JOURNAL OF INDUSTRIAL AND
MANAGEMENT OPTIMIZATION
Volume **9**, Number **1**, January **2013**

doi:10.3934/jimo.2013.9.153

pp. **153–169**

PROXIMAL POINT ALGORITHM FOR NONLINEAR COMPLEMENTARITY PROBLEM BASED ON THE GENERALIZED FISCHER-BURMEISTER MERIT FUNCTION

Yu-Lin Chang and Jein-Shan Chen

Department of Mathematics
National Taiwan Normal University
Taipei 11677, Taiwan

Jia Wu

School of Mathematical Sciences
Dalian University of Technology
Dalian 116024, China

(Communicated by Liqun Qi)

Abstract. This paper is devoted to the study of the proximal point algorithm for solving monotone and nonmonotone nonlinear complementarity problems. The proximal point algorithm is to generate a sequence by solving subproblems that are regularizations of the original problem. After given an appropriate criterion for approximate solutions of subproblems by adopting a merit function, the proximal point algorithm is verified to have global and superlinear convergence properties. For the purpose of solving the subproblems efficiently, we introduce a generalized Newton method and show that only one Newton step is eventually needed to obtain a desired approximate solution that approximately satisfies the appropriate criterion under mild conditions. The motivations of this paper are twofold. One is analyzing the proximal point algorithm based on the generalized Fischer-Burmeister function which includes the Fischer-Burmeister function as special case, another one is trying to see if there are relativistic change on numerical performance when we adjust the parameter in the generalized Fischer-Burmeister.

1. **Introduction.** In the last decades, people have put a lot of their energy and attention on the complementarity problem due to its various applications in operation research, economics, and engineering, see [16, 18, 30] and references therein. The nonlinear complementarity problem (NCP) is to find a point $x \in \mathbb{R}^n$ such that

$$\mathrm{NCP}(F): \ \langle F(x), x\rangle = 0, \ F(x) \in \mathbb{R}^n_+, \ x \in \mathbb{R}^n_+, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable mapping with $F := (F_1, F_2, \ldots, F_n)^T$. Many solution methods have been developed to solve NCP(F) [3, 4, 18, 19, 20, 21, 22, 24, 30, 40, 41]. For more details, please refers to the excellent monograph [14]. One of the most popular methods is to reformulate the NCP(F) as

2010 *Mathematics Subject Classification.* 49K30, 65K05, 90C33.

Key words and phrases. Complementarity problem, proximal point algorithm, approximation criterion.

Corresponding author: J.-S. Chen, Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. Research of J. Chen is supported by National Science Council of Taiwan.

a unconstrained optimization problem and then to solve the reformulated problem by the unconstrained optimization technique. This kind of methods is called the merit function approach, where the merit functions are usually constructed via some NCP-functions.

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

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

Furthermore, if $\phi(a, b) \geq 0$ for all $(a, b) \in \mathbb{R}^2$ then the NCP-function ϕ is called a nonnegative NCP-function. In addition, if a function $\Psi : \mathbb{R}^n \to \mathbb{R}_+$ satisfying $\Psi(x) = 0$ if and only if x solves the NCP, then Ψ is called a merit function for the NCP. How to construct a merit function via an NCP-function? If ϕ is an NCP-function, an easy way to construct a merit function is defining $\Psi : \mathbb{R}^n \to \mathbb{R}_+$ by

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

With this merit function, finding a solution of the NCP is equivalent to seeking a global minimum of the unconstrained minimization problem

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

with optimal value zero. In fact, many NCP-functions have been proposed in the literature. Among them, the Fischer-Burmeister (FB) function is one of the most popular NCP-functions, which is defined as

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

One generalization of FB function was given by Kanzow and Kleinmichel in [23]:

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

Another generalization proposed by Chen [3, 4, 5] and called generalized Fischer-Burmeister function is defined as

$$\phi_p(a, b) := a + b - \sqrt[p]{|a|^p + |b|^p}, \quad p \in (1, \infty), \quad \forall (a, b) \in \mathbb{R}^2. \tag{3}$$

Among all various methods for solving the NCP, we focus on the proximal point algorithm (PPA) in this paper. The PPA is known for its theoretically nice convergence properties, which was first proposed by Martinet [27] and further studied by Rockafellar [35], and was originally designed for finding a vector z satisfying $0 \in \mathcal{T}(z)$ where $\mathcal{T}$ is a maximal monotone operator. Therefore, it can be applied to a broad class of problems such as convex programming problems, monotone variational inequality problems, and monotone complementarity problems. In general, PPA generates a sequence by solving subproblems that are regularizations of the original problem. More specifically, for the case of monotone NCP(F), given the current point x^k, PPA obtains the next point x^{k+1} by approximately solving the subproblem

$$\text{NCP}(F^k) : \ \langle F^k(x), x\rangle = 0,\ F^k(x) \in \mathbb{R}^n_+,\ x \in \mathbb{R}^n_+, \tag{4}$$

where $F^k : \mathbb{R}^n \to \mathbb{R}^n$ is defined by

$$F^k(x) := F(x) + c_k(x - x^k) \quad \text{with} \quad c_k > 0. \tag{5}$$

It is obvious that F^k is strongly monotone when F is monotone. Then, by [14, Theorem 2.3.3], the subproblem NCP(F^k), which is more tractable than the original problem, always has a unique solution. Thus, PPA is well-defined. It was

pointed out in [26, 35] that with appropriate criteria for approximate solutions of subproblems (4), PPA has global and superlinear convergence property under mild conditions. Another implementation issue is how to solve subproblems efficiently and obtain an approximate solution such that the approximation criterion for the subproblem is fulfilled. A generalized Newton method proposed by De Luca et al. [25] which is used to solve subproblems. The approximation criterion under given conditions is eventually approximately fulfilled by a single Newton iteration of the generalized Newton method. As for the case of nonmonotone NCP, similar idea was also employed for solving P_0-NCP in [43].

In this paper, we look into again the PPA for monotone NCP and P_0-NCP studied in [42] and [43], respectively. We consider the PPA based on different NCP-functions, indeed based on the class of ϕ_p as in (3). Analogous theoretical analysis for PPA based on ϕ_p can be established easily. However, this is not the main purpose of doing such extension. In fact, there have been reported [3, 4, 6, 7] that changing the value of p has various influence on numerical performance for different types of algorithms. It is our intension to know whether such phenomenon occurs when employing PPA method for solving NCPs. As will be seen in Section 5, such phenomenon does not appear in PPA, more specifically, changing the parameter p does change the numerical performance for the proposed PPA, however, such change does not depend on p regularly. This together with earlier reports offer us a further understanding about the generalized Fischer-Burmeister function.

2. Preliminaries.

In this section, we review some background materials that will be used in the sequel and briefly introduce the proximal point algorithm.

2.1. Mathematical concepts.

Given a set $\Omega \in \mathbb{R}^n$ locally closed around $\bar{x} \in \Omega$, the regular normal cone to Ω at $\bar{x}$ is defined as

$$\widehat{\mathcal{N}}_\Omega(\bar{x}) := \Big\{ v \in \mathbb{R}^n \mid \limsup_{\substack{\Omega \\ x \to \bar{x}}} \frac{\langle v, x - \bar{x} \rangle}{||x - \bar{x}||} \leq 0 \Big\}.$$

The (limiting) normal cone to Ω at $\bar{x}$ is set to be

$$\mathcal{N}_\Omega(\bar{x}) := \limsup_{\substack{\Omega \\ x \to \bar{x}}} \widehat{\mathcal{N}}_\Omega(x),$$

where "limsup" is the Painlevé-Kuratowski outer limit of sets, see [36]. If Ω is the nonegative orthant $\mathbb{R}^n_+$, the normal cone to Ω at $\bar{x}$ is

$$\mathcal{N}_{\mathbb{R}^n_+}(\bar{x}) = \begin{cases} \{y \in \mathbb{R}^n | \langle \bar{x} - z, y \rangle \geq 0, \forall z \geq 0\} & \text{if } \bar{x} \geq 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

We now recall definitions of various monotonicity and P-properties of a mapping which are needed for subsequent analysis.

Definition 2.1. Let $F : \mathbb{R}^n \to \mathbb{R}^n$.

- **(a):** F is said to be monotone if $(x - y)^T (F(x) - F(y)) \geq 0$ for all $x, y \in \mathbb{R}^n$.
- **(b):** F is said to be strongly monotone with modulus $\mu > 0$ if $(x - y)^T (F(x) - F(y)) \geq \mu \|x - y\|^2$ for all $x, y \in \mathbb{R}^n$.
- **(c):** F is said to be an P_0-function if for all $x, y \in \mathbb{R}^n$ with $x \neq y$ there is an index i such that $x_i \neq y_i$ and $(x_i - y_i)[F_i(x) - F_i(y)] \geq 0$.
- **(d):** F is said to be an P-function if for all $x, y \in \mathbb{R}^n$ with $x \neq y$ there is an index i such that $(x_i - y_i)[F_i(x) - F_i(y)] > 0$.

It is well known that, when F is continuously differentiable, F is monotone if and only if $\nabla F(\zeta)$ is positive semidefinite for all $\zeta \in \mathbb{R}^n$ while F is strongly monotone if and only if $\nabla F(\zeta)$ is positive definite for all $\zeta \in \mathbb{R}^n$. For more details about monotonicity, please refer to [14]. In addition. it can be easily verified that if F is P_0-function, then the function F^k defined by (5) is P-function and the Jacobian matrices of F^k are P-matrices. F is said to be uniformly Lipschitz continuous on a set Ω with modulus $\kappa > 0$ if $\|F(x) - F(y)\| \leq \kappa\|x - y\|$ for all $x, y \in \Omega$.

Moreover, for a vector-valued Lipschitz continuous mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, the B-subdifferential of F at x denoted by $\partial_B F(x)$ is defined as

$$\partial_B F(x) := \left\{ \lim_{k\to\infty} \mathcal{J}F(x^k) \,\middle|\, x^k \to x,\ F \text{ is differentiable at } x^k \right\}.$$

The convex hull of $\partial_B F(x)$ is the Clarke's generalized Jacobian of F at x denoted by $\partial F(x)$, see [10]. We say that F is *strongly BD-regular* at x if every element of $\partial_B F(x)$ is nonsingular.

We need another important concept named semismoothness which was first introduced in [28] for functionals and was further extended in [33] to vector-valued functions. Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz continuous mapping. We say that F is *semismooth* at a point $x \in \mathbb{R}^n$ if F is directionally differentiable and for any $\Delta x \in \mathbb{R}^n$ and $V \in \partial F(x + \Delta x)$ with $\Delta x \to 0$, there has

$$F(x + \Delta x) - F(x) - V(\Delta x) = \mathrm{o}(||\Delta x||).$$

Furthermore, F is said to be *strongly semismooth* at x if F is semismooth at x and for any $\Delta x \in \mathbb{R}^n$ and $V \in \partial F(x + \Delta x)$ with $\Delta x \to 0$, there holds

$$F(x + \Delta x) - F(x) - V(\Delta x) = \mathrm{O}(||\Delta x||^2).$$

To close this subsection, we introduce the R-regularity of a solution $\bar{x}$ to NCP. For a solution $\bar{x} \in \mathbb{R}^n$ to the NCP(F), we define the following three index sets which are associated with $\bar{x}$:

$$\begin{aligned} \alpha &:= \{\, i \mid \bar{x}_i > 0, F_i(\bar{x}) = 0\}, \\ \beta &:= \{\, i \mid \bar{x}_i = F_i(\bar{x}) = 0\}, \\ \gamma &:= \{\, i \mid \bar{x}_i = 0, F_i(\bar{x}) > 0\}. \end{aligned}$$

We say that the solution $\bar{x}$ is R-regular [34] if $\nabla F_{\alpha\alpha}(\bar{x})$ is nonsingular and the Schur complement of $\nabla F_{\alpha\alpha}(\bar{x})$ in $\begin{bmatrix} \nabla F_{\alpha\alpha}(\bar{x}) & \nabla F_{\alpha\beta}(\bar{x}) \\ \nabla F_{\beta\alpha}(\bar{x}) & \nabla F_{\beta\beta}(\bar{x}) \end{bmatrix}$ is an P-matrix.

2.2. **Complementarity and merit functions.** Back to the generalized FB function ϕ_p, it has been proved in [3, 4, 5, 6] that the function ϕ_p given in (3) possess a system of favorite properties, such as strong semismoothness, Lipschitz continuity, and continuous differentiability except for the point $(0, 0)$. Due to these favorable properties, given a certain mapping $\hat{F} : \mathbb{R}^n \to \mathbb{R}^n$, the NCP($\hat{F}$) can be reformulated as the following nonsmooth system of equations:

$$\Phi_p(x) := \begin{pmatrix} \phi_p(x_1, \hat{F}_1(x)) \\ \vdots \\ \phi_p(x_i, \hat{F}_i(x)) \\ \vdots \\ \phi_p(x_n, \hat{F}_n(x)) \end{pmatrix} = 0. \tag{6}$$

It was also shown that the squared norm ϕ_p which is given by

$$\psi_p(a,b) := |\phi_p(a,b)|^2 \tag{7}$$

is continuously differentiable. Moreover, the merit functions Ψ_p induced from ψ_p

$$\Psi_p(x) := \|\Phi_p(x)\|^2 = \sum_{i=1}^{n} |\phi_p(x, \hat{F}_i(x))|^2 = \sum_{i=1}^{n} \psi_p(x_i, \hat{F}_i(x)) \tag{8}$$

possess SC^1 property (i.e., they are continuously differentiable and their gradients are semismooth) and LC^1 property (i.e., they are continuously differentiable and their gradients are Lipschitz continuous) under suitable assumptions. Note that $\Phi_p(x)$ is not differentiable at x when $x_i = \hat{F}_i(x) = 0$ for some $1 \le i \le n$. In addition, in light of [8, Lemma 3.1], we know that an element V of $\partial_B \Phi_p(x)$ can be expressed as

$$V = D_a + D_b \nabla \hat{F}(x)^T, \tag{9}$$

where D_a and D_b are diagonal matrices given by

$$\begin{aligned} &((D_a)_{ii}, (D_b)_{ii}) \\ = &\begin{cases} \left(1 - \frac{\mathrm{sgn}(x_i)|x_i|^{p-1}}{\|(x_i, \hat{F}_i(x))\|_p^{p-1}}, 1 - \frac{\mathrm{sgn}(\hat{F}_i(x))|\hat{F}_i(x)|^{p-1}}{\|(x_i, \hat{F}_i(x))\|_p^{p-1}}\right) & \text{if } (x_i, \hat{F}_i(x)) \neq (0,0), \\ (1-\eta, 1-\xi) & \text{otherwise,} \end{cases} \end{aligned} \tag{10}$$

with $\|(x_i, \hat{F}_i(x))\|_p^{p-1} = \left(\sqrt[p]{|x_i|^p + |\hat{F}_i(x)|^p}\right)^{p-1}$ and (η, ξ) being a vector satisfying $|\eta|^{\frac{p}{p-1}} + |\xi|^{\frac{p}{p-1}} = 1$.

Lemma 2.2. *Let D_a and D_b be defined as in (10). Then, the following hold.*

(a): *The diagonal matrices D_a and D_b satisfy*

$$(D_a)_{ii} + (D_b)_{ii} \ge 2 - \sqrt[p]{2}, \quad \forall i = 1, 2, \cdots, n.$$

(b): *Suppose that $\hat{F}$ is strongly monotone with modulus μ. Then,*

$$\|(D_a + D_b \nabla \hat{F}(x)^T)^{-1}\| \le \frac{1}{B_1 \mu}$$

where $B_1 = \frac{2 - \sqrt[p]{2}}{n} \max\{1, \|\nabla \hat{F}(x)\|\}$.

Proof. (a) To proceed, we discuss two cases as below.

(i) If $(x_i, \hat{F}_i(x)) \neq (0,0)$, then

$$\begin{aligned} (D_a)_{ii} &= 1 - \frac{\mathrm{sgn}(x_i)|x_i|^{p-1}}{\|(x_i, \hat{F}_i(x))\|_p^{p-1}} \\ (D_a)_{ii} &= 1 - \frac{\mathrm{sgn}(\hat{F}_i(x))|\hat{F}_i(x)|^{p-1}}{\|(x_i, \hat{F}_i(x))\|_p^{p-1}}. \end{aligned}$$

From the arguments of [7, Lemma 3.3], there proves an inequality

$$\frac{|a|^{p-1} + |b|^{p-1}}{\|(a,b)\|_p^{p-1}} \le 2^{\frac{1}{p}} \quad \text{for } p > 1.$$

With this inequality, it is clear that $(D_a)_{ii} + (D_b)_{ii} \ge 2 - \sqrt[p]{2}$ under this case.

(ii) If $(x_i, \hat{F}_i(x)) = (0,0)$, then $((D_a)_{ii}, (D_b)_{ii}) = (1-\eta, 1-\xi)$ where (η, ξ) satisfy $|\eta|^{\frac{p}{p-1}} + |\xi|^{\frac{p}{p-1}} = 1$. To prove the desired result, it is equivalent to showing $\eta + \xi \leq \sqrt[p]{2}$. This can be seen just by applying the Hölder inequality

$$\eta + \xi \leq (1^p + 1^p)^{\frac{1}{p}} (|\eta|^q + |\xi|^q)^{\frac{1}{q}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ (note that $q = \frac{p}{p-1}$). Thus, the proof is complete.

(b) The arguments are similar to [42, Corollary 2.7] which can be obtained by applying part(a) and [42, Proposition 2.6]. □

With the expression of $\partial_B \Phi_p(x)$, it is now possible to provide sufficient conditions for the strongly BD-regularity of Φ_p at a solution of the nonlinear complementarity problem. Let $\bar{x}$ be a solution to NCP($\hat{F}$). The following statements indicate under what conditions Φ_p is strongly BD-regular at $\bar{x}$. These results are important from the algorithmic point of view.

Lemma 2.3. *Let $\bar{x}$ be a solution to NCP($\hat{F}$). Suppose that either $\bar{x}$ is an R-regular solution or $\nabla \hat{F}(\bar{x})$ is an P-matrix. Then every matrix in $\partial_B \Phi_p(\bar{x})$ is nonsingular, i.e., Φ_p is strongly BD-regular at $\bar{x}$.*

Proof. The proofs are similar to those in [15, Propositon 3.2] and [22, Corollary 22, Corollary 23], we omit them here. □

Next, we talk about another merit function $\Psi : \mathbb{R}^n \to \mathbb{R}_+$ which is utilized in the PPA method studied in [42] and provides a more favorable error bound than Ψ_p:

$$\Psi(x) := \sum_{i=1}^{n} \left\{ |x_i \hat{F}_i(x)| + |\phi_{\text{NR}}(x_i, \hat{F}_i(x))| \right\},$$

where the natural residual function $\phi_{\text{NR}} : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$\phi_{\text{NR}}(a,b) := \min\{a, b\}.$$

It is clear that $\Psi(x) \geq 0$, and $\Psi(x) = 0$ if and only if x is a solution of NCP($\hat{F}$). The below lemma shows the error bound property for such function Ψ.

Lemma 2.4. [42, Lemma2.11] *Suppose that $\hat{F}$ is strongly monotone with modulus μ. Then we have*

$$\|x - \hat{x}\| \leq 2\max\{1, \|x\|\} \sqrt{\frac{\Psi(x)}{\mu}} \text{ for all } x \in \left\{ y \in \mathbb{R}^n_+ \mid \Psi(y) \leq \frac{\mu}{4} \right\}$$

where $\hat{x}$ is the unique solution of NCP($\hat{F}$).

Now, we here present some basic properties regarding generalized Fischer-Burmeister function from [7] which we will be used later.

Lemma 2.5. [7, Lemma 3.1, Proposition 3.1] *The functions $\psi_p(a,b)$ and $\Psi_p(x)$ defined by (7) and (8) have the following favorable properties:*

- **(a):** *$\psi_p(a,b)$ is an NCP-function.*
- **(b):** *$\psi_p(a,b)$ is continuously differentiable everywhere.*
- **(c):** *For $p \geq 2$, the gradient of $\psi_p(a,b)$ is Lipschitz continuous on any nonempty bounded set.*
- **(d):** *$\nabla_a \psi_p(a,b) \cdot \nabla_b \psi_p(a,b) \geq 0$ for any $(a,b) \in \mathbb{R}^2$, and the equality holds if and only if $\psi_p(a,b) = 0$.*

(e): $\nabla_a \psi_p(a,b) = 0 \Longleftrightarrow \nabla_b \psi_p(a,b) = 0 \Longleftrightarrow \psi_p(a,b) = 0$.
(f): $(2-\sqrt[p]{2})^2 \Psi_{\text{NR}}(x) \le \Psi_p(x) \le (2+\sqrt[p]{2})^2 \Psi_{\text{NR}}(x)$, *where* $\Psi_{\text{NR}}(x) = \sum_{i=1}^n \phi^2_{\text{NR}}(x_i, \hat{F}_i(x))$.

Consequently, analogous to [42, Lemma 2.8, Lemma 2.9] and [43, Proposition 2.2], the following results can be achieved immediately.

Lemma 2.6. *The mappings* Φ_p *and* Ψ_p *defined in (6) and (8) have the following properties.*

(a): *If* $\hat{F}$ *is continuously differentiable, then* Φ_p *is semismooth.*
(b): *If* $\nabla \hat{F}$ *is locally Lipschitz continuous, then* Φ_p *is strongly semismooth.*
(c): *If* $\hat{F}$ *is continuously differentiable, then* Ψ_p *is continuously differentiable everywhere.*
(d): *If* $\hat{F}$ *is monotone, then any stationary point of* Ψ_p *is a solution of NCP(*$\hat{F}$*).*
(e): *If* $\hat{F}$ *is an* P_0*-function, then every stationary point of* Ψ_p *is a solution of NCP(*$\hat{F}$*).*
(f): *If* $\hat{F}$ *is strongly monotone with modulus* μ *and Lischitz continuous with constant* L*, then* $\sqrt{\Psi_p(x)}$ *provides a global error bound for NCP(*$\hat{F}$*), that is,*

$$\|x - \hat{x}\| \le \frac{B_2(L+1)}{\mu}\sqrt{\Psi_p(x)}, \ \text{for all } x \in \mathbb{R}^n,$$

where $\hat{x}$ *is the unique solution of NCP(*$\hat{F}$*) and* B_2 *is a positive constant independent of* $\hat{F}$*.*
(g): *Let* $S \subset \mathbb{R}^n$ *be a compact set. Suppose* $\hat{F}$ *is strongly monotone with modulus* μ *and Lischitz continuous with constant* L *on* S*. Then* $\sqrt{\Psi_p(x)}$ *provides a global error bound on* S*, that is, there exists a positive constant* B_2 *such that*

$$\|x - \hat{x}\| \le \frac{B_2(L+1)}{\mu}\sqrt{\Psi_p(x)}, \ \text{for all } x \in S,$$

where $\hat{x}$ *is the unique solution of NCP(*$\hat{F}$*).*

Proof. Part(a) and (b) are clear since $\phi_p(a,b)$ is strongly semismooth. Part(c) is implied by Lemma 2.5(b) while Lemma 2.5(d)-(e) yield part(d). Part(e) is from [5, Proposition 3.4]. From Lemma 2.5(f), we have

$$\sqrt{\Psi_{\text{NR}}(x)} \le \frac{1}{2-\sqrt[p]{2}}\sqrt{\Psi_p(x)}.$$

This together with [29, Theorem 3.1] gives

$$\|x - \hat{x}\| \le \frac{(L+1)}{\mu}\sqrt{\Psi_{\text{NR}}(x)} \le \frac{(L+1)}{(2-\sqrt[p]{2})\mu}\sqrt{\Psi_p(x)},$$

where $\Psi_{\text{NR}}(x) = \sum_{i=1}^n \phi^2_{\text{NR}}(x_i, \hat{F}_i(x))$. This completes part(f) and part(g). □

2.3. **Proximal point algorithm.** Let $\mathcal{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping defined by

$$\mathcal{T}(x) := F(x) + \mathcal{N}_{\mathcal{K}}(x). \tag{11}$$

It is known that $\mathcal{T}$ is a maximal monotone mapping and NCP(F) is equivalent to the problem of finding a point x such that $0 \in \mathcal{T}(x)$. Then, for any starting point x^0, PPA generates a sequence $\{x^k\}$ by the approximate rule:

$$x^{k+1} \approx P_k(x^k),$$

where $P_k := \left(I + \frac{1}{c_k}\mathcal{T}\right)^{-1}$ is a vector-valued mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$, $\{c_k\}$ is some sequence of positive real numbers, and $x^{k+1} \approx P_k(x^k)$ means that x^{k+1} is an approximation to $P_k(x^k)$. Accordingly, $P_k(x^k)$ is given by

$$P_k(x^k) = \left(I + \frac{1}{c_k}(F + \mathcal{N}_\mathcal{K})\right)^{-1}(x^k),$$

from which we have

$$P_k(x^k) \in \mathrm{SOL}(\mathrm{NCP}(\mathrm{F^k})),$$

where F^k is defined by (5) and $\mathrm{SOL}(\mathrm{NCP}(F^k))$ denotes the solution set of $\mathrm{NCP}(F^k)$. Therefore, x^{k+1} is obtained by an approximate solution of $\mathrm{NCP}(F^k)$. As remarked in [42], when c_k is small, the subproblem is close to the original one whereas when c_k is large, a solution of the subproblem is expected to lie near x^k. Moreover, to ensure convergence of PPA, x^{k+1} must be located sufficiently near the solution $P_k(x^k)$. Among others, two general criteria for the approximate calculation of $P_k(x^k)$ proposed by Rockafellar [35] are as follows:

Criterion 1.: $\left\|x^{k+1} - P_k(x^k)\right\| \le \varepsilon_k,\ \sum_{k=0}^{\infty} \varepsilon_k < \infty.$

Criterion 2.: $\left\|x^{k+1} - P_k(x^k)\right\| \le \eta_k \|x^{k+1} - x^k\|,\ \sum_{k=0}^{\infty} \eta_k < \infty.$

As mentioned in [26, 35], there says the above Criterion 1 guarantees global convergence while Criterion 2, which is rather stringent, ensures superlinear convergence. The following two theorems elaborate this issue.

Theorem 2.7. [35, Theorem 1] *Let $\{x^k\}$ be any sequence generated by the PPA under Criterion 1 with $\{c_k\}$ bounded. Suppose NCP(F) has at least one solution. Then $\{x^k\}$ converges to a solution x^* of NCP(F).*

Theorem 2.8. [26, Theorem 2.1] *Suppose the solution set $\bar{X}$ of NCP(F) is nonempty, and let $\{x^k\}$ be any sequence generated by PPA with Criteria 1 and Criterion 2 and $c_k \to 0$. Let us also assume that*

$$\exists\, \delta > 0,\ \exists\, C > 0,\ \text{s.t. } \mathrm{dist}(x, \bar{X}) \le C\|\omega\| \text{ whenever } x \in \mathcal{T}^{-1}(\omega) \text{ and } \|\omega\| \le \delta. \quad (12)$$

Then the sequence $\{\mathrm{dist}(x^k, \bar{X})\}$ converges to 0 superlinearly.

3. A proximal point algorithm for the monotone NCP.

3.1. A proximal point algorithm.

Based on the previous discussion, in this subsection we describe PPA for solving $\mathrm{NCP}(F)$ where F is smooth and monotone. We first illustrate the related mappings that will be used in the remainder of this paper.

Here the mappings Φ_p and Ψ_p are the same as the ones defined by (6) and (8), respectively, except the mapping $\hat{F}$ is substituted by F, i.e.,

$$\Phi_p^k(x) := \begin{pmatrix} \phi_p(x_1, F_1^k(x)) \\ \vdots \\ \phi_p(x_i, F_i^k(x)) \\ \vdots \\ \phi_p(x_n, F_n^k(x)) \end{pmatrix}, \qquad \Phi_p(x) := \begin{pmatrix} \phi_p(x_1, F_1(x)) \\ \vdots \\ \phi_p(x_i, F_i(x)) \\ \vdots \\ \phi_p(x_n, F_n(x)) \end{pmatrix},$$

$$\Psi_p^k(x) := \|\Phi_p^k(x)\|^2, \qquad \Psi_p(x) := \|\Phi_p(x)\|^2,$$

and $\Psi^k(x) := \sum_{i=1}^n \psi(x_i, F_i^k(x))$ with $\psi(x_i, F_i^k(x)) = |x_i F_i^k(x)| + \left|\phi_{\text{NR}}(x_i, F_i^k(x))\right|$. Now we are in a position to describe the proximal point algorithm for solving NCP(F).

Algorithm 3.1.

Step 0.: Choose parameters $\alpha \in (0,1)$, $c_0 \in (0,1)$ and an initial point $x^0 \in \mathbb{R}^n$. Set $k := 0$.

Step 1.: If x^k satisfies $\Psi_p(x^k) = 0$, then stop.

Step 2.: Let $F^k(x) = F(x) + c_k(x - x^k)$. Get an approximation solution $\tilde{x}^{k+1}$ of NCP(F^k) that satisfies the conditions

$$\Psi^k\left([\tilde{x}^{k+1}]_+\right) \le \frac{c_k^3 \min\{1, \|x^k - \tilde{x}^{k+1}\|\}}{4\max\{1, \|[\tilde{x}^{k+1}]_+\|\}^2}. \tag{13}$$

Step 3.: Set $x^{k+1} := [\tilde{x}^{k+1}]_+$, $c_{k+1} = \alpha c_k$ and $k := k+1$. Go to Step 1.

The condition (13) is different from what was used in [42]. In fact, it can be decomposed into two parts as below

$$\Psi^k\left([\tilde{x}^{k+1}]_+\right) \le \frac{c_k^3}{4\max\{1, \|[\tilde{x}^{k+1}]_+\|\}^2} \tag{14}$$

$$\Psi^k\left([\tilde{x}^{k+1}]_+\right) \le \frac{c_k^3\|x^k - \tilde{x}^{k+1}\|}{4\max\{1, \|[\tilde{x}^{k+1}]_+\|\}^2} \tag{15}$$

which corresponds to the aforementioned Criterion 1 and Criterion 2, respectively, to ensure the global and superlinear convergence. However, as shown in the next lemma, the condition (15) implies

$$\sqrt{\Psi_p^k(\tilde{x}^{k+1})} \le c_k^3\|x^k - \tilde{x}^{k+1}\|. \tag{16}$$

Note that condition (14) and condition (16) are exactly what were used in [42]. In view of this, our modified PPA is more neat and compact.

Lemma 3.1. *Suppose that the following inequality holds*

$$\Psi^k([\tilde{x}^{k+1}]_+) \le \frac{c_k^3\|x^k - \tilde{x}^{k+1}\|}{4\max\{1, \|[\tilde{x}^{k+1}]_+\|\}^2}.$$

Then, we have

$$\sqrt{\Psi_p^k(\tilde{x}^{k+1})} \le c_k^3\|x^k - \tilde{x}^{k+1}\|.$$

Proof. First, applying Lemma 2.5(f) gives

$$\begin{aligned}
\sqrt{\Psi_p^k(\tilde{x}^{k+1})} &\le (2 + \sqrt[p]{2})\sqrt{\Psi_{\text{NR}}}(\tilde{x}^{k+1}) \\
&\le (2 + \sqrt[p]{2})\sum_{i=1}^n |\phi_{\text{NR}}(\tilde{x}_i^{k+1}, F_i^k(\tilde{x}^{k+1}))| \\
&\le 4\Psi^k([\tilde{x}^{k+1}]_+).
\end{aligned}$$

On the other hand, from the assumption

$$\Psi^k([\tilde{x}^{k+1}]_+) \le \frac{c_k^3\|x^k - \tilde{x}^{k+1}\|}{4\max\{1, \|[\tilde{x}^{k+1}]_+\|\}^2}$$

and $\max\{1, \|[\tilde{x}^{k+1}]_+\|\}^2 \geq 1$, we have

$$4\Psi^k([\tilde{x}^{k+1}]_+) \leq c_k^3 \|x^k - \tilde{x}^{k+1}\|.$$

Thus, the desired result follows. □

Theorem 3.2. *Let $\bar{X}$ be the solution set of NCP(F). If $\bar{X} \neq \emptyset$, then the sequence $\{x^k\}$ generated by Algorithm 3.1 converges to a solution x^* of NCP(F).*

Proof. Since F^k is strongly monotone with modulus $c_k > 0$, $P_k(x^k)$ is the unique solution of NCP(F^k). From (13), we know

$$\Psi^k(x^{k+1}) \leq \frac{c_k^3}{4\max\{1, \|x^{k+1}\|\}^2} \leq \frac{c_k}{4}.$$

Then, it follows from Lemma 2.4 that

$$\|x^{k+1} - P_k(x^k)\| \leq 2\max\{1, \|x^{k+1}\|\}\sqrt{\frac{\Psi^k(x^{x+1})}{c^k}}, \tag{17}$$

which together with (13) implies

$$\|x^{k+1} - P_k(x^k)\| \leq c_k. \tag{18}$$

Thus, $\{x^k\}$ satisfies Criterion 1 and the global convergence is guaranteed in light of Theorem 2.7. □

In order to obtain superlinear convergence, we need the following assumption which is connected to the condition adopted in Theorem 2.8 (see Theorem 3.4).

Assumption 3.1. *$\|\min\{x, F(x)\}\|$ provides a local error bound for NCP(F), i.e., there exist positive constants $\bar{\delta}$ and $\bar{C}$ such that*

$$\operatorname{dist}(x, \bar{X}) \leq \bar{C}\|\min\{x, F(x)\}\| \text{ for all } x \text{ with } \|\min\{x, F(x)\}\| \leq \bar{\delta}, \tag{19}$$

where $\bar{X}$ denotes the solution set of NCP(F).

Lemma 3.3. [31, Proposition 3] *If a Lipschitz continuous mapping H is strongly BD-regular at x^*, then there is a neighborhood $\mathbb{N}$ of x^* and a positive constant α such that $\forall x \in \mathbb{N}$ and $V \in \partial_B H(x)$, V is nonsingular and $\|V^{-1}\| \leq \alpha$. Furthermore, if H is semismooth at x^* and $H(x^*) = 0$, then there exists a neighborhood $\mathbb{N}'$ of x^* and a positive constant β such that $\forall x \in \mathbb{N}'$, $\|x - x^*\| \leq \beta\|H(x)\|$.*

We note that when $\nabla F(\bar{x})$ is positive definite at one solution $\bar{x}$ of NCP(F), we see that Assumption 3.1 is satisfied due to Lemma 2.3 and Lemma 3.3 in which we view $H(x) = \Phi_p(x)$. Lemma 3.3 also indicates under what conditions Assumption 3.1 holds. Now, we show that Assumption 3.1 indeed implies the condition (12).

Theorem 3.4. *Let $\mathcal{T}$ be defined by (11). If $\bar{X} \neq \emptyset$, then Assumption 3.1 implies condition (12), that is, there exist $\delta > 0$ and $C > 0$ such that*

$$\operatorname{dist}(x, \bar{X}) \leq C\|\omega\|,$$

whenever $x \in \mathcal{T}^{-1}(\omega)$ and $\|\omega\| \leq \delta$.

Proof. For all $x \in \mathcal{T}^{-1}(\omega)$ we have

$$\omega \in \mathcal{T}(x) = F(x) + \mathcal{N}_{\mathcal{K}}(x).$$

Therefore there exists $v \in \mathcal{N}_{\mathcal{K}}(x)$ such that $\omega = F(x) + v$. Because $\mathcal{K}$ is a convex set, it is easy to obtain

$$\Pi_{\mathcal{K}}(x + v) = x. \tag{20}$$

Noting that the projective mapping onto a convex set is nonexpansive, it follows from (20) that

$$\|x - \Pi_{\mathcal{K}}(x - F(x))\| = \|\Pi_{\mathcal{K}}(x + v) - \Pi_{\mathcal{K}}(x - F(x))\| \le \|v + F(x)\| = \|\omega\|.$$

From Assumption 3.1 and letting $C = \bar{C}$, $\delta = \bar{\delta}$ yield the desired condition (12) in Theorem 2.8. □

The following theorem gives the superlinear convergence of Algorithm 3.1, whose proof is based on Theorem 3.4 and can be obtained in the same way as done in Theorem 3.2.

Theorem 3.5. *Suppose that Assumption 3.1 holds. Let $\{x^k\}$ be generated by Algorithm 3.1. Then the sequence $\{\mathrm{dist}(x^k, \bar{X})\}$ converges to 0 superlinearly.*

Proof. By Theorem 3.2, $\{x^k\}$ is bounded. Hence F^k is uniformly Lipschitz continuous with modulus L on a bounded set containing $\{x^k\}$. By Lemma 2.4 and Lemma 3.1 we have

$$\|\tilde{x}^{k+1} - P_k(x^k)\| \le \frac{B_2(L+1)}{c_k}\sqrt{\Psi_p^k(\tilde{x}^{k+1})} \le B_2(L+1)c_k^2\|x^k - \tilde{x}^{k+1}\|,$$

for some positive constant B_2. Then the desired results follows by applying Theorem 2.8 and Theorem 3.4. □

3.2. **Generalized Newton method.** Although we have obtained the global and superlinear convergence properties of Algorithm 3.1 under mild conditions, we still need to care about how to obtain an approximation solution of the strongly monotone complementarity problem in its step 2 and what is the cost to ensure that Algorithm 3.1 is practically efficient. We will discuss this issue in this subsection. More specifically, we introduce the generalized Newton method proposed by De Luca, Facchinei, and Kanzow [25] for solving the subproblems in Step 2 of Algorithm 3.1. As mentioned earlier, for each fixed k, Problem (4) is equivalent to the following nonsmooth equation

$$\Phi_p^k(x) = 0. \tag{21}$$

We describe as below the generalized Newton method for solving the nonsmooth system (21), which is employed from what was introduced in [42] for solving NCP.

Algorithm 3.2 (generalized Newton method for NCP(F^k)).

Step 0.: Choose $\beta \in (0, \frac{1}{2})$ and an initial point $x^0 \in \mathbb{R}^n$. Set $j := 0$.

Step 1.: If $\|\Phi_p^k(x^j)\| = 0$, then stop.

Step 2.: Select an element $V^j \in \partial_B \Phi_p^k(x^j)$. Find the solution d^j of the system

$$V^j d = -\Phi_p^k(x^j). \tag{22}$$

Step 3.: Find the smallest nonnegative integer i_j such that

$$\Psi_p^k(x^j + 2^{-i_j}d^j) \le (1 - \beta 2^{1-i_j})\Psi_p^k(x^j).$$

Step 4.: Set $x^{j+1} := x^j + 2^{-i_j}d^j$ and $j := j + 1$. Go to Step 1.

Mimicking [25, Theorem 3.1], the following theorem which guarantees the convergence of the above algorithm can be obtained easily.

Theorem 3.6. *Suppose that F is differentiable and strongly monotone and that ∇F is Lipschitz continuous around the unique solution $\hat{x}$ of NCP(F^k). The Algorithm 3.2 globally converges to $\hat{x}$ and the rate of convergence is quadratic.*

4. **A proximal point algorithm for the P_0-NCP.** Most ideas in the PPA for monotone NCP introduced in previous section can be adopted for the nonmonotone NCP. The main concern is that the global and superlinear convergence relying on the monotone properties cannot be carried over to nonomonotone case. Fortunately, for the P_0-NCP which is a special subclass of nonmonotone NCPs, it is possible to execute PPA in such case with different conditions. This is what we will elaborate in this section. Note that it is known that if F is an P_0-function, then F^k defined by (5) is an P-function. Therefore, the subproblem NCP(F^k) always has a unique solution, so that PPA is well-defined. Here, the definitions of mappings Φ_p, Ψ_p, Φ_p^k and Ψ_p^k are the same as those defined in Section 3. The PPA for solving NCP(F) with P_0-function can be described as follows:

Algorithm 4.1.

Step 0.: Choose parameters $c_0 > 0$, $\delta_0 \in (0,1)$ and an initial point $x^0 \in \mathbb{R}^n$. Set $k := 0$.

Step 1.: If x^k satisfies $\Psi_p(x^k) = 0$, then stop.

Step 2.: Let $F^k(x) = F(x) + c_k(x - x^k)$. Get an approximation solution $\tilde{x}^{k+1}$ of NCP(F^k) that satisfies the conditions

$$\Psi_p^k(\tilde{x}^{k+1}) \le 2\delta_k^2 \min\{1, \|x^k - \tilde{x}^{k+1}\|^2\}. \tag{23}$$

Step 3.: Set $x^{k+1} := \tilde{x}^{k+1}$. Choose $c_{k+1} \in (0, c_k)$ and $\delta_{k+1} \in (0, \delta_k)$. Set $k := k + 1$. Go to Step 1.

We point it out that the condition (23) is different from the condition (13) used in Algorithm 3.1 because we are dealing with the nonmonotone NCP here. We first summarize some useful properties and give some assumptions that will be used in our convergence analysis.

Lemma 4.1. *For any a, b and c, we have $|\phi_p(a, b + c) - \phi_p(a, b)| \le 2|c|$. Moreover, $\|\Phi_p^k(x) - \Phi_p(x)\| \le 2c_k\|x - x^k\|$.*

Proof. The proof is direct and an extension of [43, Lemma 2.2], we omit it here. □

Proposition 1. *Suppose that F is an P_0-function. Then, for each k, the merit function Ψ_p^k is coercive, i.e.,*

$$\lim_{\|x\| \to \infty} \Psi_p^k(x) = +\infty.$$

Proof. Based on the property of ϕ_p given by [5, Lemma 3.1], the conclusion can be drawn in a way similar to [13, Proposition 3.4]. □

The following two assumptions are needed for the nonmonotone NCP case.

Assumption 4.1. *The sequence $\{c_k\}$ satisfies the following conditions:*

(a): $c_k(x^{k+1} - x^k) \to 0$ if $\{x^k\}$ is bounded.

(b): $c_k x^k \to 0$ if $\{x^k\}$ is unbounded.

Assumption 4.2. *The sequence $\{x^k\}$ converges to a solution x^* of NCP(F) that is strongly BD-regular with Φ_p.*

A few words about these two assumptions. As remarked in [43], we can use the following update rules for parameters c_k and δ_k in Algorithm 4.1 that satisfy both

Assumption 4.1 and step 3 in Algorithm 4.1.

$$c_k = \min\left\{1, \frac{1}{\|x^k\|^2}\right\} \min\left\{(\gamma)^k, \frac{1}{2}\Psi_p(x^k)\right\}, \ \delta_k = (\gamma)^k, \tag{24}$$

where $\gamma \in (0,1)$ is a given constant. Since F is an P_0-function, from [12, Theorem 3.2] and [32, Proposition 2.5], we know that Assumption 4.2 ensures that x^* is the unique solution of NCP(F). In addition, sufficient conditions for Assumption 4.2 have already been given in Lemma 2.3.

Now we are in a position to present the global and superlinear convergence for Algorithm 4.1 whose proofs actually can be shown in a way similar to [43, Theorem 1 and Theorem 3]. Thus, we omit them here.

Theorem 4.2. *Suppose that F is an P_0-function and the solution set of NCP(F) is nonempty and bounded. Suppose also that Assumption 4.1 holds. If $\delta_k \to 0$, then the sequence $\{x^k\}$ generated by Algorithm 4.1 is bounded and any accumulation point of $\{x^k\}$ is a solution of NCP(F).*

Theorem 4.3. *Suppose that Assumption 4.1 and Assumption 4.2 hold and the solution set of NCP(F) is bounded. Suppose also that $\delta_k \to 0$ and $c_k \to 0$. Then the sequence $\{x^k\}$ generated by Algorithm 4.1 converges superlinearly to the solution x^* of NCP(F).*

Like the monotone case, for solving subproblem NCP(F^k), we also adopt the generalized Newton method, i.e., Algorithm 3.2 described in Section 3. Since F is an P_0-function, it can be easily verified from the definition of F^k that the Jacobian matrices of F^k at any point x are P-matrices. Hence Φ_p^k is strongly BD-regular at x by Lemma 2.3. As a result, from [25, Theorem 11], we also have the following convergence theorem for Algorithm 3.2.

Theorem 4.4. *Suppose that F is continuously differentiable and is an P_0-function. Then Algorithm 3.2 globally converges to the unique solution of NCP(F^k). Furthermore, if ∇F is locally Lipschitz continuous, then the convergence rate is quadratic.*

5. **Numerical experiments.** In this section, we report numerical results of Algorithm 3.1 and Algorithm 4.1 for solving NCP(F) defined by (1). Our numerical experiments are carried out in MATLAB (version 7.8) running on a PC Intel core 2 Q8200 of 2.33GHz CPU and 2.00GB Memory for the test problems with all available starting points in MCPLIB [2].

In our numerical experiments, the stopping criteria for Algorithm 3.1 and Algorithm 4.1 are $\Psi_p(x^k) \le 1.0e-8$ and $|(x^k)^T(F(x^k))| \le 1.0e-3$. We also stop programs when the total iteration is more than 50. In Algorithm 3.1, we set the parameters as $\alpha = 0.5$ and $c_0 = 0.5$. In Algorithm 4.1, we adopt rules (24) to update the parameters where $\gamma = 0.5$. In Algorithm 3.2, we set the parameter $\beta = 10^{-4}$ and the initial point for Newton's method is selected as the current iteration point in the main algorithm. Notice that the main task of Algorithm 3.2 for solving the subproblem, at each iterate, is to solve the linear system (22). In numerical implementation, we apply the preconditioner conjugate gradient square method for solving system (22). Because there is no description about how to distinguish monotone NCP and non-monotone NCP from MCPLIB collection, we only use the iteration point to test the non-monotonicity. In summary, among 82 test problems in MCPLIB (different starting points are regarded as different test problems), 50 test problems are verified to be non-monotone. As a result, Algorithm 4.1

is implemented on the whole 82 problems, while Algorithm 3.1 is only implemented on 32 problems which may be monotone. A good news is that with a proper p, all the test problems can be solved successfully, which shows that the proximal point algorithms for solving complementarity problems are efficient.

To present an objective evaluation and comparison of the performance of Algorithm 3.1 and Algorithm 4.1 with different p, we consider the performance profile introduced in [11] as a means. In particular, we regard Algorithm 3.1 or Algorithm 4.1 corresponding to a p as a solver, and assume that there are n_s solvers and n_j test problems from the MCPLIB collection $\mathscr{J}$. We are interested in using the number of function evaluations as performance measure for Algorithm 3.1 or Algorithm 4.1 with different p. For each problem j and solver s, let

$$f_{j,s} := \text{ function evaluations required to solve problem } j \text{ by solver } s.$$

We compare the performance on problem j by solver s with the best performance by any one of the n_s solvers on this problem; that is, we adopt the performance ratio

$$r_{j,s} = \frac{f_{j,s}}{\min\{f_{j,s} : s \in \mathscr{S}\}},$$

where $\mathscr{S}$ is the set of four solvers. An overall assessment of each solver is obtained from

$$\rho_s(\tau) = \frac{1}{n_j}\text{size}\{j \in \mathscr{J} : r_{j,s} \leq \tau\},$$

which is called the performance profile of the number of function evaluations for solver s. Note that $\rho_s(\tau)$ approximates the probability for solver s that a performance ratio $r_{j,s}$ is within a factor τ of the best possible ratio.

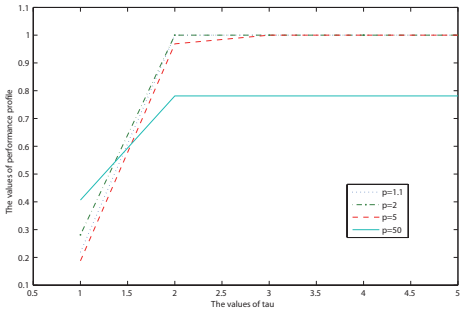

FIGURE 1. Performance profile of function evaluations of Algorithm 3.1 with four p.

Figure 1 shows the performance profile of function evaluations in Algorithm 3.1 in the range of $[0, 5]$ for four solvers on 32 test problems. The four solvers correspond to Algorithm 3.1 with $p = 1.1$, $p = 2$, $p = 5$ and $p = 50$, respectively. From this figure, we see that Algorithm 3.1 with $p = 50$ and $p = 2$ has the competitive wins (has the highest probability of being the optimal solver) and that the probabilities that it is the winner on a given monotone NCP are about 0.40 and 0.28, respectively. If we

choose being within a factor of greater than 2 of the best solver as the scope of our interest, then $p = 2$ and $p = 1.1$ would suffice, and the performance profile shows that the probability that Algorithm 3.1 with this p can solve a given monotone NCP in such range of the best solver is almost 100%. And Algorithm 3.1 with $p = 5$ has a comparable performance with $p = 2$ and $p = 1.1$ if we choose being within a factor of greater than 3 of the best solver as the scope of our interest. Although $p = 50$ has a competitive wins with $p = 2$, the probability that it can solve a given NCP within any positive factor of the best solver is lower than $p = 2$. Actually, $p = 50$ has the lowest probability within a factor of greater than 2 of the best solver. To sum up, Algorithm 3.1 with $p = 2$ have the best numerical performance than the others.

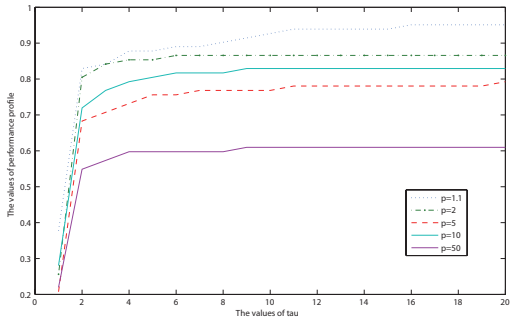

FIGURE 2. Performance profile of function evaluations of Algorithm 4.1 with five p.

Figure 2 shows the performance profile of function evaluations in Algorithm 4.1 in the range of [0, 20] for five solvers on 82 test problems. The five solvers correspond to Algorithm 4.1 with $p = 1.1$, $p = 2$, $p = 5$, $p = 10$ and $p = 50$, respectively. From this figure, we see that Algorithm 4.1 with $p = 1.1$ has the best numerical performance (has the highest probability of being the optimal solver) and that the probability that it is the winner on a given NCP is about 0.37. If we choose being within a factor of 2 or 3 of the best solver, then $p = 2$ has a comparable performance with $p = 1.1$. If we choose being within a factor of greater than 11 or 16 of the best solver as the scope of our interest, the performance profile of $p = 1.1$ shows that the probabilities that Algorithm 4.1 with this p can solve a given NCP in such range of the best solver are 93% and 95%, respectively.

To sum up, although changing the parameter p does change the numerical performance for the proposed PPA, however, such change does not depend on p regularly, unlike mentioned in [3, 4, 6, 7] which are reported for other algorithms.

Acknowledgments. The authors thank the referees for their careful reading of the paper and helpful suggestions.

REFERENCES

[1] E. D. Andersen, C. Roos and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming, **95** (2003), 249–277.

[2] S. C. Billups, S. P. Dirkse and M. C. Soares, *A comparison of algorithms for large scale mixed complementarity problems*, Computational Optimization and Applications, **7** (1997), 3–25.
[3] J.-S. Chen, *The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem*, Journal of Global Optimization, **36** (2006), 565–580.
[4] J-.S. Chen, *On some NCP-function based on the generalized Fischer-Burmeister function*, Asia-Pacific Journal of Operational Research, **24** (2007), 401–420.
[5] J.-S. Chen and S.-H. Pan, *A family of NCP functions and a desent method for the nonlinear complementarity problem*, Computational Optimization and Applications, **40** (2008), 389–404.
[6] J.-S. Chen and S.-H. Pan, *A regularization semismooth Newton method based on the generalized Fischer-Burmeister function for P_0-NCPs*, Journal of Computational and Applied Mathematics, **220** (2008), 464–479.
[7] J.-S. Chen, H.-T. Gao and S.-H. Pan, *An R-linearly convergent derivative-free algorithm for the NCPs based on the generalized Fischer-Burmeister merit function*, Journal of Computational and Applied Mathematics, **232** (2009), 455–471.
[8] J.-S. Chen, S.-H. Pan and C.-Y. Yang, *Numerical comparison of two effective methods for mixed complementarity problems*, Journal of Computational and Applied Mathematics, **234** (2010), 667–683.
[9] X. Chen and H. Qi, *Cartesian P-property and its applications to the semidefinite linear complementarity problem*, Mathematical Programming, **106** (2006), 177–201.
[10] F. H. Clarke, "Optimization and Nonsmooth Analysis," John Wiley and Sons, New York, 1983.
[11] E. D. Dolan and J. J. More, *Benchmarking optimization software with performance profiles*, Mathematical Programming, **91** (2002), 201–213.
[12] F. Facchinei, *Structural and stability properties of P_0 nonlinear complementarity problems*, Mathematics of Operations Research, **23** (1998), 735–745.
[13] F. Facchinei and C. Kanzow, *Beyond monotonicity in regularization methods for nonlinear complementarity problems*, SIAM Journal on Control and Optimization, **37** (1999), 1150–1161.
[14] F. Facchinei and J.-S. Pang, "Finite-Dimensional Variational Inequalities and Complementarity Problems," Springer Verlag, New York, 2003.
[15] F. Facchinei and J. Soares, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM Journal on Optimazation, **7** (1997), 225–247.
[16] M. C. Ferris and J.-S. Pang, *Engineering and economic applications of complementarity problems*, SIAM J. Rev., **39** (1997), 669–713.
[17] A. Fischer, *Solution of monotone complementarity problems with locally Lipschitzian functions*, Mathematical Programming, **76** (1997), 513–532.
[18] P. T. Harker and J.-S. Pang, *Finite dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications*, Mathematical Programming, **48** (1990), 161–220.
[19] 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.
[20] Z.-H. Huang and W.-Z. Gu, *A smoothing-type algorithm for solving linear complementarity problems with strong convergence properties*, Appl. Math. Optim., **57** (2008), 17–29.
[21] Z.-H. Huang, L. Qi and D. Sun, *Sub-quadratic convergence of a smoothing Newton algorithm for the P_0- and monotone LCP*, Mathematical Programming, **99** (2004), 423–441.
[22] H.-Y. Jiang, M. Fukushima and L. Qietal., *A trust region method for solving generalized complementarity problems*, SIAM Journal on Control and Optimization, **8** (1998), 140–157.
[23] C. Kanzow and H. Kleinmichel, *A new class of semismooth Newton method for nonlinear complementarity problems*, Computational Optimization and Applications, **11** (1998), 227–251.
[24] C. Kanzow, N. Yamashita and M. Fukushima, *New NCP-functions and their properties*, Journal of Optimization Theory and Applications, **94** (1997), 115–135.
[25] T. De Luca, F. Facchinei and C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Mathematical Programming, **75** (1996), 407–439.
[26] F. J. Luque, *Asymptotic convergence analysis of the proximal point algorithm*, SIAM Journal on Control and Optimization, **22** (1984), 277–293.
[27] B. Martinet, *Perturbation des méthodes d'opimisation*, RAIRO Anal. Numér., **12** (1978), 153–171.

[28] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM Journal on Control and Optimization, **15** (1997), 957–972.
[29] J.-S. Pang, *A posteriori error bounds for the linearly-constrained variational inequality problem*, Mathematics of Operations Research, **12** (1987), 474–484.
[30] J.-S. Pang, *Complementarity problems*, in " Handbook of Global Optimization" (eds. R. Horst and P. Pardalos), Kluwer Academic Publishers, Boston, Massachusetts, (1994), 271–338.
[31] J.-S. Pang and L. Qi, *Nonsmooth equations: Motivation and algorithms*, SIAM Journal on Optimization, **3** (1993), 443–465.
[32] L. Qi, *A convergence analysis of some algorithms for solving nonsmooth equations*, Mathematics of Operations Research, **18** (1993), 227–244.
[33] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming, **58** (1993), 353–367.
[34] S. M. Robinson, *Strongly regular generalized equations*, Mathematics of Operations Research, **5** (1980), 43–62.
[35] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization, **14** (1976), 877–898.
[36] R. T. Rockafellar and R. J-B. Wets, "Variational Analysis," Springer-Verlag Berlin Heidelberg, 1998.
[37] D. Sun and L. Qi, *On NCP-functions*, Computational Optimization and Applications, **13** (1999), 201–220.
[38] D. Sun and J. Sun, *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions*, Mathematical Programming, **103** (2005), 575–581.
[39] J. Wu and J.-S. Chen, *A proximal point algorithm for the monotone second-order cone complementarity problem*, Computational Optimization and Applications, **51** (2012), 1037–1063.
[40] N. Yamashita and M. Fukushima *On stationary points of the implicitm Lagrangian for nonlinear complementarity problems*, Journal of Optimization Theory and Applications, **84** (1995), 653–663.
[41] N. Yamashita and M. Fukushima *Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems*, Mathematical Programming, **76** (1997), 469–491.
[42] N. Yamashita and M. Fukushima, *The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem*, SIAM Journal on Optimization, **11** (2000), 364–379.
[43] N. Yamashita, J. Imai and M. Fukushima, *The proximal point algorithm for the P_0 complementarity problem*, in "Complementarity: Applications, Algorithms and Extensions" (eds. M. C. Ferris et al.), Kluwer Academic Publisher, (2001), 361–379.

Received May 2011; revised May 2012.

E-mail address: ylchang@math.ntnu.edu.tw
E-mail address: jschen@math.ntnu.edu.tw
E-mail address: wujia@dlut.edu.cn