Comput Optim Appl (2011) 50:49–73
DOI 10.1007/s10589-009-9315-9

# A new class of penalized NCP-functions and its properties

**J.-S. Chen · Z.-H. Huang · C.-Y. She**

Received: 16 April 2009 / Published online: 23 January 2010


**Abstract** In this paper, we consider a class of penalized NCP-functions, which includes several existing well-known NCP-functions as special cases. The merit function induced by this class of NCP-functions is shown to have bounded level sets and provide error bounds under mild conditions. A derivative free algorithm is also proposed, its global convergence is proved and numerical performance compared with those based on some existing NCP-functions is reported.

**Keywords** NCP-function · Penalized · Bounded level sets · Error bounds

## 1 Introduction

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

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

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F = (F_1, \ldots, F_n)^T$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$. We assume that $F$ is continuously differentiable throughout this paper. The

J.-S. Chen member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.
Z.-H. Huang's work is partly supported by the National Natural Science Foundation of China (Grant No. 10871144).

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

C.-Y. She
e-mail: bbgmmshe@hotmail.com

Z.-H. Huang
Department of Mathematics, Tianjin University, Tianjin 300072, China
e-mail: huangzhenghai@tju.edu.cn

NCP has attracted much attention because of its wide applications in the fields of economics, engineering, and operations research [6, 14].

Many methods have been proposed to solve the NCP; see [3, 4, 14, 16–18, 20, 23, 26, 29, 30]. For more details, please refers to the excellent monograph [9]. One of the most powerful and popular methods is to reformulate the NCP as a system of nonlinear equations [24, 25, 30], or as an unconstrained minimization problem [7, 10–12, 19, 21, 27, 29]. The objective function that can constitute an equivalent unconstrained minimization problem is called a merit function, whose global minima are coincident with the solutions of the original NCP (1). To construct a merit function, a class of functions called NCP-functions and defined below, plays a significant role.

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

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

Many NCP-functions have been proposed in the literature. Among them, the Fischer-Burmeister (FB) function is one of the most popular NCP-functions, which is defined by

$$\phi_{\mathrm{FB}}(a,b)=\sqrt{a^2+b^2}-(a+b), \quad \forall (a,b)\in\mathbb{R}^2. \tag{3}$$

Through this NCP-function $\phi_{\mathrm{FB}}$, the NCP (1) can be reformulated as a system of nonsmooth equations:

$$\Phi_{\mathrm{FB}}(x):=\begin{pmatrix}\phi_{\mathrm{FB}}(x_1,F_1(x))\\ \vdots \\ \phi_{\mathrm{FB}}(x_n,F_n(x))\end{pmatrix}=0. \tag{4}$$

Thus, the function $\Psi_{\mathrm{FB}} : \mathbb{R}^n \to \mathbb{R}_+$ defined as below is a merit function for the NCP:

$$\Psi_{\mathrm{FB}}(x):=\frac{1}{2}\|\Phi_{\mathrm{FB}}(x)\|^2=\sum_{i=1}^n \psi_{\mathrm{FB}}(x_i,F_i(x)), \tag{5}$$

where $\psi_{\mathrm{FB}} : \mathbb{R}^2 \to \mathbb{R}_+$ is the square of $\phi_{\mathrm{FB}}$, i.e.,

$$\psi_{\mathrm{FB}}(a,b)=\frac{1}{2}\left|\sqrt{a^2+b^2}-(a+b)\right|^2. \tag{6}$$

Consequently, the NCP is equivalent to the unconstrained minimization problem:

$$\min_{x\in\mathbb{R}^n} \Psi_{\mathrm{FB}}(x). \tag{7}$$

There are several generalizations of the FB function in the literature. For example, Kanzow and Kleinmichel [22] extend $\phi_{\mathrm{FB}}$ function to

$$\phi_\theta(a,b):=\sqrt{(a-b)^2+\theta ab}-(a+b), \quad \theta\in(0,4).$$

Chen, Chen, and Kanzow [2] study a penalized FB function

$$\phi_\lambda(a,b) := \lambda\phi_{\rm FB}(a,b) + (1-\lambda)a_+b_+, \quad \lambda\in(0,1).$$

Some other types of penalized FB functions are also investigated by Sun and Qi in [28]. Recently, Chen and Pan [3, 4] consider the following generalization of the FB function:

$$\phi_p(a,b) := \|(a,b)\|_p - (a+b), \tag{8}$$

where $p>1$ and $\|(a,b)\|_p$ denotes the $p$-norm of $(a,b)$, i.e., $\|(a,b)\|_p = \sqrt[p]{|a|^p+|b|^p}$. Another further generalization is proposed by Hu, Huang and Chen in [15]:

$$\phi_{\theta,p}(a,b) := \sqrt[p]{\theta(|a|^p+|b|^p)+(1-\theta)(|a-b|^p)} - (a+b), \tag{9}$$

where $p>1$, $\theta\in(0,1]$.

All the aforementioned functions are NCP-functions. The corresponding function $\psi_\theta$, $\psi_\lambda$, $\psi_p$, and $\psi_{\theta,p}$ is square of $\phi_\theta$, $\phi_\lambda$, $\phi_p$, and $\phi_{\theta,p}$, respectively, and naturally induces a merit function $\Psi_\theta$, $\Psi_\lambda$, $\Psi_p$, and $\Psi_{\theta,p}$ like what $\psi_{\rm FB}$ function does. Along this track, in this paper, we study the following merit function $\Psi_{\alpha,\theta,p}:\mathbb{R}^n\to\mathbb{R}_+$ for the NCP:

$$\Psi_{\alpha,\theta,p}(x) := \sum_{i=1}^n \psi_{\alpha,\theta,p}(x_i, F_i(x)), \tag{10}$$

where $\psi_{\alpha,\theta,p}:\mathbb{R}^2\to\mathbb{R}_+$ is an NCP-function defined by

$$\psi_{\alpha,\theta,p}(a,b) := \frac{\alpha}{2}(\max\{0,ab\})^2 + \psi_{\theta,p}(a,b) \tag{11}$$

with $\alpha\ge 0$ being a real parameter. Note that $\psi_{\alpha,\theta,p}$ includes all the above functions $\psi_{\rm FB}$, $\psi_p$, $\psi_\theta$, $\psi_{\theta,p}$ (and $\psi_7$ in [28]) as special cases. Although $\psi_{\alpha,\theta,p}$ is obtained by penalizing the function $\psi_{\theta,p}$ considered in [15], more favorable properties of $\psi_{\alpha,\theta,p}$ are explored in this work. In particular, $\Psi_{\alpha,\theta,p}$ has property of bounded level sets and provides a global error bound for the NCP under mild condition which were not studied in [15]. Thus, this paper can be viewed as a follow-up of [15]. On the other hand, as remarked in [2], penalized Fischer-Burmeister (FB) function not only possesses stronger properties than FB function but also gives extremely promising numerical performance, which is another motivation of our considering this generalization of several NCP-functions.

This paper is organized as follows. In Sect. 2, we review some definitions and preliminary results to be used in the subsequent analysis. In Sect. 3, we show some properties of the proposed merit function. In Sect. 4, we propose a derivative free algorithm based on this merit function $\Psi_{\alpha,\theta,p}$, show its global convergence, and report some numerical results. In Sect. 5, we make concluding remarks.

Throughout this paper, $\mathbb{R}^n$ denotes the space of $n$-dimensional real column vectors and $^T$ denotes transpose. For every differentiable function $f:\mathbb{R}^n\to\mathbb{R}$, $\nabla f(x)$ denotes the gradient of $f$ at $x$. For every differentiable mapping $F=(F_1,\dots,F_n)^T$ :

$\mathbb{R}^n \to \mathbb{R}^n$, $\nabla F(x) = (\nabla F_1(x) \ \ldots \ \nabla F_n(x))$ denotes the transpose Jacobian of $F$ at $x$. We use $\|x\|_p$ to denote the $p$-norm of $x$ and denote $\|x\|$ the Euclidean norm of $x$. The level set of a function $\Psi : \mathbb{R}^n \to \mathbb{R}$ is denoted by $\mathcal{L}(\Psi, c) := \{x \in \mathbb{R}^n \mid \Psi(x) \leq c\}$. In addition, we will frequently mention two merit functions. One is the natural residual merit function $\Psi_{\mathrm{NR}} : \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\Psi_{\mathrm{NR}}(x) := \frac{1}{2}\sum_{i=1}^{n} \phi_{\mathrm{NR}}^2(x_i, F_i(x)), \tag{12}$$

where $\phi_{\mathrm{NR}} : \mathbb{R}^2 \to \mathbb{R}$ denotes the minimum NCP-function $\min\{a, b\}$. Another one is $\Psi_{\theta,p} : \mathbb{R}^n \to \mathbb{R}_+$ induced by $\psi_{\theta,p}$:

$$\Psi_{\theta,p}(x) := \frac{1}{2}\sum_{i=1}^{n} \phi_{\theta,p}^2(x_i, F_i(x)). \tag{13}$$

Unless otherwise stated, in the sequel, we always suppose that $p$ is a fixed real number in $(1, \infty)$.

## 2 Preliminaries

This section briefly recalls some concepts about the mapping $F$ that will be used later. A matrix is said to be $P$-matrix if each of its principal minors is positive, and is called $P_0$-matrix if each of its principal minors is nonnegative. Obviously, $P$-matrix is a generalization of positive definite matrix, while $P_0$-matrix is a generalization of positive semidefinite matrix. Such concepts of $P$-matrix and $P_0$-function can be further extended to nonlinear mapping, which we call them $P$-function and $P_0$-function.

**Definition 2.1** Let $F = (F_1, \ldots, F_n)^T$ with $F_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, n$. We say that

(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 strongly monotone if $\langle x - y, F(x) - F(y)\rangle \geq \mu\|x - y\|^2$ for some $\mu > 0$ and for all $x, y \in \mathbb{R}^n$.
(c) $F$ 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 $x \neq y$.
(d) $F$ 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 $x, y \in \mathbb{R}^n$.
(e) $F$ is Lipschitz continuous if there exists a constant $L > 0$ such that $\|F(x) - F(y)\| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^n$.

It is well-known that every monotone function is an $P_0$ function and every strongly monotone function is a uniform $P$-function. For a continuously differentiable function $F$, if its (transpose) Jacobian $\nabla F(x)$ is an $P$-matrix then $F$ is an $P$-function (the converse may not be true), whereas the (transpose) Jacobian $\nabla F(x)$ is an $P_0$-matrix if and only if $F$ is an $P_0$-function. For more detailed properties of various monotone and $P$ ($P_0$)-function, please refer to [9].

## 3 Properties of the new NCP-function

In this section, we study some favorable properties of the merit function $\psi_{\alpha,\theta,p}$, and then present some mild conditions under which the merit function $\Psi_{\alpha,\theta,p}$ has bounded level sets and provides a global error bound, respectively. To this end, we present some technical lemmas which are needed for subsequent analysis.

**Lemma 3.1** *For $p>1$, $a>0$, $b>0$, we have $a^p+b^p \le (a+b)^p$.*

*Proof* We present two different ways to prove this lemma.

(1) For any $p>1$, $p=n+m$, where $n=[p]$ (the greatest integer less than or equal to $p$) and $m=p-n$, applying binomial theorem gives

$$\begin{aligned}(a+b)^p &= (a+b)^n(a+b)^m\\ &\ge (a^n+b^n)(a+b)^m\\ &= a^n(a+b)^m+b^n(a+b)^m\\ &\ge a^n a^m + b^n b^m\\ &= a^p+b^p.\end{aligned}$$

(2) Let $f(t)=(t+1)^p-(t^p+1)$. It is easy to verify that $f$ is increasing on $[0,\infty)$ when $p>1$. Hence, $f(a/b)\ge f(0)=0$ which yields $(a+b)^p \ge a^p+b^p$. □

**Lemma 3.2** *The function $\psi_{\alpha,\theta,p}$ defined by (11) has the following favorable properties*:

(a) *$\psi_{\alpha,\theta,p}$ is an NCP-function and $\psi_{\alpha,\theta,p}\ge 0$ for all $(a,b)\in\mathbb{R}^2$.*
(b) *$\psi_{\alpha,\theta,p}$ is continuously differentiable everywhere. Moreover, if $(a,b)\ne(0,0)$,*

$$\begin{aligned}&\nabla_a\psi_{\alpha,\theta,p}(a,b)\\ &\quad=\alpha b(ab)_+ + \left(\frac{\theta\,\mathrm{sgn}(a)\cdot|a|^{p-1}+(1-\theta)\,\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p)]^{(p-1)/p}}-1\right)\phi_{\theta,p}(a,b),\\ &\nabla_b\psi_{\alpha,\theta,p}(a,b)\\ &\quad=\alpha a(ab)_+ + \left(\frac{\theta\,\mathrm{sgn}(b)\cdot|b|^{p-1}-(1-\theta)\,\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p)]^{(p-1)/p}}-1\right)\phi_{\theta,p}(a,b),\end{aligned} \tag{14}$$

*and otherwise, $\nabla_a\psi_{\alpha,\theta,p}(0,0)=\nabla_b\psi_{\alpha,\theta,p}(0,0)=0$.*

(c) *For $p\ge 2$, the gradient of $\psi_{\alpha,\theta,p}$ is Lipschitz continuous on any nonempty bounded set $S$, i.e., there exists $L>0$ such that for any $(a,b),(c,d)\in S$,*

$$\|\nabla\psi_{\alpha,\theta,p}(a,b)-\nabla\psi_{\alpha,\theta,p}(c,d)\|\le L\|(a,b)-(c,d)\|.$$

(d) *$\nabla_a\psi_{\alpha,\theta,p}(a,b)\cdot\nabla_b\psi_{\alpha,\theta,p}(a,b)\ge 0$ for any $(a,b)\in\mathbb{R}^2$, and the equality holds if and only if $\psi_{\alpha,\theta,p}(a,b)=0$.*

(e) $\nabla_a \psi_{\alpha,\theta,p}(a,b) = 0 \Longleftrightarrow \nabla_b \psi_{\alpha,\theta,p}(a,b) = 0 \Longleftrightarrow \psi_{\alpha,\theta,p}(a,b) = 0$.
(f) *Suppose that* $\alpha > 0$. *If* $a \to -\infty$ *or* $b \to -\infty$ *or* $ab \to \infty$, *then* $\psi_{\alpha,\theta,p}(a,b) \to \infty$.

*Proof* (a) It is clear that $\psi_{\alpha,\theta,p}(a,b) \geq 0$ for all $(a,b) \in \mathbb{R}^2$ from the definition of $\psi_{\alpha,\theta,p}$. Then by [15, Proposition 2.1], we have

$$\psi_{\alpha,\theta,p}(a,b) = 0 \quad \Longleftrightarrow \quad \frac{\alpha}{2}\big(\max\{0,ab\}\big)^2 = 0 \quad \text{and}$$
$$\psi_{\theta,p}(a,b) = 0 \quad \Longleftrightarrow \quad a \geq 0, \quad b \geq 0, \quad ab = 0.$$

Hence, $\psi_{\alpha,\theta,p}$ is an NCP-function.

(b) First, direct calculations give the partial derivatives of $\psi_{\alpha,\theta,p}$. Then, using $\alpha b(ab)_+ \to (0,0)$ and $\alpha a(ab)_+ \to (0,0)$ as $(a,b) \to (0,0)$, we have $\frac{\alpha}{2}(\max\{0,ab\})^2$ is continuously differentiable everywhere. By [15, Proposition 2.5], it is known that $\psi_{\theta,p}$ is continuously differentiable everywhere. In view of the expression of $\nabla \psi_{\alpha,\theta,p}(a,b)$, $\psi_{\alpha,\theta,p}$ is also continuously differentiable everywhere.

(c) First, we claim that $a(ab)_+$ for any $a,b \in \mathbb{R}$ is Lipschitz continuous on any nonempty bounded set $S$. For any $(a,b) \in S$ and $(c,d) \in S$, without loss of generality, we may assume that $a^2+b^2 \leq k$ and $c^2+d^2 \leq k$ which imply $|a| \leq k+1$, $|b| \leq k+1$, $|c| \leq k+1$ and $|d| \leq k+1$. Then,

$$\begin{aligned}
&\big|a(ab)_+ - c(cd)_+\big| \\
&= \frac{1}{2}\big|a^2b + a|ab| - c^2d - c|cd|\big| \\
&= \frac{1}{2}\big|a^2b - a^2d + a^2d - c^2d + a|ab| - c|ab| + c|ab| - c|cd|\big| \\
&\leq \frac{1}{2}\big(|a^2b - a^2d| + |a^2d - c^2d| + \big|a|ab| - c|ab|\big| + \big|c|ab| - c|cd|\big|\big) \\
&= \frac{1}{2}\big(a^2|b-d| + |a+c||d||a-c| + |ab||a-c| + |c||ab-cd|\big) \\
&\leq \frac{1}{2}\big[k|b-d| + (|a|+|c|)|d||a-c| + k|a-c| + (k+1)|ab-ad+ad-cd|\big] \\
&\leq \frac{1}{2}\big[k|b-d| + 2(k+1)^2|a-c| + k|a-c| + (k+1)^2(|b-d|+|a-c|)\big] \\
&= \frac{1}{2}\big\{\big[2(k+1)^2 + k + (k+1)^2\big]|a-c| + \big[k+(k+1)^2\big]|b-d|\big\} \\
&\leq l\big(|a-c| + |b-d|\big) \\
&\leq \sqrt{2}l\|(a,b)-(c,d)\|,
\end{aligned}$$

where $l = 2(k+1)^2 + k + (k+1)^2$. Hence, the mapping $a(ab)_+$ is Lipschitz continuous on any nonempty bounded set $S$ and so is $\alpha a(ab)_+$. Similarly, $\alpha b(ab)_+$ is Lipschitz continuous on any nonempty bounded set $S$. All of these imply the gradient function of the function $\frac{\alpha}{2}(\max\{0, ab\})^2$ is Lipschitz continuous on any bounded set $S$. On the other hand, by [15, Theorem 2.1], the gradient function of the function $\psi_{\theta,p}$ with $p \geq 2$, $\theta \in (0, 1]$ is Lipschitz continuous. Thus, the gradient of $\psi_{\alpha,\theta,p}$ is Lipschitz continuous on any nonempty bounded set $S$.

(d) If $(a, b) = (0, 0)$, part (d) clearly holds. Now we assume that $(a, b) \neq (0, 0)$. Then,

$$\begin{aligned}
&\nabla_a \psi_{\alpha,\theta,p}(a, b) \cdot \nabla_b \psi_{\alpha,\theta,p}(a, b) \\
&\quad = cd\phi_{\theta,p}^2(a, b) + \alpha^2 ab(ab)_+{}^2 + \alpha a(ab)_+ c\phi_{\theta,p}(a, b) + \alpha b(ab)_+ d\phi_{\theta,p}(a, b),
\end{aligned} \tag{15}$$

where

$$c = \left( \frac{\theta \operatorname{sgn}(a) \cdot |a|^{p-1} + (1-\theta) \operatorname{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 1 \right),$$

$$d = \left( \frac{\theta \operatorname{sgn}(b) \cdot |b|^{p-1} - (1-\theta) \operatorname{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 1 \right).$$

From the proof of [15, Proposition 2.5 ], we know $ab(ab)_+^2 \geq 0$ and

$$\begin{aligned}
&\left( \frac{\theta \operatorname{sgn}(a) \cdot |a|^{p-1} + (1-\theta) \operatorname{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 1 \right) \leq 0, \\
&\left( \frac{\theta \operatorname{sgn}(b) \cdot |b|^{p-1} - (1-\theta) \operatorname{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 1 \right) \leq 0,
\end{aligned} \tag{16}$$

it suffices to show that the last two terms of (15) are nonnegative. For this purpose, we claim that

$$\alpha a(ab)_+ \left( \frac{\theta \operatorname{sgn}(a) \cdot |a|^{p-1} + (1-\theta) \operatorname{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 1 \right) \phi_{\theta,p}(a, b) \geq 0 \tag{17}$$

for all $(a, b) \neq (0, 0)$. If $a \leq 0$ and $b \leq 0$, then $\phi_{\theta,p}(a, b) \geq 0$, which together with the second inequality in (16) implies that (17) holds. If $a \leq 0$ and $b \geq 0$, then $(ab)_+ = 0$, which says that (17) holds. If $a > 0$ and $b > 0$, then $|a|^p + |b|^p \geq |a - b|^p$. Thus, $\phi_{\theta,p}(a, b) \leq \phi_p(a, b) \leq 0$, which together with the second inequality in (16) yields (17). If $a > 0$ and $b \leq 0$, then $(ab)_+ = 0$, and hence (17) holds. Similarly, we also have

$$\alpha b(ab)_+ \left( \frac{\theta \operatorname{sgn}(b) \cdot |b|^{p-1} - (1-\theta) \operatorname{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 1 \right) \phi_{\theta,p}(a, b) \geq 0$$

for all $(a, b) \neq (0, 0)$. Consequently, $\nabla_a \psi_{\alpha,\theta,p}(a, b) \cdot \nabla_b \psi_{\alpha,\theta,p}(a, b) \geq 0$. Besides, by the proof of [15, Proposition 2.5], we know $c = 0$ if and only if $b = 0$ and $a > 0$; $d = 0$ if and only if $a = 0$ and $b > 0$. This together with (15) says $\nabla_a \psi_{\alpha,\theta,p}(a, b) \cdot \nabla_b \psi_{\alpha,\theta,p}(a, b) = 0$ if and only if $\{\psi_{\theta,p}(a, b) = 0$ and $\alpha^2 ab(ab)_+{}^2 = 0\}$ or $\{c = 0\}$ or $\{d = 0\}$ if and only if $\{\psi_{\theta,p}(a, b) = 0$ and $ab \leq 0\}$ or $\{c = 0\}$ or $\{d = 0\}$ if and only if $\psi_{\theta,p}(a, b) = 0$ and $\frac{\alpha}{2}(\max\{0, ab\})^2 = 0$ if and only if $\psi_{\alpha,\theta,p}(a, b) = 0$.

(e) If $\psi_{\alpha,\theta,p}(a,b)=0$, then $\frac{\alpha}{2}(\max\{0,ab\})^2=0$ and $\psi_{\theta,p}(a,b)=0$, which imply $ab\le 0$ and $\phi_{\theta,p}(a,b)=0$. Hence, $\nabla_a\psi_{\alpha,\theta,p}(a,b)=0$ and $\nabla_b\psi_{\alpha,\theta,p}(a,b)=0$. Now, it remains to show that $\nabla_a\psi_{\alpha,\theta,p}(a,b)=0$ implying $\psi_{\alpha,\theta,p}(a,b)=0$. Suppose that $\nabla_a\psi_{\alpha,\theta,p}(a,b)=0$. Then,

$$\alpha b(ab)_+ = -\left(\frac{\theta\,\mathrm{sgn}(a)\cdot|a|^{p-1}+(1-\theta)\,\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p)]^{(p-1)/p}}-1\right)\phi_{\theta,p}(a,b). \tag{18}$$

We will argue that the equality (18) implies ($b=0, a\ge 0$) or ($b>0, a=0$). To see this, we let

$$c=\alpha b(ab)_+,$$
$$d=-\left(\frac{\theta\,\mathrm{sgn}(a)\cdot|a|^{p-1}+(1-\theta)\,\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p)]^{(p-1)/p}}-1\right)\phi_{\theta,p}(a,b),$$
$$e=\left(\frac{\theta\,\mathrm{sgn}(a)\cdot|a|^{p-1}+(1-\theta)\,\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p)]^{(p-1)/p}}-1\right).$$

It is not hard to observe that ($e\le 0$) and ($e=0$ implies $b=0$) which are helpful for the following discussions.

*Case 1*: $b=0$ and $a<0$. Then, $c=0$ but $d\ne 0$ which violates (18).
*Case 2*: $b<0$ and $a\ge 0$. Then, we have $e<0$, and hence $c=0$ but $d\ne 0$, which violates (18).
*Case 3*: $b<0$ and $a<0$. Then, we have $e<0$ and $\phi_{\theta,p}(a,b)>0$, which lead to $c\le 0$ but $d>0$. This contradicts to (18) too.
*Case 4*: $b>0$ and $a>0$. Then, we have $e<0$ and $\phi_{\theta,p}(a,b)<0$, which imply $c\ge 0$ but $d<0$. This contradicts to (18) too.
*Case 5*: $b>0$ and $a<0$. Similar arguments as in Case 2 cause a contradiction.

Thus, (18) implies ($b=0, a\ge 0$) or ($b>0, a=0$), and each of which always yields $\psi_{\alpha,\theta,p}(a,b)=0$. By symmetry, $\nabla_b\psi_{\alpha,\theta,p}(a,b)=0$ also implies $\psi_{\alpha,\theta,p}(a,b)=0$.

(f) If $a\to-\infty$ or $b\to-\infty$, from [15, Proposition 2.4], we know $|\phi_{\theta,p}(a,b)|\to\infty$. In addition, the fact $\frac{\alpha}{2}(\max\{0,ab\})^2\ge 0$ gives $\psi_{\alpha,\theta,p}(a,b)\to\infty$. If $ab\to\infty$, since $\alpha>0$, we have $\frac{\alpha}{2}(\max\{0,ab\})^2\to\infty$. This together with $\psi_{\theta,p}(a,b)\ge 0$ says $\psi_{\alpha,\theta,p}(a,b)\to\infty$. □

By Lemma 3.2(a), we immediately have the following theorem.

**Theorem 3.1** *Let $\Psi_{\alpha,\theta,p}$ be defined as in (10). Then $\Psi_{\alpha,\theta,p}(x)\ge 0$ for all $x\in\mathbb{R}^n$ and $\Psi_{\alpha,\theta,p}(x)=0$ if and only if $x$ solves the NCP. Moreover, if the NCP has at least one solution, then $x$ is a global minimizer of $\Psi_{\alpha,\theta,p}$ if and only if $x$ solves the NCP.*

*Proof* Since $\psi_{\theta,p}$ is an NCP-function, from [15, Proposition 2.5], we have that $x$ solving the NCP $\Longleftrightarrow$ $x\ge 0, F(x)\ge 0, \langle x,F(x)\rangle=0$ $\Longleftrightarrow$ $x\ge 0$, $F(x)\ge 0, x_iF_i(x)=0$ for all $i\in\{1,2,\dots,n\}$ $\Longleftrightarrow$ $\Psi_{\alpha,\theta,p}(x)=0$. Besides, $\Psi_{\alpha,\theta,p}(x)$ is nonnegative. Thus, if $x$ solves the NCP, then $x$ is a global minimizer of $\Psi_{\alpha,\theta,p}$. Next, we claim that if the NCP has at least one solution, then $x$ is a global minimizer of $\Psi_{\alpha,\theta,p} \implies x$ solves the NCP. Suppose $x$ does not solve the NCP. From $x$ solves the NCP $\iff \Psi_{\alpha,\theta,p}(x) = 0$ and $\Psi_{\alpha,\theta,p}(x)$ is nonnegative, it is clear $\Psi_{\alpha,\theta,p}(x) > 0$. However, by assumption, the NCP has a solution, say $y$, which makes that $\Psi_{\alpha,\theta,p}(y) = 0$. Then, we get a contradiction that $\Psi_{\alpha,\theta,p}(x) > 0 = \Psi_{\alpha,\theta,p}(y)$ and $x$ is a global minimizer of $\Psi_{\alpha,\theta,p}$. Thus, we complete the proof. □

Theorem 3.1 indicates that the NCP can be recast as the unconstrained minimization:

$$\min_{x\in\mathbb{R}^n} \Psi_{\alpha,\theta,p}(x). \tag{19}$$

In general, it is hard to find a global minimum of $\Psi_{\alpha,\theta,p}$. Therefore, it is important to know under what conditions a stationary point of $\Psi_{\alpha,\theta,p}$ is a global minimum. Using Lemma 3.2(d) and the same proof techniques as in [21, Theorem 3.5], we can establish that each stationary point of $\Psi_{\alpha,\theta,p}$ is a global minimum only if $F$ is a $P_0$-function.

**Theorem 3.2** *Let $F$ be a $P_0$-function. Then $x^* \in \mathbb{R}^n$ is a global minimum of the unconstrained optimization problem* (19) *if and only if $x^*$ is a stationary point of $\Psi_{\alpha,\theta,p}$.*

**Theorem 3.3** *The function $\Psi_{\alpha,\theta,p}$ has bounded level sets $\mathcal{L}(\Psi_{\alpha,\theta,p}, c)$ for all $c \in \mathbb{R}$, if $F$ is monotone and the NCP is strictly feasible* (*i.e., there exists $\hat{x} > 0$ such that $F(\hat{x}) > 0$*) *when $\alpha > 0$, or $F$ is a uniform $P$-function when $\alpha \geq 0$.*

*Proof* From [2], if F is a monotone function with a strictly feasible point, then the following condition holds: for every sequence $\{x^k\}$ such that $\|x^k\| \to \infty$, $(-x^k)_+ < \infty$, and $(-F(x^k))_+ < \infty$, we have $\max_{1\leq i\leq n}\{(x_i^k)_+ F_i(x^k)_+\} \to \infty$. Suppose that there exists an unbounded sequence $x^k \subseteq \mathcal{L}(\Psi_{\alpha,\theta,p}, c)$ for some $c \in \mathbb{R}$. Since $\Psi_{\alpha,\theta,p}(x^k) \leq c$, there is no index $i$ such that $x_i^k \to -\infty$ or $F_i(x^k) \to -\infty$ by Lemma 3.2(f). Hence, $\max_{1\leq i\leq n}\{(x_i^k)_+ F_i(x^k)_+\} \to \infty$. Also, there is an index $j$, and at least a subsequence $x_j^k$ such that $\{(x_j^k)_+ F_j(x^k)_+\} \to \infty$. However, this implies that $\Psi_{\alpha,\theta,p}(x^k)$ is unbounded by Lemma 3.2(f), contracting to the assumption on level sets. Another part of the proof is similar to the proof of [4, Proposition 3.5]. □

In what follows, we show that the merit functions $\Psi_{\theta,p}$, $\Psi_{\mathrm{NR}}$ and $\Psi_{\alpha,\theta,p}$ have the same order of growth behavior on every bounded set. For this purpose, we need the following crucial technical lemma.

**Lemma 3.3** *Let $\phi_{\theta,p} : \mathbb{R}^2 \to \mathbb{R}$ be defined as in* (9). *Then for any $p > 1$ and all $\theta \in (0, 1]$ we have*

$$(2 - 2^{\frac{1}{p}})|\min\{a, b\}| \leq |\phi_{\theta,p}(a, b)| \leq (2 + 2^{\frac{1}{p}})|\min\{a, b\}|. \tag{20}$$

*Proof* Without loss of generality, we assume $a \geq b$. We will prove the desired results by considering the following two cases: (1) $a + b \leq 0$ and (2) $a + b > 0$.

*Case* (1): $a+b\leq 0$. In this case, we need to discuss two subcases:

(i) $|a|^p+|b|^p\geq |a-b|^p$. In this subcase, we have

$$\begin{aligned}|\phi_{\theta,p}(a,b)| &\geq |\sqrt[p]{\theta(|a-b|^p)+(1-\theta)(|a-b|^p)}-(a+b)|\\ &= |\sqrt[p]{(|a-b|^p)}-(a+b)|\\ &= |(|a-b|-(a+b)|\\ &= |a-b-(a+b)|\\ &= |2b|\\ &= 2|\min\{a,b\}|\\ &\geq (2-2^{\frac{1}{p}})|\min\{a,b\}|. \end{aligned} \tag{21}$$

On the other hand, since $|a|^p+|b|^p\geq |a-b|^p$ and by [5, Lemma 3.2], we have

$$|\phi_{\theta,p}(a,b)|\leq |\phi_p(a,b)|\leq (2+2^{\frac{1}{p}})|\min\{a,b\}|. \tag{22}$$

(ii) $|a|^p+|b|^p<|a-b|^p$. Since $|a|^p+|b|^p<|a-b|^p$ and by [5, Lemma 3.2], we have

$$|\phi_{\theta,p}(a,b)|>|\phi_p(a,b)|\geq (2-2^{\frac{1}{p}})|\min\{a,b\}|. \tag{23}$$

On the other hand, by the discussion of Case (1),

$$|\phi_{\theta,p}(a,b)|<2|b|\leq (2+2^{\frac{1}{p}})|\min\{a,b\}|. \tag{24}$$

*Case* (2): $a+b>0$. If $ab=0$, then (20) clearly holds. Thus, we proceed the arguments by discussing two subcases:

(i) $ab<0$. In this subcases, we have $a>0$, $b<0$, $|a|>|b|$. By Lemma 3.1, $|a|^p+|b|^p\leq |a-b|^p$. Then,

$$\phi_{\theta,p}(a,b)\geq \phi_p(a,b)\geq |a|-a-b\geq -b=|\min\{a,b\}|\geq (2-2^{\frac{1}{p}})|\min\{a,b\}|. \tag{25}$$

On the other hand,

$$\phi_{\theta,p}(a,b)\leq |a-b|-(a+b)=-2b=2|\min\{a,b\}|\leq (2+2^{\frac{1}{p}})|\min\{a,b\}|. \tag{26}$$

(ii) $ab>0$. In this subcases, we have $a\geq b>0$, $|a|^p+|b|^p\geq |a-b|^p$. By Lemma 3.1, $\phi_{\theta,p}(a,b)\leq \phi_p(a,b)\leq 0$. Notice that $\phi_{\theta,p}(a,b)\geq |a-b|-(a+b)=-2b=-2\min\{a,b\}$, and hence we obtain that

$$|\phi_{\theta,p}(a,b)|\leq 2|\min\{a,b\}|\leq (2+2^{\frac{1}{p}})|\min\{a,b\}|. \tag{27}$$

On the other hand, since $\phi_{\theta,p}(a,b)\leq \phi_p(a,b)\leq 0$ , and by [5, Lemma 3.2], and hence we obtain that

$$|\phi_{\theta,p}(a,b)|\geq |\phi_p(a,b)|\geq (2-2^{\frac{1}{p}})|\min\{a,b\}|. \tag{28}$$

All the aforementioned inequalities (21)–(28) imply that (20) holds. □

**Proposition 3.1** *Let $\Psi_{\theta,p}$, $\Psi_{\mathrm{NR}}$ and $\Psi_{\alpha,\theta,p}$ be defined as in (13), (12) and (10), respectively. Let $S$ be an arbitrary bounded set. Then, for any $p > 1$, we have*

$$(2 - 2^{\frac{1}{p}})^2 \Psi_{\mathrm{NR}}(x) \le \Psi_{\theta,p}(x) \le (2 + 2^{\frac{1}{p}})^2 \Psi_{\mathrm{NR}}(x) \quad \textit{for all } x \in \mathbb{R}^n \tag{29}$$

*and*

$$\begin{aligned}&(2 - 2^{\frac{1}{p}})^2 \Psi_{\mathrm{NR}}(x) \le \Psi_{\alpha,\theta,p}(x) \le (\alpha B^2 + (2 + 2^{\frac{1}{p}})^2) \Psi_{\mathrm{NR}}(x)\\ &\textit{for all } x \in S,\end{aligned} \tag{30}$$

*where $B$ is a constant defined by*

$$B = \max_{1 \le i \le n} \left\{ \sup_{x \in S} \left\{ \max\left\{ |x_i|, |F_i(x)| \right\} \right\} \right\} < \infty.$$

*Proof* The inequality in (29) is direct by Lemma 3.3 and the definitions of $\Psi_{\theta,p}$ and $\Psi_{\mathrm{NR}}$. In addition, from Lemma 3.3 and the definition of $\Psi_{\alpha,\theta,p}$, it follows that

$$\Psi_{\alpha,\theta,p}(x) \ge \left(2 - 2^{\frac{1}{p}}\right)^2 \Psi_{\mathrm{NR}}(x) \quad \text{for all } x \in \mathbb{R}^n.$$

It remains to prove the inequality on the right hand side of (30). From the proof of [5, Proposition 3.1], we know for each $i$,

$$(x_i F_i(x))_+ \le B |\min\{x_i, F_i(x)\}| \quad \text{for all } x \in S. \tag{31}$$

By Lemma 3.3 and (31), for all $i = 1, \ldots, n$ and $x \in S$,

$$\psi_{\alpha,\theta,p}(x_i, F_i(x)) \le \frac{1}{2} \left\{ \alpha B^2 + (2 + 2^{\frac{1}{p}})^2 \right\} \min\{x_i, F_i(x)\}^2$$

holds for any $p > 1$. The proof is then complete by the definitions of $\Psi_{\alpha,\theta,p}$ and $\Psi_{\mathrm{NR}}$. □

From Proposition 3.1, we immediately obtain the following result.

**Corollary 3.1** *Let $\Psi_{\theta,p}$ and $\Psi_{\alpha,\theta,p}$ be defined by (13) and (10), respectively; and $S$ be any bounded set. Then, for any $p > 1$ and all $x \in S$, we have the following inequalities:*

$$\frac{(2 - 2^{\frac{1}{p}})^2}{(\alpha B^2 + (2 + 2^{\frac{1}{p}})^2)} \Psi_{\alpha,\theta,p}(x) \le \Psi_{\theta,p}(x) \le \frac{(2 + 2^{\frac{1}{p}})^2}{(2 - 2^{\frac{1}{p}})^2} \Psi_{\alpha,\theta,p}(x)$$

*where $B$ is the constant defined as in Proposition 3.1.*

Since $\Psi_{\theta,p}$, $\Psi_{\mathrm{NR}}$ and $\Psi_{\alpha,\theta,p}$ have the same order on a bounded set, one will provide a global error bound for the NCP as long as the other one does. As below, we show that $\Psi_{\alpha,\theta,p}$ provides a global error bound without the Lipschitz continuity of $F$ when $\alpha > 0$.

**Theorem 3.4** *Let $\Psi_{\alpha,\theta,p}$ be defined as in (10). Suppose that $F$ is a uniform $P$-function with modulus $\mu > 0$. If $\alpha > 0$, then there exists a constant $\kappa_1 > 0$ such that*

$$\|x - x^*\| \le \kappa_1 \Psi_{\alpha,\theta,p}(x)^{\frac{1}{4}} \quad \textit{for all } x \in \mathbb{R}^n;$$

*if $\alpha = 0$ and $S$ is any bounded set, there exists a constant $\kappa_2 > 0$ such that*

$$\|x - x^*\| \le \kappa_2 \big(\max\big\{\Psi_{\alpha,\theta,p}(x), \sqrt{\Psi_{\alpha,\theta,p}(x)}\big\}\big)^{\frac{1}{2}} \quad \textit{for all } x \in S;$$

*where $x^* = (x_1^*, \dots, x_n^*)$ is the unique solution for the NCP.*

*Proof* By the proof of [5, Theorem 3.4], we have

$$\mu\|x - x^*\|^2 \le \max_{1\le i\le n} \tau_i\{(x_i F_i(x))_+ + (-F_i(x))_+ + (-x_i)_+\}, \tag{32}$$

where $\tau_i := \max\{1, x_i^*, F_i(x^*)\}$. We next prove that for all $(a, b) \in \mathbb{R}^2$,

$$(-a)_+{}^2 + (-b)_+{}^2 \le [\phi_{\theta,p}(a, b)]^2. \tag{33}$$

To see this, without loss of generality, we assume $a \ge b$ and discuss three cases:

(i) If $a \ge b \ge 0$, then (33) holds obviously.
(ii) If $a \ge 0 \ge b$, then $|a|^p + |b|^p \le |a - b|^p$ by Lemma 3.1, which implies $\phi_{\theta,p}(a, b) \ge \|(a, b)\|_p - (a + b) \ge -b \ge 0$. Hence, $(-a)_+{}^2 + (-b)_+{}^2 = b^2 \le [\phi_{\theta,p}(a, b)]^2$.
(iii) If $0 \ge a \ge b$, then $(-a)_+{}^2 + (-b)_+{}^2 = a^2 + b^2 \le [\phi_{\theta,p}(a, b)]^2$. Hence, (33) follows.

Suppose that $\alpha > 0$. Using the inequality (33), we then obtain that

$$\begin{aligned}
\big[(ab)_+ + (-a)_+ + (-b)_+\big]^2 &= (ab)_+^2 + (-b)_+^2 + (-a)_+^2 + 2(ab)_+(-a)_+ \\
&\quad + 2(-a)_+(-b)_+ + 2(ab)_+(-b)_+ \\
&\le (ab)_+^2 + (-b)_+^2 + (-a)_+^2 + (ab)_+^2 + (-a)_+^2 \\
&\quad + (-a)_+^2 + (-b)_+^2 + (ab)_+^2 + (-b)_+^2 \\
&\le 3\big[(ab)_+^2 + [\phi_{\theta,p}(a, b)]^2\big] \\
&\le \tau\left[\frac{\alpha}{2}(ab)_+^2 + \frac{1}{2}\big[\phi_{\theta,p}(a, b)\big]^2\right] \\
&= \tau\psi_{\alpha,\theta,p}(a, b),
\end{aligned} \tag{34}$$

where $\tau := \max\{\frac{6}{\alpha}, 6\} > 0$. Combining (34) with (32) and letting $\hat{\tau} = \max_{1\le i\le n} \tau_i$, we get

$$\mu\|x - x^*\|^2 \le \max_{1\le i\le n} \tau_i \big\{\tau\psi_{\alpha,\theta,p}\big(x_i, F_i(x)\big)\big\}^{1/2}$$

$$
\begin{aligned}
&\le \hat{\tau}\tau^{1/2} \max_{1\le i\le n} \psi_{\alpha,\theta,p}(x_i, F_i(x))^{1/2} \\
&\le \hat{\tau}\tau^{1/2} \left\{ \sum_{i=1}^{n} \{\psi_{\alpha,\theta,p}(x_i, F_i(x))\} \right\}^{1/2} \\
&= \hat{\tau}\tau^{1/2} \Psi_{\alpha,\theta,p}(x)^{1/2}.
\end{aligned}
$$

From this, the first desired result follows immediately by setting $\kappa_1 := [\hat{\tau}\tau^{1/2}/\mu]^{1/2}$.

Suppose that $\alpha = 0$. From the proof of Proposition 3.1, the inequality (31) holds. Combining with (32)–(33), it then follows that for all $x \in S$,

$$
\begin{aligned}
\mu\|x - x^*\|^2 &\le \max_{1\le i\le n} \tau_i \big[ B|\min\{x_i, F_i(x)\}| + 2(\psi_{\theta,p}(x_i, F_i(x)))^{1/2} \big] \\
&\le \hat{\tau} \max_{1\le i\le n} \big[ \sqrt{2}\hat{B}(\psi_{\theta,p}(x_i, F_i(x)))^{1/2} + 2(\psi_{\theta,p}(x_i, F_i(x)))^{1/2} \big] \\
&\le (\sqrt{2}\hat{B} + 2)\hat{\tau}(\Psi_{\theta,p}(x))^{1/2} \\
&= (\sqrt{2}\hat{B} + 2)\hat{\tau}(\Psi_{\alpha,\theta,p}(x))^{1/2} \\
&\le (\sqrt{2}\hat{B} + 2)\hat{\tau}\big( \max\big\{ \Psi_{\alpha,\theta,p}(x), \sqrt{\Psi_{\alpha,\theta,p}(x)} \big\} \big)
\end{aligned}
$$

where $\hat{B} = B/(2 - 2^{\frac{1}{p}})$, $\hat{\tau} = \max_{1\le i\le n} \tau_i$ and the second inequality is from Lemma 3.3. Letting $\kappa_2 := [(\sqrt{2}\hat{B} + 2)\hat{\tau}/\mu]^{1/2}$, we obtain the desired result from the above inequality. □

The following lemma is needed for the proof of Proposition 3.2, which we suspect is useful in analysis of convergence rate.

**Lemma 3.4** *For all* $(a, b) \neq (0, 0)$ *and* $p > 1$, *we have the following inequality*:

$$
\left( \frac{\theta[\mathrm{sgn}(a)\cdot|a|^{p-1} + \mathrm{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 2 \right)^2 \ge \big(2 - 2^{\frac{1}{p}}\big)^2 \quad \forall \theta \in (0, 1].
$$

*Proof* If $a = 0$ or $b = 0$, the inequality holds obviously. Then we complete the proof by considering three cases: (i) $a > 0$ and $b > 0$, (ii) $a < 0$ and $b < 0$, and (iii) $ab < 0$.

*Case* (i): Since $\theta \in (0, 1]$ and $p > 1$, it follows that $\theta^{1/p} \le 1$. Now, by the proof of [5, Lemma 3.3], we have

$$
\begin{aligned}
&\frac{\theta[\mathrm{sgn}(a)\cdot|a|^{p-1} + \mathrm{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} \\
&\quad = \frac{\theta[|a|^{p-1} + |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} \\
&\quad \le \frac{\theta[|a|^{p-1} + |b|^{p-1}]}{[\theta(|a|^p + |b|^p)]^{(p-1)/p}}
\end{aligned}
$$

$$= \frac{\theta^{1/p}[|a|^{p-1} + |b|^{p-1}]}{[(|a|^p + |b|^p)]^{(p-1)/p}}$$
$$\le 2^{1/p} \quad \text{for } p > 1.$$

Therefore,

$$2 - \frac{\theta[|a|^{p-1} + |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} \ge 2 - 2^{\frac{1}{p}}$$

for $p > 1$. Squaring both sides then leads to the desired inequality.

*Case* (ii): By similar arguments as in case (i), we obtain

$$2 - 2^{\frac{1}{p}} \le 2 - \frac{\theta[|a|^{p-1} + |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}}$$
$$\le 2 + \frac{\theta[|a|^{p-1} + |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} \quad \text{for } p > 1,$$

from which the result follows immediately.

*Case* (iii): Again, we suppose $|a| \ge |b|$ and therefore have

$$2^{\frac{1}{p}} \ge \frac{\theta[|a|^{p-1} + |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}}$$
$$\ge \frac{\theta[|a|^{p-1} - |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} \quad \text{for } p > 1.$$

Thus,

$$2 - 2^{\frac{1}{p}} \le 2 - \frac{\theta[|a|^{p-1} - |b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{(p-1)/p}}$$

for $p > 1$ and the desired result is also satisfied. □

**Proposition 3.2** *Let $\psi_{\alpha,\theta,p}$ be given as in (11). Then, for all $x \in \mathbb{R}^n$ and $p > 1$,*

$$\left\| \nabla_a \psi_{\alpha,\theta,p}(x, F(x)) + \nabla_b \psi_{\alpha,\theta,p}(x, F(x)) \right\|^2 \ge 2\left(2 - 2^{\frac{1}{p}}\right)^2 \Psi_{\theta,p}(x) \quad \forall \theta \in (0,1].$$

*In particular, for all $x$ belonging to any bounded set $S$ and $p > 1$,*

$$\left\| \nabla_a \psi_{\alpha,\theta,p}(x, F(x)) + \nabla_b \psi_{\alpha,\theta,p}(x, F(x)) \right\|^2 \ge \frac{2(2 - 2^{\frac{1}{p}})^4}{(\alpha B^2 + (2 + 2^{\frac{1}{p}})^2)} \Psi_{\alpha,\theta,p}(x)$$
$$\forall \theta \in (0,1],$$

*where $B$ is defined as in Proposition 3.1 and*

$$\begin{aligned} \nabla_a \psi_{\alpha,\theta,p}(x, F(x)) &:= \left(\nabla_a \psi_{\alpha,\theta,p}(x_1, F_1(x)), \ldots, \nabla_a \psi_{\alpha,\theta,p}(x_n, F_n(x))\right)^T, \\ \nabla_b \psi_{\alpha,\theta,p}(x, F(x)) &:= \left(\nabla_b \psi_{\alpha,\theta,p}(x_1, F_1(x)), \ldots, \nabla_b \psi_{\alpha,\theta,p}(x_n, F_n(x))\right)^T. \end{aligned} \tag{35}$$

*Proof* The second part of the conclusions is direct by Corollary 3.1 and the first part. Thus, it remains to show the first part. From the definitions of $\nabla_a \psi_{\alpha,\theta,p}(x, F(x))$, $\nabla_b \psi_{\alpha,\theta,p}(x, F(x))$ and $\Psi_{\theta,p}(x)$, showing the first part is equivalent to proving that the following inequality

$$\big(\nabla_a \psi_{\alpha,\theta,p}(a,b) + \nabla_b \psi_{\alpha,\theta,p}(a,b)\big)^2 \geq 2\big(2 - 2^{\frac{1}{p}}\big)^2 \psi_{\theta,p}(a,b) \tag{36}$$

holds for all $(a,b) \in \mathbb{R}^2$. When $(a,b) = (0,0)$, the inequality (36) clearly holds. Suppose $(a,b) \neq (0,0)$. Then, it follows from (14) that

$$\begin{aligned}
&\big(\nabla_a \psi_{\alpha,\theta,p}(a,b) + \nabla_b \psi_{\alpha,\theta,p}(a,b)\big)^2 \\
&\quad = \Bigg\{\alpha(a+b)(ab)_+ \\
&\qquad + (\phi_{\theta,p}(a,b))\left(\frac{\theta[\operatorname{sgn}(a)\cdot|a|^{p-1} + \operatorname{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 2\right)\Bigg\}^2 \\
&\quad = \alpha^2(a+b)^2(ab)_+^2 \\
&\qquad + (\phi_{\theta,p}(a,b))^2\left(\frac{\theta[\operatorname{sgn}(a)\cdot|a|^{p-1} + \operatorname{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 2\right)^2 \\
&\qquad + 2\alpha(a+b)(ab)_+(\phi_{\theta,p}(a,b)) \\
&\qquad \times \left(\frac{\theta[\operatorname{sgn}(a)\cdot|a|^{p-1} + \operatorname{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 2\right).
\end{aligned} \tag{37}$$

Now, we claim that for all $(a,b) \neq (0,0) \in \mathbb{R}^2$,

$$2\alpha(a+b)(ab)_+(\phi_{\theta,p}(a,b))\left(\frac{\theta[\operatorname{sgn}(a)\cdot|a|^{p-1} + \operatorname{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 2\right) \geq 0. \tag{38}$$

If $ab \leq 0$, then $(ab)_+ = 0$ and the inequality (38) is clear. If $a, b > 0$, then by the proof of Lemma 3.4, we have

$$\left(\frac{\theta[\operatorname{sgn}(a)\cdot|a|^{p-1} + \operatorname{sgn}(b)\cdot|b|^{p-1}]}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p)]^{(p-1)/p}} - 2\right) \leq 0, \quad \forall (a,b) \neq (0,0) \in \mathbb{R}^2 \tag{39}$$

and $\phi_{\theta,p}(a,b) \leq 0$, which imply the inequality (38) also holds. If $a, b < 0$, then $\phi_{\theta,p}(a,b) \geq 0$, which together with (39) yields the inequality (38). Thus, we obtain that the inequality (38) holds for all $(a,b) \neq (0,0)$. Now using Lemma 3.4 and (37)–(38), we readily obtain the inequality (36) holds for all $(a,b) \neq (0,0)$. The proof is thus complete. □

## 4 Algorithm and numerical experiments

In this section, we investigate a derivative free algorithm for complementarity problems based on the new family of NCP-function and its related merit function. In addition, we prove the global convergence of the algorithm.

**Algorithm 4.1** (A Derivative Free Algorithm)

**Step 0** *Given real numbers* $\alpha > 0$, $p > 1, \theta \in (0, 1]$ *and* $x^0 \in \mathbb{R}^n$. *Choose* $\sigma \in (0, 1)$ *and* $\rho, \gamma \in (0, 1)$. *Set* $k := 0$.
**Step 1** *If* $\Psi_{\alpha,\theta,p}(x^k) = 0$, *stop*, *otherwise go to Step* 2.
**Step 2** *Find the smallest nonnegative integer* $m_k$ *such that*

$$\Psi_{\alpha,\theta,p}(x^k + \rho^{m_k} d_k(\gamma^{m_k})) \le (1 - \sigma\rho^{2m_k})\Psi_{\alpha,\theta,p}(x^k), \tag{40}$$

*where*

$$d_k(\gamma^{m_k}) := -\frac{\partial\Psi_{\alpha,\theta,p}(x^k, F(x^k))}{\partial b} - \gamma^{m_k}\frac{\partial\Psi_{\alpha,\theta,p}(x^k, F(x^k))}{\partial a}.$$

**Step 3** *Set* $x^{k+1} := x^k + \rho^{m_k} d_k(\gamma^{m_k}), k := k + 1$ *and go to Step* 1.

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

*Proof* The proof is similar to the one given in [4, Lemma 4.1]. □

**Proposition 4.2** *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.*

*Proof* We only need to show that if $\{x^k\}$ has an accumulation point, then the corresponding $\{d_k\}$ has also an accumulation point. In fact, under this condition, $\{x^k\}$ is bounded. Without loss of generality, we may assume $x^k \to x_*$. So,

$$\left\{\frac{\partial\Psi_{\alpha,\theta,p}(x^k, F(x^k))}{\partial b}\right\} \quad \text{and} \quad \left\{\frac{\partial\Psi_{\alpha,\theta,p}(x^k, F(x^k))}{\partial a}\right\}$$

are bounded since $\Psi_{\alpha,\theta,p}$ is continuously differentiable. This together with the fact $\gamma \in (0, 1)$ gives that the direction sequence $\{d_k\}$ is bounded. The rest of the proof are similar to those given in [4, Proposition 4.1]. □

In the following, we implement Algorithm 4.1 for complementarity problems from MCPLIB in MATLAB 7.3 in order to see the numerical behavior of Algorithm 4.1. All numerical experiments are done on a PC with CPU of 2.4 GHz and RAM of 2.0 GB. Throughout our computational experiments, we use the following stopping rules, which were also used in [4].

- $\Psi_{\alpha,\theta,p}(x^k) \le 10^{-5}$ and $d^k \le 5.0 \times 10^{-3}$; or
- $\Psi_{\alpha,\theta,p}(x^k) \le 3.0 \times 10^{-7}$ and $d^k \le 3.0 \times 10^{-2}$; or
- $\Psi_{\alpha,\theta,p}(x^k) \le 3.0 \times 10^{-6}$ and $d^k \le 10^{-2}$,

where $d^k$ represents the dual gap $\mathrm{abs}((x^k)^T F(x^k))$. We also terminate the algorithm if the step length is less than $10^{-10}$ or the number of iteration is lager than $5 \times 10^6$ or $\Psi_{\alpha,\theta,p}(x^k) \le 10^{-10}$ or $d^k \le 10^{-10}$. We use the nonmonotone line search scheme described in [13] instead of the standard monotone line search, i.e., we compute the smallest nonnegative integer $h$ such that

$$\Psi_{\alpha,\theta,p}(x^k + \rho^h d_k) \le C_k - \sigma\rho^{2h}\Psi_{\alpha,\theta,p}(x^k),$$

where

$$C_k = \max_{i=k-m_k,\dots,k} \Psi_{\alpha,\theta,p}(x^i) \quad \text{and} \quad m_k = \begin{cases} 0 & \text{if } k \le s, \\ \min\{m_{k-1}+1, \hat{m}\} & \text{otherwise.} \end{cases}$$

Throughout the experiments, the parameters we used are:

$$\hat{m} = 5, \quad s = 5, \quad \rho = 0.25, \quad \sigma = 0.5, \quad \gamma = 0.5,$$

and $\alpha = 0.1$ for the problem $a$ marked by $a^-$ and $\alpha = 2$ for others. In order to improve the numerical results, we scale some problems, i.e., divide the function $F$ by 20, in our numerical implement. It is easy to verify that such a modification does not destroy any results we obtained earlier.

We test problems from MCPLIB [1] to compare the function proposed in this paper with $\phi_{\theta,p}$ in [15], $\phi_\lambda$ in [2] as well as $\phi_\alpha$ in [31]. In order to facility the comparison, we use similar numerical computation structure given in [15]. The numerical results are listed in Tables 1–3, respectively. Among these tables, Problem denotes the problem of MCPLIB tested; GAP denotes the dual gap $\mathrm{abs}((x^k)^T F(x^k))$ when the algorithm terminates; NF denotes the number of function value computation; IT denotes the number of iteration; * denotes the algorithm fails to get an optimizer; and $^+$ denotes the underlying problem is scaled. We use HHC-Algo to denote Algorithm 4.1 with the merit function introduced in [15] being used and Pen-Algo to denote Algorithm 4.1 with the penalty merit function introduced in this paper being used.

The numerical results for $\phi_{\theta,p}$ have only very slightly differences from those in [15], which are not listed here considering the length of this paper; the numerical results for $\phi_\alpha$ are given in Tables 1–3, i.e., the cross of row $\theta = 1$ and column $p = 2$ for every cases in the tables; and the numerical results for $\phi_\lambda$ with $\lambda = 0.95$ are listed in the row cck for every cases in Tables 1–3. The reason for choosing $\lambda = 0.95$ is that [2] reports encouraging numerical results for $\phi_\lambda$ when $\lambda = 0.95$. Since $\phi_\lambda$ works better for these problems when the problems unscaled than scaled, we present the numerical results of $\phi_\lambda$ with the tested problems unscaled.

From Tables 1–3 and the numerical results for $\phi_{\theta,p}$, we have the following numerical results.

- *Comparison of $\psi_{\alpha,\theta,p}$ with $\phi_{\theta,p}$.* Altogether, 324 cases of 18 problems were tested. There is one case for which both Pen-Algo and HHC-Algo fail; and there are

2 cases for which the iterative numbers of both Pen-Algo and HHC-Algo are equal. In the rest 321 cases, there are 3 cases for which Pen-Algo fails but HHC-Algo successes; while there are 2 cases for which HHC-Algo fails but Pen-Algo successes. For the rest 316 cases, there are 236 cases in which the iterative numbers of using Pen-Algo to solve is less than those by using HHC-Algo; while there are 80 cases in which the iterative numbers of using HHC-Algo to solve is less than those by using Pen-Algo. In addition, there are similar numerical results for the dual gap. Since the data is large in the comparison of $\psi_{\alpha,\theta,p}$ with $\phi_{\theta,p}$, we will give the performance profiles with respect to the iterative number and the dual gap at the end of this section.

- *Comparison of* $\psi_{\alpha,\theta,p}$ *with* $\phi_\alpha$. In Tables 1–3, the numerical results of the algorithm using $\phi_\alpha$ are listed in rows of $\theta = 1$ and $p = 2$. For 16 problems of total 18 test problems, there are many cases of the algorithm using $\psi_{\alpha,\theta,p}$ whose iterative numbers are less than the one of the algorithm using $\phi_\alpha$ for the same problems. Consider one system of testings of the algorithm using $\psi_{\alpha,\theta,p}$, say, the case of $\theta = 0.75$ and $p = 2$. Then, it is easy to see that there are 10 problems for which the iterative numbers of the algorithm using $\psi_{\alpha,\theta,p}$ in the case of $\theta = 0.75$ and $p = 2$ are less than the one of the algorithm using $\phi_\alpha$; there are 7 problems for converse case; and there is one problem for which the iterative numbers of both algorithms are equal.
- *Comparison of* $\psi_{\alpha,\theta,p}$ *with* $\phi_\lambda$. In Tables 1–3, the numerical results of the algorithm using $\phi_\lambda$ are listed rows of cck. There is one problem for which the algorithm using $\phi_\lambda$ fails. For 11 problems of other 17 test problems, there are many cases of the algorithm using $\psi_{\alpha,\theta,p}$ whose iterative numbers are less than the one of the algorithm using $\phi_\lambda$ for the same problems. Consider one system of testings of the algorithm using $\psi_{\alpha,\theta,p}$, say, the case of $\theta = 1$ and $p = 3$. Then, it is easy to see that there are 9 problems for which the iterative numbers of the algorithm using $\psi_{\alpha,\theta,p}$ in the case of $\theta = 1$ and $p = 3$ are less than the one of the algorithm using $\phi_\lambda$; there are 8 problems for converse case; and there is one problem for which the iterative number the algorithm using $\psi_{\alpha,\theta,p}$ is 114, but the algorithm using $\phi_\lambda$ fails.

These demonstrates that the algorithm using the function proposed in this paper is comparable to the algorithm using the functions proposed in [2, 15, 31].

It is well known that the performance profiles are important for the comparisons, when the number of the test problems is large. In our numerical computation, the numbers of the test problems using Pen-Algo and HHC-Algo are large. Thus, in the following, we use performance profile introduced in [8] to obtain an overall assessment of the performance of every solver of Pen-Algo and HHC-Algo. For this purpose, we first simply introduce the method of performance profile. We use $\mathcal{S}$ to denote the set of solvers and $\mathcal{P}$ to denote the set of the test problems; and assume that we have $n_s$ solvers and $n_q$ problems. Then, $\mathcal{S} = \{\text{Pen-Algo}, \text{HHC-Algo}\}$; $n_s = 2$; and $n_q = 324$. For every problem $q$ and solver $s$, we use $t^1_{q,s}$ (respectively, $t^2_{q,s}$) to denote the iterative number (respectively, the dual gap) of solver $s$ solving problem $q$, and define

$$r^i_{q,s} := \frac{t^i_{q,s}}{\min\{t^i_{q,s} : s \in \mathcal{S}\}} \quad \text{and} \quad \rho^i_s(\tau) = \frac{1}{n_q}\text{size}\{q \in \mathcal{P} : r^i_{q,s} \le \tau\}, \quad i \in \{1, 2\},$$

**Table 1** GAP($10^{-3}$)

| Problem | $\theta$ | $p=1.5$ | | | $p=2$ | | | $p=3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | GAP | NF | IT | GAP | NF | IT | GAP | NF | IT |
| spppe(1) | 0.1 | 10 | 105919 | 13934 | 9.97 | 102046 | 13422 | 9.85 | 107490 | 14018 |
| | 0.25 | 9.52 | 103677 | 14215 | | * | | 9.91 | 100913 | 13374 |
| | 0.5 | 9.95 | 95390 | 13060 | 9.97 | 99992 | 13119 | 8.58 | 98053 | 12987 |
| | 0.75 | 9.99 | 108289 | 13996 | 10 | 94214 | 12602 | 9.97 | 114051 | 14615 |
| | 0.9 | 9.61 | 121949 | 15408 | 8.86 | 137203 | 17436 | 9.79 | 132537 | 16897 |
| | 1 | 9.82 | 151883 | 19077 | 9.57 | 172375 | 21112 | 9.03 | 192192 | 23098 |
| | cck | | | | 9.47 | 15284 | 4118 | | | |
| sppe($2^-$) | 0.1 | 9.07 | 20598 | 4308 | 9.96 | 22748 | 4772 | 9.69 | 21906 | 4534 |
| | 0.25 | 9.89 | 22618 | 4789 | 9.82 | 22573 | 4717 | 9.61 | 23629 | 4964 |
| | 0.5 | 9.94 | 25578 | 5243 | 10 | 29320 | 5899 | 9.76 | 30405 | 6072 |
| | 0.75 | 9.9 | 28778 | 5951 | 9.66 | 28370 | 5847 | 9.59 | 30838 | 6355 |
| | 0.9 | 9.62 | 31251 | 6518 | 9.85 | 33034 | 6793 | 9.98 | 31250 | 6532 |
| | 1 | 9.65 | 31530 | 6647 | 9.88 | 29844 | 6343 | 9.98 | 36351 | 7555 |
| | cck | | | | 9.83 | 14179 | 3860 | | | |
| nash($1^-$) | 0.1 | 3.58 | 921 | 219 | 3.55 | 878 | 207 | 1.82 | 918 | 215 |
| | 0.25 | 4.1 | 669 | 162 | 4.87 | 422 | 102 | 3.76 | 615 | 150 |
| | 0.5 | 2.81 | 983 | 233 | 1.5 | 2228 | 523 | 1.56 | 2352 | 552 |
| | 0.75 | 0.626 | 481 | 127 | 6 | 585 | 154 | 8.63 | 494 | 129 |
| | 0.9 | 0.441 | 788 | 216 | 0.0886 | 561 | 156 | 1.08 | 395 | 107 |
| | 1 | 8.09 | 435 | 122 | 1.43 | 1069 | 301 | 0.931 | 600 | 169 |
| | cck | | | | 2.49 | 38 | 18 | | | |
| nash($2^\pm$) | 0.1 | 0.485 | 429 | 183 | 2.15 | 493 | 210 | 2.13 | 510 | 219 |
| | 0.25 | 1.22 | 2859 | 1228 | 1.22 | 6187 | 2655 | 1.22 | 6482 | 2782 |
| | 0.5 | 1.56 | 773 | 356 | 1.69 | 712 | 327 | 1.71 | 708 | 325 |
| | 0.75 | 1.47 | 330 | 186 | 1.5 | 234 | 133 | 9.91 | 219 | 124 |
| | 0.9 | 8.73 | 193 | 121 | 7.78 | 225 | 134 | 8.91 | 198 | 125 |
| | 1 | 4.16 | 203 | 134 | 7.14 | 228 | 160 | 7.06 | 228 | 162 |
| | cck | | | | 2.11 | 36 | 18 | | | |
| cycle$^\pm$ | 0.1 | 2.27 | 9 | 8 | 2.26 | 9 | 8 | 2.26 | 9 | 8 |
| | 0.25 | 4.21 | 9 | 8 | 4.14 | 9 | 8 | 4.14 | 9 | 8 |
| | 0.5 | 3.74 | 10 | 9 | 3.64 | 10 | 9 | 3.64 | 10 | 9 |
| | 0.75 | 3.24 | 11 | 10 | 3.13 | 11 | 10 | 3.13 | 11 | 10 |
| | 0.9 | 4.76 | 11 | 10 | 4.6 | 11 | 10 | 4.59 | 11 | 10 |
| | 1 | 5.95 | 11 | 10 | 5.74 | 11 | 10 | 5.73 | 11 | 10 |
| | cck | | | | 7.45 | 5 | 4 | | | |
| explcp | 0.1 | 0.646 | 182 | 110 | 2.02 | 179 | 107 | 1.82 | 214 | 129 |
| | 0.25 | 1.29 | 177 | 106 | 1.36 | 192 | 114 | 1.44 | 192 | 114 |
| | 0.5 | 1.21 | 229 | 139 | 2.57 | 314 | 194 | 1.53 | 326 | 202 |
| | 0.75 | 1.35 | 144 | 91 | 1.23 | 166 | 112 | 0.699 | 157 | 107 |
| | 0.9 | 0.913 | 112 | 80 | 1.19 | 53 | 37 | 1.48 | 39 | 26 |
| | 1 | 0.964 | 79 | 56 | 1.49 | 31 | 18 | 0.0642 | 24 | 11 |
| | cck | | | | 0.0829 | 13 | 7 | | | |

**Table 2** GAP($10^{-3}$)

| Problem | $\theta$ | $p=1.5$ | | | $p=2$ | | | $p=3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | GAP | NF | IT | GAP | NF | IT | GAP | NF | IT |
| gafni($1^+$) | 0.1 | 0.181 | 1098 | 359 | 0.175 | 761 | 248 | 0.201 | 502 | 162 |
| | 0.25 | 0.205 | 161 | 56 | 0.212 | 252 | 90 | 0.186 | 196 | 70 |
| | 0.5 | 0.192 | 141 | 51 | 0.167 | 167 | 61 | 0.287 | 169 | 62 |
| | 0.75 | 0.08 | 154 | 57 | 0.412 | 162 | 63 | 0.233 | 178 | 70 |
| | 0.9 | 0.263 | 164 | 63 | 0.072 | 133 | 52 | 0.0273 | 250 | 101 |
| | 1 | 0.369 | 240 | 98 | 0.208 | 224 | 93 | 0.234 | 272 | 114 |
| | cck | | | | | * | | | | |
| gafni($2^+$) | 0.1 | 0.191 | 538 | 172 | 0.177 | 438 | 141 | 0.199 | 408 | 131 |
| | 0.25 | 0.21 | 226 | 78 | 0.212 | 237 | 84 | 0.186 | 210 | 74 |
| | 0.5 | 0.175 | 149 | 54 | 0.154 | 159 | 58 | 0.192 | 162 | 60 |
| | 0.75 | 0.0422 | 128 | 49 | 0.622 | 136 | 54 | 0.515 | 128 | 51 |
| | 0.9 | 0.174 | 156 | 62 | 0.146 | 212 | 88 | 0.0245 | 200 | 82 |
| | 1 | 0.394 | 235 | 98 | 0.139 | 172 | 71 | 0.121 | 146 | 59 |
| | cck | | | | 5.7 | 7496 | 1751 | | | |
| gafni($3^+$) | 0.1 | 0.181 | 953 | 308 | 0.168 | 451 | 145 | 0.195 | 486 | 158 |
| | 0.25 | 0.192 | 320 | 109 | 0.176 | 262 | 89 | 0.19 | 234 | 81 |
| | 0.5 | 0.186 | 228 | 80 | 0.135 | 195 | 68 | 0.168 | 242 | 86 |
| | 0.75 | 0.029 | 180 | 62 | 0.712 | 164 | 58 | 0.645 | 210 | 80 |
| | 0.9 | 0.303 | 213 | 76 | 0.156 | 316 | 123 | 0.0156 | 206 | 75 |
| | 1 | 0.377 | 367 | 138 | 0.172 | 269 | 102 | 0.153 | 194 | 73 |
| | cck | | | | 9.14 | 2968 | 728 | | | |
| josephy($1^+$) | 0.1 | 2.89 | 1309 | 654 | 2.89 | 1739 | 869 | 2.89 | 1749 | 874 |
| | 0.25 | 2.79 | 806 | 403 | 2.86 | 886 | 443 | 2.82 | 896 | 448 |
| | 0.5 | 0.408 | 212 | 99 | 0.477 | 395 | 188 | 0.699 | 169 | 85 |
| | 0.75 | 0.608 | 347 | 175 | 0.0026 | 87 | 45 | 0.237 | 99 | 53 |
| | 0.9 | 0.0337 | 75 | 41 | 1.12 | 78 | 45 | 0.667 | 58 | 34 |
| | 1 | 1.6 | 38 | 24 | 2.09 | 30 | 22 | 2.07 | 29 | 19 |
| | cck | | | | 2.02 | 200 | 94 | | | |
| josephy($2^+$) | 0.1 | 2.91 | 1324 | 659 | 2.9 | 1754 | 874 | 2.89 | 1754 | 874 |
| | 0.25 | 2.82 | 818 | 406 | 2.83 | 892 | 443 | 2.83 | 912 | 453 |
| | 0.5 | 0.516 | 238 | 107 | 3.19 | 185 | 89 | 3.13 | 185 | 89 |
| | 0.75 | 0.214 | 111 | 52 | 0.357 | 90 | 42 | 0.00878 | 90 | 42 |
| | 0.9 | 1.01 | 150 | 78 | 0.57 | 78 | 40 | 2.18 | 81 | 42 |
| | 1 | 1.7 | 67 | 36 | 2.72 | 48 | 28 | 1.85 | 54 | 35 |
| | cck | | | | 1.99 | 175 | 83 | | | |
| josephy($3^+$) | 0.1 | 2.89 | 1464 | 717 | 2.9 | 1955 | 957 | 2.9 | 1897 | 940 |
| | 0.25 | 2.85 | 919 | 445 | 2.82 | 1044 | 512 | 2.8 | 1088 | 523 |
| | 0.5 | 0.476 | 409 | 178 | 0.476 | 536 | 250 | 0.404 | 539 | 252 |
| | 0.75 | 0.219 | 273 | 128 | 0.315 | 301 | 145 | 0.262 | 312 | 152 |
| | 0.9 | 0.136 | 260 | 124 | 1.39 | 298 | 136 | 1.25 | 258 | 118 |
| | 1 | 0.0118 | 276 | 146 | 2.23 | 82 | 47 | 2.48 | 36 | 23 |
| | cck | | | | 2.02 | 302 | 153 | | | |

**Table 3** GAP($10^{-3}$)

| Problem | $\theta$ | $p=1.5$ | | | $p=2$ | | | $p=3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | GAP | NF | IT | GAP | NF | IT | GAP | NF | IT |
| josephy($4^+$) | 0.1 | 4.67 | 82 | 40 | 4.83 | 82 | 40 | 4.83 | 82 | 40 |
| | 0.25 | 3.8 | 104 | 51 | 3.96 | 104 | 51 | 3.97 | 104 | 51 |
| | 0.5 | 0.893 | 254 | 126 | 0.821 | 252 | 126 | 0.881 | 250 | 125 |
| | 0.75 | 0.162 | 87 | 45 | 0.0856 | 90 | 46 | 0.216 | 70 | 36 |
| | 0.9 | 0.306 | 57 | 32 | 0.632 | 38 | 22 | 0.29 | 54 | 32 |
| | 1 | 0.767 | 36 | 29 | 0.662 | 51 | 40 | 1.17 | 40 | 34 |
| | cck | | | | 2.02 | 215 | 101 | | | |
| josephy($5^+$) | 0.1 | 2.8 | 91 | 49 | 1.3 | 68 | 38 | 1.38 | 97 | 51 |
| | 0.25 | 0.0479 | 62 | 36 | 1.09 | 60 | 37 | 2.37 | 52 | 31 |
| | 0.5 | 0.0562 | 41 | 26 | 0.851 | 29 | 17 | 0.645 | 29 | 17 |
| | 0.75 | 0.711 | 19 | 12 | 0.249 | 24 | 18 | 1.25 | 23 | 18 |
| | 0.9 | 1.19 | 30 | 23 | 0.736 | 35 | 29 | 0.511 | 37 | 31 |
| | 1 | 1.15 | 46 | 34 | 1.1 | 56 | 43 | 0.929 | 59 | 46 |
| | cck | | | | 2.77 | 132 | 60 | | | |
| josephy(6) | 0.1 | 1.26 | 377 | 117 | 1.08 | 313 | 97 | 1.02 | 321 | 99 |
| | 0.25 | 1.21 | 377 | 118 | 0.859 | 275 | 86 | 0.982 | 249 | 78 |
| | 0.5 | 1.43 | 195 | 66 | 1.21 | 195 | 68 | 0.959 | 145 | 50 |
| | 0.75 | 5.45 | 168 | 64 | 2.28 | 421 | 161 | 1.24 | 1159 | 457 |
| | 0.9 | 7.46 | 196 | 73 | 2.75 | 340 | 131 | 7.5 | 219 | 82 |
| | 1 | 4.09 | 147 | 56 | 3.18 | 670 | 267 | 7.41 | 200 | 76 |
| | cck | | | | 3.15 | 162 | 75 | | | |
| kojshin(1) | 0.1 | 2 | 501 | 158 | 2.09 | 330 | 104 | 2.09 | 330 | 104 |
| | 0.25 | 8.22 | 276 | 91 | 5.54 | 280 | 93 | 5.56 | 280 | 93 |
| | 0.5 | 6.36 | 151 | 54 | 8.79 | 230 | 84 | 8.3 | 209 | 73 |
| | 0.75 | 9.06 | 178 | 66 | 9.46 | 208 | 77 | 9.32 | 203 | 74 |
| | 0.9 | 8.39 | 234 | 87 | 8.56 | 235 | 88 | 8.58 | 231 | 87 |
| | 1 | 2.38 | 361 | 141 | 5.74 | 186 | 70 | 9.17 | 249 | 96 |
| | cck | | | | 4.13 | 790 | 270 | | | |
| kojshin($2^-$) | 0.1 | 2.51 | 337 | 107 | 2.75 | 726 | 169 | | * | |
| | 0.25 | 3.38 | 339 | 119 | 3.18 | 358 | 122 | 2.74 | 534 | 194 |
| | 0.5 | 3.64 | 234 | 86 | 5.45 | 237 | 88 | 6.96 | 255 | 96 |
| | 0.75 | 1.27 | 841 | 351 | 3.7 | 498 | 207 | 3.51 | 482 | 198 |
| | 0.9 | 2.05 | 210 | 81 | 6.97 | 300 | 123 | 5.93 | 335 | 132 |
| | 1 | 5.53 | 252 | 119 | 5.31 | 447 | 209 | 5.21 | 355 | 159 |
| | cck | | | | 4.94 | 160 | 86 | | | |
| kojshin($3^+$) | 0.1 | 4.68 | 532 | 211 | 7.42 | 561 | 221 | 5.7 | 585 | 227 |
| | 0.25 | 6.12 | 567 | 218 | 6.98 | 519 | 202 | 6.31 | 544 | 216 |
| | 0.5 | 8.15 | 693 | 278 | 7.19 | 598 | 239 | 7.73 | 716 | 292 |
| | 0.75 | 9.35 | 729 | 297 | 8.26 | 837 | 349 | 8.38 | 915 | 389 |
| | 0.9 | | * | | 4.32 | 660 | 289 | | * | |
| | 1 | 6.86 | 551 | 276 | 6.54 | 285 | 184 | 6.66 | 185 | 137 |
| | cck | | | | 4.94 | 262 | 142 | | | |

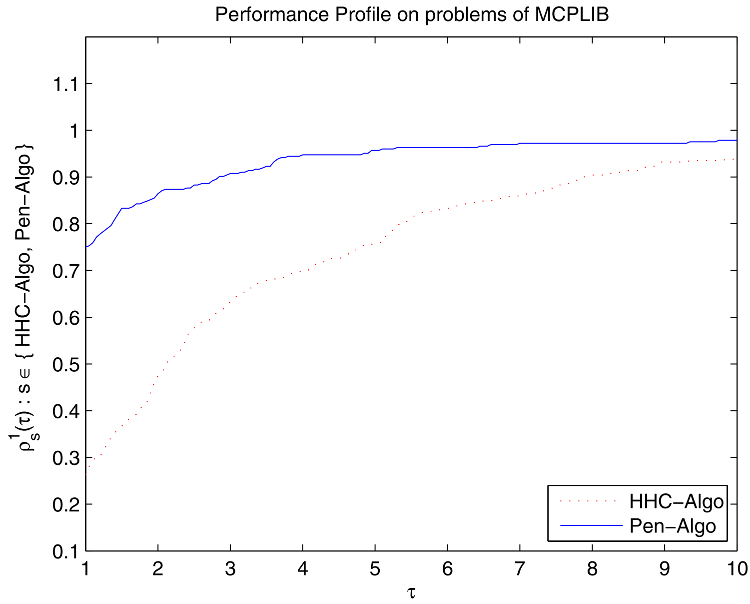


**Fig. 1** Performance profile with respect to the iterative number: local

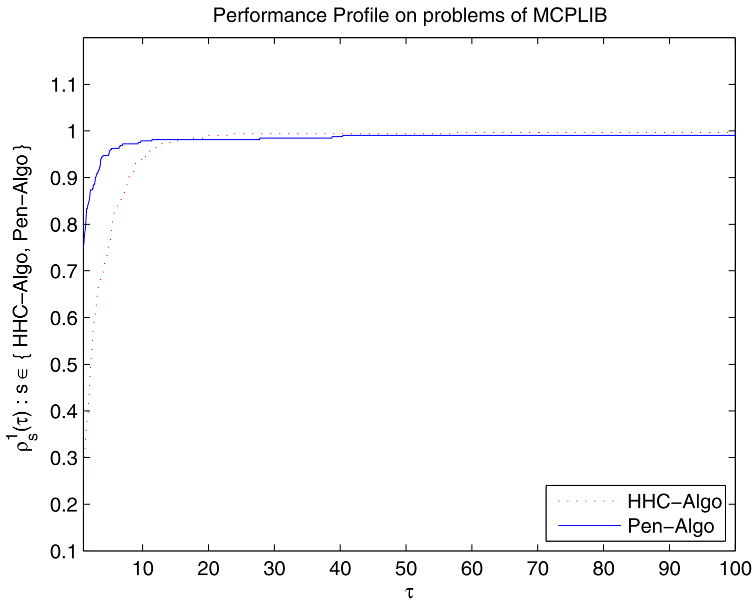


**Fig. 2** Performance profile with respect to iterative number: global

then every $\rho_s^i(\tau)$ is the distribution function for $r_{q,s}^i$ satisfying $r_{q,s}^i \leq \tau$. Thus, a plot of the performance profile reveals all of the major performance characteristics. The

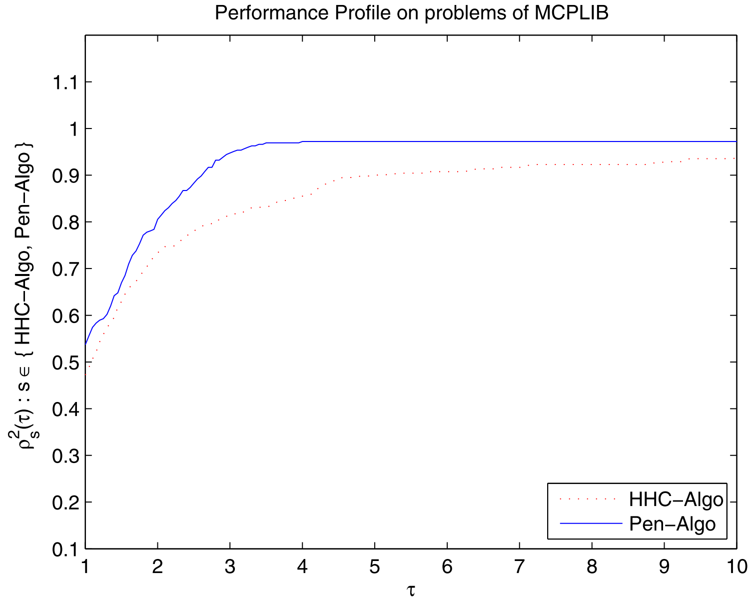


**Fig. 3** Performance profile with respect to the dual gap: local

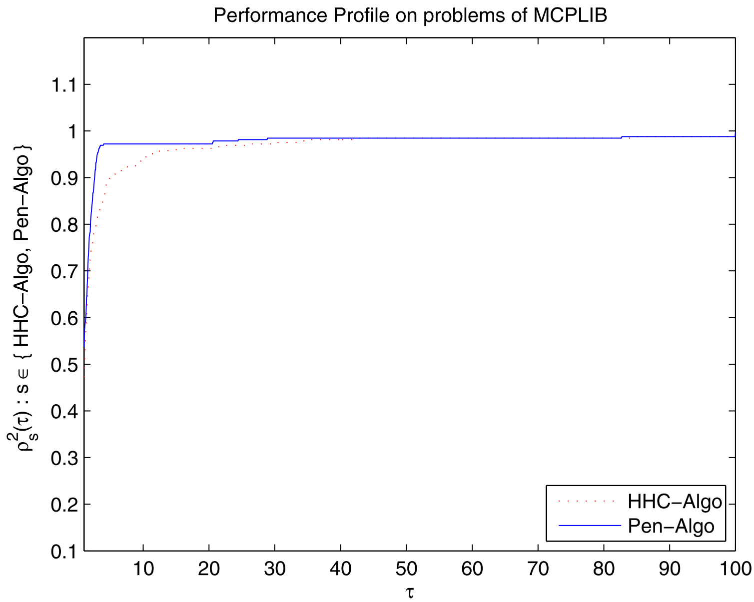


**Fig. 4** Performance profile with respect to the dual gap: global

performance profile with respect to the iterative number (respectively, the dual gap) have been mapped to Figs. 1 and 2 (respectively, Figs. 3 and 4). It is well known that solvers with large probability $\rho_s^i(\tau)$ are to be preferred. Thus, from Figs. 1–4, it is easy to see that the numerical results of the algorithm using the function proposed in this paper is comparable to the algorithm using the function proposed in [15].

## 5 Conclusions

In this paper, we have considered the merit function $\Psi_{\alpha,\theta,p}$ which includes many existing well-known merit functions for NCP (1) in the literature. Although this merit $\Psi_{\alpha,\theta,p}$ is obtained by penalizing another merit function $\Psi_{\theta,p}$ studied in [15], we have explored more properties for $\Psi_{\alpha,\theta,p}$, see Sect. 3. It is worth to point out that Lemma 3.3 not only plays an important role in Proposition 3.1, but also may be very useful in other contexts. We also suspect that Proposition 3.2 may be crucial in analyzing convergence rate of certain algorithms which is our future topic.

## 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 (1977)
2. Chen, B., Chen, X., Kanzow, C.: A penalized Fischer-Burmeister NCP-function: theoretical investigation and numerical results. Math. Programm. **88**, 211–216 (2000)
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. Chen, J.-S., Pan, S.: A family of NCP-functions and a descent method for the nonlinear complementarity problem. Comput. Optim. Appl. **40**, 389–404 (2008)
5. Chen, J.-S. Gao, H.-T., Pan, S.: A $R$-linearly convergent derivative-free algorithm for the NCPs based on the generalized Fischer-Burmeister merit function. J. Comput. Appl. Math. **232**, 455–471 (2009)
6. Cottle, R.W., Pang, J.-S., Stone, R.-E.: The Linear Complementarity Problem. Academic Press, New York (1992)
7. Dafermos, S.: An iterative scheme for variational inequalities. Math. Programm. **26**, 40–47 (1983)
8. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles. Math. Programm. **91**, 201–213 (2002)
9. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
10. Facchinei, F., Soares, J.: A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. **7**, 225–247 (1997)
11. Fischer, A.: A special Newton-type optimization methods. Optimization **24**, 269–284 (1992)
12. Geiger, C., Kanzow, C.: On the resolution of monotone complementarity problems. Comput. Optim. Appl. **5**, 155–173 (1996)
13. Grippo, L., Lampariello, F., Lucidi, S.S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
14. Harker, P.T., Pang, J.-S.: Finite dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications. Math. Programm. **48**, 161–220 (1990)
15. Hu, S.-L., Huang, Z.-H., Chen, J.-S.: Properties of a family of generalized NCP-functions and a derivative free algorithm for complementarity problems. J. Comput. Appl. Math. **230**, 69–82 (2009)
16. Huang, Z.-H.: The global linear and local quadratic convergence of a non-interior continuation algorithm for the LCP. IMA J. Numer. Anal. **25**, 670–684 (2005)
17. Huang, Z.-H., Gu, W.-Z.: A smoothing-type algorithm for solving linear complementarity problems with strong convergence properties. Appl. Math. Optim. **57**, 17–29 (2008)
18. Huang, Z.-H., Qi, L., Sun, D.: Sub-quadratic convergence of a smoothing Newton algorithm for the $P_0$-and monotone LCP. Math. Programm. **99**, 423–441 (2004)
19. Jiang, H.: Unconstrained minimization approaches to nonlinear complementarity problems. J. Glob. Optim. **9**, 169–181 (1996)

20. Jiang, H., Fukushima, M., Qi, L.: A trust region method for solving generalized complementarity problems. SIAM J. Optim. **8**, 140–157 (1998)
21. Kanzow, C.: Nonlinear complementarity as unconstrained optimization. J. Optim. Theory Appl. **88**, 139–155 (1996)
22. Kanzow, C., Kleinmichel, H.: A new class of semismooth Newton method for nonlinear complementarity problems. Comput. Optim. Appl. **11**, 227–251 (1998)
23. Kanzow, C., Yamashita, N., Fukushima, M.: New NCP-functions and their properties. J. Optim. Theory Appl. **94**, 115–135 (1997)
24. Mangasarian, O.L.: Equivalence of the complementarity problem to a system of nonlinear equations. SIAM J. Appl. Math. **31**, 89–92 (1976)
25. Pang, J.-S.: Newton's method for $B$-differentiable equations. Math. Oper. Res. **15**, 311–341 (1990)
26. Pang, J.-S.: Complementarity problems. In: Horst, R., Pardalos, P. (eds.) Handbook of Global Optimization, pp. 271–338. Kluwer Academic, Boston (1994)
27. Pang, J.-S., Chan, D.: Iterative methods for variational and complemantarity problems. Math. Programm. **27**(99), 284–313 (1982)
28. Sun, D., Qi, L.: On NCP-functions. Comput. Optim. Appl. **13**, 201–220 (1999)
29. Yamashita, N., Fukushima, M.: On stationary points of the implicit Lagrangian for nonlinear complementarity problems. J. Optim. Theory Appl. **84**, 653–663 (1995)
30. Yamashita, N., Fukushima, M.: Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems. Math. Programm. **76**, 469–491 (1997)
31. Yamada, K., Yamashita, N., Fukushima, M.: A new derivative-free descent method for the nonlinear complementarity problem. In: Di Pillo, G., Giannessi, F. (eds.) Nonlinear Optimization and Related Topics, pp. 463–487. Kluwer Academic, Amsterdam (2000)