

# The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem

Jein-Shan Chen <sup>1</sup>

Department of Mathematics  
National Taiwan Normal University  
Taipei 11677, Taiwan

February 1, 2005

(revised March 11, 2006)

**Abstract** This paper is a follow-up of the work [1] where an NCP-function and a descent method were proposed for the nonlinear complementarity problem. An unconstrained reformulation was formulated due to a merit function based on the proposed NCP-function. We continue to explore properties of the merit function in this paper. In particular, we show that the gradient of the merit function is globally Lipschitz continuous which is important from computational aspect. Moreover, we show that the merit function is  $SC^1$  function which means it is continuously differentiable and its gradient is semismooth. On the other hand, we provide an alternative proof, which uses the new properties of the merit function, for the convergence result of the descent method considered in [1].

**Key words.** Complementarity,  $SC^1$  function, merit function, semismooth function, descent method.

## 1 Introduction

In the past decades, the well-known nonlinear complementarity problem (NCP) has attracted much attention due to its various applications in operations research, economics, and engineering [7, 12, 18]. The NCP is to find a point  $x \in \mathbb{R}^n$  such that

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

where  $\langle \cdot, \cdot \rangle$  is the Euclidean inner product and  $F = (F_1, F_2, \dots, F_n)^T$  maps from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ . We assume that  $F$  is continuously differentiable throughout this paper.

---

<sup>1</sup>E-mail: jschen@math.ntnu.edu.tw, TEL: 886-2-29320206, FAX: 886-2-29332342.

There have been many methods proposed for solving the NCP [10, 12, 18]. Among which, one of the most popular approaches that has been studied intensively recently is to reformulate the NCP as an unconstrained minimization problem [6, 8, 11, 14, 15, 30]. Such a function that can constitute an equivalent unconstrained minimization problem for the NCP is called a merit function. In other words, a merit function is a function whose global minima are coincident with the solutions of the original NCP. For constructing a merit function, the class of functions, so-called NCP-functions and defined as below, serves an important role.

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

$$\phi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0. \quad (2)$$

A popular NCP-function intensively studied recently is the well-known Fischer-Burmeister NCP-function [8, 9, 26] defined as

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

Let  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be

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

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

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

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

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

In the paper [1], an NCP-function which is an extension of the Fischer-Burmeister function (3) was studied. More specifically, they define  $\phi_p : \mathbb{R}^2 \rightarrow \mathbb{R}$  by

$$\phi_p(a, b) := \|(a, b)\|_p - (a + b), \quad (7)$$

where  $\|(a, b)\|_p$  denotes the  $p$ -norm of  $(a, b)$ , i.e.,  $\|(a, b)\|_p = \sqrt[p]{|a|^p + |b|^p}$ . In other words, in the function  $\phi_p$ , the 2-norm of  $(a, b)$  in the Fischer-Burmeister function (3) is replaced by more generally a  $p$ -norm of  $(a, b)$  with  $p \geq 2$ . This function  $\phi_p$  is still an NCP-function

as was noted in Tseng's paper [28]. Nonetheless, there was no further study on this NCP-function even for  $p = 3$  until the recent paper [1] by the author. Following the function  $\phi_p$ , we can further define  $\psi_p : \mathbb{R}^2 \rightarrow \mathbb{R}_+$  by

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

The function  $\psi_p$  is a nonnegative NCP-function and smooth on  $\mathbb{R}^2$  with some favorable properties, see [1]. In this paper, we continue to explore properties of  $\psi_p$  as will be seen in Sec. 3. Analogous to  $\Phi$ , the function  $\Phi_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$  given by

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

yields a merit function  $\Psi_p : \mathbb{R}^n \rightarrow \mathbb{R}_+$  for the NCP where

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

As shown in [1],  $\Psi_p$  is a continuously differentiable merit function for the NCP. Therefore, classical iterative methods such as Newton method can be applied to the unconstrained smooth minimization of the NCP, i.e.,

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

On the other hand, derivative-free methods have also attracted much attention which do not require computation of derivatives of  $F$  [11, 14, 29]. Derivative-free methods, taking advantages of particular properties of a merit function, are suitable for problems where the derivatives of  $F$  are not available or expensive. In this paper, we also study a derivative-free descent algorithm for solving the NCP based on the merit function  $\Psi_p$  in Sec. 4. Indeed, the descent method was considered in [1], we apply the new properties of  $\psi_p$  explored in this paper to provide an alternative proof for the convergence result.

Throughout this paper,  $\mathbb{R}^n$  denotes the space of  $n$ -dimensional real column vectors and  $T$  denotes transpose. For any differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\nabla f(x)$  denotes the gradient of  $f$  at  $x$ . For any differentiable mapping  $F = (F_1, \dots, F_m)^T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $\nabla F(x) = [\nabla F_1(x) \ \cdots \ \nabla F_m(x)]$  denotes the transpose Jacobian of  $F$  at  $x$ . We write  $z = o(\alpha)$  with  $\alpha \in \mathbb{R}$  and  $z \in \mathbb{R}^n$  to mean  $\|z\|/|\alpha|$  tends to zero as  $\alpha \rightarrow 0$ . Also, we denote by  $\|x\|_p$  the  $p$ -norm of  $x$  and by  $\|x\|$  the Euclidean norm of  $x$ . In the whole paper, we assume  $p$  is an integer greater than or equal to 2.

## 2 Preliminaries

In this section, we recall some background concepts and review some known materials which are crucial to the subsequent analysis. We begin with the monotonicity of a mapping. Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , then  $F$  is monotone if  $\langle x - y, F(x) - F(y) \rangle \geq 0$ , for all  $x, y \in \mathbb{R}^n$ ;  $F$  is strictly monotone if  $\langle x - y, F(x) - F(y) \rangle > 0$ , for all  $x, y \in \mathbb{R}^n$  and  $x \neq y$ ; and  $F$  is strongly monotone with modulus  $\mu > 0$  if  $\langle x - y, F(x) - F(y) \rangle \geq \mu \|x - y\|^2$ , for all  $x, y \in \mathbb{R}^n$ . Next, we recall the so-called semismooth functions. First, we say that  $F$  is strictly continuous (also called ‘locally Lipschitz continuous’) at  $x \in \mathbb{R}^n$  [25, 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  $\delta$  can be taken to be  $\infty$ , then  $F$  is Lipschitz continuous with Lipschitz constant  $\kappa$ . Define the function  $\text{lip}F : \mathbb{R}^n \rightarrow [0, \infty]$  by

$$\text{lip}F(x) := \limsup_{\substack{y, z \rightarrow x \\ y \neq z}} \frac{\|F(y) - F(z)\|}{\|y - z\|}.$$

Then  $F$  is strictly continuous at  $x$  if and only if  $\text{lip}F(x)$  is finite. We say  $F$  is directionally differentiable at  $x \in \mathbb{R}^n$  if

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

and  $F$  is directionally differentiable if  $F$  is directionally differentiable at every  $x \in \mathbb{R}^n$ .  $F$  is differentiable (in the Fréchet sense) at  $x \in \mathbb{R}^n$  if there exists a linear mapping  $\nabla F(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$  such that

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

We say that  $F$  is continuously differentiable if  $F$  is differentiable at every  $x \in \mathbb{R}^n$  and  $\nabla F$  is continuous.

If  $F$  is strictly continuous, then  $F$  is almost everywhere differentiable by Rademacher’s Theorem—see [4] and [25, Sec. 9J]. In this case, the generalized Jacobian  $\partial F(x)$  of  $F$  at  $x$  (in the Clarke sense) can be defined as the convex hull of the generalized Jacobian  $\partial_B F(x)$ , where

$$\partial_B F(x) := \left\{ \lim_{x^j \rightarrow x} \nabla F(x^j) \mid F \text{ is differentiable at } x^j \in \mathbb{R}^n \right\}.$$

The notation  $\partial_B$  is adopted from [20]. In [25, Chap. 9], the case of  $n = 1$  is considered and the notations “ $\bar{\nabla}$ ” and “ $\bar{\partial}$ ” are used instead of, respectively, “ $\partial_B$ ” and “ $\partial$ ”.

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

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

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

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

The following lemma, proven by Sun and Sun [27, Thm. 3.6] using the definition of generalized Jacobian, (Sun and Sun did not consider the case of  $o(\|h\|)$  but their argument readily applies to this case.) enables one to study the semismooth property of  $F$  by examining only those points  $x \in \mathbb{R}^n$  where  $F$  is differentiable and thus work only with the Jacobian of  $F$ , rather than the generalized Jacobian.

**Lemma 2.1** *Suppose  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is strictly continuous and directionally differentiable in a neighborhood of  $x \in \mathbb{R}^n$ . Then, for any  $0 < \rho < \infty$ , the following two statements (where  $O(\cdot)$  depends on  $F$  and  $x$  only) are equivalent:*

(a) *For any  $h \in \mathbb{R}^n$  and any  $V \in \partial F(x + h)$ ,*

$$F(x + h) - F(x) - Vh = o(\|h\|) \quad (\text{respectively, } O(\|h\|^{1+\rho})).$$

(b) *For any  $h \in \mathbb{R}^n$  such that  $F$  is differentiable at  $x + h$ ,*

$$F(x + h) - F(x) - \nabla F(x + h)h = o(\|h\|) \quad (\text{respectively, } O(\|h\|^{1+\rho})).$$

We say  $F$  is semismooth (respectively,  $\rho$ -order semismooth) if  $F$  is semismooth (respectively,  $\rho$ -order semismooth) at every  $x \in \mathbb{R}^n$ . We say  $F$  is strongly semismooth if it is 1-order semismooth. Convex functions and piecewise continuously differentiable functions are examples of semismooth functions. The composition of two (respectively,  $\rho$ -order) semismooth functions is also a (respectively,  $\rho$ -order) semismooth function. The property of semismoothness plays an important role in nonsmooth Newton methods [20, 22] as well as in some smoothing methods. For extensive discussions of semismooth functions, see [9, 16, 22].

Now, we review some useful properties about  $\phi_p, \psi_p$  defined as in (7) and (8), respectively which will be used for the analysis in the subsequent sections. We notice that the function  $\phi_p$  reduces to the Fischer-Burmeister function given as in (3) when  $p = 2$ . Thus, most properties are extensions of properties of Fischer-Burmeister function. For detailed proofs of them, please refer to [1].

**Lemma 2.2** *([1, Prop. 3.1]) Let  $\phi_p : \mathbb{R}^2 \rightarrow \mathbb{R}$  be defined as (7) where  $p \geq 2$ . Then*

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

- (b)  $\phi_p$  is sub-additive, i.e.,  $\phi_p(w + w') \leq \phi_p(w) + \phi_p(w')$  for all  $w, w' \in \mathbb{R}^2$ .
- (c)  $\phi_p$  is positively homogeneous, i.e.,  $\phi_p(\alpha w) = \alpha \phi_p(w)$  for all  $w \in \mathbb{R}^2$  and  $\alpha \geq 0$ .
- (d)  $\phi_p$  is convex, i.e.,  $\phi_p(\alpha w + (1 - \alpha)w') \leq \alpha \phi_p(w) + (1 - \alpha)\phi_p(w')$  for all  $w, w' \in \mathbb{R}^2$  and  $\alpha \in (0, 1)$ .
- (e)  $\phi_p$  is Lipschitz continuous with  $L_1 = 1 + \sqrt{2}$ , i.e.,  $|\phi_p(w) - \phi_p(w')| \leq L_1 \|w - w'\|$ ; or with  $L_2 = 1 + 2^{(1-1/p)}$ , i.e.,  $|\phi_p(w) - \phi_p(w')| \leq L_2 \|w - w'\|_p$  for all  $w, w' \in \mathbb{R}^2$ .

Lemma 2.2(b) and (c) imply that  $\phi_p$  is sublinear, i.e., it satisfies

$$\phi_p(\alpha w + \beta w') \leq \alpha \phi_p(w) + \beta \phi_p(w')$$

for all  $w, w' \in \mathbb{R}^2$  and  $\alpha, \beta \geq 0$ . This can be seen by the fact [3, Prop. 3.11] that a function from  $\mathbb{R}^n$  to  $\mathbb{R}$  is sublinear if and only if it is positively homogeneous and sub-additive. Note that the sublinear condition is stronger than convexity. In fact, under Lemma 2.2(c), Lemma 2.2(b) is equivalent to Lemma 2.2(d). This is from [24, Thm. 4.7] that a positively homogeneous function is convex if and only if it is sub-additive.

**Lemma 2.3** ([1, Prop. 3.2]) *Let  $\phi_p, \psi_p$  be defined as (7) and (8), respectively, where  $p \geq 2$ . Then*

- (a)  $\psi_p$  is an NCP-function, i.e., it satisfies (2).
- (b)  $\psi_p(a, b) \geq 0$  for all  $(a, b) \in \mathbb{R}^2$ .
- (c)  $\psi_p$  is continuously differentiable everywhere. Moreover,  $\nabla_a \psi_p(0, 0) = \nabla_b \psi_p(0, 0) = 0$  and

$$\begin{aligned} \nabla_a \psi_p(a, b) &= \left( \frac{a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \phi_p(a, b) \\ \nabla_b \psi_p(a, b) &= \left( \frac{b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \phi_p(a, b), \end{aligned} \quad (12)$$

for  $(a, b) \neq (0, 0)$  with  $p$  is even, whereas

$$\begin{aligned} \nabla_a \psi_p(a, b) &= \left( \frac{\text{sgn}(a) \cdot a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \phi_p(a, b) \\ \nabla_b \psi_p(a, b) &= \left( \frac{\text{sgn}(b) \cdot b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \phi_p(a, b), \end{aligned} \quad (13)$$

for  $(a, b) \neq (0, 0)$  with  $p$  is odd.

- (d)  $\nabla_a \psi_p(a, b) \cdot \nabla_b \psi_p(a, b) \geq 0$  for all  $(a, b) \in \mathbb{R}^2$ . The equality holds if and only if  $\phi_p(a, b) = 0$ .

$$(e) \quad \nabla_a \psi_p(a, b) = 0 \iff \nabla_b \psi_p(a, b) = 0 \iff \phi_p(a, b) = 0.$$

**Lemma 2.4** ([1, Prop. 3.5]) *Let  $\Psi_p : \mathbb{R}^n \rightarrow \mathbb{R}$  be defined as (10) where  $p \geq 2$ . Assume  $F$  is either strongly monotone or uniform  $P$ -function, then the level sets  $\mathcal{L}(\Psi_p, \gamma)$  are bounded for all  $\gamma \in \mathbb{R}$ .*

In addition to the above properties of  $\phi_p$  and  $\psi_p$ , we still need the following two lemmas for the analysis in the subsequent sections.

**Lemma 2.5** ([13, (1.3)]) *Let  $x \in \mathbb{R}^n$  and  $1 < p_1 < p_2$ . Then*

$$\|x\|_{p_2} \leq \|x\|_{p_1} \leq n^{(1/p_1 - 1/p_2)} \|x\|_{p_2}.$$

**Lemma 2.6** *If  $F : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  has a second derivative at each point of a convex set  $D_0 \subseteq D$ , then*

$$\|\nabla F(y) - \nabla F(x)\| \leq \sup_{0 \leq t \leq 1} \|\nabla^2 F(x + t(y - x))\| \cdot \|y - x\|.$$

**Proof.** This is Theorem 3.3.5 of [17] (page 78).  $\square$

### 3 The semismooth-related properties of the NCP and merit functions

In this section, we study some semismooth-related properties of  $\phi_p$  including semismooth and almost smooth properties as well as  $SC^1$  and  $LC^1$  properties of  $\psi_p$ . The semismooth property is very important from the computational point of view. In particular, it plays a fundamental role in the superlinear convergence analysis of generalized Newton methods, see [20, 22, 31]. The classes of  $SC^1$  and  $LC^1$  functions have been a subject of interest in relation to the development of minimization algorithm. We will introduce their definitions later. We begin this section by showing that the functions  $\phi_p$  and  $\Phi_p$  are semismooth (in fact, they are strongly semismooth as shown in Corollary 3.1). Its proof is easy and routine.

**Proposition 3.1** *The function  $\Phi_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$  defined as (9) is semismooth.*

**Proof.** We notice that  $\phi_p$  is convex by Lemma 2.2(d), and hence is a semismooth function. We also observe that each component of  $\Phi_p(x)$  is the composite of the convex function  $\phi_p : \mathbb{R}^2 \rightarrow \mathbb{R}$  and the differentiable function  $(x_i, F_i(x))^T : \mathbb{R}^n \rightarrow \mathbb{R}^2$ . Since convex and differentiable functions are semismooth and the composition of semismooth functions is semismooth, it yields that  $\Phi_p$  is semismooth.  $\square$

An important concept in relation to semismooth function is the  $SC^1$  function, so we next introduce its definition as below.

**Definition 3.1** *A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be an  $SC^1$  function if  $f$  is continuously differentiable and its gradient is semismooth.*

We can view  $SC^1$  functions are functions lying between  $C^1$  and  $C^2$  functions. By defining  $SC^1$  functions, many results regarding the minimization of  $C^2$  functions can be extended to the minimization of  $SC^1$  functions, see [19] and references therein. For applications and more details of  $SC^1$  functions, please refer to the excellent book [5]. Prop. 3.2 shows that  $\psi_p$  is an  $SC^1$  function; hence, if every  $F_i$  is  $SC^1$  function then so is  $\Psi_p$ . Before presenting its proof, we need a very important and crucial technical lemma, which states  $\nabla\psi_p$  is globally Lipschitz continuous. The lemma will not only be used in the proof of Prop. 3.2 but also for the analysis of convergence result of the descent algorithm in Sec. 4.

**Lemma 3.1** *The gradient of the function  $\psi_p$  defined as (8) is Lipschitz continuous, that is, there exists  $L > 0$  such that*

$$\|\nabla\psi_p(a, b) - \nabla\psi_p(c, d)\| \leq L\|(a, b) - (c, d)\|, \quad (14)$$

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

**Proof.** Following the gradient of  $\psi_p$  given as in (12) and (13) and then applying the chain rule and quotient rule (the computation is routine though tedious, so we omit the details), we have the following two cases.

If  $p$  is even and  $(a, b) \neq (0, 0)$ , then

$$\begin{aligned} \nabla_{aa}^2\psi_p(a, b) &= \left( \frac{a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right)^2 + \frac{(p-1)a^{p-2}b^p}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p - (a+b) \right), \\ \nabla_{ab}^2\psi_p(a, b) &= \nabla_{ba}^2\psi_p(a, b) = \left( \frac{a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \left( \frac{b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right), \\ &\quad - \frac{(p-1)a^{p-1}b^{p-1}}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p - (a+b) \right), \\ \nabla_{bb}^2\psi_p(a, b) &= \left( \frac{b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right)^2 + \frac{(p-1)a^p b^{p-2}}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p - (a+b) \right). \end{aligned}$$

It is clear that  $\frac{|a|^{p-1}}{\|(a, b)\|_p^{p-1}} \leq 1$  and it also follows

$$|a|^{p-2} \cdot |b|^p \leq \left( \max\{|a|, |b|\} \right)^{2p-2} \leq \left( \sqrt[p]{|a|^p + |b|^p} \right)^{2p-2} \leq \|(a, b)\|_p^{2p-2},$$

that

$$\frac{|a|^{p-2}|b|^p}{\|(a, b)\|_p^{2p-2}} \leq 1. \quad \text{Similarly,} \quad \frac{|a|^p|b|^{p-2}}{\|(a, b)\|_p^{2p-2}} \leq 1. \quad (15)$$

On the other hand, by Lemma 2.5, we have

$$|a| + |b| \leq \sqrt{2}\sqrt{a^2 + b^2} = \sqrt{2}\|(a, b)\|_2 \leq \sqrt{2} \cdot 2^{(1/2-1/p)}\|(a, b)\|_p = 2^{(1-1/p)}\|(a, b)\|_p.$$

Applying all the above, we can give an upper bound for  $\nabla_{aa}^2 \psi_p(a, b)$  as below.

$$\begin{aligned} & \left| \nabla_{aa}^2 \psi_p(a, b) \right| \\ & \leq \left( \frac{a^{p-1}}{\|(a, b)\|_p^{p-1}} + 1 \right)^2 + \frac{(p-1)|a|^{p-2}|b|^p}{\|(a, b)\|_p^{2p-2}} + \frac{(p-1)|a|^{p-2}|b|^p \cdot (|a| + |b|)}{\|(a, b)\|_p^{2p-1}} \\ & \leq 4 + (p-1) + \frac{(p-1)|a|^{p-2}|b|^p \cdot 2^{(1-1/p)}\|(a, b)\|_p}{\|(a, b)\|_p^{2p-1}} \\ & \leq 4 + (p-1) + (p-1)2^{(1-1/p)} \\ & = 4 + (p-1) \left[ 1 + 2^{(1-1/p)} \right], \end{aligned}$$

where the last inequality holds due to (15). By the same arguments, we also have

$$\left| \nabla_{bb}^2 \psi_p(a, b) \right| \leq 4 + (p-1) \left[ 1 + 2^{(1-1/p)} \right].$$

Now, we estimate the upper bound for  $\nabla_{ab}^2 \psi_p(a, b) = \nabla_{ba}^2 \psi_p(a, b)$  as below.

$$\begin{aligned} & \left| \nabla_{ab}^2 \psi_p(a, b) \right| = \left| \nabla_{ba}^2 \psi_p(a, b) \right| \\ & \leq \left| \frac{a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right| \cdot \left| \frac{b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right| \\ & \quad + \frac{(p-1)|a|^{p-1}|b|^{p-1}}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p + (|a| + |b|) \right) \\ & \leq \left( \frac{|a|^{p-1}}{\|(a, b)\|_p^{p-1}} + 1 \right) \left( \frac{|b|^{p-1}}{\|(a, b)\|_p^{p-1}} + 1 \right) \\ & \quad + \frac{(p-1)|a|^{p-1}|b|^{p-1}}{\|(a, b)\|_p^{2p-2}} + \frac{(p-1)|a|^{p-1}|b|^{p-1} \cdot (|a| + |b|)}{\|(a, b)\|_p^{2p-1}} \\ & \leq 4 + (p-1) + \frac{(p-1)|a|^{p-1}|b|^{p-1} \cdot 2^{(1-1/p)}\|(a, b)\|_p}{\|(a, b)\|_p^{2p-1}} \\ & \leq 4 + (p-1) + (p-1)2^{(1-1/p)} \\ & = 4 + (p-1) \left[ 1 + 2^{(1-1/p)} \right], \end{aligned}$$

where the third and fourth inequalities are true by the similar result as (15), that is,  $\frac{|a|^{p-1}|b|^{p-1}}{\|(a, b)\|_p^{2p-2}} \leq 1$ .

If  $p$  is odd and  $(a, b) \neq (0, 0)$ , then we obtain

$$\begin{aligned}\nabla_{aa}^2 \psi_p(a, b) &= \left( \frac{\operatorname{sgn}(a) \cdot a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right)^2 + \frac{\operatorname{sgn}(a)\operatorname{sgn}(b) \cdot (p-1)a^{p-2}b^p}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p - (a+b) \right), \\ \nabla_{ab}^2 \psi_p(a, b) &= \nabla_{ba}^2 \psi_p(a, b) = \left( \frac{\operatorname{sgn}(a) \cdot a^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right) \left( \frac{\operatorname{sgn}(b) \cdot b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right), \\ &\quad - \frac{\operatorname{sgn}(a)\operatorname{sgn}(b) \cdot (p-1)a^{p-1}b^{p-1}}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p - (a+b) \right), \\ \nabla_{bb}^2 \psi_p(a, b) &= \left( \frac{\operatorname{sgn}(b) \cdot b^{p-1}}{\|(a, b)\|_p^{p-1}} - 1 \right)^2 + \frac{\operatorname{sgn}(a)\operatorname{sgn}(b) \cdot (p-1)a^p b^{p-2}}{\|(a, b)\|_p^{2p-1}} \left( \|(a, b)\|_p - (a+b) \right).\end{aligned}$$

In fact, the upper bounds for  $\nabla_{aa}^2 \psi_p(a, b)$ ,  $\nabla_{ab}^2 \psi_p(a, b)$ ,  $\nabla_{bb}^2 \psi_p(a, b)$  remain the same by following exactly the same steps as in the case where  $p$  is even. Thus, there exist a constant  $L > 0$  independent of  $(a, b)$  such that

$$\|\nabla^2 \psi_p(a, b)\| \leq L, \quad \forall (a, b) \neq (0, 0) \in \mathbb{R}^2.$$

Then, by Lemma 2.6, we have

$$\|\nabla \psi_p(a, b) - \nabla \psi_p(c, d)\| \leq L\|(a, b) - (c, d)\|, \quad (16)$$

for all  $(a, b), (c, d) \in \mathbb{R}^2$  with  $(0, 0) \notin [(a, b), (c, d)]$ . Moreover, (16) also holds in case  $(a, b) = (c, d) = (0, 0)$  since  $\nabla_a \psi_p(a, b) = \nabla_b \psi_p(a, b) = 0$ . Therefore, we can assume  $(a, b) \neq (0, 0)$ . From Lemma 2.3(c),  $\psi_p$  is continuously differentiable for all  $(a, b) \in \mathbb{R}^2$  with  $\nabla \psi_p(0, 0) = (0, 0)$ ; then using a continuity argument, we obtain (16) remains true for all  $(c, d) \in \mathbb{R}^2$ . Thus, (16) holds for all  $(a, b), (c, d) \in \mathbb{R}^2$  which says  $\psi_p$  is globally Lipschitz continuous.  $\square$

**Proposition 3.2** *The function  $\psi_p$  defined as in (8) is an  $SC^1$  function. Hence, if every  $F_i$  is an  $SC^1$  function, then the function  $\Psi_p$  given as (10) is also an  $SC^1$  function.*

**Proof.** It is known by Lemma 2.3(c) that  $\psi_p$  is continuously differentiable, it remains to show that the gradient of  $\psi_p$  is semismooth. From Lemma 3.1,  $\nabla \psi_p$  is Lipschitz continuous; hence is strictly continuous (locally Lipschitz continuous). Therefore, to check semismoothness of  $\nabla \psi_p$ , we only need to show that  $\nabla \psi_p$  satisfies Lemma 2.1(b). More specifically, we only need to check semismoothness at  $(0, 0)$  because at other points  $\nabla \psi_p$  is continuously differentiable (see the proof of Lemma 3.1), hence is semismooth. For this purpose, we will

have to verify that the equation in Lemma 2.1(b) is satisfied, i.e., for any  $(h_1, h_2) \in \mathbb{R}^2$  such that  $\nabla\psi_p$  is differentiable at  $(h_1, h_2)$ , we have

$$\nabla\psi_p(h_1, h_2) - \nabla\psi_p(0, 0) - \nabla^2\psi_p(h_1, h_2) \cdot h = o(\|(h_1, h_2)\|). \quad (17)$$

To prove (17), we have two cases where  $p$  is even and  $p$  is odd.

For  $p$  is even, we denote  $(\Xi_1, \Xi_2)$  the left-hand side of (17). Then, we have

$$\begin{aligned} \begin{bmatrix} \Xi_1 \\ \Xi_2 \end{bmatrix} &:= \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \cdot \phi_p(h_1, h_2) - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &- \begin{bmatrix} k_1^2 + \left(\frac{(p-1)h_1^{p-2}h_2^p}{\|(h_1, h_2)\|_p^{2p-1}}\right)\phi_p(h_1, h_2) & k_1 \cdot k_2 - k_3\phi_p(h_1, h_2) \\ k_1 \cdot k_2 - k_3\phi_p(h_1, h_2) & k_2^2 + \left(\frac{(p-1)h_1^p h_2^{p-2}}{\|(h_1, h_2)\|_p^{2p-1}}\right)\phi_p(h_1, h_2) \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \end{aligned} \quad (18)$$

where

$$\begin{aligned} k_1 &= \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right), \\ k_2 &= \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right), \\ k_3 &= \frac{(p-1)h_1^{p-1}h_2^{p-1}}{\|(h_1, h_2)\|_p^{2p-1}}. \end{aligned} \quad (19)$$

By plugging (19) into (18) and writing out  $\Xi_1$  and  $\Xi_2$ , we obtain that  $\Xi_1 = 0$  and  $\Xi_2 = 0$ . To see this, we compute  $\Xi_1$  as below:

$$\begin{aligned} \Xi_1 &= \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \phi_p(h_1, h_2) - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right)^2 h_1 \\ &- \frac{(p-1)h_1^{p-1}h_2^p}{\|(h_1, h_2)\|_p^{2p-1}} \cdot \phi_p(h_1, h_2) - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_2 \\ &+ \frac{(p-1)h_1^{p-1}h_2^p}{\|(h_1, h_2)\|_p^{2p-1}} \cdot \phi_p(h_1, h_2) \\ &= \phi_p(h_1, h_2) \left[ \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) - \frac{(p-1)h_1^{p-1}h_2^p}{\|(h_1, h_2)\|_p^{2p-1}} + \frac{(p-1)h_1^{p-1}h_2^p}{\|(h_1, h_2)\|_p^{2p-1}} \right] \\ &- \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right)^2 h_1 - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_2 \\ &= \phi_p(h_1, h_2) \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right)^2 h_1 \\ &- \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_2 \\ &= \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left[ \phi_p(h_1, h_2) - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_1 - \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_2 \right] \end{aligned}$$

$$\begin{aligned}
&= \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left[ \|(h_1, h_2)\|_p - \frac{h_1^p + h_2^p}{\|(h_1, h_2)\|_p^{p-1}} \right] \\
&= \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \cdot 0 \\
&= 0,
\end{aligned}$$

where the second-to-last equality is true since  $h_1^p + h_2^p = \|(h_1, h_2)\|_p^p$  when  $p$  is even. Similarly,

$$\begin{aligned}
\Xi_2 &= \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \phi_p(h_1, h_2) - \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right)^2 h_2 \\
&\quad - \frac{(p-1)h_1^p h_2^{p-1}}{\|(h_1, h_2)\|_p^{2p-1}} \cdot \phi_p(h_1, h_2) - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_1 \\
&\quad + \frac{(p-1)h_1^p h_2^{p-1}}{\|(h_1, h_2)\|_p^{2p-1}} \cdot \phi_p(h_1, h_2) \\
&= \phi_p(h_1, h_2) \left[ \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) - \frac{(p-1)h_1^p h_2^{p-1}}{\|(h_1, h_2)\|_p^{2p-1}} + \frac{(p-1)h_1^p h_2^{p-1}}{\|(h_1, h_2)\|_p^{2p-1}} \right] \\
&\quad - \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right)^2 h_2 - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_1 \\
&= \phi_p(h_1, h_2) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) - \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right)^2 h_2 \\
&\quad - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_1 \\
&= \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left[ \phi_p(h_1, h_2) - \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_1 - \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) h_2 \right] \\
&= \left( \frac{h_2^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \left[ \|(h_1, h_2)\|_p - \frac{h_1^p + h_2^p}{\|(h_1, h_2)\|_p^{p-1}} \right] \\
&= \left( \frac{h_1^{p-1}}{\|(h_1, h_2)\|_p^{p-1}} - 1 \right) \cdot 0 \\
&= 0,
\end{aligned}$$

where the second-to-last equality is true since  $h_1^p + h_2^p = \|(h_1, h_2)\|_p^p$  when  $p$  is even. From the above two expressions of  $\Xi_1$  and  $\Xi_2$ , it implies that (17) is satisfied. Thus,  $\nabla\psi_p$  is semismooth at  $(0, 0)$  for the case where  $p$  is even.

For  $p$  is odd, following the same arguments leads to the same verifications. Therefore, we complete proving that  $\nabla\psi_p$  is semismooth, and hence  $\psi_p$  is  $SC^1$  function. The second statement follows immediately from this result.  $\square$

We want to point out one thing that, for  $p = 2$ ,  $\psi_p$  was already proved an  $SC^1$  function in [5, 6] (Indeed, it was first formally shown in [6]). Prop. 3.2 is a general extension for any

$p \geq 2$  and its proof is much more complicated than the case of  $p = 2$ . In addition to  $SC^1$  functions, we also introduce  $LC^1$  functions here.

**Definition 3.2** *A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is called an  $LC^1$  function if  $f$  is continuously differentiable and its gradient is locally Lipschitz continuous.*

The class of  $LC^1$  minimization problems was studied in [21], where the local, superlinear convergence of an approximate Newton method was established under a semismoothness assumption on the gradient function at a solution point. It is obvious that any  $SC^1$  function is an  $LC^1$  function. With the results of Lemma 3.1 and Prop. 3.2, we therefore has the following corollaries.

**Corollary 3.1** *If every  $F_i$  is an  $LC^1$  function, then the function  $\Phi_p$  given as (9) is strongly semsmooth.*

**Proof.** We know that  $\phi_p$  is semsmooth, indeed, it is strongly semsmooth. This can be seen by Lemma 2.2(c), Lemma 3.1 and Theorem 7 of [23]. Also every  $LC^1$  function is strongly semsmooth. Thus, the result follows.  $\square$

**Corollary 3.2** *The function  $\psi_p$  defined as in (8) is an  $LC^1$  function. Hence, if every  $F_i$  is an  $LC^1$  function, then the function  $\Psi_p$  given as (10) is also an  $LC^1$  function.*

Some other issues related to semsmooth functions are concepts of piecewise smooth and almost smooth functions. It is well-known that piecewise smooth functions are examples of semsmooth functions and there have emerged other examples of semsmooth functions that are not piecewise smooth recently, see [23] and references therein. In particular, these examples include the  $p$ -norm function with  $1 < p < \infty$  defined on  $\mathbb{R}^n$  where  $n \geq 2$ , the Euclidean norm function, pseudo-smooth NCP-functions, smoothing functions, etc.. To close this section, we point out that the NCP-function studied in this paper and [1] is indeed strongly almost smooth since it is based on the  $p$ -norm function. We briefly state definition of almost smooth functions and the result as below.

**Definition 3.3** *The almost smooth (respectively, strongly almost smooth) functions are functions that are semsmooth (respectively, strongly semsmooth) on the whole space  $\mathbb{R}^n$  and smooth everywhere except on sets with "dimension" less than  $n - 1$  in the sense that the sets do not locally partition  $\mathbb{R}^n$  into multiple connected components.*

By applying Lemma 2.2(c), 3.1 and a result in [23], we immediately have an interesting property in relation to strongly almost smoothness for  $\Phi_p$ . For more details regarding to almost smooth and strongly almost smooth functions, please refer to the recent paper [23].

**Proposition 3.3** *If every  $F_i$  is an  $LC^1$  function, then the function  $\Phi_p$  defined as (9) is strongly almost smooth function.*

**Proof.** This result follows by Lemma 2.2(c), Prop. 3.1, and Theorem 7 of [23].  $\square$

## 4 A descent method

In this section, we study an almost the same descent method as in Sec. 4 of [1] for solving the unconstrained minimization (11), which does not require the derivative of  $F$  involved in the NCP. In fact, we consider the same search direction for the algorithm as in [1]:

$$d^k := -\nabla_b \psi_p(x^k, F(x^k)), \quad (20)$$

except the way to obtain the step-size is slightly different (see Step 3). Such a way to find step-size can also be found in the literature, for instance in [11]. Using the property of  $\psi_p$  being globally Lipschitz continuous (see Lemma 3.1), we have an alternative proof for the convergence result of the same descent method considered as in [1]. We state the detailed steps as below.

### Algorithm 4.1

(Step 0) Choose  $x^0 \in \mathbb{R}^n$ ,  $\varepsilon \geq 0$ ,  $\sigma \in (0, 1)$ ,  $\beta \in (0, 1)$  and set  $k := 0$ .

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

(Step 2) Let

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

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

$$\Psi_p(x^k + \beta^m d^k) \leq \Psi_p(x^k) - \sigma \beta^{2m} \|d^k\|^2. \quad (21)$$

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

We wish to show the global convergence result for Algorithm 4.1 under the strongly monotone assumption of  $F$ . The following lemmas plus Lemma 3.1 will enable the convergence result for the algorithm. In what follows, we assume that the parameter  $\varepsilon$  used in Algorithm 4.1 is set to be zero and Algorithm 4.1 generates an infinite sequence  $\{x^k\}$ .

**Lemma 4.1** (*[1, Lem. 4.1]*) *Let  $x^k \in \mathbb{R}^n$  and  $F$  be a monotone function. Then the search direction defined as (20) satisfies the descent condition  $\nabla \Psi_p(x^k)^T d^k < 0$  as long as  $x^k$  is not a solution of the NCP. Moreover, if  $F$  is strongly monotone with modulus  $\mu > 0$  then  $\nabla \Psi_p(x^k)^T d^k \leq -\mu \|d^k\|^2$ .*

**Lemma 4.2** *If  $F$  is strongly monotone, then the NCP has at most one solution.*

**Proof.** Suppose there are two solutions  $\zeta^*, x^* \in \mathbb{R}^n$  such that

$$\begin{cases} \langle F(\zeta^*), \zeta^* \rangle = 0, \\ F(\zeta^*) \geq 0, \zeta^* \geq 0 \end{cases} \quad \text{and} \quad \begin{cases} \langle F(x^*), x^* \rangle = 0, \\ F(x^*) \geq 0, x^* \geq 0. \end{cases}$$

By  $F$  is strongly monotone, we have  $\langle F(\zeta^*) - F(x^*), \zeta^* - x^* \rangle > 0$ . However,

$$\begin{aligned} & \langle F(\zeta^*) - F(x^*), \zeta^* - x^* \rangle \\ &= \langle F(\zeta^*), \zeta^* \rangle + \langle F(x^*), x^* \rangle - \langle F(\zeta^*), x^* \rangle - \langle F(x^*), \zeta^* \rangle \\ &= -\langle F(\zeta^*), x^* \rangle - \langle F(x^*), \zeta^* \rangle \\ &\leq 0, \end{aligned}$$

where the inequality is due to  $F(\zeta^*), \zeta^*, F(x^*), x^*$  are all nonnegative. Hence, it is a contradiction and therefore there is at most one solution for the NCP.  $\square$

**Proposition 4.1** *Suppose that  $F$  is continuously differentiable and strongly monotone with modulus  $\mu > 0$ . Let  $x^0 \in \mathbb{R}^n$  be any starting point and  $\mathcal{L}(x^0)$  denote its level set. Assume  $\nabla F$  is Lipschitz continuous in  $\mathcal{L}(x^0)$ . Then the sequence  $\{x^k\}$  generated by Algorithm 4.1 converges to the unique solution of the NCP.*

**Proof.** From Lemma 3.1 and the assumption of  $\nabla F$  being Lipschitz continuous, we obtain  $\nabla \Psi_p$  is also Lipschitz continuous in  $\mathcal{L}(x^0)$ , i.e.,

$$\|\nabla \Psi_p(x) - \nabla \Psi_p(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{L}(x^0)$$

and some constant  $L > 0$ . Let  $x^{k+1} = x^k + td^k$ ,  $0 \leq t \leq 1$  and  $\theta^k = x^k + \rho^k(x^{k+1} - x^k)$ , where  $\rho^k \in (0, 1)$ . Then, Lemma 4.1 and the Mean Value Theorem yield

$$\begin{aligned} & \Psi_p(x^{k+1}) - \Psi_p(x^k) \\ &= \nabla \Psi_p(\theta^k)^T (x^{k+1} - x^k) \\ &= t \nabla \Psi_p(\theta^k)^T d^k \\ &= t \nabla \Psi_p(x^k)^T d^k + t(\nabla \Psi_p(\theta^k) - \nabla \Psi_p(x^k))^T d^k \\ &\leq -t\mu \|d^k\|^2 + tL \|\theta^k - x^k\| \cdot \|d^k\| \\ &\leq (-t\mu + t^2L) \|d^k\|^2, \end{aligned}$$

where the last inequality holds since

$$\|\theta^k - x^k\| \leq \|x^{k+1} - x^k\| \leq t\|d^k\|.$$

In other words, we have

$$\Psi_p(x^k + td^k) \leq \Psi_p(x^k) - \sigma t^2 \|d^k\|^2, \quad \forall 0 \leq t \leq \min\left\{1, \frac{\mu}{\sigma + L}\right\}.$$

Hence, the step-size  $t_k$  obtained in Step 3 of the algorithm is bounded from below by

$$t_k \geq \min\left\{\beta, \frac{\beta\mu}{\sigma + L}\right\}. \quad (22)$$

Thus, a step length  $t_k > 0$  satisfying the Armijo's rule (21) can be always found, so the algorithm is well-defined. Now, since  $\{\Psi_p(x^k)\}$  is decreasing and nonnegative, it follows from

$$\Psi_p(x^k + td^k) \leq \Psi_p(x^k) - \sigma t_k \|d^k\|^2,$$

and the inequality (22) that

$$\lim_{k \rightarrow \infty} \|d^k\|^2 = 0.$$

This implies that

$$\lim_{k \rightarrow \infty} \nabla_b \Psi_p(x^k, F(x^k)) = 0 \quad (23)$$

because of the definition of  $d^k$  as in (20). Then, by Lemma 2.3(e) and (23), any accumulation point of  $\{x^k\}$  is a solution of the NCP. On the other hand,  $\{x^k\}$  is in  $\mathcal{L}(x^0)$  which is bounded by Lemma 2.4, so there exists at least one accumulation point  $x^*$ . At last, due to Lemma 4.2, the NCP has a unique solution, so the entire sequence  $\{x^k\}$  converges to  $x^*$ .  $\square$

## 5 Conclusions

In this paper, we have shown properties of the NCP-function  $\phi_p$  which is an extension of the Fischer-Burmeister function as well as properties of  $\psi_p$  which is a merit function for the NCP formed by  $\phi_p$ . They were first proposed in [28] and further studied recently in [1]. In this paper, we continued to explore some favorable properties of  $\phi_p$  and  $\psi_p$  including semismoothness,  $SC^1$ -property,  $LC^1$ -property, and almost smoothness which are important from computational view of point. Moreover, we have also provided an alternative proof, which applies the new properties discovered in this paper, for the convergence result of the descent method considered in [1]. Since  $\psi_p$  is shown  $SC^1$  and  $LC^1$  function, it is an interesting future topic to extend some appropriate Newton methods for  $SC^1$  and  $LC^1$  minimization problems (see [19]) to the equivalent the minimization based on the merit function  $\psi_p$ .

On the other hand, some other NCP-functions based on the generalized Fischer-Burmeister function  $\phi_p$  are also recently studied in [2] by the author. According to the theoretical properties built herein and therein, the numerical implementation of related algorithms including comparison with other existing well-known algorithms for NCP definitely deserve a systematic study. We will leave it in a subsequent topic.

## References

- [1] CHEN, J.-S. AND PAN, S. (2006), *A family of NCP-function and a descent method for the nonlinear complementarity problem*, TO APPEAR IN COMPUTATIONAL OPTIMIZATION AND APPLICATIONS.
- [2] CHEN, J.-S. (2006), *On some NCP-functions based on the generalized Fischer-Burmeister function*, TO APPEAR IN ASIA-PACIFIC JOURNAL OF OPERATIONAL RESEARCH.
- [3] BORWEIN, J. M. AND LEWIS, A.-S. (2000), *Convex Analysis and Nonlinear Optimization: Theory and Examples*, SPRINGER-VERLAG, NEW YORK.
- [4] CLARKE, F. H. (1983), *Optimization and Nonsmooth Analysis*, WILEY, NEW YORK.
- [5] FACCHINEI, F. AND PANG, J.-S. (2003), *Finite-Dimensional Variational Inequalities and Complementarity Problems*, VOLUMES I AND II, SPRINGER-VERLAG, NEW YORK.
- [6] FACCHINEI, F. AND SOARES, J. (1997), *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM JOURNAL ON OPTIMIZATION, VOL. 7, PP. 225-247.
- [7] FERRIS, M. C. AND PANG, J.-S. (1997), *Engineering and economic applications of complementarity problems*, SIAM REVIEW, VOL. 39, PP. 669-713.
- [8] FISCHER, A. (1992), *A special Newton-type optimization methods*, OPTIMIZATION, VOL. 24, PP. 269-284.
- [9] FISCHER, A. (1997), *Solution of the monotone complementarity problem with locally Lipschitzian functions*, MATHEMATICAL PROGRAMMING, VOL. 76, PP. 513-532.
- [10] FUKUSHIMA, M. (1996), *Merit functions for variational inequality and complementarity problems*, NONLINEAR OPTIMIZATION AND APPLICATIONS, EDITED BY G. DI PILLO AND F. GIANNESI, PLENUM PRESS, NEW YORK, PP. 155-170.
- [11] GEIGER, C. AND KANZOW, C. (1996), *On the resolution of monotone complementarity problems*, COMPUTATIONAL OPTIMIZATION AND APPLICATIONS, VOL. 5, PP. 155-173.
- [12] HARKER, P. T. AND PANG, J.-S. (1990), *Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications*, MATHEMATICAL PROGRAMMING, VOL. 48, PP. 161-220.
- [13] HIGHAM, N. J. (1992), *Estimating the matrix p-norm*, NUMERICAL MATHEMATICS, VOL. 62, PP. 539-555.

- [14] JIANG, H. (1996), *Unconstrained minimization approaches to nonlinear complementarity problems*, JOURNAL OF GLOBAL OPTIMIZATION, VOL. 9, PP. 169-181.
- [15] KANZOW, C. (1996), *Nonlinear complementarity as unconstrained optimization*, JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS, VOL. 88, PP. 139-155.
- [16] MIFFLIN, R. (1977), *Semismooth and semiconvex functions in constrained optimization*, SIAM JOURNAL ON CONTROL OPTIMIZATION, 15, PP. 959-972.
- [17] ORTEGA, J. M. AND RHEINBOLDT, W. (2000), *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, PHILADELPHIA.
- [18] PANG, J.-S. (1994), *Complementarity problems*, HANDBOOK OF GLOBAL OPTIMIZATION, EDITED BY R. HORST AND P. PARDALOS, KLUWER ACADEMIC PUBLISHERS, BOSTON, MASSACHUSETTS, PP. 271-338.
- [19] PANG, J.-S. AND QI, L. (1995), *A globally convergent Newton methods for convex  $SC^1$  minimization problems*, JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS, 85, PP. 633-648.
- [20] QI, L. (1993), *Convergence analysis of some algorithms for solving nonsmooth equations*, MATHEMATICS OF OPERATIONS RESEARCH, 18, PP. 227-244.
- [21] QI, L. (1994), *Superlinearly convergent approximate Newton methods for  $LC^1$  optimization problems*, MATHEMATICAL PROGRAMMING, 64, PP. 277-294.
- [22] QI, L. AND SUN, J. (1993), *A nonsmooth version of Newton's method*, MATHEMATICAL PROGRAMMING, 58, PP. 353-367.
- [23] QI, L. AND TSENG, P. (2002), *An Analysis of piecewise smooth functions and almost smooth functions*, SUBMITTED TO MATHEMATICS OF OPERATIONS RESEARCH.
- [24] ROCKAFELLAR, R. T. (1970), *Convex Analysis*, PRINCETON UNIVERSITY PRESS, PRINCETON NJ.
- [25] ROCKAFELLAR, R. T. AND WETS, J.-B. (1998), *Variational Analysis*, SPRINGER-VERLAG, BERLIN.
- [26] SUN, D. AND QI, L. (1999), *On NCP-functions*, COMPUTATIONAL OPTIMIZATION AND APPLICATIONS, VOL. 13, PP. 201-220.
- [27] SUN, D. AND SUN, J. (2002), *Semismooth matrix valued functions*, MATHEMATICS OF OPERATIONS RESEARCH, VOL. 27, PP. 150-169.
- [28] TSENG, P. (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.

- [29] YAMADA, K., YAMASHITA, N. AND FUKUSHIMA, M. (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. GIANNESI, KLUWER ACADEMIC PUBLISHERS, NETHERLANDS, PP. 463–489.
- [30] YAMASHITA, N. AND FUKUSHIMA, M. (1995), *On stationary points of the implicit Lagrangian for the nonlinear complementarity problems*, *JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS*, VOL. 84, PP. 653-663.
- [31] YAMASHITA, N. AND FUKUSHIMA, M. (1997), *Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems*, *MATHEMATICAL PROGRAMMING*, VOL. 76, PP. 469-491.