Theoretical and Numerical Investigation of the D-Gap Function for Box Constrained Variational Inequalities

Christian Kanzow[1] and Masao Fukushima[2,*]

[1] University of Hamburg
Institute of Applied Mathematics
Bundesstrasse 55
D-20146 Hamburg
Germany
e-mail: kanzow@math.uni-hamburg.de

[2] Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma
Nara 630-01
Japan
e-mail: fuku@is.aist-nara.ac.jp

March 6, 1996

Abstract. The D-gap function, recently introduced by Peng and further studied by Yamashita et al., allows a smooth unconstrained minimization reformulation of the general variational inequality problem. This paper is concerned with the D-gap function for variational inequality problems over a box or, equivalently, mixed complementarity problems. The purpose of this paper is twofold. First we investigate theoretical properties in depth of the D-gap function, such as the optimality of stationary points, bounded level sets, global error bounds and generalized Hessians. Next we present a nonsmooth Gauss-Newton type algorithm for minimizing the D-gap function, and report extensive numerical results for the whole set of problems in the MCPLIB test problem collection.

Key words: Variational inequality problem, complementarity problem, optimization reformulation, D-gap function

Abbreviated title: The D-Gap Function for Variational Inequalities.

---

* The work of this author was supported in part by the Scientific Research Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

1. Introduction

Given a continuously differentiable function $F : \mathbb{R}^n \to \mathbb{R}^n$ and a closed convex set $S \subseteq \mathbb{R}^n$, the variational inequality problem, denoted by $\mathrm{VI}(S, F)$, is to find a vector $x^* \in S$ such that

$$F(x^*)^T(x - x^*) \geq 0 \quad \text{for all } x \in S.$$

In this paper, we focus on the box constrained variational inequality problem, i.e., we assume that S is a rectangular set described by

$$S = \prod_{i=1}^{n} [l_i, u_i],$$

where $l_i \in \mathbb{R} \cup \{-\infty\}$ and $u_i \in \mathbb{R} \cup \{+\infty\}$ represent the (possibly infinite) lower and upper bounds on the variables, respectively. Although this is a very special case of the general $\mathrm{VI}(S, F)$, this problem still has numerous important applications especially in economic equilibrium and engineering problems, see [29, 20]. Actually, the box constrained variational inequality problem can easily be shown to be equivalent to what is called the mixed complementarity problem by, e.g., Dirkse and Ferris [11], see also [5, 6, 16, 26]. Various real-world problems have already been formulated as box constrained $\mathrm{VI}(S, F)$ and many problems are currently available for testing purposes, see [11, 12].

Recently there has been active research on reformulating the variational inequality problem as an equivalent optimization problem [24]. The approach given in this paper is based on the so-called D-gap function which has been introduced very recently by Peng [44] and subsequently been generalized by Yamashita, Taji and Fukushima [55] and further analyzed in [25, 49].

In order to define the D-gap function, we first recall that the *gap function* introduced by Auslender [3] (see also [38]) is defined by

$$g(x) := \sup_{y \in S} \left\{ F(x)^T(x - y) \right\}.$$

It is not difficult to see that g is nonnegative on the feasible set S and that $g(x) = 0$ and $x \in S$ if and only if x solves $\mathrm{VI}(S, F)$. Hence the variational inequality problem $\mathrm{VI}(S, F)$ can be rewritten as the constrained minimization problem

$$\min \quad g(x) \quad \text{s.t.} \quad x \in S.$$

However, the gap function g is not differentiable in general. Moreover, when S is unbounded, g is in general not finite-valued. These drawbacks can be avoided by using the function

$$f_\alpha(x) := \max_{y \in S} \left\{ F(x)^T(x - y) - \frac{\alpha}{2} \|x - y\|^2 \right\}, \tag{1.1}$$

where α is any positive parameter. This function was independently introduced by Auchmuty [2] and Fukushima [23], and has been called the *regularized gap function*. The regularized gap function f_α is also nonnegative on the set S and x solves $\mathrm{VI}(S, F)$ if and only if $f_\alpha(x) = 0$ and $x \in S$. Hence $\mathrm{VI}(S,F)$ is again equivalent to the constrained optimization problem

$$\min \ f_\alpha(x) \quad \text{s.t.} \quad x \in S.$$

In addition, f_α is shown to be continuously differentiable. Moreover, since the function

$$\Psi_\alpha(x,y) := F(x)^T(x-y) - \frac{\alpha}{2}\|x-y\|^2 \tag{1.2}$$

is strongly concave with respect to y, there is a unique maximizer $y_\alpha(x)$ such that

$$\begin{aligned} y_\alpha(x) &= \arg\max \{\Psi_\alpha(x,y) \mid y \in S\} \\ &= \mathrm{Proj}_S\left(x - \frac{1}{\alpha}F(x)\right), \end{aligned} \tag{1.3}$$

where $\mathrm{Proj}_S(\,\cdot\,)$ denotes the projection operator onto the set S. Therefore, the regularized gap function f_α is finite-valued everywhere. The *D-gap function* introduced by Peng [44] and further investigated in [55] can now be defined as the difference (hence the "D" in D-gap function) of two regularized gap functions, namely

$$g_{\alpha\beta}(x) := f_\alpha(x) - f_\beta(x), \tag{1.4}$$

where α and β are two arbitrary parameters such that $0 < \alpha < \beta$. It can be shown that $g_{\alpha\beta}$ is nonnegative on the whole space $\mathbb{R}^n$ and that $g_{\alpha\beta}(x) = 0$ if and only if x is a solution of the variational inequality problem $\mathrm{VI}(S,F)$. Therefore $\mathrm{VI}(S,F)$ is equivalent to the unconstrained minimization problem

$$\min \ g_{\alpha\beta}(x) \quad \text{s.t.} \quad x \in \mathbb{R}^n.$$

From the corresponding properties of the regularized gap function, it follows immediately that the D-gap function is also continuously differentiable and finite-valued everywhere. Interestingly, as shown by Peng and Yuan [46], when the variational inequality problem reduces to the nonlinear complementarity problem, the D-gap function boils down to Mangasarian and Solodov's implicit Lagrangian [37], provided that the parameters α and β are chosen such that $\beta = 1/\alpha$ and $\alpha \in (0,1)$. See also [54, 36, 33, 30, 18, 45] for further studies of the implicit Lagrangian.

We note that the D-gap function is not the first function which allows an unconstrained minimization reformulation of $\mathrm{VI}(S,F)$. In fact, Yamashita and Fukushima [53] presented such reformulations. Their reformulations, however, are mainly of theoretical interest, whereas the D-gap function will turn out to be of interest also from the numerical point of view, as demonstrated in this paper. For this, however, it is important to consider a rectangular set S, because otherwise computations with the D-gap function become very expensive.

Besides the gap function approaches just reviewed, we note that there are a number of other efficient methods for the solution of problem $\mathrm{VI}(S,F)$. Most of them are based on a suitable reformulation of the Karush-Kuhn-Tucker conditions of $\mathrm{VI}(S,F)$, see, e.g., [1, 6, 14, 15, 16, 40, 41, 51, 52] for some recent contributions.

The organization of the paper is as follows. In Section 2, we collect some basic definitions and summarize some background material. In Section 3, we consider conditions under which a stationary point of the D-gap function is a global minimum and hence a solution of VI(S, F). In Section 4, we turn our attention to conditions under which the level sets of $g_{\alpha\beta}(x)$ are bounded. The fact that the D-gap function provides a global error bound for VI(S, F) under suitable assumptions will be shown in Section 5. A disadvantage of the D-gap function is that it is in general only once, rather than twice, continuously differentiable. In Section 6, we therefore give an explicit formula to calculate a generalized Hessian of $g_{\alpha\beta}(x)$. This generalized Hessian is used in Section 7 in order to design a Newton-type algorithm. It will be shown that this algorithm is globally and superlinearly convergent under suitable assumptions. Extensive numerical results for this algorithm are reported in Section 8. We conclude with some final remarks in Section 9.

For the function $F : \mathbb{R}^n \to \mathbb{R}^n$, we denote its (transposed) Jacobian at x by $\nabla F(x)$, i.e., $\nabla F(x)$ is the matrix whose (i,j)th element is $\partial F_j(x)/\partial x_i$. Each component of $F(x)$ is denoted by $F_i(x)$, and $\nabla F_i(x)$ is the ith column of $\nabla F(x)$. For a matrix $Z \in \mathbb{R}^{n\times n}$, we denote its ith column by $[Z]_i$. Thus $[\nabla F(x)]_i = \nabla F_i(x)$ for each i.

2. Mathematical Background

The regularized gap function f_α defined by (1.1) is differentiable and its gradient is represented as

$$G_\alpha(x) := \nabla f_\alpha(x) = F(x) + \nabla F(x)(x - y_\alpha(x)) - \alpha(x - y_\alpha(x)), \tag{2.1}$$

where $y_\alpha(x)$ is given by (1.3) [23]. Thus the D-gap function $g_{\alpha\beta}$ defined by (1.4) is also differentiable and its gradient is given by

$$\nabla g_{\alpha\beta}(x) = G_\alpha(x) - G_\beta(x), \tag{2.2}$$

or

$$\nabla g_{\alpha\beta}(x) = \nabla F(x)(y_\beta(x) - y_\alpha(x)) + \beta(x - y_\beta(x)) - \alpha(x - y_\alpha(x)). \tag{2.3}$$

The vector $y_\alpha(x)$, which is actually the projection of $x - \alpha^{-1}F(x)$ onto the set S, plays an important role in the subsequent discussions. For now, we state the following proposition [29, Proposition 2.3], which says that x is a solution of VI(S, F) if and only if x is a fixed point of the mapping $y_\alpha : \mathbb{R}^n \to \mathbb{R}^n$.

Proposition 2.1 *Let $\alpha > 0$ and $x \in \mathbb{R}^n$ be arbitrary and let $y_\alpha(x)$ be defined by (1.3). Then x is a solution of VI(S, F) if and only if $x = y_\alpha(x)$.*

The next proposition gives an alternative characterization of the box constrained VI(S, F). It is likely that this result is stated elsewhere, but we have found no explicit reference. So we include a short proof here.

Proposition 2.2 *A vector $x^* \in \mathbb{R}^n$ is a solution of VI(S, F) if and only if*

$$F_i(x^*)(x_i - x_i^*) \geq 0$$

for all $x \in S$ and all $i = 1, \dots, n$.

Proof. First assume that x^* solves VI(S, F), i.e.

$$F(x^*)^T(x - x^*) \geq 0 \quad \text{for all } x \in S. \tag{2.4}$$

Assume there exist a vector $x \in S$ and an index i_0 such that

$$F_{i_0}(x^*)(x_{i_0} - x_{i_0}^*) < 0.$$

Choose a vector $y \in \mathbb{R}^n$ as follows:

$$y := (x_1^*, \dots, x_{i_0-1}^*, x_{i_0}, x_{i_0+1}^*, \dots, x_n^*)^T.$$

In view of the box structure of S, we have $y \in S$. Moreover, it follows from the definition of y that

$$F(x^*)^T(y - x^*) = \sum_{i=1}^n F_i(x^*)(y_i - x_i^*) = F_{i_0}(x^*)(x_{i_0} - x_{i_0}^*) < 0,$$

a contradiction to (2.4). The proof of the converse direction is obvious. □

Corollary 2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function. If $x^* \in \mathbb{R}^n$ is a minimizer of the nonlinear program*

$$\mathit{min} \;\; f(x) \quad \mathit{s.t.} \quad x \in S, \tag{2.5}$$

then we have

$$\frac{\partial f(x^*)}{\partial x_i}(x_i - x_i^*) \geq 0$$

for all $x \in S$ and all $i = 1, \dots, n$.

Proof. Since x^* solves the program (2.5), it satisfies the optimality condition

$$\nabla f(x^*)^T(x - x^*) \geq 0 \quad \text{for all } x \in S.$$

Since this is exactly the variational inequality problem VI($S, \nabla f$), the assertion follows immediately from Proposition 2.2. □

In the rest of this section, we summarize some basic concepts which will be frequently used in this paper. We call a matrix $M \in \mathbb{R}^{n \times n}$ a *P-matrix* if its principal minors are all positive, which is known to be equivalent to the condition

$$\max_{1 \leq i \leq n} z_i(Mz)_i > 0 \quad \text{for all } z \in \mathbb{R}^n \;\, (z \neq 0).$$

We call $F : \mathbb{R}^n \to \mathbb{R}^n$ a *P-function* if

$$\max_{1\le i\le n} (x_i - y_i)(F_i(x) - F_i(y)) > 0 \quad \text{for all } x, y \in \mathbb{R}^n \ (x \neq y),$$

and a *uniform P-function* if

$$\max_{1\le i\le n} (x_i - y_i)(F_i(x) - F_i(y)) \geq \mu \|x - y\|^2 \quad \text{for all } x, y \in \mathbb{R}^n$$

for some constant $\mu > 0$. If F is an affine function $F(x) = Mx + q$, then F is a P-function if and only if its Jacobian $\nabla F(x) = M^T$ is a P-matrix. When F is nonlinear, F is a P-function if $\nabla F(x)$ is a P-matrix for every $x \in \mathbb{R}^n$, but the converse is not true in general. However, we can say that, if F is a uniform P-function, then $\nabla F(x)$ is a P-matrix for all $x \in \mathbb{R}^n$.

Let $G : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz function. By Rademacher's theorem, G is differentiable almost everywhere. Let $\Omega := \{x \in \mathbb{R}^n \mid \nabla G(x) \text{ exists}\}$. Then the *B-subdifferential* of G at x is defined by

$$\partial_B G(x) = \{ V \in \mathbb{R}^{n\times n} \mid V = \lim_{k\to\infty} \nabla G(x^k)^T, \ \ x^k \to x, \, x^k \in \Omega \}$$

and the *generalized Jacobian* of G at x is defined by

$$\partial G(x) = \operatorname{conv} \partial_B G(x),$$

where conv denotes the convex hull of a set. We say that G is *semismooth* at x if

$$\lim_{\substack{V\in\partial G(x+td') \\ d'\to d,\, t\downarrow 0}} V d'$$

exists for any $d \in \mathbb{R}^n$. If G is semismooth at x, then G is directionally differentiable at x and the directional derivative $G'(x; d)$ is equal to the above limit. For a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we say that f is an *SC*^{1} *function* if $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is semismooth. Furthermore, a continuously differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is called an *SC*^{1} *function* if each component function $F_i : \mathbb{R}^n \to \mathbb{R}$ is an SC^1 function. We refer the reader to [48, 47] for some basic properties of semismooth functions. For a recent survey of methods for semismooth equations and SC^1 optimization problems, see [32].

The following result, due to Facchinei [13], says that for SC^1 functions, the full stepsize of 1 will be accepted by an Armijo line search rule under some reasonable assumptions.

Proposition 2.3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be an SC^1 function and suppose that we have a sequence $\{x^k\}$ converging to a stationary point x^* of f as well as a sequence of directions $\{d^k\}$. Assume that*

$$\lim_{k\to\infty} \frac{\|x^k + d^k - x^*\|}{\|x^k - x^*\|} = 0$$

and that, for all k sufficiently large and some $\rho > 0$,

$$\nabla f(x^k)^T d^k \leq -\rho \|d^k\|^2.$$

Then, for any $\sigma \in (0, 1/2)$, there is an index k_0 such that

$$f(x^k + d^k) \leq f(x^k) + \sigma \nabla f(x^k)^T d^k$$

for all $k \geq k_0$.

3. Stationary Points

Now let us return to the box constrained VI(S, F) which is to find a vector $x^* \in S$ such that

$$F(x^*)^T(x - x^*) \geq 0 \qquad \text{for all } x \in S,$$

where $S = \prod_{i=1}^n [l_i, u_i]$ with $l_i \leq u_i$ for $i = 1, \ldots, n$. In this section, we give conditions under which a stationary point of the D-gap function $g_{\alpha\beta}$ is a solution of the original problem VI(S, F). This result is important because most existing iterative methods for solving minimization problems are only guaranteed to converge to a stationary point.

Let $x \in \mathbb{R}^n$ be arbitrary and let $y_\alpha(x)$ and $y_\beta(x)$ be defined as in (1.3). Then, by the box structure of the set S, the vectors $y_\gamma(x)$, $\gamma = \alpha, \beta$, can be written componentwise as follows:

$$\begin{aligned} y_{\gamma,i}(x) &= \mathrm{Proj}_{[l_i,u_i]}\left(x_i - \frac{1}{\gamma}F_i(x)\right) \\ &= \mathrm{mid}\left\{x_i - \frac{1}{\gamma}F_i(x), l_i, u_i\right\} \qquad i = 1, \ldots, n, \end{aligned} \tag{3.1}$$

where $y_{\gamma,i}(x)$ is the ith component of the vector $y_\gamma(x)$ and mid $\{a, b, c\}$ denotes the median of three (possibly infinite) numbers a, b, c, i.e., if $b \leq c$, then

$$\mathrm{mid}\,\{a, b, c\} = \begin{cases} b & \text{if } a < b, \\ a & \text{if } b \leq a \leq c, \\ c & \text{if } c < a. \end{cases}$$

The main result in this section is now stated.

Theorem 3.1 *Let $x^* \in \mathbb{R}^n$ be a stationary point of the D-gap function $g_{\alpha\beta}$ and suppose that $\nabla F(x^*)$ is a P-matrix. Then x^* is a solution of VI(S, F).*

Proof. To simplify the notation, we denote

$$\begin{aligned} p &= x^* - \tfrac{1}{\alpha}F(x^*), \\ q &= x^* - \tfrac{1}{\beta}F(x^*), \\ p^+ &= y_\alpha(x^*), \\ q^+ &= y_\beta(x^*). \end{aligned} \tag{3.2}$$

Note that, by (2.3) and (3.2), we can write

$$\nabla g_{\alpha\beta}(x^*) = \nabla F(x^*)(q^+ - p^+) + \beta(q - q^+) - \alpha(p - p^+).$$

Since x^* is a stationary point of $g_{\alpha\beta}$, i.e.,

$$\nabla g_{\alpha\beta}(x^*) = 0, \tag{3.3}$$

we have

$$\nabla F(x^*)(q^+ - p^+) + \beta(q - q^+) - \alpha(p - p^+) = 0. \tag{3.4}$$

Now assume that $q^+ - p^+ \neq 0$. Then, since $\nabla F(x^*)$ is a P-matrix, there is an index i such that

$$(q_i^+ - p_i^+)\left(\nabla F(x^*)(q^+ - p^+)\right)_i > 0. \tag{3.5}$$

Premultiplying the ith component of (3.4) by $(q_i^+ - p_i^+)$ yields

$$(q_i^+ - p_i^+)\left(\nabla F(x^*)(q^+ - p^+)\right)_i + \beta(q_i^+ - p_i^+)(q_i - q_i^+) - \alpha(q_i^+ - p_i^+)(p_i - p_i^+) = 0. \tag{3.6}$$

Since the projection theorem implies

$$(q_i^+ - p_i^+)(q_i - q_i^+) \geq 0 \tag{3.7}$$

and

$$(p_i^+ - q_i^+)(p_i - p_i^+) \geq 0, \tag{3.8}$$

the last two terms of (3.6) are nonnegative. However, by (3.5), the first term of (3.6) is positive. Hence the equality (3.6) is a contradiction. Therefore we must have

$$q^+ - p^+ = y_\beta(x^*) - y_\alpha(x^*) = 0. \tag{3.9}$$

Now in view of the original formula (2.3) for $\nabla g_{\alpha\beta}(x)$, we can deduce from (3.3) and (3.9) that

$$(\beta - \alpha)(x^* - y_\alpha(x^*)) = 0,$$

i.e., $x^* = y_\alpha(x^*)$. By Proposition 2.1, x^* is a solution of the VI(S, F). □

From this theorem, we can readily deduce that, if F is a uniform P-function, then any stationary point of $g_{\alpha\beta}$ is a solution of VI(S, F). Moreover, it is the unique solution of VI(S, F).

4. Bounded Level Sets

In this section, we study the global behavior of the D-gap function. In particular, we are interested in conditions under which the D-gap function has bounded level sets. We will establish two results here: The first result requires the uniform P-property of the function F, while the second requires the boundedness of the set S. The second result, in particular, improves upon the recent result for the general VI by Yamashita et al. [55], who proved that the D-gap function has bounded level sets if F is strongly monotone and S is bounded.

To begin with, define the function $\varphi_{\alpha\beta} : \mathbb{R}^2 \to \mathbb{R}$ by

$$\begin{aligned}\varphi_{\alpha\beta}(a, b; l, u) \;:=\; & b\left(\mathrm{mid}\{a - \frac{1}{\beta}b, l, u\} - \mathrm{mid}\{a - \frac{1}{\alpha}b, l, u\}\right) \\ & + \frac{\beta}{2}\left(a - \mathrm{mid}\{a - \frac{1}{\beta}b, l, u\}\right)^2 - \frac{\alpha}{2}\left(a - \mathrm{mid}\{a - \frac{1}{\alpha}b, l, u\}\right)^2,\end{aligned}$$

where α and β are, of course, given real parameters such that $0 < \alpha < \beta$. Here, we temporarily let l and u be two possibly infinite real parameters, i.e., $l \in \mathbb{R} \cup \{-\infty\}$ and $u \in \mathbb{R} \cup \{+\infty\}$. As shown in Lemma 4.3 below, the function $\varphi_{\alpha\beta}$ plays a fundamental role in forming the D-gap function. We summarize some important properties of this function in the following lemmas.

Lemma 4.1 *We have $\varphi_{\alpha\beta}(a, b; l, u) \geq 0$ for all $a, b \in \mathbb{R}$.*

Proof. Note that the D-gap function $g_{\alpha\beta}$ for $\mathrm{VI}(S, F)$ with $n = 1$ reduces to the function $\varphi_{\alpha\beta}$ with the identification $(x, F(x)) \leftrightarrow (a, b)$. Then the desired result follows from the nonnegativity of the D-gap function. □

Lemma 4.2 *Let $\{a^k\}, \{b^k\} \subseteq \mathbb{R}$ be two sequences such that $|a^k| \to \infty$ and $|b^k| \to \infty$. Then the sequence $\{\varphi_{\alpha\beta}(a^k, b^k; l, u)\}$ is unbounded.*

Proof. We consider several cases.
Case 1. The bounds l and u are both infinite, i.e., $l = -\infty$ and $u = \infty$.
Then $\mathrm{mid}\{a^k - \frac{1}{\beta}b^k, l, u\} = a^k - \frac{1}{\beta}b^k$ and $\mathrm{mid}\{a^k - \frac{1}{\alpha}b^k, l, u\} = a^k - \frac{1}{\alpha}b^k$. We therefore get after some algebraic manipulations

$$\varphi_{\alpha\beta}(a^k, b^k; l, u) = \frac{\beta - \alpha}{2\alpha\beta}(b^k)^2.$$

Since $\{b^k\}$ is unbounded and $0 < \alpha < \beta$, it follows that $\{\varphi_{\alpha\beta}(a^k, b^k; l, u)\}$ is also unbounded.
Case 2. The bound l is finite and the bound u is infinite, i.e., $u = +\infty$.
For simplicity of notation, we assume without loss of generality that $l = 0$. Then we can write

$$\begin{aligned} \varphi_{\alpha\beta}(a, b; 0, u) &= b\left(\max\{a - \frac{1}{\beta}b, 0\} - \max\{a - \frac{1}{\alpha}b, 0\}\right) \\ &\quad + \frac{\beta}{2}\left(a - \max\{a - \frac{1}{\beta}b, 0\}\right)^2 - \frac{\alpha}{2}\left(a - \max\{a - \frac{1}{\alpha}b, 0\}\right)^2. \end{aligned}$$

Taking subsequences if necessary, we may consider the following four subcases separately.
Case 2.1. $\max\{a^k - \frac{1}{\beta}b^k, 0\} = 0$ and $\max\{a^k - \frac{1}{\alpha}b^k, 0\} = 0$ for all k.
A simple calculation shows that

$$\varphi_{\alpha\beta}(a^k, b^k; 0, u) = \frac{\beta - \alpha}{2}(a^k)^2.$$

Hence the sequence $\{\varphi_{\alpha\beta}(a^k, b^k; 0, u)\}$ is unbounded by the unboundedness of the sequence $\{a^k\}$ and the fact that $\beta > \alpha$.
Case 2.2. $\max\{a^k - \frac{1}{\beta}b^k, 0\} = a^k - \frac{1}{\beta}b^k$ and $\max\{a^k - \frac{1}{\alpha}b^k, 0\} = a^k - \frac{1}{\alpha}b^k$ for all k.
In this case, the argument is exactly the same as in Case 1.

Case 2.3. $\max\{a^k - \frac{1}{\beta}b^k, 0\} = a^k - \frac{1}{\beta}b^k$ and $\max\{a^k - \frac{1}{\alpha}b^k, 0\} = 0$ for all k.
After some algebraic manipulations, we get

$$\begin{aligned} \varphi_{\alpha\beta}(a^k, b^k; 0, u) &= a^k b^k - \frac{1}{2\beta}(b^k)^2 - \frac{\alpha}{2}(a^k)^2 \\ &= \frac{b^k}{2}\left(a^k - \frac{1}{\beta}b^k\right) + \frac{\alpha a^k}{2}\left(\frac{1}{\alpha}b^k - a^k\right). \end{aligned} \tag{4.1}$$

We want to show that the right-hand side is unbounded. To this end, first note that in the subcase under consideration we have $\frac{1}{\alpha}b^k \geq a^k \geq \frac{1}{\beta}b^k$ and $a^k - \frac{1}{\beta}b^k \geq 0 \geq a^k - \frac{1}{\alpha}b^k$, the latter being equivalent to $b^k \geq 0$. Hence we have $b^k \to \infty$ and therefore also $a^k \to \infty$. Now assume that $\gamma^k := a^k - \frac{1}{\beta}b^k \to 0$ and $\delta^k := \frac{1}{\alpha}b^k - a^k \to 0$ on a subsequence. We then get $\gamma^k + \delta^k = \frac{\beta-\alpha}{\alpha\beta}b^k \to 0$, i.e., $b^k \to 0$, a contradiction. Therefore, since $\gamma^k \geq 0$ and $\delta^k \geq 0$, there exists a constant $c > 0$ such that $a^k - \frac{1}{\beta}b^k \geq c$ for all k or $\frac{1}{\alpha}b^k - a^k \geq c$ for all k. In view of (4.1), this implies the unboundedness of the sequence $\{\varphi_{\alpha\beta}(a^k, b^k; 0, u)\}$.
Case 2.4. $\max\{a^k - \frac{1}{\beta}b^k, 0\} = 0$ and $\max\{a^k - \frac{1}{\alpha}b^k, 0\} = a^k - \frac{1}{\alpha}b^k$ for all k.
In this case, we get

$$\begin{aligned} \varphi_{\alpha\beta}(a^k, b^k; 0, u) &= \frac{1}{2\alpha}(b^k)^2 - a^k b^k + \frac{\beta}{2}(a^k)^2 \\ &= \frac{b^k}{2}\left(\frac{1}{\alpha}b^k - a^k\right) + \frac{\beta a^k}{2}\left(a^k - \frac{1}{\beta}b^k\right) \end{aligned}$$

with $\frac{1}{\alpha}b^k - a^k \leq 0$ and $a^k - \frac{1}{\beta}b^k \leq 0$. By a similar reasoning to Subcase 2.3, we can deduce that $b^k \to -\infty$ and $a^k \to -\infty$. Moreover we can also show that there exists a constant $c > 0$ such that $\frac{1}{\alpha}b^k - a^k \leq -c$ for all k or $a^k - \frac{1}{\beta}b^k \leq -c$ for all k. So it follows that the sequence $\{\varphi_{\alpha\beta}(a^k, b^k; 0, u)\}$ is unbounded.
Case 3. The bound l is infinite, i.e., $l = -\infty$ and the bound u is finite.
Without loss of generality, we may assume that $u = 0$. Then the function $\varphi_{\alpha\beta}$ reduces to

$$\begin{aligned} \varphi_{\alpha\beta}(a, b; l, 0) = \; & b\left(\min\{a - \frac{1}{\beta}b, 0\} - \min\{a - \frac{1}{\alpha}b, 0\}\right) \\ & + \frac{\beta}{2}\left(a - \min\{a - \frac{1}{\beta}b, 0\}\right)^2 - \frac{\alpha}{2}\left(a - \min\{a - \frac{1}{\alpha}b, 0\}\right)^2. \end{aligned}$$

Similar to Case 2, we may consider four subcases. Since the arguments in these subcases are very similar to the ones in Case 2, we omit the details here.
Case 4. The two bounds l and u are finite.
In this case, the sequences $\{\text{mid}\{a^k - \frac{1}{\beta}b^k, l, u\}\}$ and $\{\text{mid}\{a^k - \frac{1}{\beta}b^k, l, u\}\}$ remain bounded. Then it is easy to see from the definition of $\varphi_{\alpha\beta}$ that the sequence $\{\varphi_{\alpha\beta}(a^k, b^k; l, u)\}$ is unbounded. This completes the proof. □

Now let l and u again denote the n-dimensional vectors (with possibly infinite components) which constitute the lower and upper bounds of the feasible set S.

Lemma 4.3 *We have* $g_{\alpha\beta}(x) = \sum_{i=1}^n \varphi_{\alpha\beta}(x_i, F_i(x); l_i, u_i)$ *for all* $x \in \mathbb{R}^n$.

Proof. This follows immediately from the very definitions of the functions $g_{\alpha\beta}$ and $\varphi_{\alpha\beta}$. □

We are now in a position to prove the first result regarding the boundedness of level sets of $g_{\alpha\beta}$. The proof of the theorem is similar to a proof by Geiger and Kanzow [27] as well as Facchinei and Soares [19] in the context of complementarity problems using the squared Fischer function, see [22]. Notice that this theorem does not rely upon the differentiability of F.

Theorem 4.1 *Assume that* $F : \mathbb{R}^n \to \mathbb{R}^n$ *is a uniform* P*-function. Then the level sets*

$$\mathcal{L}(c) := \{x \in \mathbb{R}^n \,|\, g_{\alpha\beta}(x) \leq c\}$$

are bounded for all $c \geq 0$.

Proof. Let $c \geq 0$. Assume that there is a sequence $\{x^k\} \subseteq \mathcal{L}(c)$ such that $\|x^k\| \to \infty$. Define the index set $J \subseteq \{1, 2, \ldots, n\}$ by

$$J := \{i \,|\, \{x_i^k\} \text{ is unbounded}\}.$$

By assumption, we then have $J \neq \emptyset$. We now define a second sequence $\{y^k\} \subseteq \mathbb{R}^n$ as follows:

$$y_i^k := \begin{cases} 0 & \text{if} \quad i \in J, \\ x_i^k & \text{if} \quad i \notin J. \end{cases}$$

From the definition of $\{y^k\}$ and the fact that F is a uniform P-function, we obtain for some constant $\mu > 0$

$$\begin{aligned} \mu \sum_{i \in J} (x_i^k)^2 &= \mu \|x^k - y^k\|^2 \\ &\leq \max_{1 \leq i \leq n} (x_i^k - y_i^k)(F_i(x^k) - F_i(y^k)) \\ &= \max_{i \in J} x_i^k (F_i(x^k) - F_i(y^k)) \\ &\leq \sqrt{\sum_{i \in J} (x_i^k)^2} \sum_{i \in J} |F_i(x^k) - F_i(y^k)|, \end{aligned}$$

from which we get

$$\mu \sqrt{\sum_{i \in J} (x_i^k)^2} \leq \sum_{i \in J} |F_i(x^k) - F_i(y^k)|. \tag{4.2}$$

By definition, the sequence $\{y^k\}$ remains bounded. From the continuity of F, it follows that the sequence $\{F_i(y^k)\}$ is also bounded for every $i = 1, \ldots, n$. Hence we deduce from (4.2) that there is at least one index $i_0 \in J$ such that

$$|F_{i_0}(x^k)| \to \infty$$

on a subsequence. On the other hand, since $i_0 \in J$, we also have

$$|x_{i_0}^k| \to \infty,$$

which implies, by Lemma 4.2, that the sequence $\{\varphi_{\alpha\beta}(x_{i_0}^k, F_{i_0}(x^k); l_{i_0}, u_{i_0})\}$ is unbounded. By Lemma 4.1 and Lemma 4.3, however, we then have

$$g_{\alpha\beta}(x^k) \to \infty,$$

thus contradicting the assumption that $\{x^k\} \subseteq \mathcal{L}(c)$. □

As an immediate consequence from Theorem 3.1 and Theorem 4.1, we reobtain the following existence result [29, Theorem 3.9] for the box constrained VI(S, F). For completeness, we shall provide a brief proof.

Corollary 4.1 *Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ is a uniform P-function. Then VI(S, F) has a unique solution.*

Proof. Since $g_{\alpha\beta}$ is continuous, it is easily deduced from Theorem 4.1 and Weierstrass' theorem that $g_{\alpha\beta}$ attains its infimum on $\mathbb{R}^n$. Since such a minimizer x^* is also a stationary point of $g_{\alpha\beta}$ and the Jacobian $\nabla F(x^*)$ is a P-matrix, it follows from Theorem 3.1 that it is a solution of VI(S, F). The uniqueness of a solution can easily be proved using Proposition 2.2, so we omit the proof. □

In the proof of this corollary, we used the differentiability of $g_{\alpha\beta}$ and F. However the existence of a solution can be proved without the differentiability of F; what we need are the continuity and the uniform P-property of F [29, Theorem 3.9]. Also note that, if we assume the existence of a solution, then the uniqueness of the solution is ensured under the assumption that F is a P-function.

Now we turn to establishing the second result on the boundedness of level sets of the D-gap function. As mentioned at the beginning of this section, the next theorem improves upon the recent result by Yamashita et al. [55] (see also the comment after the theorem).

Theorem 4.2 *If S is bounded, then the level sets $\mathcal{L}(c)$ of the D-gap function $g_{\alpha\beta}$ are bounded for all $c \geq 0$.*

Proof. Suppose that there exists an unbounded sequence $\{x^k\}$ in the level set $\mathcal{L}(c)$. Consider the corresponding sequence $y_\beta(x^k)$. Since $y_\beta(x^k) \in S$ for all k and since S is bounded by assumption, the sequence $\{y_\beta(x^k)\}$ remains bounded. So we have $\|x^k - y_\beta(x^k)\| \to \infty$. However, by [55, Proposition 3.1], we have

$$g_{\alpha\beta}(x^k) \geq \frac{\beta - \alpha}{2}\|x^k - y_\beta(x^k)\|^2$$

for all k. Since $\{x^k\} \subseteq \mathcal{L}(c)$, we obtain the contradiction

$$c \geq g_{\alpha\beta}(x^k) \geq \frac{\beta - \alpha}{2}\|x^k - y_\beta(x^k)\|^2 \to \infty.$$

This completes the proof. □

Note that this theorem remains true even if F is not assumed to be continuous. (Of course, when F is not continuous, $g_{\alpha\beta}$ is in general discontinuous and hence the level sets $\mathcal{L}(c)$ need not be compact.) Also note that this theorem does not depend upon the box structure of S. Yamashita et al. [55] have proved for a general VI(S, F), where S need not be a box, that if F is strongly monotone and if either F is Lipschitz continuous or S is bounded, then the D-gap function $g_{\alpha\beta}$ provides a global error bound for VI(S, F) and hence has bounded level sets. More recently, Kanzow et al. [35] have shown that a class of merit functions for the nonlinear complementarity problem have bounded level sets if F is monotone and the problem is strictly feasible. For other merit functions in variational inequalities to provide global error bounds and, as a consequence, to have bounded level sets, it has been generally assumed that F is both strongly monotone and Lipschitz continuous [24]. In this regard, Theorem 4.2 appears to be the first result that does not require any condition on F to guarantee the boundedness of level sets of a merit function for variational inequalities.

5. Global Error Bounds

In this section, we show that the function $\sqrt{g_{\alpha\beta}}$ provides a global error bound for VI(S, F) under suitable assumptions. The following lemma, motivated by Proposition 3.5 in [50], is the main ingredient for our global error bound result.

Lemma 5.1 *Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ is a uniform P-function. Further suppose that either F is Lipschitz continuous or S is bounded. Let x^* be the unique solution of VI(S, F). Then there exists a constant $\hat{c} > 0$ such that*

$$\|x - x^*\| \leq \hat{c}\,\|x - y_\beta(x)\|$$

holds for all $x \in \mathbb{R}^n$.

Proof. First recall that, by Corollary 4.1, VI(S, F) has a unique solution under the given assumptions. By definition, for any $x \in \mathbb{R}^n$, $y_\beta(x)$ is the unique solution of the problem

$$\min \quad -\Psi_\beta(x, y) \quad \text{s.t.} \quad y \in S,$$

where Ψ_β is defined by (1.2) with α replaced by β. Hence $y_\beta(x)$ satisfies, for any $x \in \mathbb{R}^n$, the following optimality condition:

$$-\nabla_y \Psi_\beta(x, y_\beta(x))^T (u - y_\beta(x)) = (F(x) - \beta(x - y_\beta(x)))^T\,(u - y_\beta(x)) \geq 0$$

for all $u \in S$. In view of Corollary 2.1, this implies, for any $x \in \mathbb{R}^n$

$$(F_i(x) - \beta(x_i - y_{\beta,i}(x)))\,(u_i - y_{\beta,i}(x)) \geq 0$$

for all $u \in S$ and $i = 1, \dots, n$, where $y_{\beta,i}(x)$ denotes the ith element of the vector $y_\beta(x)$. In particular, putting $u = x^* \in S$, we have

$$(F_i(x) - \beta(x_i - y_{\beta,i}(x)))\,(x_i^* - y_{\beta,i}(x)) \ge 0 \tag{5.1}$$

for all $x \in \mathbb{R}^n$ and $i = 1, \dots, n$. On the other hand, since x^* solves VI(S, F), we get from Proposition 2.2

$$F_i(x^*)(u_i - x_i^*) \ge 0$$

for all $u \in S$ and $i = 1, \dots, n$. Since $y_\beta(x) \in S$ for any $x \in \mathbb{R}^n$, we therefore have

$$F_i(x^*)(y_{\beta,i}(x) - x_i^*) \ge 0 \tag{5.2}$$

for all $x \in \mathbb{R}^n$ and $i = 1, \dots, n$. Adding the two inequalities (5.1) and (5.2), we obtain

$$(F_i(x) - F_i(x^*) - \beta(x_i - y_{\beta,i}(x)))\,(x_i^* - y_{\beta,i}(x)) \ge 0$$

for all $x \in \mathbb{R}^n$ and $i = 1, \dots, n$. This is equivalent to

$$\begin{array}{rcl} (F_i(x) - F_i(x^*) - \beta(x_i - y_{\beta,i}(x)))\,(x_i^* - x_i) & & \\ +\,(F_i(x) - F_i(x^*) - \beta(x_i - y_{\beta,i}(x)))\,(x_i - y_{\beta,i}(x)) & \ge & 0 \end{array}$$

for all $x \in \mathbb{R}^n$ and $i = 1, \dots, n$. From this, it follows that

$$\begin{array}{rcl} (F_i(x) - F_i(x^*))\,(x_i - x_i^*) & \le & -\beta\,(x_i - y_{\beta,i}(x))\,(x_i^* - x_i) \\ & & +\,(F_i(x) - F_i(x^*))\,(x_i - y_{\beta,i}(x)) \\ & & -\beta\,(x_i - y_{\beta,i}(x))^2 \\ & \le & -\beta\,(x_i - y_{\beta,i}(x))\,(x_i^* - x_i) \\ & & +\,(F_i(x) - F_i(x^*))\,(x_i - y_{\beta,i}(x)) \end{array}$$

for all $x \in \mathbb{R}^n$ and $i = 1, \dots, n$.

Now suppose that F is Lipschitz continuous with constant $L > 0$. Let $\mu > 0$ be the modulus of uniform P-property for F. Then we have

$$\begin{array}{rcl} \mu\|x - x^*\|^2 & \le & \max\limits_{1 \le i \le n} (F_i(x) - F_i(x^*))\,(x_i - x_i^*) \\ & \le & \max\limits_{1 \le i \le n} \{-\beta\,(x_i - y_{\beta,i}(x))\,(x_i^* - x_i) + (F_i(x) - F_i(x^*))\,(x_i - y_{\beta,i}(x))\} \\ & \le & \beta\|x - y_\beta(x)\|\,\|x - x^*\| + L\|x - x^*\|\,\|x - y_\beta(x)\| \end{array}$$

for all $x \in \mathbb{R}^n$. Dividing both sides by $\|x - x^*\|$, we get

$$\mu\|x - x^*\| \le (L + \beta)\|x - y_\beta(x)\| \tag{5.3}$$

for all $x \in \mathbb{R}^n$. Setting $\hat{c} = (L + \beta)/\mu$ yields the desired inequality.

The case where S is bounded can be proved in a way similar to the proof of [55, Lemma 4.1 (b)]. For completeness, we give an outline of the proof. Let $b > 0$ be a constant such that $S \subseteq \{x \mid \|x\| \le b\}$, and denote $\mathcal{B} := \{x \mid \|x\| \le 3b\}$. Choose $x \in \mathbb{R}^n$ arbitrarily. We may consider two cases (i) $x \in \mathcal{B}$ and (ii) $x \notin \mathcal{B}$. In case (i), since F is Lipschitz continuous on the sphere $\mathcal{B}$, we obtain the inequality (5.3) for such x, by using the arguments in the previous paragraph. In case (ii), since $\|x^*\| \le b, \|y_\beta(x)\| \le b$ and $\|x\| > 3b$, we have $\|x - x^*\| \le 2\|x - y_\beta(x)\|$. Thus, putting the two cases together shows the existence of a constant $\hat{c} > 0$ satisfying the desired inequality for all $x \in \mathbb{R}^n$. □

We are now ready to state the main result of this section.

Theorem 5.1 *Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ is a uniform P-function. Further suppose that either F is Lipschitz continuous or S is bounded. Let x^* be the unique solution of VI(S, F). Then there is a constant $c > 0$ such that*

$$\|x - x^*\| \le c\sqrt{g_{\alpha\beta}(x)}$$

holds for all $x \in \mathbb{R}^n$.

Proof. By [55, Proposition 3.1], we have

$$g_{\alpha\beta}(x) \ge \frac{\beta - \alpha}{2}\|x - y_\beta(x)\|^2$$

for all $x \in \mathbb{R}^n$. Combining this with the inequality stated in Lemma 5.1 yields the desired result. □

6. Generalized Hessian

In this section, we are concerned with calculating a generalized Hessian of the D-gap function $g_{\alpha\beta}$, i.e., a generalized Jacobian of $\nabla g_{\alpha\beta}$. The results obtained in this section will turn out to be essential in designing a Newton-type algorithm to minimize the D-gap function $g_{\alpha\beta}$.

In this and the next sections, unless otherwise stated, we assume that the function $F : \mathbb{R}^n \to \mathbb{R}^n$ is twice continuously differentiable. We shall denote $G := \nabla g_{\alpha\beta}$ and $G_\gamma := \nabla f_\gamma$ for $\gamma = \alpha, \beta$. (The subscript γ will frequently be used to unify various expressions associated with the two parameters α and β.) By the definition (1.4) of the D-gap function, the gradient of $g_{\alpha\beta}$ is given by

$$G(x) = G_\alpha(x) - G_\beta(x),$$

where $G_\gamma = \nabla f_\gamma$ are represented as (2.1). By the box structure of S, it is easily seen from (3.1) that the functions $y_\gamma : \mathbb{R}^n \to \mathbb{R}^n$ are piecewise smooth. This implies that the gradient mappings $G_\gamma, \gamma = \alpha, \beta$, and hence G are also piecewise smooth. Note that any piecewise smooth function is semismooth. In particular, by direct calculation from the formula (2.1) for ∇f_γ, the generalized Jacobian of G_γ is given by

$$\begin{aligned}\partial G_\gamma(x) = \{V \in \mathbb{R}^{n\times n} \mid V = {} & \nabla F(x)^T + \nabla^2 F(x)(x - y_\gamma(x)) \\ & - \gamma Z_\gamma(x)(I - W_\gamma),\ W_\gamma \in \partial y_\gamma(x)\},\end{aligned}$$

where

$$\nabla^2 F(x)(x - y_\gamma(x)) = \sum_{i=1}^{n} \nabla^2 F_i(x)(x_i - y_{\gamma,i}(x)),$$

$\partial y_\gamma(x)$ is the generalized Jacobian of y_γ at x, and the matrix $Z_\gamma(x)$ is defined by

$$Z_\gamma(x) := I - \frac{1}{\gamma}\nabla F(x). \tag{6.1}$$

Let $[Z_\gamma(x)]_i$ denote the ith column of the matrix $Z_\gamma(x)$, i.e.,

$$[Z_\gamma(x)]_i = e^i - \frac{1}{\gamma}\nabla F_i(x),$$

where e^i is the ith column of the $n \times n$ identity matrix I.

Define the index sets $I_\gamma^s(x) \subseteq \{1, 2, \dots, n\}$ for $s = 1, 2, \dots, 5$ and $\gamma = \alpha, \beta$ as follows:

$$\begin{array}{rcl} I_\gamma^1(x) & := & \{i \mid x_i - \frac{1}{\gamma}F_i(x) < l_i\}, \\ I_\gamma^2(x) & := & \{i \mid x_i - \frac{1}{\gamma}F_i(x) = l_i\}, \\ I_\gamma^3(x) & := & \{i \mid l_i < x_i - \frac{1}{\gamma}F_i(x) < u_i\}, \\ I_\gamma^4(x) & := & \{i \mid x_i - \frac{1}{\gamma}F_i(x) = u_i\}, \\ I_\gamma^5(x) & := & \{i \mid x_i - \frac{1}{\gamma}F_i(x) > u_i\}. \end{array} \tag{6.2}$$

Now consider an arbitrary point $\bar{x} \in \mathbb{R}^n$. Let us derive an explicit formula for a generalized Hessian of $g_{\alpha\beta}$ at $\bar{x}$. To this end, the following regularity assumption plays an important role:

Assumption (A): The vectors $[Z_\alpha(\bar{x})]_i$ $(i \in I_\alpha^2(\bar{x}) \cup I_\alpha^4(\bar{x}))$ and $[Z_\beta(\bar{x})]_i$ $(i \in I_\beta^2(\bar{x}) \cup I_\beta^4(\bar{x}))$ are linearly independent.

Note that Assumption (A) requires the linear independence of the vectors $[Z_\gamma(\bar{x})]_i$ corresponding to the degenerate indices only. Therefore this assumption may be considered fairly mild, since the number of degenerate indices is usually small relative to n.

Let I_γ^-, I_γ^0, and I_γ^+ be arbitrary index sets such that

$$\begin{array}{rcl} I_\gamma^- & \subseteq & I_\gamma^1(\bar{x}) \cup I_\gamma^2(\bar{x}), \\ I_\gamma^0 & \subseteq & I_\gamma^2(\bar{x}) \cup I_\gamma^3(\bar{x}) \cup I_\gamma^4(\bar{x}), \\ I_\gamma^+ & \subseteq & I_\gamma^4(\bar{x}) \cup I_\gamma^5(\bar{x}), \end{array} \tag{6.3}$$

and

$$I_\gamma^- \cup I_\gamma^0 \cup I_\gamma^+ = \{1, 2, \dots, n\}, \quad I_\gamma^- \cap I_\gamma^0 = \emptyset, \quad I_\gamma^0 \cap I_\gamma^+ = \emptyset, \tag{6.4}$$

for $\gamma = \alpha, \beta$. Then, under Assumption (A), we can choose a sequence $\{x^k\}$ such that $x^k \to \bar{x}$ and

$$\begin{array}{ll} x_i^k - \frac{1}{\gamma}F_i(x^k) < l_i & \text{if} \quad i \in I_\gamma^-, \\ l_i < x_i^k - \frac{1}{\gamma}F_i(x^k) < u_i & \text{if} \quad i \in I_\gamma^0, \\ x_i^k - \frac{1}{\gamma}F_i(x^k) > u_i & \text{if} \quad i \in I_\gamma^+, \end{array} \tag{6.5}$$

for all k and $\gamma = \alpha, \beta$. Clearly, we have

$$\{x^k\} \subseteq \Omega_\alpha \cap \Omega_\beta \subseteq \Omega$$

and

$$\nabla G(x^k) = \nabla G_\alpha(x^k) - \nabla G_\beta(x^k) \qquad \text{for all} \quad k,$$

where

$$\begin{array}{rcl} \Omega & = & \{x \in \mathbb{R}^n \mid \nabla G(x) \text{ exists}\}, \\ \Omega_\gamma & = & \{x \in \mathbb{R}^n \mid \nabla G_\gamma(x) \text{ exists}\} \quad \text{for } \gamma = \alpha, \beta. \end{array}$$

Note that $\nabla G_\gamma(x^k)$ can be written as

$$\nabla G_\gamma(x^k)^T = \nabla F(x^k)^T + \nabla^2 F(x^k)(x^k - y_\gamma(x^k)) - \gamma Z_\gamma(x^k)(I - \nabla y_\gamma(x^k)^T),$$

and the ith column of $\nabla y_\gamma(x^k)$ is given by

$$\left[\nabla y_\gamma(x^k)\right]_i = \begin{cases} e^i - \frac{1}{\gamma}\nabla F_i(x^k) & \text{if } \; i \in I_\gamma^0, \\ 0 & \text{if } \; i \in I_\gamma^- \cup I_\gamma^+. \end{cases}$$

Since $\{x^k\}$ converges to $\bar{x}$, we have

$$\nabla y_\gamma(x^k)^T \to W_\gamma \in \partial_B y_\gamma(\bar{x}) \subseteq \partial y_\gamma(\bar{x})$$

and

$$\nabla G_\gamma(x^k)^T \to V_\gamma \in \partial_B G_\gamma(\bar{x}) \subseteq \partial G_\gamma(\bar{x})$$

for $\gamma = \alpha, \beta$, where W_γ is the matrix such that the ith column of W_γ^T is given by

$$\left[W_\gamma^T\right]_i = \begin{cases} e^i - \frac{1}{\gamma}\nabla F_i(\bar{x}) & \text{if } \; i \in I_\gamma^0, \\ 0 & \text{if } \; i \in I_\gamma^- \cup I_\gamma^+ \end{cases} \tag{6.6}$$

and V_γ is the matrix given by

$$V_\gamma = \nabla F(\bar{x})^T + \nabla^2 F(\bar{x})(\bar{x} - y_\gamma(\bar{x})) - \gamma Z_\gamma(\bar{x})(I - W_\gamma). \tag{6.7}$$

Therefore we have

$$\nabla G(x^k)^T = \nabla G_\alpha(x^k)^T + \nabla G_\beta(x^k)^T \;\to\; V_\alpha - V_\beta \in \partial_B G(\bar{x}) \subseteq \partial G(\bar{x}). \tag{6.8}$$

Consequently, under Assumption (A), any partitions $(I_\alpha^-, I_\alpha^0, I_\alpha^+)$ and $(I_\beta^-, I_\beta^0, I_\beta^+)$ of the index set $\{1, 2, \ldots, n\}$ satisfying (6.3) and (6.4) yield an element of the generalized Jacobian of G. More specifically, by (6.7) and (6.8), the matrix

$$\begin{array}{rcl} V & := & V_\alpha - V_\beta \\ & = & \nabla^2 F(\bar{x})(y_\beta(\bar{x}) - y_\alpha(\bar{x})) + \beta Z_\beta(\bar{x})(I - W_\beta) - \alpha Z_\alpha(\bar{x})(I - W_\alpha) \end{array} \tag{6.9}$$

with W_γ determined by (6.6), is an element of the generalized Jacobian of G, i.e., the generalized Hessian of $g_{\alpha\beta}$, at $\bar{x}$. We shall denote by $\mathcal{V}(\bar{x})$ the set of all matrices $V \in \mathbb{R}^{n \times n}$ which can be represented as (6.9). The above arguments indicate that $\mathcal{V}(\bar{x}) \subseteq \partial_B G(\bar{x})$. More generally, we can state the following lemma, which will be useful subsequently.

Lemma 6.1 *Let $\bar{x} \in \mathbb{R}^n$ be an arbitrary vector satisfying Assumption (A). Then there is a $\delta > 0$ such that*

$$\mathcal{V}(x) \subseteq \partial_B G(x)$$

for all $x \in \mathbb{R}^n$ such that $\|x - \bar{x}\| \leq \delta$.

Proof. The inclusion for $x = \bar{x}$ holds as pointed out in the paragraph preceding the lemma. Moreover, it is easy to see that Assumption (A) remains to hold in a neighborhood of $\bar{x}$, from which the desired result follows immediately. □

Note that, without Assumption (A), there may not exist a sequence $\{x^k\}$ satisfying (6.5) for some partitions $(I_\alpha^-, I_\alpha^0, I_\alpha^+)$ and $(I_\beta^-, I_\beta^0, I_\beta^+)$. Thus the matrix $V_\alpha - V_\beta$ determined from such partitions is not guaranteed to be a member of $\partial_B G(\bar{x})$ any more.

Now let us proceed to study conditions for nonsingularity of the generalized Hessian of $g_{\alpha\beta}$ at a solution x^* of VI(S, F). Define the following five index sets:

$$\begin{aligned}
I^1(x^*) &:= \{i \mid x_i^* = l_i, F_i(x^*) > 0\}, \\
I^2(x^*) &:= \{i \mid x_i^* = l_i, F_i(x^*) = 0\}, \\
I^3(x^*) &:= \{i \mid l_i < x_i^* < u_i, F_i(x^*) = 0\}, \\
I^4(x^*) &:= \{i \mid x_i^* = u_i, F_i(x^*) = 0\}, \\
I^5(x^*) &:= \{i \mid x_i^* = u_i, F_i(x^*) < 0\}.
\end{aligned}$$

Obviously, since x^* is a solution of VI(S, F), these five index sets form a partitioning of the set $\{1, \ldots, n\}$. Note that the indices in $I^2(x^*)$ and $I^4(x^*)$ are the degenerate indices. Moreover, we have

$$I^i(x^*) = I_\alpha^i(x^*) = I_\beta^i(x^*) \qquad \text{for } i = 1, 2, \ldots, 5, \tag{6.10}$$

with $I_\alpha^i(x^*)$ and $I_\beta^i(x^*)$ being defined by (6.2). Therefore Assumption (A) at a solution x^* may be restated as follows: The vectors $[Z_\alpha(x^*)]_i$ $(i \in I^2(x^*) \cup I^4(x^*))$ and $[Z_\beta(x^*)]_i$ $(i \in I^2(x^*) \cup I^4(x^*))$ are linearly independent. A sufficient condition for Assumption (A) to hold at x^* is given in the following proposition.

Proposition 6.1 *Let $x^* \in \mathbb{R}^n$ be a solution of VI(S, F). If the vectors e^i $(i \in I^2(x^*) \cup I^4(x^*))$ and $\nabla F_i(x^*)$ $(i \in I^2(x^*) \cup I^4(x^*))$ are linearly independent, then Assumption (A) is satisfied at x^*.*

Proof. In view of the definition (6.1) of $Z_\gamma(x)$, we show that

$$\begin{aligned}
0 &= \sum_{i \in I^2(x^*) \cup I^4(x^*)} \{\lambda_i [Z_\alpha(x^*)]_i + \mu_i [Z_\beta(x^*)]_i\} \\
&= \sum_{i \in I^2(x^*) \cup I^4(x^*)} \left\{ (\lambda_i + \mu_i) e^i - \left(\frac{\lambda_i}{\alpha} + \frac{\mu_i}{\beta} \right) \nabla F_i(x^*) \right\}
\end{aligned} \tag{6.11}$$

holds only if $\lambda_i = \mu_i = 0$ for all $i \in I^2(x^*) \cup I^4(x^*)$. Under the given linear independence assumption, we have for all $i \in I^2(x^*) \cup I^4(x^*)$

$$\lambda_i + \mu_i = 0, \quad \frac{\lambda_i}{\alpha} + \frac{\mu_i}{\beta} = 0,$$

from which it readily follows that $\lambda_i = \mu_i = 0$. □

Let $\hat{\mathcal{V}}(x^*)$ be the subset of $\mathcal{V}(x^*)$ which consists of matrices V given by (6.9) with W_α^* and W_β^* determined from (6.6) using particular partitions of $\{1, \ldots, n\}$ such that $(I_\alpha^-, I_\alpha^0, I_\alpha^+) = (I_\beta^-, I_\beta^0, I_\beta^+)$. It follows from (6.10) that, for any $V \in \hat{\mathcal{V}}(x^*)$, the corresponding partition (I^-, I^0, I^+) satisfies not only

$$(I^-, I^0, I^+) = (I_\alpha^-, I_\alpha^0, I_\alpha^+) = (I_\beta^-, I_\beta^0, I_\beta^+)$$

but also

$$\begin{array}{rcccl} I^1(x^*) & \subseteq & I^- & \subseteq & I^1(x^*) \cup I^2(x^*), \\ I^3(x^*) & \subseteq & I^0 & \subseteq & I^2(x^*) \cup I^3(x^*) \cup I^4(x^*), \\ I^5(x^*) & \subseteq & I^+ & \subseteq & I^4(x^*) \cup I^5(x^*). \end{array} \tag{6.12}$$

Now we make the following assumption, which is crucial to guarantee the nonsingularity of generalized Hessians $V \in \hat{\mathcal{V}}(x^*) \subseteq \partial_B G(x^*)$.

Assumption (B): For any partition (I^-, I^0, I^+) of $\{1, \ldots, n\}$ satisfying (6.12), the vectors e^i $(i \in I^- \cup I^+)$ and $\nabla F_i(x^*)$ $(i \in I^0)$ are linearly independent.

This assumption might appear to have some connection with Assumption (A). However, Assumption (A) is only concerned with nondegenerate indices, while Assumption (B) is related to the entire index set $\{1, \ldots, n\}$. Moreover Assumption (A) is related to an arbitrary point in $\mathbb{R}^n$, while Assumption (B) is only concerned with a solution of VI(S, F).

The following theorem gives conditions under which each matrix V chosen from the set $\hat{\mathcal{V}}(x^*)$ is a nonsingular matrix.

Theorem 6.1 *Let $x^* \in \mathbb{R}^n$ be a solution of VI(S, F) satisfying Assumption (B). Then each matrix $V \in \hat{\mathcal{V}}(x^*)$ is nonsingular.*

Proof. Let V be an arbitrary element of $\hat{\mathcal{V}}(x^*)$. Then the matrix V can be written as

$$V = \beta Z_\beta(x^*)(I - W_\beta^*) - \alpha Z_\alpha(x^*)(I - W_\alpha^*), \tag{6.13}$$

where W_α^* and W_β^* are formed by (6.6) with $\bar{x}$ being replaced by x^*, for some partitioning (I^-, I^0, I^+) of $\{1, \ldots, n\}$ such that $(I^-, I^0, I^+) = (I_\alpha^-, I_\alpha^0, I_\alpha^+) = (I_\beta^-, I_\beta^0, I_\beta^+)$. Note that the fact $y_\alpha(x^*) = y_\beta(x^*)$ deduced from Proposition 2.1 has been used to derive (6.13) from (6.9).

To prove the nonsingularity of V, it suffices to show that $q = 0$ is the only solution of the equation

$$V^T q = 0. \tag{6.14}$$

Note that, by the definition (6.1) of $Z_\gamma(x)$, we have

$$\gamma Z_\gamma(x^*)(I - W_\gamma^*) = \gamma(I - W_\gamma^*) - \nabla F(x^*) + \nabla F(x^*) W_\gamma^*.$$

So it follows from (6.13) and (6.14) that

$$0 = V^T q = \beta(I - (W_\beta^*)^T)q + (W_\beta^*)^T p - \alpha(I - (W_\alpha^*)^T)q - (W_\alpha^*)^T p, \tag{6.15}$$

where we set $p := \nabla F(x^*)^T q$. Now, taking into account the definition (6.6) of the matrices $W_\gamma^*, \gamma = \alpha, \beta$ and the fact that we can write

$$(W_\gamma^*)^T q = \sum_{i=1}^n [(W_\gamma^*)^T]_i q_i,$$

we obtain

$$\begin{aligned} \gamma(I - (W_\gamma^*)^T)q + (W_\gamma^*)^T p &= \sum_{i \in I^0} \left(q_i \nabla F_i(x^*) + p_i e^i \right) \\ &\quad + \gamma \sum_{i \in I^- \cup I^+} q_i e^i - \frac{1}{\gamma} \sum_{i \in I^0} p_i \nabla F_i(x^*) \end{aligned}$$

for $\gamma = \alpha, \beta$. Therefore from (6.15), we get

$$(\beta - \alpha) \sum_{i \in I^- \cup I^+} q_i e^i + \left(\frac{1}{\alpha} - \frac{1}{\beta} \right) \sum_{i \in I^0} p_i \nabla F_i(x^*) = 0.$$

In view of Assumption (B), this yields

$$\begin{aligned} (\beta - \alpha)\, q_i &= 0 \qquad \text{for all} \quad i \in I^- \cup I^+, \\ \left(\frac{1}{\alpha} - \frac{1}{\beta} \right) p_i &= 0 \qquad \text{for all} \quad i \in I^0. \end{aligned}$$

Since $p = \nabla F(x^*)^T q$ and $\alpha \neq \beta$, this can be rewritten as

$$\begin{aligned} q^T e^i &= 0 \qquad \text{for all} \quad i \in I^- \cup I^+, \\ q^T \nabla F_i(x^*) &= 0 \qquad \text{for all} \quad i \in I^0. \end{aligned}$$

Using Assumption (B) once again, we obtain $q = 0$. We conclude that the matrix V is nonsingular. □

In the nondegenerate case where $I^2(x^*) = I^4(x^*) = \emptyset$, Assumption (A) is vacuously satisfied at x^*. In fact, it is seen that G is differentiable at x^* and its Jacobian is represented as

$$\nabla G(x^*) = \beta Z_\beta(x^*)(I - \nabla y_\beta(x^*)) - \alpha Z_\alpha(x^*)(I - \nabla y_\alpha(x^*)),$$

where the matrices $\nabla y_\gamma(x^*)$ are given by

$$[\nabla y_\gamma(x^*)]_i = \begin{cases} e^i - \frac{1}{\gamma} \nabla F_i(x^*) & \text{if} \quad i \in I^3(x^*), \\ 0 & \text{if} \quad i \in I^1(x^*) \cup I^5(x^*), \end{cases}$$

for $\gamma = \alpha, \beta$. The sets $\partial G(x^*), \partial_B G(x^*), \mathcal{V}(x^*)$ and $\hat{\mathcal{V}}(x^*)$ all reduce to the singleton $\{\nabla G(x^*)^T\}$. Moreover, Assumption (B) is simplified as follows: The vectors e^i $(i \in I^1(x^*) \cup I^5(x^*))$ and $\nabla F_i(x^*)$ $(i \in I^3(x^*))$ are linearly independent.

With these observations, we obtain the following corollary to Theorem 6.1. Note that, in the particular case where VI(S, F) reduces to NCP, the D-gap function (with $\beta = 1/\alpha$ and $\alpha \in (0, 1)$) reduces to the implicit Lagrangian by Mangasarian and Solodov [37], so this corollary generalizes a corresponding result in [37].

Corollary 6.1 *Let $x^* \in \mathbb{R}^n$ be a nondegenerate solution of VI(S, F). Suppose that the vectors e^i $(i \in I^1(x^*) \cup I^5(x^*))$ and $\nabla F_i(x^*)$ $(i \in I^3(x^*))$ are linearly independent. Then the D-gap function $g_{\alpha\beta}$ is twice differentiable at x^* and the Hessian $\nabla^2 g_{\alpha\beta}(x^*)$ $(= \nabla G(x^*))$ is positive definite.*

Proof. The nonsingularity of $\nabla G(x^*)$ follows from Theorem 6.1. To see the positive definiteness, it suffices to observe that a solution of VI(S, F) minimizes $g_{\alpha\beta}$ and hence the Hessian is at least positive semidefinite. □

7. Truncated Gauss-Newton-Type Method

In this section, we present a truncated Gauss-Newton-type method for the solution of box constrained variational inequality problems, which is based on the unconstrained minimization of the D-gap function $g_{\alpha\beta}$. The method is called truncated because we solve the linear systems which occur at each iteration only inexactly. It is also called Gauss-Newton-type since we do not use matrices $V \in \mathcal{V}(x)$; instead we deal with matrices $\tilde{V}$ belonging to the following set:

$$\tilde{\mathcal{V}}(x) := \left\{ \tilde{V} \in \mathbb{R}^{n \times n} \,\middle|\, \tilde{V} = V - \nabla^2 F(x)(y_\beta(x) - y_\alpha(x)) \;\text{ for any } V \in \mathcal{V}(x) \right\},$$

i.e., the matrices in $\tilde{\mathcal{V}}(x)$ are identical to those in $\mathcal{V}(x)$ except that we cancel out the term containing the second derivatives of F. From a practical point of view, this is very important because second order information about the function F is often not available (like in the two test problem collections MCPLIB and GAMSLIB [11, 12]) or difficult to obtain. In [49], a similar Gauss-Newton strategy has also been adopted in a trust region algorithm for minimizing the D-gap function. Note that, if F is an affine function, then there is no difference between the two sets $\mathcal{V}(x)$ and $\tilde{\mathcal{V}}(x)$. We also note that if x is a solution of VI(S, F), then these two sets also coincide, cf. Proposition 2.1.

Based on the set $\tilde{\mathcal{V}}(x)$, we now state the Gauss-Newton-type algorithm. Recall that $G(x)$ denotes the gradient $\nabla g_{\alpha\beta}(x)$ of the D-gap function at a point x.

Algorithm 7.1 *(Truncated Gauss-Newton-type algorithm)*

(S.0) *Choose $x^0 \in \mathbb{R}^n, \rho > 0, \kappa > 2, \sigma \in (0, 1/2), \delta \in (0, 1), \bar{\eta} \in (0, 1), \varepsilon_{\text{TOL}} \geq 0, \eta_0 \in [0, \bar{\eta}]$ and set $k := 0$.*

(S.1) *If $\|G(x^k)\| \leq \varepsilon_{\text{TOL}}$, stop.*

(S.2) *Select a matrix $\tilde{V}_k \in \tilde{\mathcal{V}}(x^k)$. Find a vector d^k such that*

$$\tilde{V}_k d^k = -G(x^k) + r^k \tag{7.1}$$

with $\|r^k\| \leq \eta_k \|G(x^k)\|$. If such d^k does not exist or does not satisfy the descent condition

$$G(x^k)^T d^k \leq -\rho \|d^k\|^\kappa, \tag{7.2}$$

then set $d^k := -G(x^k)$.

(S.3) Let t_k be the largest number in the set $\{1, \delta, \delta^2, \delta^3, \ldots\}$ such that

$$g_{\alpha\beta}(x^k + t_k d^k) \le g_{\alpha\beta}(x^k) + t_k \sigma G(x^k)^T d^k.$$

(S.4) Choose $\eta_{k+1} \in [0, \bar{\eta}]$, set $x^{k+1} := x^k + t_k d^k, k := k+1$, and go to step (S.1).

In order to prove global and local convergence results for Algorithm 7.1, we need some lemmas. In those lemmas, we will assume that a solution x^* of VI(S, F) satisfies Assumption (A). Note that, as pointed out in the paragraph preceding Proposition 6.1, this assumption is restated as: The vectors $[Z_\alpha(x^*)]_i$ $(i \in I^2(x^*) \cup I^4(x^*))$ and $[Z_\beta(x^*)]_i$ $(i \in I^2(x^*) \cup I^4(x^*))$ are linearly independent. Proposition 6.1 gives a sufficient condition for this assumption to hold. We also note that, in this and the preceding sections, we assume the twice continuous differentiability of F.

Lemma 7.1 *The D-gap function $g_{\alpha\beta}$ is an SC^1 function.*

Proof. Since F is assumed to be a C^2 function, F is obviously an SC^1 function. Since the sum and the composite of semismooth functions are again semismooth and since a piecewise smooth function is semismooth, it follows that the gradient G is semismooth. Hence $g_{\alpha\beta}$ itself is an SC^1 function. □

In the following, we say that x^* is a *BD-regular* solution of VI(S, F) if all matrices in the B-subdifferential $\partial_B G(x^*)$ are nonsingular.

Lemma 7.2 *Let $x^* \in \mathbb{R}^n$ be a BD-regular solution of VI(S, F) satisfying Assumption (A). Then there exist a $\delta > 0$ and a constant $\tilde{\mu} > 0$ such that $\tilde{V}$ is nonsingular and*

$$\|\tilde{V}^{-1}\| \le \tilde{\mu}$$

for all $\tilde{V} \in \tilde{\mathcal{V}}(x)$ and all $x \in \mathbb{R}^n$ such that $\|x - x^*\| \le \delta$.

Proof. By definition, for any $\tilde{V} \in \tilde{\mathcal{V}}(x)$, there is a corresponding matrix $V \in \mathcal{V}(x)$ such that

$$\tilde{V} = V - \nabla^2 F(x)(y_\beta(x) - y_\alpha(x)). \tag{7.3}$$

Note that, by Proposition 2.1, we have $y_\alpha(x^*) = y_\beta(x^*) = x^*$. Moreover, in view of (1.3) and the continuity of the projection operator, we see from (7.3) that the matrix $\tilde{V}$ is arbitrarily close to V, whenever x is sufficiently close to x^*. The assertion is therefore an immediate consequence of, e.g., Theorem 3.1.4 in [9], if we can show that there exist constants $\delta > 0$ and $\mu > 0$ such that V is nonsingular and $\|V^{-1}\| \le \mu$ for all $V \in \mathcal{V}(x)$ and all $x \in \mathbb{R}^n$ with $\|x - x^*\| \le \delta$. In view of the assumed BD-regularity of the solution x^*, this indeed follows from Lemma 2.6 in [47] and the inclusion $\mathcal{V}(x) \subseteq \partial_B G(x)$ deduced from Lemma 6.1 with $\bar{x} = x^*$. □

Lemma 7.3 *Let $x^* \in \mathbb{R}^n$ be a solution of VI(S, F) satisfying Assumption (A). If all matrices in $\partial_B G(x^*)$ are positive definite, then there exist a $\delta > 0$ and a constant $\tilde{\lambda} > 0$ such that*

$$d^T \tilde{V} d \ge \tilde{\lambda} \|d\|^2$$

for all $\tilde{V} \in \tilde{\mathcal{V}}(x)$ and all $x \in \mathbb{R}^n$ such that $\|x - x^*\| \le \delta$.

Proof. As shown by Jiang and Qi [31], the given positive definiteness assumption guarantees the existence of $\delta > 0$ and $\lambda > 0$ such that $d^T V d \geq \lambda \|d\|^2$ for all $V \in \partial_B G(x)$ and all $x \in \mathbb{R}^n$ with $\|x - x^*\| \leq \delta$. Then the lemma can be proved in a way similar to Lemma 7.2. $\square$

Lemma 7.4 *Let $x^* \in \mathbb{R}^n$ be a solution of VI(S, F) satisfying Assumption (A). Then we have*

$$G(x) - \tilde{V}(x - x^*) = o(\|x - x^*\|)$$

for any $\tilde{V} \in \tilde{\mathcal{V}}(x)$.

Proof. For any $\tilde{V} \in \tilde{\mathcal{V}}(x)$, let V be the corresponding matrix in the set $\mathcal{V}(x)$ such that (7.3) holds. We first note that

$$G(x) - \tilde{V}(x - x^*) = G(x) - V(x - x^*) + (V - \tilde{V})(x - x^*). \tag{7.4}$$

Since G is semismooth and, by Lemma 6.1, $V \in \mathcal{V}(x) \subseteq \partial_B G(x)$ for any x sufficiently close to x^*, we have

$$G(x) - V(x - x^*) = o(\|x - x^*\|) \tag{7.5}$$

by Proposition 1 in [43]. Furthermore, since F is continuously differentiable, there exists a Lipschitz-constant $L > 0$ such that

$$|F_i(x) - F_i(x^*)| \leq L\|x - x^*\|$$

for all $i = 1, \ldots, n$ and for all x sufficiently close to x^*. Taking into account the nonexpansiveness of the projection operator, we therefore obtain

$$\begin{aligned} |y_{\alpha,i}(x) - y_{\alpha,i}(x^*)| &= \left|\mathrm{Proj}_{[l_i,u_i]}\left(x_i - \frac{1}{\alpha}F_i(x)\right) - \mathrm{Proj}_{[l_i,u_i]}\left(x_i^* - \frac{1}{\alpha}F_i(x^*)\right)\right| \\ &\leq \left|x_i - x_i^* + \frac{1}{\alpha}(F_i(x^*) - F_i(x))\right| \\ &\leq |x_i - x_i^*| + \frac{1}{\alpha}|F_i(x) - F_i(x^*)| \\ &\leq \left(1 + \frac{L}{\alpha}\right)\|x - x^*\| \end{aligned} \tag{7.6}$$

and similarly

$$|y_{\beta,i}(x) - y_{\beta,i}(x^*)| \leq \left(1 + \frac{L}{\beta}\right)\|x - x^*\| \leq \left(1 + \frac{L}{\alpha}\right)\|x - x^*\| \tag{7.7}$$

for all $i = 1, \ldots, n$ and all x in a small neighborhood of x^*. Moreover, since F is twice continuously differentiable, there exists a constant $c > 0$ such that

$$\|\nabla^2 F_i(x)\| \leq c \tag{7.8}$$

for all $i = 1, \ldots, n$, and all x in a neighborhood of x^*. From (7.6)–(7.8) and Proposition 2.1, we now obtain

$$\begin{aligned} \|V - \tilde{V}\| &= \left\| \sum_{i=1}^{n} \nabla^2 F_i(x)(y_{\beta,i}(x) - y_{\alpha,i}(x)) \right\| \\ &\leq c \sum_{i=1}^{n} |y_{\beta,i}(x) - y_{\alpha,i}(x)| \\ &\leq c \sum_{i=1}^{n} \{ |y_{\beta,i}(x) - y_{\beta,i}(x^*)| + |y_{\alpha,i}(x) - y_{\alpha,i}(x^*)| \} \\ &\leq 2nc \left(1 + \frac{L}{\alpha}\right) \|x - x^*\|. \end{aligned} \tag{7.9}$$

Combining (7.4), (7.5) and (7.9), we get

$$G(x) - \tilde{V}(x - x^*) = o(\|x - x^*\|)$$

for any $\tilde{V} \in \tilde{\mathcal{V}}(x)$. □

Lemma 7.5 *Let $x^* \in \mathbb{R}^n$ be a solution of VI(S, F) satisfying Assumption (A). Let $\{x^k\} \subseteq \mathbb{R}^n$ be any sequence converging to x^*, and let $\tilde{V}_k \in \tilde{\mathcal{V}}(x^k)$ be chosen arbitrarily for each k. Then the sequence $\{\tilde{V}_k\}$ remains bounded.*

Proof. For any $\tilde{V}_k \in \tilde{\mathcal{V}}(x^k)$, let V_k be the corresponding matrix in $\mathcal{V}(x^k)$ such that (7.3) holds. By Lemma 6.1, we have $\mathcal{V}(x^k) \subseteq \partial_B G(x^k) \subseteq \partial G(x^k)$ for all k sufficiently large. Since $\partial G(x^*)$ is a compact set and the point-to-set mapping ∂G is upper semicontinuous (see Clarke [7]), it follows that the sequence $\{V_k\}$ remains bounded. Since F is twice continuously differentiable, it therefore follows from (7.3) that the sequence $\{\tilde{V}_k\}$ is also bounded. □

In the following theorem, we present local and global convergence results for the proposed truncated Gauss-Newton-type method. For simplicity of notation, we assume that the termination parameter ε_{TOL} in Algorithm 7.1 is equal to 0, and that an infinite sequence is generated.

Theorem 7.1 *Let $\{x^k\} \subseteq \mathbb{R}^n$ be any sequence generated by Algorithm 7.1. Then the following statements hold:*

(a) Each accumulation point of the sequence $\{x^k\}$ is a stationary point of the D-gap function $g_{\alpha\beta}$.

(b) If an accumulation point, say x^*, of the sequence $\{x^k\}$ is an isolated solution of VI(S, F) satisfying Assumption (A), then the whole sequence $\{x^k\}$ converges to x^*.

(c) If an accumulation point, say x^*, of the sequence $\{x^k\}$ is a solution of VI(S, F) satisfying Assumption (A), all matrices in $\partial_B G(x^*)$ are positive definite, and $\eta_k \to 0$, then

(i) the whole sequence $\{x^k\}$ converges to x^*;

(ii) d^k is the truncated Gauss-Newton direction for all k large enough (i.e., the negative gradient is never used eventually);

(iii) the rate of convergence is Q-superlinear.

Proof. Parts (a) and (b) of Theorem 7.1 can be verified in a similar way as the corresponding parts of Theorem 3.1 in [8], see also [18, 14].

In order to prove part (c), we first note that all matrices in $\partial_B G(x^*)$ being positive definite implies that x^* is a BD-regular solution of VI(S,F). Hence, by Lemma 7.1 and [43, Proposition 3], it follows that x^* is a locally unique solution of the system $G(x)=0$. Since every solution of VI(S,F) is necessarily a global minimum and hence a stationary point of $g_{\alpha\beta}$, it follows that x^* is an isolated solution of VI(S,F). Hence the whole sequence $\{x^k\}$ converges to x^* by part (b).

We next show that the search direction d^k is the truncated Gauss-Newton direction for all k sufficiently large. To this end, we first note that it is always possible to find a vector d^k satisfying (7.1) with

$$\|r^k\| \le \eta_k \|G(x^k)\|. \tag{7.10}$$

This follows immediately from Lemma 7.2 and the fact that x^* is a BD-regular solution. We now want to show that the truncated Gauss-Newton direction d^k eventually satisfies the descent condition (7.2) for an arbitrary $\rho > 0$. This can be seen as follows: Let $m>0$ and $M>0$ be constants such that

$$m\|d^k\| \le \|\tilde{V}_k d^k\| \le M\|d^k\| \tag{7.11}$$

for all k sufficiently large. The existence of such a constant m follows from Lemma 7.2, since the latter lemma implies $\|d^k\| = \|\tilde{V}_k^{-1}\tilde{V}_k d^k\| \le \|\tilde{V}_k^{-1}\|\,\|\tilde{V}_k d^k\| \le \tilde{\mu}\|\tilde{V}_k d^k\|$ for some $\tilde{\mu} > 0$. On the other hand, the existence of the constant $M>0$ directly follows from Lemma 7.5. From (7.11) and (7.1), we now obtain

$$m\|d^k\| \le \|\tilde{V}_k d^k\| = \|G(x^k) - r^k\| \le M\|d^k\|,$$

so that

$$\frac{m}{2}\|d^k\| \le \|G(x^k)\| \le 2M\|d^k\| \tag{7.12}$$

for all k sufficiently large by (7.10) and the assumption that $\eta_k \to 0$. Moreover, it follows from Lemma 7.3 that there is a constant $\tilde{\lambda} > 0$ such that

$$(d^k)^T \tilde{V}_k d^k \ge \tilde{\lambda}\|d^k\|^2 \tag{7.13}$$

for all k sufficiently large. Hence we obtain from (7.1), (7.10), (7.12) and (7.13) that

$$\begin{aligned}
G(x^k)^T d^k &= -(d^k)^T \tilde{V}_k d^k + (r^k)^T d^k \\
&\le -\tilde{\lambda}\|d^k\|^2 + \|r^k\|\,\|d^k\| \\
&= -\tilde{\lambda}\|d^k\|^2 + o(\|d^k\|^2) \\
&\le -\frac{\tilde{\lambda}}{2}\|d^k\|^2
\end{aligned}$$

for all k large enough, where the second equality follows from $\eta_k \to 0$. Since $\kappa > 2$ and $d^k \to 0$ by (7.12), it therefore follows that the descent condition (7.2) is eventually satisfied for arbitrary $\rho > 0$.

The next step of the proof is to show that the stepsize t_k computed by the Armijo-rule in step (S.3) of Algorithm 7.1 is equal to 1 for all k sufficiently large. To this end, let $\{x^k\}$ be any sequence converging to x^* and let d^k denote the truncated Gauss-Newton direction satisfying (7.1) with (7.10). Since $\eta_k \to 0$, it follows from (7.10), Lemma 7.4 and Lemma 7.5 that

$$\|r^k\| = o(\|G(x^k)\|) = o(\|x^k - x^*\|).$$

Hence we obtain from (7.1), Lemma 7.4 and Lemma 7.2

$$\begin{aligned} \|x^k + d^k - x^*\| &\leq \|\tilde{V}_k^{-1}\| \, \|G(x^k) - \tilde{V}_k(x^k - x^*) - r^k\| \\ &\leq \tilde{\mu} \left(\|G(x^k) - \tilde{V}_k(x^k - x^*)\| + \|r^k\| \right) \\ &= o(\|x^k - x^*\|), \end{aligned} \tag{7.14}$$

where $\tilde{\mu} > 0$ is a constant as specified in Lemma 7.2. Since $g_{\alpha\beta}$ is an SC^1 function by Lemma 7.1, it follows from Proposition 2.3 that we eventually have $t_k = 1$. This means that we have $x^{k+1} = x^k + d^k$ for all k large enough. Thus the superlinear convergence follows immediately from (7.14). □

Note that a sufficient condition for a stationary point of $g_{\alpha\beta}$ to be a solution of VI(S, F) was given in Section 3. Furthermore, we note that the positive definiteness assumption used in part (c) of Theorem 7.1 is satisfied, e.g., if x^* is a nondegenerate solution of VI(S, F) such that Assumption (B) holds, see Theorem 6.1. (Note that Assumption (B) itself simplifies to some extent at a nondegenerate solution.)

We admit that the positive definiteness assumption in Theorem 7.1 is usually stronger than, e.g., the BD-regularity condition. However, if $g_{\alpha\beta}$ is a convex function, then its gradient mapping G is monotone and hence all matrices in $\partial_B G(x^*)$ are positive semidefinite, see [31, Proposition 2.3]. Moreover, under the BD-regularity condition, each matrix in $\partial_B G(x^*)$ is, as a limit of symmetric Hessian matrices, necessarily symmetric and nonsingular, and hence is positive definite. Unfortunately, the convexity of $g_{\alpha\beta}$ is in general hard to verify. Peng [45] gives a condition that guarantees the convexity of the function $g_{\alpha\beta}$ for the particular case where VI(S, F) reduces to the nonlinear complementarity problem so that $g_{\alpha\beta}$ becomes the implicit Lagrangian for a suitable choice of the parameters α and β. Finally we remark that the matrices in $\partial_B G(x^*)$ are not necessarily positive semidefinite even if x^* is a minimum of $g_{\alpha\beta}$. A second-order necessary condition as well as a second-order sufficient condition for unconstrained SC^1 minimization may be found in [13].

8. Numerical Results

We implemented our algorithm in MATLAB and tested it on a SUN SPARC station using all test problems and all different starting points which are available for the MATLAB-version of MCPLIB, see [12].

Some details about our implementation: The algorithm terminates at iteration k if

$$g^s_{\alpha\beta}(x^k) := g_{\alpha\beta}(x^k)/n < \varepsilon_1 \text{ or } \|G^s(x^k)\| := \|G(x^k)\|/\sqrt{n} < \varepsilon_2 \text{ or } k > k_{\max},$$

i.e., we use scaled function and gradient values in order to make the accuracy of the termination criterion more independent of the dimension of the problem. The parameters used in the implementation are as follows:

$$\varepsilon_1 = 10^{-9}, \varepsilon_2 = 10^{-5}, \rho = 10^{-8}, \kappa = 2.1, \sigma = 10^{-4}, \delta = 10^{-1}, k_{\max} = 300, \alpha = 0.9, \beta = 1.1.$$

In step (S.2) of our algorithm we solve the generalized Newton equation

$$\tilde{V}_k d = -G(x^k)$$

exactly if an estimated condition number of $\tilde{V}_k$ is of reasonable size; otherwise we use an iterative solver in such a way that (7.1) holds with

$$\eta_k := \min\{\gamma_1/(k+1), \|G(x^k)\|^{\gamma_2}\},$$

where

$$\gamma_1 = 10^{-2} \text{ and } \gamma_2 = 1.5.$$

Although the matrices $\tilde{V}_k$ are in general only symmetric and not positive definite, we follow the recommodation in the book [4] and use the CG-algorithm as iterative solver (we also made some test runs with the CG-method being replaced by the MINRES-method, but the results were usually not as good as for the CG-method).

We also replaced the monotone line search in step (S.3) of Algorithm 7.1 by the following nonmonotone Armijo-rule [28]: Calculate $t_k = \max\{1, \delta, \delta^2, \delta^3, \ldots\}$ such that

$$g_{\alpha\beta}(x^k + t_k d^k) \leq \mathcal{R}_k + \sigma t_k G(x^k)^T d^k,$$

where

$$\mathcal{R}_k := \max\{g_{\alpha\beta}(x^j) \,|\, j = k+1-\ell_k, \ldots, k\}$$

and ℓ_k denotes a positive integer which is adjusted at each iteration in the following way:

1. Set $\ell_k = 1$ for $k = 0, 1, 2, 3, 4$, i.e., start the iteration process with a monotone Armijo rule during the first five iterations.
2. Set $\ell_k = 1$ whenever $d^k = -G(x^k)$ at iteration k, i.e., use a monotone line search whenever a gradient step has been taken.
3. Let $\ell_{k+1} = \min\{\ell_k + 1, \ell\}$ at all remaining iterations k, where ℓ is a prespecified constant ($\ell = 10$ in our implementation).

Finally, we note that our algorithm assumes F to be well-defined everywhere, whereas in many (if not most) examples of MCPLIB the function F may not be defined outside of S or even on the boundary of S. Hence we included the following heuristic in our implementation: If x^k is the current iterate and d^k denotes the calculated search direction, we first compute a backtracking factor $\tilde{t}_k := \max\{1, 0.1, 0.01, \ldots\}$ such that $F(x^k + \tilde{t}_k d^k)$

is well-defined. In the following line search procedure, we then first try the steplength $t_k := \tilde{t}_k$ and reduce it further if necessary in order to satisfy the nonmonotone Armijo rule. We terminate our iteration, however, if

$$t_k < t_{\min}$$

with $t_{\min} = 10^{-9}$ in our implementation.

The results we have obtained with this algorithm are summarized in Table 1. For each example, we report the name of the test problem in MCPLIB, the dimension n of the test problem, the number m of box constraints, the starting vector chosen, the number of iterations (it), the number of F-evaluations (F-eval.), the final value of $g_{\alpha\beta}(x)/n$, the final value of $\|G(x)\|/\sqrt{n}$ and the number of Newton (N), truncated Newton (TN) and gradient steps (G) taken.

As can be seen from Table 1, the algorithm was able to solve most problems without any difficulties. Here we only give some more details on the failures: For problems colvdual and colvnlp (first starting vectors), the algorithm converged to a stationary point of the D-gap function which is not a solution of the underlying VI(S, F)-problem. Problem ehl_kost, on the other hand, seems to be very difficult for our algorithm: The iteration terminated after 300 iterations with a value of $g^s_{\alpha\beta}(x)$ of approximately 10^{-3}. Also for problems pgvon105 and pgvon106 the algorithm terminated after the maximum number of iterations. The final value of $g^s_{\alpha\beta}(x)$ for problem pgvon106, however, was around 10^{-8}, so that this problem may be regarded as solved essentially. The failures on the problem scarfbnum, on the other hand, were due to line search failures: The steplength t_k became too small during the iteration.

9. Final Remarks

In this paper, we investigated the D-gap function by Peng [44] and Yamashita et al. [55] when applied to box constrained variational inequalities. We proved several new results regarding stationary points, bounded level sets and global error bounds. We also reported extensive numerical results of a Gauss-Newton type method which is based on the D-gap function. Since the D-gap function is an extension of Mangasarian and Solodov's [37] implicit Lagrangian, these results may also be viewed as the first extensive numerical testing of this implicit Lagrangian.

There are a number of interesting questions which remain to be answered. For example, is it possible to show that the D-gap function is convex under certain assumptions? Such a result would have some influence to our local convergence theory, see the comments at the end of Section 7. Moreover, in view of the piecewise smoothness of the D-gap function, is it possible to show that our algorithm (when solving the linear systems exactly) terminates after finitely many iterations if F is an affine mapping? We hope that a positive answer to these questions will be found in the not-too-far future.

Acknowledgement. The authors gratefully acknowledge some helpful comments of Dr. Andreas Fischer regarding the implementation of the algorithm.

Table 1: Numerical results on the test problems from MCPLIB

problem	n	m	S.P.	it	F-eval.	$g^s_{\alpha\beta}(x)$	$\|G^s(x)\|$	N	TN	G
bertsekas	15	15	1	1	2	1.2e-3	7.7e-18	0	1	0
bertsekas	15	15	2	2	3	1.2e-3	1.8e-17	1	1	0
bertsekas	15	15	3	2	3	1.2e-3	1.2e-17	1	1	0
bert_oc	5000	2000	1	9	16	1.1e-15	2.1e-8	9	0	0
billups	1	1	1	6	11	1.2e-6	1.5e-8	6	0	0
bratu	5625	11250	1	24	41	2.8e-10	4.8e-5	24	0	0
choi	13	26	1	3	4	6.6e-11	1.3e-6	3	0	0
colvdual	20	20	1	–	–	–	–	–	–	–
colvdual	20	20	2	85	196	8.3e-15	2.4e-6	30	54	1
colvnlp	15	15	1	–	–	–	–	–	–	–
colvnlp	15	15	2	29	60	4.4e-20	1.1e-8	17	12	0
cycle	1	1	1	5	7	5.9e-10	1.6e-5	5	0	0
ehl_kost	101	100	1	–	–	–	–	–	–	–
ehl_kost	101	100	2	–	–	–	–	–	–	–
ehl_kost	101	100	3	–	–	–	–	–	–	–
explcp	16	16	1	30	57	8.0e-36	1.8e-18	30	0	0
freebert	15	10	1	1	2	1.2e-3	1.1e-17	0	1	0
freebert	15	10	2	33	72	1.5e-19	2.3e-8	28	4	1
freebert	15	10	3	1	2	1.2e-3	1.3e-16	0	1	0
freebert	15	10	4	1	2	1.2e-3	7.7e-18	0	1	0
freebert	15	10	5	27	55	1.6e-16	8.5e-7	19	8	0
freebert	15	10	6	1	2	1.2e-3	1.6e-16	0	1	0
gafni	5	10	1	23	49	1.7e-21	5.0e-9	23	0	0
gafni	5	10	2	32	65	7.9e-20	3.4e-8	31	0	1
gafni	5	10	3	33	65	2.8e-12	2.0e-4	33	0	0
hanskoop	14	14	1	31	62	2.1e-5	3.5e-7	28	3	0
hanskoop	14	14	2	33	92	4.1e-6	1.5e-7	31	2	0
hanskoop	14	14	3	8	17	1.1e-4	3.3e-6	5	3	0
hanskoop	14	14	4	56	106	7.4e-5	4.8e-7	52	4	0
hanskoop	14	14	5	126	231	9.7e-5	6.8e-7	66	59	1
hydroc06	29	11	1	17	30	1.8e-9	9.8e-7	13	4	0
hydroc20	99	39	1	51	56	3.2e-8	6.0e-6	1	50	0
jel	6	6	1	9	15	4.4e-12	6.8e-5	9	0	0
josephy	4	4	1	18	30	3.5e-15	3.4e-7	17	1	0
josephy	4	4	2	10	17	2.3e-17	2.8e-7	10	0	0
josephy	4	4	3	14	24	8.1e-2	1.1e-14	12	2	0
josephy	4	4	4	3	4	1.1e-15	1.9e-7	3	0	0

Table 1 (continued): Numerical results on the test problems from MCPLIB

problem	n	m	S.P.	it.	F-eval.	$g^s_{\alpha\beta}(x)$	$\|G^s(x)\|$	N	TN	G
josephy	4	4	5	3	4	1.1e-15	1.9e-7	3	0	0
josephy	4	4	6	13	20	3.5e-16	1.1e-7	13	0	0
kojshin	4	4	1	70	125	9.4e-3	3.9e-6	69	1	0
kojshin	4	4	2	10	18	6.6e-12	2.0e-5	10	0	0
kojshin	4	4	3	71	126	9.4e-3	3.9e-6	70	1	0
kojshin	4	4	4	3	4	1.1e-15	1.9e-7	3	0	0
kojshin	4	4	5	3	4	1.1e-15	1.9e-7	3	0	0
kojshin	4	4	6	7	10	5.0e-12	1.3e-5	7	0	0
mathinum	3	3	1	5	6	1.5e-10	2.1e-5	5	0	0
mathinum	3	3	2	4	5	1.1e-16	1.9e-8	4	0	0
mathinum	3	3	3	7	10	2.7e-31	1.1e-15	6	1	0
mathinum	3	3	4	5	6	8.9e-15	1.6e-7	5	0	0
mathisum	4	4	1	6	8	1.1e-16	1.6e-8	6	0	0
methan08	31	15	1	7	8	1.6e-10	1.6e-6	0	7	0
nash	10	10	1	6	7	3.8e-18	3.7e-8	6	0	0
nash	10	10	2	27	53	1.6e-15	6.2e-7	27	0	0
obstacle	2500	2500	1	9	10	6.0e-11	2.2e-5	9	0	0
opt_cont31	1024	1024	1	17	26	2.2e-30	2.4e-16	17	0	0
opt_cont127	4096	4096	1	24	46	8.8e-31	2.2e-16	24	0	0
opt_cont255	8192	8192	1	19	44	2.3e-30	2.1e-16	19	0	0
pgvon105	105	105	1	–	–	–	–	–	–	–
pgvon106	106	106	1	–	–	–	–	–	–	–
pies	42	52	1	37	90	5.4-5	2.0e-11	0	37	0
powell	16	16	1	4	5	1.4e-12	7.5e-5	0	4	0
powell	16	16	2	5	9	2.8e-13	4.2e-6	0	5	0
powell	16	16	3	6	7	2.3e-11	1.2e-5	0	6	0
powell	16	16	4	5	6	8.6e-11	7.1e-6	0	5	0
powell_mcp	8	0	1	5	6	6.9e-14	9.8e-7	5	0	0
powell_mcp	8	0	2	6	7	2.2e-14	5.6e-7	6	0	0
powell_mcp	8	0	3	7	8	4.2e-10	1.0e-4	7	0	0
powell_mcp	8	0	4	6	7	9.0e-10	1.4e-4	6	0	0
scarfanum	13	13	1	7	8	4.5e-10	2.8e-4	5	2	0
scarfanum	13	13	2	8	16	6.1e-10	3.1e-4	7	1	0
scarfanum	13	13	3	9	15	4.8e-16	2.2e-7	9	0	0
scarfasum	14	14	1	9	16	5.7e-20	6.7e-9	9	0	0
scarfbnum	39	39	1	–	–	–	–	–	–	–
scarfbnum	39	39	2	–	–	–	–	–	–	–

Table 1 (continued): Numerical results on the test problems from MCPLIB

problem	n	m	S.P.	it.	F-eval.	$g^s_{\alpha\beta}(x)$	$\|G^s(x)\|$	N	TN	G
scarfbsum	40	40	1	128	230	1.0e-3	7.6e-6	0	128	0
scarfbsum	40	40	2	29	30	4.1e-3	1.2e-6	0	29	0
sppe	27	27	1	27	33	6.4e-13	4.2e-7	4	23	0
sppe	27	27	2	9	12	7.4e-15	9.0e-8	1	6	2
tobin	42	42	1	58	78	1.1e-11	5.5e-6	57	1	0
tobin	42	42	2	30	43	1.5e-13	6.1e-7	29	1	0

References

[1] R. Andreani, A. Friedlander and J.M. Martínez, "On the solution of finite-dimensional variational inequalities using smooth optimization with simple bounds," Technical Report, Department of Applied Mathematics, University of Campinas (Campinas, Brazil, September 1995).

[2] G. Auchmuty, "Variational principles for variational inequalities," *Numerical Functional Analysis and Optimization* 10 (1989) 863–874.

[3] A. Auslender, *Optimisation: Méthodes Numériques* (Masson, Paris, 1976).

[4] R. Barrett, M. Berry, T.F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods* (SIAM, Philadelphia, PA, 1994).

[5] S.C. Billups, S.P. Dirkse and M.C. Ferris, "A comparison of algorithms for large scale mixed complementarity problems," Technical Report, Computer Sciences Department, University of Wisconsin (Madison, WI, October 1995).

[6] C. Chen and O.L. Mangasarian, "A class of smoothing functions for nonlinear and mixed complementarity problems," *Computational Optimization and Applications* 5 (1996) 97–138.

[7] F.H. Clarke, *Optimization and Nonsmooth Analysis* (John Wiley and Sons, New York, 1983, reprinted by SIAM, Philadelphia, PA, 1990).

[8] T. De Luca, F. Facchinei and C. Kanzow, "A semismooth equation approach to the solution of nonlinear complementarity problems," Preprint 93, Institute of Applied Mathematics, University of Hamburg (Hamburg, Germany, January 1995, revised August and December 1995).

[9] J.E. Dennis, Jr., and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1983).

[10] 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) 123–156.

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

[12] S.P. Dirkse, M.C. Ferris, P.V. Preckel and T. Rutherford, "The GAMS callable program library for variational and complementarity solvers," Technical Report 94-07, Computer Sciences Department, University of Wisconsin (Madison, WI, 1994).

[13] F. Facchinei, "Minimization of SC^1 functions and the Maratos effect," *Operations Research Letters* 17 (1995) 131–137.

[14] F. Facchinei, A. Fischer and C. Kanzow, "Inexact Newton methods for semismooth equations with applications to variational inequality problems," to appear in: G. Di Pillo and F. Giannessi, eds., *Nonlinear Optimization and Applications* (Plenum Press, New York, 1996).

[15] F. Facchinei, A. Fischer and C. Kanzow, "A semismooth Newton method for variational inequalities: Theoretical results and preliminary numerical experience," Preprint 102, Institute of Applied Mathematics, University of Hamburg (Hamburg, Germany, December 1995).

[16] F. Facchinei, A. Fischer and C. Kanzow, "A semismooth Newton method for variational inequalities: The case of box constraints," Preprint 103, Institute of Applied Mathematics, University of Hamburg (Hamburg, Germany, December 1995).

[17] F. Facchinei and C. Kanzow, "A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems," to appear in *Mathematical Programming*.

[18] F. Facchinei and C. Kanzow, "On unconstrained and constrained stationary points of the implicit Lagrangian," to appear in *Journal of Optimization Theory and Applications*.

[19] F. Facchinei and J. Soares, "A new merit function for nonlinear complementarity problems and a related algorithm," to appear in *SIAM Journal on Optimization*.

[20] M.C. Ferris and J.-S. Pang, "Engineering and economic applications of complementarity problems," Technical Report 95-07, Computer Sciences Department, University of Wisconsin (Madison, WI, May 1995).

[21] M.C. Ferris and T.F. Rutherford, "Accessing realistic mixed complementarity problems within MATLAB," to appear in: G. Di Pillo and F. Giannessi, eds., *Nonlinear Optimization and Applications* (Plenum Press, New York, 1996).

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

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

[24] M. Fukushima, "Merit functions for variational inequality and complementarity problems," to appear in: G. Di Pillo and F. Giannessi, eds., *Nonlinear Optimization and Applications* (Plenum Press, New York, 1996).

[25] M. Fukushima and J.-S. Pang, "Minimizing and stationary sequences of merit functions for complementarity problems and variational inequalities," Technical Report IS-TR-95026, Nara Institute of Science and Technology (Nara, Japan, September 1995).

[26] S.A. Gabriel and J.J. Moré, "Smoothing of mixed complementarity problems," Preprint P541, Mathematics and Computer Science Division, Argonne National Laboratory (Argonne, IL, September 1995).

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

[28] L. Grippo, F. Lampariello and S. Lucidi, "A nonmonotone linesearch technique for Newton's method," *SIAM Journal on Numerical Analysis* 23 (1986) 707–716.

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

[30] H. Jiang, "Unconstrained minimization approaches to nonlinear complementarities," Technical Report 94/33, School of Mathematics, University of New South Wales (Sydney, Australia, October 1994).

[31] H. Jiang and L. Qi, "Local uniqueness and iterative methods for nonsmooth variational inequalities," *Journal of Mathematical Analysis and Applications* 196 (1995) 314–331.

[32] H. Jiang, L. Qi, X. Chen and D. Sun, "Semismoothness and superlinear convergence in nonsmooth optimization and nonsmooth equations," to appear in: G. Di Pillo and F. Giannessi, eds., *Nonlinear Optimization and Applications* (Plenum Press, New York, 1996).

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

[34] C. Kanzow and M. Fukushima, "Equivalence of the generalized complementarity problem to differentiable unconstrained minimization," to appear in *Journal of Optimization Theory and Applications*.

[35] C. Kanzow, N. Yamashita and M. Fukushima, "New NCP-functions and their properties," Technical Report IS-TR-96010, Nara Institute of Science and Technology (Nara, Japan, February 1996).

[36] Z.-Q. Luo, O.L. Mangasarian, J. Ren and M.V. Solodov, "New error bounds for the linear complementarity problem," *Mathematics of Operations Research* 19 (1994) 880–892.

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

[38] P. Marcotte and J.-P. Dussault, "A note on a globally convergent Newton method for solving monotone variational inequalities," *Operations Research Letters* 6 (1987) 35–42.

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

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

[41] J.-S. Pang, "A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems," *Mathematical Programming* 51 (1991) 101–131.

[42] J.-S. Pang, "Complementarity problems," in: R. Horst and P.M. Pardalos, *Handbook of Global Optimization* (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995) 271–338.

[43] J.-S. Pang and L. Qi, "Nonsmooth equations: Motivation and algorithms," *SIAM Journal on Optimization* 3 (1993) 443–465.

[44] J.-M. Peng, "Equivalence of variational inequality problems to unconstrained minimization," Technical Report, Institute of Computational Mathematics and Scientific Computing, Academic Sinica (Beijing, China, 1995).

[45] J.-M. Peng, "The convexity of the implicit Lagrangian," Technical Report, Institute of Computational Mathematics and Scientific Computing, Academic Sinica (Beijing, China, 1995).

[46] J.-M. Peng and Y.-X. Yuan, "Unconstrained methods for generalized complementarity problems," Technical Report, Institute of Computational Mathematics and Scientific Computing, Academic Sinica (Beijing, China, 1994).

[47] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations," *Mathematics of Operations Research* 18 (1993) 227–244.

[48] L. Qi and J. Sun, "A nonsmooth version of Newton's method," *Mathematical Programming* 58 (1993) 353–367.

[49] D. Sun, M. Fukushima and L. Qi, "A computable generalized Hessian of the D-gap function and Newton-type methods for variational inequality problems," Technical Report, School of Mathematics, University of New South Wales (Sydney, Australia, October 1995).

[50] J.H. Wu, M. Florian and P. Marcotte, "A general descent framework for the monotone variational inequality problem," *Mathematical Programming* 61 (1993) 281–300.

[51] B. Xiao and P.T. Harker, "A nonsmooth Newton method for variational inequalities, I: theory," *Mathematical Programming* 65 (1994) 151–194.

[52] B. Xiao and P.T. Harker, "A nonsmooth Newton method for variational inequalities, II: numerical results," *Mathematical Programming* 65 (1994) 195–216.

[53] N. Yamashita and M. Fukushima, "Equivalent unconstrained minimization and global error bounds for variational inequality problems," to appear in *SIAM Journal on Control and Optimization*.

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

[55] N. Yamashita, K. Taji and M. Fukushima, "Unconstrained optimization reformulations of variational inequality problems," to appear in *Journal of Optimization Theory and Applications*.