentiable to Nondifferentiable Problems, *SIAM J. Control Optim.* 21, 179–204.
[19] Qi, L. and Sun, J. (1993), A Nonsmooth Version of Newton's Methods, *Math. Prog.* 58, 353–367.
[20] Sun, D. (1999), A Regularization Newton Method for Solving Nonlinear Complementarity Problems, *Appl. Math. Optim.* 40, 315–339.
[21] Sun, D. and Qi, L. (1999), On NCP-Functions, *Comput. Optim. Appl.* 13, 201–220.
[22] Sun, D. and Womersley, R.S. (1999), A New Unconstrained Differentiable Merit Function for Box Constrained Variational Inequality Problems and a Damped Gauss-Newton Method, *SIAM J. Optim.* 9, 388–413.

[23] Yamashita, N. and Fukushima, M. (1997), Modified Newton Methods for Solving a Semismooth Reformulation of Monotone Complementarity Problems, *Math*. *Prog*. 76, 469–491.

[24] Yamashita, N. and Fukushima, M. (1998), A New Merit Function and a Descent Method for Semidefinite Complementarity Problems, in Fukushima and Qi, L. (eds.), *Reformulation*: *Nonsmooth*, *Piecewise Smooth*, *Semismooth and Smoothing Methods*, Kluwer Academic Publishers, Dordrecht, pp. 381–404.