The Linear Convergence of a Derivative-Free Descent Method for Nonlinear Complementarity Problems

Wei-Zhe Gu

Department of Mathematics, School of Science

Tianjin University, Tianjin 300072, P.R. China

Email: guweizhe@tju.edu.cn

Li-Yong Lu *

Department of Mathematics, School of Science

Tianjin University of Technology, Tianjin 300384, P.R. China

Email: lylumath@gmail.com

October 9, 2013

Abstract

Recently, Hu, Huang and Chen [Properties of a family of generalized NCP-functions and a derivative free algorithm for complementarity problems, J. Comput. Appl. Math. 230 (2009): 69-82] introduced a family of generalized NCP-functions, which include many existing NCP-functions as special cases. They obtained several favorite properties of the functions; and by which, they showed that a derivative-free descent method is globally convergent under suitable assumptions. However, no result on convergent rate of the method was reported. In this paper, we further investigate some properties of this family of generalized NCP-functions. In particular, we show that, under suitable assumptions, the iterative sequence generated by the descent method discussed in their paper converges globally at a linear rate to a solution of the nonlinear complementarity problem. Some preliminary numerical results are reported, which verify the theoretical results obtained.

Key words: Nonlinear complementarity problems, merit function, derivative-free descent method, linear convergence.

*Corresponding author

AMS subject classifications: 90C33, 65K10

Running Title: A Derivative-Free Method for Complementarity Problems

1 Introduction

The nonlinear complementarity problem (NCP for short) is to find a vector $x \in \Re^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad \text{and} \quad x^T F(x) = 0,$$

where $F : \Re^n \to \Re^n$ is a given function. The NCP has been studied extensively due to its various applications in many fields. Such as mathematical programming, economics, engineering and mechanics (see, for example, [7, 9, 11]). We refer the interested readers to see the excellent monograph by Facchinei and Pang [7]. Various methods for solving the NCP have been proposed in the literature (see, for example, [2, 3, 4, 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23, 24]). Among which, one of the most popular and powerful approaches is to reformulate the NCP as an unconstrained minimization problem [2, 4, 5, 7, 8, 9, 18, 19, 20, 24]. This kind of methods is called the merit function method, where the merit function is generally constructed by some NCP-function.

Definition 1.1 *A function $\phi : \Re^2 \to \Re$ is called an NCP-function [17], if it satisfies*

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

Furthermore, if $\phi(a, b) \geq 0$, then the NCP-function ϕ is called a nonnegative NCP-function. In addition, if a function $\Psi : \Re^n \to \Re$ is nonnegative and $\Psi(x) = 0$ if and only if x solves the NCP, then Ψ is called a merit function for the NCP.

If the NCP-function ϕ is nonnegative on $\Re^2$, then it is easy to see that the function $\Psi : \Re^n \to \Re$ defined by $\Psi(x) = \sum_{i=1}^n \phi(x_i, F_i(x))$ is a merit function for the NCP. Thus, finding a solution of the NCP is equivalent to finding a global minimum of the unconstrained minimization $\min_{x \in \Re^n} \Psi(x)$ with the objective function value being zero.

Recently, Hu, Huang, and Chen [12] proposed the merit function

$$\Psi_{\theta p}(x) = \sum_{i=1}^n \psi_{\theta p}(x_i, F_i(x)) = \frac{1}{2} \sum_{i=1}^n \phi^2_{\theta p}(x_i, F_i(x)), \quad \forall x \in \Re^n, \tag{1.1}$$

where $\phi_{\theta p} : \Re^2 \to \Re$ is defined by

$$\phi_{\theta p}(a, b) = \sqrt[p]{\theta \left(|a|^p + |b|^p\right) + (1 - \theta)|a - b|^p} - (a + b), \quad \forall (a, b) \in \Re^2, \tag{1.2}$$

with $p > 1$ and $\theta \in (0, 1]$. It is easy to show that $\phi_{\theta p}(\cdot, \cdot)$ is an NCP-function. Thus, finding a solution of the NCP is equivalent to finding a global minimum of the unconstrained minimization

$$\min_{x \in \Re^n} \Psi_{\theta p}(x)$$

with the objective function value being zero. It is easy to see that if $\theta = 0$ and $p = 2$, the function $\Psi_{\theta p}(\cdot)$ defined by (1.1) reduces to the natural residual merit function $\Psi_{NR} : \Re^n \to \Re$ given by

$$\Psi_{NR}(x) = \frac{1}{2} \sum_{i=1}^{n} \phi_{NR}^2(x_i, F_i(x)), \tag{1.3}$$

where $\phi_{NR} : \Re^2 \to \Re$ is an NCP-function given by $\phi_{NR}(a, b) = -2 \min\{a, b\}$. If $\theta = 1$, the function $\Psi_{\theta p}(\cdot)$ defined by (1.1) reduces to the merit function $\Psi_p : \Re^n \to \Re$ given by

$$\Psi_p(x) = \frac{1}{2} \sum_{i=1}^{n} \phi_p^2(x_i, F_i(x)), \tag{1.4}$$

where $\phi_p : \Re^2 \to \Re$ is an NCP-function given by $\phi_p(a, b) = \|(a, b)\|_p - (a + b)$. The NCP-function ϕ_p was introduced by Luo and Tseng [19], and further studied by Chen [2] and Chen and Pan [4, 5]. Obviously, when $p = 2$, the NCP-function ϕ_p reduces the Fischer-Burmeister NCP-function $\phi_{FB}(a, b) = \sqrt{a^2 + b^2} - (a + b)$.

Recently, the so-called derivative-free methods have attracted much attention, which do not require computation of derivatives of F [12]. Derivative-free methods are particularly suitable for problems where the derivatives of F are not available or are extremely expensive to compute. For the NCP, the authors in [12] investigated a derivative-free descent method based on the merit function $\Psi_{\theta p}$ and the method is showed to be globally convergent under the assumption that F is strongly monotone. In this paper, our object is to discuss the rate of convergence for the derivative-free descent method discussed by [12] based on the merit function $\Psi_{\theta p}(x)$ defined by (1.1).

The paper is organized as follows. In Section 2, we review some definitions and preliminary results that will be used in the sequent analysis. In Section 3, we first discuss some properties of the merit function $\Psi_{\theta p}$, including the growth behavior of the merit function $\Psi_{\theta p}$, the estimation on the upper bound of $\Psi_{\theta p}$, and the Lipschitz continuity of the gradient function $\nabla \Psi_{\theta p}$; and then, we establish the linear rate of convergence of the derivative-free descent method discussed in [12]. Some final remarks are given in Section 4.

Throughout this paper, $\Re^n$ denotes the space of n-dimensional real column vectors and A^T denotes the transpose of the real-valued matrix A. For any differentiable function $F : \Re^n \to \Re$, $\bigtriangledown F(x)$ denotes the gradient of F at x. For any differentiable mapping $F = (F_1, \ldots, F_m)^T : \Re^n \to \Re^m$, $\bigtriangledown F(x) = [\bigtriangledown F_1(x), \cdots, \bigtriangledown F_m(x)]^T$ denotes the Jacobian of F at

x. The norm $\|\cdot\|$ denotes the Euclidean norm. $\lceil z \rceil$ denotes the smallest integer no less than z, for any $z \in \Re$. The level set of a function $\Psi : \Re^n \to \Re$ is denoted by $\mathcal{L}(\Psi, \gamma) = \{x \in \Re^n \mid \Psi(x) \leq \gamma\}$. We denote the set of positive integers by $\mathcal{N}^+$.

2 Preliminaries

In this section, we recall some definitions and preliminary results, which will be used in the sequent analysis. The derivative-free descent method discussed in [12] is also given in this section.

Definition 2.1 *Given the continuously differentiable function $F : \Re^n \to \Re^n$, we say that*

(i) F is monotone if $\langle x - y, F(x) - F(y)\rangle \geq 0$ for all $x, y \in \Re^n$;

(ii) F is a strongly monotone if, for all $x, y \in \Re^n$, F satisfies that $\langle x - y, F(x) - F(y)\rangle \geq \lambda\|x - y\|^2$, or, equivalently, $\langle \nabla F(x)y, y\rangle \geq \lambda\|y\|^2$, for some $\lambda > 0$;

(iii) F is a uniform P-function with modulus $\kappa > 0$ if $\max_{\substack{1\leq i\leq n \\ x_i \neq y_i}} (x_i - y_i)(F_i(x) - F_i(y)) \geq \kappa\|x - y\|^2$ for all $x, y \in \Re^n$;

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

(v) F is Lipschitz continuous if there exists a constant $L > 0$ such that $\|F(x) - F(y)\| \leq L\|x - y\|$ for all $x, y \in \Re^n$.

It is well-known that every monotone function is a P_0-function and every strongly monotone function is a uniform P-function. For a continuously differentiable function F, if its (transpose) Jacobian $\nabla F(x)$ is a P-matrix then it is a P-function (the converse may not be true); and the (transpose) Jacobian $\nabla F(x)$ is a P_0-matrix if and only if F is a P_0-function. For more properties of various monotone and $P(P_0)$-functions, please refer to [7].

Definition 2.2 *[22] Let the sequence $\{x^k\}$ converge to x^*.*

(i) The sequence $\{x^k\}$ is said to be Q-linearly convergent, if there is a constant $\beta \in (0, 1)$, which is independent of the iterative number k, such that $\lim_{k\to\infty} \frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|} = \beta$.

(ii) The sequence $\{x^k\}$ is said to be R-linearly convergent, if there is a sequence of non-negative scalars $\{q^k\}$ such that $\|x^k - x^*\| \leq q^k$ for all k, and $\{q^k\}$ converges Q-linearly to zero.

Lemma 2.1 *([12, Proposition 2.3(v) and Proposition 2.5 (v)]) Let* $\psi_{\theta p}(\cdot,\cdot)$ *and* $\phi_{\theta p}(\cdot,\cdot)$ *be defined by* (1.1) *and* (1.2)*, respectively, with* $p > 1$ *and* $1 > \theta \geq 0$*, then we have*

(i) $\phi_{\theta p}$ *is continuously differentiable on* $\Re^2 \setminus \{(0,0)\}$ *and when* $(a,b) \neq (0,0)$*,*

$$\frac{\partial \phi_{\theta p}(a,b)}{\partial a} = \frac{\theta sgn(a)|a|^{p-1} + (1-\theta)sgn(a-b)|a-b|^{p-1}}{\left[\theta\left(|a|^p + |b|^p\right) + (1-\theta)|a-b|^p\right]^{\frac{p-1}{p}}} - 1,$$
$$\frac{\partial \phi_{\theta p}(a,b)}{\partial b} = \frac{\theta sgn(b)|b|^{p-1} - (1-\theta)sgn(a-b)|a-b|^{p-1}}{\left[\theta\left(|a|^p + |b|^p\right) + (1-\theta)|a-b|^p\right]^{\frac{p-1}{p}}} - 1;$$

(ii) $\frac{\partial \psi_{\theta p}}{\partial a}(a,b)\frac{\partial \psi_{\theta p}}{\partial b}(a,b) \geq 0$ *for all* $(a,b)^T \in \Re^2$*, where the equality holds if and only if* $\phi_{\theta p}(a,b) = 0$*. And when* $(a,b) \neq (0,0)$*,*

$$\frac{\partial \psi_{\theta p}(a,b)}{\partial a} = \frac{\partial \phi_{\theta p}(a,b)}{\partial a}\phi_{\theta p}(a,b), \quad \frac{\partial \psi_{\theta p}(a,b)}{\partial b} = \frac{\partial \phi_{\theta p}(a,b)}{\partial b}\phi_{\theta p}(a,b).$$

Algorithm 2.1 *([12])(A Derivative-Free Descent Method).*

Step 0 *Given* $p > 1$, $\theta \in (0,1]$, $x^0 \in \Re^n$ *and any small number* $\varepsilon > 0$*. Choose* $\sigma, \gamma, \eta \in (0,1)$ *with* $\gamma < \eta$*. Set* $k := 0$*.*

Step 1 *If* $\Psi_{\theta p}(x^k) \leq \varepsilon$*, then stop. Otherwise go to Step 2.*

Step 2 *Let*

$$d^k(\eta^t) := -\nabla_b \Psi_{\theta p}(x^k, F(x^k)) - \eta^t \nabla_a \Psi_{\theta p}(x^k, F(x^k)).$$

Set $x^{k+1} := x^k + \gamma^{t_k} d^k(\eta^{t_k})$*, where* t_k *is the smallest nonnegative integer* t *satisfying*

$$\Psi_{\theta p}(x^k + \gamma^t d^k(\eta^t)) \leq (1 - \sigma\gamma^{2t})\Psi_{\theta p}(x^k). \tag{2.1}$$

Step 3 *Set* $k := k+1$ *and go to Step 1.*

The following error bound result from [21] will be used in our analysis later.

Theorem 2.1 *Let* F *be strongly monotone with modulus with* λ *and Lipschitz continuous with* $L > 0$ *on* $\Re^n$*. Then*

$$\|x - x^*\| \leq \frac{L+1}{\lambda}\sqrt{\sum_{i=1}^{n} |\min\{x_i, F_i(x)\}|^2} = \frac{L+1}{\sqrt{2}\lambda}\Psi_{NR}(x)^{\frac{1}{2}}$$

holds for all $x \in \Re^n$*, where* x^* *is the unique solution of the NCP.*

3 The Rate of Convergence

In this section, we first investigate several properties of the merit function $\Psi_{\theta p}$; and then, we discuss the linear convergence of Algorithm 2.1.

Firstly, we investigate the growth behavior of the two merit functions: $\Psi_{\theta p}$ and Ψ_{NR}.

Lemma 3.1 *Let $\phi_{\theta p} : \Re^2 \to \Re$ be defined in (1.2). Then, for any $p > 1$ and $\theta \in (0, 1]$,*

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

Proof: Without loss of generality, we suppose $a \geq b$ in the following proof. If $ab = 0$, it is trivial. We will prove the desired results by considering the following two cases: $ab > 0$; and $ab < 0$.

Case(i): Suppose that $ab > 0$. In this case, we have the following two subcases:

(a) Suppose that $a > 0$ and $b > 0$. Then,

$$\begin{aligned} \phi_{\theta p}(a, b) &= \sqrt[p]{\theta(|a|^p + |b|^p) + (1-\theta)|a - b|^p} - (a + b) \\ &\leq \sqrt[p]{|a|^p + |b|^p} - (a + b) < 0. \end{aligned}$$

Hence,

$$\begin{aligned} |\phi_{\theta p}(a, b)| &= (a + b) - \sqrt[p]{\theta(|a|^p + |b|^p) + (1-\theta)|a - b|^p} \\ &= b\left[\frac{a}{b} + 1 - \sqrt[p]{\theta(|\frac{a}{b}|^p + 1) + (1-\theta)|\frac{a}{b} - 1|^p}\right]. \end{aligned}$$

Let

$$H(t) = t + 1 - \sqrt[p]{\theta(|t|^p + 1) + (1-\theta)|t - 1|^p} \qquad t \in [1, +\infty),$$

then,

$$H'(t) = 1 - \frac{\theta|t|^{p-1} + (1-\theta)|t-1|^{p-1}}{[\theta(|t|^p + 1) + (1-\theta)|t-1|^p]^{\frac{p-1}{p}}} \geq 0.$$

Denote

$$\begin{aligned} H(+\infty) &:= \lim_{t\to+\infty} H(t) \\ &= \lim_{u\to 0^+} \frac{1+u-\sqrt[p]{\theta(1+u^p)+(1-\theta)(1-u)^p}}{u} \\ &= \lim_{u\to 0^+} \left(1 - \frac{\theta u^{p-1}-(1-\theta)(1-u)^{p-1}}{[\theta(1+u^p)+(1-\theta)(1-u)^p]^{\frac{p-1}{p}}}\right) = 2 - \theta. \end{aligned}$$

Hence, we obtain that

$$\left(2 - (2\theta)^{\frac{1}{p}}\right) |\min\{a, b\}| = H(1)b \leq |\phi_{\theta p}(a, b)| \leq H(+\infty)b = (2-\theta)b \leq 2 \left|\min\{a, b\}\right|.$$

(b) Suppose that $a < 0$ and $b < 0$. From the definition of $\phi_{\theta p}$ in (1.2), we have $\phi_{\theta p}(a,b) > 0$. Thus,

$$\begin{aligned} |\phi_{\theta p}(a,b)| &= \sqrt[p]{\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p} - (a+b) \\ &= -b\left[\sqrt[p]{\theta(|\frac{a}{b}|^p + 1) + (1-\theta)|\frac{a}{b} - 1|^p} + \frac{a}{b} + 1\right]. \end{aligned}$$

Let

$$H(t) = \sqrt[p]{\theta(|t|^p + 1) + (1-\theta)|t-1|^p} + t + 1, \qquad t \in (0,1],$$

then,

$$H'(t) = \frac{\theta|t|^{p-1} - (1-\theta)|t-1|^{p-1}}{[\theta(|t|^p + 1) + (1-\theta)|t-1|^p]^{\frac{p-1}{p}}} + 1 \geq 0.$$

Hence, we obtain that

$$2|\min\{a,b\}| = H(0)(-b) \leq |\phi_{\theta p}(a,b)| \leq H(1)(-b) = \left(2 + (2\theta)^{\frac{1}{p}}\right)|\min\{a,b\}|.$$

Case(ii): Suppose that $ab < 0$. It follows that $a > 0 > b$. From the definition of $\phi_{\theta p}$ in (1.2), we have $\phi_{\theta p}(a,b) > 0$ in this case. Then,

$$\begin{aligned} |\phi_{\theta p}(a,b)| &= \sqrt[p]{\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p} - (a+b) \\ &= -b\left[\sqrt[p]{\theta(|\frac{a}{b}|^p + 1) + (1-\theta)|\frac{a}{b} - 1|^p} + \frac{a}{b} + 1\right]. \end{aligned}$$

Let

$$H(t) = \sqrt[p]{\theta(|t|^p + 1) + (1-\theta)|t-1|^p} + t + 1, \qquad t \in (-\infty, 0),$$

then,

$$H'(t) = \frac{\theta|t|^{p-1} - (1-\theta)|t-1|^{p-1}}{[\theta(|t|^p + 1) + (1-\theta)|t-1|^p]^{\frac{p-1}{p}}} + 1 \geq 0.$$

Denote

$$\begin{aligned} H(-\infty) &:= \lim_{t\to-\infty} H(t) \\ &= \lim_{u\to 0^-} \frac{1+u-\sqrt[p]{\theta(1+(-u)^p)+(1-\theta)(1-u)^p}}{u} \\ &= \lim_{u\to 0^-}\left(1 - \frac{-\theta(-u)^{p-1}-(1-\theta)(1-u)^{p-1}}{[\theta(1+(-u)^p)+(1-\theta)(1-u)^p]^{\frac{p-1}{p}}}\right) = 2-\theta. \end{aligned}$$

Hence, we obtain that

$$\begin{aligned} \left(2 - (2\theta)^{\frac{1}{p}}\right)|\min\{a,b\}| \leq (2-\theta)(-b) = H(-\infty)(-b) &\leq |\phi_{\theta p}(a,b)| \\ &\leq H(0)(-b) = 2|\min\{a,b\}|. \end{aligned}$$

Combining the results of **Cases (i)-(ii)**, we complete the proof. □

From the definitions of $\Psi_{\theta p}$,Ψ_{NR} and Lemma 3.1, we immediately get the following lemma.

Lemma 3.2 *Let $\Psi_{\theta p}$ and Ψ_{NR} be defined by* (1.1) *and* (1.3)*, respectively, with $p > 1$ and $1 \geq \theta > 0$. Then, for any $x \in \Re^n$,*

$$\left(\frac{2-(2\theta)^{\frac{1}{p}}}{2}\right)^2 \Psi_{NR}(x) \leq \Psi_{\theta p}(x) \leq \left(\frac{2+(2\theta)^{\frac{1}{p}}}{2}\right)^2 \Psi_{NR}(x).$$

Secondly, we give an estimation on the upper bound of $\Psi_{\theta p}$.

Lemma 3.3 *For all $(a,b) \neq (0,0)$ and $p > 1$, $\theta \in (0,1]$, we have the following inequality:*

$$\left(\frac{\theta|a|^{p-1}sgn(a)+\theta|b|^{p-1}sgn(b)}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p]^{\frac{p-1}{p}}}-2\right)^2 \geq \left(2-(2\theta)^{\frac{1}{p}}\right)^2.$$

Proof: Without loss of generality, we assume $a \geq b$. If $ab = 0$, it is trivial. We will prove the desired result by considering the following three cases: $a > 0$ and $b > 0$; $a < 0$ and $b < 0$; and $ab < 0$.

Case(a): Suppose that $a > 0$ and $b > 0$. Then,

$$\begin{aligned}\frac{\theta|a|^{p-1}\mathrm{sgn}(a)+\theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p]^{\frac{p-1}{p}}} &= \frac{\theta|a|^{p-1}+\theta|b|^{p-1}}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \\ &= \frac{\theta[1+(\frac{b}{a})^{p-1}]}{[\theta(1+(\frac{b}{a})^p)+(1-\theta)(1-\frac{b}{a})^p]^{\frac{p-1}{p}}} \geq 0.\end{aligned}$$

Let

$$L(t) = \frac{\theta[1+t^{p-1}]}{[\theta(1+t^p)+(1-\theta)(1-t)^p]^{\frac{p-1}{p}}}, \quad t \in (0,1],$$

by direct calculation, we get

$$L'(t) = \frac{\theta t^{p-2}(1-t)+(1-\theta)t^{p-2}(1-t)^p+(1-\theta)(1-t)^{p-1}+(1-\theta)(1-t)^{p-1}t^{p-1}}{[\theta(1+t^p)+(1-\theta)(1-t)^p]^{\frac{2p-1}{p}}}\theta(p-1).$$

Thus, for all $t \in (0,1]$, we have $L'(t) \geq 0$. Furthermore,

$$L(t) \leq L(1) = (2\theta)^{\frac{1}{p}},$$

which in turn implies that

$$2-\frac{\theta|a|^{p-1}\mathrm{sgn}(a)+\theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p+|b|^p)+(1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \geq 2-(2\theta)^{\frac{1}{p}}.$$

Squaring both sides leads to the desired inequality.

Case(b): Suppose that $a < 0$ and $b < 0$. Then, $-a > 0$ and $-b > 0$. From **Case(a)**, we have

$$0 \geq \frac{\theta|a|^{p-1}\mathrm{sgn}(a) + \theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \quad = \quad -\frac{\theta|a|^{p-1} + \theta|b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \geq -(2\theta)^{\frac{1}{p}}.$$

Then,

$$\left(\frac{\theta|a|^{p-1}\mathrm{sgn}(a) + \theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} - 2\right)^2 \geq 2^2 \geq \left(2 - (2\theta)^{\frac{1}{p}}\right)^2.$$

Case(c): Suppose that $ab < 0$. In this case, we have the following two subcases:
(i) Suppose that $|a| \geq |b|$, then

$$\begin{aligned} 0 \leq \frac{\theta|a|^{p-1}\mathrm{sgn}(a) + \theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \quad &= \quad \frac{\theta|a|^{p-1} - \theta|b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \\ &\leq \quad \frac{\theta|a|^{p-1} + \theta|b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} \\ &\leq \quad (2\theta)^{\frac{1}{2}}. \end{aligned}$$

Hence,

$$\left(\frac{\theta|a|^{p-1}\mathrm{sgn}(a) + \theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} - 2\right)^2 \geq \left(2 - (2\theta)^{\frac{1}{p}}\right)^2.$$

(ii) Suppose that $|a| < |b|$, we have

$$\frac{\theta|a|^{p-1}\mathrm{sgn}(a) + \theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} < 0,$$

then

$$\left(\frac{\theta|a|^{p-1}\mathrm{sgn}(a) + \theta|b|^{p-1}\mathrm{sgn}(b)}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} - 2\right)^2 \geq 2^2 \geq \left(2 - (2\theta)^{\frac{1}{p}}\right)^2.$$

Combining cases (a)-(c), we complete the proof. □

Proposition 3.1 *Let $\Psi_{\theta p}$ be defined in (1.1) with $p > 1$ and $0 < \theta \leq 1$. Then, for all $x \in \Re^n$, we have*

$$\|\nabla_a \Psi_{\theta p}(x, F(x)) + \nabla_b \Psi_{\theta p}(x, F(x))\|^2 \geq 2\left(2 - (2\theta)^{\frac{1}{p}}\right)^2 \Psi_{\theta p}(x),$$

where

$$\nabla_a \Psi_{\theta p}(x, F(x)) := \left(\frac{\partial \psi_{\theta p}(x_1, F_1(x))}{\partial a}, \ldots, \frac{\partial \psi_{\theta p}(x_n, F_n(x))}{\partial a}\right),$$

$$\nabla_b \Psi_{\theta p}(x, F(x)) := \left(\frac{\partial \psi_{\theta p}(x_1, F_1(x))}{\partial b}, \ldots, \frac{\partial \psi_{\theta p}(x_n, F_n(x))}{\partial b}\right).$$

Proof: If $(a,b) \neq (0,0)$, then from Lemma 2.1, we obtain

$$
\begin{aligned}
\frac{\partial \psi_{\theta p}(a,b)}{\partial a} &= \left(\frac{\theta \mathrm{sgn}(a)|a|^{p-1} + (1-\theta)\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} - 1 \right) \phi_{\theta p}(a,b), \\
\frac{\partial \psi_{\theta p}(a,b)}{\partial b} &= \left(\frac{\theta \mathrm{sgn}(b)|b|^{p-1} - (1-\theta)\mathrm{sgn}(a-b)|a-b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} - 1 \right) \phi_{\theta p}(a,b),
\end{aligned}
$$

therefore,

$$
\frac{\partial \psi_{\theta p}(a,b)}{\partial a} + \frac{\partial \psi_{\theta p}(a,b)}{\partial b} = \left(\frac{\theta \mathrm{sgn}(a)|a|^{p-1} + \theta \mathrm{sgn}(b)|b|^{p-1}}{[\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p]^{\frac{p-1}{p}}} - 2 \right) \phi_{\theta p}(a,b).
$$

From Lemma 3.3, we get

$$
\begin{aligned}
\|\nabla_a \Psi_{\theta p}(x, F(x)) + \nabla_b \Psi_{\theta p}(x, F(x))\|^2 &\geq \left(2 - (2\theta)^{\frac{1}{p}}\right)^2 \sum_{i=1}^{n} \phi_{\theta p}^2(x_i, F_i(x)) \\
&= 2\left(2 - (2\theta)^{\frac{1}{p}}\right)^2 \Psi_{\theta p}(x),
\end{aligned}
$$

and hence, we complete the proof. □

Thirdly, we investigate the Lipschitz continuity of the gradient function $\nabla \Psi_{\theta p}$. To this end, we need the result on the boundedness of the level set for the function $\Psi_{\theta p}$.

Lemma 3.4 *Let $\Psi_{\theta p}$ be defined by* (1.1) *with $p > 1$ and $1 > \theta \geq 0$. Suppose that F is either a strongly monotone function or a uniform P-function. Then the level set*

$$
\mathcal{L}(\Psi_{\theta p}, \Psi_{\theta p}(x^0)) := \{x \in \Re^n | \Psi_{\theta p}(x) \leq \Psi_{\theta p}(x^0)\}
$$

is bounded, where x^0 is the starting point in Algorithm 2.1.

Proof: This result can be directly obtained from [12, Proposition 3.4]. □

Since $\nabla \Psi_{\theta p}$ and F are continuous on $\Re^n$, then

$$
D(x^0) := \sup\{\|d(x)\| \mid x \in \mathcal{L}(\Psi_{\theta p}, \Psi_{\theta p}(x^0))\}
$$

is finite, where $d(x)$ is the search direction in Algorithm 2.1 computed at the point x. Therefore, the set

$$
\mathcal{B}(x^0) := \mathcal{L}(\Psi_{\theta p}, \Psi_{\theta p}(x^0)) + \{x \mid \|x\| \leq D(x^0)\}
$$

is bounded and closed.

Lemma 3.5 *([12, Theorem 2.1]) The gradient function of the function $\psi_{\theta p}$ defined by* (1.1) *with $p \geq 2$ and $1 > \theta \geq 0$ is Lipschitz continuous, that is, there exists a positive constant L_1 such that*

$$
\|\nabla \psi_{\theta p}(a,b) - \nabla \psi_{\theta p}(c,d)\| \leq L_1 \|(a,b) - (c,d)\|
$$

holds for all $(a,b), (c,d) \in \Re^2$.

Under the assumption that F and ∇F are Lipschitz continuous with some constant $L > 0$ on $\Re^n$, and from Lemma 3.5, we can further get the next lemma.

Lemma 3.6 *Let $\Psi_{\theta p}$ be defined by (1.1) with $p \geq 2$ and $1 > \theta \geq 0$. Suppose that F and ∇F are Lipschitz continuous with some constant $L > 0$ on $\mathcal{B}(x^0)$. Then $\nabla \Psi_{\theta p}$ is Lipschitz continuous on $\mathcal{B}(x^0)$, i.e., there exists a positive constant $\bar{L}$ such that*

$$\|\nabla \Psi_{\theta p}(x) - \nabla \Psi_{\theta p}(y)\| \leq \bar{L}\|x - y\|$$

holds for all $x, y \in \mathcal{B}(x^0)$.

Proof: Because $\nabla_b \Psi_{\theta p}$, F and ∇F are continuous on the bounded and closed set $\mathcal{B}(x^0)$, then there exist some constants $C, \rho > 0$ such that

$$\|\nabla_b \Psi_{\theta p}(x, F(x))\| \leq C \quad \text{and} \quad \|\nabla F(x)\| \leq \rho, \quad \forall x \in \mathcal{L}(\Psi_{\theta p}, \Psi_{\theta p}(x^0)).$$

For all $x, y \in \mathcal{B}(x^0)$, from Lemma 3.5, we have

$$\begin{aligned}
& \|\nabla_a \Psi_{\theta p}(x, F(x)) - \nabla_a \Psi_{\theta p}(y, F(y))\|^2 \\
\leq\ & \|\nabla_a \Psi_{\theta p}(x, F(x)) - \nabla_a \Psi_{\theta p}(y, F(y))\|^2 + \|\nabla_b \Psi_{\theta p}(x, F(x)) - \nabla_b \Psi_{\theta p}(y, F(y))\|^2 \\
\leq\ & L_1^2(\|x - y\|^2 + \|F(x) - F(y)\|^2) \\
\leq\ & L_1^2(1 + L^2)\, \|x - y\|^2 .
\end{aligned}$$

Denote $L_2 := L_1\sqrt{1 + L^2}$, then it follows that

$$\|\nabla_a \Psi_{\theta p}(x, F(x)) - \nabla_a \Psi_{\theta p}(y, F(y))\| \leq L_2\, \|x - y\| .$$

By the same way, we obtain

$$\|\nabla_b \Psi_{\theta p}(x, F(x)) - \nabla_b \Psi_{\theta p}(y, F(y))\| \leq L_2\, \|x - y\| .$$

Then, we have

$$\begin{aligned}
& \|\nabla \Psi_{\theta p}(x) - \nabla \Psi_{\theta p}(y)\| \\
=\ & \|\nabla_a \Psi_{\theta p}(x, F(x)) + \nabla F(x) \nabla_b \Psi_{\theta p}(x, F(x)) - \nabla_a \Psi_{\theta p}(y, F(y)) - \nabla F(y) \nabla_b \Psi_{\theta p}(y, F(y))\| \\
\leq\ & \|\nabla_a \Psi_{\theta p}(x, F(x)) - \nabla_a \Psi_{\theta p}(y, F(y))\| + \|\nabla F(x) \nabla_b \Psi_{\theta p}(x, F(x)) - \nabla F(x) \nabla_b \Psi_{\theta p}(y, F(y))\| \\
& + \|\nabla F(x) \nabla_b \Psi_{\theta p}(y, F(y)) - \nabla F(y) \nabla_b \Psi_{\theta p}(y, F(y))\| \\
\leq\ & \|\nabla_a \Psi_{\theta p}(x, F(x)) - \nabla_a \Psi_{\theta p}(y, F(y))\| + \|\nabla F(x)\|\, \|\nabla_b \Psi_{\theta p}(x, F(x)) - \nabla_b \Psi_{\theta p}(y, F(y))\| \\
& + \|\nabla_b \Psi_{\theta p}(y, F(y))\|\, \|\nabla F(x) - \nabla F(y)\| \\
\leq\ & L_2\, \|x - y\| + \rho L_2\, \|x - y\| + CL\, \|x - y\| \\
=\ & (L_2 + \rho L_2 + CL)\, \|x - y\| ,
\end{aligned}$$

where the second inequality is from the consistency of the matrix norm and vector norm. Let $\bar{L} := \max\{L_2 + \rho L_2 + CL, 1/2\}$. Then, $\|\nabla \Psi_{\theta p}(x) - \nabla \Psi_{\theta p}(y)\| \leq \bar{L}\|x - y\|$. □

Finally, by using the properties of the function $\Psi_{\theta p}$, we show the linear convergence of Algorithm 2.1, for which the following lemma is helpful.

Lemma 3.7 *Let $\gamma, \eta \in (0,1)$ be chosen in Algorithm 2.1 with $\gamma < \eta$. Then there exist $K \in \mathcal{N}^+$ and $\varrho > 0$, such that the inequalities*

$$\varrho \leq \eta^K, \quad \varrho \leq \lambda, \quad \eta^K\lambda - (\eta^K\rho/2)^2 \geq \varrho\eta^K(\rho+1) + \lambda\varrho, \tag{3.1}$$

$$2\left(2-(2\theta)^{\frac{1}{p}}\right)^2(\varrho - \frac{\bar{L}\gamma^K}{2}) \geq \sigma\gamma^K \tag{3.2}$$

hold.

Proof: Let $K := \max\left\{-\left\lceil \log_{\frac{\gamma}{\eta}}\left(\frac{4\sigma}{2\left(2-(2\theta)^{\frac{1}{p}}\right)^2} + 2\bar{L}\right)\right\rceil, \lceil \log_\eta \frac{2\lambda}{\rho^2}\rceil, \lceil \log_\eta \frac{\lambda}{1+\rho}\rceil, 1\right\}$, then

$$\eta^K \leq \frac{2\lambda}{\rho^2}, \tag{3.3}$$

$$\eta^K \leq \frac{\lambda}{1+\rho}, \tag{3.4}$$

$$\left(\frac{\gamma}{\eta}\right)^K \leq \frac{1}{\frac{4\sigma}{2\left(2-(2\theta)^{\frac{1}{p}}\right)^2} + 2\bar{L}}. \tag{3.5}$$

As $\bar{L} \geq \frac{1}{2}$ from Lemma 3.6, then we have $\frac{4\sigma}{2\left(2-(2\theta)^{\frac{1}{p}}\right)^2} + 2\bar{L} > 1$. Set $\varrho := \frac{\eta^K}{4} > 0$. We show that K and ϱ satisfy the inequalities (3.1) and (3.2).

First, we prove (3.1) holds. We have that

$$\begin{aligned} \varrho = \tfrac{\eta^K}{4} \leq \eta^K, \quad \varrho = \tfrac{\eta^K}{4} \leq \tfrac{\lambda}{4(1+\rho)} < \lambda, & \qquad \text{(by (3.4))}; \\ \eta^K\lambda - (\eta^K\rho/2)^2 \geq \tfrac{1}{2}\eta^K\lambda, & \qquad \text{(by (3.3))}; \\ \varrho\eta^K(\rho+1) + \lambda\varrho \leq 2\lambda\varrho = \tfrac{1}{2}\eta^K\lambda, & \qquad \text{(by (3.3) and } \varrho = \tfrac{\eta^K}{4}), \end{aligned}$$

that is, (3.1) holds.

By a direct calculation and noting that $\varrho = \frac{\eta^K}{4}$, it follows that (3.5) implies (3.2). □

Theorem 3.1 *Let $\Psi_{\theta p}$ be defined by (1.1) with $p \geq 2$ and $1 > \theta \geq 0$. Suppose F is continuously differentiable and strongly monotone with modulus $\lambda > 0$. Let $x^0 \in \Re^n$ be any given starting point, and suppose that F and ∇F are Lipschitz continuous with some constant $L > 0$ on $\mathcal{B}(x^0)$. Then, for the sequence $\{x^k\}$ generated by Algorithm 2.1, it holds that the sequence $\{\Psi_{\theta p}(x^k)\}$ converges to zero Q-linearly, and $\{x^k\}$ converges R-linearly to the solution of NCP.*

Proof: For the sequence $\{x^k\}$ generated by Algorithm 2.1, the sequence $\{\Psi_{\theta p}(x^k)\}$ is non-increasing from (2.1). Hence $\{x^k\}$ is contained in $\mathcal{L}(\Psi_{\theta p}, \Psi_{\theta p}(x^0))$. For $\gamma^K \in (0,1)$, where

K is defined in Lemma 3.7, we have that $x^k, x^k + \gamma^K d^k(\eta^K) \in \mathcal{B}(x^0)$, and

$$
\begin{aligned}
& \Psi_{\theta p}(x^k + \gamma^K d^k) - \Psi_{\theta p}(x^k) \\
= \ & \int_0^{\gamma^K} \left\langle \nabla \Psi_{\theta p}(x^k + \mu d^k), d^k \right\rangle d\mu \\
= \ & \gamma^K \left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle + \int_0^{\gamma^K} \left\langle \nabla \Psi_{\theta p}(x^k + \mu d^k) - \nabla \Psi_{\theta p}(x^k), d^k \right\rangle d\mu \\
\leq \ & \gamma^K \left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle + \int_0^{\gamma^K} \left\langle \bar{L}\mu d^k, d^k \right\rangle d\mu \\
= \ & \gamma^K \left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle + \bar{L} \int_0^{\gamma^K} \mu \|d^k\|^2 d\mu \\
= \ & \gamma^K \left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle + \frac{\bar{L}\gamma^{2K}}{2} \|d^k\|^2.
\end{aligned}
$$

Hence, we have

$$
\Psi_{\theta p}(x^k) - \Psi_{\theta p}(x^k + \gamma^K d^k) \ \geq \ -\gamma^K \langle \nabla \Psi_{\theta p}(x^k), d^k \rangle - \frac{\bar{L}\gamma^{2K}}{2} \|d^k\|^2. \tag{3.6}
$$

In the following, we give estimations of two items in the right-hand side of (3.6).

Firstly, we estimate the item $-\gamma^K \langle \nabla \Psi_{\theta p}(x^k), d^k \rangle$ in the right-hand side of the inequality (3.6). It follows that

$$
\begin{aligned}
-\left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle \ = \ & -\Big\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)) + \nabla F(x^k) \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \ , \\
& -\nabla_b \Psi_{\theta p}(x^k, F(x^k)) - \eta^K \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
= \ & \Big\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)) + \nabla F(x^k) \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \ , \\
& \nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \eta^K \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
= \ & \Big\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
& + \eta^K \Big\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
& + \Big\langle \nabla F(x^k) \nabla_b \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
& + \eta^K \Big\langle \nabla F(x^k) \nabla_b \Psi_{\theta p}(x^k, F(x^k)), \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
\geq \ & \eta^K \Big\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
& + \Big\langle \nabla F(x^k) \nabla_b \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
& + \eta^K \Big\langle \nabla F(x^k) \nabla_b \Psi_{\theta p}(x^k, F(x^k)), \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\rangle \\
\geq \ & \eta^K \Big\| \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\|^2 + \lambda \Big\| \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \Big\|^2 \\
& - \eta^K \rho \Big\| \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \Big\| \Big\| \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \Big\|,
\end{aligned} \tag{3.7}
$$

where the first inequality follows from $\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \rangle \geq 0$ by Lemma 2.1(ii), and the second inequality follows from the strong monotonicity of F by Definition 2.1 (ii) and Cauchy-Schwarz inequality.

Below we show that the following inequality

$$\begin{aligned}\eta^K \left\|\nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2 \quad &+ \quad \lambda \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k))\right\|^2 - \eta^K \rho \left\|\nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\| \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k))\right\| \\ &\geq \quad \varrho \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2 \qquad (3.8)\end{aligned}$$

holds, where $\varrho > 0$ set in Lemma 3.7. By the Cauchy-Schwarz inequality, it is sufficient to show that

$$\begin{aligned}(\eta^K - \varrho) \left\|\nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2 + (\lambda - \varrho) \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k))\right\|^2 \\ -(\eta^K \rho + 2\varrho) \left\|\nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\| \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k))\right\| \geq 0.\end{aligned}$$

This above inequality holds if and only if

$$\eta^K - \varrho \geq 0, \lambda - \varrho \geq 0, \quad \text{and} \quad \triangle = (\eta^K \rho + 2\varrho)^2 - 4(\eta^K - \varrho)(\lambda - \varrho) \leq 0,$$

that is,

$$\varrho \leq \eta^K, \varrho \leq \lambda \quad \text{and} \quad \eta^K \lambda - (\eta^K \rho/2)^2 \geq \varrho \eta^K (\rho + 1) + \lambda \varrho. \qquad (3.9)$$

From Lemma 3.7, we have (3.9) holds, then the inequality (3.8) holds. Combining (3.7) and (3.8), we obtain

$$-\left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle \geq \varrho \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2. \qquad (3.10)$$

Secondly, we estimate the item $-\frac{\bar{L}\gamma^{K2}}{2}\|d^k\|^2$ in the right-hand side of the inequality (3.6). It follows that

$$\begin{aligned}-\left\|d^k\right\|^2 \quad &= \quad -\left\|-\nabla_b \Psi_{\theta p}(x^k, F(x^k)) - \eta^K \nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2 \\ &= \quad -\left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k))\right\|^2 - 2\eta^K \left\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \right\rangle \\ &\quad\quad -\eta^{2K} \left\|\frac{\partial \Psi_{\theta p}}{\partial a}(x^k, F(x^k))\right\|^2 \\ &\geq \quad -\left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k))\right\|^2 - 2 \left\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \right\rangle \\ &\quad\quad - \left\|\nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2 \\ &= \quad -\left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2. \qquad (3.11)\end{aligned}$$

where the inequality follows from $\eta^K \in (0, 1)$, and $\left\langle \nabla_a \Psi_{\theta p}(x^k, F(x^k)), \nabla_b \Psi_{\theta p}(x^k, F(x^k)) \right\rangle \geq 0$ from Lemma 2.1(ii).

Now, by combining (3.10) with (3.11), we obtain that

$$\begin{aligned}\Psi_{\theta p}(x^k) - \Psi_{\theta p}(x^k + \gamma^K d^k) \quad &\geq \quad -\gamma^K \left\langle \nabla \Psi_{\theta p}(x^k), d^k \right\rangle - \frac{\bar{L}\gamma^{K2}}{2} \left\|d^k\right\|^2 \\ &\geq \quad \gamma^K \varrho \left\|\nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \nabla_a \Psi_{\theta p}(x^k, F(x^k))\right\|^2\end{aligned}$$

$$
\begin{aligned}
& -\frac{\bar{L}\gamma^{2K}}{2} \left\| \nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \right\|^2 \\
= \ & \gamma^K (\varrho - \frac{\bar{L}\gamma^K}{2}) \left\| \nabla_b \Psi_{\theta p}(x^k, F(x^k)) + \nabla_a \Psi_{\theta p}(x^k, F(x^k)) \right\|^2 \\
\geq \ & 2\left(2 - (2\theta)^{\frac{1}{p}}\right)^2 \gamma^K (\varrho - \frac{\bar{L}\gamma^K}{2}) \Psi_{\theta p}(x^k). \qquad (3.12)
\end{aligned}
$$

the last inequality is from Proposition 3.1 and $\varrho - \frac{\bar{L}\gamma^K}{2} > 0$. In fact, from Lemma 3.7, we have $2\left(2 - (2\theta)^{\frac{1}{p}}\right)^2 (\varrho - \frac{\bar{L}\gamma^K}{2}) \geq \sigma\gamma^K > 0$, so $\varrho - \frac{\bar{L}\gamma^K}{2} > 0$.

It follows from the inequality (3.12) that Step 2 of Algorithm 2.1 is satisfied whenever

$$2\left(2 - (2\theta)^{\frac{1}{p}}\right)^2 \gamma^K (\varrho - \frac{\bar{L}\gamma^K}{2}) \geq \sigma\gamma^{2K},$$

i.e.,

$$2\left(2 - (2\theta)^{\frac{1}{p}}\right)^2 \gamma^K (\varrho - \frac{\bar{L}\gamma^K}{2}) \geq \sigma\gamma^{2K}.$$

From Lemma 3.7, we have the above inequality holds. That is, Step 2 of Algorithm 2.1 is satisfied for $t_k = K$.

Denote $\bar{\eta} := \eta^K, \bar{\gamma} := \gamma^K$, then it follows that $x^{k+1} = x^k + \gamma^{t_k} d^k$ with $\gamma^{t_k} \geq \bar{\gamma} > 0$. Thus, we obtain that

$$\Psi_{\theta p}(x^k) - \Psi_{\theta p}(x^{k+1}) \quad \geq \quad \sigma\bar{\gamma}^2 \Psi_{\theta p}(x^k).$$

This implies

$$\left(1 - \sigma\bar{\gamma}^2\right) \Psi_{\theta p}(x^k) \quad \geq \quad \Psi_{\theta p}(x^{k+1}) \geq 0,$$

which means that $\{\Psi_{\theta p}(x^k)\}$ converges Q-linearly to zero.

From Theorem 2.1 and Lemma 3.2, we can further get

$$\|x^k - x^*\| \leq \frac{L+1}{\sqrt{2}\lambda} \Psi_{NR}(x^k)^{\frac{1}{2}} \leq \frac{\sqrt{2}(L+1)}{(2 - (2\theta)^{\frac{1}{p}})\lambda} \Psi_{\theta p}(x^k)^{\frac{1}{2}}.$$

Since the sequence $\{\Psi_{\theta p}(x^k)\}$ converges Q-linearly to zero, the sequence $\{x^k\}$ converges R-linearly to the solution x^* of the NCP . □

4 Numerical results

Some numerical results of Algorithm 2.1 for complementarity problems from MCPLIB[1] are reported in [12]. From the results, we see that Algorithm 2.1 works well for the tested problem in MCPLIB[1]. In this section, we implement Algorithm 2.1 in MATLAB 7.11 for some complementarity problems from MCPLIB[1] to observe the convergence of Algorithm 2.1. All numerical experiments are done at a PC with CPU of 2.4 GHz and RAM of 256 MB. The algorithm was terminated whenever one of the following conditions was satisfied:

(1) $\Psi_{\theta p}(x^k) \le 10^{-9}$ and $d \le 10^{-3}$;

(2) the steplength γ^{t_k} is less than 10^{-9};

(3) the number of iterations is more than 100000.

We tests some problems for two purse: one is to investigate the convergence of Algorithm 2.1 with different p; the other is to investigate the convergence of Algorithm 2.1 with different θ.

Firstly, we take "gafni(1)" for example with different p. The parameters are chosen as follows:

$$\eta = 0.8,\ \sigma = 0.5,\ \gamma = 0.6 \quad \text{and} \quad \theta = 0.5.$$

Figure 1, 2, 3 describe the detail iteration process of Algorithm 2.1 with $p = 1.1$, $p = 10$, $p = 100$, respectively.

Secondly, we take "josephy(5)" for example with different θ. The parameters are chosen as follows:

$$\eta = 0.8,\ \sigma = 0.5,\ \gamma = 0.6 \quad \text{and} \quad p = 10.$$

Figure 4, 5,6 describe the detail iteration process of Algorithm 2.1 with $\theta = 0$, $\theta = 0.5$ and $\theta = 1$, respectively.

Some interesting phenomenon can be obtained from Figures 1-6:

- For the tested problems, the sequence $\{\Psi_{\theta p}(x^k)\}$, got from Algorithm 2.1, converges Q-linearly to zero. This phenomenon indeed verified the conclusions of Theorem 3.1. We also test other problems form MCPLIB[1], the convergent behaviors of Algorithm 2.1 are almost the same.
- From Figure 1-3 we may see that the merit function $\Psi_{\theta p}$ in case when $p = 100$ has a faster decrease than the case when $p = 10$; $\Psi_{\theta p}$ in case when $p = 10$ has a faster decrease than the case when $p = 1.1$. We may get a conclusion that the convergence rate of Algorithm 2.1 becomes better when p increases.
- From Figure 4-6 we may see that the merit function $\Psi_{\theta p}$ in case when $\theta = 0$ has a faster decrease than the case when $\theta = 0.5$; $\Psi_{\theta p}$ in case when $\theta = 0.5$ has a faster decrease than the case when $\theta = 1$. We may get a conclusion that the convergence rate of Algorithm 2.1 becomes worse when θ increases.

5 Some Final Remarks

In this paper, we obtained several favorite properties of the merit function $\Psi_{\theta p}$, including the growth behavior of the merit function $\Psi_{\theta p}$, the estimation on the upper bound of $\Psi_{\theta p}$, and the Lipschitz continuity of the gradient function $\nabla\Psi_{\theta p}$. In particular, we showed that the iterative point

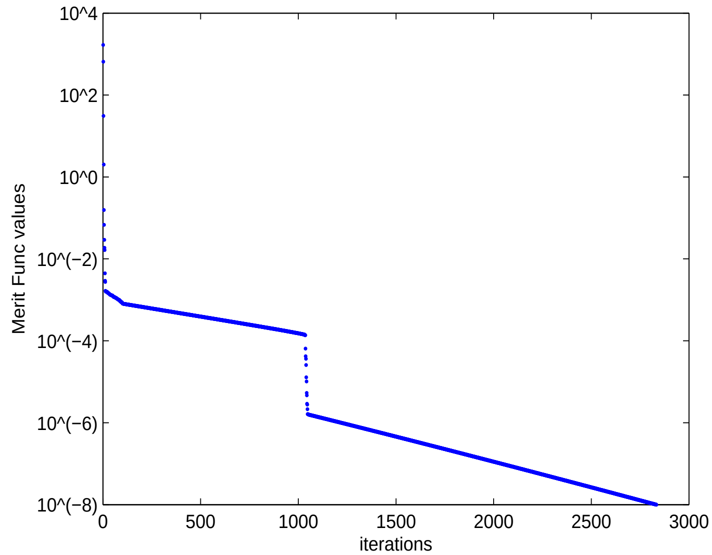

Figure 1: Convergence behavior of “gafni(1)” with $p = 1.1$.

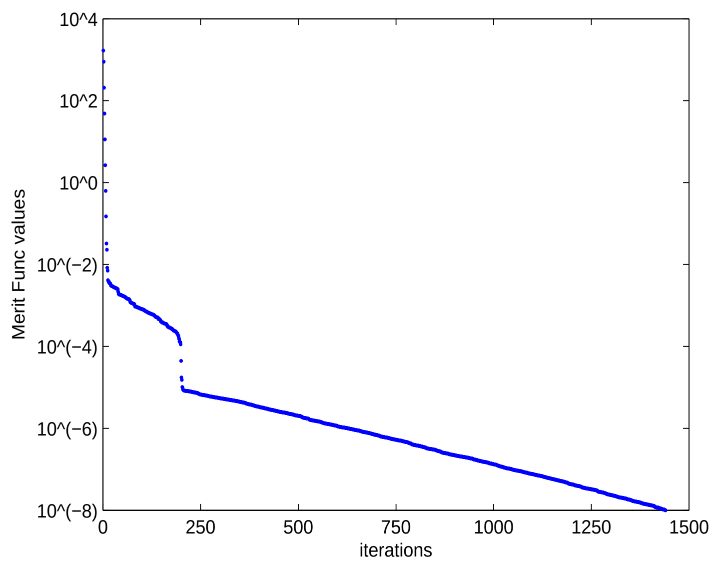

Figure 2: Convergence behavior of “gafni(1)” with $p = 10$.

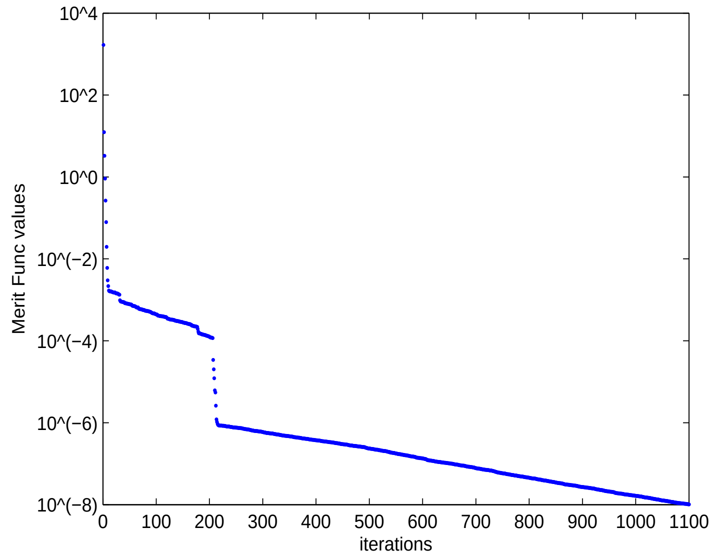

Figure 3: Convergence behavior of “gafni(1)” with $p = 100$.

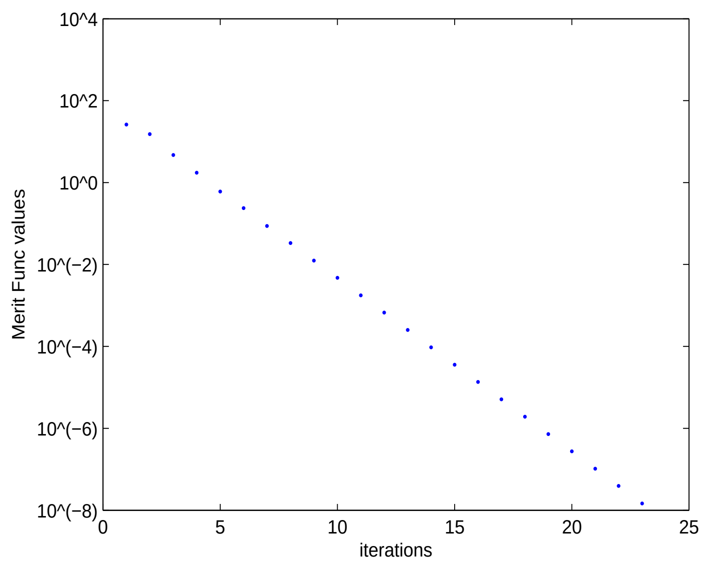

Figure 4: Convergence behavior of “josephy(5)” with $\theta = 0$.

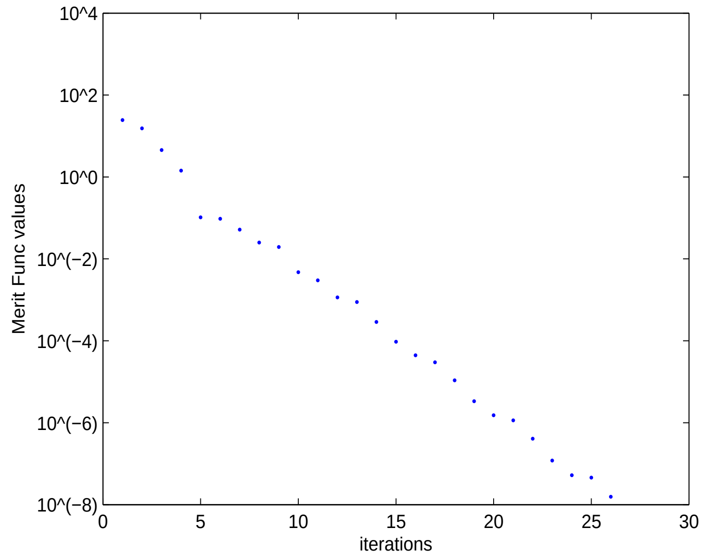

Figure 5: Convergence behavior of "josephy(5)" with $\theta = 0.5$.

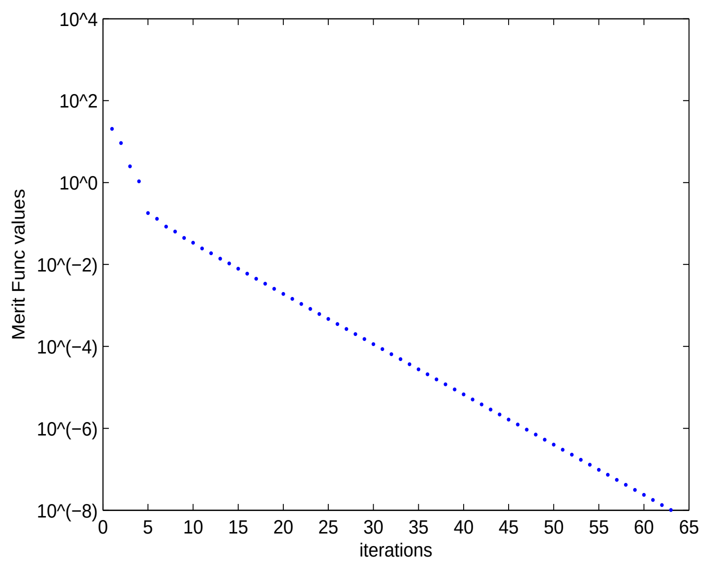

Figure 6: Convergence behavior of "josephy(5)" with $\theta = 1$.

sequence $\{x^k\}$ generated by Algorithm 2.1 is globally R-linearly convergent and the corresponding merit function sequence $\{\Psi_{\theta p}(x^k)\}$ is globally Q-linearly convergent under suitable assumptions.

It is interesting whether the NCP-function $\phi_{\theta p}$ and the merit function $\Psi_{\theta p}$ can be extended to the case of the symmetric cone or not. If yes, whether the properties of these functions and the convergence results of the derivative-free descent method, obtained in [12] and this paper, are still satisfied. These are our subjects of future research.

References

[1] S.C. Billups, S.P. Drikse and M.C. Soares, A comparison of algorithm for large scale mixed complementartiy problems, Comput. Optim. Appl. 7 (1977): 3–25.

[2] J.-S. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, J. Global Optim. 36 (2006): 565–580.

[3] J.-S. Chen, Z.H. Huang, C.-Y. She, A new class of penalized NCP-functions and its properties, Comput. Optim. Appl. 50 (2011): 49–73.

[4] J.-S. Chen, S.H. Pan, A regularization semismooth Newton method based on the generalized Fischer-Burmeister function for P_0-NCPs, J. Comput. Appl. Math. 220 (2008): 464–479.

[5] J.-S. Chen, S.H. Pan, A family of NCP-functions and a descent method for the nonlinear complementarity problem, Comput. Optim. Appl. 40 (2008): 389–404.

[6] J.-S. Chen, H.-T. Gao and S.H. Pan, An R-linearly convergent derivative-free algorithm for nonlinear complementarity problems based on the generalized Fisher-Burmeister merit function, Comput. Optim. Appl. 232 (2009): 455–471.

[7] F. Facchinei and J.S. Pang , Finite-dimensional variational inequalities and complementarity problems. Springer Verlag, New York. 2003.

[8] F. Facchinei and J. Soares , A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. 7 (1997): 225–247.

[9] M.C. Ferris, J.S. Pang, Engineering and economic applications of complementarity problems, SIAM Rev. 39 (1997): 669–713.

[10] C. Geiger and C. Kanzow, On the resolution of monotone complementarity problems. Comput. Optim. Appl. 5 (1996): 155–173.

[11] P.T. Harker, J.S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications, Math. Program. 48 (1990): 161–220.

[12] S.L. Hu, Z.H. Huang, J.-S. Chen, Properties of a family of generalized NCP-functions and a derivative free algorithm for complementarity problems, J. Comput. Appl. Math. 230 (2009): 69–82.

[13] S.L. Hu, Z.H. Huang, N. Lu, Smoothness of a class of generalized merit functions for the second-order cone complementarity problem, Pacific J. Optim. 6 (2010): 551–571.

[14] Z.H. Huang, The global linear and local quadratic convergence of a non-interior continuation algorithm for the LCP, IMA J. Numer. Anal. 25 (2005): 670–684.

[15] Z.H. Huang, L. Qi, D. Sun, Sub-quadratic convergence of a smoothing Newton algorithm for the P_0- and monotone LCP, Math. Program. 99 (2004): 423–441.

[16] H.Y. Jiang, M. Fukushima, L. Qi, et al., A trust region method for solving generalized complementarity problems, SIAM. J. Optim. 8 (1998) 140–157.

[17] C. Kanzow, H. Kleinmichel, A new class of semismooth Newton method for nonlinear complementarity problems, Comput. Optim. Appl. 11 (1998): 227–251.

[18] L.Y. Lu, Z.H. Huang, S.L. Hu, Properties of a family of merit functions and a merit function method for the NCP, Appl. Math. – J. Chinese Univ. 25 (2010): 379–390.

[19] Z.Q. Luo, P. Tseng, A new class of merit functions for the nonlinear complementarity problem, In Complementarity and Variational Problems: State of the Art, Edited by M. C. Ferris and J.-S. Pang, SIAM, Philadelphia, 204–225, 1997.

[20] O.L. Mangasarian, M.V. Solodov, A linearly convergent derivative-free descent method for strongly monotone complementarity problems, Comput. Optim. Appl. 14 (1999): 5–16.

[21] J.S. Pang, A posteriori error bounds for the linearly-constrained variational inequality problem, Math. Oper. Res. 12 (1987): 474–484.

[22] J.M. Ortega, W. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, SIAM, Philadelphia, 2000.

[23] L. Qi, D. Sun, G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequality problems, Math. Program. 87 (2000): 1–35.

[24] K. Yamada, N. Yamashita, M. Fukushima, A new derivative-free descent method for the nonlinear complementarity problem. In: Pillo, G.D., Giannessi, F. (eds.) Nonlinear Optimization and Related Topics, pp. 463–489, 2000. Kluwer Academic, Dordrecht.