Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2012.m10027

Vol. **5**, No. 2, pp. 186-204
May 2012

A Regularization Semismooth Newton Method for P_0-NCPs with a Non-monotone Line Search

Li-Yong Lu[1], Wei-Zhe Gu[2,*] and Wei Wang[3]

[1] *Department of Mathematics, School of Science, Tianjin University of Technology, Tianjin 300384, China.*
[2] *Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, China.*
[3] *School of Public Administration, Tianjin University of Commerce, Tianjin 300134, China.*

Received 6 August 2010; Accepted (in revised version) 23 May 2011

Available online 27 March 2012

Abstract. In this paper, we propose a regularized version of the generalized NCP-function proposed by Hu, Huang and Chen [J. Comput. Appl. Math., 230 (2009), pp. 69–82]. Based on this regularized function, we propose a semismooth Newton method for solving nonlinear complementarity problems, where a non-monotone line search scheme is used. In particular, we show that the proposed non-monotone method is globally and locally superlinearly convergent under suitable assumptions. We test the proposed method by solving the test problems from MCPLIB. Numerical experiments indicate that this algorithm has better numerical performance in the case of $p=5$ and $\theta \in [0.25, 075]$ than other cases.

AMS subject classifications: 90C33, 65K10

Key words: Nonlinear complementarity problem, non-monotone line search, semismooth Newton method, global convergence, local superlinear convergence.

1. Introduction

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

$$x \geq 0, \quad f(x) \geq 0, \quad x^T f(x) = 0, \tag{1.1}$$

where $f : \Re^n \to \Re^n$ is a continuously differentiable mapping with $f := (f_1, f_2, \cdots, f_n)^T$. If f is a P_0-function, i.e.,

$$\max_{1 \leq i \leq n \, x_i \neq y_i} (x_i - y_i)(f_i(x) - f_i(y)) \geq 0$$

*Corresponding author. *Email address:* `weizhegu@yahoo.com.cn` (W.-Z. Gu)

holds for all $x, y \in \Re^n$ and $x \neq y$, then we call (1.1) the P_0-NCP. The NCP has various applications in operation research, economics, and engineering (see, for example, [1–4]). Various methods for solving the NCP have been proposed in the literature (see, for example, [5–18]). In all the mentioned methods, the so-called NCP-function, i.e., $\phi(a,b)=0$ if and only if $a \geq 0, b \geq 0, ab = 0$, plays an important role.

Recently, a family of new NCP-functions was proposed in [19], which is defined by

$$\omega_{\theta p}(a,b) := \sqrt[p]{\theta(|a|^p + |b|^p) + (1-\theta)|a-b|^p} - a - b, \tag{1.2}$$

where $p \in (1, +\infty), \theta \in (0,1]$, and $(a,b) \in \Re^2$. When $\theta = 1$, the function $\omega_{\theta p}$ reduces to the function in [20, 21], and when $\theta = 1$ and $p = 2$, the function $\omega_{\theta p}$ reduces to the Fischer-Burmeister function [10]. It was showed in [19, Propositions 2.1 and 2.3] that the function $\omega_{\theta p}(\cdot,\cdot)$ is an NCP function and a semismooth function in $\Re^2$. Moreover, it is known that $\omega^2_{\theta p}(\cdot,\cdot)$ is continuously differentiable and strongly semismooth in $\Re^2$ (see, e.g., [19, Proposition 2.5]).

In this paper, by using the symmetrically perturbed technique proposed in [22, 23], we give a regularized version of the generalized NCP-function (1.2), which is defined by

$$\begin{aligned}
&\phi_{\theta p}(\mu,a,b)\\
&= \sqrt[p]{\theta(|\mu a+b|^p + |a+\mu b|^p) + (1-\theta)|\mu a + b - (a+\mu b)|^p} - ((\mu a + b) + (a + \mu b))\\
&= \sqrt[p]{\theta(|\mu a+b|^p + |a+\mu b|^p) + (1-\theta)|(1-\mu)(a-b)|^p} - (1+\mu)(a+b),
\end{aligned} \tag{1.3}$$

where $(\mu,a,b) \in \Re_+ \times \Re \times \Re$; and $\theta \in [0,1]$ and $p \in (1,+\infty)$ are two given parameters. It is obvious that $\phi_{\theta p}(0,\cdot,\cdot)$ is an NCP function. For all $z := (\mu, x) \in \Re_+ \times \Re^n$, we define

$$H_{\theta p}(z) := \begin{pmatrix} \mu \\ \Phi_{\theta p}(z) \end{pmatrix} \qquad \Psi_{\theta p}(z) := \|H_{\theta p}(z)\|^2 = \mu^2 + \|\Phi_{\theta p}(z)\|^2, \tag{1.4}$$

where

$$\Phi_{\theta p}(z) := \begin{pmatrix} \phi_{\theta p}(\mu, x_1, f_1(x)) \\ \vdots \\ \phi_{\theta p}(\mu, x_n, f_n(x)) \end{pmatrix}. \tag{1.5}$$

It is easy to see that $z := (\mu, x)$ is a solution of $H_{\theta p}(z) = 0$ if and only if $\mu = 0$ and x solves the NCP (1.1). We will show that the function $H_{\theta p}(z)$ defined in (1.4) is coercive with respect to z. Such a property can improve the global convergence of the semismooth Newton method (see [24, 25]; also see [26] for comparisons about the conditions of the global convergence). It should be noted that the function $H_{\theta p}(z)$ defined in (1.4) is not coercive with respect to z if the function $\phi_{\theta p}(\cdot,\cdot,\cdot)$ in the definition of $H_{\theta p}(\cdot)$ is replaced by the function $\omega_{\theta p}(\cdot,\cdot)$ given in (1.2).

Many numerical methods based on the NCP-function (or the smoothed NCP-function) have not only good convergence, but also good numerical results, such as the semismooth

Newton method and smoothing Newton algorithm. Generally, the theoretical analysis of these methods was given based the method with a monotone line search; while numerical experiments were done based on the method with a non-monotone line search in order to improve the numerical results (see, for example, [18,23–25]). Recently, the theoretical analysis of non-monotone smoothing Newton algorithms were given in [14] for the NCP (1.1) and in [27] for system of equalities and inequalities.

In this paper, based on the new function $\phi_{\theta p}(\cdot,\cdot,\cdot)$ defined by (1.3), we propose a semismooth Newton method with a non-monotone line search for solving $H_{\theta p}(z)=0$. We show that the method is globally and locally superlinearly (quadratically) convergent under suitable assumptions. We test the proposed method through solving the test problems from MCPLIB [28]. Numerical experiments indicate that this method has better numerical performance in the case of $p=5$ and $\theta \in [0.25, 075]$ than other cases. It should be noted that the case of $p=5$ and $\theta \in [0.25, 075]$ is not contained in those known cases given in the literature (for example, [10,20,21]). In addition, in order to see that *how the regularized parameter μ would affect the numerical performance when θ and p are fixed*, we also do some numerical experiments by changing the value of μ_0. The numerical results show that the proposed method has worse numerical performance in the case of $\mu_0=0$ than the other cases, which demonstrates that the regularized technique used in this paper is useful for the numerical computation of the method.

The rest of this paper is organized as follows. In Section 2, we discuss some properties of the new function. In Section 3, we give a non-monotone semismooth Newton method for the NCP. The global and local superlinear (quadratic) convergence of the method are discussed in Section 4. Numerical results are reported in Section 5. The conclusions are given in Section 6.

Throughout this paper, $\mathscr{K} := \{1,2,\cdots\}$ and $\mathscr{I} = \{1,2,\cdots,n\}$; $|J|$ denotes the cardinality of an index set J; $\Re^n$ denotes the space of n-dimensional real column vectors and $\Re^n_+$ (respectively, $\Re^n_{++}$) denotes the non-negative (respectively, positive) orthant in $\Re^n$; the superscript ^{T} denotes transpose; and the sign$(\cdot)$ denotes sign function. We use vec$\{u_i : i \in \mathscr{I}\}$ to denote the vector u, and use diag$\{u_i : i \in \mathscr{I}\}$ to denote the diagonal matrix whose i-th diagonal element is u_i. For any two vectors $x \in \Re^l$ and $s \in \Re^r$, where l and r are any two positive integers, we write $(x^T, s^T)^T$ as (x,s) for simplicity. If $\{\alpha_k\}$ and $\{\beta_k\}$ are two sequences in $\Re$ with $\alpha_k, \beta_k > 0$ for all $k \in \mathscr{K}$, $\alpha_k = O(\beta_k)$ means $\limsup_{k\to+\infty} \alpha_k/\beta_k < +\infty$; and $\alpha_k = o(\beta_k)$ means $\limsup_{k\to+\infty} \alpha_k/\beta_k = 0$. For a matrix $A \in \Re^{m\times n}$, and two index sets $\alpha \subseteq \{1,\cdots,m\}$ and $\beta \subseteq \{1,\cdots,m\}$, $A_{\alpha\beta}$ denotes a submatrix of A with $A_{\alpha\beta} \in \Re^{|\alpha|\times|\beta|}$. For any $(\mu,x) \in \Re \times \Re^n$, we always use the following notation unless stated otherwise:

$$z := (\mu, x), \quad z^k := (\mu_k, x^k),\ \forall\ k \in \mathscr{K}, \quad z_i^k := (\mu_k, x_i^k),\ \forall\ k \in \mathscr{K},\ i \in \mathscr{I}.$$

For any fixed $\theta \in [0,1]$ and $p \in (1,+\infty)$, and any $(\mu,a,b) \in \Re_+ \times \Re \times \Re$, we denote

$$\begin{aligned} h_{\theta p}(\mu,a,b) :=& \sqrt[p]{\theta(|\mu a + b|^p + |a+\mu b|^p) + (1-\theta)|(1-\mu)(a-b)|^p}, \\ h(\mu,a,b) :=& \theta\left[a|\mu a+b|^{p-1}\mathrm{sign}(\mu a+b) + b|a+\mu b|^{p-1}\mathrm{sign}(a+\mu b)\right] \\ & - (1-\theta)(a-b)|(1-\mu)(a-b)|^{p-1}\mathrm{sign}((1-\mu)(a-b)). \end{aligned}$$

2. Preliminaries

In this section, we give some basic concepts and preliminary results which will be used in our analysis.

A matrix $M \in \Re^{n\times n}$ is called a P_0-matrix if, for every $x \in \Re^n$ with $x \neq 0$, there is an index $i_0 = i_0(x)$ with $x_{i_0} \neq 0$ and $x_{i_0}[Mx]_{i_0} \geq 0$; and a P-matrix if, for every $x \in \Re^n$ with $x \neq 0$, it holds that $\max_i x_i[Mx]_i > 0$.

The concept of semismoothness plays an important role in the analysis on local fast convergence of some Newton-type methods. Such a concept was originally introduced by Mifflin [29] for functionals and was extended to vector valued functions in [30] by Qi and Sun. A function $f : \Re^n \to \Re^n$ is said to be semismooth (or strongly semismooth) at $x \in \Re^n$ if it is directionally differentiable at x and $Vd - f'(x;d) = o(\|d\|)$ (or$= O(\|d\|^2)$) holds for any $d \to 0$ and $V \in \partial f(x+d)$, where $\partial f(x)$ denotes the Clarke's generalized Jacobian of f at x [31].

By the definition of $\phi_{\theta p}$, similar to the proof of [19, Proposition 2.3], we can obtain the following lemma.

Lemma 2.1. *Let $(\mu, a, b) \in \Re^3$ and $\phi_{\theta p}$ be defined by* (1.3). *Then the function $\phi^2_{\theta p}(\cdot,\cdot,\cdot)$ is continuously differentiable in $\Re^3$; and the function $\phi_{\theta p}(\cdot,\cdot,\cdot)$ is strongly semismooth in $\Re^3$.*

Next, we give the expression of the generalized Jacobian of the function $H_{\theta p}$.

Proposition 2.1. *Suppose that f is a continuously differentiable P_0-function. Given $\theta \in [0,1]$, $p \in (1,+\infty)$, let $\partial H_{\theta p}(z)$ denote the Jacobian matrix of $H_{\theta p}$ defined by* (1.4)*, then for any $z = (\mu, x) \in \Re_+ \times \Re^n$, we have*

$$(\partial H_{\theta p}(z))^T \subseteq \begin{bmatrix} 1 & v_{\theta p}(z)^T \\ 0 & E_{\theta p}(z) \end{bmatrix},$$

where

$$v_{\theta p}(z) := \mathrm{vec}\left\{ \frac{h(\mu, x_i, f_i(x))}{h_{\theta p}(\mu, x_i, f_i(x))^{p-1}} - (x_i + f_i(x)) : i \in \mathscr{I} \right\}, \tag{2.1a}$$

$$\begin{aligned} E_{\theta p}(z) :=& \theta\mu A(z) + \theta B(z) + (1-\theta)(1-\mu)C(z) - (1+\mu)I \\ &+ \nabla f(x)[\theta A(z) + \theta\mu B(z) - (1-\theta)(1-\mu)C(z) - (1+\mu)I], \end{aligned} \tag{2.1b}$$

where I denotes the $n\times n$ identity matrix and $A(z), B(z), C(z)$ are possibly multi-valued $n\times n$ diagonal matrices with the i-th diagonal elements given by

$$A_{ii}(z) := \frac{|\mu x_i + f_i(x)|^{p-1} sign(\mu x_i + f_i(x))}{h_{\theta p}(\mu, x_i, f_i(x))^{p-1}}, \tag{2.2a}$$

$$B_{ii}(z) := \frac{|x_i + \mu f_i(x)|^{p-1} sign(x_i + \mu f_i(x))}{h_{\theta p}(\mu, x_i, f_i(x))^{p-1}}, \tag{2.2b}$$

$$C_{ii}(z) := \frac{|(1-\mu)(x_i - f_i(x))|^{p-1} sign((1-\mu)(x_i - f_i(x)))}{h_{\theta p}(\mu, x_i, f_i(x))^{p-1}} \tag{2.2c}$$

if $(x_i, f_i(x)) \neq (0,0)$*; and by*

$$A_{ii}(z) := \zeta_i, \quad B_{ii}(z) := \eta_i, \quad C_{ii}(z) := \xi_i \tag{2.3}$$

for any (ζ_i, η_i, ξ_i) *such that* $|\zeta_i| \le 1$, $|\eta_i| \le 1$, $|\xi_i| \le 1$, *and*

$$\theta(|\zeta_i|^{p/(p-1)} + |\eta_i|^{p/(p-1)}) + (1-\theta)|\xi_i|^{p/(p-1)} \le 1 \tag{2.4}$$

if $(x_i, f_i(x)) = (0,0)$.

Proof. For any $z := (\mu, x) \in \Re_+ \times \Re^n$, it follows from [31, Proposition 2.6.2(e)] that

$$\partial H_{\theta p}(z)^T \subseteq \partial H_{\theta p,1}(z) \times \partial H_{\theta p,2}(z) \times \cdots \times \partial H_{\theta p,n+1}(z), \tag{2.5}$$

where the right-hand side denotes a set of matrices whose j-th column belongs to $\partial H_{\theta p,j}(z)$, and $H_{\theta p,j}(z)$ is the j-th component function. Firstly, from a direct computation, we obtain that $\partial H_{\theta p,1}(z) = (1,0,0,\cdots,0)^T \in \Re^{1+n}$. Secondly, for $j \in \{2,3,\cdots,n+1\}$, let $i = j-1$, if $(x_i, f_i(x)) \neq (0,0)$, by a direct computation, we obtain that the first component of $\partial H_{\theta p,i}(z)$ is

$$\frac{h_1(\mu, x_i, f_i(x))}{h_{\theta p}(\mu, x_i, f_i(x))^{p-1}} - (x_i + f_i(x));$$

and the last n components of $\partial H_{\theta p,i}(z)$ are

$$\begin{aligned}
&\Big[\theta|\mu x_i + f_i(x)|^{p-1} sign(\mu x_i + f_i(x))(\mu e_i + \nabla f_i(x)) \\
&\quad +\theta|x_i + \mu f_i(x)|^{p-1} sign(x_i + \mu f_i(x))(e_i + \mu\nabla f_i(x)) \\
&\quad +(1-\theta)(1-\mu)|(1-\mu)(x_i - f_i(x))|^{p-1} sign((1-\mu)(x_i - f_i(x)))(e_i - \nabla f_i(x))\Big] \\
&\quad \times h_{\theta p}(\mu, x_i, f_i(x))^{1-p} - (1+\mu)(e_i + \nabla f_i(x)) \\
=&\Big[h_{\theta p}(\mu, x_i, f_i(x))^{1-p}\{\theta\mu|\mu x_i + f_i(x)|^{p-1} sign(\mu x_i + f_i(x)) \\
&\quad +\theta|x_i + \mu f_i(x)|^{p-1} sign(x_i + \mu f_i(x)) \\
&\quad +(1-\theta)(1-\mu)|(1-\mu)(x_i - f_i(x))|^{p-1} sign((1-\mu)(x_i - f_i(x)))\} - (1+\mu)\Big] e_i \\
&\quad +\Big[h_{\theta p}(\mu, x_i, f_i(x))^{1-p}\{\theta|\mu x_i + f_i(x)|^{p-1} sign(\mu x_i + f_i(x)) \\
&\quad +\theta\mu|x_i + \mu f_i(x)|^{p-1} sign(x_i + \mu f_i(x)) \\
&\quad -(1-\theta)(1-\mu)|(1-\mu)(x_i - f_i(x))|^{p-1} sign((1-\mu)(x_i - f_i(x)))\} - (1+\mu)\Big] \nabla f_i(x),
\end{aligned}$$

where $e_i \in \Re^n$ denotes the vector of the i-th component is 1 and the other components are zeros. Hence (2.1) holds, where

$$A(z) = diag\{A_{ii}(z)\}, \quad B(z) = diag\{B_{ii}(z)\}, \quad C(z) = diag\{C_{ii}(z)\}$$

with every $A_{ii}(z), B_{ii}(z), C_{ii}(z)$ being given by (2.2). In addition, if $(x_i, f_i(x)) = (0,0)$, then it is easy to obtain that (2.1) holds, where

$$A(z) = diag\{A_{ii}(z)\}, \quad B(z) = diag\{B_{ii}(z)\}, \quad C(z) = diag\{C_{ii}(z)\}$$

with every $A_{ii}(z), B_{ii}(z), C_{ii}(z)$ being given by (2.3) for any (ζ_i, η_i, ξ_i) satisfying $|\zeta_i| \le 1$, $|\eta_i| \le 1$, $|\xi_i| \le 1$, and

$$\theta(|\zeta_i|^{p/(p-1)} + |\eta_i|^{p/(p-1)}) + (1-\theta)|\xi_i|^{p/(p-1)} \le 1.$$

The proof is complete. □

Proposition 2.2. *Suppose that f is a continuously differentiable P_0-function. Given $\theta \in [0,1]$ and $p \in (1,+\infty)$. Let $E_{\theta p}(z)$ be defined by* (2.1) *where $z = (\mu, x)$, then for any $\mu > 0$, $-E_{\theta p}(z)$ is a P-matrix.*

Proof. Since f is a continuously differentiable P_0-function, it follows that $\nabla f(x)$ is a P_0-matrix, i.e., for any $x \in \Re^n \backslash \{0\}$, there exists an $i_0 \in \{i : x_i \neq 0\}$ such that

$$x_{i_0}(\nabla f(x)^T x)_{i_0} \ge 0.$$

In the following, we show that for the same index i_0, the inequality

$$x_{i_0}(-E_{\theta p}(z)^T x)_{i_0} > 0$$

holds for any $\mu > 0$. From the definition of $E_{\theta p}(z)$, we obtain that

$$\begin{aligned}
&x_{i_0}(E_{\theta p}(z)^T x)_{i_0}\\
=&(\theta\mu A_{i_0 i_0} + \theta B_{i_0 i_0} + (1-\theta)(1-\mu)C_{i_0 i_0} - (1+\mu))x_{i_0}^2\\
&+ [\theta A_{i_0 i_0} + \theta\mu B_{i_0 i_0} - (1-\theta)(1-\mu)C_{i_0 i_0} - (1+\mu)]x_{i_0}(\nabla f(x)^T x)_{i_0},
\end{aligned} \tag{2.6}$$

where

$$\begin{aligned}
&A_{i_0 i_0} := A_{i_0 i_0}(z) = \frac{|a|^{p-1}\mathrm{sign}(a)}{d^{p-1}}, \ \ B_{i_0 i_0} := B_{i_0 i_0}(z) = \frac{|b|^{p-1}\mathrm{sign}(b)}{d^{p-1}},\\
&C_{i_0 i_0} := C_{i_0 i_0}(z) = \frac{|c|^{p-1}\mathrm{sign}(c)}{d^{p-1}}, \ \ a = \mu x_{i_0} + f_{i_0}(x), \ b = x_{i_0} + \mu f_{i_0}(x),\\
&c = (1-\mu)(x_{i_0} - f_{i_0}(x)), \ d = h_{\theta p}(\mu, x_{i_0}, f_{i_0}(x)) = [\theta|a|^p + \theta|b|^p + (1-\theta)|c|^p]^{\frac{1}{p}}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&|\theta\mu A_{i_0 i_0} + \theta B_{i_0 i_0} + (1-\theta)(1-\mu)C_{i_0 i_0}|\\
\le&\theta\mu|a|^{p-1} + \theta|b|^{p-1} + (1-\theta)|1-\mu||c|^{p-1} \times d^{1-p}\\
=&\mu\theta^{\frac{1}{p}}(\theta^{\frac{p-1}{p}}|a|^{p-1}) + \theta^{\frac{1}{p}}(\theta^{\frac{p-1}{p}}|b|^{p-1}) + |1-\mu|(1-\theta)^{\frac{1}{p}}((1-\theta)^{\frac{p-1}{p}}|c|^{p-1}) \times d^{1-p}\\
\le&\left[(\mu\theta^{\frac{1}{p}})^p + (\theta^{\frac{1}{p}})^p + (|1-\mu|(1-\theta)^{\frac{1}{p}})^p\right]^{\frac{1}{p}}\left[(\theta^{\frac{p-1}{p}}|a|^{p-1})^{\frac{p}{p-1}} + (\theta^{\frac{p-1}{p}}|b|^{p-1})^{\frac{p}{p-1}}\right.\\
&\left. + ((1-\theta)^{\frac{p-1}{p}}|c|^{p-1})^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \times d^{1-p}\\
=&\left[\mu^p\theta + \theta + |1-\mu|^p(1-\theta)\right]^{\frac{1}{p}} \times \left[\theta|a|^p + \theta|b|^p + (1-\theta)|c|^p\right]^{\frac{p-1}{p}} \times \left[\theta|a|^p + \theta|b|^p + (1-\theta)|c|^p\right]^{\frac{1-p}{p}}\\
=&\left[\mu^p\theta + \theta + |1-\mu|^p(1-\theta)\right]^{\frac{1}{p}}.
\end{aligned}$$

where the first inequality follows from the triangle inequality; and the second inequality follows from the well-known Hölder inequality. For $\mu > 0$, it is easy to show that

$$g(\theta) := \left[\mu^p \theta + \theta + |1-\mu|^p (1-\theta)\right]^{\frac{1}{p}}$$

is a monotonically increasing function. Then for any $\theta \in [0,1]$,

$$g(\theta) \leq g(1) = (1+\mu^p)^{\frac{1}{p}} < 1+\mu.$$

Thus,

$$\theta\mu A_{i_0 i_0} + \theta B_{i_0 i_0} + (1-\theta)(1-\mu) C_{i_0 i_0} - (1+\mu) < 0.$$

Similarly,

$$\theta A_{i_0 i_0} + \theta\mu B_{i_0 i_0} - (1-\theta)(1-\mu) C_{i_0 i_0} - (1+\mu) < 0.$$

These, together with the fact that $x_{i_0}(\nabla f(x)^T x)_{i_0} \geq 0$, imply that $x_{i_0}(E_{\theta p}(z)^T x)_{i_0} < 0$ for any $\mu > 0$. Therefore, for any $\mu > 0$, $-E_{\theta p}(z)$ is a P-matrix. □

From Propositions 2.1 and 2.2, the following result can be easily obtained.

Corollary 2.1. *Suppose that f is a continuously differentiable P_0-function. Given $\theta \in [0,1]$ and $p \in (1,+\infty)$. Then, for any $z = (\mu, x) \in \Re \times \Re^n$ with $\mu > 0$, all $V \in \partial H_{\theta p}(z)$ are nonsingular.*

If $\mu = 0$, then the result obtained above does not hold in general. In the following, we give a condition for which all generalized Jacobians of $H_{\theta p}$ at a solution of the NCP are nonsingular.

Let $z^* = (\mu_*, x^*) \in \Re_+ \times \Re^n$ be a solution of $H_{\theta p}(z^*) = 0$. Then, $\mu_* = 0$ and x^* is a solution of (1.1). Associated to the solution x^*, we define three index sets:

$$\bar{\alpha} := \{i | x_i^* > 0\}, \ \bar{\beta} := \{i | x_i^* = f_i(x^*) = 0\}, \ \bar{\gamma} := \{i | f_i(x^*) > 0\}. \tag{2.7}$$

We say that the solution x^* is R-regular if $\nabla f_{\bar{\alpha}\bar{\alpha}}(x^*)$ is nonsingular and the Schur-complement of $\nabla f_{\bar{\alpha}\bar{\alpha}}(x^*)$ in

$$M = \begin{pmatrix} \nabla f_{\bar{\alpha}\bar{\alpha}}(x^*) & \nabla f_{\bar{\alpha}\bar{\beta}}(x^*) \\ \nabla f_{\bar{\beta}\bar{\alpha}}(x^*) & \nabla f_{\bar{\beta}\bar{\beta}}(x^*) \end{pmatrix} \tag{2.8}$$

is a P-matrix.

Proposition 2.3. *Suppose that $z^* = (\mu_*, x^*) \in \Re_+ \times \Re^n$ is a solution of $H_{\theta p}(z^*) = 0$, and x^* is an R-regular solution of the NCP, then all $V \in \partial H_{\theta p}(z^*)$ are nonsingular.*

Proof. From Proposition 2.1, it can be seen that for any $V \in \partial H_{\theta p}(z^*)$, there exists a $W(z^*) := (v_{\theta p}(z^*), E_{\theta p}(z^*)) \in \Re^n \times \Re^{n \times n}$ such that

$$V = \begin{bmatrix} 1 & 0 \\ v_{\theta p}(z^*) & E_{\theta p}(z^*) \end{bmatrix}, \tag{2.9}$$

where

$$E_{\theta p}(z^*) = \theta B(z^*) + (1-\theta)C(z^*) - I + \nabla f(x^*)[\theta A(z^*) - (1-\theta)C(z^*) - I], \tag{2.10}$$

with $\mu_* = 0$ and $A(z^*)$, $B(z^*)$, $C(z^*)$ being characterized as in Proposition 2.1. By the definition of V, it is easy to see that nonsingularity of V is equivalent to nonsingularity of $E_{\theta p}(z^*)$. With the expression of $\nabla f(x^*)$, $E_{\theta p}(z^*)$ can be written in the following partitioned form:

$$E_{\theta p}(z^*) = \begin{pmatrix} (\theta-2)\nabla f_{\bar{\alpha}\bar{\alpha}} & \nabla f_{\bar{\alpha}\bar{\beta}}\left[\theta A_{\bar{\beta}\bar{\beta}} - (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] & 0_{\bar{\alpha}\bar{\gamma}} \\ (\theta-2)\nabla f_{\bar{\beta}\bar{\alpha}} & \left[\theta B_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] + \delta_{\theta\bar{\beta}} & 0_{\bar{\beta}\bar{\gamma}} \\ (\theta-2)\nabla f_{\bar{\gamma}\bar{\alpha}} & \nabla f_{\bar{\gamma}\bar{\beta}}\left[\theta A_{\bar{\beta}\bar{\beta}} - (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] & (\theta-2)I_{\bar{\gamma}\bar{\gamma}} \end{pmatrix}.$$

where

$$\begin{aligned} &A_{\bar{\beta}\bar{\beta}} := A_{\bar{\beta}\bar{\beta}}(z^*), \quad B_{\bar{\beta}\bar{\beta}} := B_{\bar{\beta}\bar{\beta}}(z^*), \quad C_{\bar{\beta}\bar{\beta}} := C_{\bar{\beta}\bar{\beta}}(z^*), \\ &\delta_{\theta\bar{\beta}} = \nabla f_{\bar{\beta}\bar{\beta}}\left[\theta A_{\bar{\beta}\bar{\beta}} - (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right]. \end{aligned}$$

Thus, $E_{\theta p}(z^*)$ is nonsingular if and only if the matrix

$$U = \begin{pmatrix} (\theta-2)\nabla f_{\bar{\alpha}\bar{\alpha}} & \nabla f_{\bar{\alpha}\bar{\beta}}\left[\theta A_{\bar{\beta}\bar{\beta}} - (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] \\ (\theta-2)\nabla f_{\bar{\beta}\bar{\alpha}} & \left[\theta B_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] + \delta_{\theta\bar{\beta}} \end{pmatrix} \tag{2.11}$$

is nonsingular; which is equivalent to that if the zero vector is the unique solution of the system of equations

$$Uy = U\begin{pmatrix} y_{\bar{\alpha}} \\ y_{\bar{\beta}} \end{pmatrix} = 0. \tag{2.12}$$

It is easy to see that this system can be rewritten as

$$(\theta-2)\nabla f_{\bar{\alpha}\bar{\alpha}}y_{\bar{\alpha}} + \nabla f_{\bar{\alpha}\bar{\beta}}\left[\theta A_{\bar{\beta}\bar{\beta}} - (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right]y_{\bar{\beta}} = 0, \tag{2.13a}$$

$$(\theta-2)\nabla f_{\bar{\beta}\bar{\alpha}}y_{\bar{\alpha}} + \left(\left[\theta B_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] + \delta_{\theta\bar{\beta}}\right)y_{\bar{\beta}} = 0. \tag{2.13b}$$

Suppose that x^* is an R-regular solution of the NCP, then $\nabla f_{\bar{\alpha}\bar{\alpha}}$ is nonsingular. Thus, solving the first equation with respect to $y_{\bar{\alpha}}$ and substituting it into the second equation, we have

$$y_{\bar{\alpha}} = -((\theta-2)\nabla f_{\bar{\alpha}\bar{\alpha}})^{-1}\nabla f_{\bar{\alpha}\bar{\beta}}\left[\theta A_{\bar{\beta}\bar{\beta}} - (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right]y_{\bar{\beta}}, \tag{2.14a}$$

$$\begin{aligned} &\left(\nabla f_{\bar{\beta}\bar{\beta}} - \nabla f_{\bar{\beta}\bar{\alpha}}\nabla f_{\bar{\alpha}\bar{\alpha}}^{-1}\nabla f_{\bar{\alpha}\bar{\beta}}\right)\left[I_{\bar{\beta}\bar{\beta}} - \theta A_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}}\right]y_{\bar{\beta}} \\ &\quad = \left[\theta B_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right]y_{\bar{\beta}}, \end{aligned} \tag{2.14b}$$

where $\nabla f_{\bar{\beta}\bar{\beta}} - \nabla f_{\bar{\beta}\bar{\alpha}}\nabla f_{\bar{\alpha}\bar{\alpha}}^{-1}\nabla f_{\bar{\alpha}\bar{\beta}}$ is the Schur-complement of $\nabla f_{\bar{\alpha}\bar{\alpha}}$ in the matrix M defined by Eq. (2.8). Consequently, it is a P-matrix by the R-regularity assumption. Furthermore, showing the nonsingularity of U is equivalent to showing that the unique solution of the second equation of (2.14a) is the zero vector. We proceed by contradiction. Suppose that there exists a solution $y_{\bar{\beta}} \neq 0$, and consider the two cases:

(a) Suppose that $\left[I_{\bar{\beta}\bar{\beta}} - \theta A_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}}\right] y_{\bar{\beta}} = 0$. Define $\hat{\mathscr{I}} = \{i | y_{\bar{\beta}_i} \neq 0\}$. Then for any $i \in \hat{\mathscr{I}}$, $1 - \theta A_{ii} + (1-\theta)C_{ii} = 0$. We assume that $\theta B_{ii} + (1-\theta)C_{ii} - 1 = 0$ for the same index i. Then, it follows that $\theta(A_{ii} + B_{ii}) = 2$, this is a contradiction with $|A_{ii}| \leq 1, |B_{ii}| \leq 1$ and

$$\theta(|A_{ii}|^{p/(p-1)} + |B_{ii}|^{p/(p-1)}) + (1-\theta)|C_{ii}|^{p/(p-1)} \leq 1$$

from Proposition 2.1. So, $\theta B_{ii} + (1-\theta)C_{ii} - 1 \neq 0$, and hence,

$$\left[\theta B_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}} - I_{\bar{\beta}\bar{\beta}}\right] y_{\bar{\beta}} \neq 0.$$

This and the assumption lead to a contradiction to the second equation of (2.14a).

(b) Suppose that $\left[I_{\bar{\beta}\bar{\beta}} - \theta A_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}}\right] y_{\bar{\beta}} \neq 0$. Since for any $i \in \mathscr{I}$, $|A_{ii}| \leq 1, |B_{ii}| \leq 1, |C_{ii}| \leq 1$, $1 - \theta A_{ii} + (1-\theta)C_{ii}$ and $\theta B_{ii} + (1-\theta)C_{ii} - 1$ which are both nonzero (if any) have opposite signs. Thus,

$$\begin{aligned}&\left(\left[I_{\bar{\beta}\bar{\beta}} - \theta A_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}}\right] y_{\bar{\beta}}\right)_i \left(\left(\nabla f_{\bar{\beta}\bar{\beta}} - \nabla f_{\bar{\beta}\bar{\alpha}} \nabla f_{\bar{\alpha}\bar{\alpha}}^{-1} \nabla f_{\bar{\alpha}\bar{\beta}}\right)\right.\\ &\qquad \left.\left[I_{\bar{\beta}\bar{\beta}} - \theta A_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}}\right] y_{\bar{\beta}}\right)_i \leq 0. \end{aligned} \tag{2.15}$$

Since $\nabla f_{\bar{\beta}\bar{\beta}} - \nabla f_{\bar{\beta}\bar{\alpha}} \nabla f_{\bar{\alpha}\bar{\alpha}}^{-1} \nabla f_{\bar{\alpha}\bar{\beta}}$ is a P-matrix, we have

$$\left[I_{\bar{\beta}\bar{\beta}} - \theta A_{\bar{\beta}\bar{\beta}} + (1-\theta)C_{\bar{\beta}\bar{\beta}}\right] y_{\bar{\beta}} = 0,$$

which is a contradiction to the assumption. Therefore, the proof is complete. □

The following result is about the coerciveness of $H_{\theta p}(\cdot)$, which will be used in our analysis on the convergence of the algorithm.

Proposition 2.4. *Suppose that f is a continuously differentiable P_0-function. For any sequence $\{z^k\}$ satisfying that $\|z^k\| \to +\infty$ as $k \to +\infty$ and $\mu_k \in [\hat{\mu}, \tilde{\mu}]$, where $\hat{\mu}$ and $\tilde{\mu}$ are positive scalars with $\hat{\mu} < \tilde{\mu}$, it follows that $\|H_{\theta p}(z^k)\| \to +\infty$ as $k \to +\infty$.*

Proof. By the definition of $H_{\theta p}(\cdot)$, we only need to prove that $\lim_{k\to+\infty} \|\Phi_{\theta p}(z^k)\| = +\infty$. Suppose that the proposition is not true, then there exists a sequence $\{(\mu_k, x^k)\}$ such that

$$0 < \hat{\mu} \leq \mu_k \leq \tilde{\mu}, \ \|\Phi_{\theta p}(z^k)\| \leq c, \ \|x^k\| \to +\infty, \tag{2.16}$$

where $c > 0$ is certain constant. Since the sequence $\{x^k\}$ is unbounded, the index set $N := \{i \in \{1, \cdots, n\} : \{x_i^k\}$ is unbounded$\}$ is nonempty. Without loss of generality, we can assume that $\{|x_i^k|\} \to +\infty$ for any $i \in N$. Let the sequence $\{\hat{x}^k\}$ be defined by

$$\hat{x}_i^k = 0 \ \text{ if } \ i \in N \quad \text{and} \quad \hat{x}_i^k = x_i^k \ \text{ if } \ i \notin N. \tag{2.17}$$

Then, $\{\hat{x}^k\}$ is obviously bounded. Noting that f is a P_0-function, we have

$$0 \leq \max_{1\leq i\leq n \, x_i^k \neq \hat{x}_i^k} (x_i^k - \hat{x}_i^k)[f_i(x^k) - f_i(\hat{x}^k)] = x_{i_0}^k [f_{i_0}(x^k) - f_{i_0}(\hat{x}^k)], \tag{2.18}$$

where $i_0 \in N$ is one of the indices for which $\max_{i\in N} x_i^k [f_i(x^k) - f_i(\hat{x}^k)]$ is attained, and i_0 is assumed, without loss of generality, to be independent of k. Noting that $i_0 \in N$, it follows that $|x_{i_0}^k| \to +\infty$ as $k \to +\infty$. We now consider the following two cases:

Case 1 Suppose that $x_{i_0}^k \to +\infty$ as $k \to +\infty$. Since $\{f_{i_0}(\hat{x}^k)\}$ is bounded by the continuity of f_{i_0} and $\{\hat{x}^k\}$ is bounded, it follows from (2.18) that $\{f_{i_0}(x^k)\}$ is bounded below. Since $0 < \hat{\mu} < \mu_k < \tilde{\mu}$, we have $\mu_k x_{i_0}^k + f_{i_0}(x^k) \to +\infty$ and $x_{i_0}^k + \mu_k f_{i_0}(x^k) \to +\infty$. Hence, by [19, Proposition 2.4] and (1.5), we get $\|\Phi_{\theta p}(z^k)\| \to +\infty$ as $k \to +\infty$.

Case 2 Suppose that $x_{i_0}^k \to -\infty$ as $k \to +\infty$. Since $f_{i_0}(\hat{x}^k)$ is bounded, it follows from (2.18) that $f_{i_0}(x^k) \le f_{i_0}(\hat{x}^k)$ for a sufficiently large number $k \in \mathscr{K}$. Since $0 < \hat{\mu} < \mu_k < \tilde{\mu}$, we have $\mu_k x_{i_0}^k + f_{i_0}(x^k) \to -\infty$ and $x_{i_0}^k + \mu_k f_{i_0}(x^k) \to -\infty$. Hence, by [19, Proposition 2.4] and (1.5), we get $\|\Phi_{\theta p}(z^k)\| \to +\infty$ as $k \to +\infty$.

In either case, we obtain $\|\Phi_{\theta p}(z^k)\| \to +\infty$ as $k \to +\infty$, which is a contradiction to the boundedness of $\{\Psi_{\theta p}(z^k)\}$. This completes the proof. $\square$

Assumption 2.1. *The solution set* $S = \{x \in \Re^n : x \ge 0, f(x) \ge 0, x^T f(x) = 0\}$ *of the NCP* (1.1) *is nonempty and bounded.*

3. A Semismooth Newton Method

In this section, we propose a semismooth Newton method for solving $H_{\theta p}(z) = 0$, and give some basic results.

Algorithm 3.1. *(A semismooth Newton method with a non-monotone line search)*

Step 0 *Given any* $p \in (1, +\infty)$*, and choose* $\delta \in (0,1), \sigma \in (0,1/2), \theta \in [0,1], t \in [1/2,1]$*, and a positive integer* M*. Let* $z^0 := (\mu_0, x^0) \in \Re_{++} \times \Re^n$ *be an arbitrary vector. Choose* $\gamma \in (0,1)$ *such that* $\gamma\mu_0 < 1$*. Set* $e^0 := (1, 0, \cdots, 0) \in \Re^{1+n}$*,* $C_0 := \Psi_{\theta p}(z^0)$*,* $\beta_{\theta p}(z^{-1}) := \gamma$*, and* $Q_0 := 1$*. Choose* $\eta_0 \in [0,1]$ *and a sufficiently small positive number* ε*. Set* $m_0 := 1$ *and* $k := 0$*.*

Step 1 *If* $\Psi_{\theta p}(z^k) = 0$*, stop. Otherwise, let*

$$\beta_{\theta p}(z^k) := \min\{\gamma, \gamma\Psi_{\theta p}(z^k)^t, \beta_{\theta p}(z^{k-1})\}. \tag{3.1}$$

Step 2 *Compute* $\Delta z^k := (\Delta\mu_k, \Delta x^k) \in \Re \times \Re^n$ *by*

$$V\Delta z^k = -H_{\theta p}(z^k) + \mu_0\beta_{\theta p}(z^k)e_0, \tag{3.2}$$

where $V \in \partial H_{\theta p}(z^k)$*.*

Step 3 *Let* α_k *be the maximum of the values* $1, \delta, \delta^2, \cdots$ *such that*

$$\Psi_{\theta p}(z^k + \alpha_k\Delta z^k) \le C_k - 2\sigma(1 - \gamma\mu_0)\alpha_k\Psi_{\theta p}(z^k). \tag{3.3}$$

Step 4 *Set* $z^{k+1} := z^k + \alpha_k\Delta z^k$*, and* $m_k := \min\{k, M\}$*. If*

$$\sum_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^{k-i}) \leq \Sigma_{i=1}^{m_k-1}\eta_{k-i}\Psi_{\theta p}(z^k), or,\ \Psi_{\theta p}(z^k) < \varepsilon, \tag{3.4}$$

we set $\eta_k := 0$*; otherwise, we choose* $\eta_k \in (0,1]$*. Set*

$$C_k := \frac{\eta_k \sum_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^{k-i}) + \Psi_{\theta p}(z^k)}{Q_k}. \tag{3.5}$$

where $Q_k := 1 + \eta_k \sum_{i=1}^{m_k-1} \eta_{k-i}$*, and* $k := k+1$*, Go to step 1.*

In Algorithm 3.1, a non-monotone line search scheme is adopted. Such a non-monotone line search scheme were originally introduced by Hu, Huang, and Wang [14] in a smoothing Newton algorithm for the NCP; and by Hu, Huang, and Lu [32] in a descent algorithm for unconstrained optimization. It is easy to see that C_k is a convex combination of $\Psi_{\theta p}(z^0), \Psi_{\theta p}(z^1), \cdots, \Psi_{\theta p}(z^k)$. The choice of η_k controls the degree of the non-monotonicity. If $\eta_k = 0$ for all $k \in \mathscr{K}$, then the line search reduces to the usual monotone Armijo line search. It should be noted that Algorithm 3.1 is a non-monotone semismooth Newton method, which is different from those non-monotone algorithms mentioned above.

Lemma 3.1. *Suppose that* f *is a continuously differentiable* P_0*-function and the sequence* $\{z^k\}$ *is generated by Algorithm 3.1. Then the following results hold.*

(a) For all $k \in \mathscr{K}$*,* $\Psi_{\theta p}(z^k) \leq C_k$*.*

(b) The sequence $\{\beta_{\theta p}(z^k)\}$ *is monotonically decreasing.*

(c) Let $\mathscr{N} := \{(\mu, x) \in \Re_{++} \times \Re^n : \beta_{\theta p}(z)\mu_0 \leq \mu_0\}$*. Then* $z^k \in \mathscr{N}$ *for any* $k \in \mathscr{K}$ *and* $0 < \mu_{k+1} \leq \mu_k$*.*

(d) Algorithm 3.1 is well-defined.

Proof. (a) By Step 0, we know that $C_0 := \Psi_{\theta p}(z^0)$. In the following, we assume that $k > 0$. When

$$\sum_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^{k-i}) \leq \sum_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^k)$$

or $\Psi_{\theta p}(z^k) < \varepsilon$, we have $\eta_k := 0$, and hence, $C_k := \Psi_{\theta p}(z^k)$ by (3.5). Otherwise, i.e., when

$$\sum_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^{k-i}) > \sum_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^k),$$

by(3.5), we have that

$$\begin{aligned} C_k =& \frac{\eta_k \sum\limits_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^{k-i}) + \Psi_{\theta p}(z^k)}{Q_k} > \frac{\eta_k \sum\limits_{i=1}^{m_k-1} \eta_{k-i}\Psi_{\theta p}(z^k) + \Psi_{\theta p}(z^k)}{Q_k} \\ =& \frac{(\eta_k \sum\limits_{i=1}^{m_k-1} \eta_{k-i} + 1)\Psi_{\theta p}(z^k)}{Q_k} = \Psi_{\theta p}(z^k). \end{aligned} \tag{3.6}$$

Then, $\Psi_{\theta p}(z^k) \le C_k$ holds for any $k \in \mathscr{K}$.

(b) From the definition of $\beta_{\theta p}$ in (3.1), the result (b) is obviously satisfied.

(c) When $k = 0$,

$$\mu_0 \beta_{\theta p}(z^0) = \mu_0 \min\{\gamma, \gamma \Psi_{\theta p}(z^0)^t, \beta_{\theta p}(z^{-1})\} \le \mu_0 \beta_{\theta p}(z^{-1}) = \mu_0 \gamma < \mu_0,$$

i.e., $z^0 \in \mathscr{N}$ holds. Without lose of generality, suppose that $z^k \in \mathscr{N}$. Then,

$$\begin{aligned}
\mu_0 \beta_{\theta p}(z^{k+1}) - \mu_{k+1} &= \mu_0 \beta_{\theta p}(z^{k+1}) - (\mu_k + \alpha_k \Delta \mu_k) \\
&= \mu_0 \beta_{\theta p}(z^{k+1}) - [(1-\alpha_k)\mu_k + \alpha_k \mu_0 \beta_{\theta p}(z^k)] \\
&\le \mu_0 \beta_{\theta p}(z^{k+1}) - [(1-\alpha_k)\mu_0 \beta_{\theta p}(z^k) + \alpha_k \mu_0 \beta_{\theta p}(z^k)] \\
&= \mu_0 [\beta_{\theta p}(z^{k+1}) - \beta_{\theta p}(z^k)] \le 0.
\end{aligned} \tag{3.7}$$

That is, $z^{k+1} \in \mathscr{N}$. Therefore, $z^k \in \mathscr{N}$ holds for all $k \in \mathscr{K}$.

Next, we show that $0 < \mu_{k+1} \le \mu_k$ for any $k \in \mathscr{K}$. Obviously, $\mu_0 > 0$. We assume that $\mu_k > 0$. From the first equation in (3.2), we have $\Delta \mu_k = -\mu_k + \mu_0 \beta_{\theta p}(z^k)$. Thus,

$$\mu_{k+1} = \mu_k + \alpha_k \Delta \mu_k = (1-\alpha_k)\mu_k + \alpha_k \mu_0 \beta_{\theta p}(z^k) > 0, \tag{3.8a}$$

$$\mu_{k+1} = (1-\alpha_k)\mu_k + \alpha_k \mu_0 \beta_{\theta p}(z^k) \le (1-\alpha_k)\mu_k + \alpha_k \mu_k = \mu_k. \tag{3.8b}$$

Thus, the results in (c) hold.

(d) Firstly, from (c) we know that $\mu_k > 0$ for all $k \in \mathscr{K}$. Thus, it follows form Corollary 2.1 that the system of equations (3.2) is solvable. Secondly, we show that Step 3 is well defined. For any $z \in \Re_{++} \times \Re^n$, we denote $\varphi_{\theta p}(z) := \|\Phi_{\theta p}(z)\|^2$. For any $z^k \in \Re_{++} \times \Re^n$ with $\mu_k > 0$ and $k \in \mathscr{K}$, there exists a $W_k \in \partial \varphi_{\theta p}(z^k)$ such that

$$\begin{aligned}
&(\nabla \varphi_{\theta p}(z^k))^T \Delta z^k = 2(\Phi_{\theta p}(z^k))^T W_k \Delta z^k \\
&= -2(\Phi_{\theta p}(z^k))^T \Phi_{\theta p}(z^k) = -2\varphi_{\theta p}(z^k).
\end{aligned} \tag{3.9}$$

Since $t \in [1/2, 1]$, we can show that $\beta_{\theta p}(z^k) \le \gamma \Psi_{\theta p}(z^k)^{1/2}$. Thus, using $\mu_k + \alpha \Delta \mu_k = (1-\alpha)\mu_k + \alpha \mu_0 \beta_{\theta p}(z^k)$, we have

$$\begin{aligned}
(\mu_k + \alpha \Delta \mu_k)^2 &= (1-\alpha)^2 \mu_k^2 + 2(1-\alpha)\alpha \mu_0 \mu_k \beta_{\theta p}(z^k) + \mu_0^2 \beta_{\theta p}(z^k)^2 \alpha^2 \\
&\le (1-\alpha)^2 \mu_k^2 + 2\alpha \mu_0 \mu_k \gamma \Psi_{\theta p}(z^k)^{1/2} + O(\alpha^2) \\
&\le (1-\alpha)^2 \mu_k^2 + 2\alpha \mu_0 \gamma \|H_{\theta p}(z^k)\| \Psi_{\theta p}(z^k)^{1/2} + O(\alpha^2) \\
&= (1-\alpha)^2 \mu_k^2 + 2\alpha \mu_0 \gamma \Psi_{\theta p}(z^k) + O(\alpha^2).
\end{aligned} \tag{3.10}$$

Define $h(\alpha) = \varphi_{\theta p}(z^k + \alpha \Delta z^k) - \varphi_{\theta p}(z^k) - \alpha(\nabla \varphi_{\theta p}(z^k))^T \Delta z^k$. Since $\varphi_{\theta p}(\cdot)$ is continuously differentiable at any $z^k \in \Re^{n+1}$, we have $h(\alpha) = o(\alpha)$, and hence,

$$\begin{aligned}
&\|\Phi_{\theta p}(z^k + \alpha \Delta z^k)\|^2 \\
&= \varphi_{\theta p}(z^k + \alpha \Delta z^k) \\
&= \varphi_{\theta p}(z^k) + \alpha(\nabla \varphi_{\theta p}(z^k))^T \Delta z^k + h(\alpha) \\
&= \varphi_{\theta p}(z^k) - 2\alpha \varphi_{\theta p}(z^k) + o(\alpha) \quad (by\ (3.9)) \\
&= (1-2\alpha)\varphi_{\theta p}(z^k) + o(\alpha) \le (1-2\alpha)\Psi_{\theta p}(z^k) + o(\alpha).
\end{aligned} \tag{3.11}$$

Therefore, we obtain

$$\begin{aligned}\Psi_{\theta p}(z^k+\alpha\Delta z^k) &= (\mu_k+\alpha\Delta\mu_k)^2+\|\Phi_{\theta p}(z^k+\alpha\Delta z^k)\|^2\\ &\le (1-\alpha)^2\mu_k^2+2\alpha\mu_0\gamma\Psi_{\theta p}(z^k)+O(\alpha^2)+(1-2\alpha)\Psi_{\theta p}(z^k)+o(\alpha)\\ &\le \Psi_{\theta p}(z^k)-2(1-\gamma\mu_0)\alpha\Psi_{\theta p}(z^k)+o(\alpha)\\ &\le C_k-2(1-\gamma\mu_0)\alpha\Psi_{\theta p}(z^k)+o(\alpha),\end{aligned} \tag{3.12}$$

where the last inequality holds from the result (a). The above inequality implies that there exists $\bar{\alpha}\in(0,1]$ such that

$$\Psi_{\theta p}(z^k+\hat{\alpha}\Delta z^k)\le C_k-2\sigma(1-\gamma\mu_0)\hat{\alpha}\Psi_{\theta p}(z^k)$$

for all $\hat{\alpha}\in[0,\bar{\alpha}]$, which implies that Step 3 is well defined. Therefore, Algorithm 3.1 is well defined. □

4. Convergence of Algorithm 3.1

In this section, we will consider the global and local superlinear convergence of Algorithm 3.1.

Lemma 4.1. *Suppose that f is a P_0-function and Assumption 2.1 holds. Suppose that $\{\mu_k\}$ and $\{\nu_k\}$ are two infinite sequences such that for each $k\in\mathscr{K}$, $\mu_k>0,\nu_k\ge 0$ satisfying*

$$\lim_{k\to+\infty}\mu_k=0 \quad \textit{and} \quad \lim_{k\to+\infty}\nu_k=0. \tag{4.1}$$

For each $k\in\mathscr{K}$, let $x^k\in\mathfrak{R}^n$ satisfying

$$\sum_{i=1}^n\phi_{\theta p}^2(\mu_k,x_i^k,f_i(x^k))\le\nu_k. \tag{4.2}$$

Then $\{x^k\}$ remains bounded and every accumulation point of $\{x^k\}$ is a solution of the NCP (1.1).

Proof. The proof is similar to the one in [33, Theorem 5.4]. We omit its proof. □

Lemma 4.2. *Let $H_{\theta p}(\cdot)$ and $\beta_{\theta p}(\cdot)$ be defined by* (1.4) *and* (3.1)*, respectively. $\{z^k\}$ is the infinite iteration sequence generated by Algorithm 3.1. Then*

$$\lim_{k\to+\infty}\beta_{\theta p}(z^k)=0 \quad \textit{and} \quad \lim_{k\to+\infty}\Psi_{\theta p}(z^k)=0. \tag{4.3}$$

Proof. By Lemma 3.1 (b)(c), we obtain that sequences $\{\beta_{\theta p}(z^k)\}$ and $\{\mu_k\}$ are monotonically decreasing and $z^k\in\mathscr{N}$ for all $k\in\mathscr{K}$. It is not difficult to see that both $\{\mu_k\}$ and $\{\beta_{\theta p}(z^k)\}$ are convergent. We denote their limit points by μ_* and $\beta_{\theta p}^*$, respectively. Then we have $\mu_*\ge\beta_{\theta p}^*\ge 0$. Suppose that $\beta_{\theta p}^*\ne 0$, then $\mu_*\ge\beta_{\theta p}^*>0$. Since

$$0<\beta_{\theta p}(z^*)\mu_0\le\beta_{\theta p}(z^{k+1})\mu_0\le\beta_{\theta p}(z^k)\mu_0\le\mu_k\le\mu_{k-1}\le\cdots\le\mu_0, \tag{4.4}$$

by the definition of $\beta_{\theta p}(\cdot)$, we can obtain that the sequence $\{\Psi_{\theta p}(z^k)\}$ is bounded. Thus, by Proposition 2.4, we obtain that the sequence $\{z^k\}$ must be bounded. Without loss of generality, we assume that $\lim_{k\to\infty}z^k=z^*:=(\mu_*,x^*)$. Then,

$$\lim_{k\to+\infty}\Psi_{\theta p}(z^k)=\Psi_{\theta p}(z^*)>0 \quad \text{and} \quad \lim_{k\to+\infty}\beta_{\theta p}(z^k)=\beta_{\theta p}(z^*)>0. \tag{4.5}$$

Let α_k be the step-length at the k-th step iteration. We consider the following two cases.

Case 1 Suppose that $\alpha_k \geq d > 0$ for all k, where d is a constant. Then, by combining

$$\begin{aligned}\Psi_{\theta p}(z^k + \alpha_k \Delta z^k) \leq & C_k - 2\sigma(1-\gamma\mu_0)\alpha_k \Psi_{\theta p}(z^k) + o(\alpha) \\ \leq & C_k - 2\sigma(1-\gamma\mu_0) d \Psi_{\theta p}(z^k) + o(\alpha)\end{aligned} \tag{4.6}$$

with the fact that $\limsup_{k\to+\infty} C_k = \Psi_{\theta p}(z^*)$ by (3.5), we have

$$\Psi_{\theta p}(z^*) \leq \Psi_{\theta p}(z^*) - 2\sigma(1-\gamma\mu_0) d \Psi_{\theta p}(z^*), \tag{4.7}$$

i.e., $1 \leq 1 - 2\sigma(1-\gamma\mu_0)d$, which contradicts the fact that $d \in (0,1], 2\sigma \in (0,1)$, and $\gamma\mu_0 < 1$.

Case 2 Suppose that $\lim_{k\to+\infty} \alpha_k = 0$. Then, the stepsize $\hat{\alpha}_k := \alpha_k/\delta$ does not satisfy the line search criterion (3.3) for any sufficiently large k, i.e.

$$\begin{aligned}\Psi_{\theta p}(z^k + \hat{\alpha}_k \Delta z^k) > & C_k - 2\sigma(1-\gamma\mu_0)\hat{\alpha}_k \Psi_{\theta p}(z^k) \\ \geq & [1 - 2\sigma(1-\gamma\mu_0)\hat{\alpha}_k] \Psi_{\theta p}(z^k).\end{aligned} \tag{4.8}$$

Hence $[\Psi_{\theta p}(z^k + \hat{\alpha}_k \Delta z^k) - \Psi_{\theta p}(z^k)]/\hat{\alpha}_k > -2\sigma(1-\gamma\mu_0)\Psi_{\theta p}(z^k)$. Furthermore,

$$\begin{aligned}-2\sigma(1-\gamma\mu_0)\Psi_{\theta p}(z^*) \leq & 2H_{\theta p}(z^*)^T V \Delta z^* \\ = & 2H_{\theta p}(z^*)^T [-H_{\theta p}(z^*) + \mu_0 \beta_{\theta p}(z^*) e_0] \\ = & -2H_{\theta p}(z^*)^T H_{\theta p}(z^*) + 2\mu_0 \beta_{\theta p}(z^*) H_{\theta p}(z^*)^T e_0 \\ \leq & 2(-1+\gamma\mu_0)\Psi_{\theta p}(z^*),\end{aligned} \tag{4.9}$$

where in the second step we have used Eq. (3.2) and in the last step we have used Eq. (3.1). By $\Psi_{\theta p}(z^*) > 0$ and $\gamma\mu_0 < 1$, we have $-2\sigma(1-\gamma\mu_0) \leq -2(1-\gamma\mu_0)$ which implies $\sigma \geq 1$. This contradicts the fact that $\sigma \in (0, 1/2)$.

By combining **Case 1** with **Case 2**, we obtain $\beta_{\theta p}(z^*) = 0$, i.e., the first result of the lemma holds. This, together with the definition of the function $\beta_{\theta p}(\cdot)$, implies that there exists a subsequence $\{z^{k_n}\}$ such that the sequence $\{\Psi_{\theta p}(z^{k_n})\}$ converges to zero. Thus, by (3.4) given in Algorithm 3.1, we know that $\eta_k = 0$ for all sufficiently large k. Thus, $\Psi_{\theta p}(z^{k+1}) \leq \Psi_{\theta p}(z^k)$ for all sufficiently large k. Therefore, the sequence $\{\Psi_{\theta p}(z^k)\}$ converges to zero, i.e., the second result of the lemma holds. □

Combining Lemmas 4.1 and 4.2, we obtain the global convergence of Algorithm 3.1.

Theorem 4.1. *Suppose that f is a continuously differentiable P_0-function and the sequence $\{(\mu_k, x^k)\}$ is generated by Algorithm* 3.1. *If Assumption* 2.1 *is satisfied, then the infinite iteration sequence $\{(\mu_k, x^k)\}$ is bounded and every accumulation point (μ_*, x^*) of the sequence satisfies that $\mu_* = 0$ and x^* is a solution of the NCP* (1.1).

Now, we investigate the local superlinear (quadratic) convergence of Algorithm 3.1. By the definition of $\beta_{\theta p}(\cdot)$ in (3.1), we have that $\beta_{\theta p}(z^k) \leq \gamma \Psi_{\theta p}(z^k)^t$ for all $k \in \mathscr{K}$. By a similar way as the one in [25, Theorem 5.1], we obtain the superlinear (quadratic) convergence of Algorithm 3.1 as follows.

Theorem 4.2. *Assume that f is a continuously differentiable P_0-function and Assumption* 2.1 *is satisfied. Let $t = 1$. Suppose that $z^* := (\mu_*, x^*) \in \Re_+ \times \Re^n$ is an accumulation point of the infinte sequence*

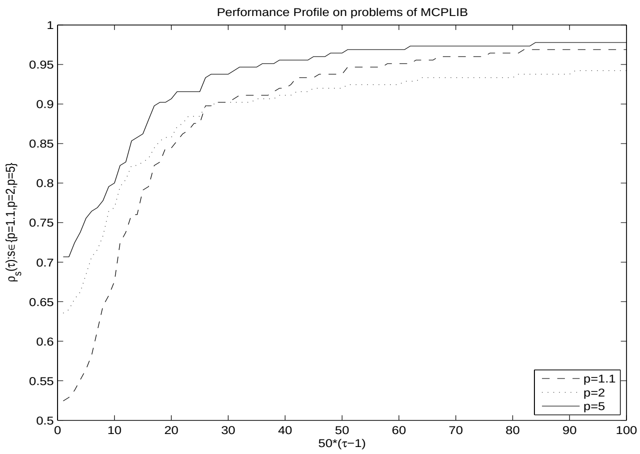

Figure 1: Performance profile for test problems in MCPLIB with the schemes: $p = 1.1$, $p = 2$, $p = 5$.

$\{z^k\}$ generated by Algorithm 3.1 *and all $V \in \partial H_{\theta p}(z^*)$ are nonsingular. Then the whole sequence $\{z^k\}$ converges to z^* with*

$$\|z^{k+1} - z^k\| = o(\|z^k - z^*\|), \quad \mu_{k+1} = o(\mu_k). \tag{4.10}$$

Furthermore, if ∇f is locally Lipschitz continuous around x^*, then

$$\|z^{k+1} - z^k\| = O(\|z^k - z^*\|^2), \quad \mu_{k+1} = O(\mu_k^2). \tag{4.11}$$

By Proposition 2.3, if the assumption that all $V \in \partial H_{\theta p}(z^*)$ are nonsingular is replaced by that the NCP is R-regular at x^*, then all conclusions of Theorem 4.2 hold.

5. Numerical Results

In this section, we implement Algorithm 3.1 for solving complementarity problems from MCPLIB [28] by the codes in Matlab. All experiments were done at an PC with CPU of 2.8 GHz and RAM of 2.00GB. In our computational experiments, the parameters used in the algorithm are chosen as follows:

$$\sigma = 10^{-4}, \quad \gamma = 0.02, \quad \delta = 0.5, \quad t = 0.75, \quad \mu_0 = 0.1, \quad M = 5, \quad \varepsilon = 10^{-6}.$$

And if (3.4) holds, set $\eta_k = 0$; otherwise, set $\eta_k = 0.85$. The starting points x_0 are taken according to those given in [28] and we use $\|H_{\theta p}(z^k)\| \leq 10^{-6}$ as the stopping rule. We test all most problems in MCPLIB [28] to see the numerical behavior of Algorithm 3.1 on three specific values of p, i.e., p=1.1, 2, 5; and five specific values of θ, i.e., $\theta = 0, 0.25, 0.5, 0.75, 1$.

To compare the performance profile of the cases: $p = 1.1$, $p = 2$, $p = 5$ with respect to the iterative number, we give some numerical analysis based on the performance profile proposed in [34]. Now, we give a brief introduction of this method. Let $\mathscr{S} := \{p = 1.1, p = 2, p = 5\}$, the schemes to be compared; $\mathscr{B}$ be the set of the 225 problems from MCPLIB; and $t_{b,s}$ be the number of iterations needed to solve problem b by scheme s. Then, we compute

$$\rho_s(\tau) := \frac{1}{225} \text{ size } \left\{ b \in \mathscr{B} : r_{b,s} \leq \tau \right\},$$

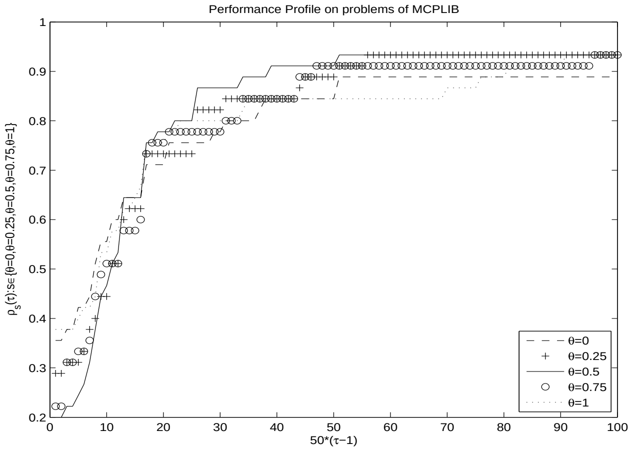

Figure 2: Performance profile for test problems in MCPLIB with the schemes: $\theta = 0$, $\theta = 0.25$, $\theta = 0.5$, $\theta = 0.75$, $\theta = 1$.

where $r_{b,s} := \frac{t_{b,s}}{\min\{t_{b,s}: s \in \mathscr{S}\}}$. $\rho_s(\tau)$ is the probability for scheme $s \in \mathscr{S}$ that a performance ratio $r_{b,s}$ is within a factor $\tau \in \mathfrak{R}$ of the possible ratio, and it is the distribution function for the performance ratio $r_{b,s}$. The analysis results are mapped in Fig. 1. It was pointed out in [34] that schemes with large probability $\rho_s(\tau)$ are to be preferred. From Fig. 1, we can see that Algorithm 3.1 works better for the scheme of $p = 5$ than the schemes of $p = 1.1$ and $p = 2$.

Then, we also compare the performance profile of the cases: $\theta = 0$, $\theta = 0.25$, $\theta = 0.5$, $\theta = 0.75$, $\theta = 1$ when $p = 5$. The analysis results are mapped in Fig. 2. From Fig. 2, we can see that Algorithm 3.1 in the case of $\theta \in [0.25, 0.75]$ is comparable to the cases of $\theta = 0$ and $\theta = 1$.

From the above numerical results, we see that the proposed method has better numerical performance in the case of $p = 5$ and $\theta \in [0.25, 075]$ than other cases.

In addition, it is easy to see that the regularized parameter μ introduced in $\phi_{\theta p}$ defined by (1.3)

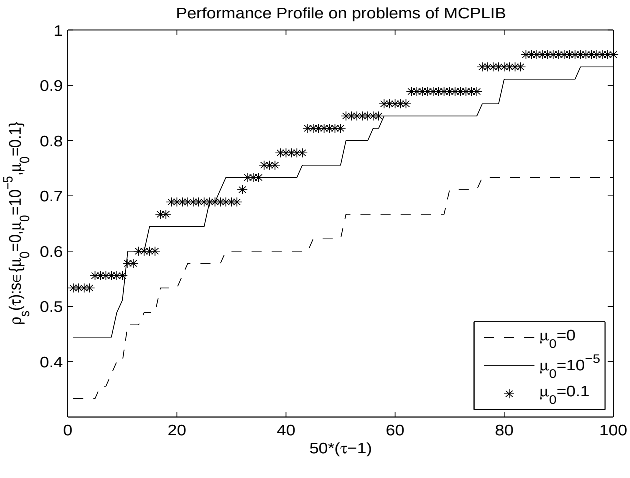

Figure 3: Performance profile for test problems in MCPLIB with the schemes: $\mu_0 = 0$, $\mu_0 = 10^{-5}$, $\mu_0 = 0.1$ when $\theta = 0.5$ and $p = 5$.

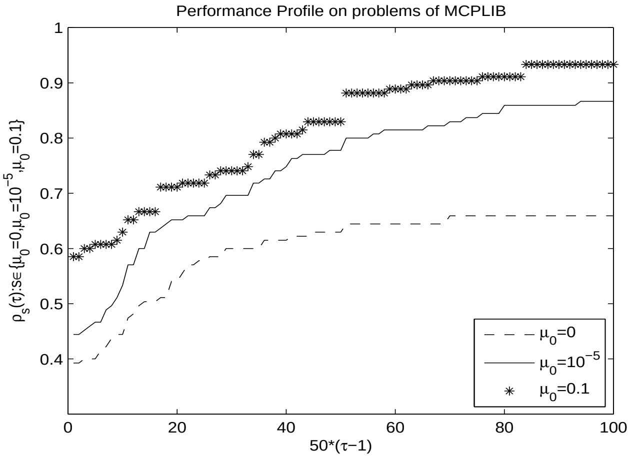

Figure 4: Performance profile for test problems in MCPLIB with the schemes: $\mu_0 = 0$, $\mu_0 = 10^{-5}$, $\mu_0 = 0.1$ when $\theta = 0, 0.5, 1$ and $p = 5$.

plays an important role in the proof of the global convergence of Algorithm 3.1. A natural question is *How the parameter μ affect the implementation of Algorithm 3.1 depends on the initial value of μ, that is* μ_0. In the following, we give the performance profile of the testings of Algorithm 3.1 for the same problems as above through choosing the parameter μ_0 in the cases of $\mu_0 = 0$, $\mu_0 = 10^{-5}$, and $\mu_0 = 0.1$, respectively, where for the case of $p = 5$ and $\theta = 0.5$, the numerical results are mapped to Fig. 3; and for the case of $p = 5$ and $\theta = 0$, 0.5, 1.0, the numerical results are mapped to Fig. 4. From Figs. 3 and 4, we can get that Algorithm 3.1 works best for the case of $\mu_0 = 0.1$ and worst for the case of $\mu_0 = 0$. This demonstrates that it is helpful for the numerical computation to introduce the regularized parameter μ into the NCP function $\omega_{\theta p}$ defined by (1.2). We have also tested some other cases, and find that if μ_0 is too large, the number of iterations becomes large as μ_k converges to zero. Thus, from the view of computation, it is not suitable for Algorithm 3.1 to choose large μ_0.

6. Conclusions

Based on a symmetrically perturbed function of the generalized NCP-function in [19], we proposed a regularized semismooth Newton method with a non-monotone line search for solving the P_0-NCP. We showed that the proposed method is globally and locally superlinearly (quadratically) convergent under suitable assumptions. We also reported some numerical results, which demonstrate the proposed method is effective for solving the problems from MCPLIB. Numerical experiments indicate that the proposed method has better numerical performance in the case of $p = 5$ and $\theta \in [0.25, 075]$ than other cases; while the case of $p = 5$ and $\theta \in [0.25, 075]$ is not contained in those known cases given in literature. In addition, the method has better numerical performance in the case of $\mu_0 = 0.1$ than the other cases. Thus, it is valuable to investigate the symmetrically perturbed function $\phi_{\theta p}$ defined by (1.3) and the semismooth Newton method for the P_0-NCP. It is interesting whether the function $\phi_{\theta p}$ can be extended to the case of symmetric cones or not; and by which some methods can be designed to solving the symmetric cone complementarity problem.

References

[1] M.C. FERRIS, J.S. PANG, *Engineering and economic applications of complementarity problems*, SIAM Rev., vol. 39 (1997), pp. 669–713.

[2] P.T. HARKER, J.S. PANG, *Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications*, Math. Program. , vol. 48 (1990), pp. 161–220.

[3] F. FACCHINEI, J.S. PANG, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Verlag, New York, 2003.

[4] J.Y. HAN, N.H. XIU, H.D. QI, *Theory and methods for nonlinear complementarity problems*, Shanghai Science and Technology Press. 2006(in Chinese).

[5] B. CHEN, X. CHEN, C. KANZOW, *A penalized Fischer-Burmeister NCP-function*, Math. Program., vol. 88 (2000), pp. 211–216.

[6] F. FACCHINEI, C. KANZOW, *On unconstrained and constrained stationary points of the implicit Lagrangian*, J. Optim. Theory Appl., vol. 92 (1997), pp. 99–115.

[7] F. FACCHINEI, J. SOARES, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM J. Optim., vol. 7 (1997), pp. 225-247.

[8] B. CHEN, N.H. XIU, *A global linear and local quadratic non-interior continuation method for NCP based on Chen-Mangasarian smoothing functions*, SIAM J. Optim. , vol.9 (1999), pp. 605–623.

[9] B. CHEN, N.H. XIU, *A superlinear noninterior one-step continuation method for monotone LCP in the absence of strict complementarity*, J. Optim. Theory Appl., vol. 108 (2001), pp. 317–332.

[10] A. FISCHER, *A special Newton-type optimization method*, Optim., vol. 24 (1992), pp. 269–284.

[11] L.P. ZHANG, S.Y. WU, T.R. GAO, *Improved smoothing Newton methods for P_0 nonlinear complementarity problems*, Appl. Math. Comput., vol. 215 (2009), pp. 324–332.

[12] C. GEIGER, C. KANZOW, *On the resolution of monotone complementarity problems*, Comput. Optim. Appl., vol. 5 (1996), pp. 155–173.

[13] Z.H. HUANG, L.Q. QI, D.F. SUN, *Sub-quadratric convergence of a smoothing Newton algorithm for the P_0-and monotone LCP*, Math. Program., vol. 99 (2004), pp. 423–441.

[14] S.L. HU, Z.H. HUANG, P. WANG, *A non-monotone smoothing Newton algorithm for solving nonlinear complementarity problems*, Optim. Methods Softw., vol.24 (2009), pp. 447–460.

[15] C. KANZOW, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., vol. 17 (1996), pp. 851–868.

[16] C. KANZOW, H. KLEINMICHEL, *A new class of semismooth Newton method for nonlinear complementarity problems*, Comput. Optim. Appl., vol.11 (1998), pp. 227-251.

[17] C. KANZOW, N. YAMASHITA, M. FUKUSHIMA, *New NCP-functions and their properties*, J. Optim. Theory Appl., vol.94 (1997), pp. 115–135.

[18] L.Q. QI, D.F. SUN, G.L. ZHOU, *A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequality problems*, Math. Program., vol. 87 (2000), pp. 1–35.

[19] 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., vol. 230 (2009), pp. 69–82.

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

[21] J.-S. CHEN, S. PAN, *A family of NCP-functions and a descent method for the nonlinear complementarity problem*, Comput. Optim. Appl., vol. 40 (2008), pp. 389–404.

[22] Z.H. HUANG, J. HAN, D.C. XU, L.P. ZHANG, *The non-interior continuation methods for solving the P_0 function nonlinear complementarity problem*, Science in China, vol.44 (2001), pp. 1107-

1114.
[23] Z.H. Huang, J. Han, Z. Chen, *A predictor-corrector smoothing Newton algorithm, based on a new smoothing function, for solving the nonlinear complementarity problem with a P_0 function*, J. Optim. Theory Appl., vol. 117 (2003), pp. 39–68.
[24] J.-S. Chen, S. Pan, *A regularization semismooth Newton method based on the generalized Fischer-Burmeister function for P0-NCPs*, J. Comput. Appl. Math., vol. 220 (2008), pp. 464–479.
[25] D.F. Sun, A regularization Newton method for solving nonlinear complementarity problems, Appl. Math. Optim., vol. 40 (1999), pp. 315–339.
[26] Z.H. Huang, *Sufficient conditions on nonemptiness and boundedness of the solution set of the P_0 function nonlinear complementarity problem*, Oper. Res. Lett., vol. 30 (2002), pp. 202–210.
[27] Y. Zhang, Z.H. Huang, *A nonmonotone smoothing-type algorithm for solving a system of equalities and inequalities*, J. Comput. Appl. Math., vol. 233 (2010), pp. 2312–2321.
[28] S.C. Billups, S.P. Dirkse, M.C. Ferris, *A comparison of algorithms for large scale mixed complementarity problems*, Comput. Optim. Appl., vol. 7 (1997), pp. 3–25.
[29] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., vol. 15 (1977), pp. 957–972.
[30] L.Q. Qi, J. Sun, *A nonsmooth version of Newton's method*, Math. Program., vol. 58 (1993), pp. 353–367.
[31] F.H. Clarke, *Optimization and Nonsmooth Analysis,* Wiley, New York, 1983.
[32] S.L. Hu, Z.H. Huang, N. Lu, *A non-monotone line search algorithm for unconstrained optimization*, J. Sci. Comput., vol.42 (2010), pp. 38–53.
[33] F. Facchinei, C. Kanzow, *Beyond monotonicity in regularization methods for the nonlinear complementarity problems*, SIAM J. Optim., vol. 37 (1999), pp. 1150–1161.
[34] E.D. Dolan, J.J. Moré, *Benchmarking optimization software with performance profiles*, Math. Program., vol. 91 (2002), pp. 201–213.