A family of NCP functions and a descent method for the nonlinear complementarity problem

Jein-Shan Chen · Shaohua Pan

Abstract In last decades, there has been much effort on the solution and the analysis of the nonlinear complementarity problem (NCP) by reformulating NCP as an unconstrained minimization involving an NCP function. In this paper, we propose a family of new NCP functions, which include the Fischer-Burmeister function as a special case, based on a p-norm with p being any fixed real number in the interval $(1, +\infty)$, and show several favorable properties of the proposed functions. In addition, we also propose a descent algorithm that is indeed derivative-free for solving the unconstrained minimization based on the merit functions from the proposed NCP functions. Numerical results for the test problems from MCPLIB indicate that the descent algorithm has better performance when the parameter p decreases in $(1, +\infty)$. This implies that the merit functions associated with $p \in (1, 2)$, for example $p = 1.5$, are more effective in numerical computations than the Fischer-Burmeister merit function, which exactly corresponds to $p = 2$.

Keywords NCP · NCP function · Merit function · Descent method

1 Introduction

The nonlinear complementarity problem (NCP) is to find a point $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad \langle x, F(x) \rangle = 0, \tag{1}$$

J.-S. Chen (✉)
Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan
e-mail: jschen@math.ntnu.edu.tw

S. Pan
School of Mathematical Sciences, South China University of Technology, Guangzhou 510641, China
e-mail: shhpan@scut.edu.cn

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F = (F_1, F_2, \ldots, F_n)^T$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$. We assume that F is continuously differentiable throughout this paper. The NCP has attracted much attention due to its various applications in operations research, economics, and engineering [6, 12, 18].

There have been many methods proposed for solving the NCP [9, 12, 18]. Among which, one of the most popular and powerful approaches that has been studied intensively recently is to reformulate the NCP as a system of nonlinear equations [17, 24] or as an unconstrained minimization problem [5, 7, 10, 14–16, 23]. Such a function that can constitute an equivalent unconstrained minimization problem for the NCP is called a merit function. In other words, a merit function is a function whose global minima are coincident with the solutions of the original NCP. For constructing a merit function, the class of functions, so-called NCP-functions and defined as below, serves an important role.

Definition 1.1 *A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an NCP-function if it satisfies*

$$\phi(a, b) = 0 \quad \Longleftrightarrow \quad a \geq 0, \ b \geq 0, \ ab = 0. \tag{2}$$

Over the past two decades, a variety of NCP-functions have been studied, see [9, 20] and references therein. Among which, a popular NCP-function intensively studied recently is the well-known Fischer-Burmeister NCP-function [7, 8] defined as

$$\phi(a, b) = \sqrt{a^2 + b^2} - (a + b). \tag{3}$$

With the above characterization of ϕ, the NCP is equivalent to a system of nonsmooth equations:

$$\Phi(x) = \begin{pmatrix} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{pmatrix} = 0. \tag{4}$$

Then the function $\Psi : \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\Psi(x) := \frac{1}{2} \|\Phi(x)\|^2 = \frac{1}{2} \sum_{i=1}^n \phi(x_i, F_i(x))^2 \tag{5}$$

is a merit function for the NCP, i.e., the NCP can be recast as an unconstrained minimization:

$$\min_{x \in \mathbb{R}^n} \Psi(x). \tag{6}$$

In this paper, we propose and investigate a family of new NCP functions based on the Fischer-Burmeister function (3). In particular, we define $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ by

$$\phi_p(a, b) := \|(a, b)\|_p - (a + b), \tag{7}$$

where p is any fixed real number in the interval $(1, +\infty)$ and $\|(a, b)\|_p$ denotes the p-norm of (a, b), i.e., $\|(a, b)\|_p = \sqrt[p]{|a|^p + |b|^p}$. In other words, in the function ϕ_p,

we replace the 2-norm of (a, b) in the Fischer-Burmeister function (3) by a more general p-norm with $p \in (1, +\infty)$. The function ϕ_p is still an NCP-function as was noted in Tseng's paper [21]. Nonetheless, to our knowledge, there was no further study on this family of NCP functions except for $p = 2$. We aim to explore and study properties of ϕ_p in this paper. More specifically, we define $\psi_p : \mathbb{R}^2 \to \mathbb{R}_+$ by

$$\psi_p(a, b) := \frac{1}{2}|\phi_p(a, b)|^2. \tag{8}$$

For any given $p > 1$, the function ψ_p is a nonnegative NCP-function and smooth on $\mathbb{R}^2$ as will be seen in Sect. 3. Analogous to Φ, the function $\Phi_p : \mathbb{R}^n \to \mathbb{R}^n$ given as

$$\Phi_p(x) = \begin{pmatrix} \phi_p(x_1, F_1(x)) \\ \vdots \\ \phi_p(x_n, F_n(x)) \end{pmatrix} \tag{9}$$

yields a family of merit functions $\Psi_p : \mathbb{R}^n \to \mathbb{R}$ for the NCP for which

$$\Psi_p(x) := \frac{1}{2}\|\Phi_p(x)\|^2 = \frac{1}{2}\sum_{i=1}^{n}\phi_p(x_i, F_i(x))^2 = \sum_{i=1}^{n}\psi_p(x_i, F_i(x)). \tag{10}$$

As will be seen later, Ψ_p for any given $p > 1$ is a continuously differentiable merit function for the NCP. Therefore, classical iterative methods such as Newton method can be applied to the unconstrained smooth minimization of the NCP, i.e.,

$$\min_{x \in \mathbb{R}^n} \Psi_p(x). \tag{11}$$

On the other hand, derivative-free methods [22] have also attracted much attention which do not require computation of derivatives of F. Derivative-free methods, taking advantages of particular properties of a merit function, are suitable for problems where the derivatives of F are not available or expensive.

In this paper, we also study a derivative-free descent algorithm for solving the NCP based on the merit function Ψ_p. The algorithm is shown to be convergent for strongly monotone NCPs. In addition, we also do numerical experiments with three specific merit functions $\Psi_{1.5}$, Ψ_2 and Ψ_3 for the test problems from MCPLIB. Numerical results show that the descent algorithm has better performance as p decreases in the interval $(1, +\infty)$. This means that a more effective NCP function than the Fischer-Burmeister function, at lest in numerical computations, can be obtained by setting $p \in (1, 2)$ in $\phi_p(a, b)$.

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors and ^{T} denotes transpose. For any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x. For any differentiable mapping $F = (F_1, \ldots, F_m)^T : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x) \cdots \nabla F_m(x)]$ denotes the transpose Jacobian of F at x. We denote by $\|x\|_p$ the p-norm of x and by $\|x\|$ the Euclidean norm of x. In addition, unless otherwise stated, we always assume p in the sequel is any fixed real number in $(1, +\infty)$.

2 Preliminaries

In this section, we recall some background concepts and materials which will play an important role in the subsequent analysis.

Definition 2.1 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$, then*

(a) *F is monotone if $\langle x - y, F(x) - F(y) \rangle \geq 0$, for all $x, y \in \mathbb{R}^n$.*

(b) *F is strictly monotone if $\langle x - y, F(x) - F(y) \rangle > 0$, for all $x, y \in \mathbb{R}^n$ and $x \neq y$.*

(c) *F is strongly monotone with modulus $\mu > 0$ if $\langle x - y, F(x) - F(y) \rangle \geq \mu \|x - y\|^2$, for all $x, y \in \mathbb{R}^n$.*

(d) *F is a P_0-function if $\max_{\substack{1 \leq i \leq n \\ x_i \neq y_i}} (x_i - y_i)(F_i(x) - F_i(y)) \geq 0$, for all $x, y \in \mathbb{R}^n$ and $x \neq y$.*

(e) *F is a P-function if $\max_{1 \leq i \leq n}(x_i - y_i)(F_i(x) - F_i(y)) > 0$, for all $x, y \in \mathbb{R}^n$ and $x \neq y$.*

(f) *F is a uniform P-function with modulus $\mu > 0$ if $\max_{1 \leq i \leq n}(x_i - y_i)(F_i(x) - F_i(y)) \geq \mu \|x - y\|^2$, for all $x, y \in \mathbb{R}^n$.*

(g) *$\nabla F(x)$ is uniformly positive definite with modulus $\mu > 0$ if $d^T \nabla F(x) d \geq \mu \|d\|^2$, for all $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$.*

(h) *F is Lipschitz continuous if there exists a constant $L > 0$ such that $\|F(x) - F(y)\| \leq L \|x - y\|$, for all $x, y \in \mathbb{R}^n$.*

From the above definitions, it is obvious that strongly monotone functions are strictly monotone, and strictly monotone functions are monotone. Moreover, F is a P_0-function if F is monotone and F is a uniform P-function with modulus $\mu > 0$ if F is strongly monotone with modulus $\mu > 0$. In addition, when F is continuously differentiable, we have the following conclusions:

1. F is monotone if and only if $\nabla F(x)$ is positive semidefinite for all $x \in \mathbb{R}^n$.
2. F is strictly monotone if $\nabla F(x)$ is positive definite for all $x \in \mathbb{R}^n$.
3. F is strongly monotone if and only if $\nabla F(x)$ is uniformly positive definite.

Next, we recall the definition of P_0-matrix and P-matrix.

Definition 2.2 *A matrix $M \in \mathbb{R}^{n \times n}$ is a*

(a) *P_0-matrix if each of its principal minors is nonnegative.*

(b) *P-matrix if each of its principal minors is positive.*

It is obvious that every P-matrix is also a P_0-matrix. Furthermore, it is known that the Jacobian of every continuously differentiable P_0-function is a P_0-matrix.

Finally, we state one of the characterizations of P_0-matrices that will be used later, and for more properties about P-matrix and P_0-matrix, please refer to [4].

Lemma 2.1 *A matrix $M \in \mathbb{R}^{n \times n}$ is a P_0-matrix if and only if for every nonzero vector x there exists an index i such that $x_i \neq 0$ and $x_i(Mx)_i \geq 0$.*

3 A family of NCP functions and their properties

In this section, we study a family of NCP functions ϕ_p defined as (7) with $p > 1$, which are indeed variants of Fischer-Burmeister function, and show that these functions have several favorable properties analogous to what Fischer-Burmeister function has. We first present some similar properties of ϕ_p to those for Fischer-Burmeister function.

Proposition 3.1 *Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be defined as (7) with p being any fixed real number in the interval $(1, +\infty)$. Then*

(a) *ϕ_p is an NCP-function, i.e., it satisfies (2).*
(b) *ϕ_p is sub-additive, i.e., $\phi_p(w + w') \leq \phi_p(w) + \phi_p(w')$ for all $w, w' \in \mathbb{R}^2$.*
(c) *ϕ_p is positive homogeneous, i.e., $\phi_p(\alpha w) = \alpha \phi_p(w)$ for all $w \in \mathbb{R}^2$ and $\alpha \geq 0$.*
(d) *ϕ_p is convex, i.e., $\phi_p(\alpha w + (1 - \alpha)w') \leq \alpha \phi_p(w) + (1 - \alpha)\phi_p(w')$ for all $w, w' \in \mathbb{R}^2$ and $\alpha \in 0$.*
(e) *ϕ_p is Lipschitz continuous with $L_1 = \sqrt{2} + 2^{(1/p-1/2)}$ when $1 < p < 2$, and with $L_2 = 1 + \sqrt{2}$ when $p \geq 2$. In other words, $|\phi_p(w) - \phi_p(w')| \leq L_1 \|w - w'\|$ when $1 < p < 2$ and $|\phi_p(w) - \phi_p(w')| \leq L_2 \|w - w'\|$ when $p \geq 2$ for all $w, w' \in \mathbb{R}^2$.*

Proof (a) The proof can be seen in [21, page 20]. For completeness, we here include it. Consider any $a \geq 0$ and $b \geq 0$ satisfying $ab = 0$. Then, we have either $a = 0$ or $b = 0$. This implies that $\phi_p(a, b) = \sqrt[p]{|a|^p} - a$ or $\phi_p(a, b) = \sqrt[p]{|b|^p} - b$, and consequently $\phi_p(a, b) = 0$. Conversely, consider any $(a, b) \in \mathbb{R}^2$ satisfying $\phi_p(a, b) = 0$. Then there must hold $a \geq 0$ and $b \geq 0$, otherwise we have $\sqrt[p]{|a|^p + |b|^p} > (a + b)$ and hence $\phi_p(a, b) > 0$. Now we prove that one of a and b must be 0. If not, then $\|(a, b)\|_p < \|(a, b)\|_1 = a + b$, which obviously contradicts the fact that $\phi_p(a, b) = 0$. The two sides show that ϕ_p is indeed an NCP-function.

(b) Let $w = (a, b)$ and $w' = (c, d)$. Then

$$\phi_p(w + w') = \|(a, b) + (c, d)\|_p - (a + b + c + d)$$

$$\leq \|(a, b)\|_p + \|(c, d)\|_p - (a + b) - (c + d)$$

$$= \phi_p(a, b) + \phi_p(c, d) = \phi_p(w) + \phi_p(w'),$$

where the inequality is true since the triangle inequality holds for p-norm when $p > 1$.

(c) Let $w = (a, b) \in \mathbb{R}^2$ and $\alpha > 0$. Then the proof follows by

$$\phi_p(\alpha w) = \sqrt[p]{|\alpha a|^p + |\alpha b|^p} - (\alpha a + \alpha b) = \alpha \sqrt[p]{|a|^p + |b|^p} - \alpha(a + b) = \alpha \phi_p(w).$$

(d) This is true by part (b) and part (c).

(e) Let $w = (a, b)$ and $w' = (c, d)$, we have

$$|\phi_p(w) - \phi_p(w')| = |\|(a, b)\|_p - (a + b) - \|(c, d)\|_p + (c + d)|$$

$$\leq |\|(a, b)\|_p - \|(c, d)\|_p| + |a - c| + |b - d|$$

$$\leq \|(a, b) - (c, d)\|_p + \sqrt{2}\sqrt{|a - c|^2 + |b - d|^2}$$

$$\leq \|(a, b) - (c, d)\|_p + \sqrt{2}\|(a, b) - (c, d)\|$$

$$= \|w - w'\|_p + \sqrt{2}\|w - w'\|.$$

Then, from the inequality as below (see [13, (1.3)]),

$$\|x\|_{p_2} \leq \|x\|_{p_1} \leq n^{(1/p_1 - 1/p_2)}\|x\|_{p_2} \quad \text{for } x \in \mathbb{R}^n \text{ and } 1 < p_1 < p_2,$$

we obtain the desired results.   □

As below, ϕ_p has more further properties which are key to proving results of the subsequent section.

Lemma 3.1 *Let* $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ *be defined as* (7) *where* $p > 1$. *If* $\{(a^k, b^k)\} \subseteq \mathbb{R}^2$ *with* $(a^k \to -\infty)$ *or* $(b^k \to -\infty)$ *or* $(a^k \to \infty$ *and* $b^k \to \infty)$, *then we have* $|\phi_p(a^k, b^k)| \to \infty$ *for* $k \to \infty$.

Proof This result has been mentioned in [21, p. 20].   □

Next, we study another family of NCP functions $\psi_p : \mathbb{R}^2 \to \mathbb{R}_+$ defined by (8). This class of functions will lead the NCP to a reformulation of unconstrained minimization. In other words, they are a family of merit functions for the NCP. Furthermore, they have some favorable properties shown as below. Particularly, ψ_p for any given $p > 1$ is continuously differentiable everywhere whereas ϕ_p is not differentiable everywhere.

Proposition 3.2 *Let* ϕ_p, ψ_p *be defined as* (7) *and* (8), *respectively, where* p *is any fixed real number in the interval* $(1, +\infty)$. *Then*

(a) ψ_p *is an NCP-function, i.e., it satisfies* (2).
(b) $\psi_p(a, b) \geq 0$ *for all* $(a, b) \in \mathbb{R}^2$.
(c) ψ_p *is continuously differentiable everywhere.*
(d) $\nabla_a \psi_p(a, b) \cdot \nabla_b \psi_p(a, b) \geq 0$ *for all* $(a, b) \in \mathbb{R}^2$. *The equality holds if and only if* $\phi_p(a, b) = 0$.
(e) $\nabla_a \psi_p(a, b) = 0 \iff \nabla_b \psi_p(a, b) = 0 \iff \phi_p(a, b) = 0$.

Proof (a) Since $\psi_p(a, b) = 0$ if and only if $\phi_p(a, b) = 0$, the desired result is satisfied by Proposition 3.1(a).

(b) It is clear by definition of ψ_p.

(c) From direct computation, we obtain $\nabla_a \psi_p(0, 0) = \nabla_b \psi_p(0, 0) = 0$. For $(a, b) \neq (0, 0)$,

$$\nabla_a \psi_p(a, b) = \left(\frac{sgn(a) \cdot |a|^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \phi_p(a, b),$$

$$\nabla_b \psi_p(a, b) = \left(\frac{sgn(b) \cdot |b|^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \phi_p(a, b) \tag{12}$$

where $sgn(\cdot)$ is the sign function. Clearly,

$$\left| \frac{sgn(a) \cdot |a|^{p-1}}{\|(a,b)\|_p^{p-1}} \right| \leq 1 \quad \text{and} \quad \left| \frac{sgn(b) \cdot |b|^{p-1}}{\|(a,b)\|_p^{p-1}} \right| \leq 1 \tag{13}$$

(i.e., uniformly bounded) and moreover $\phi_p(a,b) \to 0$ as $(a,b) \to (0,0)$. Therefore, we have $\nabla_a \psi_p(a,b) \to 0$ and $\nabla_b \psi_p(a,b) \to 0$ as $(a,b) \to (0,0)$. This means that ψ_p is continuously differentiable everywhere.

(d) From part (c), we know that if $(a,b) = (0,0)$, it is clear that $\nabla_a \psi_p(a,b) \cdot \nabla_b \psi_p(a,b) = 0$ and $\psi_p(a,b) = 0$. Now we assume that $(a,b) \neq (0,0)$. Then,

$$\nabla_a \psi_p(a,b) \cdot \nabla_b \psi_p(a,b)$$
$$= \left(\frac{sgn(a) \cdot |a|^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 \right) \left(\frac{sgn(b) \cdot |b|^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 \right) \phi_p^2(a,b). \tag{14}$$

Again, from (13), it follows immediately that $\nabla_a \psi_p(a,b) \cdot \nabla_b \psi_p(a,b) \geq 0$ for all $(a,b) \in \mathbb{R}^2$. The equality holds if and only if $\phi_p(a,b) = 0$, $\frac{sgn(a) \cdot |a|^{p-1}}{\|(a,b)\|_p^{p-1}} = 1$ or $\frac{sgn(b) \cdot |b|^{p-1}}{\|(a,b)\|_p^{p-1}} = 1$. In fact, if $\frac{sgn(a) \cdot |a|^{p-1}}{\|(a,b)\|_p^{p-1}} = 1$, then we have $a > 0$ and $|a| = \|(a,b)\|_p$, which leads to $b = 0$ and hence $\phi_p(a,b) = \sqrt[p]{|a|^p} - a = a - a = 0$. Similarly, we have $\phi_p(a,b) = 0$ if $\frac{sgn(b) \cdot |b|^{p-1}}{\|(a,b)\|_p^{p-1}} = 1$. Thus, we conclude that the equality holds if and only if $\phi_p(a,b) = 0$.

(e) It is already seen in the last part of proof for part (d). □

It was shown that if F is monotone [10] or a P_0-function [5], then any stationary point of Ψ is a global minima of the unconstrained minimization $\min_{x \in \mathbb{R}^n} \Psi(x)$, and hence solves the NCP. Moreover, it was also shown that if F is strongly monotone [10] or uniform P-function [5], then the level sets of Ψ are bounded. In what follows, we will present and prove analogous results for Ψ_p under the same conditions as in [5, 10]. The ideas for proving the following propositions are borrowed from those analogous results in [5, 10].

Proposition 3.3 *Let $\Psi_p : \mathbb{R}^n \to \mathbb{R}$ be defined as (10) with $p > 1$. Then $\Psi_p(x) \geq 0$ for all $x \in \mathbb{R}^n$ and $\Psi_p(x) = 0$ if and only if x solves the NCP (1). Moreover, suppose that the NCP (1) has at least one solution. Then x is a global minimizer of Ψ_p if and only if x solves the NCP (1).*

Proof The results directly follow from Proposition 3.2. □

Proposition 3.4 *Let $\Psi_p : \mathbb{R}^n \to \mathbb{R}$ be defined as (10) with $p > 1$. Assume F is either monotone or P_0-function, then every stationary point of Ψ_p is a global minima of (11); and therefore solves the NCP (1).*

Proof (I) For the assumption of monotonicity of F, suppose that x^* is a stationary point of Ψ_p. Then we have $\nabla \Psi_p(x^*) = 0$ which implies that

$$\sum_{i=1}^{n} (\nabla_a \psi_p(x_i^*, F_i(x^*))e_i + \nabla_b \psi_p(x_i^*, F_i(x^*))\nabla F_i(x^*)) = 0, \qquad (15)$$

where $e_i = (0, \ldots, 1, \ldots, 0)^T$. We denote $\nabla_a \psi_p(x^*, F(x^*)) = (\ldots, \nabla_a \psi_p(x_i^*, F_i(x^*)), \ldots)^T$ and $\nabla_b \psi_p(x^*, F(x^*)) = (\ldots, \nabla_b \psi_p(x_i^*, F_i(x^*)), \ldots)^T$, respectively. Then (15) can be abbreviated as

$$\nabla_a \psi_p(x^*, F(x^*)) + \nabla F(x^*)\nabla_b \psi_p(x^*, F(x^*)) = 0. \qquad (16)$$

Now, multiplying (16) by $\nabla_b \psi_p(x^*, F(x^*))^T$ leads to

$$\sum_{i=1}^{n} (\nabla_a \psi_p(x_i^*, F_i(x^*)) \cdot \nabla_b \psi_p(x_i^*, F_i(x^*)))$$
$$+ \nabla_b \psi_p(x^*, F(x^*))^T \nabla F(x^*)\nabla_b \psi_p(x^*, F(x^*)) = 0. \qquad (17)$$

Since F is monotone, $\nabla F(x^*)$ is positive semidefinite, the second term of (17) is nonnegative. Moreover, each term in the first summation of (17) is nonnegative as well due to Proposition 3.2(d). Therefore, we have

$$\nabla_a \psi_p(x_i^*, F_i(x^*)) \cdot \nabla_a \psi_p(x_i^*, F_i(x^*)) = 0, \quad \forall i = 1, 2, \ldots, n,$$

which yields $\phi_p(x_i^*, F_i(x^*)) = 0$ for all $i = 1, 2, \ldots, n$ by Proposition 3.2(e). Thus, $\Psi_p(x^*) = 0$ which says x^* is a global minimizer of (11).

(II) If F is P_0-function and suppose x^* is a stationary point of Ψ_p. Then $\nabla \Psi_p(x^*) = 0$ which yields (16). Notice that $\nabla_a \psi_p(a, b)$ and $\nabla_b \psi_p(a, b)$ are given as forms of (12). If we denote $A(x^*)$ and $B(x^*)$ the possibly multivalued $n \times n$ diagonal matrices whose diagonal elements are given by

$$A_{ii}(x^*) = \frac{sgn(x_i^*) \cdot |x_i^*|^{p-1}}{\|(x_i^*, F_i(x^*))\|_p^{p-1}} \quad \text{if } (x_i^*, F_i(x^*)) \neq (0, 0)$$

and

$$B_{ii}(x^*) = \frac{sgn(F_i(x^*)) \cdot |F_i(x^*)|^{p-1}}{\|(x_i^*, F_i(x^*))\|_p^{p-1}} \quad \text{if } (x_i^*, F_i(x^*)) \neq (0, 0).$$

If $(x_i^*, F_i(x^*)) = (0, 0)$ then we let $A(x^*) = B(x^*) = I$, i.e., the $n \times n$ identity matrix. With the notions of $A(x^*)$, $B(x^*)$ and (12), (16) can be rewritten as

$$\left[(A(x^*) - I) + \nabla F(x^*)(B(x^*) - I) \right] \Phi_p(x^*) = 0. \qquad (18)$$

We want to prove that $\Phi_p(x^*) = 0$ (and hence $\Psi_p(x^*) = 0$). Suppose not, i.e., $\Phi_p(x^*) \neq 0$. Recall that $\Phi_p(x^*) = 0$ if and only if (1) is satisfied and the ith component of $\Phi_p(x^*)$ is $\phi_p(x_i^*, F_i(x^*))$. Thus, $\phi_p(x_i, F_i(x^*)) \neq 0$ means one of the following occurs:

1. $x_i^* \neq 0$ and $F_i(x^*) \neq 0$.
2. $x_i^* = 0$ and $F_i(x^*) < 0$.
3. $x_i^* < 0$ and $F_i(x^*) = 0$.

In every case, we have $B_{ii}(x^*) \neq 1$ (since $B_{ii}(x^*) = 1$ if and only if $\phi_p(x_i^*, F_i(x^*)) = 0$ by Proposition 3.2(d, e)), so that $(B_{ii}(x^*) - 1) \cdot \phi_p(x_i^*, F_i(x^*)) \neq 0$. Similar arguments apply for the vector $(A(x^*) - I)\Phi_p(x^*)$. Thus, from the above, we can easily verify that if $\Phi_p(x^*) \neq 0$ then $(B(x^*) - I)\Phi_p(x^*)$ and $(A(x^*) - I)\Phi_p(x^*)$ are both nonzero. Moreover, both of their nonzero elements are in the same positions, and such nonzero elements have the same sign. But, for (18) to hold, it would be necessary that $\nabla F(x^*)$ "revert the sign" of all the nonzero elements of $(B(x^*) - I)\Phi_p(x^*)$, which contradicts the fact that $\nabla F(x^*)$ is a P_0-matrix by Lemma 2.1. $\qquad \square$

Proposition 3.5 *Let* $\Psi_p : \mathbb{R}^n \to \mathbb{R}$ *be defined as* (10) *with* $p > 1$. *Assume F is either strongly monotone or uniform P-function, then the level sets*

$$\mathcal{L}(\Psi_p, \gamma) := \{x \in \mathbb{R}^n \mid \Psi_p(x) \leq \gamma\}$$

are bounded for all $\gamma \in \mathbb{R}$.

Proof (I) First, we consider the assumption of strong monotonicity of F. Suppose there exists an unbounded sequence $\{\|x^k\|\}_{k \in K} \to \infty$ with $\{x^k\}_{k \in K} \subseteq \mathcal{L}(\Psi_p, \gamma)$ for some $\gamma \geq 0$, where K is a subset of N. We define the index set

$$J := \{i \in \{1, 2, \ldots, n\} \mid \{x_i^k\} \text{ is unbounded}\}.$$

Since $\{x^k\}$ is unbounded, $J \neq \emptyset$. Let $\{z^k\}$ denote a bounded sequence defined by

$$z_i^k = \begin{cases} 0, & \text{if } i \in J, \\ x_i^k, & \text{if } i \notin J. \end{cases}$$

Then from the definition of $\{z^k\}$ and the strong monotonicity of F, we obtain

$$\begin{aligned} \mu \sum_{i \in J}(x_i^k)^2 &= \mu \|x^k - z^k\|^2 \\ &\leq \langle x^k - z^k, F(x^k) - F(z^k) \rangle \\ &= \sum_{i=1}^n (x_i^k - z_i^k)(F_i(x^k) - F_i(z^k)) \\ &= \sum_{i \in J} x_i^k(F_i(x^k) - F_i(z^k)) \\ &\leq \left(\sum_{i \in J}(x_i^k)^2\right)^{1/2} \sum_{i \in J} |F_i(x^k) - F_i(z^k)|. \end{aligned} \qquad (19)$$

Since $\sum_{i \in J}(x_i^k)^2 \neq 0$ for $k \in K$, then dividing by $\sum_{i \in J}(x_i^k)^2$ on both sides of (19) yields

$$\mu \left(\sum_{i \in J}(x_i^k)^2 \right)^{1/2} \leq \sum_{i \in J} |F_i(x^k) - F_i(z^k)|, \quad k \in K. \tag{20}$$

On the other hand, we know $\{F_i(z^k)\}_{k \in K}$ is bounded ($i \in J$) due to $\{z^k\}_{k \in K}$ is bounded and F is continuous. Therefore, from (20), we have

$$\{|F_{i_0}(x^k)|\} \to \infty \quad \text{for some } i_0 \in J.$$

Also, $\{\|x_{i_0}^k\|\} \to \infty$ by the definition of the index set J. Thus, Lemma 3.1 yields

$$\phi_p(x_{i_0}^k, F_{i_0}(x^k)) \to \infty \quad \text{as } k \to \infty.$$

But this contradicts $\{x^k\} \subseteq \mathcal{L}(\Psi_p, \gamma)$.

(II) If F is uniform P-function, then the proof almost follows the same arguments as above. In particular, (19) is replaced by

$$\begin{aligned}
\mu \sum_{i \in J}(x_i^k)^2 &= \mu \|x^k - z^k\|^2 \\
&\leq \max_{1 \leq i \leq n} (x_i^k - z_i^k)(F_i(x^k) - F_i(z^k)) \\
&= \max_{i \in J} x_i^k (F_i(x^k) - F_i(z^k)) \\
&= x_{j_0}^k (F_{j_0}(x^k) - F_{j_0}(z^k)) \\
&\leq |x_{j_0}^k| \cdot |F_{j_0}(x^k) - F_{j_0}(z^k)|,
\end{aligned} \tag{21}$$

where j_0 is one of the indices for which the max is attained. Then dividing by $|x_{j_0}^k|$ on both sides of (21) and the proof follows. $\qquad\square$

4 A descent method

In this section, we study a descent method for solving the unconstrained minimization (11), which does not require the derivative of F involved in the NCP. In addition, we prove a global convergence result for this derivative-free descent algorithm. More precisely, we consider the search direction as below:

$$d^k := -\nabla_b \psi_p(x^k, F(x^k)), \tag{22}$$

where $\nabla_b \psi_p(x^k, F(x^k)) = (\nabla_b \psi_p(x_1^k, F(x_1^k)), \dots, \nabla_b \psi_p(x_n^k, F(x_n^k)))^T$. From the following lemma, we see that d^k is a descent direction of Ψ_p at x^k under monotonicity assumption.

Lemma 4.1 *Let $x^k \in \mathbb{R}^n$ and F be a monotone function. Then the search direction defined as (22) satisfies the descent condition $\nabla \Psi_p(x^k)^T d^k < 0$ as long as x^k is not a solution of the NCP (1). Moreover, if F is strongly monotone with modulus $\mu > 0$, then $\nabla \Psi_p(x^k)^T d^k \leq -\mu \|d^k\|^2$.*

Proof Since $\nabla \Psi_p(x^k) = \nabla_a \psi_p(x^k, F(x^k)) + \nabla F(x^k) \nabla_b \psi_p(x^k, F(x^k))$, we have that

$$\nabla \Psi_p(x^k)^T d^k = -\sum_{i=1}^n \nabla_a \psi_p(x_i^k, F_i(x^k)) \cdot \nabla_b \psi_p(x_i^k, F_i(x^k))$$
$$- (d^k)^T \nabla F(x^k)(d^k). \tag{23}$$

From the monotonicity of F, it follows that $\nabla F(x^k)$ is positive semidefinite. Therefore, the second term of (23) is nonnegative. Also, by Proposition 3.2(d), the first term of (23) is nonnegative. Therefore, $\nabla \Psi_p(x^k)^T d^k \leq 0$. We next prove that $\nabla \Psi_p(x^k)^T d^k < 0$ by contradiction. Assume that $\nabla \Psi_p(x^k)^T d^k = 0$. Then $\nabla_a \psi_p(x_i^k, F_i(x^k)) \cdot \nabla_b \psi_p(x_i^k, F_i(x^k)) = 0$ for all i which, by Proposition 3.2(d) again, yields $\phi_p(x_i^k, F_i(x^k)) = 0$. Thus, $\Phi_p(x^k, F(x^k)) = 0$ and $\Psi_p(x^k, F(x^k)) = 0$. Consequently, x^k solves the NCP (1). This obviously contradicts our assumption that x^k is not a solution of the NCP (1).

If F is strongly monotone with modulus $\mu > 0$, then we have that

$$\nabla \Psi_p(x^k)^T d^k \leq -(d^k)^T \nabla F(x^k)(d^k) \leq -\mu \|d^k\|^2,$$

where the first inequality follows from (23) and Proposition 3.2(d). □

The above lemma motivates the following descent algorithm.

Algorithm 4.1

(Step 0) Given a real number $p > 1$ and $x^0 \in \mathbb{R}^n$. Choose the parameters $\varepsilon \geq 0$, $\sigma \in (0, 1)$ and $\beta \in (0, 1)$. Set $k := 0$.

(Step 1) If $\Psi_p(x^k) \leq \varepsilon$, then Stop.

(Step 2) Let

$$d^k := -\nabla_b \psi_p(x^k, F(x^k)).$$

(Step 3) Compute a step-size $t_k := \beta^{m_k}$, where m_k is the smallest nonnegative integer m satisfying the Armijo-type condition:

$$\Psi_p(x^k + \beta^m d^k) \leq (1 - \sigma \beta^{2m}) \Psi_p(x^k). \tag{24}$$

(Step 4) Set $x^{k+1} := x^k + t_k d^k$, $k := k + 1$ and Go to Step 1.

We next show the global convergence result for Algorithm 4.1 under the strongly monotone assumption of F. To this end, we assume that the parameter ε used in Algorithm 4.1 is set to be zero and Algorithm 4.1 generates an infinite sequence $\{x^k\}$.

Proposition 4.1 *Suppose that F is strongly monotone. Then the sequence $\{x^k\}$ generated by Algorithm 4.1 has at least one accumulation point and any accumulation point is a solution of the NCP (1).*

Proof Firstly, we show that there exists a nonnegative integer m_k in Step 3 of Algorithm 4.1 whenever x^k is not a solution. Assume that the conclusion does not hold. Then for any $m > 0$,

$$\Psi_p(x^k + \beta^m d^k) - \Psi_p(x^k) > -\sigma\beta^{2m}\Psi_p(x^k).$$

Dividing by β^m on both sides and taking the limit $m \to \infty$ yield

$$\langle \nabla\Psi_p(x^k), d^k \rangle \geq 0.$$

Since F is strongly monotone, this obviously contradicts Lemma 4.1. Hence, we can find an integer m_k in Step 3.

Secondly, we show that the sequence $\{x^k\}$ generated by Algorithm 4.1 has at least one accumulation point. By the descent property of Algorithm 4.1, the sequence $\{\Psi_p(x^k)\}$ is decreasing. Thus, by Proposition 3.5, the generated sequence is bounded and hence it has at least one accumulation point.

Finally, we prove that every accumulation point is a solution of the NCP (1). Let x^* be an arbitrary accumulation point of the generated sequence $\{x^k\}$. Then there exists a subsequence $\{x^k\}_{k \in K}$ converging to x^*. We know that $-\nabla_b\psi_p(\,\cdot\,, F(\cdot))$ is continuous since ψ_p is continuously differentiable, therefore, $\{d^k\}_{k \in K} \to d^*$. Next, we need to discuss two cases. First, we consider the case where there exists a constant $\bar{\beta}$ such that $\beta^{m_k} \geq \bar{\beta} > 0$ for all $k \in K$. Then, from (24), we obtain

$$\Psi_p(x^{k+1}) \leq (1 - \sigma\bar{\beta}^2)\Psi_p(x^k)$$

for all $k \in K$ and the entire sequence $\{\Psi_p(x^k)\}$ is decreasing. Thus, we have $\Psi_p(x^*) = 0$ (by taking the limit) which says x^* is a solution of the NCP (1). Now, we consider the other case where there exists a further subsequence such that $\beta^{m_k} \to 0$. Note that by Armijo's rule (24) in Step 3, we have

$$\Psi_p(x^k + \beta^{m_k-1} \cdot d^k) - \Psi_p(x^k) > -\sigma\beta^{2(m_k-1)}\Psi_p(x^k).$$

Dividing both sides by β^{m_k-1} and passing to the limit on the subsequence, we obtain

$$\langle \nabla\Psi_p(x^*), d^* \rangle \geq 0$$

which implies that x^* is a solution of the NCP (1) by Lemma 4.1. □

5 Numerical experiments

We implemented Algorithm 4.1 with our code in MATLAB 6.1 for all test problems with all available starting points in MCPLIB [1]. All numerical experiments were done at a PC with CPU of 2.8 GHz and RAM of 512 MB. In order to improve the

numerical behavior of Algorithm 4.1, we replaced the standard (monotone) Armijo-rule by nonmonotone line search as described in [11], i.e., we computed the smallest nonnegative integer l such that

$$\Psi_p(x^k + \beta^l d^k) \leq \mathcal{W}_k - \sigma\beta^{2l}\Psi_p(x^k),$$

where $\mathcal{W}_k$ is given by

$$\mathcal{W}_k = \max_{j=k-m_k,\ldots,k} \Psi_p(x^j)$$

and where, for given nonnegative integers $\hat{m}$ and s, we set

$$m_k = \begin{cases} 0 & \text{if } k \leq s, \\ \min\{m_{k-1} + 1, \hat{m}\} & \text{otherwise.} \end{cases}$$

Throughout the experiments, we use $\hat{m} = 5$ and $s = 5$. Moreover, we use the parameters $\sigma = 1.0e-10$ and $\beta = 0.2$ in Algorithm 4.1. We terminated our iteration when the number of iteration is over $500\,000$ or the steplength is less than $1.0e-10$ or one of the following conditions is satisfied:

(C1) $\Psi_p(x^k) \leq 1.0e-5$ and $(x^k)^T F(x^k) \leq 5.0e-3$;
(C2) $\Psi_p(x^k) \leq 3.0e-7$ and $(x^k)^T F(x^k) \leq 3.0e-2$;
(C3) $\Psi_p(x^k) \leq 3.0e-6$ and $(x^k)^T F(x^k) \leq 1.0e-2$.

Our computational results are summarized in Tables 1, 2, 3 (see the Appendix). In these tables, the first column lists the name of the problems and the starting point number in MCPLIB, **Gap** denotes the value of $x^T F(x)$ at the final iteration, **NF** indicates the number of function evaluations of the merit function Ψ_p for solving each problem, and **Time** represents the CPU time in seconds for solving each problem.

The results reported in Tables 1–3 show that our descent method based on the merit function $\Psi_{1.5}(x)$, $\Psi_2(x)$ or $\Psi_3(x)$ was able to solve most complementarity problems in MCPLIB. More precisely, there are seven failures (**pgvon105, pgvon106, powell, scarfanum, scarfasum, scarfbnum, scarfbsum**) for Algorithm 4.1 due to a too small steplength. After a careful check, we find the direction d defined in Algorithm 4.1 is not a descent one for these problems. In fact, the seven problems are regarded as difficult ones for those Newton type algorithms [19, 20]. In addition, we may see that the descent algorithm using the merit function $\Psi_{1.5}(x)$ has better numerical results than using the Fischer-Burmeister function. Particularly, it appears from Tables 1–3 that the descent algorithm based on $\Psi_p(x)$ will take more function evaluations and yield larger value of **Gap** when the parameter p increases. A reasonable interpretation for this is that the value of $\Psi_p(x)$ become smaller when p increases and hence causes some difficulty for the descent Algorithm 4.1. This also implies that the performance of Algorithm 4.1 will become worse when the parameter p increases. This is an important new discovery, which has big contribution in constructing new NCP-functions, not found in the literature to our best knowledge.

6 Final remarks

In this paper, we have studied a family of NCP-functions $\phi_p(a, b)$ which include the well-known Fischer-Burmeister function as a special case and have shown that this class of functions enjoy some favorable properties as other NCP-functions do. In addition, we propose a descent method for the unconstrained minimization (11) which is a reformulation of the NCP via the proposed NCP-functions. Numerical results for the test problem in MCPLIB have shown this method is promising when $\Psi_p(x)$ is specified as $\Psi_{1.5}(x)$, $\Psi_2(x)$ or $\Psi_3(x)$. Moreover, from our numerical implementations, there indicates that the performance of the descent method become better when p decreases, which is a new and important discovery. This implies that there does exist new NCP-function which is better than Fischer-Burmeister function. It is yet unknown whether similar phenomena happens in different algorithm, which is an interesting future topic.

There still are many issues for this NCP-function to be explored like those for other NCP-functions done in the literature. For instance, it would be of interest to know the semismoothness property of ψ_p and the Lipschitz continuous property of $\nabla \psi_p$. In fact, some of them are recently studied in [2, 3]. In addition, it is interesting to know whether this class of NCP-functions can be used for SDCP and SOCCP. Some researchers have started this issue but no update reports by now. We leave them for future research topics.

Appendix

Table 1   Numerical results for MCPLIB problems based on $\Psi_{1.5}(x)$, $\Psi_2(x)$ and $\Psi_3(x)$

Problem	$\Psi_{1.5}(x)$			$\Psi_2(x)$			$\Psi_3(x)$		
	Gap	NF	Time	Gap	NF	Time	Gap	NF	Time
bertsekas(1)	9.99e–3	63 094	30.74	3.00e–2	86 826	36.42	3.00e–2	71 127	35.92
bertsekas(2)	3.00e–2	63 764	34.92	3.00e–2	65 801	31.39	1.00e–2	89 556	49.22
bertsekas(3)	3.00e–2	318 308	176.4	3.00e–2	322 751	161.5	3.00e–2	416 869	231.8
billups	3.35e–19	25	0.00	3.35e–19	25	0.00	3.35e–19	25	0.00
colvdual(1)	1.48e–2	69 675	41.84	2.33e–2	70 393	36.72	3.00e–2	181 627	109.2
colvdual(2)	1.00e–2	34 266	21.36	9.98e–3	49 436	24.81	9.94e–3	53 895	32.13
colvnlp(1)	1.08e–2	206 856	104.1	1.43e–2	207 529	93.99	3.00e–2	221 400	117.23
colvnlp(2)	9.99e–3	11 390	5.37	9.99e–3	11 753	4.84	9.96e–3	11 964	5.83
cycle	1.14e–3	7	0.00	9.05e–6	5	0.00	2.81e–4	4	0.00
explcp	2.43e–3	5895	3.14	2.58e–3	6001	2.70	2.88e–3	6008	3.13
gafni(1)	2.15e–3	202	0.08	2.68e–3	203	0.06	4.65e–3	203	0.08
gafni(2)	2.08e–3	236	0.09	2.63e–3	229	0.08	4.64e–3	227	0.08
gafni(3)	2.05e–3	250	0.09	2.61e–3	240	0.08	4.62e–3	238	0.08

Table 2 Numerical results for MCPLIB problems based on $\Psi_{1.5}(x)$, $\Psi_2(x)$ and $\Psi_3(x)$

Problem	$\Psi_{1.5}(x)$			$\Psi_2(x)$			$\Psi_3(x)$		
	Gap	NF	Time	Gap	NF	Time	Gap	NF	Time
hanskoop(1)	2.86e–3	763	0.45	2.80e–3	2 322	1.58	2.77e–3	1 783	0.94
hanskoop(2)	2.86e–3	858	0.48	2.77e–3	2 224	1.36	2.77e–3	1 685	0.91
hanskoop(3)	2.03e–3	453	0.25	1.96e–3	1 105	0.66	1.50e–3	974	0.63
hanskoop(4)	2.04e-3	440	0.27	1.96e–3	1 082	0.66	1.20e–3	1 860	1.02
hanskoop(5)	*	*	*	*	*	*	*	*	*
josephy(1)	3.72e–3	1 112	0.31	4.46e–3	1 150	0.31	6.91e–3	1 172	0.50
josephy(2)	1.00e–2	1 393	0.42	1.00e–2	5 104	1.36	1.00e–2	43 398	13.08
josephy(3)	1.00e–2	632	0.17	1.00e–2	1 373	0.36	1.00e–2	1 211	0.34
josephy(4)	1.00e–2	3 106	0.88	1.00e–2	15 801	4.67	1.25e–2	500 027	150.8
josephy(5)	2.34e–3	48	0.03	2.33e–3	47	0.02	2.42e–2	45	0.02
josephy(6)	1.00e–2	1 399	0.41	1.00e–2	5 114	1.36	1.00e–2	177 730	52.36
kojshin(1)	2.43e–2	16 827	4.89	2.42e–2	162 856	45.84	9.76e–2	500 001	148.8
kojshin(2)	2.42e–2	283 119	84.59	9.22e–2	500 001	142.6	6.51e–1	500 001	154.0
kojshin(3)	2.43e–2	23 533	6.98	2.42e–2	254 703	66.86	1.22e–1	500 001	158.0
kojshin(4)	6.98e–3	90	0.08	4.42e–3	553	0.13	3.73e–3	97	0.03
kojshin(5)	9.99e–3	4 340	1.23	9.76e–3	26 286	7.67	1.57e–2	500 001	152.3
kojshin(6)	2.41e–2	182 387	54.64	6.81e–2	50 0001	141.9	3.23e–1	500 001	150.8

The * in Table 2 means that the method fails for this problem.

Table 3 Numerical results for MCPLIB problems based on $\Psi_{1.5}(x)$, $\Psi_2(x)$ and $\Psi_3(x)$

Problem	$\Psi_{1.5}(x)$			$\Psi_2(x)$			$\Psi_3(x)$		
	Gap	NF	Time	Gap	NF	Time	Gap	NF	Time
mathinum(1)	9.36e–3	167	0.05	9.64e–3	125	0.03	8.87e–3	122	0.03
mathinum(2)	6.27e–3	133	0.03	8.00e–3	127	0.03	5.60e–3	113	0.03
mathinum(3)	9.15e–3	111	0.03	8.62e–3	81	0.03	8.41e–3	83	0.02
mathinum(4)	8.99e–3	252	0.08	9.49e–3	136	0.03	8.71e–3	128	0.05
mathisum(1)	1.18e–3	109 142	31.84	2.70e–3	113 416	32.72	7.80e–3	500 056	152.7
mathisum(2)	1.20e–3	90 185	26.97	2.80e–3	102 412	29.98	7.90e–3	500 005	155.3
mathisum(3)	1.19e–3	109 323	32.34	2.84e–3	102 416	29.81	7.90e–3	500 004	153.5
mathisum(4)	1.20e–3	90 285	27.30	2.84e–3	102 417	29.81	7.90e–3	500 003	159.2
nash(1)	1.52e–2	4764	2.50	1.52e–2	4837	2.30	1.52e–2	4842	2.47
nash(2)	1.52e–2	4633	2.33	1.52e–2	4669	2.19	1.52e–2	4669	2.58
sppe(1)	3.00e–2	88 251	66.86	3.00e–2	88 378	57.52	3.00e–2	88 487	73.67
sppe(2)	3.00e–2	87 482	66.58	3.00e–2	87 460	59.30	3.00e–2	87 523	71.06
tobin(1)	2.55e–2	14 187	15.48	2.55e–2	14 348	14.44	2.56e-2	14 392	16.50
tobin(2)	2.72e–2	14 174	16.12	3.00e–2	14 467	13.61	2.10e–2	20 330	24.39

References

1. Billups, S.C., Dirkse, S.P., Soares, M.C.: A comparison of algorithms for large scale mixed complementarity problems. Comput. Optim. Appl. **7**, 3–25 (1997)
2. Chen, J.-S.: On some NCP-functions based on the generalized Fischer-Burmeister function. Asia-Pac. J. Oper. Res. **24**, 401–420 (2007)
3. Chen, J.-S.: The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem. J. Glob. Optim. **36**, 565–580 (2006)
4. Cottle, R.W., Pang, J.-S., Stone, R.-E.: The Linear Complementarity Problem. Academic Press, New York (1992)
5. Facchinei, F., Soares, J.: A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. **7**, 225–247 (1997)
6. Ferris, M.C., Mangasarian, O.L., Pang, J.-S. (eds.): Complementarity: Applications, Algorithms and Extensions. Kluwer Academic, Dordrecht (2001)
7. Fischer, A.: A special Newton-type optimization methods. Optimization **24**, 269–284 (1992)
8. Fischer, A.: Solution of the monotone complementarity problem with locally Lipschitzian functions. Math. Program. **76**, 513–532 (1997)
9. Fukushima, M.: Merit functions for variational inequality and complementarity problems. In: Di Pillo, G., Giannessi, F. (eds.) Nonlinear Optimization ans Applications, pp. 155–170. Plenum Press, New York (1996)
10. Geiger, C., Kanzow, C.: On the resolution of monotone complementarity problems. Comput. Optim. Appl. **5**, 155–173 (1996)
11. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
12. Harker, P.T., Pang, J.-S.: Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications. Math. Program. **48**, 161–220 (1990)
13. Higham, N.J.: Estimating the matrix p-norm. Num. Math. **62**, 539–555 (1992)
14. Jiang, H.: Unconstrained minimization approaches to nonlinear complementarity problems. J. Glob. Optim. **9**, 169–181 (1996)
15. Kanzow, C.: Nonlinear complementarity as unconstrained optimization. J. Optim. Theory Appl. **88**, 139–155 (1996)
16. Mangasarian, O.L., Solodov, M.V.: Nonlinear complementarity as unconstrained and constrained minimization. Math. Program. **62**, 277–297 (1993)
17. Mangasarian, O.L.: Equivalence of the complementarity problem to a system of nonlinear equations. SIAM J. Appl. Math. **31**, 89–92 (1976)
18. Pang, J.-S.: Complementarity problems. In: Horst, R., Pardalos, P. (eds.) Handbook of Global Optimization, pp. 271–338. Kluwer Academic, Boston (1994)
19. Qi, H.-D., Liao, L.-Z.: A smoothing Newton method for gneral nonlinear complementarity problems. Comput. Optim. Appl. **17**, 231–253 (2000)
20. Sun, D., Qi, L.-Q.: On NCP-functions. Comput. Optim. Appl. **13**, 201–220 (1999)
21. Tseng, P.: Global behaviour of a class of merit functions for the nonlinear complementarity problem. J. Optim. Theory Appl. **89**, 17–37 (1996)
22. Yamada, K., Yamashita, N., Fukushima, M.: A new derivative-free descent method for the nonlinear complementarity problems. In: Pillo, G.D., Giannessi, F. (eds.) Nonlinear Optimization and Related Topics, pp. 463–489. Kluwer Academic, Dordrecht (2000)
23. Yamashita, N., Fukushima, M.: On stationary points of the implicit Lagrangian for the nonlinear complementarity problems. J. Optim. Theory Appl. **84**, 653–663 (1995)
24. Yamashita, N., Fukushima, M.: Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems. Math. Program. **76**, 469–491 (1997)