A Superlinearly Convergent Algorithm for the Monotone Nonlinear Complementarity Problem without Uniqueness and Nondegeneracy Conditions[1]

Hiroshige Dan[2], Nobuo Yamashita[2] and Masao Fukushima[2]

April 9, 2001

Abstract

The purpose of this paper is to present an algorithm for solving the monotone nonlinear complementarity problem (NCP) that enjoys superlinear convergence in a genuine sense without the uniqueness and nondegeneracy conditions. Recently, Yamashita and Fukushima proposed a method based on the proximal point algorithm (PPA) for monotone NCP. The method has the favorable property that a generated sequence converges to the solution set of NCP superlinearly. However, when a generated sequence converges to a degenerate solution, the subproblems may become computationally expensive and the method does not have genuine superlinear convergence. More recently, Yamashita, Dan and Fukushima presented a technique to identify whether a solution is degenerate or not. Using this technique, we construct a differentiable system of nonlinear equations whose solution is a solution of the original NCP. Moreover, we propose a hybrid algorithm which is based on the PPA and uses this system. We show that the proposed algorithm has a genuine quadratic or superlinear rate of convergence even if it converges to a solution that is neither unique nor nondegenerate.

[1]This research was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan.

[2]Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan.

1 Introduction

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

$$\mathrm{NCP}(F): \quad \bar{x}_i \geq 0, \quad F_i(\bar{x}) \geq 0, \quad \bar{x}_i F_i(\bar{x}) = 0, \quad i = 1, \ldots, n,$$

where F is a mapping from $\Re^n$ to $\Re^n$. When F is affine, NCP(F) is called the linear complementarity problem (LCP). NCP can be found in various fields, e.g., operations research, engineering, finance and so on [7]. Throughout this paper, F is assumed to be continuously differentiable and monotone.

For a solution $\bar{x}$ of NCP(F), let $P(\bar{x}), N(\bar{x})$ and $C(\bar{x})$ be the index sets defined by

$$\begin{aligned} P(\bar{x}) &:= \{i \mid \bar{x}_i > 0, F_i(\bar{x}) = 0\}, \\ N(\bar{x}) &:= \{i \mid \bar{x}_i = 0, F_i(\bar{x}) > 0\}, \\ C(\bar{x}) &:= \{i \mid \bar{x}_i = 0, F_i(\bar{x}) = 0\}, \end{aligned}$$

respectively. Note that $P(\bar{x}) \cup N(\bar{x}) \cup C(\bar{x}) = \{1, \ldots, n\}$ and these index sets are mutually disjoint. If $C(\bar{x}) \neq \emptyset$, we call $\bar{x}$ a degenerate solution, otherwise we call $\bar{x}$ a nondegenerate solution. In this paper, we will focus on these index sets for a particular solution, and develop an algorithm that is designed to solve NCP(F) by estimating the correct index sets.

Various methods for solving NCPs, such as the generalized Newton method (GNM) [4], the smoothing method [2] and the regularization method [6], have been proposed and shown to have nice convergence properties. However, those methods generally require the local uniqueness of a solution for a superlinear rate of convergence. Recently, Yamashita and Fukushima [14] proposed a method, henceforth called the PPA, based on the proximal point algorithm, and showed that it has a superlinear rate of convergence without the local uniqueness of a solution. However, when a sequence $\{x^k\}$ generated by the PPA converges to a degenerate solution, subproblems may become computationally expensive. This difficulty comes from the fact that we do not know in advance whether or not $\{x^k\}$ converges to a degenerate solution. More recently, Yamashita, Dan and Fukushima [15] presented a technique that enables us to identify $P^* = P(x^*), N^* = N(x^*)$ and $C^* = C(x^*)$ when x^k enters a certain vicinity of x^*. Once we identify P^*, N^* and C^*, we can find a solution of NCP(F) by solving the system of nonlinear

equations

$$G_{x^*}(x) := \begin{pmatrix} F_{P^*}(x) \\ x_{N^*} \\ F_{C^*}(x) \\ x_{C^*} \end{pmatrix} = 0, \tag{1.1}$$

where $F_{P^*}(x)$ is the vector with components $F_i(x)$, $i \in P^*$, x_{N^*} is the vector with components x_i, $i \in N^*$, and so on. In fact, if a solution $\hat{x}$ of (1.1) is sufficiently near to x^*, then we have $\hat{x}_{P^*} > 0$ and $F_{N^*}(\hat{x}) > 0$, and hence $\hat{x}$ is also a solution of NCP(F). Moreover, since the mapping $G_{x^*} : \Re^n \to \Re^{n+|C^*|}$ is differentiable, we can use any Newton-type method to solve (1.1), and we can expect that such a method has a quadratic or superlinear rate of convergence no matter whether x^* is degenerate or nondegenerate.

In this paper, we construct a differentiable system of nonlinear equations (1.1) by using the technique presented in [15]. Moreover, we propose a hybrid algorithm that generates a sequence $\{x^k\}$ by the PPA primarily and also tries to find a solution of the system of nonlinear equations (1.1) by using the inexact Levenberg-Marquardt method (ILMM) proposed in [3], when an iterative point x^k is judged to lie sufficiently close to a solution. We show that the proposed method has a quadratic or superlinear convergence property if $\|G_{x^*}(x)\|$ provides a local error bound for (1.1), i.e., there exists constants $b_G > 0$ and $c_G > 0$ such that

$$c_G \mathrm{dist}\,(x, X_G^*) \leq \|G_{x^*}(x)\| \quad \forall x \in B\,(x^*, b_G)\,, \tag{1.2}$$

where X_G^* is the solution set of (1.1) and $B\,(x^*, b_G) := \{x \in \Re^n \mid \|x - x^*\| \leq b_G\}$. More specifically, we show that either a sequence $\{x^k\}$ generated by the proposed algorithm converges to the solution set of NCP quadratically or $\{\|G_{x^*}(x^k)\|\}$ converges to 0 superlinearly, without assuming the nondegeneracy of a solution.

This paper is organized as follows: In Section 2, we review some fundamental concepts and results such as error bounds for NCP, the PPA [14] and the ILMM [3]. In Section 3, we construct a differentiable system of nonlinear equations whose solution is a solution of NCP(F). Moreover, we propose the hybrid algorithm based on (1.1) and show its global and superlinear convergence. In Section 4, we make some concluding remarks and discuss future research topics.

2 Preliminaries

In this section, we review some concepts and results that are used in the subsequent discussion.

2.1 Error bound for NCP

First, we recall some mathematical concepts related to a mapping F.

Definition 2.1 *The mapping $F : \Re^n \to \Re^n$ is called*
(i) *monotone if*

$$(x-y)^T (F(x)-F(y)) \geq 0 \qquad \forall x, y \in \Re^n,$$

(ii) *strongly monotone with modulus $\mu > 0$ if*

$$(x-y)^T (F(x)-F(y)) \geq \mu \|x-y\|^2 \qquad \forall x, y \in \Re^n.$$

Next, we consider a system of nonlinear equations equivalent to NCP(F). In order to reformulate an NCP, we may use a function $\phi : \Re^2 \to \Re$ that has the following property:

$$\phi(a,b) = 0 \iff a \geq 0,\ b \geq 0,\ ab = 0. \tag{2.1}$$

Such a function ϕ is called an NCP-function. Using an NCP-function ϕ, define the mapping $H_F : \Re^n \to \Re^n$ by

$$H_F(x) = \begin{pmatrix} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{pmatrix}.$$

From (2.1), it is easily seen that NCP(F) is equivalent to the system of nonlinear equations

$$H_F(x) = 0.$$

The following two NCP-functions are well-known:

$$\begin{aligned} \phi_{NR}(a,b) &= \min\{a,b\}, \\ \phi_{FB}(a,b) &= \sqrt{a^2+b^2} - a - b. \end{aligned}$$

The functions ϕ_{NR} and ϕ_{FB} are called the natural residual function and the Fischer-Burmeister function [8], respectively. The function ϕ_{FB} is equivalent to ϕ_{NR} in the sense that they satisfy the following inequalities [13]:

$$(2-\sqrt{2})\,|\phi_{NR}(a,b)| \ \leq\ |\phi_{FB}(a,b)| \ \leq\ (2+\sqrt{2})\,|\phi_{NR}(a,b)| \quad \forall (a,b)^T \in \Re^2.$$

Throughout this paper, we assume that the mapping H_F is given by any NCP-function ϕ satisfying

$$\nu_1 |\phi_{NR}(a,b)| \ \leq \ |\phi(a,b)| \ \leq \ \nu_2 |\phi_{NR}(a,b)| \quad \forall (a,b)^T \in \Re^2, \tag{2.2}$$

where ν_1 and ν_2 are some positive constants.

The next theorem states some error bound properties of $\|H_F(x)\|$ for NCP, which were essentially proved by Pang [11] and Robinson [12].

Theorem 2.1 (i) *Suppose that F is strongly monotone with modulus $\mu > 0$ and Lipschitz continuous with constant $L > 0$. Then $\|H_F(x)\|$ provides a global error bound for NCP(F), that is,*

$$\|x - \hat{x}\| \leq \frac{K_1(L+1)}{\mu} \|H_F(x)\| \qquad \forall x \in \Re^n,$$

where $\hat{x}$ is the unique solution of NCP(F) and $K_1 > 0$ is a constant independent of F.
(ii) *Suppose that F is affine and NCP(F) has a solution. Then $\|H_F(x)\|$ provides a local error bound for NCP(F), that is, there exist positive constants K_2 and K_3 such that*

$$\|H_F(x)\| \leq K_2 \implies \mathrm{dist}(x, X^*) \leq K_3 \|H_F(x)\|,$$

where X^* denotes the solution set of NCP(F).

2.2 The PPA and identification of the index sets

The PPA presented in [14] is stated as follows:

Algorithm PPA

Step 0: Choose parameters $\beta \in (0,1), c_0 \in (0,1)$ and an initial point $x^0 \in \Re^n$. Set $k := 0$.

Step 1: If x^k satisfies a stopping criterion, then stop.

Step 2: Let $F^k : \Re^n \to \Re^n$ be given by

$$F^k(x) = F(x) + c_k(x - x^k).$$

Find an approximate solution x^{k+1} of NCP(F^k) such that

$$\|H_{F^k}(x^{k+1})\| \leq (\beta)^k \min\left\{1, \|x^{k+1} - x^k\|\right\}. \tag{2.3}$$

Step 3: Choose $c_{k+1} > 0$ and set $k := k + 1$. Go to Step 2. □

This algorithm enjoys nice convergence properties. Specifically, a sequence generated by the PPA converges to a solution of NCP(F) globally, and the distance between the generated sequence and the solution set of NCP(F) converges to 0 superlinearly under mild assumptions. In what follows, $\{x^k\}$ denotes a sequence generated by Algorithm PPA, and x^* denotes the limit point of $\{x^k\}$. The following convergence theorem has been established in [14].

Theorem 2.2 *Suppose that F is monotone and Lipschitzian, and that NCP(F) has a solution. Then, $\{x^k\}$ generated by the PPA converges to a solution x^* of NCP(F) whenever $\{c_k\}$ is bounded. Moreover, if $\|H_F(x)\|$ provides a local error bound in a neighborhood of x^* and $c_k \to 0$, then $\{\mathrm{dist}(x^k, X^*)\}$ converges to 0 superlinearly, where X^* is the solution set of NCP(F).*

The most expensive task in Algorithm PPA is to solve NCP(F^k) in Step 2. In [14], it is proposed that NCP(F^k) be solved using the Generalized Newton Method (GNM) [4]. Note that F^k is strongly monotone when F is monotone. Then, the GNM can find an approximate solution of subproblem NCP(F^k) rapidly when x^* is nondegenerate, as stated in the following theorem [14].

Theorem 2.3 *Suppose that x^* is a nondegenerate solution of NCP(F) and $\|H_F(x)\|$ provides a local error bound in a neighborhood of x^*. Then, a single iteration of the GNM for NCP(F^k) yields a point x^{k+1} that satisfies (2.3), provided k is sufficiently large.*

When x^* is a degenerate solution, Theorem 2.3 no longer guarantees that the GNM can find a solution of subproblem NCP(F^k) in a few iterations, and hence it may take much time to solve subproblems. To overcome this difficulty, we will use the technique proposed in [15] to identify the index sets $P(x^*), N(x^*)$ and $C(x^*)$. The following theorem has been established in [15].

Theorem 2.4 *Suppose that $\|H_F(x)\|$ provides a local error bound for NCP(F) and we let $c_k = (\alpha)^k$ for all k in Algorithm PPA, where α is a constant such that $\beta < \alpha < 1$. Let a function sequence $\{\rho^k\}$ be defined by*

$$\rho^k(x) = \sqrt{\frac{\|H_F(x)\|}{(\alpha)^k}},$$

where the index sets P^k, N^k and C^k are defined by

$$\begin{aligned} P^k &:= \left\{ i \mid x_i^k > \rho^k(x^k), F_i(x^k) \leq \rho^k(x^k) \right\}, \\ N^k &:= \left\{ i \mid x_i^k \leq \rho^k(x^k), F_i(x^k) > \rho^k(x^k) \right\}, \\ C^k &:= \left\{ i \mid x_i^k \leq \rho^k(x^k), F_i(x^k) \leq \rho^k(x^k) \right\}, \end{aligned} \tag{2.4}$$

respectively. Then, we have

$$P^k = P(x^*), \quad N^k = N(x^*), \quad C^k = C(x^*).$$

for sufficiently large k.

2.3 Inexact Levenberg-Marquardt method

The Levenberg-Marquardt method (LMM) [1, 9] is a method for solving the system of nonlinear equations

$$G(y) = 0, \tag{2.5}$$

where $G : \Re^n \to \Re^m$ is a continuously differentiable mapping. The LMM generates a sequence $\{y^l\}$ by $y^{l+1} := y^l + \hat{d}^l$, where $\hat{d}^l$ is a solution of the system of linear equations

$$\left(\nabla G(y^l) \nabla G(y^l)^T + \mu_l I \right) d = -\nabla G(y^l) G(y^l). \tag{2.6}$$

Here $\nabla G(y) \in \Re^{n \times m}$ is the transposed Jacobian of G at y, μ_l is a positive parameter, and I is the identity matrix. Since $\nabla G(y^l) \nabla G(y^l)^T + \mu_l I$ is positive definite, (2.6) has a unique solution. However, it is expensive to find an exact solution of (2.6) when n is large. In that case, the inexact Levenberg-Marquardt method (ILMM) [3, 5] is useful. The ILMM uses an approximate solution d^l of (2.6) as a search direction, and generates a sequence $\{y^l\}$ by $y^{l+1} := y^l + d^l$.

Dan, Yamashita and Fukushima [3] show that the ILMM has a quadratic rate of convergence under a local error bound condition, which is milder than the nonsingularity condition at a solution. This result is stated in the following theorem.

Theorem 2.5 *Let $\{y^l\}$ be a sequence generated by the ILMM and y^* be a solution of (2.5). Suppose that the following two conditions hold:*
(i) *There exist constants $b_1 \in (0, 1)$ and $\kappa_1 \in (0, \infty)$ such that*

$$\|G(y) - G(x) - \nabla G(x)(y - x)\| \leq \kappa_1 \|y - x\|^2 \quad \forall x, y \in B(y^*, b_1),$$

where $B(y^*, b_1) := \{y \in \Re^n \mid \|y - y^*\| \leq b_1\}$.
(ii) $\|G(y)\|$ *provides an error bound for (2.5) on* $B(y^*, b_1)$*, i.e., there exists a constant* $\kappa_2 > 0$ *such that*

$$\kappa_2 \operatorname{dist}(y, Y^*) \leq \|G(y)\| \quad \forall y \in B(y^*, b_1),$$

where Y^* *is the solution set of (2.5).*
Let the parameters μ_l *be chosen as* $\mu_l = \|G(y^l)\|^2$ *and the residual vectors* r^l *be defined by*

$$r^l := \left(\nabla G(y^l)\nabla G(y^l)^T + \mu_l I\right) d^l + \nabla G(y^l) G(y^l). \tag{2.7}$$

If an initial point y^0 *is sufficiently close to* y^* *and*

$$\frac{\|r^l\|}{\mu_l} = O\left(\operatorname{dist}(y^l, Y^*)^2\right)$$

holds for all l*, then there exists a constant* $c > 0$ *such that* $\{y^l\} \subseteq B(y^*, c\|y^0 - y^*\|)$*. Moreover,* $\{\operatorname{dist}(y^l, Y^*)\}$ *converges to* 0 *quadratically.*

The assumption (i) of Theorem 2.5 holds when ∇G is locally Lipschitzian [10, Theorem 3.2.12].

3 Proposed algorithm and its convergence properties

In this section, we describe the proposed algorithm and show that it has at least a superlinear rate of convergence for NCP, without assuming the local uniqueness and the nondegeneracy of a solution.

3.1 Differentiable system of nonlinear equations

First, we make the following assumption to guarantee that the PPA has a superlinear convergence and the technique to identify the index sets works well.

Assumption 3.1 (i) ∇F *is locally Lipschitzian.*
(ii) $\|H_F(x)\|$ *provides a local error bound for NCP(F), i.e., there exist constants* $b_H > 0$ *and* $c_H > 0$ *such that*

$$\|H_F(x)\| \leq c_H \implies b_H \operatorname{dist}(x, X^*) \leq \|H_F(x)\|,$$

where X^* *is the solution set of NCP(F).*

Sufficient conditions under which $\|H_F(x)\|$ provides a local error bound for NCP(F) are given in Theorem 2.1.

We consider the mapping $G_{x^*} : \Re^n \to \Re^{n+|C^*|}$ defined by (1.1), where x^* is a solution of NCP(F). Since $x_{P^*} > 0$ and $F_{N^*}(x) > 0$ in a sufficiently small neighborhood of x^*, any solution $\hat{x}$ of (1.1) also solves NCP(F) if $\hat{x}$ is sufficiently close to x^*. We note that G_{x^*} is differentiable and ∇G_{x^*} is locally Lipschitzian, and hence the assumption (i) of Theorem 2.5 holds for G_{x^*}. So, the ILMM applied to the system of equations (1.1) has a quadratic rate of convergence if $\|G_{x^*}(x)\|$ provides a local error bound. Accordingly, we make the following assumption.

Assumption 3.2 *$\|G_{x^*}(x)\|$ provides a local error bound for (1.1) in a neighborhood of x^*, i.e., there exist positive constants b_G and c_G such that (1.2) holds.*

Assumption 3.2 holds when F_{P^*} and F_{C^*} are affine. Though this assumption does not necessarily hold when F is nonlinear, it does not seem very restrictive. Assumptions 3.1 and 3.2 are closely related. In fact, as shown in the next lemma, Assumption 3.2 is implied by Assumption 3.1 under some normal circumstances.

Lemma 3.1 *Suppose that Assumption 3.1 holds. If there exists a constant $r_1 > 0$ such that $X^* \cap B(x^*, r_1) = X^*_G \cap B(x^*, r_1)$, where X^*_G is the solution set of (1.1), then $\|G_{x^*}(x)\|$ provides a local error bound for (1.1) in a neighborhood of x^*, i.e., there exist constants b_G and c_G such that (1.2) holds. In particular, if x^* is a locally unique solution of NCP(F), then $\|G_{x^*}(x)\|$ provides a local error bound for (1.1) in a neighborhood of x^*.*

Proof: Choosing $r_2 > 0$ sufficiently small yields that, for any $x \in B(x^*, r_2)$,

$$\begin{aligned} x_i &\geq F_i(x) \quad \forall i \in P^*, \\ x_i &\leq F_i(x) \quad \forall i \in N^*. \end{aligned}$$

It then follows that, for any $x \in B(x^*, r_2)$,

$$\begin{aligned} \|H_F(x)\| &= \sqrt{\sum_{i \in P^* \cup N^* \cup C^*} \phi^2(x_i, F_i(x))} \\ &\leq \nu_2 \sqrt{\sum_{i \in P^* \cup N^* \cup C^*} |\min\{x_i, F_i(x)\}|^2} \\ &\leq \nu_2 \sqrt{\sum_{i \in P^*} F_i^2(x) + \sum_{i \in N^*} x_i^2 + \sum_{i \in C^*} (x_i^2 + F_i^2(x))} \\ &= \nu_2 \|G_{x^*}(x)\|, \end{aligned} \tag{3.1}$$

where ν_2 is the positive constant given in (2.2). Moreover, by choosing $r_3 > 0$ sufficiently small, we have

$$\|G_{x^*}(x)\| \leq c_H/\nu_2 \quad \forall x \in B(x^*, r_3). \tag{3.2}$$

Let $r := \min\{r_1, r_2, r_3\}$. It then follows from (3.1) and (3.2) that

$$\|H_F(x)\| \leq c_H \quad \forall x \in B(x^*, r).$$

Therefore, from Assumption 3.1 (ii) and (3.1), we get

$$b_H \text{dist}(x, X^*) \leq \|H_F(x)\| \leq \nu_2 \|G_{x^*}(x)\| \quad \forall x \in B(x^*, r). \tag{3.3}$$

Let $x \in B\left(x^*, \frac{r}{2}\right)$, and let $\bar{x}_G$ and $\bar{x}$ be one of the nearest points from x in X_G^* and X^*, respectively. Since $x^* \in X_G^*$ and $x^* \in X^*$, we have, for any $x \in B\left(x^*, \frac{r}{2}\right)$,

$$\begin{aligned} \|\bar{x}_G - x\| &\leq \|x^* - x\| \leq \frac{r}{2}, \\ \|\bar{x} - x\| &\leq \|x^* - x\| \leq \frac{r}{2}, \end{aligned}$$

and hence

$$\begin{aligned} \|\bar{x}_G - x^*\| &\leq \|\bar{x}_G - x\| + \|x^* - x\| \leq r, \\ \|\bar{x} - x^*\| &\leq \|\bar{x} - x\| + \|x^* - x\| \leq r. \end{aligned}$$

Therefore we have $\bar{x}_G \in X_G^* \cap B(x^*, r)$ and $\bar{x} \in X^* \cap B(x^*, r)$. It then follows from the assumption $X^* \cap B(x^*, r) = X_G^* \cap B(x^*, r)$ that $\text{dist}(x, X^*) = \text{dist}(x, X_G^*)$ for any $x \in B\left(x^*, \frac{r}{2}\right)$. Consequently, by (3.3), we have

$$\nu_2^{-1} b_H \text{ dist}(x, X_G^*) \leq \|G_{x^*}(x)\| \quad \forall x \in B\left(x^*, \frac{r}{2}\right),$$

i.e., $\|G_{x^*}(x)\|$ provides a local error bound for (1.1) in a neighborhood of x^*. □

3.2 The algorithm

Let $\{x^k\}$ be a sequence generated by the PPA, x^* be the limit of $\{x^k\}$, and P^k, N^k, C^k be defined by (2.4). Then, as stated in Subsection 2.2, $P^k = P(x^*), N^k = N(x^*)$ and $C^k = C(x^*)$ hold when k is sufficiently large, and hence, the system of nonlinear equations

$$G^k(x) := \begin{pmatrix} F_{P^k}(x) \\ x_{N^k} \\ F_{C^k}(x) \\ x_{C^k} \end{pmatrix} = 0 \tag{3.4}$$

coincides with the system of nonlinear equations (1.1). In this case, a solution $\hat{x}$ of (3.4) is a solution of NCP(F) if $\hat{x}$ is sufficiently close to x^*. Then we naturally come up with the following method: We generate a sequence $\{x^k\}$ by the PPA primarily, and if an iterative point is judged to be close to the solution x^*, then we solve (3.4) by the ILMM. Based on this idea, we propose the following hybrid algorithm.

Algorithm Hybrid

Step 0: Choose parameters $M_1 > 0, M_2 > 0, 0 < \beta < \alpha < 1, 0 < \gamma < 1$, and an initial point x^0. Let the initial index sets be $P^0 = N^0 = C^0 = \emptyset$. Set $k := 1$.

Step 1: Let $c_k := (\alpha)^k$ and obtain x^k by applying a single iteration of Algorithm PPA. Determine the index sets P^k, N^k, C^k by (2.4).

Step 2: If $\|H_F(x^k)\| \leq M_1, P^k = P^{k-1}, N^k = N^{k-1}, C^k = C^{k-1}$ and $P^k \cup N^k \cup C^k = \{1, \ldots, n\}$, then go to Step 3. Otherwise, set $k := k + 1$ and go to Step 1.

Step 3: [ILMM for (3.4)]

Step 3.0: Set $y^{k,0} := x^k$, $\mu_{k,0} := \|G^k(y^{k,0})\|$, and $l := 0$.

Step 3.1: If $y^{k,l}$ solves NCP(F), then terminate. Otherwise, find an approximate solution $d^{k,l}$ of the system of linear equations

$$\left(\nabla G^k(y^{k,l})\nabla G^k(y^{k,l})^T + \mu_{k,l} I\right) d = -\nabla G^k(y^{k,l})^T G^k(y^{k,l}). \tag{3.5}$$

Step 3.2: If the conditions

$$\begin{aligned}
(y^{k,l} + d^{k,l})_{P^k} &> 0, && (3.6)\\
F_{N^k}(y^{k,l} + d^{k,l}) &> 0, && (3.7)\\
\|G^k(y^{k,l} + d^{k,l})\| &\leq (\gamma)^l \|G^k(y^{k,l})\| && (3.8)\\
\|x^k - (y^{k,l} + d^{k,l})\| &\leq M_2 && (3.9)
\end{aligned}$$

are not satisfied, then set $k := k + 1$ and go to Step 1.

Step 3.3: Set $y^{k,l+1} := y^{k,l} + d^{k,l}, \mu_{k,l+1} := \|G^k(y^{k,l+1})\|$ and $l := l + 1$. Go to Step 3.1. □

In Algorithm Hybrid, we try to identify the three index sets $P(x^*), N(x^*)$ and $C(x^*)$ in Step 2. The conditions in Step 2 will be satisfied when k is sufficiently large. Note, however, that those conditions do not guarantee that the index sets P^*, N^* and C^* are identified correctly. Therefore, we check conditions (3.6) – (3.9) in Step 3.2

to ensure that the sequence $\{y^{k,l}\}_{l=0,1,2,\ldots}$ is converging to a solution of NCP(F). The conditions (3.6) and (3.7) check whether P^k and N^k contain any wrong indices. The condition (3.8) ensures that $\{\|G^k(y^{k,l})\|\}_{l=0,1,2,\ldots}$ is converging to 0 superlinearly, and the condition (3.9) ensures that $\{y^{k,l}\}_{l=0,1,2,\ldots}$ is not diverging.

3.3 Convergence theorem

Let the residual vector $r^{k,l}$ associated with an approximate solution $d^{k,l}$ of the system of linear equations (3.5) be defined by

$$r^{k,l} := \left(\nabla G^k(y^{k,l})\nabla G^k(y^{k,l})^T + \mu_{k,l} I\right) d^{k,l} + \nabla G^k(y^{k,l}) G^k(y^{k,l}).$$

Now we show the following convergence theorem for Algorithm Hybrid.

Theorem 3.1 *Suppose that Assumptions 3.1 and 3.2 hold. Suppose also that, for any k, the sequence $\{y^{k,l}\}_{l=0,1,2,\ldots}$ generated in Step 3 of Algorithm Hybrid satisfies*

$$\frac{\|r^{k,l}\|}{\mu_{k,l}} = O\left(\mathrm{dist}(y^{k,l}, X^*_{G^k})^2\right), \tag{3.10}$$

where $X^*_{G^k}$ is the solution set of (3.4). Then, there exists a positive integer k for which either of the following statements is true:

(a): *Any accumulation point of the sequence $\{y^{k,l}\}_{l=0,1,2,\ldots}$ is a solution of NCP(F), and $\{\|G^k(y^{k,l})\|\}_{l=0,1,2,\ldots}$ converges to 0 superlinearly.*

(b): *$\{y^{k,l}\}_{l=0,1,2,\ldots}$ converges to a solution of NCP(F), and $\{\mathrm{dist}(y^{k,l}, X^*_G)\}_{l=0,1,2,\ldots}$ converges to 0 quadratically.*

Proof: First, we consider the case where the inner loop in Step 3 cycles infinitely for some k, that is, an infinite sequence $\{y^{k,l}\}_{l=0,1,2,\ldots}$ satisfies (3.6) – (3.9), no matter whether the index sets $P(x^*), N(x^*)$ and $C(x^*)$ have yet to be identified correctly. In this case, it follows from (3.9) that $\{y^{k,l}\}_{l=0,1,2,\ldots}$ has accumulation points, and from (3.8), $\{\|G^k(y^{k,l})\|\}_{l=0,1,2,\ldots}$ converges to 0 superlinearly. Moreover, from (3.6) and (3.7), any accumulation point of $\{y^{k,l}\}_{l=0,1,2,\ldots}$ is a solution of NCP(F). Therefore the statement (a) holds.

Next, we show that, even if (a) does not hold, the statement (b) holds eventually. In fact, for sufficiently large k, $\|H_F(x^k)\|$ becomes sufficiently small by Theorem 2.2, and $P^k = P(x^*), N^k = N(x^*), C^k = C(x^*)$ hold by Theorem 2.4. Hence, the system (3.4) coincides with the system (1.1) for sufficiently large k. Note that, by Assumptions 3.1

and 3.2, G_{x^*} and x^* satisfy the assumptions (i) and (ii) in Theorem 2.5, where G and y^* are regarded as G_{x^*} and x^*, respectively. In what follows, we consider a sequence $\{z^{k,l}\}_{l=0,1,2,\ldots}$ generated by the ILMM for the equation $G_{x^*}(x) = 0$ with an initial point $z^{k,0} := x^k$ for sufficiently large k, and show that $z^{k,l}$ satisfies the conditions (3.6)–(3.9) for all l. By Lemma 2.3 in [3] and (3.10), if k is sufficiently large and $z^{k,0} = x^k$ is sufficiently close to x^*, then we have $\|z^{k,l+1} - \bar{z}^{k,l+1}\| \leq c \, \|z^{k,l} - \bar{z}^{k,l}\|^2$ for all l, where c is a positive constant and $\bar{z}^{k,l}$ is one of the nearest points from $z^{k,l}$ in X^*_G. Then, from Assumption 3.2, we have

$$\begin{aligned}
\frac{\|G_{x^*}(z^{k,l+1})\|}{\|G_{x^*}(z^{k,l})\|^2} &= \frac{\|G_{x^*}(z^{k,l+1}) - G_{x^*}(\bar{z}^{k,l+1})\|}{\|G_{x^*}(z^{k,l})\|^2} \\
&\leq \frac{L_{x_*}\|z^{k,l+1} - \bar{z}^{k,l+1}\|}{c_G\|z^{k,l} - \bar{z}^{k,l}\|^2} \\
&\leq \frac{cL_{x^*}}{c_G} \qquad l = 0, 1, 2, \ldots,
\end{aligned}$$

where L_{x^*} is a Lipschitz constant of the mapping G_{x^*}. When k is sufficiently large, we have $\|G_{x^*}(z^{k,0})\| \leq (\gamma c_G)/(cL_{x^*})$, and it follows that for each $l = 0, 1, 2, \ldots$

$$\begin{aligned}
\|G_{x^*}(z^{k,l+1})\| &\leq \frac{cL_{x^*}}{c_G}\|G_{x^*}(z^{k,l})\|^2 \\
&\leq \left(\frac{cL_{x^*}}{c_G}\|G_{x^*}(z^{k,l-1})\|\right)^2 \|G_{x^*}(z^{k,l})\| \\
&\vdots \\
&\leq \left(\frac{cL_{x^*}}{c_G}\|G_{x^*}(z^{k,0})\|\right)^{2^l} \|G_{x^*}(z^{k,l})\| \\
&\leq (\gamma)^{2^l}\|G_{x^*}(z^{k,l})\| \\
&\leq (\gamma)^l\|G_{x^*}(z^{k,l})\|.
\end{aligned}$$

Therefore, (3.8) holds when k is sufficiently large. Let $\hat{r}^k$ be the distance between $y^{k,0}$ and x^*. When $\hat{r}^k$ is sufficiently small, Theorem 2.5 says that a generated sequence $\{z^{k,l}\}_{l=0,1,2,\ldots}$ satisfies $z^{k,l} \in B(x^*, c_3\hat{r}^k)$ for all l, where $c_3 > 0$ is a constant. Moreover, choosing sufficiently large k if necessary, $\hat{r}^k$ can be made arbitrarily small. Then, it follows from $x^*_{P^*} > 0$ and $F_{N^*}(x^*) > 0$ that (3.6), (3.7) and (3.9) hold for sufficiently large k. Consequently, (3.6)–(3.9) are satisfied for sufficiently large k. Then, from Theorem 2.5, $\{\mathrm{dist}(z^{k,l}, X^*_G)\}_{l=0,1,2,\ldots}$ converges to 0 quadratically, i.e., the case (b) holds. This completes the proof. □

Some remarks about Algorithm Hybrid are in order.

- The system of nonlinear equations (3.4) has only $|P^k|$ variables actually, since $x_{N^k} = 0$ and $x_{C^k} = 0$. Hence, (3.4) becomes smaller than the original problem.

- The condition $\|H_F(x^k)\| \leq M_1$ in Step 2 is not needed to establish Theorem 3.1. However, using this condition, we can skip useless calculation which may occur when

$$P(x^*) \neq P^k = P^{k-1},\ N(x^*) \neq N^k = N^{k-1},\ C(x^*) \neq C^k = C^{k-1},$$
$$P^k \cup N^k \cup C^k = \{1, \ldots, n\}.$$

- We may replace the condition in Step 2 by

$$P^k = \cdots = P^{k-j},\ N^k = \cdots = N^{k-j},\ C^k = \cdots = C^{k-j},$$
$$P^k \cup N^k \cup C^k = \{1, \ldots, n\},$$

  where j is any positive integer. This modification may also enable us to omit unnecessary calculation when the index sets are not identified correctly.

- By Theorem 2.3, the PPA has a genuine superlinear rate of convergence if x^* is nondegenerate. Consequently, when we judge x^* to be nondegenerate in Step 2, we may immediately return to Step 1 instead of proceeding to Step 3.

4 Concluding Remarks

In this paper, we have constructed the system of nonlinear equations which is useful in dealing with degenerate NCPs, and proposed an algorithm based on it. We have shown that the proposed algorithm has a quadratic or superlinear rate of convergence even if the NCP has a degenerate solution.

Finally, we mention some future research topics.

(i) In order to guarantee that Algorithm Hybrid has a quadratic or superlinear rate of convergence, we need Assumption 3.2. In Lemma 3.1, we have given a sufficient condition for Assumption 3.2 to hold, but this condition is imposed on the solution set itself. It is an interesting and important subject to find a milder and/or simpler sufficient condition for Assumption 3.2.

(ii) We have used the technique proposed in [15] to identify P^*, N^* and C^*. However, it is difficult in practice to confirm that $P^k = P^*, N^k = N^*$ and $C^k = C^*$ are attained. If we have a criterion to make sure that the correct identification of the index sets has been accomplished, we may design an algorithm that avoids useless calculations in Step 3 of Algorithm Hybrid.

References

[1] D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1995.

[2] X. Chen, Smoothing methods for complementarity problems and their applications: A survey, *Journal of the Operations Research Society of Japan* 43 (2000) 32–47.

[3] H. Dan, N. Yamashita and M. Fukushima, Convergence properties of the inexact Levenberg-Marquardt method under local error bound conditions, Technical Report 2001-001, Department of Applied Mathematics and Physics, Kyoto University (January 2001).

[4] T. De Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, *Mathematical Programming* 75 (1996) 407–439.

[5] F. Facchinei and C. Kanzow, A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, *Mathematical Programming* 76 (1997) 493–512.

[6] F. Facchinei and C. Kanzow, Beyond monotonicity in regularization methods for complementarity problems, *SIAM Journal on Control and Optimization* 37 (1999) 1150–1161.

[7] M.C. Ferris and J.-S. Pang, Engineering and economic applications of complementarity problem, *SIAM Review* 39 (1997) 669–713.

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

[9] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, NY, 1987.

[10] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY, 1970.

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

[12] S.M. Robinson, Some continuity properties of polyhedral multifunctions, *Mathematical Programming Study* 14 (1981) 206–214.

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

[14] N. Yamashita and M. Fukushima, The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem, *SIAM Journal on Optimization* 11 (2001) 364-379.

[15] N. Yamashita, H. Dan and M. Fukushima, On the identification of degenerate indices in the nonlinear complementarity problem with the proximal point algorithm, Technical Report 2001-003, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University (February 2001).