SIAM J. OPTIM.
Vol. 8, No. 1, pp. 140–157, February 1998

007

A TRUST REGION METHOD FOR SOLVING GENERALIZED COMPLEMENTARITY PROBLEMS*

HOUYUAN JIANG†, MASAO FUKUSHIMA‡, LIQUN QI§, AND DEFENG SUN§

Abstract. Based on a semismooth equation reformulation using Fischer's function, a trust region algorithm is proposed for solving the generalized complementarity problem (GCP). The algorithm uses a generalized Jacobian of the function involved in the semismooth equation and adopts the squared natural residual of the semismooth equation as a merit function. The proposed algorithm is applicable to the nonlinear complementarity problem because the latter problem is a special case of the GCP. Global convergence and, under a nonsingularity assumption, local Q-superlinear (or quadratic) convergence of the algorithm are established. Moreover, calculation of a generalized Jacobian is discussed and numerical results are presented.

Key words. generalized complementarity problem, nonlinear complementarity problem, semismooth equation, trust region method, global and superlinear convergence

AMS subject classifications. 90C33, 65K10

PII. S1052623495296541

1. Introduction.

The GCP is to find a vector $x \in \Re^n$ such that

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

where $F, G : \Re^n \to \Re^n$ are assumed to be continuously differentiable.

When $G(x) \equiv x - E(x)$, where $E : \Re^n \to \Re^n$, the GCP reduces to the so-called implicit complementarity problem [8, 28, 27]. In particular, if $G(x) \equiv x$, the GCP reduces to the nonlinear complementarity problem (NCP), which is a general framework for optimality conditions of mathematical programs as well as variational inequalities. Moreover, if F is an affine function, i.e., $F(x) = Mx + q$ with a matrix $M \in \Re^{n\times n}$ and a vector $q \in \Re^n$, then the NCP reduces to the linear complementarity problem (LCP), which in turn contains linear and quadratic programming problems as special cases.

There are many Newton-based methods for solving the NCP. We do not intend to give a short survey of numerical methods for solving the NCP. The interested reader is referred to two survey papers [18, 29] and to [1, 2, 15, 22, 30] for recent progress. In order to enlarge the domain of convergence of Newton-based methods, line search is usually used on certain merit functions derived from different reformulations of the NCP. A trust region strategy is used in [26] for solving a bound-constrained nonlinear least squares reformulation of the NCP. Global convergence of this trust region method was established, and, under the strict complementarity condition and other conditions, superlinear convergence was also given.

---

*Received by the editors December 21, 1995; accepted for publication (in revised form) December 17, 1996. The research of the first, third, and fourth authors was supported by the Australian Research Council; the research of the second author was supported in part by the Scientific Research Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

http://www.siam.org/journals/siopt/8-1/29654.html

†Department of Mathematics, The University of Melbourne, Parkville, Victoria 3052, Australia (jiang@mundoe.maths.mu.oz.au).

‡Department of Applied Mathematics and Physics, Graduate School of Engineering, Kyoto University, Kyoto 606-01, Japan (fuku@kuamp.kyoto-u.ac.jp).

§School of Mathematics, The University of New South Wales, Sydney 2052, Australia (qi@alpha.maths.unsw.edu.au, sun@alpha.maths.unsw.edu.au).

In this paper, we shall propose a trust region method for solving an unconstrained least squares reformulation of the GCP. It is surely applicable to the NCP since the GCP is a generalization of the NCP. The main difference between the proposed trust region method and the classical trust region methods lies in the updating rule for the trust region radius at the beginning of each iteration. More precisely, at the beginning of each iteration, the trust region radius is always set greater than a fixed (small) positive constant rather than solely updated from the final trust region radius of the last iteration as in the classical trust region methods. This type of updating rule for the trust region radius has been used in recent literature [14, 17, 23] for optimization problems and systems of nonsmooth equations. We will show that the proposed special strategy of updating trust region radii enables us to recover local superlinear convergence under some conditions in spite of the fact that the functions involved in the system of equations are only semismooth. (For globally and superlinearly convergent trust region methods for systems of smooth equations, see [25].)

The remainder of the paper is organized as follows. In the next section, the GCP will be converted into a system of semismooth equations and an unconstrained differentiable minimization problem. These two reformulations are equivalent to the GCP in a certain sense. In section 3, a trust region method is proposed for solving the GCP based on these two reformulations. In section 4, we discuss regularity conditions which ensure that a stationary point of the reformulated unconstrained minimization problem is a solution of the GCP. Section 5 is devoted to proving global convergence of the algorithm. In section 6, local superlinear convergence of the algorithm will be established under some nonsingularity condition. In section 7, we present a method for calculating a generalized Jacobian, which is needed in the algorithm. Numerical results are presented in section 8. We conclude our paper by giving some remarks in the final section.

A few words about our notation. The inner product of vectors $x, y \in \Re^n$ is denoted by $x^T y$. Let $\|\cdot\|$, $\|\cdot\|_\infty$, and $\|\cdot\|_2$ denote any fixed norm, the l_∞ norm, and the Euclidean norm, respectively, in a space of appropriate dimension. The Jacobian of a vector function F at a point x is denoted by $\nabla F(x)^T$. If $M = (M_{ij})$ is an $n \times n$ matrix and α, β are two subsets of $\{1, \ldots, n\}$, then $M_{\alpha\beta}$ denotes the submatrix of M consisting of elements M_{ij} with $i \in \alpha$ and $j \in \beta$.

2. Reformulation and preliminaries.

Throughout this paper, we assume that F and G are continuously differentiable on $\Re^n$.

Let $\phi : \Re^2 \to \Re$ be defined by

$$\phi(a, b) := \sqrt{a^2 + b^2} - a - b. \tag{2.1}$$

A basic property of this function is that

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

From this property, the GCP (1.1) can readily be recast as the system of nonsmooth equations

$$H(x) = \begin{pmatrix} H_1(x) \\ \vdots \\ H_n(x) \end{pmatrix} := \begin{pmatrix} \phi(F_1(x), G_1(x)) \\ \vdots \\ \phi(F_n(x), G_n(x)) \end{pmatrix} = 0 \tag{2.2}$$

in the sense that x solves (1.1) if and only if x solves (2.2). The function ϕ was first introduced by Fischer [11] (but attributed to Burmeister) and later used to study various methods for solving the NCP and related problems; see survey paper [12] for details. Recently the function ϕ has been used to reformulate the GCP [20]. (See also [35] for other reformulations of the GCP.)

Note that H is locally Lipschitz on $\Re^n$ and Fréchet differentiable on the set Ω, where

$$\Omega := \{x \in \Re^n \mid F_i(x)^2 + G_i(x)^2 > 0, i = 1, \ldots, n\}.$$

However, H is not necessarily differentiable at $x \notin \Omega$. Let $\partial H(x)$ denote the Clarke's generalized Jacobian of H at $x \in \Re^n$ [3], which can be defined as the convex hull of the set $\partial_B H(x)$ [32], where

$$\partial_B H(x) = \left\{ V \in \Re^{n\times n} \mid V = \lim_{x^k \to x} \nabla H(x^k)^T, H \text{ is differentiable at } x^k \text{ for all } k \right\}.$$

Similar to the discussions of [9, 19], we have for any $x \in \Re^n$

$$\partial H(x) \subseteq \mathcal{D}^F(x)\nabla F(x)^T + \mathcal{D}^G(x)\nabla G(x)^T, \tag{2.3}$$

where $\mathcal{D}^F(x)$ and $\mathcal{D}^G(x)$ are sets of $n \times n$ diagonal matrices such that, for each pair $(D^F, D^G) = (\text{diag }\{D_1^F, \ldots, D_n^F\}, \text{diag }\{D_1^G, \ldots, D_n^G\}) \in \mathcal{D}^F(x) \times \mathcal{D}^G(x)$, the following conditions are satisfied:

$$(D_i^F + 1)^2 + (D_i^G + 1)^2 \leq 1, \quad i = 1, \ldots, n. \tag{2.4}$$

In particular, if $F_i(x)^2 + G_i(x)^2 > 0$ for all $i = 1, \ldots, n$, then we have $\mathcal{D}^F(x) = \{D^F\}$ and $\mathcal{D}^G(x) = \{D^G\}$ with

$$D_i^F = \frac{F_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1, \quad D_i^G = \frac{G_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1, \quad i = 1, \ldots, n, \tag{2.5}$$

and $\nabla H(x)$ exists and is given by

$$\nabla H(x)^T = D^F \nabla F(x)^T + D^G \nabla G(x)^T.$$

Define a merit function $\Phi : \Re^n \to \Re$ as

$$\Phi(x) := \frac{1}{2}\sum_{i=1}^{n} \phi(F_i(x), G_i(x))^2 = \frac{1}{2}\|H(x)\|_2^2.$$

When the GCP has a solution, solving (1.1) is equivalent to finding a global minimum point of the unconstrained minimization problem

$$\min_{x \in \Re^n} \Phi(x). \tag{2.6}$$

A favorable property of Φ is that it is continuously differentiable on the whole space $\Re^n$, although H itself is not continuously differentiable in general [20]. It is easy to verify that for any $x \in \Re^n$ and any $V \in \partial H(x)$

$$\nabla \Phi(x) = V^T H(x) = \nabla F(x) D^F H(x) + \nabla G(x) D^G H(x), \tag{2.7}$$

where D^F and D^G are diagonal matrices satisfying (2.4) such that $V = D^F \nabla F(x)^T + D^G \nabla G(x)^T$. The unconstrained minimization reformulation (2.6) as well as the semismooth equation reformulation (2.2) will be the basis of the proposed trust region method for solving the GCP.

We now introduce some definitions which will be useful in the subsequent analysis. Let $E : \Re^n \to \Re^n$ be locally Lipschitz around $x \in \Re^n$. E is said to be semismooth at x if

$$\lim_{\substack{V \in \partial E(x+th') \\ h' \to h, t \downarrow 0}} Vh'$$

exists for any $h \in \Re^n$ [33]. It has been proved [33] that $E'(x; h)$, the directional derivative of E at x along the direction h, exists for any $h \in \Re^n$ if E is semismooth at x. Moreover, E is said to be strongly semismooth at x if E is semismooth at x and for any $V \in \partial E(x + h)$, $h \to 0$,

$$Vh - E'(x; h) = O(\|h\|^2).$$

A matrix $M \in \Re^{n\times n}$ is called a P-matrix if its principal minors are all positive, a P_0-matrix if its principal minors are all nonnegative, and an S_0-matrix if

$$\{x \in \Re^n \mid x \geq 0, x \neq 0, Mx \geq 0\} \neq \emptyset.$$

It is known [4] that every P-matrix is a P_0-matrix and every P_0-matrix is an S_0-matrix.

3. Algorithm.

The proposed trust region algorithm for solving the GCP (1.1) is now formally stated.

ALGORITHM.

Step 0. Let α_1, α_2, α_3, α_4, ρ_1, ρ_2, $\Delta_{\min}$, and Δ_1 be such that $0 < \alpha_1 < \alpha_2 < 1 < \alpha_3 < \alpha_4$, $0 < \rho_1 < \rho_2 < 1$, $\Delta_{\min} > 0$, and $\Delta_1 > 0$. Let $x^1 \in \Re^n$ be a starting point. Set $k := 1$.

Step 1. If $\nabla\Phi(x^k) = 0$, stop. Otherwise, let $\hat{\Delta} := \max\{\Delta_{\min}, \Delta_k\}$ and choose $V_k \in \partial_B H(x^k)$.

Step 2. Let $\hat{s}$ be a solution of the minimization problem

$$\begin{aligned} &\min \nabla\Phi(x^k)^T s + \frac{1}{2} s^T V_k^T V_k s \\ &\text{s.t.} \|s\| \leq \hat{\Delta}. \end{aligned} \tag{3.1}$$

Step 3. Let

$$\hat{r} := \frac{\Phi(x^k + \hat{s}) - \Phi(x^k)}{\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k)}.$$

If $\hat{r} \geq \rho_1$, then let $s^k := \hat{s}$, $x^{k+1} := x^k + s^k$, $\delta_k := \hat{\Delta}$. For a fixed $\alpha \in [\alpha_3, \alpha_4]$, let

$$\Delta_{k+1} := \begin{cases} \hat{\Delta} & \text{if} \quad \rho_1 \leq \hat{r} < \rho_2, \\ \alpha\hat{\Delta} & \text{if} \quad \hat{r} \geq \rho_2. \end{cases}$$

Let k be replaced by $k + 1$, and return to Step 1. Otherwise, choose $\Delta \in [\alpha_1\hat{\Delta}, \alpha_2\hat{\Delta})$. Let $\hat{\Delta} := \Delta$, and repeat Step 2.

Remarks. (i) In Step 2 of the algorithm, the minimization problem (3.1) may have several solutions. From the first equality in (2.7), it is easy to see that, for any $s \in \Re^n$,

$$\frac{1}{2}\|H(x^k) + V_k s\|_2^2 - \Phi(x^k) = \nabla\Phi(x^k)^T s + \frac{1}{2} s^T V_k^T V_k s. \tag{3.2}$$

(ii) The main difference between the proposed trust region method and the classical trust region methods lies in the updating rule for the trust region radius at the beginning of each iteration. More precisely, at the beginning of each iteration k, the trust region radius is always set greater than the fixed positive constant $\Delta_{\min}$ rather than solely updated from the final trust region radius of iteration $k-1$ as in the classical trust region methods. This type of updating rule for the trust region radius is also used in recent literature [14, 17, 23]. As we shall see, this special strategy enables us not only to establish global convergence but also to recover local superlinear convergence of the algorithm under some conditions, despite that (2.2) is a system of nonsmooth equations.

(iii) Note that δ_k does not play any role in the algorithm. But it will be useful in proving convergence theorems in the subsequent sections.

PROPOSITION 3.1. *The algorithm cannot cycle between Step 2 and Step 3 infinitely, provided that x^k is not a stationary point of Φ.*

Proof. Suppose the algorithm does cycle for some k and x^k is not a stationary point of Φ; i.e., $\nabla\Phi(x^k) \neq 0$. Then the continuous differentiability of Φ implies

$$\begin{aligned}\Phi(x^k+\hat{s}) - \Phi(x^k) &= \nabla\Phi(x^k)^T\hat{s} + o(\|\hat{s}\|)\\ &\le \nabla\Phi(x^k)^T\hat{s} + \frac{1}{2}\hat{s}^T V_k^T V_k \hat{s} + o(\|\hat{s}\|)\\ &= \frac{1}{2}\|H(x^k)+V_k\hat{s}\|_2^2 - \Phi(x^k) + o(\|\hat{s}\|),\end{aligned} \tag{3.3}$$

where the last equality follows from (3.2). Let $s^* = -\frac{\nabla\Phi(x^k)}{\|\nabla\Phi(x^k)\|}$. Clearly, $\tau\hat{\Delta}s^*$ is a feasible solution of (3.1) for any $\tau \in [0,1]$. Since $\hat{s}$ solves (3.1), it follows from (3.2) that for $0 \le \tau \le 1$

$$\begin{aligned}\frac{1}{2}\|H(x^k)+V_k\hat{s}\|_2^2 - \Phi(x^k) &\le \tau\hat{\Delta}\nabla\Phi(x^k)^T s^* + \frac{1}{2}\tau^2\hat{\Delta}^2(s^*)^T V_k^T V_k s^*\\ &= -\tau\hat{\Delta}\|\nabla\Phi(x^k)\| + \frac{1}{2}\tau^2\hat{\Delta}^2(s^*)^T V_k^T V_k s^*.\end{aligned}$$

By the assumption that the algorithm cycles between Steps 2 and 3, both $\|\hat{s}\|$ and $\hat{\Delta}$ tend to zero, while $\hat{r} < \rho_1$ is maintained. Consequently, there exist $\tau^* > 0$ and $\Delta^* > 0$ such that when $\hat{\Delta} \le \Delta^*$

$$\frac{1}{2}\|H(x^k)+V_k\hat{s}\|_2^2 - \Phi(x^k) \le -\tau^*\hat{\Delta}\|\nabla\Phi(x^k)\|.$$

This, together with the definition of $\hat{r}$ and the relation (3.3), implies that

$$\hat{r} \ge 1 - \frac{o(\|\hat{s}\|)}{\tau^*\hat{\Delta}\|\nabla\Phi(x^k)\|} \ge 1 - \frac{o(\|\hat{\Delta}\|)}{\tau^*\hat{\Delta}\|\nabla\Phi(x^k)\|} \to 1 \qquad \text{as} \quad \hat{\Delta} \to 0,$$

which indicates that we eventually have $\hat{r} \ge \rho_1$. This is a contradiction and the desired result follows. □

4. Regularity condition. Most unconstrained minimization methods normally generate a sequence converging to a local minimizer or a stationary point rather than a global minimizer. As discussed in section 2, if the GCP has a solution, then solving the GCP is equivalent to finding a global minimizer of the unconstrained minimization problem (2.6). It is therefore crucial to study under what conditions a stationary point of (2.6) becomes a solution of the GCP. To this end, let us introduce some index sets: $\mathcal{C}(x), \mathcal{R}(x), \mathcal{P}(x)$, and $\mathcal{N}(x)$, which stand for complementarity indices, residual indices, positive indices, and negative indices, respectively; i.e.,

$$\begin{aligned}
\mathcal{C}(x) &:= \{i \in \{1, \dots, n\} \mid F_i(x) \geq 0,\ G_i(x) \geq 0,\ F_i(x)G_i(x) = 0\}, \\
\mathcal{R}(x) &:= \{i \in \{1, \dots, n\} \mid i \notin \mathcal{C}(x)\}, \\
\mathcal{P}(x) &:= \{i \in \mathcal{R}(x) \mid F_i(x) > 0,\ G_i(x) > 0\}, \\
\mathcal{N}(x) &:= \mathcal{R}(x) \setminus \mathcal{P}(x).
\end{aligned}$$

In the rest of the paper, we shall simply denote these sets as $\mathcal{C}, \mathcal{R}, \mathcal{P}$, and $\mathcal{N}$; the point x under consideration will always be clear from the context.

It is easy to verify the following result using (2.3), (2.4), and (2.5). The proof is omitted.

PROPOSITION 4.1. *Let diagonal matrices D^F and D^G satisfy* (2.4) *such that $D^F \nabla F(x)^T + D^G \nabla G(x)^T \in \partial H(x)$. Then for each $x \in \Re^n$, we have the following relations:*

$$\begin{array}{lclcl}
(D^F H(x))_i > 0 & \iff & (D^G H(x))_i > 0 & \iff & i \in \mathcal{P}, \\
(D^F H(x))_i = 0 & \iff & (D^G H(x))_i = 0 & \iff & i \in \mathcal{C}, \\
(D^F H(x))_i < 0 & \iff & (D^G H(x))_i < 0 & \iff & i \in \mathcal{N}.
\end{array}$$

Note that, since D^F and D^G are diagonal matrices, we have $(D^F H(x))_i = D_i^F H_i(x)$ and $(D^G H(x))_i = D_i^G H_i(x)$ for each $i = 1, \dots, n$.

DEFINITION 4.2. *The GCP is said to be regular at a point x if for any two vectors $z^1 \neq 0$, $z^2 \neq 0$ in $\Re^n$ satisfying*

$$\begin{array}{lll}
z^1_{\mathcal{C}} = 0, & z^1_{\mathcal{P}} > 0, & z^1_{\mathcal{N}} < 0, \\
z^2_{\mathcal{C}} = 0, & z^2_{\mathcal{P}} > 0, & z^2_{\mathcal{N}} < 0,
\end{array}$$

we have

$$\nabla F(x) z^1 + \nabla G(x) z^2 \neq 0.$$

Moreover, a stationary point x of Φ is called a regular stationary point if the GCP is regular at x.

It is not hard to see that, when the GCP reduces to the NCP, the regularity condition introduced here is slightly weaker than that defined in [5]. Under the present regularity condition, the following theorem establishes the equivalence between solutions of the GCP and stationary points of the merit function Φ.

THEOREM 4.3. *x is a solution of the GCP if and only if x is a regular stationary point of Φ.*

Proof. If x is a solution of the GCP, then x is a stationary point of Φ, and, hence, $\mathcal{P} = \mathcal{N} = \emptyset$. By definition, x is a regular point of Φ.

Conversely, suppose x is a regular stationary point of Φ. Let D^F and D^G satisfy the conditions of Proposition 4.1. Let $z^1 = D^F H(x)$ and $z^2 = D^G H(x)$. Then

$$0 = \nabla \Phi(x) = \nabla F(x) D^F H(x) + \nabla G(x) D^G H(x) = \nabla F(x) z^1 + \nabla G(x) z^2. \tag{4.1}$$

If x is not a solution of the GCP, it follows from Proposition 4.1 that $z^1 \neq 0$, $z^2 \neq 0$. By the definition of regularity and Proposition 4.1, for z^1 and z^2,

$$\nabla F(x) z^1 + \nabla G(x) z^2 \neq 0,$$

which contradicts (4.1). Therefore x is a solution of the GCP. □

Next we present a sufficient condition for ensuring regularity of the GCP. The proof is based on a result of [5] for the NCP. In fact, when the GCP happens to be the NCP, the obtained result boils down to that of [5].

Let $D \in \Re^{|\mathcal{R}|\times|\mathcal{R}|}$ denote a diagonal matrix with diagonal elements $D_1, \dots, D_{|\mathcal{R}|}$ defined by

$$D_i = \begin{cases} 1 & \text{if } i \in \mathcal{P}, \\ -1 & \text{if } i \in \mathcal{N}. \end{cases}$$

Evidently, DD is the $|\mathcal{R}| \times |\mathcal{R}|$ identity matrix. Using this notation, we establish the following proposition.

PROPOSITION 4.4. *Assume that $\nabla G(x)$ is invertible and*

$$D(\nabla F(x)^T(\nabla G(x)^T)^{-1})_{\mathcal{R}\mathcal{R}}D$$

is an S_0-matrix. Then the GCP is regular at x.

Proof. Since $D(\nabla F(x)^T(\nabla G(x)^T)^{-1})_{\mathcal{R}\mathcal{R}}D$ is an S_0-matrix, there exists a vector $y^1_{\mathcal{R}} \neq 0$ such that

$$y^1_{\mathcal{R}} \geq 0, \quad D(\nabla F(x)^T(\nabla G(x)^T)^{-1})_{\mathcal{R}\mathcal{R}}Dy^1_{\mathcal{R}} \geq 0. \tag{4.2}$$

Let vectors $y^2 \in \Re^n$ and $y \in \Re^n$ be such that

$$y^2_{\mathcal{C}} = 0, \quad y^2_{\mathcal{R}} = Dy^1_{\mathcal{R}}, \tag{4.3}$$

$$y = (\nabla G(x)^{-1})^T y^2. \tag{4.4}$$

Clearly,

$$y^2_{\mathcal{R}} \neq 0, \quad y^2_{\mathcal{P}} \geq 0, \quad y^2_{\mathcal{N}} \leq 0. \tag{4.5}$$

For any two vectors $z^1 \neq 0$, $z^2 \neq 0$ in $\Re^n$ satisfying

$$\begin{array}{lll} z^1_{\mathcal{C}} = 0, & z^1_{\mathcal{P}} > 0, & z^1_{\mathcal{N}} < 0, \\ z^2_{\mathcal{C}} = 0, & z^2_{\mathcal{P}} > 0, & z^2_{\mathcal{N}} < 0, \end{array}$$

it follows from (4.3), (4.4), and (4.5) that

$$y^T\nabla G(x)z^2 = (y^2)^T z^2 = (y^2_{\mathcal{R}})^T z^2_{\mathcal{R}} > 0. \tag{4.6}$$

By the definition of y, y^2, and z^1, we have

$$\begin{array}{rcl} y^T\nabla F(x)z^1 & = & (y^2)^T\nabla G(x)^{-1}\nabla F(x)z^1 \\ & = & (y^2_{\mathcal{R}})^T(\nabla G(x)^{-1}\nabla F(x))_{\mathcal{R}\mathcal{R}}z^1_{\mathcal{R}} \\ & = & (y^1_{\mathcal{R}})^T D(\nabla G(x)^{-1}\nabla F(x))_{\mathcal{R}\mathcal{R}}DDz^1_{\mathcal{R}}. \end{array}$$

Then (4.2) and the fact that $Dz^1_{\mathcal{R}} > 0$ imply that

$$y^T\nabla F(x)z^1 \geq 0,$$

which together with (4.6) yields

$$y^T(\nabla F(x)z^1 + \nabla G(x)z^2) > 0.$$

Consequently,

$$\nabla F(x)z^1 + \nabla G(x)z^2 \neq 0.$$

Therefore the GCP is regular at x. ∎

It has been proved in [20] that any stationary point of Φ is a solution of the GCP if $\nabla G(x)^T$ is invertible and $\nabla F(x)^T(\nabla G(x)^T)^{-1}$ is a P_0-matrix. It is not difficult to see that this result is a consequence of Proposition 4.4 by using the fact that a matrix M is a P_0-matrix if and only if the matrix DMD is a P_0-matrix for any nonsingular diagonal matrix D and the fact that any P_0-matrix is an S_0-matrix.

5. Global convergence. We now suppose that the algorithm generates an infinite sequence $\{x^k\}$; i.e., the stopping test in Step 1 of the algorithm is never fulfilled. Let $c > 0$ be a constant such that

$$\|y\| \leq c\|y\|_2$$

for all $y \in \Re^n$. We first present a standard result.

LEMMA 5.1. *Let $\hat{s}$ be a solution of* (3.1). *Then*

$$\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k) \leq -\frac{1}{2}\|\nabla\Phi(x^k)\|_2 \min\left\{\frac{\hat{\Delta}}{c}, \frac{\|\nabla\Phi(x^k)\|_2}{\|V_k^TV_k\|_2}\right\}.$$

Proof. Suppose $\tilde{s}$ is a solution of the following minimization problem

$$\begin{aligned} \min \quad & \nabla\Phi(x^k)^T s + \frac{1}{2}s^T V_k^T V_k s \\ \text{s.t.} \quad & \|s\|_2 \leq \frac{\hat{\Delta}}{c}. \end{aligned}$$

Then from Theorem 4 in [31],

$$\frac{1}{2}\|H(x^k) + V_k\tilde{s}\|_2^2 - \Phi(x^k) \leq -\frac{1}{2}\|\nabla\Phi(x^k)\|_2 \min\left\{\frac{\hat{\Delta}}{c}, \frac{\|\nabla\Phi(x^k)\|_2}{\|V_k^TV_k\|_2}\right\}. \tag{5.1}$$

Since $\|\tilde{s}\| \leq c\|\tilde{s}\|_2 \leq \hat{\Delta}$, $\tilde{s}$ is a feasible solution of (3.1). Since $\hat{s}$ solves (3.1), the desired result follows from (3.2) and (5.1). ∎

LEMMA 5.2. *Suppose x^* is the limit of a subsequence $\{x^k\}_{k\in K}$. If x^* is not a stationary point of Φ, then there exist $\hat{k}$ and $\Delta > 0$ such that for all $k \geq \hat{k}$ $(k \in K)$*

$$\hat{r} := \frac{\Phi(x^k + \hat{s}) - \Phi(x^k)}{\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k)} \geq \rho_1$$

whenever $\hat{\Delta} \in (0, \Delta)$, where $\hat{s}$ is a solution of (3.1).

Proof. First note that

$$\hat{r} = 1 + \frac{-\frac{1}{2}\hat{s}^T V_k^T V_k \hat{s} + o(\|\hat{s}\|_2)}{\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k)}. \tag{5.2}$$

Since x^* is not a stationary point and ∂H is upper semicontinuous, there exist positive constants β_1 and β_2 such that

$$\|\nabla\Phi(x^k)\|_2 \geq \beta_1, \quad \frac{\|\nabla\Phi(x^k)\|_2}{\|V_k^TV_k\|_2} \geq \beta_2 \tag{5.3}$$

for all sufficiently large $k \in K$. By Lemma 5.1,

$$\frac{1}{2}\|H(x^k) + V_k \hat{s}\|_2^2 - \Phi(x^k) \leq -\frac{1}{2}\beta_1 \min\left\{\frac{\hat{\Delta}}{c}, \beta_2\right\} \tag{5.4}$$

for all sufficiently large $k \in K$. Then the desired result is a direct consequence of (5.2), (5.3), and (5.4). ∎

LEMMA 5.3. *Suppose x^* is the limit of a subsequence $\{x^k\}_{k \in K}$. If x^* is not a stationary point of Φ, then*

$$\liminf_{k \in K,\ k \to \infty} \delta_k > 0,$$

where δ_k is defined as in the algorithm.

Proof. By Lemma 5.2, there exists a $\Delta > 0$ such that $\hat{r} \geq \rho_1$ whenever $\hat{\Delta} < \Delta$ at each iteration $k \in K$ sufficiently large. Thus, by the updating rule of the trust region radius in the algorithm, we have $\delta_k \geq \alpha_1 \Delta$ for all sufficiently large $k \in K$. The desired result follows. ∎

Now we are ready to establish a global convergence theorem for the proposed trust region algorithm.

THEOREM 5.4. *Let $\{x^k\}$ be generated by the algorithm. Then any accumulation point of $\{x^k\}$ is a stationary point of Φ. Moreover, it is a solution of the GCP if the regularity condition holds at this point.*

Proof. Suppose x^* is an accumulation point of $\{x^k\}$, say $\lim_{k \in K,\ k \to \infty} x^k = x^*$. If x^* is not a stationary point of Φ, then Lemmas 5.1 and 5.3 imply that there exist $\beta_1 > 0$, $\beta_2 > 0$, $\delta > 0$, and $\hat{k}$ such that (5.4) holds and $\delta_k \geq \delta$ for all $k \geq \hat{k}$ $(k \in K)$. By the algorithm, Lemma 5.1, (5.4), and the fact that Φ is nonnegative, we obtain

$$\begin{aligned}
\Phi(x^1) &\geq \sum_{k=1}^{\infty} \left[\Phi(x^k) - \Phi(x^{k+1})\right] \\
&\geq \sum_{k=1}^{\infty} \rho_1 \left[\Phi(x^k) - \frac{1}{2}\|H(x^k) + V_k s^k\|_2^2\right] \\
&\geq \sum_{\substack{k \geq \hat{k} \\ k \in K}} \rho_1 \left[\Phi(x^k) - \frac{1}{2}\|H(x^k) + V_k s^k\|_2^2\right] \\
&\geq \rho_1 \sum_{\substack{k \geq \hat{k} \\ k \in K}} \frac{1}{2}\beta_1 \min\left\{\frac{\delta_k}{c}, \beta_2\right\} \\
&\geq \frac{1}{2}\rho_1 \beta_1 \sum_{\substack{k \geq \hat{k} \\ k \in K}} \min\left\{\frac{\delta}{c}, \beta_2\right\} \ = \ \infty.
\end{aligned}$$

This is impossible. Therefore, x^* is a stationary point of Φ. ∎

We now turn to the case where $\{x^k\}$ does not necessarily have an accumulation point.

THEOREM 5.5. *Let $\{x^k\}$ be generated by the algorithm. If $\{V_k\}$ is bounded, then $\{\nabla\Phi(x^k)\}$ is not bounded away from zero; that is,*

$$\liminf_{k \to \infty} \|\nabla\Phi(x^k)\| = 0. \tag{5.5}$$

Proof. Suppose (5.5) does not hold. Then the boundedness of $\{V_k\}$ and Lemma 5.1 imply the existence of β_1 and β_2 such that (5.4) holds for all k. By the algorithm and (5.4), we obtain

$$\begin{aligned}\Phi(x^1) &\geq \sum_{k=1}^{\infty}\left[\Phi(x^k) - \Phi(x^{k+1})\right] \\ &\geq \sum_{k=1}^{\infty}\rho_1\left[\Phi(x^k) - \frac{1}{2}\|H(x^k) + V_k s^k\|_2^2\right] \\ &\geq \frac{1}{2}\rho_1\beta_1\sum_{k=1}^{\infty}\min\left\{\frac{\delta_k}{c}, \beta_2\right\}.\end{aligned}$$

This implies

$$\sum_{k=1}^{\infty}\delta_k < \infty,$$

and, hence,

$$\sum_{k=1}^{\infty}\|x^{k+1} - x^k\| < \infty.$$

This implies that $\{x^k\}$ converges to a point x^*. By Theorem 5.4, x^* is a stationary point of Φ; i.e, $\nabla\Phi(x^*) = 0$. This contradicts the assumption that

$$\liminf_{k\to\infty}\|\nabla\Phi(x^k)\| > 0. \qquad \square$$

Remark. If ∇F and ∇G are bounded on $\Re^n$, then ∂H is bounded on $\Re^n$ and the boundedness assumption on $\{V_k\}$ is satisfied.

This theorem says that the sequence $\{x^k\}$ generated by the algorithm contains a stationary subsequence $\{x^k\}_{k\in K}$ in the sense that $\lim_{k\to\infty, k\in K}\|\nabla\Phi(x^k)\| = 0$, even if $\{x^k\}$ is unbounded. However, a stationary sequence is not necessarily a minimizing sequence of Φ in general. Conditions under which any stationary sequence of Φ is a minimizing sequence have been studied in [16] for the NCP.

6. Superlinear convergence. In this section, we shall be concerned with the rate of convergence of the algorithm. It is known [33, 32] that semismoothness and certain nonsingularity conditions at a solution of the system of nonsmooth equations play a crucial role in establishing superlinear convergence of some generalized Newton methods. Recall that H is said to be BD-regular at a point x if all the elements in $\partial_B H(x)$ are nonsingular [32]. We now derive some sufficient conditions for ensuring BD-regularity at a solution of the GCP.

Suppose $\nabla G(\bar{x})^T$ is invertible at a solution $\bar{x}$ of the GCP. Let

$$M = \begin{pmatrix} M_{\alpha\alpha} & M_{\alpha\beta} & M_{\alpha\gamma} \\ M_{\beta\alpha} & M_{\beta\beta} & M_{\beta\gamma} \\ M_{\gamma\alpha} & M_{\gamma\beta} & M_{\gamma\gamma}\end{pmatrix} := \nabla F(\bar{x})^T(\nabla G(\bar{x})^T)^{-1},$$

where

$$\begin{aligned}\alpha &:= \{i \in \{1,\dots,n\} \mid F_i(\bar{x}) = 0, G_i(\bar{x}) > 0\}, \\ \beta &:= \{i \in \{1,\dots,n\} \mid F_i(\bar{x}) = 0, G_i(\bar{x}) = 0\}, \\ \gamma &:= \{i \in \{1,\dots,n\} \mid F_i(\bar{x}) > 0, G_i(\bar{x}) = 0\}.\end{aligned}$$

The GCP is said to be R-regular at $\bar{x}$ if $M_{\alpha\alpha}$ is nonsingular and the Schur complement of $M_{\alpha\alpha}$ in the matrix

$$\begin{pmatrix} M_{\alpha\alpha} & M_{\alpha\beta} \\ M_{\beta\alpha} & M_{\beta\beta} \end{pmatrix};$$

i.e., $M_{\beta\beta} - M_{\beta\alpha}M_{\alpha\alpha}^{-1}M_{\alpha\beta}$ is a P-matrix, see [34].

PROPOSITION 6.1. *Let $\bar{x}$ be a solution of the GCP. Suppose $\nabla G(\bar{x})^T$ is invertible. If the GCP is R-regular at $\bar{x}$, then any $V \in \partial H(\bar{x})$ is nonsingular.*

Proof. Since $\nabla G(\bar{x})^T$ is invertible, by (2.3) any $V \in \partial H(\bar{x})$ can be represented as

$$\begin{aligned} V &= D^F \nabla F(\bar{x})^T + D^G \nabla G(\bar{x})^T \\ &= \left(D^F \nabla F(\bar{x})^T (\nabla G(\bar{x})^T)^{-1} + D^G\right) \nabla G(\bar{x})^T \\ &= (D^F M + D^G) \nabla G(\bar{x})^T, \end{aligned}$$

where D^F and D^G satisfy the condition (2.4) and $M = \nabla F(\bar{x})^T(\nabla G(\bar{x})^T)^{-1}$. Then V is nonsingular if and only if $D^F M + D^G$ is nonsingular. By some standard analysis (see, e.g., [9]), the nonsingularity of $D^F M + D^G$ can be deduced from the fact that the GCP is R-regular at x. ∎

COROLLARY 6.2. *Suppose $\nabla G(\bar{x})^T$ is invertible. If $\nabla F(\bar{x})^T(\nabla G(\bar{x})^T)^{-1}$ is a P-matrix, then any $V \in \partial H(\bar{x})$ is nonsingular.*

Proof. Note that every principal submatrix of a P-matrix is a P-matrix, hence nonsingular, and the Schur complement of every principal submatrix of a P-matrix is a P-matrix. Therefore, the GCP is R-regular at $\bar{x}$. The desired result follows from Proposition 6.1. ∎

LEMMA 6.3. (a) *H is semismooth on $\Re^n$ if F and G are continuously differentiable on $\Re^n$;* (b) *H is strongly semismooth on $\Re^n$ if ∇F and ∇G are Lipschitz continuous on $\Re^n$.*

Proof. Apparently, H is (strongly) semismooth at x if and only if each component of H is (strongly) semismooth at x. Therefore it suffices to prove that H_1 is (strongly) semismooth at any $x \in \Re^n$ under different assumptions.

Note that H_1 can be regarded as a composition of the function $\phi : \Re^2 \to \Re$ defined by (2.1) and the function $h : \Re^n \to \Re^2$ with $h(x) = (F_1(x), G_1(x))$; i.e., $H_1(x) = \phi(h(x))$. It is known that the composition of semismooth functions are semismooth [24, Theorem 5], and the composition of strongly semismooth functions are strongly semismooth [13, Theorem 5.7]. By Lemma 5.6 in [13], ϕ is strongly semismooth, hence semismooth on $\Re^2$. Therefore (a) holds if F and G are semismooth on $\Re^n$, and (b) holds if F and G are strongly semismooth on $\Re^n$. On the other hand, the semismoothness of F and G follows from the continuous differentiability of F and G, and the strong semismoothness of F and G follows from the Lipschitz continuity of the Jacobians of F and G. ∎

LEMMA 6.4. *Let $\{x^k\}$ be generated by the algorithm and x^* be an accumulation point of $\{x^k\}$. If the BD-regularity condition holds at x^*, then x^* is a solution of the GCP and there exists an open neighborhood $\mathcal{N}(x^*)$ of x^* such that when $x^k \in \mathcal{N}(x^*)$, we have*

$$x^{k+1} = x^k - V_k^{-1}H(x^k) \in \mathcal{N}(x^*).$$

Proof. By the BD-regularity condition and Theorem 5.4, x^* is a solution of the GCP and $V_k^{-1}H(x^k) \to 0$ as $x^k \to x^*$. Therefore, when x^k is sufficiently close to x^*, we have $\|V_k^{-1}H(x^k)\| \le \Delta_{\min}$ and $-V_k^{-1}H(x^k)$ is a solution of (3.1) if $\hat{\Delta} \ge \Delta_{\min}$. On the other hand, we have $\|H(x^k - V_k^{-1}H(x^k))\| = o(\|H(x^k)\|)$ as $x^k \to x^*$ by Theorem 3.1 in [32]. This implies that $\hat{r} \ge \rho_1$ holds for some $\hat{\Delta} \ge \Delta_{\min}$ when x^k is sufficiently close to x^*. This in turn implies that, by the updating rule of the trust region radius, $\Delta_k \ge \Delta_{\min}$ for all k large enough. Therefore we have $x^{k+1} = x^k - V_k^{-1}H(x^k)$ for all k sufficiently large, and $\|x^{k+1} - x^*\| = o(\|x^k - x^*\|)$ as $x^k \to x^*$ from Theorem 3.1 in [32]. The conclusion follows easily. □

We are now in a position to present rate of convergence results for the proposed algorithm.

THEOREM 6.5. *Suppose that the BD-regularity condition holds at an accumulation point x^* of the sequence $\{x^k\}$ generated by the algorithm. Then the entire sequence $\{x^k\}$ converges to x^* Q-superlinearly if F and G are continuously differentiable on $\Re^n$. Moreover, the rate of convergence is Q-quadratic if ∇F and ∇G are Lipschitz continuous on $\Re^n$.*

Proof. Since x^* is an accumulation point of $\{x^k\}$, there exists a subsequence $\{x^k\}_{k\in K}$ such that $\lim_{k\in K, k\to\infty} x^k = x^*$. By Lemma 6.4, there exists $\hat{k} \in K$ such that when $k \ge \hat{k}$

$$x^{k+1} = x^k - V_k^{-1}H(x^k).$$

It follows that the algorithm reduces to the generalized Newton method considered in [32]. Therefore, Q-superlinear convergence is guaranteed by Theorem 3.1 in [32]. Note that H is strongly semismooth if ∇F and ∇G are Lipschitz continuous. This implies Q-quadratic convergence again by Theorem 3.1 in [32]. □

7. Computation of a generalized Jacobian.

In the algorithm, we assumed that a generalized Jacobian of H is available at any point x. We now present a method for calculating a generalized Jacobian of H at x.

Define

$$I(x) := \{i \in \{1, \dots, n\} \mid F_i(x)^2 + G_i(x)^2 = 0,\ \|\nabla F_i(x)\| + \|\nabla G_i(x)\| > 0\}.$$

If $I(x) \ne \emptyset$, then we may assume, without loss of generality, that $I(x) = \{1, \dots, k\}$ for some $k \le n$ and that $\nabla F_i(x) \ne 0$ for each $i \in I(x)$. Now we shall construct a vector $d \in \Re^n$ such that

$$\nabla F_i(x)^T d \ne 0 \quad \text{for } i \in I(x). \tag{7.1}$$

First let $d := \nabla F_1(x)$ and

$$J(x) := \{i \in I(x) \mid \nabla F_i(x)^T d = 0\}.$$

If $J(x) = \emptyset$, then d satisfies (7.1). Otherwise, choose $j \in J(x)$, let

$$\bar{d} := \frac{\min_{i\in I(x)\setminus J(x)} |\nabla F_i(x)^T d|}{2\max_{i\in J(x)} \|\nabla F_i(x)\|\|\nabla F_j(x)\|} \nabla F_j(x),$$

and put $\hat{d} := d + \bar{d}$. Then it is clear that $\nabla F_j(x)^T \hat{d} \ne 0$. Moreover, for $i \in I(x) \setminus J(x)$, we have

$$\nabla F_i(x)^T \hat{d} = \nabla F_i(x)^T d + \frac{\min_{i\in I(x)\setminus J(x)} |\nabla F_i(x)^T d|}{2\max_{i\in J(x)} \|\nabla F_i(x)\|\|\nabla F_j(x)\|} \nabla F_i(x)^T \nabla F_j(x).$$

Since the absolute value of the second term on the right-hand side is never greater than that of the first term, we have $\nabla F_i(x)^T \hat{d} \neq 0$ for $i \in I(x) \setminus J(x)$. Define $\hat{J}(x)$ by

$$\hat{J}(x) := \{i \in I(x) \mid \nabla F_i(x)^T \hat{d} = 0\}.$$

The above arguments indicate that $\hat{J}(x) \subseteq J(x) \setminus \{j\}$. Put $d := \hat{d}$ and $J(x) := \hat{J}(x)$. If $J(x) = \emptyset$, then d satisfies (7.1). Otherwise, choose an index from $J(x)$ and repeat the above procedure. After at most k steps, we will have $J(x) = \emptyset$ and a vector d satisfying (7.1).

For small $t > 0$, let

$$y(t) := x + td + \hat{d}(t),$$

where $\|\hat{d}(t)\| = o(t)$ and d is a vector satisfying (7.1). Then H is differentiable at $y(t)$ with appropriately chosen $\hat{d}(t)$. Letting t tend to zero, we obtain a matrix $V := \lim_{t \to 0^+} \nabla H(y(t))^T$, which is an element of $\partial_B H(x)$ with the form $D^F \nabla F(x)^T + D^G \nabla G(x)^T$, where $D^F = \mathrm{diag}\{D_1^F, \ldots, D_n^F\}$ and $D^G = \mathrm{diag}\{D_1^G, \ldots, D_n^G\}$ are determined by

$$D_i^F = \begin{cases} \dfrac{F_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1 & \text{if } F_i(x)^2 + G_i(x)^2 > 0, \\ \xi_i & \text{if } F_i(x) = G_i(x) = 0, \text{ and } \\ & \quad \|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0, \\ \dfrac{\nabla F_i(x)^T d}{\sqrt{(\nabla F_i(x)^T d)^2 + (\nabla G_i(x)^T d)^2}} - 1 & \text{if } F_i(x) = G_i(x) = 0, \text{ and } \\ & \quad \|\nabla F_i(x)\| + \|\nabla G_i(x)\| > 0, \end{cases}$$

$$D_i^G = \begin{cases} \dfrac{G_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1 & \text{if } F_i(x)^2 + G_i(x)^2 > 0, \\ \eta_i & \text{if } F_i(x) = G_i(x) = 0, \text{ and } \\ & \quad \|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0, \\ \dfrac{\nabla G_i(x)^T d}{\sqrt{(\nabla F_i(x)^T d)^2 + (\nabla G_i(x)^T d)^2}} - 1 & \text{if } F_i(x) = G_i(x) = 0, \text{ and } \\ & \quad \|\nabla F_i(x)\| + \|\nabla G_i(x)\| > 0 \end{cases}$$

for $i = 1, \ldots, n$, and ξ_i and η_i are some constants. Note that we have not specified any fixed numbers for D_i^F and D_i^G if $F_i(x) = G_i(x) = 0$ and $\|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0$. In practice, however, this does not cause any problem in calculating a generalized Jacobian $V \in \partial_B H(x)$. In fact, as shown easily, $\|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0$ implies that the ith row of any $V \in \partial_B H(x)$ becomes a zero vector.

8. Numerical experiments. In this section, we present some numerical experiments for the proposed algorithm. We have chosen the l_∞-norm in the constraint set of the subproblem (3.1) so that (3.1) becomes a linear least squares problem with box constraints. We implemented a nonmonotone version of the algorithm in the sense that $\hat{r}$ in Step 3 of the algorithm is defined as

$$\hat{r} := \frac{\Phi(x^k + s) - \Psi_k}{\frac{1}{2}\|H(x^k) + V_k \hat{s}\|_2^2 - \Phi(x^k)},$$

where $\Psi_k := \max\{\Phi(x^l) \mid l = k - l_k, \ldots, k\}$ and l_k is a nonnegative integer. This means that the objective function value is decreased compared with the maximum of the objective function values in the last $l_k + 1$ iterations, not necessarily decreased compared with the objective function value at the very last iteration. The nonmonotone version of the algorithm reduces to the algorithm in section 3 if $l_k \equiv 0$ for any k. In the code, we simply let $l_k = 3$ for $k \geq 4$ and $l_k = k - 1$ for $k = 2, 3$. The motivation to use the nonmonotone version of the algorithm is that the nonmonotone strategy can advance computational efficiency for complementarity problems, as observed in [6, 5, 21]. This has also been confirmed by our experience.

The algorithm was implemented in MATLAB and run on a SPARC 10 workstation. The subproblem (3.1) was solved by the **qp.m** inside MATLAB. Throughout the computational experiments, the parameters used in the algorithm were set as $\Delta_{\min} = 1.0$, $\Delta_1 = 100$, $\rho_1 = 10^{-4}$, $\rho_2 = 0.75$. The trust region radius is updated as follows: If $\hat{r} \geq \rho_2$, then $\Delta_{k+1} := 2\hat{\Delta}$; if $\hat{r} < \rho_1$, then $\hat{\Delta} := \hat{\Delta}/2$. We used

$$\min\{\|\min\{F(x), G(x)\}\|_\infty, \|\nabla\Phi(x)\|_2\} \leq 10^{-6}$$

as a stopping rule, where $\min\{F(x), G(x)\}$ denotes the vector with components min $\{F_i(x), G_i(x)\}$, $i = 1, \ldots, n$. Note that the second term on the left-hand side of the above stopping rule is used as a safeguard against the case that an accumulation point of the sequence generated by the algorithm is a mere stationary point of Φ, which is not a solution of the NCP or the GCP.

The code stated above was tested on the problems from libraries GAMSLIB and MCPLIB [1, 7, 10]. For our purpose, we tested all linear and nonlinear complementarity problems from the libraries and leave other problems such as mixed complementarity problems in the sense of [1], nonlinear equations, and the KKT conditions of nonlinear programming problems for the future investigation when the corresponding theoretical results are justified. We have noted that some of the test problems such as **cammcp, hansmcp**, and **vonthmcp** in GAMSLIB, and **choi** and **powell_mcp** in MCPLIB are actually not NCPs, although they were originally classified as NCPs [1]. Therefore **cammcp, hansmcp, vonthmcp, choi**, and **powell_mcp** were not tested for our code. On the other hand, we tested some problems such as those with the suffix "mge" in GAMSLIB, and **colvdual, colvnlp, nash**, and **powell** in MCPLIB. These problems were not originally classified as NCPs.

The numerical results are summarized in Tables 8.1 and 8.2 for the problems from the libraries GAMSLIB and MCPLIB, respectively. In Tables 8.1 and 8.2, **Dim** denotes the number of the variables in the problem, **Iter** denotes the number of iterations, which is also equal to the number of Jacobian evaluations for the function F and G, **NF** denotes the number of function evaluations for the functions F and G, and Φ denotes the final value of Φ at the found solution.

Two generalized complementarity problems in [27] were tested too. The interested reader is referred to [27] for full details of these two examples. The results are shown in Table 8.3. Three different starting points have been used, as distinguished by **Start** in Table 8.3.

Outrata–Zowe first problem [27]. Here $n = 4$. It is an implicit complementarity problem with both F and G being linear functions. We used the same three starting points as in [27].

Starting points: (a) $(0, 0, 0, 0)$, (b) $(-0.5, -0.5, -0.5, -0.5)$, (c) $(-1, -1, -1, -1)$.

Outrata–Zowe second problem [27]. Here $n = 4$. This problem is a modification of the last problem in which F is unchanged, but G is a nonlinear function.

TABLE 8.1
Numerical results for the problems from GAMSLIB.

Problem	Dim	Iter	NF	Φ
cafemge	101	31	95	5.1899e-13
cammge	128	0	1	5.09253e-13
co2mge	208	1	2	1.2755e-14
dmcmge	170	–	–	–
etamge	114	276	911	4.07511e-14
finmge	153	0	1	2.19764e-14
hansmge	43	13	23	3.46003e-23
harkmcp*	32	9	12	4.89955e-09
kehomge	9	11	19	8.14523e-20
mr5mcp	350	23	43	1.54433e-14
nsmge	212	21	42	1.4332e-16
oligomcp	6	6	7	5.41969e-17
sammge	23	0	1	0
scarfmcp	18	11	16	1.68098e-16
scarfmge	18	13	19	1.63513e-16
shovmge	51	1	2	5.58898e-14
threemge	9	0	1	0
transmcp	11	13	17	5.87216e-15
two3mcp	6	11	16	3.59176e-15
unstmge	5	8	19	1.56619e-16
vonthmge	80	–	–	–

TABLE 8.2
Numerical results for the problems from MCPLIB.

Problem	Dim	Iter	NF	Φ
bertsekas	15	17	42	1.97117e-16
billups	1	81	1085	4.97509e-05
colvdual	20	260	3458	1.08274e-4
colvnlp	15	27	61	5.79401e-14
cycle	1	3	10	1.70007e-21
explcp	16	19	42	3.6185e-14
hanskoop	14	20	54	4.56762e-18
josephy	4	26	45	2.97784e-14
kojshin	4	13	14	4.59541e-13
mathinum	3	4	5	4.33461e-17
mathisum	3	6	14	3.70779e-22
nash	10	8	9	9.61443e-20
pgvon105	105	18	38	1.25115e-13
pgvon106	106	–	–	–
powell	16	9	17	7.80373e-18
scarfanum	13	11	21	1.68104e-16
scarfasum	14	7	19	1.78446e-16
scarfbnum	39	25	47	5.58127e-15
scarfbsum	40	17	25	2.76111e-18
sppe	27	8	9	3.92819e-18
tobin	42	12	15	2.59193e-24

Starting points: (a) $(0,0,0,0)$, (b) $(-0.5,-0.5,-0.5,-0.5)$, (c) $(-1,-1,-1,-1)$.

The numerical results presented in Tables 8.1, 8.2, and 8.3 show that the proposed method is viable for solving most NCPs from GAMSLIB and MCPLIB as well as the two GCPs efficiently. In Table 8.1, our code failed to solve **dmcmge** and **vonthmge** within 500 iterations. The problem **harkmcp** was solved by using some

TABLE 8.3
Numerical results for the two GCPs.

Problem	Dim	Start	Iter	NF	Φ
Outrata-Zowe 1	4	(a)	5	17	7.64995e-18
	4	(b)	4	16	9.7148e-15
	4	(c)	5	11	3.4347e-24
Outrata-Zowe 2	4	(a)	5	17	1.0474e-18
	4	(b)	4	16	4.88604e-15
	4	(c)	5	11	7.054e-22

minor modification of the code. Specifically, the Hessian of the objective function in the quadratic programming subproblem (3.1) was perturbed by adding $10^{-10} \times I$, where I is the identity matrix of an appropriate dimension, if the condition number of the Hessian is greater than 10^{15}. The reason we made this change in the code for **harkmcp** is that we noticed that the quadratic programming code **qp.m** failed to produce a solution when no perturbation was adopted. In Table 8.2, the code reached to a stationary point but it was an approximate solution point for the problem **colvdual**. The code also failed to solve the problem **pgvon106** because the machine was unable to evaluate the objective function, i.e., $\Phi = \text{NaN}$, after the seventh iteration. However, we mention that **pgvon106** is not an NCP since some lower bounds of the variable x are 10^{-7} rather than zero and that the Jacobian of the function F is highly ill conditioned when components of x are close to zero. Notice in Table 8.1 that our code terminated without proceeding any iteration for some problems in GAMSLIB; i.e. Iter= 0. This is because the starting point provided in GAMSLIB is very close to the solution of the corresponding problem, which can also be observed from the value of Φ in the last column of Table 8.1.

9. Conclusions. In this paper, we have proposed a trust region method for solving the generalized complementarity problem by using both semismooth equation and differentiable optimization reformulations. The special trust region updating rule enables us to establish not only global convergence but also local superlinear convergence of the algorithm under some suitable conditions. We remark again that our trust region method is very different from other existing methods of using line search schemes as far as the globalization strategy is concerned. The proposed algorithm was implemented in MATLAB and was tested for all the NCPs from GAMSLIB and MCPLIB libraries. The preliminary numerical results presented show the viability of this method. The code successfully solved most of the test problems in a reasonably small number of function and Jacobian evaluations, although it failed or converged slowly in some cases. The latter fact suggests that there remain more issues to be addressed. This may be regarded as a further research topic. As expressed, we only tested NCPs from these two libraries. Therefore a possible future topic is to extend the proposed method to the mixed nonlinear complementarity problem which contains the variational inequality problem as a special case if the KKT reformulation of the latter problem is used.

Acknowledgments. The authors are grateful to Dr. Kouichi Taji for his helpful comment on Definition 4.2. We are thankful to Professor Michael Ferris for his kind help regarding the GAMSLIB and MCPLIB libraries. We are indebted to Professor Olvi Mangasarian and anonymous referees for their constructive comments which improved our numerical experiments. We also thank the Department of Electrical and Electronic Engineering at the University of Melbourne for providing access to the

SPARC 10 workstation used in numerical tests.

REFERENCES

[1] S. C. Billups, S. P. Dirkse, and M. C. Ferris, *A comparison of large scale mixed complementarity problem solvers*, Comput. Optim. Appl., 7 (1997), pp. 3–25.

[2] C. Chen and O. L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Comput. Optim. Appl., 5 (1996), pp. 97–138.

[3] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

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

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

[6] S. P. Dirkse and M. C. Ferris, *The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems*, Optimization Methods and Software, 5 (1995), pp. 123–156.

[7] S. P. Dirkse and M. C. Ferris, *MCPLIB: A collection of nonlinear mixed complementarity problems*, Optimization Methods and Software, 5 (1995), pp. 319–345.

[8] I. C. Dolcetta and U. Mosco, *Implicit complementarity problems and quasi-variational inequalities*, in Variational Inequalities and Complementarity Problems: Theory and Applications, R. W. Cottle, F. Giannessi, and J.-L. Lions, eds., Wiley, New York, 1980, pp. 75–87.

[9] F. Facchinei and J. Soares, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM J. Optim., 7 (1997), pp. 225–247.

[10] M. C. Ferris and T. F. Rutherford, *Accessing realistic mixed complementarity problems within MATLAB*, in Nonlinear Optimization and Applications, G. Di Pillo and F. Giannessi, eds., Plenum Press, New York, 1996, pp. 141–153.

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

[12] A. Fischer, *An NCP-function and its use for the solution of complementarity problems*, in Recent Advances in Nonsmooth Optimization, D. Du, L. Qi, and R. Womersley, eds., World Scientific Publishers, River Edge, NJ, 1995, pp. 88–105.

[13] A. Fischer, *Solution of monotone complementarity problems with locally Lipschitzian functions*, Math. Programming, 76 (1997), pp. 513–532.

[14] A. Friedlander, J. M. Martínez, and A. Santos, *A new algorithm for bound constrained minimization*, J. Appl. Math. Optim., 30 (1994), pp. 235–266.

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

[16] M. Fukushima and J.-S. Pang, *Minimizing and stationary sequences of merit functions for complementarity problems and variational inequalities*, in Complementarity and Variational Problems: State of the Art, M. C. Ferris and J.-S. Pang, eds., SIAM, Philadelphia, PA, 1997, pp. 91–104.

[17] S. A. Gabriel and J.-S. Pang, *A trust region method for constrained nonsmooth equations*, in Large Scale Optimization: State of the Art, W. W. Hager, D. W. Hearn, and P. M. Pardalos, eds., Kluwer Academic Publishers, Boston, MA, 1994, pp. 159–186.

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

[19] H. Jiang and L. Qi, *A new nonsmooth equations approach to nonlinear complementarity problems*, SIAM J. Control Optim., 35 (1997), pp. 178–193.

[20] C. Kanzow and M. Fukushima, *Equivalence of the generalized complementarity problem to differentiable unconstrained minimization*, J. Optim. Theory Appl., 90 (1996), pp. 581–603.

[21] C. Kanzow and M. Fukushima, *Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities*, Math. Programming, to appear.

[22] Z.-Q. Luo and P. Tseng, *A new class of merit functions for the nonlinear complementarity problem*, in Complementarity and Variational Problems: State of the Art, M. C. Ferris and J.-S. Pang, eds., SIAM, Philadelphia, PA, 1997, pp. 204–225.

[23] J. M. Martínez and A. Santos, *A trust region strategy for minimization on arbitrary domains*, Math. Programming, 68 (1995), pp. 267–302.

[24] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., 15 (1977), pp. 959–972.

[25] J. J. Moré, *Recent developments in algorithms and software for trust region methods*, in Mathematical Programming: The State of the Art, A. Bachem, M. Grötschel, and B. Korte, eds., Springer-Verlag, Berlin, 1983, pp. 258–287.

[26] J. J. Moré, *Global methods for nonlinear complementarity problems*, Math. Oper. Res., 21 (1996), pp. 589–614.

[27] J. V. Outrata and J. Zowe, *A Newton method for a class of quasi-variational inequalities*, Comput. Optim. Appl., 4 (1995), pp. 5–21.

[28] J.-S. Pang, *The implicit complementarity problem*, in Nonlinear Programming 4, O. L. Mangasarian, S. M. Robinson, and R. R. Meyer, eds., Academic Press, New York, 1981, pp. 487–518.

[29] J.-S. Pang, *Complementarity problems*, in Handbook of Global Optimization, R. Horst and P. Pardalos, eds., Kluwer Academic Publishers, Boston, MA, 1995, pp. 271–338.

[30] J.-S. Pang and S. A. Gabriel, *NE/SQP: A robust algorithm for the nonlinear complementarity problem*, Math. Programming, 60 (1993), pp. 295–337.

[31] M. J. D. Powell, *Convergence properties of a class of minimization algorithms*, in Nonlinear Programming 2, O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, eds., Academic Press, New York, 1975, pp. 1–27.

[32] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res., 18 (1993), pp. 227–244.

[33] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Math. Programming, 58 (1993), pp. 353–368.

[34] S. M. Robinson, *Generalized equations*, in Mathematical Programming: The State of the Art, A. Bachem, M. Grötschel, and B. Korte, eds., Springer-Verlag, Berlin, 1983, pp. 346–367.

[35] P. Tseng, N. Yamashita, and M. Fukushima, *Equivalence of complementarity problems to differentiable minimization: A unified approach*, SIAM J. Optim., 6 (1996), pp. 446–460.