A Semismooth Active-set Algorithm for Degenerate Nonlinear Complementarity Problems

Haodong Yu
School of Mathematics and Information, Shanghai Lixin University of Commerce, Shanghai, China
Email: nianchuixiao@msn.com

***Abstract*— We propose a semismooth active-set Newton algorithm for solving the nonlinear complementarity problems with degenerate solutions. This method introduces the active-set technique to identify the degenerate set. At each iteration, the search direction is obtained by two reduced linear systems. Instead of employing gradient steps as adjustments to guarantee the sufficient reduction of the merit function, the algorithm employs a Newton-type direction, which is more efficient than gradient direction, in the adjustment step. This method has globally convergence. When near the solution, the degenerate set will be identified correctly, and only one reduced linear system is solved at each iteration. Under some mild assumptions, locally superlinear convergence is obtained as well. Numerical experiments on MATLAB shows the efficiency of the method.**

***Index Terms*— NCP; semismooth method; degenerate solution; active-set strategy; superlinear convergence; MATLAB**

I. Introduction

Consider the following nonlinear complementarity problem(denoted by NCP(F)):

$$F(x) \geq 0 \quad x \geq 0 \quad and \quad x^T F(x) = 0, \qquad (1)$$

where $x \in \mathbb{R}^n$, and $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function. The nonlinear complementarity problem is an important problem and has attracted strong interests due to its important applications in economics, engineering, and mechanics, etc (see [1] for a review).

By using the following F-B (Fischer-Burmeister) function $\phi : \mathbb{R}^2 \to \mathbb{R}$:

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

the complementarity problem (1) can be reformulated to the next equivalent nonlinear equations

$$\Phi(x) = \begin{pmatrix} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{pmatrix} = 0 \qquad (2)$$

Clearly, $x^* \in \mathbb{R}^n$ solves the NCP(F) if and only if it is a solution of the equation (2). The nature merit function for (2) is $\Psi(x) := \frac{1}{2}\|\Phi(x)\|^2$.

This work is supported by Foundation for University youth teachers by Shanghai Municipality(No. shlx014).

For any solution x^* of NCP(F), denote $\mathcal{I} := \{1, \cdots, n\}$ be its index set. Then $\mathcal{I}$ can be divided into three parts:

$$\begin{aligned} \alpha = \alpha(x^*) := &\quad \{i \in \mathcal{I} \mid x_i^* > 0\}, \\ \beta = \beta(x^*) := &\quad \{i \in \mathcal{I} \mid x_i^* = 0 = F_i(x^*)\}, \\ \chi = \chi(x^*) := &\quad \{i \in \mathcal{I} \mid F_i(x^*) > 0\}. \end{aligned}$$

Definition 1: Suppose x^* is a solution of (1), then x^* is called a degenerate solution if $\beta(x^*)$ is not empty; if $\beta(x^*) = \emptyset$, then x^* is called a non-degenerate solution.

The F-B function is continuously differentiable for any pair $(a, b) \neq (0, 0)$, and has only semismooth at the zero point. For this reason, the system (2) is also semismooth at its degenerate solution and can only be solved by generalized Jacobian technique or smoothing method. From both theoretical and practical point of view, the identification of the degenerate set β of the solution is very important. If the degenerate indices can be identified before exactly knowing x^*, then we only have to solve a reduced form of the equation (2) in a small neighbor of x^*. The original degenerate NCP(F) will also be transformed to a non-degenerate problem. (See [2], [3] for a reference.)

Active-set method is an important technique in nonlinear optimization. In [4], Facchinei, Fischer and Kanzow presented a new method for the accurate identification of the active constraints for nonlinear programs with inequality constraints. The main idea of their technique is to define an identification function $\rho(x, \lambda)$ which has "slower" convergence rate to the KKT set than the pair (x, λ) does. In [5], Kanzow and Qi proposed an active-set QP-free Newton method for variational inequality problems (denoted VIP). They used the technique presented in [4] to identify the active constraints of the solution of VIP and incorporate this technique into a global algorithm.

The method in [5] employs gradient steps as safeguards in those cases when the Newton-type step does not sufficiently reduce the merit function. However, as is well known, the gradient step may have low efficiency of iteration in many cases, especially when it meet with so-called "zigzag phenomenon". Moreover, the gradient step doesn't have local superlinearly convergence. Hence, the numerical efficiency of the gradient step may be slow, which will destroy the whole convergence rate of the algorithm.

In this paper, we propose a semismooth active-set

doi:10.4304/jsw.8.7.1568-1576

method for degenerate nonlinear complementarity problems. We apply the active-set technology to identify the degenerate indices of the solution. Instead of employing gradient steps as adjustments to guarantee the sufficient reduction of the merit function, the algorithm employs a Newton-type direction, which is more efficient than gradient direction, in the adjustment step. We will show in Section 3 that the Newtonian adjustment step is a kind of asymptotically best adjustment.

Our method has global convergence and each accumulation point is at least a stationary point of (2). Under some regularity assumptions, the estimate set is eventually equal to the degenerate set of the solution, and the local superlinear convergence of the algorithm is obtained as well.

The organization of this paper is as follows. In Section 2, some definitions and preliminary results are given. An active-set semismooth algorithm is presented in Section 3. In Section 4, we prove the proposed algorithm is well defined and establish its global convergence. The local convergence and the analysis of the identification technique is established in Section 5. Finally, some numerical results on MATLAB are listed to show the efficiency of our algorithm.

II. Preliminary

For any $x \in \mathbb{R}^n$ and the index set $\mathcal{J} \subseteq \{1, \cdots n\}$, we denote $x_{\mathcal{J}}$ as the vector with components $x_i, i \in \mathcal{J}$. Let mapping G: $\mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitzian. Assume $x \in \mathbb{R}^n$ be any differentiable point of G, denote $G'(x)$ be the Jacobian of G at x, and $\nabla G(x)$ be its transposed Jacobian.

The function $\Phi(x)$ defined by (2) is then locally Lipschitzian. The generalized Jacobian $\partial\Phi(x)$ of Φ at x (in the Clarke sense) also exists and has the following property:

$$\partial\Phi(x) \subseteq D_a(x) + D_b(x)F'(x) \tag{3}$$

where $D_a(x) = diag(a_1(x), \cdots, a_n(x))$, $D_b(x) = diag(b_1(x), \cdots, b_n(x))$ with their elements be

$$a_i(x) = \frac{x_i}{\sqrt{x_i^2 + F_i^2(x)}} - 1, \quad b_i(x) = \frac{F_i(x)}{\sqrt{x_i^2 + F_i^2(x)}} - 1$$

when $(x_i, F_i(x)) \neq (0, 0)$ and

$$a_i(x) = \xi_i - 1, b_i(x) = \eta_i - 1, \quad (\xi_i, \eta_i) \in \mathbb{R}^2, \|(\xi_i, \eta_i)\| \leq 1$$

when $(x_i, F_i(x)) = (0, 0)$. The next property shows the semismoothness of Φ.

Proposition 2: ([6]) Assume that $\{x^k\} \subseteq \mathbb{R}^n$ is a convergent sequence with a limit point $x^* \in \mathbb{R}^n$. Then the following statements hold.
(i) The function Φ is semismooth, which implies that for any $V_k \in \partial\Phi(x^k)$,

$$\|\Phi(x^k) - \Phi(x^*) - V_k(x^k - x^*)\| = o(\|x^k - x^*\|).$$

(ii) If F' is Lipschitz continuous, then the function Φ is strongly semismooth, which implies that for any $V_k \in \partial\Phi(x^k)$,

$$\|\Phi(x^k) - \Phi(x^*) - V_k(x^k - x^*)\| = O(\|x^k - x^*\|^2).$$

Proposition 3: ([5]) The merit function Ψ is continuously differentiable with $\nabla\Psi(x) = H^T\Phi(x)$ for an arbitrary element $H \in \partial\Phi(x)$.

Definition 4: A matrix M is said to be a P matrix if for all $x \in \mathbb{R}^n, x \neq 0$, there exists a component $x_k \neq 0$ such that

$$x_k(Mx)_k > 0. \tag{4}$$

Now we introduce Robinson's [7] strong regularity condition of the solution of NCP(F) and its error bound property, see [7] and [8] for details.

Definition 5: Assume that $\tilde{x}$ is a solution of the NCP(F), then $\tilde{x}$ is said to be a R-regular solution if $F'(\tilde{x})_{\alpha\alpha}$ is nonsingular and the Schur complement of it

$$F'(\tilde{x})_{\beta\beta} - F'(\tilde{x})_{\beta\alpha}F'(\tilde{x})_{\alpha\alpha}^{-1}F'(\tilde{x})_{\alpha\beta}$$

is a P-matrix.

Proposition 6: Assume that $\tilde{x}$ is a R-regular solution of the NCP(F), then
(a) There exists $c_0 > 0$ and $\delta_1 > 0$ such that for all x that satisfy $\|x - \tilde{x}\| \leq \delta_1$, the matrices $H \in \partial\Phi(x)$ are nonsingular and

$$\|H^{-1}\| \leq c_0 \tag{5}$$

(b) there exists $c_1 > 0$ and $b_1 > 0$ such that

$$\|\Phi(x)\| \geq c_1\|x - \tilde{x}\| \tag{6}$$

for all x which satisfies that $\|x - \tilde{x}\| \leq b_1$.

III. Description of Algorithm

In this section, we give the description of our algorithm for nonlinear complementarity problems. For convenience of expression, we denote $g_k := \nabla\Psi(x^k)$. We also use the index set $\mathcal{I} = \{1, \cdots, n\}$ for the variable x. We introduce a continuous forcing function $\rho : \mathbb{R} \to \mathbb{R}$, which satisfies two properties: (a) $\rho(t) \geq 0$ for all $t \in \mathbb{R}$; (b) $\rho(t) = 0$ if and only if $t = 0$.

Now we give the description of our algorithm in details.

Algorithm 3.1

Step 0. Initialization:
Select $x^0 \in \mathbb{R}^n$, and $\lambda, \eta, \nu, \sigma \in (0, 1)$, $\delta > 0$, and the toleration $\epsilon \geq 0$. Set $k := 0$.
Step 1. If $\|\nabla\Psi(x^k)\| \leq \epsilon$, then terminates.
Step 2. Set $\delta_k = \min\{\delta, (\|\Phi(x^k)\|)^\nu\}$, and

$$I_k := \{i \in \mathcal{I} \,\big|\, \|(x_i^k, F_i(x^k))\| \leq \delta_k\}. \tag{7}$$

Step 3. Choose H_k be any element of $\partial\Phi(x^k)$ and write $H_k = (H^k_{\cdot I_k}, H^k_{\cdot \overline{I}_k})$. Correspondingly, denote $g_k = ((g^k_{I_k})^T, (g^k_{\overline{I}_k})^T)^T$. For simplicity, we write $g_k = (g^k_{I_k}, g^k_{\overline{I}_k})$. Compute $d^k_{\overline{I}_k}$ be the solution of the following reduced system:

$$\left((H^k_{\cdot \overline{I}_k})^T H^k_{\cdot \overline{I}_k} + \rho(\Psi(x^k))E_1\right) d_{\overline{I}_k} = -g^k_{\overline{I}_k} \tag{8}$$

where E_1 denotes the identical matrix with $|I_k|$- dimensions. Denote the trail step: $d^k_t = (-x^k_{I_k}, d^k_{\overline{I}_k})^T$. (The subscript 't' means 'trial'.)

Step 4.(Fast Step) If

$$\Psi(x^k + d_t^k) \leq \eta^2 \Psi(x^k) \tag{9}$$

then accept $d^k = d_t^k$, and set $x^{k+1} = x^k + d^k$. Then, goto Step 7. If (9) is not satisfied, goto Step 5.

Step 5.(Adjustment Step) Compute $d_{I_k}^k$ by solving the following system:

$$\begin{aligned} &\left((H_{\cdot I_k}^k)^T H_{\cdot I_k}^k + \rho(\Psi(x^k))E_2\right) d_{I_k} + (H_{\cdot I_k}^k)^T H_{\cdot \overline{I}_k}^k d_{\overline{I}_k}^k \\ =& - g_{I_k}^k \end{aligned} \tag{10}$$

where E_2 is the identical matrix with $|\overline{I}_k|$- dimensions. Set $d^k = (d_{I_k}^k, d_{\overline{I}_k}^k)^T$.

Step 6. Search t_k as the maximal element in $\{\lambda^s \,|\, s = 0, 1, 2, \cdots\}$ such that

$$\Psi(x^k + t_k d^k) \leq (1 - \sigma t_k^2)\Psi(x^k). \tag{11}$$

Then set $x^{k+1} = x^k + t_k d^k$.

Step 7. Set $k := k + 1$, and return to Step 1.

Now we give some explanations about the proposed algorithm. At each iteration, the estimate set I_k is used to approximate the degenerate indices of the solution. We will analyze its property in Section 5.

When the estimate set is constructed, the reduced system (8) is solved to generate the trial direction. This system is motivated by the standard generalized Newton equation

$$H_k d = -\Phi(x^k) \tag{12}$$

and its Levenberg-Marquardt regularized form:

$$(H_k^T H_k + \rho(\Psi(x^k))I)d = -\nabla\Psi(x^k). \tag{13}$$

In Step 4, the algorithm constructs a trial direction d_t^k by setting $x_{I_k}^{k+1} = 0$. If d_t^k gives the merit function $\Psi(\cdot)$ a satisfying descent, then the Fast Step is accepted. If, on the other hand, the descent test (9) is not satisfied, then the Adjustment Step and a line search is carried out to obtain a descent direction.

In [5], the adjustment step is generated through a gradient step. That is, set $d_{I_k}^k = -\min\{g_{I_k}^k, x_{I_k}^k\}$. As is mentioned in Section 1, this designation may cause both theoretical and numerical problems.

In our algorithm, $d_{I_k}^k$ is adjusted by solving the reduced linear system (10). Notice that (10) is a Newtonian equation. Indeed, it is the regularized form of the following reduced system:

$$\begin{aligned} &(H_{\cdot I_k}^k)^T H_{\cdot I_k}^k d_{I_k} + (H_{\cdot I_k}^k)^T H_{\cdot \overline{I}_k}^k d_{\overline{I}_k}^k \\ =& -(H_{\cdot I_k}^k)^T \Phi(x^k) = -g_{I_k}^k \end{aligned} \tag{14}$$

where d_{I_k} is the unknown and $d_{\overline{I}_k}^k$ is the solution of (8). Moreover, note that the standard Newton equation (12) can be rewritten as follows,

$$\begin{aligned} (H_k)^T H_k d &= \begin{pmatrix} (H_{\cdot I_k}^k)^T H_{\cdot I_k}^k & (H_{\cdot I_k}^k)^T H_{\cdot \overline{I}_k}^k \\ (H_{\cdot \overline{I}_k}^k)^T H_{\cdot I_k}^k & (H_{\cdot \overline{I}_k}^k)^T H_{\cdot \overline{I}_k}^k \end{pmatrix} \begin{pmatrix} d_{I_k}^k \\ d_{\overline{I}_k}^k \end{pmatrix} \\ &= -(H_k)^T \Phi(x^k) = -g_k. \end{aligned} \tag{15}$$

Hence the reduced linear system (14) is in fact the second part of (15) or (12). Furthermore, combining (8) with (10), we obtain that if the adjustment step is carried out, the search direction d^k is actually the solution of following equation:

$$\begin{aligned} &\left[\begin{pmatrix} (H_{\cdot I_k}^k)^T H_{\cdot I_k}^k & (H_{\cdot I_k}^k)^T H_{\cdot \overline{I}_k}^k \\ 0 & (H_{\cdot \overline{I}_k}^k)^T H_{\cdot \overline{I}_k}^k \end{pmatrix} + \rho(\Psi(x^k))I\right] d \\ =& -\nabla\Psi(x^k) \end{aligned} \tag{16}$$

which is actually an approximation of (13). Since the standard Newton method has locally superlinear convergence, this approximation can be used to develop a new approach to analyze the local convergence of the algorithm. (16) also illustrates that the total computation amount on solving two reduced linear systems is not more than the standard Newton step (12).

We end this section by the following property.

Proposition 7: For any $d_{\overline{I}_k}^k$ given, suppose that the matrix $H_{\cdot I_k}^k$ is full rank, and d_{I_k} satisfies (14), then d_{I_k} is the solution of the following problem:

$$\min_{d \in \mathbb{R}^{I_k}} \|\Phi(x^k) + H_{\cdot \overline{I}_k}^k d_{\overline{I}_k}^k + H_{\cdot I_k}^k d\|^2. \tag{17}$$

Proof: Since the matrix $H_{\cdot I_k}^k$ is full rank, the matrix $(H_{\cdot I_k}^k)^T H_{\cdot I_k}^k$ is positive definite and the quadratic problem (17) is strictly convex. Notice that d_{I_k} satisfies (14), which is just the optimal condition of the (17). Thus, d_{I_k} is the solution of the quadratic problem. □

Obviously, if the iterative sequence converges to a solution of NCP(F), then $\Psi(x^k)$ must tends to zero. Therefore, Proposition 7 implies that $d_{I_k}^k$ is an asymptotically best adjustment on d^k in sense of the least square norms.

IV. Global Convergence

In this section, we analyze the global properties of the proposed algorithm. Firstly, we will show the algorithm is well defined. Secondly, we will prove the global convergence of the algorithm. Without loss of generality, we assume that the terminate toleration $\epsilon = 0$, and the iteration doesn't terminate in finite steps.

Lemma 8: For any $k \in \mathbb{N}$, if d^k is computed by the 'Adjustment Step', then d^k is a descent direction of the merit function Ψ, i.e,

$$\nabla\Psi^T(x^k)d^k < 0.$$

Proof: Since $d_{\overline{I}_k}^k$ and $d_{I_k}^k$ are the solutions of the reduced equations (8) and (10) respectively, we have

$$\begin{aligned} &\nabla\Psi^T(x^k)d^k = ((H_k)^T\Phi(x^k))^T d^k \\ =\ & (d_{I_k}^k)^T (H_{\cdot I_k}^k)^T \Phi(x^k) + (d_{\overline{I}_k}^k)^T (H_{\cdot \overline{I}_k}^k)^T \Phi(x^k) \\ =\ & -(d_{I_k}^k)^T [\left((H_{\cdot I_k}^k)^T H_{\cdot I_k}^k + \rho(\Psi(x^k))E_2\right) d_{I_k}^k + (H_{\cdot I_k}^k)^T H_{\cdot \overline{I}_k}^k d_{\overline{I}_k}^k] \\ & -(d_{\overline{I}_k}^k)^T [(H_{\cdot \overline{I}_k}^k)^T H_{\cdot \overline{I}_k}^k + \rho(\Psi(x^k))E_1] d_{\overline{I}_k}^k \end{aligned}$$

The third equality holds because of (8) and (10). For simplicity, we denote $z_{I_k}^k = H_{\cdot I_k}^k d_{I_k}^k$, and $z_{\overline{I}_k}^k = H_{\cdot \overline{I}_k}^k d_{\overline{I}_k}^k$.

Hence,

$$\begin{aligned} & \nabla\Psi^T(x^k)d^k \\ = & -[(z^k_{I_k})^T z^k_{I_k} + (z^k_{I_k})^T z^k_{\overline{I}_k} + (z^k_{\overline{I}_k})^T z^k_{\overline{I}_k}] \\ & -\rho(\Psi(x^k))[\|d^k_{\overline{I}_k}\|^2 + \|d^k_{I_k}\|^2] \\ = & -\frac{1}{2}(\|z^k_{I_k}\|^2 + \|z^k_{\overline{I}_k}\|^2 + \|z^k_{I_k} + z^k_{\overline{I}_k}\|^2) \\ & -\rho(\Psi(x^k))\|d^k\|^2 \\ < & 0 \end{aligned} \tag{18}$$

The last inequality holds since the iteration doesn't terminate in finite steps. □

From the proof of this lemma, we also get the following conclusion.

Lemma 9: For any x^* such that $\Psi(x^*) \neq 0$, assume that H^* is any element of $\partial\Phi(x^*)$. Denote $H^* = (H_{\cdot I_*}, H_{\cdot \overline{I}_*})$. Correspondingly, denote $d^* = (d_{I_*}, d_{\overline{I}_*})$, where $d_{\overline{I}_*}$ and d_{I_*} are the solution of (8) and (10) respectively, then $\nabla\Psi^T(x^*)d^* = 0$ if and only if $d^* = 0$.

Proof: If $d^* = 0$, we immediately have $\nabla\Psi^T(x^*)d^* = 0$. Conversely, if $\nabla\Psi^T(x^*)d^* = 0$, denote $z_{I_*} = H_{\cdot I_*}d_{I_*}$, and $z_{\overline{I}_*} = H_{\cdot \overline{I}_*}d_{\overline{I}_*}$, then by (18), we have

$$0 = \nabla\Psi^T(x^*)d^* = -\frac{1}{2}(\|z_{I_*}\|^2 + \|z_{\overline{I}_*}\|^2 + \|z_{I_*} + z_{\overline{I}_*}\|^2) - \rho(\Psi(x^*))\|d^*\|^2$$

Therefore, we obtain that $\rho(\Psi(x^*))\|d^*\|^2 = 0$. Since $\Psi(x^*) \neq 0$, then $d^* = 0$. □

Now we show that Algorithm 3.1 is well defined.

Proposition 10: Algorithm 3.1 is well defined.

Proof: It suffices to prove that, if the 'Adjustment Step' is carried out, then the line search (11) in Step 6 of the algorithm is well defined. By contradiction, suppose that for any $s \geq 0, s \in \mathbb{N}$, we have

$$\Psi(x^k + \lambda^s d^k) > (1 - \sigma\lambda^{2s})\Psi(x^k).$$

This inequality can be rewritten as

$$\frac{\Psi(x^k + \lambda^s d^k) - \Psi(x^k)}{\lambda^s} > -\sigma\lambda^s\Psi(x^k).$$

Forcing $s \to +\infty$, we obtain that

$$\nabla\Psi^T(x^k)d^k \geq 0.$$

However, as is shown in Lemma 8, $\nabla\Psi^T(x^k)d^k < 0$. This is a contradiction. Hence, the line search (11) is well defined. □

We now come to prove the global convergence of the algorithm. We shall show that any accumulation point of the iterative sequence is at least a stationary point of the system (2).

Theorem 11: Every accumulation point of the iterative sequence $\{x^k\}$ generated by Algorithm 3.1 is a stationary point of the equation (2).

Proof: Obviously, $\{\Psi(x^k)\}$ is a positive and strictly decreasing sequence. Then it must converge to some limit $\Psi^* \geq 0$. For any accumulation point x^*, it must be $\Psi(x^*) = \Psi^*$. If the 'Fast Step' is accepted infinitely, by (9), we get $\Psi^* = 0$. Hence, we have $\Psi(x^*) = 0$. That is, every accumulation point is a solution of the NCP(F).

If, on the other hand, d^k is eventually computed always by the 'Adjustment Step', we have to discuss two cases. Firstly, in the case when $\Psi^* = 0$, we still have the conclusion that any accumulation point is a solution of the problem. Secondly, if $\Psi^* > 0$, the proof is by contradiction.

Suppose that x^* is an accumulation point of the iterative sequence with a subsequence $\{x^k\}_K$ converging to it. Now we assume that $\nabla\Psi(x^*) \neq 0$. We can choose a subsequence $\{x^k\}_{K_1}$, $K_1 \subseteq K$, such that for all $k \in K_1$, I_k equals to a fixed set $J \subseteq I$. Since $\partial\Phi$ is upper semicontinuous and closed at x^* (see Proposition 2.6.2 of [9]), we can also find a subsequence $\{x^k\}_{K_2}$, $K_2 \subseteq K_1$, such that $\{H_k\}_{K_2} \to H^*$, of which the limit $H^* \in \partial\Phi(x^*)$.

Combining (8) with (10), we obtain that d^k is the unique solution of the following linear system

$$\left[\begin{pmatrix} (H^k_{\cdot I_k})^T H^k_{\cdot I_k} & (H^k_{\cdot I_k})^T H^k_{\cdot \overline{I}_k} \\ 0 & (H^k_{\cdot \overline{I}_k})^T H^k_{\cdot \overline{I}_k} \end{pmatrix} + \rho(\Psi(x^k))I\right] d = -\nabla\Psi(x^k) \tag{19}$$

Taking the limit $k \to \infty, k \in K_2$, and let d^* be the unique solution of the next equation,

$$\left[\begin{pmatrix} (H^*_{\cdot J})^T H^*_{\cdot J} & (H^*_{\cdot J})^T H^*_{\cdot \overline{J}} \\ 0 & (H^*_{\cdot \overline{J}})^T H^*_{\cdot \overline{J}} \end{pmatrix} + \rho(\Psi^*)I\right] d = -\nabla\Psi(x^*) \tag{20}$$

Therefore, $\{d^k\}_{K_2} \to d^*$. Since $\nabla\Psi(x^*) \neq 0$, this immediately implies that $d^* \neq 0$.

On the other hand, by the linear search (11), we obtain that

$$\lim_{k\in K_2, k\to\infty} t_k^2 \leq \lim_{k\in K_2, k\to\infty} \frac{\Psi(x^k) - \Psi(x^{k+1})}{\sigma\Psi(x^k)} = 0$$

which implies that $\{t_k\}_{K_2} \to 0$. Consequently, for all $k \in K_2$ large enough, we have

$$\frac{\Psi(x^k + \lambda^{t_k - 1}d^k) - \Psi(x^k)}{\lambda^{t_k-1}} > -\sigma\lambda^{t_k-1}\Psi(x^k).$$

Forcing $k \to \infty, k \in K_2$, we get

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

This, together with Lemma 8, implies that $\nabla\Psi^T(x^*)d^* = 0$. By (20), this forces that $d^* = 0$. However, we have shown that $d^* \neq 0$. Hence, this is a contradiction. □

V. Local Convergence

In this section, we establish the locally superlinear convergence of Algorithm 3.1. We begin with the following lemma.

Lemma 12: Let $\{x^k\}$ be a sequence of generated by Algorithm 3.1, and assume that x^* is an R-regular solution of NCP(F). Then, there exists some scalar c_2 such that

$$\|\left((H^k_{\cdot I_k})^T H^k_{\cdot I_k}\right)^{-1}\| \leq c_2, \|\left((H^k_{\cdot \overline{I}_k})^T H^k_{\cdot \overline{I}_k}\right)^{-1}\| \leq c_2$$
$$\|\left((H^k)^T H^k\right)^{-1}\| \leq c_2$$

for all $x^k \in \mathbb{R}^n$ which is sufficiently close to x^*.

Proof: It follows from Proposition 6(a) that, in a small neighbor of x^*, the matrices H_k are all nonsingular. Hence, the matrices $(H^k_{\cdot I_k})^T H^k_{\cdot I_k}$, $(H^k_{\cdot \overline{I}_k})^T H^k_{\cdot \overline{I}_k}$, and $(H^k)^T H^k$ are all nonsingular. The remainder of this proof is the same as Lemma 5 in [5]. □

Lemma 13: Assume that x^* is an accumulation point and is an R-regular solution of NCP(F). Furthermore, suppose that $\{x^k\}_K$ is a subsequence converges to x^* and the corresponding sequence $\{d^k\}_K$ are always computed by 'Adjustment Step', then there exists some $c_3 > 0$ such that for all $k \in K$ large enough

$$\|d^k\| \le c_3 \|\Phi(x^k)\|.$$

Proof: By (8), we have

$$\begin{aligned} 0 &= (d^k_{\overline{I}_k})^T [H_k^T \Phi(x^k)]_{\overline{I}_k} + (d^k_{\overline{I}_k})^T [(H^k_{\cdot I_k})^T H^k_{\cdot \overline{I}_k} \\ &\quad + \rho(\Psi(x^k)) E_1] d^k_{\overline{I}_k} \\ &\ge \|H^k_{\cdot \overline{I}_k} d^k_{\overline{I}_k}\|^2 + (d^k_{\overline{I}_k})^T (H^k_{\cdot \overline{I}_k})^T \Phi(x^k) \\ &\ge \|H^k_{\cdot \overline{I}_k} d^k_{\overline{I}_k}\|^2 - \|H^k_{\cdot \overline{I}_k} d^k_{\overline{I}_k}\| \|\Phi(x^k)\| \end{aligned}$$

Therefore, for all x^k close to x^* enough, we have

$$\|H^k_{\cdot \overline{I}_k} d^k_{\overline{I}_k}\| \le \|\Phi(x^k)\|.$$

Together with lemma 12, we have

$$\begin{aligned} \|d^k_{\overline{I}_k}\| &\le \|\left((H^k_{\cdot I_k})^T H^k_{\cdot I_k}\right)^{-1}\| \|(H^k_{\cdot I_k})^T H^k_{\cdot I_k} d^k_{\overline{I}_k}\| \\ &\le c_2 \|H^k_{\cdot I_k}\| \|H^k_{\cdot I_k} d^k_{\overline{I}_k}\| \\ &\le c_2 \|H_k\| \|\Phi(x^k)\| \end{aligned} \tag{21}$$

Since the generalized Jacobian has upper semicontinuity, there exists a scalar $\kappa > 0$ such that $\|H_k\| \le \kappa$ for all $k \in K$ large enough. Let $c_4 = c_2\kappa$, we obtain

$$\|d^k_{\overline{I}_k}\| \le c_4 \|\Phi(x^k)\| \tag{22}$$

for $k \in K$ large enough.

On the other hand, by the upper semicontinuity of $\partial\Phi$, the sequence $\{\|(H^k_{\cdot I_k})^T H^k_{\cdot \overline{I}_k}\|\}_K$ is bounded. Suppose its upper bound is $\kappa_1 > 0$, then we get

$$\|(H^k_{\cdot I_k})^T H^k_{\cdot \overline{I}_k} d^k_{\overline{I}_k}\| \le \kappa_1 c_4 \|\Phi(x^k)\|, \quad k \in K.$$

Moreover, there exists a constant $\kappa_2 > 0$ such that

$$\|g^k_{I_k}\| \le \|g_k\| \le \|H_k\| \|\Phi(x^k)\| \le \kappa_2 \|\Phi(x^k)\|.$$

Therefore, by (10), we have

$$\begin{aligned} \|d^k_{I_k}\| \le &(\kappa_1 c_4 + \kappa_2) \| \left((H^k_{\cdot I_k})^T H^k_{\cdot I_k} + \rho(\Psi(x^k)) E_2\right)^{-1} \| \\ &\cdot \|\Phi(x^k)\|, \quad k \in K. \end{aligned}$$

Notice that $\rho(\Psi(x^k)) \to 0, k \in K$, it follows from Lemma 12 that

$$\|d^k_{I_k}\| \le 2c_2(\kappa_1 c_4 + \kappa_2) \|\Phi(x^k)\| \tag{23}$$

for $k \in K$ large enough. From (22) and (23), we get

$$\|d^k\| \le \|d^k_{I_k}\| + \|d^k_{\overline{I}_k}\| \le [c_4 + 2c_2(\kappa_1 c_4 + \kappa_2)] \|\Phi(x^k)\|.$$

Then set $c_3 := c_4 + 2c_2(\kappa_1 c_4 + \kappa_2)$, the proof is completed. □

Now, we prove the convergence of entire iterative sequence $\{x^k\}$. We first cite the following conclusion, of which the proof can be found in [10].

Proposition 14: ([5]) Suppose that $w^* \in \mathbb{R}^t$ is an isolated accumulation point of the sequence $\{w^k\} \subseteq \mathbb{R}^t$. If for any subsequence $\{w^k\}_J$ converging to w^*, there exists an infinite subsequence $\tilde{J} \subseteq J$ such that $\{\|w^{k+1} - w^k\|\}_{\tilde{J}} \to 0$, then the entire sequence $\{w^k\}$ converges to w^*.

Lemma 15: Assume that x^* is an R-regular solution of NCP(F) and is an accumulation point of the iterative sequence $\{x^k\}$ generated by Algorithm 3.1, then the whole sequence $\{x^k\}$ converges to x^*.

Proof: Suppose that $\{x^k\}_K$ is a subsequence converging to x^*. Obviously, we have

$$\{\|\Phi(x^k)\|\}_K \to \|\Phi(x^*)\| = 0 \tag{24}$$

If there exists an infinite set K_1 for which the adjustment steps are taken. By Lemma 13, we get

$$\{\|d^k\|\}_{K_1} \to 0.$$

This immediately shows that

$$\{\|x^{k+1} - x^k\|\}_{K_1} \to 0.$$

If the adjustment steps are carried out only finite times on K. Without loss of generality, suppose that the fast steps are always taken on $k \in K$, and I_k equals to a fixed set $J \subseteq I$. By (22), we obtain

$$\{\|d^k_{\overline{J}}\|\}_K \to 0.$$

Since $\{x^k\}_K$ converges to x^*, it holds that $\{x^k_J\}_K \to x^*_J$. Therefore, we also obtain that

$$\{\|x^{k+1} - x^k\|\}_K \to 0.$$

Hence, there always exists a subsequence K on which $x^{k+1} - x^k \to 0$. It follows from Proposition 14 that the entire sequence converges to x^*. □

We next establish the relationship between the estimate set I_k and the degenerate set β, that is, I_k is eventually equal to β under the R-regular condition.

Lemma 16: ([3]) Suppose that x^* is an R-regular solution of NCP(F), and the entire sequence $\{x^k\}$ generated by Algorithm 3.1 converges to x^*, then $I_k = \beta$ for all k large enough.

Assumption 17: The sequence $\{H_k\}$ satisfies

$$\frac{\|P_k H^k_{\cdot \beta} d^k_\beta\|}{\|d^k\|} \to 0$$

where P_k is the projector in $\mathbb{R}^n$ onto $H^k_{\cdot \overline{\beta}}$, that is,

$$P_k = H^k_{\cdot \overline{\beta}} [(H^k_{\cdot \overline{\beta}})^T H^k_{\cdot \overline{\beta}}]^{-1} (H^k_{\cdot \overline{\beta}})^T.$$

To analyze the locally superlinear convergence, we need the following result.

Proposition 18: ([11]) Let $G : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitzian and $x^* \in \mathbb{R}^n$ with $G(x^*) = 0$ such that all elements in $\partial G(x^*)$ are nonsingular, and assume that there are two sequence $\{x^k\} \subset \mathbb{R}^n$ and $\{d^k\} \subset \mathbb{R}^n$ with

$$\lim_{k \to \infty} x^k = x^* \qquad \text{and} \qquad \|x^k + d^k - x^*\| = o(\|x^k - x^*\|)$$

Then

$$\|G(x^k + d^k)\| = o(\|G(x^k)\|).$$

Denote $K = \{k|k \in \mathbb{N}$,the 'Adjustment Step' is taken at the kth step$\}$. We will prove in Theorem 20 that if Assumption 17 holds, then K must be a finite set.

Lemma 19: Suppose that x^* is an R-regular solution of NCP(F), and the index set K is infinite. Denote the sequence $\{\tilde{d}^k\}_K$, where $\tilde{d}^k$ is generated by solving the following equation

$$((H_k)^T H_k + \rho(\Psi(x^k))I)d = -(H_k)^T \Phi(x^k). \quad (25)$$

Suppose Assumption 17 holds, then $\tilde{d}^k = d^k + o(\|d^k\|), (k \in K)$, where d^k is computed by the 'Adjustment Step' of Algorithm 3.1.

Proof: Since Assumption 17 holds, then

$$(H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\beta} d^k_\beta = o(\|d^k\|). \quad (26)$$

It follows from (19) and (25) that, for $k \in K$ large enough,

$$\begin{aligned}
& ((H_k)^T H_k + \rho(\Psi(x^k))I)d^k \\
= & \left[\begin{pmatrix} (H^k_{\cdot\beta})^T H^k_{\cdot\beta} & (H^k_{\cdot\beta})^T H^k_{\cdot\bar{\beta}} \\ (H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\beta} & (H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\bar{\beta}} \end{pmatrix} + \rho(\Psi(x^k))I\right] \begin{pmatrix} d^k_\beta \\ d^k_{\bar{\beta}} \end{pmatrix} \\
= & \left[\begin{pmatrix} (H^k_{\cdot\beta})^T H^k_{\cdot\beta} & (H^k_{\cdot\beta})^T H^k_{\cdot\bar{\beta}} \\ 0 & (H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\bar{\beta}} \end{pmatrix} + \rho(\Psi(x^k))I\right] \begin{pmatrix} d^k_\beta \\ d^k_{\bar{\beta}} \end{pmatrix} \\
& +o(\|d^k\|) \\
= & -\nabla\Psi(x^k) + o(\|d^k\|) \\
= & ((H_k)^T H_k + \rho(\Psi(x^k))I)\tilde{d}^k + o(\|d^k\|).
\end{aligned}$$

Since x^* is an R-regular solution of NCP(F), it follows from Lemma 12 and $\{\rho(\Psi(x^k))\}_K \to 0$ that

$$d^k = \tilde{d}^k + o(\|d^k\|),$$

which is just what we want to prove. □

Theorem 20: Suppose that Assumption 17 holds and x^* is an accumulation point of the sequence $\{x^k\}$ generated by Algorithm 3.1. If x^* is an R-regular solution of NCP(F), then

(i) The whole sequence $\{x^k\}$ converges to x^*.

(ii) For all k large enough, the 'Fast Step' of Algorithm 3.1 is always accepted.

(iii) The sequence $\{x^k\}$ converges to x^* Q-superlinearly. Furthermore, if F' is Lipschitz continuous and $\rho(\Psi(x^k)) = O(\|\Phi(x^k)\|)$, the rate of convergence is Q-quadratic.

Proof: The first assertion is just Lemma 15. To prove the second statement, we first assume by contradiction that K is an infinite set. From (25), we obtain that for $k \in K$

$$\begin{aligned}
& (H_k^T H_k + \rho(\Psi(x^k))I)(x^k + \tilde{d}^k - x^*) \\
= & (H_k^T H_k + \rho(\Psi(x^k))I)(x^k - x^*) - H_k^T \Phi(x^k) \\
= & -H_k^T[\Phi(x^k) - \Phi(x^*) - H_k(x^k - x^*)] \\
& +\rho(\Psi(x^k))(x^k - x^*) \quad (27)
\end{aligned}$$

Consequently, by Proposition 2 and Lemma 12 we obtain

$$\|x^k + \tilde{d}^k - x^*\| = o(\|x^k - x^*\|), \quad k \in K. \quad (28)$$

Therefore, it follows from Lemma 19 that

$$\|x^k + d^k - x^* + o(\|d^k\|)\| = o(\|x^k - x^*\|), \quad k \in K.$$

which implies that

$$\|x^k + d^k - x^*\| = o(\|x^k - x^*\|), \quad k \in K. \quad (29)$$

Furthermore, we obtain from (29) that $\|x^k_\beta + d^k_\beta - x^*_\beta\| = o(\|x^k - x^*\|)$, $(k \in K)$. Recall that in Step 3 of Algorithm 3.1, we defined the trial step $d^k_t = (-x^k_\beta, d^k_{\bar{\beta}})^T$. It follows that $\|x^k + d^k_t - x^*\| = o(\|x^k - x^*\|)$, for $k \in K$. Thus, by Proposition 18, we can say that

$$\|\Phi(x^k + d^k_t)\| = o(\|\Phi(x^k)\|). \quad (30)$$

Therefore, for all $k \in K$ large enough, condition (9) of Step 4 is always satisfied, which means $k \notin K$. This is a contradiction. Hence, the index set K is finite, and the second statement holds.

Finally, we prove the locally superlinear convergence of the algorithm. Since the 'Fast Step' is eventually always accepted, then we have $x^k_\beta = 0$ and

$$\|x^k + d^k - x^*\| = \|x^k_{\bar{\beta}} + d^k_{\bar{\beta}} - x^*_{\bar{\beta}}\|. \quad (31)$$

for $k \in \mathbb{N}$ large enough.

On the other hand, from (8), we have

$$\begin{aligned}
& [(H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\bar{\beta}} + \rho(\Psi(x^k))E_1](x^k_{\bar{\beta}} + d^k_{\bar{\beta}} - x^*_{\bar{\beta}}) \\
= & -[H_k^T \Phi(x^k)]_{\bar{\beta}} + (H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\bar{\beta}}(x^k_{\bar{\beta}} - x^*_{\bar{\beta}}) + \rho(\Psi(x^k))(x^k_{\bar{\beta}} - x^*_{\bar{\beta}}) \\
= & -(H^k_{\cdot\bar{\beta}})^T[\Phi(x^k) - \Phi(x^*) - H^k_{\cdot\bar{\beta}}(x^k_{\bar{\beta}} - x^*_{\bar{\beta}})] \\
& +\rho(\Psi(x^k))(x^k_{\bar{\beta}} - x^*_{\bar{\beta}}) \\
= & -(H^k_{\cdot\bar{\beta}})^T[\Phi(x^k) - \Phi(x^*) - H_k(x^k - x^*)] \\
& +\rho(\Psi(x^k))(x^k_{\bar{\beta}} - x^*_{\bar{\beta}})
\end{aligned}$$

Notice that $\rho(\Psi(x^k)) \to 0$, we obtain from Lemma 12 that for $k \in \mathbb{N}$ large enough,

$$\|((H^k_{\cdot\bar{\beta}})^T H^k_{\cdot\bar{\beta}} + \rho(\Psi(x^k))E_1)^{-1}\| \le 2c_2.$$

Therefore, it follows from Proposition 2 that

$$\begin{aligned}
\|x^k_{\bar{\beta}} + d^k_{\bar{\beta}} - x^*_{\bar{\beta}}\| \le & 2c_2\Big[\|H^k_{\cdot\bar{\beta}}\| \cdot \|\Phi(x^k) - \Phi(x^*) - H_k(x^k - x^*)\| \\
& +\rho(\Psi(x^k))\|x^k - x^*\|\Big] \\
= & o(\|x^k - x^*\|). \quad (32)
\end{aligned}$$

Furthermore, if F' is Lipschitz continuous and $\rho(\Psi(x^k)) = O(\|\Phi(x^k)\|)$, from (31), (32) and Proposition 2(ii), it holds that

$$\|x^k + d^k - x^*\| = O(\|x^k - x^*\|)^2.$$

Hence, the last statement is proved. □

VI. Numerical Results

MATLAB is one of the main platforms used to test numerical methods on NCP(F). In this section, we implemented the program code of the proposed algorithm in MATLAB 7.5 on a Intel Pentium 4 and report the numerical results on some complementarity problems to show the efficiency of the proposed algorithm. We set the parameters as $\lambda = 0.5$, $\eta = 0.8$, $\nu = 0.6$, $\delta = 1$, $\sigma = 0.15$, and set the toleration $\epsilon = 1.0e-6$. The forcing function is set as $\rho(t) = \sqrt{t}$.

The generalized Jacobian $H \in \partial\Phi(x)$ can be chosen as follows:
(i) Let $\xi = \{j | x_j = 0 = F_j(x)\}$.
(ii) Choose any $z \in \mathbb{R}^n$ such that $z_j \neq 0$ if $j \in \xi$.
(iii) Let $H = (H_1, H_2, \cdots, H_n)^T$ where

$$H_j = (\frac{x_j}{\sqrt{x_j^2 + F_j(x)^2}} - 1)e_j + (\frac{F_j(x)}{\sqrt{x_j^2 + F_j(x)^2}} - 1)\nabla F_j(x)$$

if $j \notin \xi$, and

$$\begin{aligned} H_j \quad =& \Big(\frac{z_j}{\sqrt{z_j^2 + (\nabla F_j(x)^T z)^2}} - 1\Big)e_j \\ &+ \Big(\frac{\nabla F_j(x)^T z}{\sqrt{z_j^2 + (\nabla F_j(x)^T z)^2}} - 1\Big)\nabla F_j(x) \end{aligned}$$

if $j \in \xi$.

This technique is from [12], in which the matrix H constructed above is proved to be an element of $\partial\Phi(x)$. For practical computation, we further redefine the set ξ,i.e, $\xi = \{j | \sqrt{x_j^2 + F_j(x)^2} \leq 10^{-6}\}$.

Similar with [5], [13], we introduce the nonmonotone linear search to improve the computation efficiency. That is, we replace the linear search condition (11) by the next condition

$$\Psi(x^k + t_k d^k) \leq M_k - \sigma t_k^2 \Psi(x^k). \tag{33}$$

where M_k is defined by

$$M_k = \max_{l=k-m_k,\cdots,k} \Psi(x^l),$$

and $m_{k+1} = \min\{m_k + 1, 5\}$, $m_0 = 0$.

We first describe some computational results of some degenerate NCPs of small dimensions. The test problems are as follows:

Problem 1. This is an linear complementarity problem(LCP). Here $n = 4$ and define $F(\cdot)$ as

$$F(x) = \begin{bmatrix} 1 - x_1 + x_2 + x_3 \\ x_1 - 1 \\ x_4 - 1 \\ 1 + x_3 - x_4 \end{bmatrix}$$

This problem is from [14] and has a unique degenerate solution $x^* = (1, 0, 0, 1)^T$.

Problem 2. This example is Kojima-Shindo problem from [15]. Let $n = 4$ and

$$F(x) = \begin{bmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{bmatrix}.$$

It has two solutions: a degenerate solution $x^1 = (\frac{\sqrt{6}}{2}, 0, 0, 0.5)^T$ and a non-degenerate solution $x^2 = (1, 0, 3, 0)^T$.

Problem 3. Here $n = 2$. This problem is example 6.1 from [16], it has a degenerate solution $x^1 = (1, 0)^T$ and a non-degenerate solution $x^2 = (0, \frac{\sqrt{5}-1}{2})^T$. The function $F(\cdot)$ is defined as follows

$$F(x) = \begin{bmatrix} (x_1 - 1)^2 \\ x_1 + x_2 + x_2^2 - 1 \end{bmatrix}.$$

Problem 4. Here $n = 3$ and set $F(\cdot)$ as

$$F(x) = \begin{bmatrix} x_1 - 2 \\ x_2 - x_1 - x_3 + x_2^3 + 3 \\ x_2 + x_3 + 2x_3^3 - 3 \end{bmatrix}.$$

This problem has a degenerate solution $(2, 0, 1)^T$.

Problem 5. This example is modified from Mathiesen's [17]. It has a family of solutions $(\varpi, 0, 0, 0)^T$, where $\varpi \in [0, 3]$. If $\varpi = 0, 1, 3$, the solution is degenerate.

$$F(x) = \begin{bmatrix} -x_2 + x_3 + x_4 \\ x_1 - (4.5x_3 + 2.7x_4)(x_2 + 1) \\ 1 - x_1 - (0.5x_2 + 0.3x_4)(x_3 + 1) \\ 3 - x_1 \end{bmatrix}.$$

Problem 6. In this problem, we set $n = 3$. The only solution of this example is $x^* = (1, 3, 0)^T$, which is degenerate. The definition of $F(\cdot)$ is as follows.

$$F(x) = \begin{bmatrix} x_2 - x_1 - 2 \\ x_1^2 - x_3 - 1 \\ 3x_1^3 - x_2 + x_3^2 \end{bmatrix}.$$

The computational results of these examples are illustrated in Table 1, in which **Iter** means the number of iterations, **NF** denotes the number of evaluations of the function F, the column **Fast** lists the number of 'Fast Step' taken during the iteration, and '$\boldsymbol{I_k = \beta}$' denotes the iteration from which the degenerate set is correctly estimated.

The numerical results in Table 1 shows that our proposed algorithm is very effective. As is shown in **Iter** and **NF**, for most initial points, the algorithm converges to the solution very quickly and the amount of computation on the function F is very small. On the other hand, the column $\boldsymbol{I_k = \beta}$ shows that the active-set method plays an important role on improving the efficiency of the algorithm. In general, the degenerate indices of the solution are identified only after several iterations. Moreover, the column **Fast** shows that 'Fast Step' is accepted for most iteration steps.

We also test our algorithm for all the NCPs in MCPLIB, which is a collection of large scale nonlinear mixed complementarity problems (see [18]). For comparison, we also implemented the numerical experiments on the underlying active-set semismooth method which employing the gradient step in the "adjustment step". For simplification, the algorithm using the gradient step is denoted as Algorithm II. The results are listed in Table II and III respectively, in which the meaning of every columns are the same as in Table 1.

TABLE I.
NUMERICAL RESULTS FOR ALGORITHM 3.1

Problem	Start Point	Iter	NF	$\Psi(x)$	Fast	$I_k = \beta$
Pro 1	(2,4,1,5)	16	17	1.5e-15	15	3
Pro 1	(5,5,-5,0)	11	11	3.5e-15	11	2
Pro 1	(100,1,100,1)	19	38	2.4e-15	13	7
Pro 1	(10,10,10,10)	15	15	1.4e-15	15	3
Pro 2	(1,2,3,4)	11	16	2.0e-09	6	8
Pro 2	(5,0,0,5)	19	22	8.9e-14	14	12
Pro 2	(-5,3,-1,-5)	11	15	4.7e-18	4	5
Pro 2	(1,8,2,10)	54	58	1.3e-17	16	50
Pro 3	(1.5,-0.5)	4	5	3.3e-33	3	3
Pro 3	(3,3)	6	6	5.8e-20	4	4
Pro 3	(8,2)	5	5	1.2e-27	4	3
Pro 3	(4,6)	5	5	8.5e-17	5	0
Pro 4	(-1,-3,-5)	12	12	5.5e-19	10	0
Pro 4	(0,4,0)	9	10	5.2e-27	6	5
Pro 4	(-100,100,100)	12	12	1.1e-28	9	4
Pro 4	(6,6,6)	11	11	2.0e-16	8	0
Pro 5	(-1,-2,-3,-4)	7	7	6.2e-16	7	6
Pro 5	(5,5,5,5)	8	8	1.1e-28	8	4
Pro 5	(8,6,4,2)	8	10	3.6e-21	6	6
Pro 5	(2,4,6,8)	49	49	7.3e-16	49	3
Pro 6	(-3,6,-5)	26	29	1.3e-14	9	15
Pro 6	(3,2,1)	11	11	1.1e-14	9	7
Pro 6	(2,2,2)	4	4	8.1e-18	4	3
Pro 6	(9,9,9)	30	34	1.1e-14	7	9

As is shown in Table II and III, Algorithm 3.1 can solve almost all the MCPLIB test problems, and the numerical efficiency is high. For Algorithm 3.1, there are 3 problems fails to compute a result, while for Algorithm II, there are 4 problems fails to compute a result. Furthermore, most of the test problems are solved with fewer iterations and function evaluations by Algorithm 3.1 than by Algorithm II. Finally, from the Column $\boldsymbol{I_k = \beta}$, Algorithm 3.1 has better robustness on identifying the degenerate indices. The numerical result of Table II is consistent with the theoretical analysis. That is, the Newton step is much better than the gradient step in the 'Adjustment Step'.

VII. CONCLUSION

In this paper, we apply the active-set technology to identify the degenerate indices of the nonlinear complementarity problems, and incorporate this technique into a semismooth algorithm for solving NCP(F). In the new proposed method, we introduce a Newtonian adjustment direction, instead of a gradient direction, when the trial search direction can't guarantee the sufficient reduction of the merit function. Numerical results illustrates that the new algorithm is efficient.

We believe that this method can be used to solve other problems, e.g. the mixed nonlinear complementarity problems. We will focus on this problem in further research.

TABLE II.
NUMERICAL RESULTS OF ALGORITHM 3.1 FOR THE MCPLIB PROBLEMS

Problem	Alg3.1			
	Iter	NF	Fast	$I_k = \beta$
bertsekas(1)	43	319	35	41
bertsekas(2)	42	351	34	32
bertsekas(3)	17	21	8	13
billups	134	179	121	0
colvnlp(1)	44	52	25	41
colvnlp(2)	25	27	10	14
colvnlp(3)	21	23	14	18
colvdual	31	48	22	28
cycle	3	4	1	0
explcp	6	6	0	4
hanshoop	-	-	-	-
josephy(1)	24	31	14	20
josephy(2)	7	10	2	0
kojshin	1	11	1	11
mathisum	9	9	1	4
nash(1)	8	8	0	0
nash(2)	9	14	4	0
ne-hard	7	8	1	4
powell(1)	15	16	13	11
powell(2)	13	13	1	12
powell(3)	10	10	1	8
pgvon106(1)	47	324	46	47
pgvon106(2)	26	41	22	26
pgvon106(3)	-	-	-	-
scarfanum	-	-	-	-
scarfbnum	4	4	0	0
sppe(1)	10	10	0	9
sppe(2)	6	6	0	4
sppe(3)	10	10	0	6
tobin(1)	10	11	3	6
tobin(2)	9	10	1	5
tobin(3)	2	2	0	0

REFERENCES

[1] M. C. Ferris and J. S. Pang, "Engineering and Economic Applications of Complementarity Problems," *SIAM Rev*, vol. 39, no. 4, pp. 669–713, 1997.

[2] N. Yamashita, H. Dan, and M. Fukushima, "On the identification of degenerate indices in the nonlinear complementarity problem with the proximal point algorithm," *Math. Program*, vol. 99, no. 2, pp. 377–397, 2004.

[3] H. D. Yu and D. G. Pu, "Smoothing newton method for ncp with the identification of degenerate indices," *J. Comput. Appl. Math*, vol. 234, no. 12, pp. 3424–3435, 2010.

[4] F. Facchinei, A. Fischer, and C. Kanzow, "On the accurate identification of active constraints," *SIAM. J. Optim*, vol. 9, no. 1, pp. 14–32, 1998.

[5] C. Kanzow and H. D. Qi, "A qp-free constrained newton-type method for variational inequality problems," *Math. Program*, vol. 85, no. 1, pp. 81–106, 1999.

[6] C. Kanzow and H. Pieper, "Jacobian smoothing methods for nonlinear complementarity problems," *SIAM. J. Optim*, vol. 9, no. 2, pp. 342–373, 1999.

[7] S. M. Robinson, "Strongly regular generalized equations," *Math. Oper. Res*, vol. 5, no. 1, pp. 43–62, 1980.

[8] F. Facchinei, A. Fischer, and C. Kanzow, "Regularity properties of a semismooth reformulation of variational inequalities," *SIAM. J. Optim*, vol. 8, no. 3, pp. 850–869, 1998.

[9] F. H. Clark, *Optimization and Nonsmooth Analysis*. New York: John Wiley, 1983.

[10] J. J. Moré and D. C. Sorensen, "Computing a trust region

TABLE III.
NUMERICAL RESULTS OF ALGORITHM II FOR THE MCPLIB PROBLEMS

Problem	Alg II			
	Iter	NF	Fast	$I_k = \beta$
bertsekas(1)	48	369	41	35
bertsekas(2)	39	302	32	23
bertsekas(3)	110	439	82	105
billups	134	179	121	0
colvnlp(1)	59	236	41	56
colvnlp(2)	40	69	17	31
colvnlp(3)	69	338	53	61
colvdual	2523	34896	2515	2515
cycle	3	4	1	0
explcp	6	6	0	4
hanshoop	-	-	-	-
josephy(1)	10	13	3	6
josephy(2)	7	10	2	0
kojshin	12	12	1	12
mathisum	11	11	2	4
nash(1)	8	8	0	0
nash(2)	9	15	5	0
ne-hard	7	8	1	4
powell(1)	25	86	11	21
powell(2)	7	7	0	5
powell(3)	25	25	1	8
pgvon106(1)	79	434	77	72
pgvon106(2)	-	-	-	-
pgvon106(3)	84	899	81	84
scarfanum	-	-	-	-
scarfbnum	-	-	-	-
sppe(1)	10	10	0	9
sppe(2)	7	7	0	4
sppe(3)	11	11	0	11
tobin(1)	11	14	2	6
tobin(2)	12	11	1	7
tobin(3)	3	3	0	0

step," *SIAM.J.Sci.Stat.Comput*, vol. 4, no. 2, pp. 553–572, 1983.

[11] Y. F. Yang and L. Q. Qi, "Smoothing trust region methods for nonlinear complementarity problems with p_0-functions," *Annuals Operations Research*, vol. 133, no. 1-4, pp. 99–117, 2005.

[12] T. D. Luca, F. Facchinei, and C. Kanzow, "A semismooth equation approach to the solution of nonlinear complementarity problems," *Math Program*, vol. 75, no. 3, pp. 407–439, 1996.

[13] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for newton's method," *SIAM. J. Numer. Anal*, vol. 23, no. 4, pp. 707–716, 1986.

[14] C. Oberlin and S. J. Wright, "An accelerated newton method for equations with semismooth jacobians and nonlinear complementarity problems," *Math. Program.*, vol. 117, no. 1-2, pp. 355–386, 2009.

[15] J. S. Pang and S. A. Gabriel, "Ne/sqp: A robust algorithm for the nonlinear complementarity problem," *Math. Program*, vol. 60, no. 1-3, pp. 295–337, 1993.

[16] A. N. Daryina, A. F. Izmailov, and M. V. Solodov, "A class of active-set newton methods for mixed complementarity problems," *SIAM J. Optim*, vol. 15, no. 2, pp. 409–429, 2004.

[17] L. Mathiesen, "An algorithm based on a sequence of a linear complementarity problems applied to a walrasion equilibrium model: An example," *Math. Program*, vol. 37, no. 1, pp. 1–18, 1987.

[18] S. Derkse and M. Ferris, "Mcplib: a collection of nonlinear mixed complementarity problems. optim," *Methods Softw*, vol. 5, no. 4, pp. 319–345, 1995.

Haodong Yu received his Ph.D. degree in applied mathematics from Tongji University, Shanghai in 2011. He received his MS degree in operational research from Tongji University in 2008.

He is currently a lecturer in School of Mathematics and Information, Shanghai Lixin University of Commerce. His research interests include numerical algorithms in optimization and statistics.