Asia-Pacific Journal of Operational Research, vol.24, pp. 401-420, 2007

On some NCP-functions based on the generalized Fischer-Burmeister function

Jein-Shan Chen [1]
Department of Mathematics
National Taiwan Normal University
Taipei, Taiwan 11677

June 2, 2005
(revised November 18, 2005)
(second revised January 11, 2006)

Abstract In this paper, we study several NCP-functions for the nonlinear complementarity problem (NCP) which are indeed based on the generalized Fischer-Burmeister function, $\phi_p(a,b) = \|(a,b)\|_p - (a+b)$. It is well known that the NCP can be reformulated as an equivalent unconstrained minimization by means of merit functions involving NCP-functions. Thus, we aim to investigate some important properties of these NCP-functions that will be used in solving and analyzing the reformulation of the NCP.

Key words. NCP-function, complementarity, merit function, bounded level sets, stationary point.

1 Introduction

The nonlinear complementarity problem (NCP) [13, 20] 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 $F = (F_1, F_2, \cdots, F_n)^T$ maps 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 [9, 13, 20].

There have been many methods proposed for solving the NCP [13, 20]. 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] or as an

[1]E-mail: jschen@math.ntnu.edu.tw, TEL: 886-2-29325417, FAX: 886-2-29332342.

unconstrained minimization problem [8, 10, 14]. 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.

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

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

Many NCP-functions and merit functions have been explored during the past two decades [6, 16, 25, 26]. Among which, a popular NCP-function intensively studied recently is the well-known Fischer-Burmeister NCP-function [10, 11] defined as

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

With the above characterization of ϕ_{FB}, the NCP is equivalent to a system of nonsmooth equations:

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

For each NCP-function, there is a natural merit function, $\Psi_{\text{FB}} : \mathbb{R}^n \to \mathbb{R}_+$ given by

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

from which the NCP can be recast as an unconstrained minimization:

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

In this paper, we are particularly interested in the generalized Fischer-Burmeister function, i.e., $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ given by

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

where p is a positive integer greater than one and $\|(a, b)\|_p = \sqrt[p]{|a|^p + |b|^p}$ means the p-norm of (a, b). Notice that ϕ_p reduces to the well known Fischer-Burmeister function ϕ_{FB} when $p = 2$ and its related properties were recently presented in [3, 4]. Corresponding to ϕ_p, we define $\psi_p : \mathbb{R}^2 \to \mathbb{R}_+$ by

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

Then both ϕ_p and ψ_p are NCP-functions and yield a merit function

$$\Psi_p(x) := \sum_{i=1}^{n} \psi_p(x_i\ , F_i(x)) = \frac{1}{2}\sum_{i=1}^{n} \phi_p(x_i\ , F_i(x))^2, \tag{9}$$

from which the NCP can be reformulated as an unconstrained minimization:

$$\min_{x\in\mathbb{R}^n} \Psi_p(x). \tag{10}$$

However, there has some limitations for the (generalized) Fischer-Burmeister functions and some of its variants when dealing with monotone complementarity problem. In particular, its natural merit function Ψ_p does not guarantee bounded level sets for this class of problem which is an important class (see page 4 of [1]). Some modifications to the Fischer-Burmeister have been proposed to conquer the above problem, see [16, 25]. In this paper, we extend these modifications to the generalized Fischer-Burmeister function ϕ_p. More specifically, we study the following NCP-functions:

$$\begin{array}{lll} \phi_1(a,b) := & \phi_p(a,b) - \alpha a_+ b_+, & \alpha > 0, \\ \phi_2(a,b) := & \phi_p(a,b) - \alpha (ab)_+, & \alpha > 0, \\ \phi_3(a,b) := & \sqrt{[\phi_p(a,b)]^2 + \alpha (a_+ b_+)^2}, & \alpha > 0, \\ \phi_4(a,b) := & \sqrt{[\phi_p(a,b)]^2 + \alpha [(ab)_+]^2}, & \alpha > 0, \end{array} \tag{11}$$

The function ϕ_1 is called penalized Fischer-Burmeister function when $p = 2$ and was studied in [1]. The functions ϕ_2, ϕ_3, ϕ_4 generalize the merit functions of $p = 2$, which were discussed in [25, 27]. Note that for $i = 1, 2, 3, 4$, we have

$$\phi_i(a,b) \equiv \phi_p(a,b) \tag{12}$$

for all $(a, b) \in N_-$ (this notation is used in [25]) where

$$N_- := \{(a,b)|\ ab \le 0\}. \tag{13}$$

Thus, ϕ_i where $i = 1, 2, 3, 4$ are only different in the first or third quadrant.

Similarly, for each ϕ_i there is an associated $\psi_i : \mathbb{R}^2 \to \mathbb{R}_+$ given by

$$\psi_i(a,b) := \frac{1}{2}|\phi_i(a,b)|^2 \quad i = 1, 2, 3, 4, \tag{14}$$

which is also an NCP-function for every i. Moreover, for $\phi \in \{\phi_1, \phi_2, \phi_3, \phi_4\}$, we can define

$$\Phi(x) = \begin{pmatrix} \phi(x_1\ , F_1(x)) \\ \cdot \\ \cdot \\ \cdot \\ \phi(x_n\ , F_n(x)) \end{pmatrix}, \tag{15}$$

from which the NCP is equivalent to the unconstrained minimization:

$$\min_{x\in\mathbb{R}^n} \Psi(x) \tag{16}$$

where

$$\Psi(x) := \frac{1}{2}\|\Phi(x)\|^2 = \frac{1}{2}\sum_{i=1}^{n} \phi(x_i\ , F_i(x))^2 \tag{17}$$

is the natural merit function corresponding to $\phi \in \{\phi_1, \phi_2, \phi_3, \phi_4\}$.

The paper is organized as follows. In Sec. 2, we review some background definitions including monotonicity, P_0-function, semismoothness, e.t.c. and known results about Ψ_p and its related properties. In Sec. 3, we show that all $(\phi_i)^2, i \in \{1, 2, 3, 4\}$ are continuously differentiable and investigate properties of the merit function Ψ constructed via ϕ_i with $i \in \{1, 2, 3, 4\}$. In particular, it provides bounded level sets for a monotone NCP with a strictly feasible point. In addition, we give conditions under which a stationary point of Ψ is a solution of the NCP. In general, the analytic techniques used in this paper are similar to those in [1, 8, 25] since the work is somewhat considered the extensions of NCP-functions studied in those literatures.

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 $F = (F_1, ..., F_m)^T : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x)\ \cdots\ \nabla F_m(x)]$ 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 this whole paper, we assume p is a positive integer greater than one.

2 Preliminaries

In this section, we recall some background concepts and materials which will play an important role in the subsequent analysis. Let $F : \mathbb{R}^n \to \mathbb{R}^n$. Then, (1) F is monotone if $\langle x - y, F(x) - F(y)\rangle \geq 0$, for all $x, y \in \mathbb{R}^n$. (2) F is strictly monotone if $\langle x - y, F(x) - F(y)\rangle > 0$, for all $x, y \in \mathbb{R}^n$ and $x \neq y$. (3) F is strongly monotone with modulus $\mu > 0$ if $\langle x - y, F(x) - F(y)\rangle \geq \mu\|x - y\|^2$, for all $x, y \in \mathbb{R}^n$. (4) F is a P_0-function if $\max\limits_{\substack{1\leq i\leq n\\ x_i\neq y_i}}(x_i - y_i)(F_i(x) - F_i(y)) \geq 0$, for all $x, y \in \mathbb{R}^n$ and $x \neq y$. (5) F is a P-function if $\max\limits_{1\leq i\leq n}(x_i - y_i)(F_i(x) - F_i(y)) > 0$, for all $x, y \in \mathbb{R}^n$ and $x \neq y$. (6) F is a uniform P-function with modulus $\mu > 0$ if $\max\limits_{1\leq i\leq n}(x_i - y_i)(F_i(x) - F_i(y)) \geq \mu\|x - y\|^2$, for all $x, y \in \mathbb{R}^n$. (7) F is a R_0-function if for every sequence $\{x^k\}$ satisfying $\{\|x^k\|\} \to \infty$, $\liminf\limits_{k\to\infty} \dfrac{\min_i x_i^k}{\|x\|^k} \geq 0$, and $\liminf\limits_{k\to\infty} \dfrac{\min_i F_i(x^k)}{\|x\|^k} \geq 0$, there exists an index j such that

$\{x_j^k\} \to \infty$ and $\{F_j(x^k)\} \to \infty$.

It is clear 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: (i) F is monotone if and only if $\nabla F(x)$ is positive semi-definite for all $x \in \mathbb{R}^n$. (ii) F is strictly monotone if $\nabla F(x)$ is positive definite for all $x \in \mathbb{R}^n$. (iii) F is strongly monotone if and only if $\nabla F(x)$ is uniformly positive definite. An R_0-function can be viewed as a generalization of a uniform P-function since every uniform P-function is an R_0-function (see [2, Prop. 3.11]).

A matrix $M \in \mathbb{R}^{n\times n}$ is a P_0-matrix if every of its principal minors is nonnegative, and it is a P-matrix if every of its principal minors is positive. In addition, it is said to be a R_0-matrix if the following system has only zero solution:

$$\begin{aligned} x &\geq 0, \\ M_i x &= 0 \quad if \quad x_i > 0, \\ M_i x &\geq 0 \quad if \quad x_i = 0, \end{aligned}$$

It is obvious that every P-matrix is also a P_0-matrix and it is known that the Jacobian of every continuously differentiable P_0-function is a P_0-matrix. For more properties about P-matrix and P_0-matrix, please refer to [7]. It is also known that F is an R_0-function if and only if M is an R_0-matrix when F is an affine function, see [2, Prop. 3.10].

Next, we recall the definition of semismoothness. First, we introduce that F is strictly continuous (also called 'locally Lipschitz continuous') at $x \in \mathbb{R}^n$ [24, Chap. 9] if there exist scalars $\kappa > 0$ and $\delta > 0$ such that

$$\|F(y) - F(z)\| \leq \kappa \|y - z\| \quad \forall y, z \in \mathbb{R}^n \text{ with } \|y - x\| \leq \delta,\ \|z - x\| \leq \delta;$$

and F is strictly continuous if F is strictly continuous at every $x \in \mathbb{R}^n$. If δ can be taken to be ∞, then F is Lipschitz continuous with Lipschitz constant κ. We say F is directionally differentiable at $x \in \mathbb{R}^n$ if

$$F'(x; h) \ := \ \lim_{t \to 0^+} \frac{F(x + th) - F(x)}{t} \quad \text{exists} \qquad \forall h \in \mathbb{R}^n;$$

and F is directionally differentiable if F is directionally differentiable at every $x \in \mathbb{R}^n$.

Assume $F : \mathbb{R}^n \to \mathbb{R}^m$ is strictly continuous. We say F is semismooth at x if F is directionally differentiable at x and, for any $V \in \partial F(x + h)$ (the generalized Jacobian), we have

$$F(x + h) - F(x) - Vh = o(\|h\|).$$

We say F is ρ-order semismooth at x $(0 < \rho < \infty)$ if F is semismooth at x and, for any $V \in \partial F(x+h)$, we have

$$F(x+h) - F(x) - Vh = O(\|h\|^{1+\rho}).$$

We say F is semismooth (respectively, ρ-order semismooth) if F is semismooth (respectively, ρ-order semismooth) at every $x \in \mathbb{R}^k$. We say F is *strongly semismooth* if it is 1-order semismooth. Convex functions and piecewise continuously differentiable functions are examples of semismooth functions. Examples of strongly semismooth functions include piecewise linear functions and LC^1 *functions* meaning smooth functions with gradients being locally Lipschitz continuous (strictly continuous) [7, 22]. The composition of two (respectively, ρ-order) semismooth functions is also a (respectively, ρ-order) semismooth function. The property of semismoothness plays an important role in nonsmooth Newton methods [21, 23] as well as in some smoothing methods mentioned in the previous section. For extensive discussions of semismooth functions, see [11, 18, 23].

To end this section, we collect some useful properties of ϕ_p, ψ_p defined as in (7) and (8), respectively, that will be used in the subsequent analysis. All the proofs can be found in [3].

Property 2.1 *([3, Prop. 3.1, Lem. 3.1]) Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be defined as (7). Then*

(a) *ϕ_p is an NCP-function, i.e., it satisfies (2).*

(b) *ϕ_p is sub-additive, i.e., $\phi_p(w + w') \le \phi_p(w) + \phi(w')$ 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 $\alpha \ge 0$.*

(d) *ϕ_p is convex, i.e., $\phi_p(\alpha w + (1-\alpha)w') \le \alpha\phi_p(w) + (1-\alpha)\phi_p(w')$ for all $w, w' \in \mathbb{R}^2$ and $\alpha \ge 0$.*

(e) *ϕ_p is Lipschitz continuous with $L_1 = 1 + \sqrt{2}$, i.e., $|\phi_p(w) - \phi_p(w')| \le L_1\|w - w'\|$; or with $L_2 = 1 + 2^{(1-1/p)}$, i.e., $|\phi_p(w) - \phi_p(w')| \le L_2\|w - w'\|_p$ for all $w, w' \in \mathbb{R}^2$.*

(f) *ϕ_p is semismooth.*

(g) *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 $k \to \infty$.*

Property 2.2 *([3, Prop. 3.2]) Let ϕ_p, ψ_p be defined as (7) and (8), respectively. Then*

(a) *ψ_p is an NCP-function, i.e., it satisfies (2).*

(b) *$\psi_p(a, b) \ge 0$ 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 $\Longleftrightarrow$ $\phi_p(a,b) = 0$.*

(e) *$\nabla_a \psi_p(a,b) = 0 \Longleftrightarrow \nabla_b \psi_p(a,b) = 0 \Longleftrightarrow \phi_p(a,b) = 0$.*

From these properties, it was proved in [3] that $\Psi_p(x) \geq 0$ for all $x \in \mathbb{R}^n$ and $\Psi_p(x) = 0$ if and only if x solves the NCP (1), where $\Psi_p : \mathbb{R}^n \to \mathbb{R}$ is defined as (9). Moreover, suppose that the NCP has at least one solution. Then x is a global minimizer of Ψ_p if and only if x solves the NCP. In addition, it was also shown in [3] that if F is either monotone or P_0-function, then every stationary point of Ψ_p is a global minima of (10); and therefore solves the original NCP. We will investigate the analogous results for the merit function Ψ which is based on ϕ_i studied in this paper. On the other hand, as mentioned the natural merit function induced from the generalized Fischer-Burmeister (which behaves like the Fischer-Burmeister function) does not guarantee bounded level sets under the assumption of F being monotone. Instead, there needs that F is strongly monotone or uniform P-function to ensure that the property is held. Another main purpose of this work is to obtain same results for the merit function Ψ studied in this paper under the weaker assumption that F is monotone only (see Sec. 4).

3 Properties of ϕ and ψ

In this section, we investigate properties of $\phi \in \{\phi_1, \phi_2, \phi_3, \phi_4\}$ and $\psi \in \{\psi_1, \psi_2, \psi_3, \psi_4\}$ defined as in (11) and (14), respectively. These include strong semismoothness of ϕ and continuous differentiability of ψ. First, we denote

$$N_\phi := \{(a,b) |\ a \geq 0, b \geq 0, ab = 0\}. \tag{18}$$

This notation is adopted from [1] and it is easy to see that $(a,b) \in N_\phi$ if and only if (a,b) satisfies (2). Now we are ready to show the favorable properties of ϕ and ψ.

Proposition 3.1 *Let $\phi \in \{\phi_1, \phi_2, \phi_3, \phi_4\}$ be defined as in (11). Then*

(a) *$\phi(a,b) = 0 \Longleftrightarrow (a,b) \in N_\phi \Longleftrightarrow (a,b)$ satisfies (2).*

(b) *ϕ is strongly semismooth.*

(c) *Let $\{a^k\}, \{b^k\} \subseteq \mathbb{R}$ be any two sequences such that either $a_+^k b_+^k \to \infty$ or $a^k \to -\infty$ or $b^k \to -\infty$. Then $|\phi(a^k, b^k)| \to \infty$ for $k \to \infty$.*

Proof. (a) It is enough to prove the first equivalence. Suppose $\phi(a,b) = 0$, for $i = 2,3,4$, $\phi_i(a,b) = 0$ yields $\phi_p(a,b) = 0$ which says $(a,b) \in N_\phi$ by Property 2.1(a). For $i = 1$, $\phi_1(a,b) = 0$ implies $\phi_p(a,b) = \alpha a_+ b_+$. Since α could be any arbitrary positive number,

the above leads to $\phi_p(a,b) = a_+b_+ = 0$ which which says $(a,b) \in N_\phi$ by Property 2.1(a) again.

On the other hand, suppose $(a,b) \in N_\phi$ then $\phi_p(a,b) = 0$ by by Property 2.1(a). Since $a \geq 0, b \geq 0$, we obtain $a_+b_+ = ab = 0$. Hence we see that all $\phi_i(a,b) = 0, i = 1,2,3,4$.

(b) The verification of strong semismoothness of ϕ is a routine work which can be done as in [25, Lemma 1]. We omit it.

(c) This follows from Property 2.1(g) and definition of $(\cdot)_+$. $\square$

Proposition 3.2 *Let Φ be defined as in (15) with $\phi \in \{\phi_1, \phi_2, \phi_3, \phi_4\}$. Then*

(a) *Φ is semismooth.*

(b) *Φ is strongly semismooth if every F_i is LC^1 function.*

Proof. By using Prop. 3.1(b) and the fact that every LC^1 function is strongly semismooth, the results follow. $\square$

The following is a technical lemma which describes the generalized gradients of all $\phi_i, i = 1,2,3,4$ defined as in (11). It will be used for proving Prop. 3.3.

Lemma 3.1 *Let $\phi_1, \phi_2, \phi_3, \phi_4$ be defined as (11).*

(a) *The generalized gradient $\partial\phi_1(a,b)$ of ϕ_1 at a point (a,b) is equal to the set of all (v_a, v_b) such that*

$$(v_a, v_b) = \begin{cases} \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right) - \alpha\Big(b_+\partial a_+, a_+\partial b_+\Big), \\ \qquad\qquad if\ (a,b) \neq (0,0)\ and\ p\ is\ even, \\ \left(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right) - \alpha\Big(b_+\partial a_+, a_+\partial b_+\Big), \\ \qquad\qquad if\ (a,b) \neq (0,0)\ and\ p\ is\ odd, \\ (\xi - 1,\ \zeta - 1), \qquad if\ (a,b) = (0,0), \end{cases} \tag{19}$$

where (ξ, ζ) is any vector satisfying $|\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \leq 1$ and

$$\partial z_+ = \begin{cases} 1, & if\ z > 0, \\ [0,1], & if\ z = 0, \\ 0, & if\ z < 0. \end{cases}$$

(b) *The generalized gradient $\partial\phi_2(a,b)$ of ϕ_2 at a point (a,b) is equal to the set of all (v_a, v_b) such that*

$$
\begin{aligned}
(v_a, v_b) &= \Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) - \alpha(b,a), \\
&\quad if\ (a,b) \neq (0,0), ab > 0\ and\ p\ is\ even, \\
(v_a, v_b) &= \Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) - \alpha(b,a)\cdot[0,1], \\
&\quad if\ (a,b) \neq (0,0), ab = 0\ and\ p\ is\ even, \\
(v_a, v_b) &= \Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big), \\
&\quad if\ (a,b) \neq (0,0), ab < 0\ and\ p\ is\ even, \\
(v_a, v_b) &= \Big(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) - \alpha(b,a), \\
&\quad if\ (a,b) \neq (0,0), ab > 0\ and\ p\ is\ odd, \\
(v_a, v_b) &= \Big(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) - \alpha(b,a)\cdot[0,1], \\
&\quad if\ (a,b) \neq (0,0), ab = 0\ and\ p\ is\ odd, \\
(v_a, v_b) &= \Big(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big), \\
&\quad if\ (a,b) \neq (0,0), ab < 0\ and\ p\ is\ odd, \\
(v_a, v_b) &= (\xi - 1, \zeta - 1) - \alpha(b,a)\cdot[0,1], \\
&\quad if\ (a,b) = (0,0),
\end{aligned}
\tag{20}
$$

where (ξ,ζ) is any vector satisfying $|\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \leq 1$.

(c) *ϕ_3 is continuously differentiable everywhere except at $(0,0)$ with*

$$
\nabla_a\phi_3(a,b) = \begin{cases}
\dfrac{\phi_p(a,b)\cdot\left[\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right] + \alpha(a_+)(b_+)^2}{\phi_3(a,b)}, \\
\qquad if\ (a,b) \neq (0,0),\ and\ p\ is\ even, \\
\dfrac{\phi_p(a,b)\cdot\left[\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right] + \alpha(a_+)(b_+)^2}{\phi_3(a,b)}, \\
\qquad if\ (a,b) \neq (0,0),\ and\ p\ is\ odd,
\end{cases}
\tag{21}
$$

$$
\nabla_b\phi_3(a,b) = \begin{cases}
\dfrac{\phi_p(a,b)\cdot\left[\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right] + \alpha(a_+)^2(b_+)}{\phi_3(a,b)}, \\
\qquad if\ (a,b) \neq (0,0),\ and\ p\ is\ even, \\
\dfrac{\phi_p(a,b)\cdot\left[\frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right] + \alpha(a_+)^2(b_+)}{\phi_3(a,b)}, \\
\qquad if\ (a,b) \neq (0,0),\ and\ p\ is\ odd,
\end{cases}
\tag{22}
$$

and $\partial\phi_3(0,0) = (v_a, v_b)$ *where* $(v_a, v_b) \in (-\infty, \infty)$.

(d) ϕ_4 *is continuously differentiable everywhere except at* $(0,0)$ *with*

$$\nabla_a \phi_4(a,b) = \begin{cases} \dfrac{\phi_p(a,b)\cdot\left[\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}}-1\right]+\alpha(ab)_+\cdot b}{\phi_4(a,b)}, \\ \qquad if\ (a,b)\neq(0,0),\ \ and\ p\ is\ even, \\ \dfrac{\phi_p(a,b)\cdot\left[\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}}-1\right]+\alpha(ab)_+\cdot b}{\phi_4(a,b)}, \\ \qquad if\ (a,b)\neq(0,0),\ \ and\ p\ is\ odd, \end{cases} \tag{23}$$

$$\nabla_b \phi_4(a,b) = \begin{cases} \dfrac{\phi_p(a,b)\cdot\left[\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}}-1\right]+\alpha(ab)_+\cdot a}{\phi_4(a,b)}, \\ \qquad if\ (a,b)\neq(0,0),\ \ and\ p\ is\ even, \\ \dfrac{\phi_p(a,b)\cdot\left[\frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}}-1\right]+\alpha(ab)_+\cdot a}{\phi_4(a,b)}, \\ \qquad if\ (a,b)\neq(0,0),\ \ and\ p\ is\ odd, \end{cases} \tag{24}$$

and $\partial\phi_4(0,0) = (v_a, v_b)$ *where* $(v_a, v_b) \in (-\infty, \infty)$.

Proof. (a) First, we note that ϕ_p is continuously differentiable everywhere except at $(0,0)$ (see [3]). Hence, by the Corollary to Prop. 2.2.1 in [5], ϕ_p is strictly differentiable everywhere except at the origin. Let $\phi_+(a,b) := a_+ b_+$. Then ϕ_+ is strictly differentiable at the origin as proved in [1, Prop. 2.1]. Both ϕ_1 and ϕ_+ are strongly semismooth functions, we know that they are locally Lipschitz (strictly continuous) functions. Thus, the Corollary 2 to Prop. 2.3.3 in [5] yields

$$\partial\phi_1(a,b) = \partial\phi_p(a,b) - \alpha\cdot\partial\phi_+(a,b).$$

On the other hand, the generalized gradient of ϕ_p can be verified as below (see [3]):

$$\begin{aligned} \partial\phi_p(a,b) &= \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right),\ \text{if } (a,b)\neq(0,0) \text{ and p is even}, \\ \partial\phi_p(a,b) &= \left(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1,\ \frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right),\ \text{if } (a,b)\neq(0,0) \text{ and p is odd}, \\ \partial\phi_p(a,b) &= (\xi-1,\ \zeta-1),\ \text{if } (a,b)=(0,0), \end{aligned} \tag{25}$$

where (ξ,ζ) is any vector satisfying $|\xi|^{\frac{p}{p-1}} + |\zeta|^{\frac{p}{p-1}} \leq 1$. In addition, it was already shown in [1, Prop. 2.1] that

$$\partial\phi_+(a,b) = (b_+\partial a_+, a_+\partial b_+).$$

Thus, the desired results follow.

(b) Following the same arguments as in part(a) and using the fact that

$$\partial(ab)_+ = \begin{cases} (b,a), & if\ ab > 0, \\ (0,0), & if\ ab < 0, \\ (b,a)\cdot[0,1], & if\ ab = 0, \end{cases}$$

the desired results hold.

(c) It is known that $(\phi_p)^2$ and $(a_+b_+)^2$ are continuously differentiable. Then the desired results follow by direct computations using the chain rule and the fact that

$$\partial(\sqrt{z}) = \begin{cases} \dfrac{1}{2\sqrt{z}}, & if\ z > 0, \\ (-\infty,\ \infty), & if\ z = 0. \end{cases}$$

(d) Same arguments as part(c). □

Proposition 3.3 *Let $\psi \in \{\psi_1, \psi_2, \psi_3, \psi_4\}$ be defined as in (14). Then*

(a) $\psi(a,b) = 0 \Longleftrightarrow (a,b) \in N_\phi \Longleftrightarrow (a,b)$ *satisfies (2).*

(b) ψ *is continuously differentiable on* $\mathbb{R}^2$.

(c) $\nabla_a\psi(a,b)\cdot\nabla_b\psi(a,b) \geq 0$ *for all* $(a,b) \in \mathbb{R}^2$.

(d) $\psi(a,b) = 0 \Longleftrightarrow \nabla\psi(a,b) = 0 \Longleftrightarrow \nabla_a\psi(a,b) = 0 \Longleftrightarrow \nabla_b\psi(a,b) = 0$.

Proof. (a) The proof is straightforward by the same arguments as in Prop. 3.1(a).

(b) The ideas for the proof are indeed borrowed from [8, Prop. 3.4].
For $i = 1$ and p is even, $\psi_1(a,b) = \frac{1}{2}(\phi_1(a,b))^2$. By the chain rule (see [5, Theorem 2.2.4]), we obtain $\partial\psi_1(a,b) = \partial\phi_1(a,b)^T\phi_1(a,b)$. We will show that $\partial\phi_1(a,b)^T\phi_1(a,b)$ is single-valued for all $(a,b) \in \mathbb{R}^2$ because the zero of ϕ_1 cancels the multi-valued portion of $\partial\phi_1(a,b)^T$. To see this, we discuss several cases as below.
(i) If $a > 0, b > 0$, then $(b_+\partial a_+, a_+\partial b_+) = (b,a)$ which is single-valued. Hence, by (19), it is easy to see that $\partial\phi_1(a,b)^T\phi_1(a,b)$ is single-valued.
(ii) If $a > 0, b < 0$, then $(b_+\partial a_+, a_+\partial b_+) = (0,a)$ which is single-valued. Hence, by (19), $\partial\phi_1(a,b)^T\phi_1(a,b)$ is single-valued.
(iii) If $a > 0, b = 0$, then $(b_+\partial a_+, a_+\partial b_+) = (0, a\cdot[0,1])$ which is multi-valued. However, under this case, we observe that $\phi_1(a,b) = \|(a,b)\|_p - (a+b) - \alpha a_+b_+ = 0$. Hence, $\partial\phi_1(a,b)^T\phi_1(a,b)$ is still single-valued.
(iv) If $a < 0, b > 0$ or $a < 0, b < 0$, or $a < 0, b = 0$, then $(b_+\partial a_+, a_+\partial b_+)$ all equals $(0,0)$ which is single-valued. Hence, by (19), $\partial\phi_1(a,b)^T\phi_1(a,b)$ is single-valued.
(v) If $a = 0, b > 0$, then $(b_+\partial a_+, a_+\partial b_+) = (b\cdot[0,1], 0)$ which is multi-valued. However,

under this case, we observe that $\phi_1(a,b) = \|(a,b)\|_p - (a+b) - \alpha a_+ b_+ = 0$. Hence, $\partial\phi_1(a,b)^T\phi_1(a,b)$ is still single-valued.
(vi) If $a = 0, b < 0$, then then $(b_+\partial a_+, a_+\partial b_+) = (0,0)$ which is single-valued. Hence, by (19), $\partial\phi_1(a,b)^T\phi_1(a,b)$ is single-valued.
(vii) If $a = 0, b = 0$ then $\phi_1(a,b) = 0$. Hence, $\partial\phi_1(a,b)^T\phi_1(a,b)$ is single-valued.
Thus, by applying the Corollary to Theorem 2.2.4 in [5], the above facts yield that ψ_1 is continuously differentiable everywhere. For p is odd, going over the same cases, the proof follows.

For $i = 2$, $\psi_2(a,b) = \frac{1}{2}(\phi_2(a,b))^2$. We discuss the following cases : (i) $(a,b) \neq (0,0)$ and $ab > 0$, (ii) $(a,b) \neq (0,0)$ and $ab = 0$, (iii) $(a,b) \neq (0,0)$ and $ab < 0$, (iv) $(a,b) = (0,0)$. From (20), we know that $\partial\phi_2(a,b)$ becomes multi-valued when $ab = 0$ or $(a,b) = (0,0)$. However, $\phi_2(a,b) = 0$ under these two cases which implies that $\partial\phi_2(a,b)^T\phi_2(a,b)$ is still single-valued. Hence, ψ_2 is continuously differentiable everywhere by the Corollary to Theorem 2.2.4 in [5] again.

For $i = 3, 4$, from (21)-(24), it is trivial that $\partial\phi_3(a,b), \partial\phi_4(a,b)$ are single-valued when $(a,b) \neq (0,0)$. When $(a,b) = (0,0)$, we observe that $\phi_3(a,b) = \phi_4(a,b) = 0$. Hence, $\partial\phi_3(a,b)^T\phi_3(a,b)$ and $\partial\phi_4(a,b)^T\phi_4(a,b)$ are still single-valued, which yield that ψ_3, ψ_4 are continuously differentiable everywhere by the same reason as above.

(c) For $i = 1$, $\psi_1 = \frac{1}{2}(\phi_1)^2$, we employ and go over the cases discussed as in part (b).
(i) If $a > 0, b > 0$, then $(b_+\partial a_+, a_+\partial b_+) = (b, a)$. Hence, from (19), we obtain that

$$\nabla_a\psi_1(a,b) = \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha b\right)\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\right)\phi_1(a,b),$$

for both p are even and odd. Then, $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b)$ equals

$$\Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha b\Big)\Big(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\Big)\phi_1^2(a,b).$$

Since, $\left|\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}}\right| \leq 1$, $\left|\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}}\right| \leq 1$, and $\alpha a > 0, \alpha b > 0$, we know

$$\left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha b\right) < 0 \quad \text{and} \quad \left(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\right) < 0,$$

which implies that $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) \geq 0$.
(ii) If $a > 0, b < 0$, then $(b_+\partial a_+, a_+\partial b_+) = (0, a)$. Hence, from (19), we have

$$\nabla_a\psi_1(a,b) = \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\right)\phi_1(a,b),$$

for p is even; and

$$\nabla_a\psi_1(a,b) = \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{-b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\right)\phi_1(a,b),$$

for p is odd. Again, since $\left|\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}}\right| \leq 1$, $\left|\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}}\right| \leq 1$, and $\alpha a > 0$, it can be easily verified that $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) \geq 0$.
(iii) If $a > 0, b = 0$, then $\phi_1(a,b) = 0$ which says $\nabla_a\psi_1(a,b) = 0 = \nabla_b\psi_1(a,b)$. Hence, $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) = 0$.
(iv) If $a < 0, b > 0$ or $a < 0, b < 0$, or $a < 0, b = 0$, then $(b_+\partial a_+, a_+\partial b_+) = (0,0)$. Hence, from (19), we have

$$\nabla_a\psi_1(a,b) = \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),$$

for p is even; and

$$\nabla_a\psi_1(a,b) = \left(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),$$

for p is odd. Again, by $\left|\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}}\right| \leq 1$, and $\left|\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}}\right| \leq 1$, the desired inequality holds.
(v) If $a = 0, b > 0$, then $\phi_1(a,b) = 0$ which says $\nabla_a\psi_1(a,b) = 0 = \nabla_b\psi_1(a,b)$. Hence, $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) = 0$.
(vi) If $a = 0, b < 0$, then then $(b_+\partial a_+, a_+\partial b_+) = (0,0)$. Hence, from (19), we have

$$\nabla_a\psi_1(a,b) = -\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),$$

for p is even; and

$$\nabla_a\psi_1(a,b) = -\phi_1(a,b),\ \nabla_b\psi_1(a,b) = \left(\frac{-b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\right)\phi_1(a,b),$$

for p is odd. By the same reasons as in previous discussions, we obtain that $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) \geq 0$.
(vii) If $a = 0, b = 0$, then $\phi_1(a,b) = 0$. Hence, $\nabla_a\psi_1(a,b) = 0 = \nabla_b\psi_1(a,b)$ and $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) = 0$.

For $i = 2$, $\psi_2 = \frac{1}{2}(\phi_2)^2$, we discuss discuss four cases as in part (b).
(i) If $(a,b) \neq (0,0)$ and $ab > 0$, from (20), we have

$$\nabla_a\psi_2(a,b) = \left(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha b\right)\phi_2(a,b),\ \nabla_b\psi_2(a,b) = \left(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\right)\phi_2(a,b),$$

for p is even; and

$$\nabla_a\psi_2(a,b) = \left(\frac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha b\right)\phi_2(a,b),\ \nabla_b\psi_2(a,b) = \left(\frac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1 - \alpha a\right)\phi_2(a,b),$$

for p is odd. By the same reasons as in previous discussions, it can be easily verified that $\nabla_a\psi_1(a,b)\cdot\nabla_b\psi_1(a,b) \geq 0$.

(ii) If $(a,b) \neq (0,0)$ and $ab = 0$, then $\phi_2(a,b) = 0$. Hence, $\nabla_a \psi_2(a,b) = 0 = \nabla_b \psi_2(a,b)$ and $\nabla_a \psi_2(a,b) \cdot \nabla_b \psi_2(a,b) = 0$.
(iii) If $(a,b) \neq (0,0)$ and $ab < 0$, the arguments are the same as case (iv) for $i = 1$ except that ϕ_1 is replaced by ϕ_2.
(iv) If $(a,b) = (0,0)$, then $\phi_2(a,b) = 0$. Hence, $\nabla_a \psi_2(a,b) = 0 = \nabla_b \psi_2(a,b)$ and $\nabla_a \psi_2(a,b) \cdot \nabla_b \psi_2(a,b) = 0$.

For $i = 3$, $\psi_3 = \frac{1}{2}(\phi_3)^2$, we have two cases as below.
(i) If $(a,b) \neq (0,0)$, from (21)-(22), we have

$$\begin{aligned} \nabla_a \psi_3(a,b) &= \phi_p(a,b)\Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(a_+)(b_+)^2, \\ \nabla_b \psi_3(a,b) &= \phi_p(a,b)\Big(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(a_+)^2(b_+), \end{aligned}$$

for p is even; and

$$\begin{aligned} \nabla_a \psi_3(a,b) &= \phi_p(a,b)\Big(\frac{sgn(a) \cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(a_+)(b_+)^2, \\ \nabla_b \psi_3(a,b) &= \phi_p(a,b)\Big(\frac{sgn(b) \cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(a_+)^2(b_+), \end{aligned}$$

for p is odd. Thus, $\nabla_a \psi_3(a,b) \cdot \nabla_b \psi_3(a,b)$ equals

$$\begin{aligned} &\phi_p^2(a,b)\Big(\tfrac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\Big(\tfrac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha^2(a_+)^3(b_+)^3 \\ + \quad &\phi_p(a,b)\Big(\tfrac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(a_+)^2(b_+) + \phi_p(a,b)\Big(\tfrac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(a_+)(b_+)^2 \end{aligned}$$

or

$$\begin{aligned} &\phi_p^2(a,b)\Big(\tfrac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\Big(\tfrac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha^2(a_+)^3(b_+)^3 \\ + \quad &\phi_p(a,b)\Big(\tfrac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(a_+)^2(b_+) + \phi_p(a,b)\Big(\tfrac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(a_+)(b_+)^2. \end{aligned}$$

Note that in the above expressions, it is trivial that the first and second terms are nonnegative. We also notice that

$$(a_+)(b_+) = \begin{cases} ab, & \text{if } a > 0, b > 0 \\ 0, & \text{else.} \end{cases}$$

Therefore, we only need to consider the subcase of $a > 0, b > 0$ for the third and fourth terms. In fact, summing up the third and fourth term under this subcase gives

$$\alpha ab \cdot \phi_p(a,b)\left[\Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)a + \Big(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)b\right]$$

$$
\begin{aligned}
&= \alpha ab \cdot \phi_p(a,b)\Big[\frac{a^p+b^p}{\|(a,b)\|_p^{p-1}} - (a+b)\Big] \\
&= \alpha ab \cdot \phi_p(a,b)[\|(a,b)\|_p - (a+b)] \\
&= \alpha ab \cdot \phi_p^2(a,b) \\
&\geq 0.
\end{aligned}
$$

Thus, we proved $\nabla_a \psi_2(a,b) \cdot \nabla_b \psi_2(a,b) \geq 0$.

For $i = 4$, $\psi_4 = \frac{1}{2}(\phi_4)^2$, we also have two cases as below.
(i) If $(a,b) \neq (0,0)$, from (23)-(24), we have

$$
\begin{aligned}
\nabla_a \psi_4(a,b) &= \phi_p(a,b)\Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(ab)_+ \cdot b, \\
\nabla_b \psi_4(a,b) &= \phi_p(a,b)\Big(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(ab)_+ \cdot a,
\end{aligned}
$$

for p is even; and

$$
\begin{aligned}
\nabla_a \psi_4(a,b) &= \phi_p(a,b)\Big(\frac{sgn(a) \cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(ab)_+ \cdot b, \\
\nabla_b \psi_4(a,b) &= \phi_p(a,b)\Big(\frac{sgn(b) \cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha(ab)_+ \cdot a,
\end{aligned}
$$

for p is odd. Thus, $\nabla_a \psi_4(a,b) \cdot \nabla_b \psi_4(a,b)$ equals

$$
\begin{aligned}
&\phi_p^2(a,b)\Big(\tfrac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\Big(\tfrac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha^2 (ab)_+^2 \cdot (ab) \\
+ \quad &\phi_p(a,b)\Big(\tfrac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(ab)_+ \cdot a + \phi_p(a,b)\Big(\tfrac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(ab)_+ \cdot b
\end{aligned}
$$

or

$$
\begin{aligned}
&\phi_p^2(a,b)\Big(\tfrac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\Big(\tfrac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big) + \alpha^2 (ab)_+^2 \cdot (ab) \\
+ \quad &\phi_p(a,b)\Big(\tfrac{sgn(a)\cdot a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(ab)_+ \cdot a + \phi_p(a,b)\Big(\tfrac{sgn(b)\cdot b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)\alpha(ab)_+ \cdot b.
\end{aligned}
$$

The first and second terms are nonnegative by the same reasons in previous discussions. We notice that

$$
(ab)_+ = \begin{cases} ab, & \text{if } ab > 0 \\ 0, & \text{else.} \end{cases}
$$

Thus, we only need to consider the subcase of $ab > 0$ for the third and fourth terms. In fact, summing up the third and fourth term under this subcase gives

$$
\alpha(ab)_+ \cdot \phi_p(a,b)\left[\Big(\frac{a^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)a + \Big(\frac{b^{p-1}}{\|(a,b)\|_p^{p-1}} - 1\Big)b\right]
$$

$$
\begin{aligned}
&= \alpha(ab)_+ \cdot \phi_p(a,b)\Big[\frac{a^p+b^p}{\|(a,b)\|_p^{p-1}} - (a+b)\Big] \\
&= \alpha(ab)_+ \cdot \phi_p(a,b)[\|(a,b)\|_p - (a+b)] \\
&= \alpha(ab)_+ \cdot \phi_p^2(a,b) \\
&\geq 0.
\end{aligned}
$$

The arguments hold as well for p is odd. Hence, we proved $\nabla_a\psi_2(a,b)\cdot\nabla_b\psi_2(a,b) \geq 0$.

(d) Going over exactly the same cases for each i discussed as in part (c) where $\nabla_a\psi(a,b)$ and $\nabla_b\psi(a,b)$ are formed, it is not hard to verify that the desired result is satisfied. We omit the details. □

Based on the properties of ψ stated as in Prop. 3.3 and using the same proof techniques developed in [6, 15, 16], we have the following condition for a stationary point to be a solution of the NCP. We omit the details.

Proposition 3.4 *Assume that $x^* \in \mathbb{R}^n$ is a stationary point of Ψ defined as (15)-(17) (except for Ψ induced from ψ_2) such that the Jacobian $\nabla F(x^*)$ is a P_0-matrix. Then x^* is a solution of the NCP.*

As pointed out in Prop. 3.4, if Ψ is induced from ψ_2 then Prop. 3.4 does not necessary hold for such a Ψ. The reason is that there needs $\nabla_a\psi(a,b)\cdot\nabla_b\psi(a,b) > 0$ when $\psi(a,b) \neq 0$ in the proof. However, this is not always true (we proved that $\nabla_a\psi(a,b)\cdot\nabla_b\psi(a,b) \geq 0$) for ψ_2. A counterexample for $p = 2$ was given in [25, pp. 206-207]. Hence, due to this reason, the merit function induced from ψ_2 may not be recommended even though it is continuously differentiable.

4 Bounded level sets

As mentioned earlier, the merit function Ψ_p defined as in (9) does not guarantee bounded level sets for monotone NCP. In fact, it needs that F is either strongly monotone or uniform P-function to obtain property of bounded level sets, see Prop. 3.5 of [3]. In this section, we establish that if F is either a monotone function with strictly feasible solution or a R_0-function, the merit function Ψ defined as in (17) provides the bounded level sets. This results indicates that Ψ may be a better choice of merit function for the NCP than Ψ_p in certain sense. The motivation is from [1] where a condition, under which a penalized Fischer-Burmeister function is guaranteed to have bounded level sets property, was proposed. We also adopt the condition for our merit functions studied in this paper.

Now, we begin to see under what condition the level sets

$$\mathcal{L}(\gamma) := \{x \in \mathbb{R}^n | \ \Psi(x) \leq \gamma\} \tag{26}$$

are to be bounded for all $\gamma \geq 0$. When $p = 2$ and $\psi = \psi_1$, as shown in [1], it turns out that the following condition on F is sufficient. We employ it for our ψ's and extend the existing results as for $p = 2$ to general $p \geq 2$.

Condition A For any sequence $\{x^k\}$ such that

$$\|x^k\| \to \infty, \quad [-x^k]_+ < \infty, \quad [-F(x^k)]_+ < \infty, \tag{27}$$

it holds

$$\max_i [x_i^k]_+ [F_i(x^k)]_+ \to \infty. \tag{28}$$

Proposition 4.1 *If F satisfies Condition A, then the level sets $\mathcal{L}(\gamma)$ are bounded for all $\gamma \geq 0$.*

Proof. Suppose not, then there exists an unbounded sequence $\{x^k\} \subseteq \mathcal{L}(\gamma)$ for some $\gamma \geq 0$. Since $\Psi(x^k) \leq \gamma$ for all $k \in N$ and $\Psi(x) = \dfrac{1}{2}\displaystyle\sum_{i=1}^{n} \phi^2(x_i, F_i(x))$, there is no index i such that $x_i^k \to -\infty$ or $F_i(x^k) \to -\infty$ by Prop. 3.1(c). Hence, (27) in condition A is held, which says $\max\limits_i (x_i^k)_+ (F_i(x^k))_+ \to \infty$. In other words, there is a j and at least a subsequence $\{x_j^k\}$ where $k \in K \subseteq N$ such that

$$(x_j^k)_+ (F_j(x^k))_+ \to \infty \quad k \in K.$$

However, this implies $\Psi(x^k)$ is unbounded by definition of ϕ as in (11) and Prop. 3.1(c) again which leads to a contradiction to the level sets assumption. $\square$

In fact, condition A is satisfied if F is either a monotone function with a strictly feasible point or a R_0-function ([1, Prop. 3.10]) which indicates that the condition A on F may be the weakest assumption to guarantee bounded level sets for the nonlinear complementarity problem since both F is monotone with a strictly feasible point and a R_0-function are sufficient to condition A. By the way, it is also known that if F is a P_0-function and the NCP has a nonempty and bounded solution set, then there is a strictly feasible solution for the NCP (see [1, Prop. 3.12]).

5 Final Remarks

In this paper, we have studied several new NCP-functions based on the generalized Fischer-Burmeister function and have shown that each of them enjoys all the properties possessed by their counterparts when $p = 2$. The property of error bounds is not included in this paper though we also wished to investigate conditions under which the merit functions Ψ derived from $\psi_i, i = 1, 2, 3, 4$ provide error bounds for the NCP. In fact, it was shown in [1] that if F is a uniform P-function then the aforementioned property holds for $p = 2$ and ψ_1. However, we have not established the parallel results for general $p \geq 2$ and the other ψ_is. The main reason is that we could not yet derive analogous inequalities as in [26, Lem. 3.1] for $\phi_p, p \geq 2$ which plays an important role in proving the error bounds property. We will keep an eye on this topic. On the other hand, according to the the theoretical part built in this paper (not taking it for granted before we prove it even though we think it should be true), the numerical implementation of related algorithms may be interesting for future research.

We want to point something out. During the reviewing, [3] relaxes the condition of p being positive integer greater than one to more general condition of $p > 1$. Moreover, Lemma 3.1(a) is improved in [19] where the condition $p > 1$ is considered.

Acknowledge. The author thanks the referees for their helpful comments and is grateful to Professor P. Tseng for his suggestion on studying these NCP-functions.

References

[1] B. Chen, X. Chen and C. Kanzow (2000), *A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results*, Mathematical Programming, vol. 88, pp. 211-216.

[2] B. Chen and P. T. Harker (1997), *Smooth approximations to nonlinear complementarity problems*, SIAM Journal on Optimization, vol. 2, pp. 403-420.

[3] J.-S. Chen and S. Pan (2006), *A family of NCP-functions and a descent method for the nonlinear complementarity problem*, to appear in Computational Optimization and Applications.

[4] J.-S. Chen (2006), *The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem*, Journal of Global Optimization, vol. 36, pp. 565-580.

[5] F. H. Clarke (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.

[6] T. De Luca, F. Facchinei and C. Kanzow (1996), *A semismooth equation approximation to the solution of the nonlinear complementarity problems*, Mathematical Programming, vol. 75, pp. 407-439.

[7] F. Facchinei and J.-S. Pang (2003), *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York.

[8] F. Facchinei and J. Soares (1997), *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM Journal on Optimization, vol. 7, pp. 225-247.

[9] M. C. Ferris and J.-S. Pang (1997), *Engineering and economic applications of complementarity problems*, SIAM Review, vol. 39, pp. 669-713.

[10] A. Fischer (1992), *A special Newton-type optimization methods*, Optimization, vol. 24, pp. 269-284.

[11] A. Fischer (1997), *Solution of the monotone complementarity problem with locally Lipschitzian functions*, Mathematical Programming, vol. 76, pp. 513-532.

[12] C. Geiger and C. Kanzow (1996), *On the resolution of monotone complementarity problems*, Computational Optimization and Applications, vol. 5, pp. 155-173.

[13] P. T. Harker and J.-S. Pang (1990), *Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications*, Mathematical Programming, vol. 48, pp. 161-220.

[14] C. Kanzow (1996), *Nonlinear complementarity as unconstrained optimization*, Journal of Optimization Theory and Applications, vol. 88, pp. 139-155.

[15] C. Kanzow and H. Kleinmichel (1998), *A new class of semismooth Newton-type methods for the nonlinear complementarity problems*, Computational Optimizations and Applications, vol. 11, pp. 227-251.

[16] C. Kanzow, N. Yamashita and M. Fukushima (1997), *New NCP-functions and their properties*, Journal of Optimization Theory and Applications, vol. 94, pp. 115-135.

[17] O. L. Mangasarian (1976), *Equivalence of the complementarity problem to a system of nonlinear equations*, SIAM Journal on Applied Mathematics, vol. 31, pp. 89-92.

[18] R. Mifflin (1977), *Semismooth and semiconvex functions in constrained optimization*, SIAM Journal on Control Optimization, vol. 15, pp. 959-972.

[19] S. Pan and J.-S. Chen (2007), *A family of penalized NCP functions based on the generalized Fischer-Burmeister functions: theoretical investigation and numerical results*, submitted.

[20] J.-S. Pang, *Complementarity problems* (1994), Handbook of Global Optimization, edited by R. Horst and P. Pardalos, Kluwer Academic Publishers, Boston, Massachusetts, pp. 271-338.

[21] L.-Q. Qi (1993), *Convergence analysis of some algorithms for solving nonsmooth equations*, Mathematics of Operations Research, vol. 18, pp. 227-244.

[22] L.-Q. Qi (1994), *Superlinearly convergent approximate Newton methods for LC^1 optimization problems*, Mathematical Programming, vol. 64, pp. 277-294.

[23] L.-Q. Qi and J. Sun (1993), *A nonsmooth version of Newton's method*, Mathematical Programming, vol. 58, pp. 353-367.

[24] R.T. Rockafellar and J.-B. Wets (1998), *Variational Analysis*, Springer-Verlag, Berlin.

[25] D. Sun and L.-Q. Qi (1999), *On NCP-functions*, Computational Optimization and Applications, vol. 13, pp. 201-220.

[26] P. Tseng (1996), *Global behaviour of a class of merit functions for the nonlinear complementarity problem*, Journal of Optimization Theory and Applications, vol. 89, pp. 17-37.

[27] K. Yamada, N. Yamashita, and M. Fukushima (2000), *A new derivative-free descent method for the nonlinear complementarity problems*, in Nonlinear Optimization and Related Topics edited by G.D. Pillo and F. Giannessi, Kluwer Academic Publishers, Netherlands, pp. 463–489.