

# Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems

Guoyin Li<sup>†</sup>

Kung Fu Ng<sup>‡</sup>

**Abstract.** We present some error bound results of generalized D-gap functions for nonsmooth and nonmonotone variational inequality (VI) problems. Application is given in providing a derivative-free descent method.

**Key words.** variational inequality problem, generalized D-gap function and error bound.

**AMS subject classifications.** 65H10, 90C26

## 1 Introduction

Throughout this paper, let  $P$  denote a nonempty closed convex set in an Euclidean space  $\mathbb{R}^n$  and let  $F$  be a continuous map from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ . We consider  $VI(F, P)$ , the variational inequality problem associated with the data  $F$  and  $P$ , that is, to find a vector  $x \in P$  such that

$$F(x)^T(y - x) \geq 0 \quad \text{for all } y \in P. \quad (1.1)$$

When  $P$  is the nonnegative orthant in  $\mathbb{R}^n$ ,  $VI(F, P)$  reduces to the nonlinear complementarity problem  $NCP(F)$ , i.e., finding a vector  $x \in \mathbb{R}_+^n$ ,  $F(x) \in \mathbb{R}_+^n$  such that  $F(x)^T x = 0$ . Variational inequality (VI) problems have been widely studied in various fields such as mathematical programming, game theory and economics etc.; see [8, 9, 10] and references therein for the background information and motivation of the (VI) problems covering both smooth and nonsmooth functions. In fact nonsmooth variational problems are quite abundant, see [1, 11, 12, 23] for recent developments. Below, let us only mention explicitly one of the simplest examples.

**Example 1.1** *Let  $\Omega$  be a bounded open set in  $\mathbb{R}^2$  with Lipschitz boundary  $\partial\Omega$ . Given a continuous function  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  and consider the following free boundary problem: find a twice*

---

\*Research of the second author was supported by an Earmarked Grant from the Research Grant Council of Hong Kong.

<sup>†</sup>Department of Mathematics, The Chinese University of Hong Kong, Shatin, New Territory, Hong Kong (gyli@math.cuhk.edu.hk)

<sup>‡</sup>Department of Mathematics, The Chinese University of Hong Kong, Shatin, New Territory, Hong Kong (kfung@math.cuhk.edu.hk)

continuously differentiable function  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that

$$\begin{aligned} \Delta u + \phi(u) &= 0 & \text{in } & \Omega^+, \\ u &= 0 & \text{in } & \Omega_0, \\ u = |\nabla u| &= 0 & \text{on } & \Gamma, \\ u &= 1 & \text{on } & \partial\Omega \end{aligned}$$

where  $\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$ ,  $\Omega^+ = \{z \in \Omega : u(z) > 0\}$ ,  $\Omega_0 = \{z \in \Omega : u(z) = 0\}$  and  $\Gamma = \partial\Omega_0 = \partial\Omega^+ \cap \Omega$ . In the special case when  $\phi(u) = \lambda u^p$  for some  $0 < p \leq 1$ , this model has been discussed in [1] (see also [4]). Since the corresponding domain  $\Omega^+$ ,  $\Omega_0$  and  $\partial\Omega$  depend on the solution  $u$  (and hence are unknown objects), it is hard to obtain an analytic expression for the solution in general. Thus, one often seeks for a numerical approximated solution. Using finite element approximation or finite difference approximation, we obtain a nonlinear complementarity problem  $NCP(F)$  with  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  defined by  $F(x) = Mx + p(x) + q$ , where  $n$  is a positive integer related to precision of the discretization,  $M$  is some symmetric  $n \times n$  matrix,  $q \in \mathbb{R}^n$  and  $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is defined by

$$p(x_1, \dots, x_n) = (\phi(\max\{x_1, 0\}), \dots, \phi(\max\{x_n, 0\})). \quad (1.2)$$

Note that  $F$  is nonsmooth in general. Therefore, we see that the corresponding nonlinear complementarity problem is an example of nonsmooth variational inequality problem.

In recent years, much effort has been made to reformulate the (VI) problems as equivalent optimization problems through the consideration of merit functions. Among these merit functions, D-gap functions and generalized D-gap functions for (VI) problems are particularly interesting, because they cast (VI) problems as equivalent unconstrained optimization problems (cf. [24, 33]). Peng and Fukushima [25] gave a global error bound result for D-gap functions under the assumption that  $F$  is smooth, (globally) Lipschitz on  $\mathbb{R}^n$  and strongly monotone. Their result was extended by Yamashita, Taji and Fukushima [33] to cover generalized D-gap functions. Recently, Huang and Ng [13] gave an example showing that the D-gap functions might not provide global error bounds if the globally Lipschitz assumption of  $F$  is dropped.

In this present paper, we establish some error bound results for generalized D-gap functions applicable to the case when  $F$  is not necessarily smooth and not necessarily (strongly) monotone. In particular, we show in section 4 that although a global error bound result might fail if the Lipschitz assumption of  $F$  is not assumed, a local error bound result (with a fractional exponent) still holds. As an application, we establish a convergence result in section 5 for a derivative-free descent method of Armijo type leading to the solution of the corresponding  $VI(F, P)$ .

## 2 Preliminaries

For  $x_0$  in a finite dimensional Euclidean space  $X$  and  $\delta > 0$ , let  $B(x_0, \delta)$  (resp.  $\overline{B}(x_0, \delta)$ ) denote the open (resp. closed) ball with center  $x_0$  and radius  $\delta$ . For subset  $C$  of  $X$ , we denote the interior, closure, convex hull and topological boundary of it by  $\text{int } C$ ,  $\overline{C}$ ,  $\text{co } C$  and  $\partial C$  respectively.

**Definition 2.1** Let  $m, n \in \mathbb{N}$ . A vector-valued function  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is said to be Lipschitz (with modulus  $L$ ) around  $x_0 \in \mathbb{R}^n$  if there exist two positive constants  $\delta$  and  $L$  such that

$$\|G(x_1) - G(x_2)\| \leq L\|x_1 - x_2\| \text{ for all } x_1, x_2 \in B(x_0, \delta). \quad (2.1)$$

If  $G$  is Lipschitz around each point of  $\mathbb{R}^n$ , then  $G$  is said to be locally Lipschitz (on  $\mathbb{R}^n$ ). Moreover,  $G$  is said to be Lipschitz on a given set  $C \subseteq \mathbb{R}^n$  with modulus  $L$  if (2.1) holds with  $B(x_0, \delta)$  replaced by  $C$ . Given any subset  $B$  of  $\mathbb{R}^n$  with zero (Lebesgue) measure and  $x_0 \in \mathbb{R}^n$ , define  $\partial^B G(x_0)$  by

$$\partial^B G(x_0) = \{\xi \in \mathbb{R}^{m \times n} : \xi = \lim_{i \rightarrow \infty} \nabla G(x_i), x_i \rightarrow x_0, x_i \in D_G \setminus B\}, \quad (2.2)$$

where  $D_G$  denotes the set of all differentiable points of  $G$ . Then, as in [6] (see also [30, Theorem 4]), the (Clarke) generalized Jacobian  $\partial^c G(x_0)$  is given by

$$\partial^c G(x_0) = \text{co } \partial^B G(x_0). \quad (2.3)$$

If  $m = 1$  and  $g = G$  then

$$\partial^c g(x_0) = \{\xi \in \mathbb{R}^n : \xi^T v \leq g^\circ(x_0; v) \text{ for all } v \in \mathbb{R}^n\},$$

where  $\xi^T$  denotes the transpose of  $\xi$ , and  $g^\circ(x_0; v)$  denotes the Clarke directional derivative of  $g$  at  $x_0$  in the direction  $v$  (see [5, 6]):

$$g^\circ(x_0; v) = \limsup_{y \rightarrow x_0, t \downarrow 0} \frac{g(y + tv) - g(y)}{t}.$$

Let  $U \subseteq \mathbb{R}^n$  be an open set and let  $G$  be Lipschitz on the set  $U$  with modulus  $L$ . Then for any  $x \in U$  we have

$$\|V\| \leq L \text{ for any } V \in \partial^c G(x). \quad (2.4)$$

**Definition 2.2 (cf. [8])** Let  $C$  be a subset of  $\mathbb{R}^n$ . A mapping  $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is said to be (i) coercive on  $C$  if

$$\lim_{x \in C, \|x\| \rightarrow \infty} \frac{G(x)^T(x - y)}{\|x\|} = +\infty \text{ for any } y \in C \quad (2.5)$$

(thus if  $C$  is bounded then  $G$  is always coercive on  $C$ .);

(ii) monotone on  $C$  if  $[G(x) - G(y)]^T(x - y) \geq 0$  for all vectors  $x$  and  $y$  in  $C$ ;

(iii) strongly monotone on  $C$  (with modulus  $\mu > 0$ ) if for all vectors  $x$  and  $y$  in  $C$ ,

$$[G(x) - G(y)]^T(x - y) \geq \mu\|x - y\|^2. \quad (2.6)$$

**Remark 2.1** (i) If  $G$  is locally Lipschitz on  $\mathbb{R}^n$  and strongly monotone on  $\mathbb{R}^n$  with modulus  $\mu > 0$ , then (see [14]) for any  $x \in \mathbb{R}^n$ ,  $V \in \partial^c G(x)$  we have

$$d^T V d \geq \mu\|d\|^2 \text{ for all } d \in \mathbb{R}^n. \quad (2.7)$$

(ii) If  $G$  is strongly monotone on a closed convex subset  $C$  of  $\mathbb{R}^n$ , then it is coercive on  $C$ . Indeed, suppose that (2.6) holds for all  $x, y \in C$ . Then for any  $y \in C$ ,

$$\lim_{x \in C, \|x\| \rightarrow \infty} \frac{[G(x) - G(y)]^T(x - y)}{\|x - y\|} = +\infty. \quad (2.8)$$

Note that for any fixed  $y \in C$ ,  $\{\frac{G(y)^T(x-y)}{\|x-y\|} : x \in C\}$  is bounded. It follows from (2.8) that

$$\lim_{x \in C, \|x\| \rightarrow \infty} \frac{G(x)^T(x - y)}{\|x\|} = \lim_{x \in C, \|x\| \rightarrow \infty} \frac{G(x)^T(x - y)}{\|x - y\|} = +\infty.$$

(iii) Let  $P \subseteq \mathbb{R}^n$  be a closed convex set and let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be continuous on  $P$ . Then, by [8, Proposition 2.2.3], the VI( $F, P$ ) has a solution whenever  $F$  is coercive on  $P$ . Moreover, if  $F$  is strongly monotone on  $P$ , the VI( $F, P$ ) has a unique solution (cf. [8, Theorem 2.3.3]).

Following [2, 3, 7, 8, 17], a set  $C \subseteq \mathbb{R}^n$  is said to be

(i) semianalytic, if for any  $x \in \mathbb{R}^n$ , there exists a neighbourhood  $U$  of  $x$  such that

$$C \cap U = \bigcup_{i=1}^l \bigcap_{j=1}^s \{x \in U : f_{ij}(x) = 0, g_{ij}(x) < 0\}$$

for some integers  $l, s$  and some real analytic functions  $f_{ij}, g_{ij}$  on  $\mathbb{R}^n$  ( $1 \leq i \leq l, 1 \leq j \leq s$ );

(ii) subanalytic if for any  $x \in C$ , there exist a neighbourhood  $U$  of  $x$  and a bounded semianalytic set  $Z \subseteq \mathbb{R}^{n+p}$  such that  $C \cap U = \{x \in \mathbb{R}^n : (x, y) \in Z \text{ for some } y \in \mathbb{R}^p\}$ .

Moreover, a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be subanalytic if its graph  $\text{gph} f := \{(x, f(x)) : x \in \mathbb{R}^n\}$  is subanalytic and a vector-valued function  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is said to be subanalytic if each of its component is subanalytic. We summarize below some basic properties of subanalytic sets and subanalytic functions ((S3) is easily seen from the definition):

**(S1)** (cf. [8, (p1) and (p2) P.597]) Finite union (resp. intersection) of subanalytic sets is subanalytic. The Cartesian product (resp. complement, closure) of subanalytic sets is subanalytic.

**(S2)** (cf. [3, Theorem 2.3]) Let  $C$  be a bounded subanalytic set in  $\mathbb{R}^{n+p}$  (for some positive integers  $n, p$ ) and let  $\pi : \mathbb{R}^{n+p} \rightarrow \mathbb{R}^n$  be the projection defined by  $\pi(x, y) = x$  for all  $x \in \mathbb{R}^n$  and  $y \in \mathbb{R}^p$ . Then  $\pi(C)$  is a subanalytic set of  $\mathbb{R}^n$ .

**(S3)** If  $C \cap V$  is subanalytic for each compact subanalytic set  $V$  of  $\mathbb{R}^n$ , then  $C$  is subanalytic.

**(S4)** (cf. [8, (p3) P.597])  $d(\cdot, C)$  is a subanalytic function if  $C$  is a subanalytic set where  $d(x, C) := \inf_{c \in C} \|x - c\|$  for any  $x \in \mathbb{R}^n$ .

**(S5)** (cf. [8, (p5) and (p8) P.598]) If  $f, g$  are subanalytic functions on  $\mathbb{R}^n$  and  $\lambda \in \mathbb{R}$  then  $f + g$  (resp.  $\lambda f, \max\{f, g\}$ ) is subanalytic. If  $F, G : \mathbb{R}^n \rightarrow \mathbb{R}^n$  are subanalytic vector-valued functions, then  $F^T G$  is a subanalytic function on  $\mathbb{R}^n$ .

**(S6)** (cf. [8, (p4) P.597]) If  $f$  is subanalytic and  $\lambda \in \mathbb{R}$ , then  $\{x : f(x) = \lambda\}, \{x : f(x) < \lambda\}$  and  $\{x : f(x) \leq \lambda\}$  are subanalytic.

**(S7)** (Łojasiewicz's inequality, cf. [17, Theorem 2.1.1]) If  $f, g$  are continuous subanalytic functions on compact subanalytic set  $C \subseteq \mathbb{R}^n$  such that  $f^{-1}(0) \subseteq g^{-1}(0)$  then there exist a constant  $\alpha > 0$  and a positive integer  $N$  such that

$$\alpha |g(x)|^N \leq |f(x)| \text{ for all } x \in C.$$

**Remark 2.2** If  $f$  is a continuous function on  $\mathbb{R}^n$  such that  $U_f := \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \geq r\}$  is subanalytic then  $f$  is subanalytic. Indeed, note that (by the continuity of  $f$ ),  $L_f := \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq r\} = \overline{\mathbb{R}^{n+1} \setminus U_f}$ . It follows from (S1) that  $L_f$  is also subanalytic. This implies that  $\text{gph} f$  is subanalytic and hence  $f$  is subanalytic.

### 3 Generalized D-gap functions

Given  $VI(F, P)$ , the D-gap function  $\theta_{ab}$  is defined as the difference of two regularized gap functions:

$$\theta_{ab}(x) \equiv \theta_a(x) - \theta_b(x) \text{ for all } x \in \mathbb{R}^n,$$

where  $0 < a < b$ , and  $\theta_c$  is a regularized gap function defined by

$$\theta_c(x) = \sup_{y \in P} \{F(x)^T(x - y) - \frac{c}{2}\|x - y\|^2\} \text{ for all } x \in \mathbb{R}^n, c > 0. \quad (3.1)$$

In order to have wider scope of applications as well as more efficient algorithms, Yamashita et.al. [33] replaced the function  $\frac{1}{2}\|x - y\|^2$  in (3.1) by a more general function  $\varphi$  and thereby introduced generalized D-gap functions  $\theta_{ab}^\varphi$  and generalized regularized gap functions  $\theta_c^\varphi$  respectively defined by (for  $0 < a < b$  and  $c > 0$ )

$$\theta_{ab}^\varphi(x) \equiv \theta_a^\varphi(x) - \theta_b^\varphi(x) \text{ for all } x \in \mathbb{R}^n \quad \text{and} \quad (3.2)$$

$$\theta_c^\varphi(x) = \sup_{y \in P} \{F(x)^T(x - y) - c \varphi(x, y)\} \text{ for all } x \in \mathbb{R}^n. \quad (3.3)$$

Here  $\varphi$  is a real-valued function on  $\mathbb{R}^n \times \mathbb{R}^n$  satisfying the following properties (P6 and P7 are redundant; see [13, 33]):

**P1:**  $\varphi$  is continuously differentiable on  $\mathbb{R}^n \times \mathbb{R}^n$ .

**P2:**  $\varphi(x, y) \geq 0$  for all  $x, y \in \mathbb{R}^n$  and the equality holds if and only if  $x = y$ .

**P3:**  $\{\varphi(x, \cdot) : x \in \mathbb{R}^n\}$  is uniformly strongly convex in the sense that there exists a positive real number  $\beta$  such that for all  $x \in \mathbb{R}^n$ ,

$$\varphi(x, y_1) - \varphi(x, y_2) \geq \nabla_y \varphi(x, y_2)^T(y_1 - y_2) + \beta\|y_1 - y_2\|^2 \text{ for all } y_1, y_2 \in \mathbb{R}^n, \quad (3.4)$$

where  $\nabla_y \varphi$  denotes the partial derivative of  $\varphi$  with respect to the second variable.

**P4:**  $\{\nabla_y \varphi(x, \cdot) : x \in \mathbb{R}^n\}$  is uniformly Lipschitz on  $\mathbb{R}^n$  (with modulus  $\kappa \geq 2\beta > 0$ ) that is, for all  $x \in \mathbb{R}^n$

$$\|\nabla_y \varphi(x, y_1) - \nabla_y \varphi(x, y_2)\| \leq \kappa\|y_1 - y_2\| \text{ for all } y_1, y_2 \in \mathbb{R}^n. \quad (3.5)$$

**P5:** For any  $x, y \in \mathbb{R}^n$ ,  $\nabla_x \varphi(x, y) = -\nabla_y \varphi(x, y)$ .

**P6:**  $\nabla_x \varphi(x, y) = 0 \Leftrightarrow \nabla_y \varphi(x, y) = 0 \Leftrightarrow x = y$ .

**P7:** For all  $x, y \in \mathbb{R}^n$  we have

$$\beta\|x - y\|^2 \leq \varphi(x, y) \leq (\kappa - \beta)\|x - y\|^2. \quad (3.6)$$

**Remark 3.1** For example, the function

$$\varphi(x, y) = \frac{1}{2}\|x - y\|^2 \quad (3.7)$$

satisfies properties **P1-P7**. More generally, for any twice continuously differentiable, strongly convex function  $\nu : \mathbb{R}^n \rightarrow [0, \infty)$  such that  $\nu(0) = 0$ , the function  $\varphi(x, y) := \nu(x - y)$  also has properties **P1-P7**.

Below, we list some basic properties of functions  $\theta_{ab}^\varphi$  and  $\theta_c^\varphi$  defined respectively by (3.2) and (3.3). From now and onward,  $\varphi$  denotes a function on  $\mathbb{R}^n \times \mathbb{R}^n$  satisfying P1-P7 (with the associated constants  $\beta, \kappa > 0$ ). We always assume that  $a, b, c > 0$  and  $a < b$ . The following proposition collects some useful facts for the generalized D-gap functions in which part (i) – (iii) are taken from [31], part (iv) – (vii) from [33] and part (viii) from [19].

**Proposition 3.1** The following statements are valid for  $VI(F, P)$ .

(i) For every  $x \in \mathbb{R}^n$ , there exists a unique vector  $y_c^\varphi(x) \in P$  at which the supremum in (3.3) is attained, i.e.,

$$\theta_c^\varphi(x) = F(x)^T(x - y_c^\varphi(x)) - c\varphi(x, y_c^\varphi(x)). \quad (3.8)$$

(ii)  $y_c^\varphi(\cdot)$  and  $\theta_c^\varphi(\cdot)$  are continuous on  $\mathbb{R}^n$  and  $\theta_c^\varphi(x) \geq 0$  for all  $x \in P$ .

(iii)  $[x \in P, \theta_c^\varphi(x) = 0] \Leftrightarrow [x = y_c^\varphi(x)] \Leftrightarrow x$  is a solution of  $VI(F, P)$ .

(iv)  $\theta_{ab}^\varphi$  is continuous and nonnegative on  $\mathbb{R}^n$ .

(v)  $[\theta_{ab}^\varphi(x) = 0] \Leftrightarrow x$  is a solution of  $VI(F, P) \Leftrightarrow [x = y_a^\varphi(x) = y_b^\varphi(x)]$ .

(vi) For all  $x \in \mathbb{R}^n$ , we have

$$(b - a)\varphi(x, y_b^\varphi(x)) \leq \theta_{ab}^\varphi(x) \leq (b - a)\varphi(x, y_a^\varphi(x)) \quad \text{and} \quad (3.9)$$

$$\beta(b - a)\|x - y_b^\varphi(x)\|^2 \leq \theta_{ab}^\varphi(x) \leq (\kappa - \beta)(b - a)\|x - y_a^\varphi(x)\|^2 \quad (3.10)$$

(vii) If  $F$  is continuously differentiable on  $\mathbb{R}^n$ , then  $\theta_c^\varphi$  and  $\theta_{ab}^\varphi$  are also continuously differentiable on  $\mathbb{R}^n$ .

(viii) If  $F$  is locally Lipschitz on  $\mathbb{R}^n$ , then  $y_c^\varphi$ ,  $\theta_c^\varphi$  and  $\theta_{ab}^\varphi$  are also locally Lipschitz on  $\mathbb{R}^n$ .

**Remark 3.2** (cf. [31]) If  $\varphi$  is given by (3.7), then for any  $x \in \mathbb{R}^n$ ,  $y_c^\varphi(x)$  is the projection of  $x - c^{-1}F(x)$  on  $P$ , that is

$$y_c^\varphi(x) = Pr_P(x - c^{-1}F(x)) \text{ for all } x \in \mathbb{R}^n. \quad (3.11)$$

**Theorem 3.1** Let  $F$  be a locally Lipschitz function on  $\mathbb{R}^n$ . Then we have

$$\begin{aligned} \partial^c \theta_{ab}^\varphi(x) &= \partial^c F(x)^T(y_b^\varphi(x) - y_a^\varphi(x)) \\ &\quad - a\nabla_x \varphi(x, y_a^\varphi(x)) + b\nabla_x \varphi(x, y_b^\varphi(x)) \text{ for all } x \in \mathbb{R}^n. \end{aligned} \quad (3.12)$$

**Proof.** Since  $F$  is locally Lipschitz, one can apply Proposition 3.1(viii) and the Rademacher theorem (cf. [26]) to see that  $F$ ,  $\theta_a^\varphi$  and  $\theta_b^\varphi$  are differentiable almost everywhere on  $\mathbb{R}^n$ . Hence there exists a set  $A$  of measure zero such that these functions are differentiable at each point of  $\mathbb{R}^n \setminus A$ . Applying (2.2) and (2.3) with  $\{\theta_{ab}^\varphi, 1\}$  in place of  $\{G, m\}$ , we obtain that for each  $x \in \mathbb{R}^n$ ,

$$\partial^c \theta_{ab}^\varphi(x) = \text{co} \{ \partial^A \theta_{ab}^\varphi(x) \} = \text{co} \{ \lim_{i \rightarrow \infty} \nabla \theta_{ab}^\varphi(x_i) : x_i \rightarrow x, x_i \in \mathbb{R}^n \setminus A \}. \quad (3.13)$$

For any  $\mu > 0$ , let us define  $\psi_\mu : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  by

$$\psi_\mu(x, y) := F(x)^T(x - y) - \mu \varphi(x, y) \text{ for all } x, y \in \mathbb{R}^n. \quad (3.14)$$

Then by (3.3) and (3.8),

$$\theta_\mu^\varphi(x) = \sup_{y \in P} \psi_\mu(x, y) = \psi_\mu(x, y_\mu^\varphi(x)) \text{ for all } x \in \mathbb{R}^n. \quad (3.15)$$

Since for all  $\mu > 0$ ,  $x \in \mathbb{R}^n \setminus A$  and  $v \in \mathbb{R}^n$ ,

$$\begin{aligned} \nabla_x \psi_\mu(x, y_\mu^\varphi(x)) &= \nabla \psi_\mu(\cdot, y_\mu^\varphi(x))|_x \\ &= \nabla F(\cdot)^T(\cdot - y_\mu^\varphi(x)) + F(\cdot) - \mu \nabla_x \varphi(\cdot, y_\mu^\varphi(x))|_x \\ &= \nabla F(x)^T(x - y_\mu^\varphi(x)) + F(x) - \mu \nabla_x \varphi(x, y_\mu^\varphi(x)), \end{aligned} \quad (3.16)$$

and  $\psi_\mu(x + tv, y_\mu^\varphi(x + tv)) = \sup_{y \in P} \psi_\mu(x + tv, y) \geq \psi_\mu(x + tv, y_\mu^\varphi(x))$  (thanks to (3.15)), one has

$$\begin{aligned} \nabla \theta_\mu^\varphi(x)v &= (\theta_\mu^\varphi)'(x; v) \\ &= \lim_{t \downarrow 0} \frac{\theta_\mu^\varphi(x + tv) - \theta_\mu^\varphi(x)}{t} \\ &= \lim_{t \downarrow 0} \frac{\psi_\mu(x + tv, y_\mu^\varphi(x + tv)) - \psi_\mu(x, y_\mu^\varphi(x))}{t} \\ &\geq \lim_{t \downarrow 0} \frac{\psi_\mu(x + tv, y_\mu^\varphi(x)) - \psi_\mu(x, y_\mu^\varphi(x))}{t} \\ &= \nabla_x \psi_\mu(x, y_\mu^\varphi(x))v. \end{aligned} \quad (3.17)$$

Since  $v$  is arbitrary, it follows from (3.16) that, for any  $\mu > 0$  and  $x \in \mathbb{R}^n \setminus A$ ,

$$\nabla \theta_\mu^\varphi(x) = \nabla_x \psi_\mu(x, y_\mu^\varphi(x)) = \nabla F(x)^T(x - y_\mu^\varphi(x)) + F(x) - \mu \nabla_x \varphi(x, y_\mu^\varphi(x)).$$

Consequently, by considering  $\mu = a, b$  we have

$$\begin{aligned} \nabla \theta_{ab}^\varphi(x) &= \nabla \theta_a^\varphi(x) - \nabla \theta_b^\varphi(x) \\ &= \nabla F(x)^T(y_b^\varphi(x) - y_a^\varphi(x)) \\ &\quad - a \nabla_x \varphi(x, y_a^\varphi(x)) + b \nabla_x \varphi(x, y_b^\varphi(x)), \text{ for all } x \in \mathbb{R}^n \setminus A. \end{aligned} \quad (3.18)$$

Next, let us consider a general  $x \in \mathbb{R}^n$ , and take  $\rho > 0$  such that  $F$  is Lipschitz on  $B(x, \rho)$ ; thus  $\{\nabla F(z) : z \in B(x, \rho) \setminus A\}$  is bounded. Since  $y_a^\varphi$  and  $y_b^\varphi$  are continuous (by (ii)), it follows from (2.3) that

$$\begin{aligned} &\text{co}\left\{ \lim_{i \rightarrow \infty} \nabla F(x_i)^T(y_b^\varphi(x_i) - y_a^\varphi(x_i)) : x_i \rightarrow x, x_i \in \mathbb{R}^n \setminus A \right\} \\ &= \left( \text{co}\left\{ \lim_{i \rightarrow \infty} \nabla F(x_i) : x_i \rightarrow x, x_i \in \mathbb{R}^n \setminus A \right\} \right)^T (y_b^\varphi(x) - y_a^\varphi(x)) \\ &= \partial^c F(x)^T (y_b^\varphi(x) - y_a^\varphi(x)). \end{aligned}$$

Combining this with (3.13), (3.18) and (P1), we have for all  $x \in \mathbb{R}^n$

$$\partial^c \theta_{ab}^\varphi(x) = \partial^c F(x)^T (y_b^\varphi(x) - y_a^\varphi(x)) - a \nabla_x \varphi(x, y_a^\varphi(x)) + b \nabla_x \varphi(x, y_b^\varphi(x)).$$

Thus (3.12) is shown and the proof is complete.  $\square$

## 4 Error bound results for generalized D-gap functions

In this section, we establish some error bound results for generalized D-gap functions. Denote the distance from  $x$  to a set  $C$  by  $d(x, C)$ . For a function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ , we sometimes use  $[f \leq \epsilon]$  to denote the (sub-)level set  $\{x \in \mathbb{R}^n : f(x) \leq \epsilon\}$  at level  $\epsilon$ . Following [8] and [9], we say that  $f$  has a local error bound on  $C$  if there exist two positive constants  $\tau, \epsilon$  such that for all  $x \in [f \leq \epsilon] \cap C$

$$d(x, S_f \cap C) \leq \tau \max\{f(x), 0\}, \quad (4.1)$$

where  $S_f := [f \leq 0]$ . Furthermore, we say that  $f$  has a global error bound on  $C$  if there exists a constant  $\tau > 0$  such that (4.1) holds for all  $x \in C$ .

The following result was established by Peng and Fukushima [25] in the special case when  $\varphi$  is given by (3.7) and  $F$  is strongly monotone on  $\mathbb{R}^n$  (see Remark 2.1 (ii)).

**Lemma 4.1** *Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be coercive on  $\mathbb{R}^n$ . Then for any constants  $a, b$  satisfying  $0 < a < b$ , the level set  $[\theta_{ab}^\varphi \leq \epsilon]$  is bounded for any  $\epsilon > 0$ .*

**Proof.** Suppose on the contrary that there exist a constant  $\epsilon_0 > 0$  and a sequence  $\{x_k\}_{k=1}^\infty$  such that  $\lim_{k \rightarrow \infty} \|x_k\| = \infty$  and  $\theta_{ab}^\varphi(x_k) \leq \epsilon_0$  for all  $k \in \mathbb{N}$ . Letting  $w := \left(\frac{\epsilon_0}{\beta(b-a)}\right)^{1/2}$  where  $\beta$  is defined as in (P3), it follows from (3.10) that

$$\|x_k - y_b^\varphi(x_k)\| \leq w \text{ for all } k \in \mathbb{N}. \quad (4.2)$$

Consequently, by (P6) and (3.5), we have

$$\|\nabla_y \varphi(x_k, y_b^\varphi(x_k))\| = \|\nabla_y \varphi(x_k, y_b^\varphi(x_k)) - \nabla_y \varphi(x_k, x_k)\| \leq \kappa \|x_k - y_b^\varphi(x_k)\| \leq \kappa w.$$

Since  $y_b^\varphi(x_k)$  minimizes the function  $-F(x_k)^T(x_k - \cdot) + b\varphi(x_k, \cdot)$  on  $P$  (see Proposition 3.1(i)), it follows from the first order optimality condition and the Cauchy-Schwartz inequality that, for any  $y \in P$ ,

$$\begin{aligned} 0 &\leq \left(F(x_k) + b\nabla_y \varphi(x_k, y_b^\varphi(x_k))\right)^T \left(y - y_b^\varphi(x_k)\right) \\ &\leq F(x_k)^T(y - y_b^\varphi(x_k)) + b\kappa w \|y - y_b^\varphi(x_k)\|. \end{aligned} \quad (4.3)$$

Fixing  $y \in P$  and denoting  $y + x_k - y_b^\varphi(x_k)$  by  $y_k$ , this implies that

$$0 \leq \frac{-F(x_k)^T(x_k - y_k)}{\|x_k - y_k\|} + b\kappa w \text{ for all } k \in \mathbb{N}. \quad (4.4)$$

On the other hand, since  $\{y_k\}_{k \in \mathbb{N}}$  is bounded, there exists a bounded box  $B$  such that  $\{y_k\}_{k \in \mathbb{N}} \subseteq B$ . Hence, for each  $k$ ,

$$\frac{F(x_k)^T(x_k - y_k)}{\|x_k\|} \geq \inf_{z \in B} \frac{F(x_k)^T(x_k - z)}{\|x_k\|}. \quad (4.5)$$

By a well-known theorem in linear programming, the infimum on the right hand side of the preceding inequality is attained at an extremal point (depending on  $k$ ) of  $B$ . Since  $B$  is a box

(hence has only finitely many extremal points), it follows from (4.5) and the definition of coercive that

$$\lim_{k \rightarrow \infty} \frac{F(x_k)^T(x_k - y_k)}{\|x_k\|} = +\infty. \quad (4.6)$$

Since  $\|x_k\| \rightarrow \infty$  and  $\{y_k\}_{k \in \mathbb{N}}$  is bounded, this implies that

$$\lim_{k \rightarrow \infty} \frac{F(x_k)^T(x_k - y_k)}{\|x_k - y_k\|} = +\infty.$$

This contradicts (4.4) and completes the proof.  $\square$

**Theorem 4.1** *Let  $\varphi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $P \subseteq \mathbb{R}^n$  be subanalytic. Then the following statements hold for any constants  $a, b$  satisfying  $0 < a < b$ :*

(i)  $\theta_{ab}^\varphi$  is a subanalytic function on  $\mathbb{R}^n$ .

(ii) Suppose in addition that  $F$  is coercive on  $\mathbb{R}^n$ . Then there exists  $\gamma \in (0, 1]$  such that  $(\theta_{ab}^\varphi)^\gamma$  has a local error bound on  $\mathbb{R}^n$ .

**Proof.** To see (i), by (S5) it suffices to show  $\theta_c^\varphi$  is subanalytic for any  $c > 0$ . To do this, fix a constant  $c > 0$ , a compact subanalytic set  $V$  of  $\mathbb{R}^n \times \mathbb{R}$  and let  $M := \sup_{(x,r) \in V} \|y_c^\varphi(x)\| < \infty$ . Since  $y_c^\varphi$  is continuous (by Proposition 3.1(ii)),  $M$  is finite. Define  $A_c = A_c^{(1)} \cap A_c^{(2)}$  where

$$A_c^{(1)} = \{(x, r, y) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n : F(x)^T(x - y) - c\varphi(x, y) - r \geq 0\},$$

$$A_c^{(2)} = V \times (P \cap \overline{B}(0, M)).$$

By the given assumptions together with (S1), (S5) and (S6), it is clear that  $A_c^{(1)}$ ,  $A_c^{(2)}$  and  $A_c$  are subanalytic. Define  $\pi : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}$  by

$$\pi(x, r, y) = (x, r).$$

We claim that

$$\pi(A_c) = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : \theta_c^\varphi(x) \geq r\} \cap V. \quad (4.7)$$

Granting this, (S2) implies that  $\{(x, r) \in \mathbb{R}^n \times \mathbb{R} : \theta_c^\varphi(x) \geq r\} \cap V$  is subanalytic for any compact subanalytic set  $V \subseteq \mathbb{R}^n$ . Consequently, (S3) implies that

$$\{(x, r) \in \mathbb{R}^n \times \mathbb{R} : \theta_c^\varphi(x) \geq r\} \text{ is subanalytic.} \quad (4.8)$$

and thus  $\theta_c^\varphi$  is subanalytic by Proposition 3.1(ii) and Remark 2.2. To see (4.7), let  $(x, r) \in \pi(A_c)$ . Then  $(x, r) \in V$  and there exists  $y_0 \in P \cap \overline{B}(0, M)$  such that

$$F(x)^T(x - y_0) - c\varphi(x, y_0) \geq r.$$

Hence  $\theta_c^\varphi(x) \geq r$  and  $\pi(A_c) \subseteq \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : \theta_c^\varphi(x) \geq r\} \cap V$ . Conversely, let  $(x, r) \in V$  such that  $\theta_c^\varphi(x) \geq r$ . Then, by Proposition 3.1(i) and the definition of  $M$ , we have  $(x, r, y_c^\varphi(x)) \in V \times (P \cap \overline{B}(0, M))$  and

$$F(x)^T(x - y_c^\varphi(x)) - c\varphi(x, y_c^\varphi(x)) - r = \theta_c^\varphi(x) - r \geq 0.$$

Therefore,  $(x, r, y_c^\varphi(x)) \in A_c$ ,  $(x, r) \in \pi(A_c)$  and (4.7) is shown. To prove (ii), let  $a, b$  be constants satisfying  $0 < a < b$  and let  $S$  denote the solution set of  $VI(F, P)$ . From part (i), and Proposition 3.1(iv)-(v), we see that  $\theta_{ab}^\varphi$  is a continuous subanalytic function satisfying

$$S = (d(\cdot, S))^{-1}(0) = (\theta_{ab}^\varphi)^{-1}(0). \quad (4.9)$$

In particular,  $S$  is a subanalytic subset of  $\mathbb{R}^n$  (by (S6)) and hence  $d(\cdot, S)$  is a continuous subanalytic function on  $\mathbb{R}^n$  (by (S4)). Fixing any  $\epsilon > 0$ , by Lemma 4.1, we have  $K := [\theta_{ab}^\varphi \leq \epsilon]$  is bounded hence compact. Denote the restriction of  $\theta_{ab}^\varphi$  (resp.  $d(\cdot, S)$ ) on  $K$  by  $\theta_{ab}^\varphi|_K$  (resp.  $d(\cdot, S)|_K$ ). Applying Łojasiewicz's inequality (S7) with  $\{\theta_{ab}^\varphi|_K, d(\cdot, S)|_K\}$  in place of  $\{f, g\}$  and using (4.9), we obtain a constant  $\alpha > 0$  and a positive integer  $N$  such that  $\alpha(d(x, S))^N \leq \theta_{ab}^\varphi(x)$  for all  $x \in K$ . Thus  $d(x, S) \leq \tau(\theta_{ab}^\varphi(x))^\gamma$  for all  $x \in K$ , where  $\tau = \alpha^{-\frac{1}{N}}$  and  $\gamma = \frac{1}{N} \in (0, 1]$ . This establishes part (ii) and completes the proof.  $\square$

Below, we present an example to show that the condition “ $F$  is coercive” in Theorem 4.1(ii) cannot be dropped.

**Example 4.1** Consider  $P = \mathbb{R}$ ,  $\varphi(x, y) = \frac{1}{2}\|x - y\|^2$ ,  $b = 2$ ,  $a = 1$  and  $F : \mathbb{R} \rightarrow \mathbb{R}$  is defined by  $F(x) = xe^x$ . It is clear that  $P$  is a subanalytic set, and  $F, \varphi$  are analytic (and hence subanalytic) functions. Moreover, by Remark 3.2 and (3.8), we have

$$y_c^\varphi(x) = x - c^{-1}F(x), \theta_c^\varphi(x) = \frac{\|F(x)\|^2}{2c} \text{ and } \theta_{ab}^\varphi(x) = \frac{b-a}{2ab}\|F(x)\|^2. \quad (4.10)$$

In particular, we have  $\theta_{ab}^\varphi(x) = \frac{1}{4}\|F(x)\|^2 = \frac{(xe^x)^2}{4}$  and  $(\theta_{ab}^\varphi)^{-1}(0) = \{0\}$ . Letting  $S := (\theta_{ab}^\varphi)^{-1}(0)$  and  $x_k = -k$  ( $k \in \mathbb{N}$ ), it follows that for any  $\gamma \in (0, 1]$ ,

$$\lim_{k \rightarrow \infty} \frac{d(x_k, S)}{(\theta_{ab}^\varphi(x_k))^\gamma} = \lim_{k \rightarrow \infty} \frac{k}{(\frac{1}{4}(ke^{-k})^2)^\gamma} = +\infty.$$

Since  $\theta_{ab}^\varphi(x_k) \rightarrow 0$ , this implies that  $(\theta_{ab}^\varphi)^\gamma$  fails to have a local error bound on  $\mathbb{R}$  for any  $\gamma \in (0, 1]$ . By Theorem 4.1(ii),  $F$  is not coercive (indeed

$$\liminf_{\|x\| \rightarrow \infty} \frac{F(x)^T x}{\|x\|} = \liminf_{|x| \rightarrow \infty} |x|e^x = 0).$$

As explained in [17], the exponent  $\gamma$  in Łojasiewicz's inequality (S7) (hence in Theorem 4.1) is, in general difficult to determine. Our next result gives a local error bound result with the exponent  $\gamma$  explicitly determined. Recall that the lower Hadamard directional derivative  $D^-f(x; v)$  of  $f$  at  $x \in \mathbb{R}^n$  in the direction  $v$  is defined by

$$D^-f(x; v) = \liminf_{t \downarrow 0, u \rightarrow v} \frac{f(x + tu) - f(x)}{t}. \quad (4.11)$$

It is clear that for all locally Lipschitz functions  $f$  and for all  $x, v \in \mathbb{R}^n$ ,

$$D^-f(x; v) \leq f^\circ(x; v). \quad (4.12)$$

In order to achieve our key result (Theorem 4.2), we first present several lemmas. Part (i) of the following lemma is due to [22].

**Lemma 4.2** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be a proper lower semicontinuous function. Let  $\gamma \in (0, 1]$ ,  $\epsilon > 0$  and  $\delta > 0$ . Let  $S_f := \{x : f(x) \leq 0\} \neq \emptyset$ . Then the following statements hold:

(i) Suppose that for each  $x \in \text{dom}(f) \setminus S_f$ , there exists a sequence  $\{x_n\}$  in  $\mathbb{R}^n$  such that  $x_n \rightarrow x$  and

$$0 < \|x - x_n\| \leq \delta (f^\gamma(x) - f^\gamma(x_n)) \text{ for all large } n.$$

Then  $f^\gamma$  has a global error bound with modulus  $\delta$  on  $\mathbb{R}^n$ .

(ii) Suppose that  $f$  is continuous and for each  $x \in [f \leq \epsilon] \setminus S_f$  there exists  $h_x \in \mathbb{R}^n$  with  $\|h_x\| = 1$  such that

$$D^- f(x; h_x) \leq -\delta f^{1-\gamma}(x). \quad (4.13)$$

Then  $f^\gamma$  has a global error bound with modulus  $1/(\gamma\delta)$  on  $[f \leq \epsilon]$ .

**Proof.** Part (i) is proved in [22] as noted. To prove (ii), define  $\tilde{f}$  by

$$\tilde{f}(x) = \begin{cases} f(x), & \text{if } x \in [f \leq \epsilon], \\ +\infty, & \text{else.} \end{cases} \quad (4.14)$$

Then  $\tilde{f}$  is a proper lower semicontinuous function with  $\text{dom} \tilde{f} = [f \leq \epsilon]$ . By (i), it suffices to show that for any  $\lambda > 1$  and  $x \in [f \leq \epsilon] \setminus S_f$  there exists a sequence  $\{x_n\}$  in  $\mathbb{R}^n$  such that  $x_n \rightarrow x$  and

$$0 < \|x - x_n\| \leq \frac{\lambda}{\gamma\delta} (\tilde{f}^\gamma(x) - \tilde{f}^\gamma(x_n)) \text{ for all large } n. \quad (4.15)$$

To do this, let us fix  $\lambda > 1$  and  $x \in [f \leq \epsilon] \setminus S_f$ . By assumption, there exist  $h_x \in \mathbb{R}^n$  with  $\|h_x\| = 1$  and two sequences  $t_n \downarrow 0$ ,  $u_n \rightarrow h_x$  such that

$$\lim_{n \rightarrow \infty} \frac{f(x_n) - f(x)}{t_n} = D^- f(x; h_x) \leq -\delta f^{1-\gamma}(x) < 0, \quad (4.16)$$

where  $x_n = x + t_n u_n$ . By the continuity of  $f$ , and since  $f(x) > 0$  and  $x_n \rightarrow x$ , we may assume without loss of generality that  $f(x_n) > 0$  for all  $n \in \mathbb{N}$ . Writing

$$f^\gamma(x_n) - f^\gamma(x) = \gamma f^{\gamma-1}(x) (f(x_n) - f(x)) + \alpha_n,$$

where  $\alpha_n = o(f(x_n) - f(x))$  and assuming that  $D^- f(x; h_x) > -\infty$ , we have

$$\lim_{n \rightarrow \infty} \frac{\alpha_n}{t_n} = \lim_{n \rightarrow \infty} \left( \frac{f(x_n) - f(x)}{t_n} \cdot \frac{\alpha_n}{f(x_n) - f(x)} \right) = 0.$$

This together with (4.16) implies that

$$\lim_{n \rightarrow \infty} \frac{f^\gamma(x_n) - f^\gamma(x)}{t_n} = \gamma f(x)^{\gamma-1} D^- f(x; h_x) \leq -\gamma\delta \quad (4.17)$$

(one can show similarly that (4.17) is valid also for the case when  $D^- f(x; h_x) = -\infty$ ). Noting that  $u_n \rightarrow h_x$  and  $\|h_x\| = 1$ , it follows that

$$\lim_{n \rightarrow \infty} \frac{f^\gamma(x_n) - f^\gamma(x)}{\|x_n - x\|} \leq -\gamma\delta < -\frac{\gamma\delta}{\lambda}.$$

Thus, for all large  $n$ ,

$$f^\gamma(x) - f^\gamma(x_n) > \frac{\gamma\delta}{\lambda} \|x_n - x\|,$$

(in particular  $f^\gamma(x_n) < f^\gamma(x) \leq \epsilon^\gamma$  and so  $x_n \in \text{dom} \tilde{f}$ ). Therefore (4.15) is established and this completes the proof.  $\square$

**Lemma 4.3** *Let  $b > a > 0$ . Then for all  $x \in \mathbb{R}^n$ , we have*

$$(\nabla_y \varphi(x, y_a^\varphi(x)) - \nabla_y \varphi(x, y_b^\varphi(x)))^T (y_a^\varphi(x) - y_b^\varphi(x)) \geq 2\beta \|y_a^\varphi(x) - y_b^\varphi(x)\|^2. \quad (4.18)$$

**Proof.** Considering  $y_a^\varphi(x), y_b^\varphi(x)$  in place of  $y_1, y_2$  in (3.4), we have

$$\varphi(x, y_a^\varphi(x)) - \varphi(x, y_b^\varphi(x)) \geq \nabla_y \varphi(x, y_b^\varphi(x))^T (y_a^\varphi(x) - y_b^\varphi(x)) + \beta \|y_a^\varphi(x) - y_b^\varphi(x)\|^2.$$

Similarly,

$$\varphi(x, y_b^\varphi(x)) - \varphi(x, y_a^\varphi(x)) \geq \nabla_y \varphi(x, y_a^\varphi(x))^T (y_b^\varphi(x) - y_a^\varphi(x)) + \beta \|y_a^\varphi(x) - y_b^\varphi(x)\|^2,$$

Summing these two inequalities gives (4.18).  $\square$

Let  $T(a, P)$  and  $N(a, P)$  respectively denote the tangent cone and the normal cone of  $P$  at  $a \in P$ , that is  $T(a, P) = \overline{\bigcup_{t>0} t(P - a)}$  and  $N(a, P) = (T(a, P))^\circ = \{x : x^T d \leq 0, \text{ for all } d \in T(a, P)\}$ . It is known (and easy to verify) that  $N(a, P) = \{x : x^T(p - a) \leq 0 \text{ for all } p \in P\}$ . Let  $(F(x))^\circ$  denote the negative polar of the singleton  $\{F(x)\}$ , i.e.  $(F(x))^\circ = \{d \in \mathbb{R}^n : F(x)^T d \leq 0\}$ . For a vector  $x \in \mathbb{R}^n$ , following [9], we define  $T_{ab}(x, P)$  and  $T_{ab}(x, F, P)$  by

$$T_{ab}(x, P) = \left( T(y_b^\varphi(x), P) \right) \cap \left( -T(y_a^\varphi(x), P) \right) \quad \text{and} \quad T_{ab}(x, F, P) = T_{ab}(x, P) \cap (F(x))^\circ.$$

**Remark 4.1** *Clearly,  $T_{ab}(x, P)$  and  $T_{ab}(x, F, P)$  are closed convex cones. Moreover, it is easy to verify that*

$$y_a^\varphi(x) - y_b^\varphi(x) \in T_{ab}(x, P) \quad (4.19)$$

(because  $y_a^\varphi(x), y_b^\varphi(x)$  belong to the convex set  $P$ ).

**Lemma 4.4** *Let  $b > a > 0$  and  $x \in \mathbb{R}^n$ . Then  $y_a^\varphi(x) - y_b^\varphi(x) \in T_{ab}(x, F, P)$  and*

$$\left( a \nabla_x \varphi(x, y_a^\varphi(x)) - b \nabla_x \varphi(x, y_b^\varphi(x)) \right)^T \left( y_a^\varphi(x) - y_b^\varphi(x) \right) \geq 0. \quad (4.20)$$

**Proof.** Since  $y_a^\varphi(x)$  minimizes the function  $-F(x)^T(x - \cdot) + a\varphi(x, \cdot)$  on  $P$ , the first order optimality condition implies that

$$\left( F(x) + a \nabla_y \varphi(x, y_a^\varphi(x)) \right)^T \left( y - y_a^\varphi(x) \right) \geq 0 \text{ for all } y \in P.$$

Letting  $y = y_b^\varphi(x)$ , we obtain

$$-\left( F(x) + a \nabla_y \varphi(x, y_a^\varphi(x)) \right)^T \left( y_a^\varphi(x) - y_b^\varphi(x) \right) \geq 0. \quad (4.21)$$

Interchanging the roles of  $a, b$  we similarly obtain

$$\left(F(x) + b\nabla_y\varphi(x, y_b^\varphi(x))\right)^T \left(y_a^\varphi(x) - y_b^\varphi(x)\right) \geq 0. \quad (4.22)$$

By (P5), adding the above two inequalities gives (4.20). It remains to show that  $y_a^\varphi(x) - y_b^\varphi(x) \in T_{ab}(x, F, P)$ . Let  $y_a^\varphi(x) - y_b^\varphi(x)$  be denoted by  $z$ . By (4.19), it suffices to show that  $z \in F(x)^\circ$ . Noting that

$$\begin{aligned} (F(x) + a\nabla_y\varphi(x, y_a^\varphi(x)))^T z &= \left(1 - \frac{a}{b}\right)F(x)^T z + \frac{a}{b}(F(x)^T + b\nabla_y\varphi(x, y_b^\varphi(x)))^T z \\ &\quad + a(\nabla_y\varphi(x, y_a^\varphi(x)) - \nabla_y\varphi(x, y_b^\varphi(x)))^T z, \end{aligned}$$

where the last two terms are nonnegative (thanks to (4.22) and (4.18)), it follows from (4.21) that  $0 \geq \left(1 - \frac{a}{b}\right)F(x)^T z$ . Since  $b > a > 0$ , this implies that  $F(x)^T z \leq 0$  and so  $z \in F(x)^\circ$  as required to show.  $\square$

For any constants  $a, b, c$  satisfying  $0 < a < b$  and  $c > 0$ . Define a multifunction  $\Omega_{abc} : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$  as follows:

$$\Omega_{abc}(x) := \begin{cases} y_a^\varphi(x) - y_b^\varphi(x), & \text{if } c\|x - y_a^\varphi(x)\| < \|y_a^\varphi(x) - y_b^\varphi(x)\|, \\ \{y_a^\varphi(x) - x, y_a^\varphi(x) - y_b^\varphi(x)\} & \text{if } c\|x - y_a^\varphi(x)\| = \|y_a^\varphi(x) - y_b^\varphi(x)\|, \\ y_a^\varphi(x) - x, & \text{if } c\|x - y_a^\varphi(x)\| > \|y_a^\varphi(x) - y_b^\varphi(x)\|. \end{cases} \quad (4.23)$$

We are now ready to establish our main result of this section. The first assertion of (i) in the following theorem follows immediately from Proposition 3.1(viii) and Lemma 4.1.

**Theorem 4.2** *Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be locally Lipschitz and coercive on  $\mathbb{R}^n$ . Suppose that there exist positive constants  $a, b$  with  $a < b$  such that  $\mu_{ab} > 0$ , where  $\mu_{ab}$  is defined by*

$$\mu_{ab} := \inf\{d^T V d : V \in \partial^c F(x), d \in T_{ab}(x, F, P), \|d\| = 1, \theta_{ab}^\varphi(x) \neq 0\} \quad (4.24)$$

(with the convention that  $\inf \emptyset = +\infty$ ). Then the following assertions hold.

(i) *Let  $\epsilon$  be any positive number. Then  $\theta_{ab}^\varphi$  is Lipschitz on  $[\theta_{ab}^\varphi \leq 2\epsilon]$  (with some Lipschitz constant  $L$ ). Let  $c, \delta$  be positive real numbers satisfying*

$$c \leq \min\left\{1, \frac{(b-a)\beta}{2(L+b\kappa)}\right\} \text{ and } \delta \leq \min\left\{\frac{(b-a)\beta}{2\sqrt{(\kappa-\beta)(b-a)}}, \frac{c\mu_{ab}}{\sqrt{(\kappa-\beta)(b-a)}}\right\} \quad (4.25)$$

where  $\beta, \kappa$  are defined as in (3.4) and (3.5) and let  $\Omega_{abc}$  be defined as in (4.23). For each  $x \in (\theta_{ab}^\varphi)^{-1}(0, \epsilon]$ , we have  $0 \notin \Omega_{abc}(x)$  and

$$(\theta_{ab}^\varphi)^\circ(x; h_x) \leq -\delta\sqrt{\theta_{ab}^\varphi(x)} \text{ for each } h_x = \frac{w_x}{\|w_x\|} \text{ with } w_x \in \Omega_{abc}(x). \quad (4.26)$$

(ii) *Let  $\epsilon, L$  and  $\delta$  be as in (i). Then  $d(x, [\theta_{ab}^\varphi \leq 0]) \leq \frac{2}{\delta}\sqrt{\theta_{ab}^\varphi}$  for all  $x \in [\theta_{ab}^\varphi \leq \epsilon]$ .*

(iii)  $\sqrt{\theta_{ab}^\varphi}$  has a local error bound on  $\mathbb{R}^n$ .

**Proof.** The implication (ii)  $\Rightarrow$  (iii) is by the definitions. Let  $\epsilon > 0$ . By (4.12), (4.26) entails that  $D^-\theta_{ab}^\varphi(x; h_x) \leq -\delta\sqrt{\theta_{ab}^\varphi(x)}$  for all  $x \in (\theta_{ab}^\varphi)^{-1}(0, \epsilon]$ . Hence, by virtue of Lemma 4.2(ii) (applied to  $\gamma = 1/2$  and  $\theta_{ab}^\varphi$  in place of  $f$ ), (i) implies that  $\sqrt{\theta_{ab}^\varphi}$  has an error bound with modulus  $2/\delta$  on  $[\theta_{ab}^\varphi \leq \epsilon]$ . Thus the assertion (ii) follows from (i). For (i), we only need to verify the second assertion since the first one holds as we already noted. Let us fix  $x$  with  $0 < \theta_{ab}^\varphi(x) \leq \epsilon$ ,  $w_x \in \Omega_{abc}(x)$ . Then, by Proposition 3.1(iii),  $0 \notin \Omega_{abc}(x)$  and hence  $w_x$  is nonzero. Let  $h_x = w_x/\|w_x\|$ . Note that the inequality in (4.26) is equivalent to

$$(\theta_{ab}^\varphi)^\circ(x; w_x) \leq -\delta\|w_x\|\sqrt{\theta_{ab}^\varphi(x)}. \quad (4.27)$$

Thus by [5, Proposition 2.1.2 (b)], it is sufficient to show that

$$W^T w_x \leq -\delta\|w_x\|\sqrt{\theta_{ab}^\varphi(x)}, \text{ for all } W \in \partial^c \theta_{ab}^\varphi(x). \quad (4.28)$$

To do this, let us fix  $W \in \partial^c \theta_{ab}^\varphi(x)$  and make use of (3.12) to express  $W$  in the form

$$W = -V^T z - a\nabla_x \varphi(x, y_a^\varphi(x)) + b\nabla_x \varphi(x, y_b^\varphi(x)) \quad (4.29)$$

where  $V \in \partial^c F(x)$  and  $z = y_a^\varphi(x) - y_b^\varphi(x)$ . Note that  $x \neq y_a^\varphi(x)$  since  $\theta_{ab}^\varphi(x) \neq 0$  (see Proposition 3.1(iii) and (v)). Moreover if  $T_{ab}(x, F, P) = \{0\}$  then  $y_a^\varphi(x) - y_b^\varphi(x) = 0$  (see Lemma 4.4), and hence  $c\|x - y_a^\varphi(x)\| > 0 = \|y_a^\varphi(x) - y_b^\varphi(x)\|$ . Therefore, one of the following cases must hold.

- (1 $^\circ$ )  $c\|x - y_a^\varphi(x)\| < \|y_a^\varphi(x) - y_b^\varphi(x)\|$  and  $T_{ab}(x, F, P) \neq \{0\}$ ;
- (2 $^\circ$ )  $c\|x - y_a^\varphi(x)\| > \|y_a^\varphi(x) - y_b^\varphi(x)\|$ ;
- (3 $^\circ$ )  $c\|x - y_a^\varphi(x)\| = \|y_a^\varphi(x) - y_b^\varphi(x)\|$  and  $T_{ab}(x, F, P) \neq \{0\}$ .

Suppose (1 $^\circ$ ) holds. Then, since  $w_x \in \Omega_{abc}(x)$ , (4.23) entails that  $w_x = z$  and it follows from (3.10) that

$$c\sqrt{\theta_{ab}^\varphi(x)} < \sqrt{(\kappa - \beta)(b - a)}\|z\|. \quad (4.30)$$

Thus, by the definition of  $\delta$  given in (4.25), to show (4.28) it is sufficient to show that

$$W^T z \leq -\mu_{ab}\|z\|^2, \text{ for all } W \in \partial^c \theta_{ab}^\varphi(x). \quad (4.31)$$

On the other hand, from (4.29) and (4.20) we obtain

$$W^T z \leq -z^T V z. \quad (4.32)$$

Since  $z \in T_{ab}(x, F, P) \setminus \{0\}$  (by Lemma 4.4 and  $z = w_x \neq 0$ ), it follows from (4.24) that  $-W^T z \geq \mu_{ab}\|z\|^2$ ; thus (4.31) is valid. Therefore (4.28) is verified for case (1 $^\circ$ ). Suppose next that (2 $^\circ$ ) holds. Define  $u = y_a^\varphi(x) - x$ . Then  $u = w_x \neq 0$  and

$$\|z\| < c\|u\|. \quad (4.33)$$

Thus, to show (4.28), it is sufficient to prove that

$$W^T u \leq -\delta\|u\|\sqrt{\theta_{ab}^\varphi(x)} \text{ for all } W \in \partial^c \theta_{ab}^\varphi(x). \quad (4.34)$$

From (4.29), we note that

$$W^T u = -z^T V u + \left( -a \nabla_x \varphi(x, y_a^\varphi(x)) + b \nabla_x \varphi(x, y_b^\varphi(x)) \right)^T u. \quad (4.35)$$

We claim that the following two statements hold:

$$-z^T V u < cL \|u\|^2, \quad (4.36)$$

$$\left( -a \nabla_x \varphi(x, y_a^\varphi(x)) + b \nabla_x \varphi(x, y_b^\varphi(x)) \right)^T u < \left( -(b-a)\beta + b\kappa c \right) \|u\|^2. \quad (4.37)$$

Granting this, it follows from (4.35) and the definition of  $c$  that

$$W^T u < \left( -(b-a)\beta + (L + b\kappa)c \right) \|u\|^2 \leq -(b-a)\beta \|u\|^2/2.$$

Consequently (4.34) is seen to hold since, by (3.10) and the definition of  $\delta$  one has

$$\begin{aligned} W^T u < -\frac{(b-a)\beta \|u\|^2}{2} &\leq -\frac{(b-a)\beta}{2\sqrt{(\kappa-\beta)(b-a)}} \|u\| \sqrt{\theta_{ab}^\varphi(x)} \\ &\leq -\delta \|u\| \sqrt{\theta_{ab}^\varphi(x)}. \end{aligned} \quad (4.38)$$

To see (4.36), we recall that  $F$  is Lipschitz with modulus  $L$  on  $[\theta_{ab}^\varphi \leq 2\epsilon]$  hence, by (2.4), we get that  $\|V\| \leq L$ , for all  $x \in [\theta_{ab}^\varphi < 2\epsilon]$  and  $V \in \partial^c F(x)$ . This together with the Cauchy-Schwartz inequality and (4.33) gives that

$$-z^T V u \leq L \|z\| \|u\| < cL \|u\|^2. \quad (4.39)$$

Therefore, (4.36) holds. To prove (4.37), from (P2) and (3.4), we have

$$-\varphi(x, y_a^\varphi(x)) = \varphi(x, x) - \varphi(x, y_a^\varphi(x)) \geq -\nabla_y \varphi(x, y_a^\varphi(x))^T u + \beta \|u\|^2.$$

This together with (P2) and (P5) in the definition of  $\varphi$  yields that

$$\nabla_x \varphi(x, y_2)^T u = -\nabla_y \varphi(x, y_2)^T u \leq -\beta \|u\|^2 - \varphi(x, y_a^\varphi(x)) \leq -\beta \|u\|^2. \quad (4.40)$$

By (P5), (3.5), the Cauchy-Schwartz inequality and (4.33), we have

$$\begin{aligned} \left( \nabla_x \varphi(x, y_b^\varphi(x)) - \nabla_x \varphi(x, y_a^\varphi(x)) \right)^T u &= \left( -\nabla_y \varphi(x, y_b^\varphi(x)) + \nabla_y \varphi(x, y_a^\varphi(x)) \right)^T u \\ &\leq \kappa \|z\| \|u\| < \kappa c \|u\|^2, \end{aligned}$$

and it follows from (4.40) that

$$\begin{aligned} \left( -a \nabla_x \varphi(x, y_a^\varphi(x)) + b \nabla_x \varphi(x, y_b^\varphi(x)) \right)^T u &= (b-a) \left( \nabla_x \varphi(x, y_a^\varphi(x)) \right)^T u \\ &\quad + b \left( \nabla_x \varphi(x, y_b^\varphi(x)) - \nabla_x \varphi(x, y_a^\varphi(x)) \right)^T u \\ &< \left( -(b-a)\beta + b\kappa c \right) \|u\|^2. \end{aligned}$$

Thus (4.37) holds and this verifies (4.28) for case (2°). Finally, suppose (3°) holds. Then since  $w_x \in \Omega_{abc}(x)$ , (4.23) entails that  $w_x = z$  or  $w_x = u$ . If  $w_x = z$ , then  $\|w_x\| = \|z\| = c\|u\|$  by (3°). By (3.10), it follows that

$$c\sqrt{\theta_{ab}^\varphi(x)} \leq c\sqrt{(\kappa - \beta)(b - a)}\|u\| = \sqrt{(\kappa - \beta)(b - a)}\|z\|. \quad (4.41)$$

We proceed as in (1°) and obtain that

$$W^T z \leq -\mu_{ab}\|z\|^2, \text{ for all } W \in \partial^c \theta_{ab}^\varphi(x).$$

This together with (4.41) and the definition of  $\delta$  implies that

$$W^T z \leq -\frac{\delta\sqrt{(\kappa - \beta)(b - a)}}{c}\|z\|^2 \leq -\delta\sqrt{\theta_{ab}^\varphi(x)}\|z\|,$$

and hence (4.28) holds in the case when (3°) holds and  $w_x = z$ . If  $w_x = u$ , proceed as in (2°), we see that (4.36) and (4.37) hold but replacing the strict inequalities by the corresponding non-strict inequalities. Thus one deduces that (similar to the derivation of (4.38))

$$W^T u \leq -\delta\|u\|\sqrt{\theta_{ab}^\varphi(x)}.$$

Therefore (4.28) holds when (3°) holds and  $w_x = u$ . This completes the proof.  $\square$

**Corollary 4.1** *Let  $b > a > 0$  and let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be locally Lipschitz and strongly monotone on  $\mathbb{R}^n$ . Then the function  $\sqrt{\theta_{ab}^\varphi}$  has a local error bound on  $\mathbb{R}^n$ .*

**Proof.** The assumptions imply that  $F$  is coercive (by Remark 2.1(ii)) and that  $\mu_{ab} > 0$  holds (by (2.7)), where  $\mu_{ab}$  is defined as in (4.24). Therefore, the conclusion follows from the preceding Theorem.  $\square$

**Remark 4.2** *The assumption  $\mu_{ab} > 0$  in Theorem 4.2 does not imply the solution is unique. For example consider  $P = \mathbb{R}$ ,  $\varphi(x, y) = \frac{1}{2}\|x - y\|^2$ ,  $a = 1$ ,  $b = 2$  and  $F : \mathbb{R} \rightarrow \mathbb{R}$  is defined by  $F(x) = \max\{x, 0\}$ . Note that  $\theta_{ab}^\varphi(x) = \frac{b-a}{2ab}\|F(x)\|^2 = \max\{x, 0\}^2/4$  (by (4.10)). The solution set  $S = [\theta_{ab}^\varphi = 0] = \mathbb{R}_-$ . On the other hand, since  $\{x : \theta_{ab}^\varphi(x) \neq 0\} = (0, +\infty)$ , it follows from (4.24) that*

$$\begin{aligned} \mu_{ab} &= \inf\{d^T V d : V \in \partial^c F(x), d \in T_{ab}(x, F, P), \|d\| = 1, \theta_{ab}^\varphi(x) \neq 0\} \\ &\geq \inf\{d^T V d : V \in \partial^c F(x), \|d\| = 1, x \in (0, +\infty)\} = 1. \end{aligned}$$

**Remark 4.3** *Our error bound results depend on the constant  $\mu_{ab}$ . Therefore, in general, we need some priori information to determine the constant  $\mu_{ab}$ . However, it can be determined explicitly in the following two important cases:*

- (i) *If  $F$  is strongly monotone with modulus  $\mu > 0$ . Then we can take  $\mu_{ab} = \mu$ .*
- (ii) *If  $P = \mathbb{R}_+^n$  and  $\varphi(x, y) = \frac{1}{2}\|x - y\|^2$ . Then, the cone  $T_{ab}(x, F, P)$  (and hence the constant  $\mu_{ab}$ ) can be determined explicitly (see [9, page 941-943]).*

Recall that, for a locally Lipschitz function  $f$ , we say  $x$  is a generalized unconstrained stationary point of  $f$  if  $0 \in \partial^c f(x)$ . We now summarize the connection between the stationary points of the D-gap function and the solution of the corresponding  $VI(F, P)$  in the following theorem.

**Theorem 4.3** Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be locally Lipschitz continuous on  $\mathbb{R}^n$ . Let  $a$  and  $b$  are scalars satisfying  $b > a > 0$ . Consider the following statements:

(i)  $x$  is a solution of  $VI(F, P)$ .

(ii)  $x$  is a generalized unconstrained stationary point of  $\theta_{ab}^\varphi$  and the following relation holds:

$$\left. \begin{array}{l} d \in T_{ab}(x, F, P) \\ V^T d \in T_{ab}(x, F, P)^\circ \forall V \in \partial^c F(x) \end{array} \right\} \Rightarrow F(x)^T d = 0. \quad (4.42)$$

(iii)  $x$  is a generalized unconstrained stationary point of  $\theta_{ab}^\varphi$  and  $\mu_{ab} > 0$  where  $\mu_{ab}$  is defined as in (4.24).

Then the following implication holds: (iii)  $\Rightarrow$  (ii)  $\Leftrightarrow$  (i).

**Proof.** The proof of the equivalence of (i) and (ii) is similar to [9, Theorem 10.3.4]. To see (iii)  $\Rightarrow$  (ii), let statement (iii) holds. We proceed by contradiction and suppose that the statement (ii) doesn't hold. Thus there exists  $d \in T_{ab}(x, F, P)$ ,  $V \in \partial^c F(x)$  and  $V^T d \in T_{ab}(x, F, P)^\circ$  satisfying  $d^T F(x) \neq 0$ . In particular, one has  $d \neq 0$  and

$$d^T V d \leq 0. \quad (4.43)$$

Moreover, since (i)  $\Leftrightarrow$  (ii), we may assume without loss of generality that  $x$  is not a solution of  $VI(F, P)$ . It follows from the definition of  $\mu_{ab}$  that  $d^T V d \geq \mu_{ab} \|d\|^2$ . This together with (4.43) and  $\mu_{ab} > 0$  implies that  $d = 0$ . This makes contradiction and finishes the proof.  $\square$

Below, we present three examples. The first example shows that the condition " $\mu_{ab} > 0$ " in Theorem 4.2 cannot be dropped and the second/third one shows that our Theorem 4.1 and 4.2 can be applied to some cases when  $F$  is nonsmooth/nonmonotone.

**Example 4.2** Consider  $P = \mathbb{R}$ ,  $\varphi(x, y) = \frac{1}{2} \|x - y\|^2$ ,  $a = 1$ ,  $b = 2$  and  $F : \mathbb{R} \rightarrow \mathbb{R}$  is defined by  $F(x) = x^3$ . Then as in Example 4.1,  $P, F$  and  $\varphi$  are subanalytic, and (4.10) holds. Moreover,  $F$  is clearly coercive, and Theorem 4.1(ii) implies that there exists some  $\gamma \in (0, 1]$  such that  $(\theta_{ab}^\varphi)^\gamma$  has a local error bound on  $\mathbb{R}$ . (Indeed, by (4.10),  $\theta_{ab}^\varphi(x) = \frac{b-a}{2ab} \|F(x)\|^2 = x^6/4$  and so  $(\theta_{ab}^\varphi)^{1/6}$  has a local error bound on  $\mathbb{R}$ .) Let  $x_k = 1/k$  ( $k \in \mathbb{N}$ ). Then  $\theta_{ab}^\varphi(x_k) \rightarrow 0$  and

$$\frac{d(x_k, [\theta_{ab}^\varphi = 0])}{\sqrt{\theta_{ab}^\varphi(x_k)}} = \frac{1/k}{1/(2k^3)} \rightarrow +\infty.$$

Thus  $\sqrt{\theta_{ab}^\varphi}$  has no local error bound on  $\mathbb{R}$  (by Theorem 4.2, it follows that  $\mu_{ab} = 0$ ).

**Example 4.3** Let  $\varphi(x, y) = \frac{1}{2} \|x - y\|^2$ ,  $P = \mathbb{R}_+^n$ ,  $a \in (0, 1)$ ,  $b = 1$ . Consider the nonsmooth nonlinear complementary problem discussed in Example 1.1 with  $\phi(u) = \max\{u, u^{2m+1}\}$  ( $m \in \mathbb{N}$ ). In this case, one has  $F(x) = Mx + p(x) + q$  where  $M = (M_{i,j})_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n}$  with  $M_{i,i} = 2$  and  $M_{i,i+1} = M_{i+1,i} = -1$ ,  $q \in \mathbb{R}^n$  and  $p$  is defined by

$$p(x_1, \dots, x_n) = (\max\{x_1, x_1^{2m+1}, 0\}, \dots, \max\{x_n, x_n^{2m+1}, 0\}).$$

It can be verified that  $F$  is locally Lipschitz and coercive. Moreover, since  $M$  is positive definite, for all  $d, x \in \mathbb{R}^n$ ,  $V \in \partial^c F(x)$ ,

$$d^T V d \geq d^T M d \geq \alpha \|d\|^2,$$

where  $\alpha$  is some positive number. Thus, from Theorem 4.2,  $\sqrt{\theta_{ab}^\varphi}$  has a local error bound on  $\mathbb{R}^n$ .

**Example 4.4** Consider  $\varphi(x, y) = \frac{1}{2}\|x - y\|^2$ ,  $P = \mathbb{R}_+^2$ ,  $a \in (0, 1)$ ,  $b = 1$  and  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is defined by  $F(x_1, x_2) = (F_1(x_1, x_2), F_2(x_1, x_2))$  where

$$F_1(x_1, x_2) = x_1 + \max\{x_1, 0\} \cdot \max\{x_2, 0\} \quad \text{and} \quad F_2(x_1, x_2) = x_2 + \frac{3}{2} \max\{x_1, 0\}, \quad (4.44)$$

that is,

$$F_1(x_1, x_2) = \begin{cases} x_1 + x_1x_2, & \text{if } x_1 \geq 0, x_2 \geq 0 \\ x_1, & \text{else,} \end{cases} \quad \text{and} \quad F_2(x_1, x_2) = \begin{cases} \frac{3x_1}{2} + x_2, & \text{if } x_1 \geq 0, \\ x_2, & \text{else.} \end{cases} \quad (4.45)$$

Thus  $F$  is subanalytic (see (S5)). Moreover, we have

$$\nabla F(x) = \begin{cases} \begin{pmatrix} 1 + x_2 & x_1 \\ \frac{3}{2} & 1 \end{pmatrix}, & \text{if } x_1 > 0 \text{ and } x_2 > 0. \\ \begin{pmatrix} 1 & 0 \\ \frac{3}{2} & 1 \end{pmatrix}, & \text{if } x_1 > 0 \text{ and } x_2 < 0, \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \text{if } x_1 < 0 \text{ and } x_2 > 0, \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \text{if } x_1 < 0 \text{ and } x_2 < 0. \end{cases} \quad (4.46)$$

Since  $F$  is monotone on an open convex set  $D$  if and only if  $V$  is positive semidefinite for all  $V \in \partial^B F(x)$  and  $x \in D$  (see [14, Proposition 2.3(a)]), and  $\nabla F(1, \frac{1}{4})$  is negative definite. Thus,  $F$  is not monotone on  $\mathbb{R}^2$ . Note that the corresponding VI( $F, P$ ) reduces to the following nonlinear complementary problem: find  $x = (x_1, x_2) \in \mathbb{R}_+^2$  such that

$$F_i(x) \geq 0 \text{ and } x_i F_i(x) = 0 \text{ for all } i = 1, 2.$$

Clearly the solution set  $S$  of this problem is the singleton  $\{(0, 0)\}$ . Moreover, by the following elementary inequalities,

$$\begin{aligned} x_1^2 + \frac{3}{2}x_1x_2 + x_1^2x_2 + x_2^2 &\geq x_1^2 + x_2^2 \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2 \text{ such that } x_1 \geq 0 \text{ and } x_2 \geq 0; \\ x_1^2 + \frac{3}{2}x_1x_2 + x_2^2 &\geq \frac{1}{4}(x_1^2 + x_2^2) \quad \text{for all } x_1, x_2 \in \mathbb{R}, \end{aligned}$$

it is easy to verify that

$$F(x)^T x \geq \frac{1}{4}\|x\|^2 \text{ for all } x \in \mathbb{R}^2. \quad (4.47)$$

Since  $\{\frac{F(x)^T y}{\|x\|} : x \in \mathbb{R}^2\}$  is bounded for any  $y \in \mathbb{R}^2$ , this implies that  $F$  is coercive. Thus one can apply Theorem 4.1(ii) to conclude that  $(\theta_{ab}^\varphi)^\gamma$  has a local error bound for some  $\gamma > 0$ . In fact, for the present case, Theorem 4.2 produces a better result by giving the explicit value  $\gamma = 1/2$ . To see this, it suffices to show  $\mu_{ab} > 0$ , where  $\mu_{ab}$  is defined as in (4.24). Since  $y_c^\varphi(x) = (\max\{0, x_1 - c^{-1}F_1(x)\}, \max\{0, x_2 - c^{-1}F_2(x)\})$  for any  $x := (x_1, x_2) \in \mathbb{R}^2$  (see Remark 3.2), it follows from (4.44) that  $y_b^\varphi(x) = (0, 0)$  since  $b = 1$ . Consequently,  $T(y_b^\varphi(x), P) = \mathbb{R}_+^2$  and so

$$T_{ab}(x, F, P) \subseteq \mathbb{R}_+^2. \quad (4.48)$$

Define a mapping  $M : \mathbb{R}^2 \rightarrow 2^{\mathbb{R}^n \times \mathbb{R}^n}$  as follows

$$M(x) = \begin{cases} \text{co}\left\{ \begin{pmatrix} 1+x_2 & x_1 \\ \frac{3}{2} & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \frac{3}{2} & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}, & x_1 \geq 0 \text{ and } x_2 \geq 0. \\ \text{co}\left\{ \begin{pmatrix} 1 & 0 \\ \frac{3}{2} & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}, & \text{else.} \end{cases} \quad (4.49)$$

It follows from (4.46), (2.2) and (2.3) that for any  $x \in \mathbb{R}^2$

$$\partial^c F(x) \subseteq M(x).$$

Thus  $d^T V d \geq 1$  for all  $V \in \partial^c F(x) \subseteq M(x)$  and for all  $d \in \mathbb{R}_+^2$  with  $\|d\| = 1$ . This together with (4.24) and (4.48) implies that  $\mu_{ab} \geq 1$ . In particular,  $\mu_{ab} > 0$  and hence  $\sqrt{\theta_{ab}^\varphi}$  has a local error bound on  $\mathbb{R}^2$ .

## 5 A Derivative Free Descent Method

In this section, we consider an Armijo type descent method. Denote the solution set of  $VI(F, P)$  and the initial point respectively by  $S$  and  $x_0$ . Throughout this section, we assume that  $F$  is locally Lipschitz, coercive and  $0 < a < b$  such that  $\mu_{ab} > 0$ , where  $\mu_{ab}$  is defined as in (4.24). Considering  $\epsilon = \theta_{ab}^\varphi(x_0)$ , it follows from Theorem 4.2 that there exist  $\delta, c > 0$  satisfying (4.25) such that, for each  $x \in (\theta_{ab}^\varphi)^{-1}(0, \epsilon]$  one has  $d(x, S) \leq \frac{2}{\delta} \sqrt{\theta_{ab}^\varphi(x)}$  and

$$(\theta_{ab}^\varphi)^\circ(x; h_x) \leq -\delta \sqrt{\theta_{ab}^\varphi(x)}, \quad (5.1)$$

where  $h_x = \frac{w_x}{\|w_x\|}$  and  $w_x$  is defined by:

$$w_x = \begin{cases} y_a^\varphi(x) - y_b^\varphi(x), & \text{if } c\|x - y_a^\varphi(x)\| \leq \|y_a^\varphi(x) - y_b^\varphi(x)\|, \\ y_a^\varphi(x) - x, & \text{else.} \end{cases} \quad (5.2)$$

Note from the definition of  $w_x$  and  $c$  that

$$\|w_x\| \geq c\|y_a^\varphi(x) - x\|. \quad (5.3)$$

Moreover, by the continuity of  $\theta_{ab}^\varphi(\cdot)$  and (5.1), we note that for all  $\rho \in (0, 1)$  and  $x \in (\theta_{ab}^\varphi)^{-1}(0, \epsilon]$

$$\begin{aligned} \limsup_{m \rightarrow \infty} \frac{\sqrt{\theta_{ab}^\varphi(x + \rho^m w_x)} - \sqrt{\theta_{ab}^\varphi(x)}}{\rho^m \|w_x\|} &= \limsup_{m \rightarrow \infty} \frac{\theta_{ab}^\varphi(x + \rho^m w_x) - \theta_{ab}^\varphi(x)}{\rho^m \|w_x\| [\sqrt{\theta_{ab}^\varphi(x + \rho^m w_x)} + \sqrt{\theta_{ab}^\varphi(x)}]} \\ &\leq \frac{(\theta_{ab}^\varphi)^\circ(x; \frac{w_x}{\|w_x\|})}{2\sqrt{\theta_{ab}^\varphi(x)}} \leq -\frac{\delta}{2}. \end{aligned} \quad (5.4)$$

Our algorithm is as follows:

### Algorithm

Step 0: Let  $\rho \in (0, 1)$ . Let  $x_0$  be a given vector in  $\mathbb{R}^n$ . Set  $k = 0$ .

Step 1: If  $\theta_{ab}^\varphi(x_k) = 0$  then stop (and  $x_k \in S$ ). If not then go to Step 2.

Step 2: Compute  $w_{x_k}$  by (5.2) and go to Step 3.

Step 3: Let  $m_k$  be the smallest integer  $m$  such that

$$\sqrt{\theta_{ab}^\varphi(x_k + \rho^m w_{x_k})} - \sqrt{\theta_{ab}^\varphi(x_k)} \leq -\frac{\delta}{4} \rho^m \|w_{x_k}\| \quad (5.5)$$

and set  $x_{k+1} = x_k + \rho^{m_k} w_{x_k}$ . Go to step 4.

Step 4: Return to Step 1 with  $k$  replaced by  $k + 1$ .

Note. By (5.4), the line search in Step 3 is well defined (unless  $x_k$  is already a solution of  $VI(F, P)$ ).

**Theorem 5.1** *The sequence  $\{x_k\}$  generated by the above algorithm is bounded and the limit of each of its convergent subsequences is a solution of  $VI(F, P)$ .*

**Proof.** If  $\theta_{ab}^\varphi(x_k) = 0$  then  $x_k \in S$  by Proposition 3.1(v). Suppose therefore that  $\theta_{ab}^\varphi(x_k) > 0$  (hence  $w_{x_k} \neq 0$  by (3.10) and (5.3)) for each  $k$ . Let  $\epsilon = \theta_{ab}^\varphi(x_0)$ . By (5.5), the sequence  $\{\theta_{ab}^\varphi(x_k)\}$  is decreasing and hence converges; thus  $\{x_k\}_{k=1}^\infty \subseteq (\theta_{ab}^\varphi)^{-1}(0, \epsilon]$  and

$$\lim_{k \rightarrow \infty} \rho^{m_k} \|w_{x_k}\| = 0. \quad (5.6)$$

By Lemma 4.1,  $\{x_k\}$  is bounded hence has a convergent subsequence, say  $\{x_{k_i}\}_{i=1}^\infty$ . Let  $\lim_{i \rightarrow \infty} x_{k_i} = x^*$ . We claim that  $x^* \in S$ . Indeed, if  $\{m_{k_i}\}_{i=1}^\infty$  is bounded, then (5.6) implies that  $w_{x_{k_i}} \rightarrow 0$  and hence that  $y_a^\varphi(x_{k_i}) - x_{k_i} \rightarrow 0$  as  $i \rightarrow \infty$  (see (5.3)). By the continuity of  $y_a^\varphi$  it follows that  $y_a^\varphi(x^*) - x^* = 0$  and so  $x^* \in S$  by Proposition 3.1(iii). Therefore, we may assume that  $\{m_{k_i}\}$  is unbounded. By considering a subsequence if necessary we may assume further that  $m_{k_i} \rightarrow \infty$  and

$$\frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|} \rightarrow h^* \quad (5.7)$$

for some unit vector  $h^*$ . We suppose  $x^* \notin S$ . Then

$$x^* \neq y_a^\varphi(x^*) \text{ and } x^* \in (\theta_{ab}^\varphi)^{-1}(0, \epsilon]. \quad (5.8)$$

Moreover, by the Lebourg Mean Value Theorem (cf. [5, Theorem 2.3.7]), for each  $i \in \mathbb{N}$ , there exist  $y_i$  in the line segment  $[x_{k_i}, x_{k_i} + \rho^{m_{k_i}-1} w_{x_{k_i}}]$  and  $W_{k_i} \in \partial^c \theta_{ab}^\varphi(y_i)$  such that

$$\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1} w_{x_{k_i}}) - \theta_{ab}^\varphi(x_{k_i}) = \rho^{m_{k_i}-1} W_{k_i}^T w_{x_{k_i}}.$$

Dividing both sides by  $\rho^{m_{k_i}-1}\|w_{x_{k_i}}\|$ , it follows from  $W_{k_i} \in \partial^c \theta_{ab}^\varphi(y_i)$  and [5, Proposition 2.1.2 (b)] that for each  $i \in \mathbb{N}$

$$\frac{\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}}) - \theta_{ab}^\varphi(x_{k_i})}{\rho^{m_{k_i}-1}\|w_{x_{k_i}}\|} = W_{k_i}^T \frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|} \leq (\theta_{ab}^\varphi)^\circ(y_i; \frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|}).$$

This implies that

$$\limsup_{i \rightarrow \infty} \frac{\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}}) - \theta_{ab}^\varphi(x_{k_i})}{\rho^{m_{k_i}-1}\|w_{x_{k_i}}\|} \leq \limsup_{i \rightarrow \infty} (\theta_{ab}^\varphi)^\circ(y_i; \frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|}). \quad (5.9)$$

Noting that  $y_i \rightarrow x^*$  (by  $x_{k_i} \rightarrow x^*$  and  $y_i \in [x_{k_i}, x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}}]$ ), one has from [5, Proposition 2.1.1 (b)] and (5.7) that

$$\limsup_{i \rightarrow \infty} (\theta_{ab}^\varphi)^\circ(y_i; \frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|}) \leq (\theta_{ab}^\varphi)^\circ(x^*; h^*),$$

and it follows from (5.9) that

$$\limsup_{i \rightarrow \infty} \frac{\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}}) - \theta_{ab}^\varphi(x_{k_i})}{\rho^{m_{k_i}-1}\|w_{x_{k_i}}\|} \leq (\theta_{ab}^\varphi)^\circ(x^*; h^*). \quad (5.10)$$

Since  $\sqrt{\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}})} + \sqrt{\theta_{ab}^\varphi(x_{k_i})} \rightarrow 2\sqrt{\theta_{ab}^\varphi(x^*)}$ , (5.10) implies that

$$\limsup_{i \rightarrow \infty} \frac{\sqrt{\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}})} - \sqrt{\theta_{ab}^\varphi(x_{k_i})}}{\rho^{m_{k_i}-1}\|w_{x_{k_i}}\|} \leq \frac{(\theta_{ab}^\varphi)^\circ(x^*; h^*)}{2\sqrt{\theta_{ab}^\varphi(x^*)}}.$$

Since  $-\frac{\delta}{4}\rho^{m_{k_i}-1}\|w_{x_{k_i}}\| < \sqrt{\theta_{ab}^\varphi(x_{k_i} + \rho^{m_{k_i}-1}w_{x_{k_i}})} - \sqrt{\theta_{ab}^\varphi(x_{k_i})}$  (by the line search (Step 3)), it follows that

$$(\theta_{ab}^\varphi)^\circ(x^*; h^*) \geq -\frac{\delta}{2}\sqrt{\theta_{ab}^\varphi(x^*)}. \quad (5.11)$$

Write  $\mathbb{N} = I_1 \cup I_2$ , where

$$I_1 := \{i : c\|x_{k_i} - y_a^\varphi(x_{k_i})\| \leq \|y_a^\varphi(x_{k_i}) - y_b^\varphi(x_{k_i})\|\} \text{ and } I_2 := \mathbb{N} \setminus I_1. \quad (5.12)$$

**Case 1:** Suppose  $|I_1| = +\infty$ . Since for each  $i \in I_1$ , one has

$$c\|x_{k_i} - y_a^\varphi(x_{k_i})\| \leq \|y_a^\varphi(x_{k_i}) - y_b^\varphi(x_{k_i})\| \text{ and } w_{x_{k_i}} = y_a^\varphi(x_{k_i}) - y_b^\varphi(x_{k_i}). \quad (5.13)$$

It follows (by passing to the limits) that

$$c\|x^* - y_a^\varphi(x^*)\| \leq \|y_a^\varphi(x^*) - y_b^\varphi(x^*)\|, \quad (5.14)$$

and

$$h^* = \lim_{i \in I_1, i \rightarrow \infty} \frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|} = \frac{y_a^\varphi(x^*) - y_b^\varphi(x^*)}{\|y_a^\varphi(x^*) - y_b^\varphi(x^*)\|}. \quad (5.15)$$

By (5.14) and (4.23),  $y_a^\varphi(x^*) - y_b^\varphi(x^*) \in \Omega_{abc}(x^*)$ . Since  $\theta_{ab}^\varphi(x^*) > 0$ , it follows from (4.26) and (5.15) that  $(\theta_{ab}^\varphi)^\circ(x^*; h^*) \leq -\delta\sqrt{\theta_{ab}^\varphi(x^*)}$ , contradicting (5.11).

**Case 2:** Suppose  $|I_2| = +\infty$ . Since for each  $i \in I_2$ , one has

$$c\|x_{k_i} - y_a^\varphi(x_{k_i})\| > \|y_a^\varphi(x_{k_i}) - y_b^\varphi(x_{k_i})\| \quad \text{and} \quad w_{x_{k_i}} = y_a^\varphi(x_{k_i}) - x_{k_i}. \quad (5.16)$$

It follows (by passing to the limits) that

$$c\|x^* - y_a^\varphi(x^*)\| \geq \|y_a^\varphi(x^*) - y_b^\varphi(x^*)\| \quad (5.17)$$

and

$$h^* = \lim_{i \in I_2, i \rightarrow \infty} \frac{w_{x_{k_i}}}{\|w_{x_{k_i}}\|} = \frac{y_a^\varphi(x^*) - x^*}{\|y_a^\varphi(x^*) - x^*\|}. \quad (5.18)$$

By (5.17) and (4.23),  $y_a^\varphi(x^*) - x^* \in \Omega_{abc}(x^*)$ . Since  $\theta_{ab}^\varphi(x^*) > 0$ , it follows from (4.26) and (5.18) that  $(\theta_{ab}^\varphi)^\circ(x^*; h^*) \leq -\delta\sqrt{\theta_{ab}^\varphi(x^*)}$ , again contradicting (5.11).  $\square$

**Note:** Suppose that  $F$  is smooth, strongly monotone and  $\nabla F$  is locally Lipschitz. Then the solution of  $VI(F, P)$  is unique. From the preceding theorem, we see that the  $\{x_k\}_{k \in \mathbb{N}}$  generated by our algorithm converges to the unique solution of  $VI(F, P)$  (say  $x^*$ ). Indeed, in this case, the following stronger conclusion holds:

(i)  $\{\theta_{ab}^\varphi(x_k)\}$  converges Q-linearly, i.e. either the algorithm terminates in finite steps or there exists  $r \in (0, 1)$  such that

$$\lim_{k \rightarrow \infty} \frac{\theta_{ab}^\varphi(x_k)}{\theta_{ab}^\varphi(x_{k-1})} \leq r; \quad (5.19)$$

(ii)  $\{x_k\}$  converges R-linearly, i.e. either the algorithm terminates in finite steps or there exists  $r' \in (0, 1)$  such that

$$\lim_{k \rightarrow \infty} \sqrt[k]{\|x_k - x^*\|} \leq r'. \quad (5.20)$$

Indeed, to see (5.19), we first note that  $\theta_{ab}^\varphi(x_k) \neq 0$  for all  $k \in \mathbb{N}$ . Since  $\theta_{ab}^\varphi$  is smooth and  $\nabla\theta_{ab}^\varphi$  is locally Lipschitz in this case (see Proposition 3.1(iii)-(iv) and Theorem 3.1), then from the mean-value theorem we have for all  $\alpha \in [0, 1]$

$$\begin{aligned} \theta_{ab}^\varphi(x_k + \alpha w_{x_k}) - \theta_{ab}^\varphi(x_k) &= \int_0^\alpha \nabla\theta_{ab}^\varphi(x_k + tw_{x_k})^T w_{x_k} dt \\ &= \int_0^\alpha (\nabla\theta_{ab}^\varphi(x_k + tw_{x_k}) - \nabla\theta_{ab}^\varphi(x_k))^T w_{x_k} dt \\ &\quad + \alpha \nabla\theta_{ab}^\varphi(x_k)^T w_{x_k}. \end{aligned} \quad (5.21)$$

Note that  $\sup_k \|w_{x_k}\| < +\infty$  (since  $\{x_k\}$  is bounded and  $y_c^\varphi(\cdot)$  is continuous for all  $c > 0$ ) and hence  $\{x_k + tw_{x_k} : t \in [0, 1]\} \subseteq M$  where  $M := [\theta_{ab}^\varphi(x) \leq \theta_{ab}^\varphi(x_0)] + \sup_k \|w_{x_k}\| \overline{\mathbb{B}}(0, 1)$  is bounded. There exists  $L > 0$  such that  $\|\nabla\theta_{ab}^\varphi(x_k + tw_{x_k}) - \nabla\theta_{ab}^\varphi(x_k)\| \leq tL\|w_{x_k}\|$  for all  $k$ . Thus, from (5.21) and (5.1), we have

$$\begin{aligned} \theta_{ab}^\varphi(x_k + \alpha w_{x_k}) - \theta_{ab}^\varphi(x_k) &\leq \frac{\alpha^2}{2} L \|w_{x_k}\|^2 + \alpha \nabla\theta_{ab}^\varphi(x_k)^T w_{x_k} \\ &= \frac{\alpha^2}{2} L \|w_{x_k}\|^2 + \alpha (\theta_{ab}^\varphi)^\circ(x_k; w_{x_k}) \\ &\leq \alpha \|w_{x_k}\| \left( \frac{\alpha}{2} L \|w_{x_k}\| - \delta \sqrt{\theta_{ab}^\varphi(x_k)} \right). \end{aligned}$$

This implies that

$$\begin{aligned}
\frac{\sqrt{\theta_{ab}^\varphi(x_k + \alpha w_{x_k})} - \sqrt{\theta_{ab}^\varphi(x_k)}}{\alpha \|w_{x_k}\|} &= \frac{\theta_{ab}^\varphi(x_k + \alpha w_{x_k}) - \theta_{ab}^\varphi(x_k)}{\alpha \|w_{x_k}\| (\sqrt{\theta_{ab}^\varphi(x_k + \alpha w_{x_k})} + \sqrt{\theta_{ab}^\varphi(x_k)})} \\
&\leq \frac{\frac{\alpha}{2} L \|w_{x_k}\| - \delta \sqrt{\theta_{ab}^\varphi(x_k)}}{\sqrt{\theta_{ab}^\varphi(x_k + \alpha w_{x_k})} + \sqrt{\theta_{ab}^\varphi(x_k)}} \\
&\leq \frac{\frac{\alpha}{2} L \|w_{x_k}\|}{\sqrt{\theta_{ab}^\varphi(x_k)}} - \delta.
\end{aligned} \tag{5.22}$$

On the other hand, from [20, Proposition 2.2], there exists  $\tau > 0$  such that  $\|y_a^\varphi(x_k) - y_b^\varphi(x_k)\| \leq \tau \|x_k - y_b^\varphi(x_k)\|$ . This implies that  $\|x_k - y_a^\varphi(x_k)\| \leq \|x_k - y_b^\varphi(x_k)\| + \|y_b^\varphi(x_k) - y_a^\varphi(x_k)\| \leq (1 + \tau) \|x_k - y_b^\varphi(x_k)\|$ . Thus, from the definition of  $w_x$  and (3.10), we have  $\|w_{x_k}\| \leq (1 + \tau) \|x_k - y_b^\varphi(x_k)\| \leq (1 + \tau) \sqrt{\frac{1}{\beta(b-a)}} \sqrt{\theta_{ab}^\varphi(x_k)}$ . It follows from (5.22) that

$$\frac{\sqrt{\theta_{ab}^\varphi(x_k + \alpha w_{x_k})} - \sqrt{\theta_{ab}^\varphi(x_k)}}{\alpha \|w_{x_k}\|} \leq \frac{\alpha L(1 + \tau)}{2\sqrt{\beta(b-a)}} - \delta.$$

This implies that there exists  $m \in \mathbb{N}$  such that  $m_k \leq m$  for all  $k \in \mathbb{N}$  (see the line search (5.5)). It follows from (5.5), (5.3) and (3.10) that

$$\begin{aligned}
\sqrt{\theta_{ab}^\varphi(x_k)} &= (\sqrt{\theta_{ab}^\varphi(x_k)} - \sqrt{\theta_{ab}^\varphi(x_{k-1})}) + \sqrt{\theta_{ab}^\varphi(x_{k-1})} \\
&\leq -\frac{\delta}{4} \rho^{m_{k-1}} \|w_{x_{k-1}}\| + \sqrt{\theta_{ab}^\varphi(x_{k-1})} \\
&\leq -\frac{\delta c}{4} \rho^m \|x_{k-1} - y_a^\varphi(x_{k-1})\| + \sqrt{\theta_{ab}^\varphi(x_{k-1})} \\
&\leq -\frac{\delta c}{4\sqrt{(\kappa - \beta)(b-a)}} \rho^m \sqrt{\theta_{ab}^\varphi(x_{k-1})} + \sqrt{\theta_{ab}^\varphi(x_{k-1})} \\
&= \gamma \sqrt{\theta_{ab}^\varphi(x_{k-1})},
\end{aligned} \tag{5.23}$$

where  $\gamma := 1 - \frac{\delta c \rho^m}{4\sqrt{(\kappa - \beta)(b-a)}} \in (0, 1)$  (thanks to  $\theta_{ab}^\varphi(x_k) \neq 0$  for all  $k \in \mathbb{N}$ ). Thus (5.19) is seen to hold. To see (5.20), from our error bound result and (5.23), one has

$$\|x_k - x^*\| \leq \frac{2}{\delta} \sqrt{\theta_{ab}^\varphi(x_k)} \leq \frac{2}{\delta} \gamma^{k-1} \sqrt{\theta_{ab}^\varphi(x_1)}.$$

Thus, (5.20) follows.

**Remark 5.1** (a). *Our algorithm is indeed a derivative-free type method, i.e., we do not need to calculate the (generalized) Jacobian of  $F$ . This is useful especially when the calculation of the (generalized) Jacobian of  $F$  is expensive. Although, there already exist projection-type methods (which are also derivative-free methods) providing iterative sequences that converge to a solution assuming only  $F$  is monotone and continuous (e.g., [28, 29]). Our present approach is based*

on the consideration of error bounds of the merit function  $\theta_{ab}^\varphi$ . Hence we not only have the convergence result (Theorem 5.1) but also know how far from the  $k$ th point of the iteration to the solution set.

(b). Example 4.4 shows that our algorithm can be applied to some cases when  $F$  is nonmonotone.

(c). Suppose that  $F$  is smooth, strongly monotone and  $\nabla F$  is locally Lipschitz. It follows from the preceding note that the sequence  $\{x_k\}$  generated by our algorithm converges  $R$ -linearly. Under the same assumption, another linear convergent algorithm based on the implicit Lagrangian (which is a special case of our generalized  $D$ -gap function when  $\varphi(x, y) = \frac{1}{2}\|x - y\|^2$ ,  $a = b^{-1} \in (0, 1)$  and  $P$  is the nonnegative orthant) has been proposed in [18] for solving nonlinear complementary problems (which is a particular case of variational inequality problems). However, we note that the search direction used in [18] is different from the one we used in our algorithm.

(d). Our algorithm involves the parameters  $c$  and  $\delta$  which depend on the constants  $L$  and  $\mu_{ab}$  (see (4.25)). Thus, in general, some priori information is needed to determine these two constants. However, in some special situations (e.g.  $F$  is piecewise linear and  $P$  is the nonnegative orthant), the constant  $L$  and  $\mu_{ab}$  (and hence the parameter  $c$  and  $\delta$ ) can be determined explicitly. Moreover, when these two constants are too costly to obtain, similar to [18, page 10-11] and [33, page 453], we can start our algorithm with some reasonably small  $c$  and  $\delta$ , and adapt iteratively with decreasing values of them if the algorithm does not appear to make predicted progress.

**Acknowledgements:** We would like to express our sincere thanks to the anonymous referees for many helpful comments and for pointing out the references [16, 18]. Moreover, we are indebted to Dr. L.L. Tan for many stimulating discussion and to Prof. A.S. Lewis who suggested us to look at the subanalytic case.

## References

- [1] Alefeld, G., Chen X.: A regularized projection method for complementarity problems with non-Lipschitzian functions. *Math. Comp.* **77**, 379-395(2008).
- [2] Bierstone, E., Milman, P.D.: Subanalytic geometry. Model theory, algebra, and geometry, 151-172, *Math. Sci. Res. Inst. Publ.*, **39**, Cambridge Univ. Press, Cambridge, 2000.
- [3] Bolte, J., Daniilidis, A., Lewis, A.S.: The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, *SIAM J. Optim.* **17**, 1205-1223(2006).
- [4] Chen X., Nashed Z., Qi L.Q.: Smoothing methods and semismooth methods for nondifferentiable operator equations *SIAM J. Numerical Analysis.* **38**, 1200-1216(2000).
- [5] Clarke, F.H.: *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
- [6] Clarke, F.H., Ledyaev, Y.S., Stern, R.J., Wolenski, P.R.: *Nonsmooth Analysis and Control Theory*. Graduate Texts in Mathematics, **178**. Springer-Verlag, New York, 1998.
- [7] Dries, V.D., Miller, C.: Geometries, categories and o-minimal structures, *Duke Math. J.* **84**, 497-540(1996).

- [8] Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems Volume I, Springer publication, 2003.
- [9] Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems Volume II, Springer publication, 2003.
- [10] Harker, P.T., Pang, J.S.: Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications. *Math. Program.* **48**, 161-220(1990).
- [11] Harrington, J.E., Hobbs, B.F., Pang, J.S., Liu, A., Roch, G.: Collusive Game Solutions via Optimization, *Math. Program.*, **104**, 407-435(2005).
- [12] Hobbs, B.F., Pang, J.S.: Nash-Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints, *Oper. Res.* **55**, 113-127(2007).
- [13] Huang, L.R., Ng, K.F.: Equivalent optimization formulations and error bounds for variational inequality problems, *J. Optim. Theory Appl.* **125**, **2**, 299-314(2005).
- [14] Jiang, H.Y., Qi, L.Q.: Local uniqueness and convergence of iterative methods for nonsmooth variational inequalities. *J. Math. Anal. Appl.* **196**, 314-331(1995).
- [15] Kanzow, C., Fukushima, M.: Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities. *Math. Program.* **83**, 55-87(1998).
- [16] Luo, Z.Q., Mangasarian, O.L., Ren, J., Solodov, M.V., New error bounds for the linear complementarity problems. *Mathematics of Operations Research* **19** (1994), 880-892.
- [17] Luo, Z.Q., Pang, J.S., Ralph, D.: *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- [18] Mangasarian, O.L., Solodov, M.V., A linearly convergent derivative-free descent method for strongly monotone complementarity problems. *Computational Optimization and Applications* **14**, 5-16 (1999).
- [19] Ng, K.F., Tan, L.L.: Error bounds for regularized gap function for nonsmooth variational inequality problem, *Math. Program.*, **110**, 405-429(2007).
- [20] Ng, K.F., Tan, L.L.: D-gap functions for nonsmooth variational inequality problems. *J. Optim. Theory Appl.* **133** (2007), no. 1, 77-97.
- [21] Mangasarian, O.L., Solodov, M. V.: A linearly convergent derivative-free descent method for strongly monotone complementarity problems. *Comput. Optim. Appl.* **14** 5-16(1999).
- [22] Ng, K.F., Zheng, X.Y.: Global error bounds with fractional exponents, *Math. Program.* **88**, 357-370(2000).
- [23] Pang, J.S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. *Comput. Manag. Sci.* **1**, 21-56(2005).

- [24] Peng, J. M., Equivalence of variational inequality problems to unconstrained minimization. *Math. Programming* 78 (1997), no. 3, Ser. A, 347–355.
- [25] Peng, J.M., Fukushima, M.: A hybrid Newton method for solving the variational inequality problem via the D-gap function. *Math. Program.* **86**, 367-386(1999).
- [26] Rademacher, H.: Über partielle und totale differenzierbarkeit I, *Mathematical Annals.* **89**, 340-359(1919).
- [27] Solodov, M.V.: Merit functions and error bounds for generalized variational inequalities. *J. Math. Anal. Appl.* 287, **2**, 405-414 (2003).
- [28] Solodov, M.V., Svaiter, B.F.: A new projection method for variational inequality problems. *SIAM J. Control Optim.* **37**, 765-776(1999).
- [29] Solodov, M.V., Tseng, P.: Modified projection-type methods for monotone variational inequalities. *SIAM J. Control Optim.* **34**, 1814-1830(1996).
- [30] Warga, J.: Fat homeomorphisms and unbounded derivate containers, *J. Math. Anal. Appl.* 81, **2**, 545-560(1981).
- [31] Wu, J.H., Florian, M., Marcotte, P.: A general descent framework for the monotone variational inequality problem, *Math. Program.* **61**, 281-300(1993).
- [32] Yamashita, N., Fukushima, M.: Equivalent unconstrained minimization and global error bounds for variational inequality problems, *SIAM. J. Control. Optim.* **35**, 273-284(1997).
- [33] Yamashita, N., Taji, K., Fukushima, M.: Unconstrained optimization reformulations of variational inequality problems, *J. Optim. Theory Appl.* **92**, 439-456(1997).