A New Derivative-Free Descent Method for the Nonlinear Complementarity Problem

Kenjiro Yamada[1], Nobuo Yamashita[2] and Masao Fukushima[2]

July 22, 1998

Abstract. Recently, much effort has been made in solving and analyzing the nonlinear complementarity problem (NCP) by means of a reformulation of the problem as an equivalent unconstrained optimization problem involving a merit function. In this paper, we propose a new merit function for the NCP and show several favorable properties of the proposed function. In particular, we give conditions under which the function provides a global error bound for the NCP and conditions under which its level sets are bounded. Moreover, we propose a new derivative-free descent algorithm for solving the NCP based on this function. We show that any accumulation point generated by the algorithm is a solution of the NCP under the monotonicity assumption on the problem. Also, we prove that the sequence generated by the algorithm converges linearly to the solution under the strong monotonicity assumption. The most interesting feature of the algorithm is that the search direction and the stepsize are adjusted simultaneously during the linesearch process, whereas a fixed search direction is used in the linesearch process in earlier derivative-free algorithms proposed by Geiger and Kanzow, Jiang, Mangasarian and Solodov, and Yamashita and Fukushima. Making use of this particular feature, we can establish the linear convergence of the algorithm under more practical assumptions compared with the linearly convergent derivative-free algorithm recently proposed by Mangasarian and Solodov.

Keywords. Derivative-free method, nonlinear complementarity problem, merit function, global convergence, linear convergence.

[1]IBM Japan, Ltd., 19-21, Nihonbashi Hakozaki-cho, Chuo-ku, Tokyo 103-8510 Japan. e-mail: ken2ro@jp.ibm.com

[2]Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. e-mail: {nobuo,fuku}@kuamp.kyoto-u.ac.jp

1. Introduction

The nonlinear complementarity problem (NCP) [11, 20] is to find a point $x \in R^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0, \tag{1}$$

where $F : R^n \to R^n$ is a given function. Throughout this paper, we assume that F is continuously differentiable. Applications of the NCP can be found in many important fields such as mathematical programming, economics, engineering and mechanics [2, 11].

A variety of methods have been proposed for solving the NCP [9, 11, 20]. One of the powerful approaches for solving the NCP that have been studied intensively in recent years is to reformulate the NCP as an equivalent system of nonlinear equations [16, 21, 25] or as an equivalent unconstrained optimization problem [3, 5, 6, 10, 12, 14, 18, 22, 24]. The objective function of such an equivalent unconstrained optimization problem is called a merit function. In other words, a merit function for the NCP is a function whose global minima are coincident with the solutions of the NCP.

To construct a merit function, the class of functions called NCP-functions have played an important role. A function $\phi : R^2 \to R$ is called an NCP-function if it satisfies

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

An NCP-function ϕ is said to be nonnegative if

$$\phi(a,b) \geq 0 \quad \text{for all } (a,b)^T \in R^2.$$

It is clear that if ϕ is a nonnegative NCP-function on R^2, then the function $\Psi : R^n \to R$ defined by

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

is a merit function for the NCP. In this paper, we are particularly interested in the following merit functions constructed by some well-known NCP-functions. The natural residual $\Psi_{NR} : R^n \to R^n$ [11] is a merit function defined by

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

where $\phi_{NR} : R^2 \to R$ is a nonnegative NCP-function given by

$$\phi_{NR}(a,b) = \min\{a,b\}^2.$$

Note that the function ϕ_{NR} is not differentiable everywhere, hence neither is Ψ_{NR}. So, the equivalent unconstrained optimization problem of minimizing the function Ψ_{NR} cannot be solved by using general algorithms that require the differentiability of the objective function. Mangasarian and Solodov [17] proposed the implicit Lagrangian $\Psi_{MS} : R^n \to R$ defined by

$$\Psi_{MS}(x) = \sum_{i=1}^{n} \phi_{MS}(x_i, F_i(x)),$$

where $\phi_{MS} : R^2 \to R$ is a nonnegative NCP-function given by

$$\phi_{MS}(a,b) = ab + \frac{1}{2\alpha}\left\{(-\alpha b + a)_+^2 - a^2 + (-\alpha a + b)_+^2 - b^2\right\},$$

where $\alpha > 1$ is a fixed constant and $(z)_+ = \max\{0, z\}$ for $z \in R$. The function ϕ_{MS} is continuously differentiable, and so is Ψ_{MS}. Any stationary point of Ψ_{MS} is a solution of the NCP whenever F is a P-function [4] (see also [24]). Another well-known merit function is the function $\Psi_{FB} : R^n \to R$ defined by

$$\Psi_{FB}(x) = \sum_{i=1}^{n} \phi_{FB}(x_i, F_i(x)), \tag{3}$$

where $\phi_{FB} : R^2 \to R$ is given by

$$\phi_{FB}(a,b) = \frac{1}{2}(\sqrt{a^2 + b^2} - a - b)^2. \tag{4}$$

The function ϕ_{FB} is a nonnegative NCP-function introduced by Fischer [6] and called the Fischer-Burmeister function. This function has been extensively studied and shown to have a number of favorable properties. For example, ϕ_{FB} is continuously differentiable on R^2, and hence Ψ_{FB} is also continuously differentiable. Moreover, any stationary point of Ψ_{FB} is a solution of the NCP whenever F is a P_0-function [5]. Note that this condition is weaker than the one required for the implicit Lagrangian to have the same property.

For solving an equivalent unconstrained optimization problem based on a merit function, extensions of classical iterative methods such as Newton's method have been proposed. Recently, so-called derivative-free methods have attracted much attention, which do not require computation of derivatives of F and make use of the particular properties of a merit function [8, 10, 12, 18, 24, 27]. Derivative-free methods are particularly suitable for problems where the derivatives of F are not available or are extremely expensive to compute.

For the variational inequality problem (VIP), which contains the NCP as a special case, Fukushima [8] proposed a derivative-free descent method based on the regularized gap function and showed its global convergence under the strong monotonicity assumption on the problem mapping. More recently, Yamashita, Taji and Fukushima [27] proposed a derivative-free descent method for solving the VIP based on the D-gap function and established global convergence of the method under similar assumptions.

For the NCP, Yamashita and Fukushima [24] proposed a derivative-free descent method based on the implicit Lagrangian Ψ_{MS}. They showed that their method is globally convergent under the strong monotonicity assumption on F. Jiang [12] proposed a derivative-free descent method based on the Fischer-Burmeister function. He showed that his method converges globally to a solution of the NCP if F is monotone. Geiger and Kanzow [10] proposed a similar method and proved its global convergence under somewhat stronger assumptions than [12]. Note however that convergence rate of the proposed methods has not been discussed for the above mentioned methods for the VIP and the NCP. Recently, Mangasarian and Solodov [18] slightly modified the derivative-free descent algorithm of [24] and showed that the sequence generated by the modified algorithm converges linearly to the solution of the NCP. However, for their algorithm to have a global convergence

property, the strong monotonicity assumption on F is needed. Moreover, in order to ensure the linear rate of convergence, some parameters used in the algorithm have to be specified by using usually unknown constants such as the modulus of strong monotonicity of F.

In this paper, we propose the following merit function $\Psi_\alpha : R^n \to R$ for the NCP :

$$\Psi_\alpha(x) = \sum_{i=1}^{n} \phi_\alpha(x_i, F_i(x)), \tag{5}$$

where $\phi_\alpha : R^2 \to R$ is a nonnegative NCP-function given by

$$\phi_\alpha(a,b) = \frac{\alpha}{2}(ab)_+^2 + \frac{1}{2}(\sqrt{a^2+b^2} - a - b)^2 \tag{6}$$

and $\alpha \geq 0$ is a real parameter. The function Ψ_α is a slight modification of the function proposed in [26] (see also [1]). When $\alpha = 0$, the function ϕ_α reduces to the Fischer-Burmeister function ϕ_{FB}, so ϕ_α is considered an extension of the Fischer-Burmeister function. We show that ϕ_α inherits several favorable properties of the Fischer-Burmeister function. In addition, when $\alpha > 0$, we can show that ϕ_α enjoys an important property that the Fischer-Burmeister function does not have. In particular, we give conditions under which Ψ_α provides a global error bound for the NCP and conditions under which it has bounded level sets. Moreover, we propose a new derivative-free descent algorithm for solving the NCP based on the function Ψ_α. We prove that this algorithm has a global convergence property if F is a monotone function, and that the sequence generated by the algorithm converges linearly to the solution if F is a strongly monotone function. In contrast with the algorithm proposed in [18], proper values of the parameters used in our algorithm do not rely on usually unknown constants such as the modulus of strong monotonicity of F.

This paper is organized as follows. In Section 2, we review some definitions and preliminary results that will be used in the subsequent analyses. In Section 3, we show some important properties of the proposed merit function. In Section 4, we present a derivative-free descent algorithm for solving the NCP and prove global convergence of the algorithm, while in Section 5, we prove its linear convergence. Section 6 makes some concluding remarks.

Throughout the paper, we shall adopt the following notations: For a real-valued matrix A of any dimension, A^T denotes its transpose. For a differentiable vector-valued function $F : R^n \to R^n$, ∇F denotes the transposed $n \times n$ Jacobian matrix whose columns are the gradients of the components of F and the ith component function of F is denoted by F_i. The norm $\|\cdot\|$ denotes the Euclidean norm. For $z \in R$, $\lceil z \rceil$ denotes the smallest integer no less than z. The level set of a function $\Psi : R^n \to R$ is denoted $\mathcal{L}(\Psi, c) := \{x \in R^n \mid \Psi(x) \leq c\}$.

2. Preliminaries

In this section, we recall some definitions and preliminary results that will be used in the subsequent analyses.

The following classes of functions $F : R^n \to R^n$ play an important role in the NCP.

Definition 2.1 *Let S be a convex set in R^n.*

(a) *F is monotone on S if*

$$(x - x')^T(F(x) - F(x')) \geq 0 \quad \textit{for all} \quad x, x' \in S.$$

(b) *F is a P_0-function on S if*

$$\max_{\substack{1 \leq i \leq n \\ x_i \neq x'_i}} (x_i - x'_i)(F_i(x) - F_i(x')) \geq 0 \quad \textit{for all} \quad x, x' \in S, \; x \neq x'.$$

(c) *F is a P-function on S if*

$$\max_{1 \leq i \leq n} (x_i - x'_i)(F_i(x) - F_i(x')) > 0 \quad \textit{for all} \quad x, x' \in S, \; x \neq x'.$$

(d) *F is strongly monotone with modulus $\mu > 0$ on S if*

$$(x - x')^T(F(x) - F(x')) \geq \mu \|x - x'\|^2 \quad \textit{for all} \quad x, x' \in S.$$

(e) *F is a uniform P-function with modulus $\mu > 0$ on S if*

$$\max_{1 \leq i \leq n} (x_i - x'_i)(F_i(x) - F_i(x')) \geq \mu \|x - x'\|^2 \quad \textit{for all} \quad x, x' \in S.$$

(f) *$\nabla F(x)$ is uniformly positive definite with modulus $\mu > 0$ on S if*

$$d^T \nabla F(x) d \geq \mu \|d\|^2 \quad \textit{for all} \quad x \in S, \; d \in R^n.$$

(g) *F is Lipschitz continuous on S if there exists a constant $L > 0$ such that*

$$\|F(x) - F(x')\| \leq L \|x - x'\| \quad \textit{for all} \quad x, x' \in S.$$

□

When the properties mentioned in Definition 2.1 hold with $S = R^n$, we simply omit the words "on S". Note that F is a P_0-function on S if F is monotone on S, and that F is a uniform P-function with modulus $\mu > 0$ on S if F is strongly monotone with modulus $\mu > 0$ on S. Also, when F is continuously differentiable, F is strongly monotone with modulus $\mu > 0$ on S if and only if $\nabla F(x)$ is uniformly positive definite with modulus $\mu > 0$ on S.

We recall some concepts about the NCP which will be used when we discuss boundedness of the solution set of the NCP.

Definition 2.2 *The NCP is said to be strictly feasible if there exists $\hat{x} > 0$ such that $F(\hat{x}) > 0$.* □

The error bound is an important concept that indicates how close an arbitrary point is to the solution set of the NCP. Thus, an error bound may be used to provide stopping rules for an iterative method. The following lemma shows that the natural residual provides a global error bound for the NCP under suitable conditions, which will play an important role in obtaining the linear convergence result of the algorithm to be presented in Section 4.

Lemma 2.1 *([19]) Let Ψ_{NR} be defined by (2). Suppose that F is strongly monotone with modulus $\mu > 0$ and Lipschitz continuous with constant $L > 0$ on some convex set $S \subseteq R^n$. Let x^* be the unique solution of the NCP. Then the following inequality holds:*

$$\|x - x^*\| \leq \frac{L+1}{\mu}\Psi_{NR}(x)^{\frac{1}{2}} \quad \textit{for all} \quad x \in S.$$

□

In the next lemma, we summarize some important properties of the Fischer-Burmeister function ϕ_{FB} defined by (4).

Lemma 2.2 *([10, 13, 23]) The function ϕ_{FB} defined by (4) has the following properties:*

(i) $\phi_{FB}(a,b) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0.$

(ii) $\phi_{FB}(a,b) \geq 0$ *for all* $(a,b)^T \in R^2$.

(iii) ϕ_{FB} *is continuously differentiable for all* $(a,b)^T \in R^2$.

(iv) $\dfrac{\partial \phi_{FB}}{\partial a}(a,b)\dfrac{\partial \phi_{FB}}{\partial b}(a,b) \geq 0$ *at any* $(a,b)^T \in R^2$.

(v) $\dfrac{\partial \phi_{FB}}{\partial a}(a,b) = 0 \Longleftrightarrow \dfrac{\partial \phi_{FB}}{\partial b}(a,b) = 0 \Longleftrightarrow \phi_{FB}(a,b) = 0.$

(vi) $\dfrac{1}{2}\left(\dfrac{2}{\sqrt{2}+2}\right)^2 \min\{a,b\}^2 \leq \phi_{FB}(a,b) \leq \dfrac{1}{2}(\sqrt{2}+2)^2 \min\{a,b\}^2$ *for all* $(a,b)^T \in R^2$. □

Lemma 2.2 (i) and (v) mean that ϕ_{FB}, $\frac{\partial \phi_{FB}}{\partial a}$ and $\frac{\partial \phi_{FB}}{\partial b}$ are all NCP-functions. Moreover, the above lemma indicates that the function Ψ_{FB} defined by (3) is nonnegative on R^n and that $\Psi_{FB}(x) = 0$ if and only if x solves the NCP. Important results about the merit function Ψ_{FB} are summarized in the following theorem.

Theorem 2.1 *([10, 23]) The function Ψ_{FB} defined by (3) has the following properties:*

(i) *If the NCP has at least one solution, then the NCP is equivalent to the unconstrained optimization problem* $\min_{x \in R^n} \Psi_{FB}(x)$.

(ii) *If F is a P_0-function, any stationary point of Ψ_{FB} is a solution of the NCP.*

(iii) *Let Ψ_{NR} be the natural residual defined by (2). Then the following inequalities hold:*

$$\frac{1}{2}\left(\frac{2}{\sqrt{2}+2}\right)^2 \Psi_{NR}(x) \leq \Psi_{FB}(x) \leq \frac{1}{2}(\sqrt{2}+2)^2 \Psi_{NR}(x) \quad \textit{for all} \quad x \in R^n. \tag{7}$$

Moreover, if F is strongly monotone and Lipschitz continuous, then Ψ_{FB} provides a global error bound for the NCP.

(iv) *If F is a uniform P-function, then the level set $\mathcal{L}(\Psi_{FB}, c)$ is bounded for all $c \in R$.* □

3. A New Merit Function and Its Properties

In this section, we propose a new merit function and show that the function has several favorable properties.

Let $\Psi_\alpha : R^n \to R$ be defined by (5) and (6), i.e.,

$$\Psi_\alpha(x) = \sum_{i=1}^{n} \phi_\alpha(x_i, F_i(x)),$$

where $\phi_\alpha : R^2 \to R$ is given by

$$\phi_\alpha(a,b) = \frac{\alpha}{2}(ab)_+^2 + \frac{1}{2}(\sqrt{a^2+b^2} - a - b)^2$$

and $\alpha \geq 0$ is a real parameter. Since the function ϕ_0 reduces to the Fischer-Burmeister function ϕ_{FB}, the function ϕ_α inherits several favorable properties of the Fischer-Burmeister function. Moreover, when $\alpha > 0$, ϕ_α enjoys an important property the Fischer-Burmeister function ϕ_{FB} does not have (see Lemma 3.1 (vi) below).

Lemma 3.1 *Let $\alpha \geq 0$. The function ϕ_α defined by (6) has the following properties:*

(i) $\phi_\alpha(a,b) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0.$

(ii) $\phi_\alpha(a,b) \geq 0$ *for all* $(a,b)^T \in R^2$.

(iii) ϕ_α *is continuously differentiable at any* $(a,b)^T \in R^2$ *and*

$$\frac{\partial \phi_\alpha}{\partial a}(0,0) = \frac{\partial \phi_\alpha}{\partial b}(0,0) = 0, \tag{8}$$

$$\frac{\partial \phi_\alpha}{\partial a}(a,b) = \alpha b(ab)_+ + \left(\frac{a}{\sqrt{a^2+b^2}} - 1\right)(\sqrt{a^2+b^2} - a - b) \quad \text{for } (a,b) \neq (0,0), \tag{9}$$

$$\frac{\partial \phi_\alpha}{\partial b}(a,b) = \alpha a(ab)_+ + \left(\frac{b}{\sqrt{a^2+b^2}} - 1\right)(\sqrt{a^2+b^2} - a - b) \quad \text{for } (a,b) \neq (0,0). \tag{10}$$

(iv) $\dfrac{\partial \phi_\alpha}{\partial a}(a,b)\dfrac{\partial \phi_\alpha}{\partial b}(a,b) \geq 0$ *for all* $(a,b)^T \in R^2$.

(v) $\dfrac{\partial \phi_\alpha}{\partial a}(a,b) = 0 \Longleftrightarrow \dfrac{\partial \phi_\alpha}{\partial b}(a,b) = 0 \Longleftrightarrow \phi_\alpha(a,b) = 0.$

If $\alpha > 0$, then the function ϕ_α enjoys the following additional property:

(vi) *If $a \to -\infty$ or $b \to -\infty$ or $ab \to \infty$, then $\phi_\alpha(a,b) \to \infty$.*

Proof. Since (i)–(iii) and (vi) can easily be shown, we only show (iv) and (v). We first prove (iv). When $(a,b) = (0,0)$, $\frac{\partial \phi_\alpha}{\partial a}(a,b)\frac{\partial \phi_\alpha}{\partial b}(a,b) \geq 0$ holds obviously. When $(a,b) \neq (0,0)$, it follows from (9) and (10) that

$$\begin{aligned}\frac{\partial \phi_\alpha}{\partial a}(a,b)\frac{\partial \phi_\alpha}{\partial b}(a,b) &= \alpha^2 ab(ab)_+^2 + \alpha a(ab)_+ \left(\frac{a}{\sqrt{a^2+b^2}} - 1\right)(\sqrt{a^2+b^2} - a - b) \\ &\quad + \alpha b(ab)_+ \left(\frac{b}{\sqrt{a^2+b^2}} - 1\right)(\sqrt{a^2+b^2} - a - b) \\ &\quad + \left(\frac{a}{\sqrt{a^2+b^2}} - 1\right)\left(\frac{b}{\sqrt{a^2+b^2}} - 1\right)(\sqrt{a^2+b^2} - a - b)^2. \end{aligned} \tag{11}$$

Since

$$ab(ab)_+^2 \geq 0, \quad \frac{a}{\sqrt{a^2+b^2}} - 1 \leq 0, \quad \frac{b}{\sqrt{a^2+b^2}} - 1 \leq 0, \tag{12}$$

it is sufficient to show that the second and third terms of (11) are both nonnegative for all $(a,b)^T \neq (0,0)^T \in R^2$. We first show

$$\alpha a(ab)_+ \left(\frac{a}{\sqrt{a^2+b^2}} - 1 \right) (\sqrt{a^2+b^2} - a - b) \geq 0. \tag{13}$$

Consider three cases: $a = 0$, $a < 0$ and $a > 0$. When $a = 0$, (13) holds obviously. When $a < 0$, we have $\sqrt{a^2+b^2} - a - b > 0$, and hence (13) holds. Now we consider the case where $a > 0$. If $b \leq 0$, we have $(ab)_+ = 0$. If $b > 0$, then we have $\sqrt{a^2+b^2} - a - b < 0$. It follows from (12) that (13) holds. Therefore (13) holds for all $(a,b)^T \neq (0,0)^T \in R^2$. In a similar way, we can show that the inequality

$$\alpha b(ab)_+ \left(\frac{b}{\sqrt{a^2+b^2}} - 1 \right) (\sqrt{a^2+b^2} - a - b) \geq 0 \tag{14}$$

holds for all $(a,b)^T \neq (0,0)^T \in R^2$. Consequently, it follows from (11)–(14) that

$$\frac{\partial \phi_\alpha}{\partial a}(a,b) \frac{\partial \phi_\alpha}{\partial b}(a,b) \geq 0. \tag{15}$$

We consider (v). Suppose that $\phi_\alpha(a,b) = 0$. Then it follows from the definition of ϕ_α that $ab = 0$ and $\sqrt{a^2+b^2} - a - b = 0$. Thus we obtain $\frac{\partial \phi_\alpha}{\partial a}(a,b) = 0$ and $\frac{\partial \phi_\alpha}{\partial b}(a,b) = 0$ from (9) and (10), respectively. Next we show that $\frac{\partial \phi_\alpha}{\partial a}(a,b) = 0$ implies $\phi_\alpha(a,b) = 0$. Suppose that $\frac{\partial \phi_\alpha}{\partial a}(a,b) = 0$. It then follows from (9) that

$$\alpha b(ab)_+ = - \left(\frac{a}{\sqrt{a^2+b^2}} - 1 \right) (\sqrt{a^2+b^2} - a - b). \tag{16}$$

Suppose that both sides of (16) do not vanish. Then we have $b \neq 0$. If $b < 0$, then the left hand side of (16) is not positive and the right hand side of (16) is positive. This is a contradiction. When $b > 0$, $a > 0$ must hold in order to have $(ab)_+ \neq 0$. However, in this case, the left hand side of (16) is positive and the right hand side of (16) is negative. This is a contradiction, too. Therefore both sides of (16) must be equal to zero, and then it can easily be shown that $\phi_\alpha(a,b) = 0$ holds. In a similar way, we can show that $\frac{\partial \phi_\alpha}{\partial b}(a,b) = 0$ implies $\phi_\alpha(a,b) = 0$. The proof is complete. □

As mentioned above, the Fischer-Burmeister function ϕ_{FB} does not have a property like Lemma 3.1 (vi); more precisely, $ab \to \infty$ does not necessarily imply $\phi_{FB}(a,b) \to \infty$. Lemma 3.1 (vi) will turn out to be useful in showing the boundedness of level sets of Ψ_α. From the above lemma, we can easily deduce the following theorem.

Theorem 3.1 *Let Ψ_α be defined by (5). Then, $\Psi_\alpha(x) \geq 0$ for all $x \in R^n$ and $\Psi_\alpha(x) = 0$ if and only if x solves the NCP. Moreover, suppose that the NCP has at least one solution. Then, x is a global minimizer of Ψ_α if and only if x solves the NCP.* □

The above theorem says that the NCP can be reformulated as the unconstrained minimization problem:

$$\min_{x \in R^n} \Psi_\alpha(x). \tag{17}$$

In practice, it is not easy to find a global minimum of Ψ_α. Therefore, it is important to know under what conditions any stationary point of Ψ_α is a global minimum.

Theorem 3.2 *Let F be a P_0-function. Then, $x^* \in R^n$ is a global minimum of the unconstrained optimization problem (17) if and only if x^* is a stationary point of Ψ_α.*

Proof. The theorem can be proved by using property (iv) of Lemma 3.1 and the same proof technique as that of Theorem 3.5 in [10]. □

Now we give some preliminary results that will be useful in obtaining the global and linear convergence results for the algorithm to be presented in the next section. We first show that the functions Ψ_α and Ψ_{NR} are of the same order on any bounded set.

Lemma 3.2 *Let Ψ_α and Ψ_{NR} be defined by (5) and (2), respectively, and $\alpha \geq 0$. Then for any bounded set S, the following inequality holds:*

$$\frac{1}{2}\left(\frac{2}{\sqrt{2}+2}\right)^2 \Psi_{NR}(x) \leq \Psi_\alpha(x) \leq \left\{\frac{\alpha}{2}B^2 + \frac{1}{2}(\sqrt{2}+2)^2\right\} \Psi_{NR}(x) \quad \textit{for all} \quad x \in S, \tag{18}$$

where B is a constant defined by

$$B = \sup_{x \in S}\left\{\max_{1 \leq i \leq n}\{\max\{|x_i|, |F_i(x)|\}\}\right\} < \infty.$$

Proof. By Theorem 2.1 (iii) and the definitions of Ψ_α and $\Psi_{FB} = \Psi_0$, we have

$$\Psi_\alpha(x) \geq \Psi_{FB}(x) \geq \frac{1}{2}\left(\frac{2}{\sqrt{2}+2}\right)^2 \Psi_{NR}(x) \quad \text{for all} \quad x \in R^n.$$

Therefore the left-hand inequality of (18) holds.

Next we consider the right-hand inequality of (18). Let us show that, for each i, the following inequality holds for all $x \in S$:

$$(x_i F_i(x))_+ \leq B \, |\min\{x_i, F_i(x)\}|. \tag{19}$$

We first consider the case where $F_i(x) \geq x_i$. Then $\min\{x_i, F_i(x)\} = x_i$. When $F_i(x) \geq x_i \geq 0$, it follows from the definition of B that

$$(x_i F_i(x))_+ = x_i F_i(x) = F_i(x) \, |\min\{x_i, F_i(x)\}| \leq B \, |\min\{x_i, F_i(x)\}|.$$

In the case where $F_i(x) \geq 0 \geq x_i$, we have $(x_i F_i(x))_+ = 0$ and hence (19) holds obviously. When $0 \geq F_i(x) \geq x_i$, it follows from the definition of B that

$$(x_i F_i(x))_+ = |x_i F_i(x)| \leq |x_i|^2 \leq B \, |\min\{x_i, F_i(x)\}|.$$

In a similar way, we can show that (19) also holds when $x_i \geq F_i(x)$. Therefore (19) holds for all $x \in S$. Thus it follows from Lemma 2.2 (vi) and (19) that for all $i = 1, \ldots, n$ and $x \in S$

$$\begin{aligned} \phi_\alpha(x_i, F_i(x)) &= \frac{\alpha}{2}(x_i F_i(x))_+^2 + \phi_{FB}(x_i, F_i(x)) \\ &\leq \left\{ \frac{\alpha}{2} B^2 + \frac{1}{2}(\sqrt{2}+2)^2 \right\} \min\{x_i, F_i(x)\}^2. \end{aligned} \tag{20}$$

Consequently, from the definitions of Ψ_α and Ψ_{NR}, we have the right-hand inequality of (18). The proof is complete. □

Lemma 3.2 indicates that Ψ_α is in the same order of the natural residual Ψ_{NR} on a bounded set. Thus, from Lemma 2.1, Ψ_α provides a global error bound for the NCP if F is strongly monotone and Lipschitz continuous. Moreover, we can show that the function Ψ_α provides a global error bound without assuming Lipschitz continuity of F when $\alpha > 0$. Note that Ψ_{FB} does not have the same property under the strong monotonicity of F only.

Theorem 3.3 *Let Ψ_α be defined by (5) and $\alpha > 0$. Suppose that F is a uniform P-function with modulus μ. Then there exists a positive constant ρ such that*

$$\|x - x^*\| \leq \rho \Psi_\alpha(x)^{\frac{1}{4}} \quad \textit{for all} \quad x \in R^n, \tag{21}$$

where x^* is the unique solution of the NCP.

Proof. Since F is a uniform P-function with modulus μ, we have

$$\begin{aligned} \mu\|x - x^*\|^2 &\leq \max_{1 \leq i \leq n} (x_i - x_i^*)(F_i(x) - F_i(x^*)) \\ &= \max_{1 \leq i \leq n} \{x_i F_i(x) - x_i^* F_i(x) - F_i(x^*)x_i + x_i^* F_i(x^*)\} \\ &= \max_{1 \leq i \leq n} \{x_i F_i(x) - x_i^* F_i(x) - F_i(x^*)x_i\} \\ &\leq \max_{1 \leq i \leq n} \tau_i \{(x_i F_i(x))_+ + (-F_i(x))_+ + (-x_i)_+\}, \end{aligned} \tag{22}$$

where

$$\tau_i := \max\{1, x_i^*, F_i(x^*)\}.$$

Now we show that there exists a positive constant τ such that

$$\{(ab)_+ + (-b)_+ + (-a)_+\}^2 \leq \tau \phi_\alpha(a, b) \tag{23}$$

for all $(a, b)^T \in R^2$. First we establish the inequality

$$(-a)_+^2 + (-b)_+^2 \leq \left(\sqrt{a^2 + b^2} - a - b\right)^2. \tag{24}$$

Without loss of generality, we assume $a \geq b$. In the case where $a \geq b \geq 0$, (24) holds obviously. In the case where $a \geq 0 \geq b$, we have $\sqrt{a^2 + b^2} - a - b \geq -b \geq 0$, which implies that

$$(-a)_+^2 + (-b)_+^2 = b^2 \leq \left(\sqrt{a^2 + b^2} - a - b\right)^2.$$

In the case where $0 \geq a \geq b$, since $\sqrt{a^2+b^2} - a - b \geq \sqrt{a^2+b^2}$, we obtain

$$(-a)_+^2 + (-b)_+^2 = a^2 + b^2 \leq \left(\sqrt{a^2+b^2} - a - b\right)^2.$$

We thus have shown that (24) holds for all $(a,b)^T \in R^2$. Therefore, we have for all $(a,b)^T \in R^2$

$$\begin{aligned}
\{(ab)_+ + (-b)_+ + (-a)_+\}^2 &= (ab)_+^2 + (-b)_+^2 + (-a)_+^2 \\
&\quad + 2(ab)_+(-b)_+ + 2(ab)_+(-a)_+ + 2(-a)_+(-b)_+ \\
&\leq (ab)_+^2 + (-b)_+^2 + (-a)_+^2 \\
&\quad + ((ab)_+^2 + (-b)_+^2) + ((ab)_+^2 + (-a)_+^2) + ((-a)_+^2 + (-b)_+^2) \\
&\leq 3\left\{(ab)_+^2 + (-b)_+^2 + (-a)_+^2\right\} \\
&\leq 3\left\{(ab)_+^2 + \left(\sqrt{a^2+b^2} - a - b\right)^2\right\} \\
&\leq \tau\left\{\frac{\alpha}{2}(ab)_+^2 + \frac{1}{2}\left(\sqrt{a^2+b^2} - a - b\right)^2\right\} \\
&= \tau\phi_\alpha(a,b),
\end{aligned}$$

where

$$\tau := \max\left\{\frac{6}{\alpha}, 6\right\} > 0,$$

and the third inequality follows from (24).

Consequently, letting

$$\hat{\tau} := \max_{1\leq i\leq n} \tau_i,$$

and combining (22) and (23), we obtain

$$\begin{aligned}
\mu\|x - x^*\|^2 &\leq \max_{1\leq i\leq n} \tau_i \{\tau\phi_\alpha(x_i, F_i(x))\}^{\frac{1}{2}} \\
&\leq \hat{\tau}\tau^{\frac{1}{2}} \max_{1\leq i\leq n} \phi_\alpha(x_i, F_i(x))^{\frac{1}{2}} \\
&\leq \hat{\tau}\tau^{\frac{1}{2}} \left\{\sum_{i=1}^n \phi_\alpha(x_i, F_i(x))\right\}^{\frac{1}{2}} \\
&= \hat{\tau}\tau^{\frac{1}{2}} \Psi_\alpha(x, F(x))^{\frac{1}{2}}.
\end{aligned}$$

The proof is complete. □

Next we consider the boundedness of level sets of Ψ_α. The boundedness of level sets of a merit function is important since it ensures that the sequence generated by a descent method has at least one accumulation point. The following theorem gives conditions under which Ψ_α has bounded level sets.

Theorem 3.4 *Suppose that either of the following conditions holds:*

(a) $\alpha > 0$, *F is monotone and the NCP is strictly feasible.*

(b) $\alpha \geq 0$ *and F is strongly monotone.*

Then the level set $\mathcal{L}(\Psi_\alpha, c)$ is bounded for all $c \in R$.

Proof. Under condition (a), the boundedness of the level set $\mathcal{L}(\Psi_\alpha, c)$ is easily shown in a way similar to Theorem 4.1 in [15] by using Lemma 3.1 (vi). Under condition (b), the result can be established by slightly modifying the proof of Theorem 3.8 in [10]. □

Finally we state a preliminary result which is the key to establishing both global and linear convergence of the algorithm. We abbreviate the vectors $(\frac{\partial \phi_\alpha}{\partial a}(x_1, F_1(x)), \ldots, \frac{\partial \phi_\alpha}{\partial a}(x_n, F_n(x)))^T$ and $(\frac{\partial \phi_\alpha}{\partial b}(x_1, F_1(x)), \ldots, \frac{\partial \phi_\alpha}{\partial b}(x_n, F_n(x)))^T$ as $\frac{\partial \phi_\alpha}{\partial a}(x, F(x))$ and $\frac{\partial \phi_\alpha}{\partial b}(x, F(x))$, respectively.

Lemma 3.3 *Let $\alpha \geq 0$. Then the following inequality holds for all $x \in R^n$:*

$$\| \frac{\partial \phi_\alpha}{\partial a}(x, F(x)) + \frac{\partial \phi_\alpha}{\partial b}(x, F(x)) \|^2 \geq (2 - \sqrt{2})^4 \Psi_{NR}(x).$$

Proof. Suppose $(a,b) \neq (0,0)$. It then follows from (9) and (10) that

$$\begin{aligned} \left(\frac{\partial \phi_\alpha}{\partial a}(a,b) + \frac{\partial \phi_\alpha}{\partial b}(a,b) \right)^2 &= \left\{ \alpha(a+b)(ab)_+ + (\sqrt{a^2+b^2} - a - b) \left(\frac{a+b}{\sqrt{a^2+b^2}} - 2 \right) \right\}^2 \\ &= \alpha^2 (a+b)^2 (ab)_+^2 + (\sqrt{a^2+b^2} - a - b)^2 \left(\frac{a+b}{\sqrt{a^2+b^2}} - 2 \right)^2 \\ &\quad + 2\alpha(a+b)(ab)_+ (\sqrt{a^2+b^2} - a - b) \left(\frac{a+b}{\sqrt{a^2+b^2}} - 2 \right). \end{aligned} \tag{25}$$

Now we show that the inequality

$$2\alpha(a+b)(ab)_+ (\sqrt{a^2+b^2} - a - b) \left(\frac{a+b}{\sqrt{a^2+b^2}} - 2 \right) \geq 0 \tag{26}$$

holds for all $(a,b)^T \neq (0,0)^T \in R^2$. Note that $\frac{a+b}{\sqrt{a^2+b^2}} - 2 \leq 0$ holds for all $(a,b)^T \neq (0,0)^T \in R^2$. When $ab \leq 0$, we have $(ab)_+ = 0$. When $a > 0$ and $b > 0$, we have $a+b > 0$ and $\sqrt{a^2+b^2} - a - b < 0$. When $a < 0$ and $b < 0$, we have $a + b < 0$ and $\sqrt{a^2+b^2} - a - b > 0$. Therefore (26) holds for all $(a,b)^T \neq (0,0)^T \in R^2$. Combining (25) and (26), we obtain for all $(a,b)^T \neq (0,0)^T \in R^2$

$$\begin{aligned} \left(\frac{\partial \phi_\alpha}{\partial a}(a,b) + \frac{\partial \phi_\alpha}{\partial b}(a,b) \right)^2 &\geq (\sqrt{a^2+b^2} - a - b)^2 \left(\frac{a+b}{\sqrt{a^2+b^2}} - 2 \right)^2 \\ &\geq \left(\frac{2}{\sqrt{2}+2} \right)^2 \min\{a,b\}^2 \left(\frac{a+b}{\sqrt{a^2+b^2}} - 2 \right)^2 \\ &\geq \left(\frac{2}{\sqrt{2}+2} \right)^2 (2-\sqrt{2})^2 \min\{a,b\}^2 \\ &= (2-\sqrt{2})^4 \min\{a,b\}^2, \end{aligned} \tag{27}$$

where the second inequality follows from Lemma 2.2 (vi) and the last inequality follows from the fact that

$$2 - \sqrt{2} \leq 2 - \frac{a+b}{\sqrt{a^2+b^2}} \quad \text{for all} \quad (a,b)^T \neq (0,0)^T \in R^2.$$

Note that (27) also holds trivially for $(a,b)^T = (0,0)^T$, and hence it holds for all $(a,b)^T \in R^2$. Consequently, the lemma follows from the definition of Ψ_{NR}. □

4. A Descent Method and Global Convergence

In this section, we propose a derivative-free descent method based on the function ϕ_α. Furthermore, we prove global convergence of this method.

We consider the following search direction:

$$d^k(\beta) := -\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) - \beta \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)), \tag{28}$$

where β is a parameter such that $\beta \in (0,1)$. Although this search direction is not necessarily a descent direction of Ψ_α at x^k for any $\beta \in (0,1)$, we can choose $\beta > 0$ sufficiently small so that $d^k(\beta)$ is a descent direction of Ψ_α at x^k, provided that the monotonicity assumption on F is fulfilled. We show this result in the following lemma.

Lemma 4.1 *Suppose that F is monotone and $\alpha \geq 0$. If x^k is not a solution of the NCP, then there exists $\bar{\beta}(x^k) \in (0,1)$ such that, for all $\beta \in [0, \bar{\beta}(x^k))$ the search direction $d^k(\beta)$ defined by (28) satisfies the descent condition*

$$\nabla \Psi_\alpha(x^k)^T d^k(\beta) < 0. \tag{29}$$

Proof. Since F is continuously differentiable, the function Ψ_α is also continuously differentiable by Lemma 3.1 (iii) and the gradient of Ψ_α at x^k is given by

$$\nabla \Psi_\alpha(x^k) = \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)). \tag{30}$$

From (30) and the definition of $d^k(\beta)$, we have

$$\begin{aligned} &\nabla \Psi_\alpha(x^k)^T d^k(\beta) \\ &= -\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) - \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \\ &\quad + \beta \left(-\| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \|^2 - \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \right). \end{aligned} \tag{31}$$

Now let $p(x^k)$ and $q(x^k)$ be defined by

$$p(x^k) := -\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) - \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))$$

and

$$q(x^k) := -\| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \|^2 - \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)),$$

respectively. Then (31) is rewritten as

$$\nabla \Psi_\alpha(x^k)^T d^k(\beta) = p(x^k) + \beta q(x^k).$$

By the monotonicity of F, $\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \geq 0$. If x^k is not a solution of the NCP, we have $\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) > 0$ from Lemma 3.1 (iv) and (v). Therefore $p(x^k) < 0$. Let $\bar{\beta}(x^k)$ be defined by

$$\bar{\beta}(x^k) := \begin{cases} -p(x^k)/q(x^k) & \text{if } \ q(x^k) \neq 0 \ \text{ and } \ 0 \leq -p(x^k)/q(x^k) \leq 1, \\ 1 & \text{otherwise.} \end{cases}$$

For such $\bar{\beta}(x^k)$, it is easily seen that, for all $\beta \in [0, \bar{\beta}(x^k))$, the search direction $d^k(\beta)$ satisfies the descent condition $\nabla \Psi_\alpha(x^k)^T d^k(\beta) < 0$. □

The above lemma motivates us to propose the following descent algorithm.

Algorithm 4.1

Step1. *Choose* $x^0 \in R^n, \epsilon \geq 0, \sigma \in (0,1), \beta \in (0,1)$ *and* $\gamma \in (0,1)$. *Set* $k := 0$.

Step2. *If* $\Psi_\alpha(x^k) \leq \epsilon$, *stop:* x^k *is an approximate solution of the NCP.*

Step3. *Compute* $x^{k+1} := x^k + \gamma^{l_k} d^k(\beta^{l_k})$, *where*

$$d^k(\beta^{l_k}) := -\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) - \beta^{l_k} \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))$$

with l_k *being the smallest nonnegative integer* l *satisfying*

$$\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k) \leq -\sigma(\gamma^l)^2 \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|^2. \tag{32}$$

Step4. *Set* $k := k + 1$ *and go to Step 2.*

Note that the above algorithm has no need to compute the gradients of Ψ_α, and therefore there is also no need to compute the Jacobian of F. The most remarkable feature of Algorithm 4.1 is that not only the stepsize but also the search direction itself is adjusted during the backtracking search of Armijo-type, which may be regarded as a sort of curvilinear search.

We are ready to state the global convergence result for Algorithm 4.1 under the monotonicity assumption on F. In the remainder of the paper, we assume that the parameter ϵ used in Algorithm 4.1 is set to be zero and Algorithm 4.1 generates an infinite sequence $\{x^k\}$.

Theorem 4.1 *Suppose that* F *is monotone and* $\alpha \geq 0$. *Then Algorithm 4.1 is well-defined for any initial point* x^0. *Furthermore, if* x^* *is an accumulation point of the sequence* $\{x^k\}$ *generated by Algorithm 4.1, then* x^* *is a solution of the NCP.*

Proof. We first prove that Algorithm 4.1 is well-defined. It suffices to show that Step 3 is well-defined. Assume to the contrary that there is no nonnegative integer l satisfying (32). It then follows that for any integer $l \geq 0$

$$\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k) > -\sigma(\gamma^l)^2 \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|^2.$$

Dividing the above inequality by γ^l and passing to the limit $l \to \infty$, we have

$$\lim_{l \to \infty} \frac{\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k)}{\gamma^l} \geq 0. \tag{33}$$

Since Ψ_α is continuously differentiable, Ψ_α is locally Lipschitz continuous at x^k, i.e., there exist $L_1 > 0$ and $\delta > 0$ such that

$$\|\Psi_\alpha(x^k) - \Psi_\alpha(x')\| \leq L_1 \|x^k - x'\|, \quad \text{for all } x' \text{ such that } \|x^k - x'\| < \delta.$$

Therefore, we have

$$\begin{aligned}
&\lim_{l\to\infty} \frac{\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k)}{\gamma^l} \\
&= \lim_{l\to\infty} \left\{ \frac{\Psi_\alpha(x^k + \gamma^l d^k(0)) - \Psi_\alpha(x^k)}{\gamma^l} + \frac{\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k + \gamma^l d^k(0))}{\gamma^l} \right\} \\
&= \nabla\Psi_\alpha(x^k)^T d^k(0) + \lim_{l\to\infty} \frac{\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k + \gamma^l d^k(0))}{\gamma^l} \\
&\le \nabla\Psi_\alpha(x^k)^T d^k(0) + \lim_{l\to\infty} L_1 \|d^k(\beta^l) - d^k(0)\| \\
&= \nabla\Psi_\alpha(x^k)^T d^k(0). \qquad (34)
\end{aligned}$$

It then follows from (33) and (34) that

$$\nabla\Psi_\alpha(x^k)^T d^k(0) \ge 0.$$

This is a contradiction, because from Lemma 4.1, $d^k(0)$ must be a descent direction of Ψ_α at x^k if x^k is not a solution of the NCP. Thus Algorithm 4.1 is well-defined.

Next we prove that any accumulation point x^* of $\{x^k\}$ is a solution of the NCP. Let $\{x^k\}_{k\in K}$ be a subsequence converging to x^*. Then $\{x^k\}_{k\in K}$ is bounded and thus $\{d^k(\beta^{l_k})\}_{k\in K}$ is bounded since ϕ_α is continuously differentiable. Without loss of generality, we assume $d^k(\beta^{l_k}) \to d^*$ as $k(\in K) \to \infty$. Since $\Psi_\alpha(x^k)$ decreases at each iteration, the right-hand side of (32) tends to 0. First we consider the case where $\{l_k\}_{k\in K}$ is bounded. Then $\{\gamma^{l_k}\}_{k\in K}$ does not approach 0, so we have

$$\|\frac{\partial\phi_\alpha}{\partial a}(x^*, F(x^*)) + \frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*))\|^2 = 0.$$

Therefore, it follows from Lemma 3.3 that $\Psi_{NR}(x^*) = 0$, i.e., x^* is a solution of the NCP. Next suppose that $\{l_k\}_{k\in K}$ is unbounded, which implies $\{\gamma^{l_k}\}_{k\in K} \to 0$. When $\{l_k\}_{k\in K} \to \infty$, it follows from Step 3 of Algorithm 4.1 that for all $k \in K$

$$\Psi_\alpha(x^k + \gamma^{l_k-1} d^k(\beta^{l_k-1})) - \Psi_\alpha(x^k) > -\sigma(\gamma^{l_k-1})^2 \|\frac{\partial\phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial\phi_\alpha}{\partial b}(x^k, F(x^k))\|^2.$$

Dividing the above inequality by γ^{l_k-1} and passing to the limit $k \to \infty$, we have

$$\lim_{k\to\infty} \frac{\Psi_\alpha(x^k + \gamma^{l_k-1} d^k(\beta^{l_k-1})) - \Psi_\alpha(x^k)}{\gamma^{l_k-1}} \ge 0. \qquad (35)$$

Since Ψ_α is continuously differentiable, by the Mean-Value theorem there exists some $\tilde{\gamma}^k \in [0, \gamma^{l_k-1}]$ such that

$$\Psi_\alpha(x^k + \gamma^{l_k-1} d^k(\beta^{l_k-1})) - \Psi_\alpha(x^k) = \gamma^{l_k-1} \nabla\Psi_\alpha(x^k + \tilde{\gamma}^k d^k(\beta^{l_k-1}))^T d^k(\beta^{l_k-1}). \qquad (36)$$

As $k(\in K) \to \infty$, we have $\{\gamma^{l_k-1}\} \to 0$, which in turn implies $\{\tilde{\gamma}^k\}_{k\in K} \to 0$ and hence

$$\lim_{k\to\infty} \tilde{\gamma}^k d^k(\beta^{l_k-1}) = 0. \qquad (37)$$

It then follows from (36), (37) and the given assumptions that

$$\begin{aligned}\lim_{k\to\infty}\frac{\Psi_\alpha(x^k+\gamma^{l_k-1}d^k(\beta^{l_k-1}))-\Psi_\alpha(x^k)}{\gamma^{l_k-1}} &= \lim_{k\to\infty}\nabla\Psi_\alpha(x^k+\tilde{\gamma}^k d^k(\beta^{l_k-1}))^T d^k(\beta^{l_k-1}) \\ &= \nabla\Psi_\alpha(x^*)^T d^*, \end{aligned} \tag{38}$$

where $\tilde{\gamma}^k \in [0, \gamma^{l_k-1}]$. Combining (35) and (38), we obtain

$$\nabla\Psi_\alpha(x^*)^T d^* \geq 0.$$

Since $\nabla\Psi_\alpha(x^*)^T d^* \leq 0$ holds obviously, we have $\nabla\Psi_\alpha(x^*)^T d^* = 0$. Moreover, since passing to the limit in (28) and (30) on the subsequence yields

$$d^* = -\frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*))$$

and

$$\nabla\Psi_\alpha(x^*) = \frac{\partial\phi_\alpha}{\partial a}(x^*, F(x^*)) + \nabla F(x^*)\frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*)),$$

we obtain

$$\begin{aligned}&\nabla\Psi_\alpha(x^*)^T d^* \\ &= -\frac{\partial\phi_\alpha}{\partial a}(x^*, F(x^*))^T\frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*)) - \frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*))^T\nabla F(x^*)\frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*)) \\ &= 0.\end{aligned}$$

Consequently, by the monotonicity of F and Lemma 3.1 (iv), we have

$$\frac{\partial\phi_\alpha}{\partial a}(x^*, F(x^*))^T\frac{\partial\phi_\alpha}{\partial b}(x^*, F(x^*)) = 0.$$

It then follows from Lemma 3.1 (i) and (v) that x^* is a solution of the NCP. □

Combining the above theorem and Lemma 3.4, we obtain the following global convergence result.

Theorem 4.2 *Suppose that either of the following conditions (a) and (b) holds:*

(a) $\alpha > 0$, *F is monotone and the NCP is strictly feasible.*

(b) $\alpha \geq 0$ *and F is strongly monotone.*

Then Algorithm 4.1 is well-defined for any initial point x^0. Furthermore, the sequence $\{x^k\}$ generated by Algorithm 4.1 has at least one accumulation point and any accumulation point of $\{x^k\}$ is a solution of the NCP. □

5. Linear Convergence

In this section, we show that any sequence generated by Algorithm 4.1 converges linearly to the solution if F is strongly monotone.

Before proving the linear convergence result, we give the following lemma.

Lemma 5.1 *Suppose that F is strongly monotone with modulus $\mu > 0$ and that $\alpha \geq 0$. Let $x^0 \in R^n$ be any given vector. Then there exists an integer $\hat{l} \geq 0$ such that, for each x in the level set $\mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0))$ and for all $l \geq \hat{l}$, the search direction $d^k(\beta^l)$ satisfies*

$$\nabla \Psi_\alpha(x)^T d^k(\beta^l) \leq -\frac{\beta^l}{2} \left(\| \frac{\partial \phi_\alpha}{\partial a}(x, F(x)) \| + \| \frac{\partial \phi_\alpha}{\partial b}(x, F(x)) \| \right)^2 .$$

Proof. It follows from Theorem 3.4 that the level set $\mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0))$ is bounded. Since ∇F is continuous, there exists a constant $\nu > 0$ such that

$$\|\nabla F(x)\| \leq \nu \quad \text{for all} \quad x \in \mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0)). \tag{39}$$

From (28) and (30), we have

$$\begin{aligned}
&\nabla \Psi_\alpha(x^k)^T d^k(\beta^l) \\
&= -\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) - \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \\
&\quad - \beta^l \left(\| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \|^2 + \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \right) \\
&\leq -\mu \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|^2 - \beta^l \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \|^2 \\
&\quad - \beta^l \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \nabla F(x^k) \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \\
&\leq -\mu \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|^2 - \beta^l \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \|^2 \\
&\quad + \beta^l \nu \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \| \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|,
\end{aligned} \tag{40}$$

where the first inequality follows from Lemma 3.1 (iv) and the strong monotonicity of F and, the second inequality follows from the Cauchy-Schwartz inequality and (39).

Now let $\hat{l}$ be defined by

$$\hat{l} := \left\lceil \log_\beta \frac{2\mu}{\nu^2 + 2\nu + 2} \right\rceil .$$

Then β^l satisfies

$$\beta^l \leq \frac{2\mu}{\nu^2 + 2\nu + 2} \quad \text{for all} \quad l \geq \hat{l}. \tag{41}$$

Since (41) implies $\mu - \frac{\beta^l}{2} \geq 0$, the right-hand side of (40) can be rewritten as

$$\begin{aligned}
&-\mu \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|^2 - \beta^l \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \|^2 + \beta^l \nu \| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \| \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \| \\
&= -\frac{\beta^l}{2} \left(\| \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) \| + \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \| \right)^2 - \left(\mu - \frac{\beta^l}{2} \right) \| \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \|^2
\end{aligned}$$

$$
\begin{aligned}
& -\frac{\beta^l}{2}\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\|^2 + \beta^l(\nu+1)\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\|\|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\| \\
= \quad & -\frac{\beta^l}{2}\left(\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\| + \|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|\right)^2 \\
& -\left(\sqrt{\frac{\beta^l}{2}}\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\| - \sqrt{\mu - \frac{\beta^l}{2}}\|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|\right)^2 \\
& +\left\{\beta^l(\nu+1) - \sqrt{\beta^l(2\mu-\beta^l)}\right\}\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\|\|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\| \qquad (42)
\end{aligned}
$$

for all $l \geq \hat{l}$. Since (41) implies

$$
\begin{aligned}
\sqrt{\beta^l(2\mu-\beta^l)} &\geq \sqrt{\beta^{2l}(\nu^2+2\nu+2) - \beta^{2l}} \\
&= \sqrt{\beta^{2l}(\nu+1)^2} \\
&= \beta^l(\nu+1),
\end{aligned}
$$

the last term of (42) is nonpositive for all $l \geq \hat{l}$. Consequently, it follows from (40) and (42) that for all $l \geq \hat{l}$

$$
\nabla \Psi_\alpha(x^k)^T d^k(\beta^l) \leq -\frac{\beta^l}{2}\left(\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\| + \|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|\right)^2.
$$

□

Now the linear convergence result for Algorithm 4.1 is presented.

Theorem 5.1 *Suppose that F is strongly monotone and $\alpha \geq 0$. Suppose also that F is Lipschitz continuous on any bounded set. Let the parameters γ and β be chosen to satisfy $\gamma < \beta$. Then the sequence $\{x^k\}$ generated by Algorithm 4.1 converges R-linearly to the solution of the NCP and the sequence $\{\Psi_\alpha(x^k)\}$ converges Q-linearly to zero.*

Proof. We first show that the sequence $\{\Psi_\alpha(x^k)\}$ converges Q-linearly to zero. By Theorem 3.4, the level set $\mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0))$ is bounded. Then since F, ∇F and $\nabla \Psi_\alpha$ are all Lipschitz continuous on $\mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0))$, there exists a constant $L_2 > 0$ such that

$$
\|\nabla \Psi_\alpha(x) - \nabla \Psi_\alpha(x')\| \leq L_2\|x - x'\| \quad \text{for all} \quad x, x' \in \mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0)). \qquad (43)
$$

Since the sequence $\{\Psi_\alpha(x^k)\}$ is nonincreasing, $\{x^k\} \subseteq \mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0))$. It then follows from (43) and the Cauchy-Schwartz inequality that for each $t \in [0,1]$

$$
\begin{aligned}
\Psi_\alpha(x^k + td^k(\beta^l)) - \Psi_\alpha(x^k) &= \int_0^t \nabla \Psi_\alpha(x^k + sd^k(\beta^l))^T d^k(\beta^l) ds \\
&= t\nabla \Psi_\alpha(x^k)^T d^k(\beta^l) \\
&\quad + \int_0^t \left(\nabla \Psi_\alpha(x^k + sd^k(\beta^l)) - \nabla \Psi_\alpha(x^k)\right)^T d^k(\beta^l) ds \\
&\leq t\nabla \Psi_\alpha(x^k)^T d^k(\beta^l) + L_2 \int_0^t s\|d^k(\beta^l)\|^2 ds \\
&= t\left(\nabla \Psi_\alpha(x^k)^T d^k(\beta^l) + \frac{L_2 t}{2}\|d^k(\beta^l)\|^2\right). \qquad (44)
\end{aligned}
$$

By Lemma 3.1 (iv) and $\beta^l \in (0,1)$, we have

$$\begin{aligned}
&\|d^k(\beta^l)\|^2 \\
&\quad = \| - \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) - \beta^l \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\|^2 \\
&\quad = \|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2 + \beta^{2l}\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\|^2 + 2\beta^l \frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \\
&\quad \le \|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2 + \|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\|^2 + 2\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))^T \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k)) \\
&\quad = \|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2. \qquad (45)
\end{aligned}$$

Let $\hat{l}$ be as specified in the proof of Lemma 5.1. Then, from (44), (45) and Lemma 5.1, we have for all $l \ge \hat{l}$

$$\begin{aligned}
&\Psi_\alpha(x^k + \gamma^l d^k(\beta^l)) - \Psi_\alpha(x^k) \\
&\quad \le \gamma^l \left(-\frac{\beta^l}{2} \left(\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k))\| + \|\frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\| \right)^2 + \frac{L_2 \gamma^l}{2}\|d^k(\beta^l)\|^2 \right) \\
&\quad \le \gamma^l \left(-\frac{\beta^l}{2}\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2 + \frac{L_2 \gamma^l}{2}\|d^k(\beta^l)\|^2 \right) \\
&\quad \le -\frac{\gamma^l}{2}(\beta^l - L_2\gamma^l)\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2, \qquad (46)
\end{aligned}$$

where the second inequality follows from the Cauchy-Schwartz inequality and the last inequality follows from (45). By (46), condition (32) is satisfied whenever

$$-\frac{\gamma^l}{2}(\beta^l - L_2\gamma^l) \le -\sigma(\gamma^l)^2.$$

Since this inequality can be rewritten as

$$\left(\frac{\gamma}{\beta}\right)^l \le \frac{1}{2\sigma + L_2}, \qquad (47)$$

condition (32) is satisfied for all $l \ge \bar{l}$, where $\bar{l}$ is defined by

$$\bar{l} := \max\left\{ \hat{l}, \left\lceil \log_{\frac{\beta}{\gamma}}(2\sigma + L_2) \right\rceil \right\}. \qquad (48)$$

Notice that $\bar{l}$ does not depend on k. Thus, we have $l_k \le \bar{l}$ for all k since l_k is the smallest nonnegative integer l satisfying (32). It then follows from (32) and Lemma 3.3 that

$$\begin{aligned}
\Psi_\alpha(x^{k+1}) - \Psi_\alpha(x^k) &\le -\sigma(\gamma^{l_k})^2\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2 \\
&\le -\sigma(\gamma^{\bar{l}})^2\|\frac{\partial \phi_\alpha}{\partial a}(x^k, F(x^k)) + \frac{\partial \phi_\alpha}{\partial b}(x^k, F(x^k))\|^2 \\
&\le -(2-\sqrt{2})^4 \sigma \gamma^{2\bar{l}} \Psi_{NR}(x^k). \qquad (49)
\end{aligned}$$

Since the level set $\mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0))$ is bounded, it follows from Lemma 3.2 that

$$\Psi_\alpha(x^k) \le \left\{ \frac{\alpha}{2}B^2 + \frac{1}{2}(\sqrt{2}+2)^2 \right\} \Psi_{NR}(x^k) \quad \text{for all} \quad x^k \in \mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0)), \qquad (50)$$

where B is given by

$$B = \sup \left\{ \max_{1 \le i \le n} \left\{ \max\{|x_i^k|, |F_i(x^k)|\} \right\} \;\middle|\; x \in \mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0)) \right\}.$$

Therefore from (49) and (50) we have

$$\Psi_\alpha(x^{k+1}) - \Psi_\alpha(x^k) \le -\frac{2(2-\sqrt{2})^4}{\alpha B^2 + (\sqrt{2}+2)^2} \sigma \gamma^{2\bar{l}} \Psi_\alpha(x^k). \tag{51}$$

Consequently, we obtain

$$\left(1 - \frac{2(2-\sqrt{2})^4 \sigma \gamma^{2\bar{l}}}{\alpha B^2 + (\sqrt{2}+2)^2} \right) \Psi_\alpha(x^k) \ge \Psi_\alpha(x^{k+1}) \ge 0.$$

On the other hand, since $(\sqrt{2}+2)^2 > 2(2-\sqrt{2})^4, \sigma\gamma^{2\bar{l}} \in (0,1)$ and $\alpha B^2 \ge 0$, we have

$$0 < 1 - \frac{2(2-\sqrt{2})^4 \sigma \gamma^{2\bar{l}}}{\alpha B^2 + (\sqrt{2}+2)^2} < 1.$$

Therefore the sequence $\{\Psi_\alpha(x^k)\}$ converges Q-linearly to zero.

Next we prove that the sequence $\{x^k\}$ converges R-linearly to the unique solution x^* of the NCP. From Lemma 2.1 and Lemma 3.2, there exists a positive constant κ such that

$$\|x - x^*\| \le \kappa \Psi_\alpha(x)^{\frac{1}{2}} \quad \text{for all } x \in \mathcal{L}(\Psi_\alpha, \Psi_\alpha(x^0)).$$

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

The conditions in the above theorem are weaker than those in [18]. For the linear convergence, we only need the condition $\gamma < \beta$, which does not involve unknown parameters. Also, note that some parameters used in Algorithm 4.1 do not rely on usually unknown constants such as the modulus of strong monotonicity of F.

6. Concluding Remarks

In this paper, we have proposed a merit function for the NCP and shown its properties. We have also proposed a new derivative-free descent algorithm based on this function for solving the NCP and proven global and linear convergence theorems under appropriate assumptions.

The most interesting part of Algorithm 4.1 is Step 3, where the search direction and the stepsize are adjusted simultaneously, whereas search directions in the algorithms proposed in [10, 12, 18, 24] are fixed during a linesearch process. This particular feature enables us to establish the linear convergence of the algorithm under more practical assumptions compared with the algorithm of [18].

Derivative-free descent algorithms have also been proposed for solving the VIP [8, 27]. However, the convergence rate results for those algorithms have been scarce. Also, those algorithms generally use parameters that have to be specified by using unknown constants. It is an interesting future topic to extend the results of this paper to the more general class of VIPs.

References

[1] B. Chen, X. Chen and C. Kanzow, A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results, Working paper, School of Mathematics, University of New South Wales, September 1997.

[2] R.W. Cottle, J.-S. Pang and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, New York, 1992.

[3] S. Dafermos, An iterative scheme for variational inequalities, *Mathematical Programming* 26 (1983), pp. 40-47.

[4] F. Facchinei and C. Kanzow, On unconstrained and constrained stationary points of the implicit Lagrangian, *Journal of Optimization Theory and Applications* 92 (1997), pp. 99-115.

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

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

[7] A. Fischer, An NCP-Function and its use for the solution of complementarity problems, *Recent Advances in Nonsmooth Optimization*, D.Z. Du, L. Qi, and R.S. Womersley (eds.), World Scientific Publishers, Singapore, 1995, pp. 88-105.

[8] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Mathematical Programming* 53 (1992), pp. 99-110.

[9] M. Fukushima, Merit functions for variational inequality and complementarity problems, *Nonlinear Optimization and Applications*, G. Di Pillo and F. Giannessi (eds.), Plenum Press, New York, 1996, pp. 155-170.

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

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

[12] H. Jiang, Unconstrained minimization approaches to nonlinear complementarity problems, *Journal of Global Optimization* 9 (1996), pp. 169-181.

[13] C. Kanzow, An unconstrained optimization technique for large-scale linearly constrained convex minimization problems, *Computing* 53 (1994), pp. 101-117.

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

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

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

[17] O.L. Mangasarian and M.V. Solodov, Nonlinear complementarity as unconstrained and constrained minimization, *Mathematical Programming* 62 (1993), pp. 277-297.

[18] O.L. Mangasarian and M.V. Solodov, A linearly convergent descent method for strongly monotone complementarity problems, *Computational Optimization and Applications*, to appear.

[19] J.-S. Pang, A posteriori error bounds for the linearly-constrained variational inequality problem, *Mathematics of Operations Research* 12 (1987), pp. 474-484.

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

[21] J.-S. Pang, Newton's method for B-differentiable equations, *Mathematics of Operations Research* 15 (1990), pp. 311-341.

[22] J.-S. Pang and D. Chan, Iterative methods for variational and complementarity problems, *Mathematics Programming* 24 (1982), pp. 284-313.

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

[24] N. Yamashita and M. Fukushima, On stationary points of the implicit Lagrangian for nonlinear complementarity problems, *Journal of Optimization Theory and Applications* 84 (1995), pp. 653-663.

[25] N. Yamashita and M. Fukushima, Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems, *Mathematical Programming* 76 (1997), pp. 469-491.

[26] N. Yamashita and M. Fukushima, A new merit function and a descent method for semidefinite complementarity problems, *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi (eds.), Kluwer Academic Publishers, Boston, Massachusetts, 1998, pp. 401-416.

[27] N. Yamashita, K. Taji and M. Fukushima, Unconstrained optimization reformulations of variational inequality problems, *Journal of Optimization Theory and Applications* 92 (1997), pp. 439-456.