Journal of Global Optimization
https://doi.org/10.1007/s10898-025-01558-6

Kurdyka-Łojasiewicz inequality and error bounds of D-Gap functions for nonsmooth and nonmonotone variational inequality problems

M. H. Li[1] · K. W. Meng[2] · X. Q. Yang[3]

Received: 2 April 2024 / Accepted: 14 October 2025

Abstract

In this paper, we study regularized/D-gap functions associated with a nonsmooth and nonmonotone variational inequality problem. We present some exact formulas for the subderivative, the regular subdifferential, and the limiting subdifferential of the regularized/D-gap functions respectively. By virtue of these formulas, we provide some sufficient conditions and necessary conditions for the Kurdyka-Łojasiewicz inequality property and the error bound property for the D-gap function respectively. As an application of our Kurdyka-Łojasiewicz inequality result, we show that, under certain mild assumptions, the sequence generated by a derivative-free descent algorithm with an inexact line search converges linearly to a solution of the variational inequality problem.

Keywords Nonlinear programming · Variational inequality problem · D-gap function · Kurdyka-Łojasiewicz inequality · Error bound

Mathematics Subject Classification Primary 65K10 · 65K15; Secondary 90C26 · 49M37

1 Introduction

In this paper, we consider a variational inequality problem (VIP) of finding $x \in K$ such that

$$\langle F(x), y - x \rangle \geq 0 \quad \forall y \in K,$$

✉ X. Q. Yang
mayangxq@polyu.edu.hk

M. H. Li
minghuali20021848@163.com

K. W. Meng
mengkw@swufe.edu.cn

[1] School of Mathematics and Big Data, Chongqing University of Arts and Sciences, Yongchuan, 402160 Chongqing, China

[2] School of Mathematics and Big Data Laboratory on Financial Security and Behavior, Southwestern University of Finance and Economics, Chengdu 611130, China

[3] Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong

where K is a closed and convex subset of $\mathbb{R}^n$ and the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz continuous but not necessarily monotone. (VIP) has many applications in various fields such as mathematical programming, traffic network equilibrium problems and economics. For the background information and motivations of (VIP), we refer the reader to the comprehensive book [12] by Facchinei and Pang.

One popular approach to studying (VIP) is to reformulate (VIP) as equivalent constrained or unconstrained optimization problems by introducing appropriate gap (merit) functions; see [2, 3, 9, 12, 13, 15, 17, 18, 20, 22, 27, 30–35, 39, 40, 42–44, 46, 47]. Among various reformulations in the literature, we recall that $\bar{x}$ solves (VIP) if and only if $\bar{x}$ solves the following unconstrained optimization problem with 0 as its optimal value:

$$\min_{x\in\mathbb{R}^n} \quad f_{ab}(x) := f_a(x) - f_b(x),$$

where $b > a > 0$, and for each $c > 0$,

$$f_c(x) := \max_{y\in K} \left\{ \langle F(x), x - y\rangle - \frac{c}{2}||y - x||^2 \right\}.$$

f_c is known as the regularized gap function [2, 13] with c being the regularized parameter. f_{ab} is often known as the D-gap function [33] with 'D' standing for the 'difference' of two parameterized regularized gap functions. By replacing the quadratic term in the definition of f_c with a more general one that retains similar properties to the quadratic term, the corresponding generalized regularized gap and generalized D-gap functions have also been extensively studied in the literature; see [21, 22, 42, 47].

The (generalized) differentiability properties of these regularized gap and D-gap functions have been extensively investigated, and have been utilized to study various properties of error bounds [12] and the Kurdyka-Łojasiewicz (KL, for short) inequality [11]. The latter properties have played very important roles in convergence analysis for algorithms designed based on gap functions.

Below, we summarize a few results related to the (generalized) D-gap function. Peng [33] showed that if F is continuously differentiable and strongly monotone, the D-gap function is also continuously differentiable and its square root provides a global error bound for (VIP). Yamashita et al. [47] introduced the generalized D-gap function and obtained its continuous differentiability by assuming that F is continuously differentiable. Moreover, under the assumptions that F is strongly monotone and that either F is Lipschitz continuous or K is compact, they showed that the square root of the generalized D-gap function provides a global error bound for (VIP), and that the sequence generated by a descent algorithm with an inexact line search converges to the unique solution of (VIP). Based on the D-gap function and by assuming that F is continuously differentiable and monotone, Solodov and Tseng [38] developed two unconstrained methods that are similar to the feasible direction method in Zhu and Marcotte [48] which is based on the regularized gap function. By assuming that F is locally Lipschitz continuous, Xu [45] obtained a formula for the Clarke subdifferential of the D-gap function, and a global convergence result for a descent algorithm with an inexact line search under the circumstance that F is strongly monotone and Lipschitz continuous. By using the same assumptions as in [45], Ng and Tan [28] obtained some formulas for the Clarke directional derivative and the Clarke subdifferential of the D-gap function. By assuming that F is coercive and locally Lipschitz continuous, and by introducing a condition expressed in terms of the Clarke generalized Jacobian of F, Li and Ng [21] showed that the square root of the generalized D-gap function provides a local error bound for (VIP), and by virtue of which, they proved that any cluster point of the sequence generated by a descent algorithm with an inexact line search is a solution of (VIP), and that the convergence rate is linear when F is smooth, strongly monotone and ∇F is locally Lipschitz continuous. Note that Li and Ng [21] also provided some formulas for the Clarke directional derivative and the Clarke subdifferential of the generalized D-gap function, which were very crucial for their arguments. Later Li et al. [22] established some error bound results for the generalized D-gap function by assuming that F is Lipschitz continuous, locally monotone and coercive.

From the literature review above, it is clear to see that most of the existing results for error bounds and the convergence of a descent algorithm were obtained by assuming that F is strongly monotone, and with an exception that the error bound was obtained in Li and Ng [21] when F is nonmonotone. Regarding the KL inequality, to the best of our knowledge, there is almost no direct result for the case when F is locally Lipschitz continuous. By examining the definition for the KL inequality (see Definition 2.3 below) and the theory of error bounds in [6, 24], it seems to be case that the notion of the subderivative, the regular/Fréchet subdifferential, and the general/limiting subdifferential set (see Definition 2.2) should have been employed in studying the generalized differentiability properties of the regularized gap and D-gap functions. But it is quite surprising that there is no such a related result in the literature for the case when F is locally Lipschitz continuous but not necessarily monotone.

To fill this gap, we will investigate the KL inequality and error bounds of the D-gap function for nonsmooth and nonmonotone (VIP) by deriving formulas for the subderivative and the (limiting) subdifferential of the D-gap functions respectively. As an application of our results for the KL inequality and by following the idea in the proof of the abstract convergence result [5] for inexact descent methods, we will establish the linear convergence rate for a descent algorithm with an inexact line search.

The main contributions of the paper are as follows.

(i) We obtain a number of exact formulas for the subderivatives, the regular/Fréchet subdifferentials and the general/limiting subdifferential sets of the regularized gap function f_c and the D-gap function f_{ab}, respectively. See Propositions 3.1-3.2 below. For example, we obtain the following formula for the limiting subdifferential of f_{ab} at a point $\bar{x}$:

$$\partial f_{ab}(\bar{x}) = D^* F(\bar{x})\,(\pi_b(\bar{x}) - \pi_a(\bar{x})) - b\,(\pi_b(\bar{x}) - \pi_a(\bar{x})) + (b-a)(\bar{x} - \pi_a(\bar{x})),$$

where $D^* F(\bar{x})$ denotes the coderivative of F at $\bar{x}$ (cf. Definition 2.5), and for each $c > 0$, $\pi_c(x) := P_K\left(x - \frac{F(x)}{c}\right)$ with $P_K(\cdot)$ being the projection operator onto K. To the best of our knowledge, this formula has not been seen in the literature, although, as mentioned above, exact formulas have been obtained for the Clarke directional derivatives and the Clarke subdifferentials of f_c and f_{ab}, respectively. It should be noticed that, although f_c is a marginal function and $f_{ab} = f_a - f_b$ is a difference of two marginal functions, we cannot obtain our formulas by directly applying the theory of marginal functions known from the literature [1, 19, 25, 26, 36]. As a matter of fact, our approach depends heavily on the inherent structures of f_c and f_{ab}.

(ii) By virtue of the formula obtained for the general/limiting subdifferential of the D-gap function f_{ab}, we present a few sharp results on the properties of the KL inequality and the error bounds for f_{ab}. In particular, by assuming that the following inequality holds for some $\mu > 0$ and all $x \in \mathbb{R}^n$ where F is differentiable:

$$\langle \nabla F(x)(\pi_a(x) - \pi_b(x)),\ \pi_a(x) - \pi_b(x)\rangle \geq \mu||\pi_a(x) - \pi_b(x)||^2, \quad (1)$$

which can be considered as a restricted (weaker) notion of strong monotonicity, we show that

$$d(0, \partial f_{ab}(x)) \geq \mu\|\pi_b(x) - \pi_a(x)\| \quad \forall x \in \mathbb{R}^n,$$

and that f_{ab} is a KL function with an exponent of $\frac{1}{2}$, and moreover that some local/global error bound results holds. See Theorem 4.1 below.

(iii) By assuming (1) (guaranteeing that f_{ab} is a KL function with an exponent of $\frac{1}{2}$) and the global Lipschitz continuity of f_{ab}, we establish a linear convergence rate for a derivative-free descent algorithm, which is essentially the same type of algorithm studied in [17, 21, 34, 35, 45, 47]. See Theorem 5.1 below. Starting from any initial point x_0, the algorithm generates a sequence $\{x_n\}$ via $x_{n+1} = x_n + t_n d_n$, where d_n is a search direction, being either $\pi_a(x_n) - x_n$ or $\pi_a(x_n) - \pi_b(x_n)$, and t_n is the stepsize determined by an Armijo line search. Under further mild assumptions, we show that the stepsize sequence $\{t_n\}$ is bounded below by a positive constant $t^* > 0$ (cf. Proposition 5.2 below), and moreover the following hold (cf. Proposition 5.3 below):

$$f_{ab}(x_{n+1}) - f_{ab}(x_n) \le -M_1||x_{n+1} - x_n||^2$$

and

$$d(0, \partial f_{ab}(x_n)) \le \frac{M_2}{t^*}||x_{n+1} - x_n||,$$

where M_1 and M_2 are positive constants.

The outline of the paper is as follows. Section 2 introduces the notation, terminology, and main mathematical preliminaries. In section 3, we present some exact formulas for the subderivatives, the regular/Fréchet subdifferentials, and the general/limiting subdifferentials of the regularized gap function f_c and the D-gap function f_{ab}, respectively. Using these formulas for the D-gap function, Section 4 provides sufficient conditions and necessary conditions for the error bound property and the KL inequality property respectively. As an application of our KL inequality result, we show in Section 5 that a descent algorithm (based on the D-gap function) with an inexact line search converges linearly to a solution of (VIP).

2 Notation and mathematical preliminaries

Throughout the paper we use the standard notations of variational analysis; see the seminal book [37] by Rockafellar and Wets. The Euclidean norm of a vector x is denoted by $||x||$, and the inner product of vectors x and y is denoted by $\langle x, y\rangle$. Let $A \subset \mathbb{R}^n$ be a nonempty set. We denote by conv A the convex hull of A. The polar cone of A is defined by $A^* := \{v \in \mathbb{R}^n \mid \langle v, x\rangle \le 0 \ \forall x \in A\}$. The distance from x to A is defined by $d(x, A) := \inf_{y\in A} ||y - x||$. The projection mapping P_A is defined by $P_A(x) := \{y \in A \mid \|y - x\| = d(x, A)\}$.

Definition 2.1 Let $C \subset \mathbb{R}^n$ and let $x \in C$.

(i) The tangent cone to C at x is denoted by $T_C(x)$, i.e., $w \in T_C(x)$ if there exist sequences $t_k \to 0$ and $\{w_k\} \subset \mathbb{R}^n$ with $w_k \to w$ and $x + t_k w_k \in C \ \forall k$.

(ii) The regular normal cone to C at x is denoted by $\widehat{N}_C(x)$, i.e., $v \in \widehat{N}_C(x)$ if

$$\langle v, x - \bar{x}\rangle \le o(\|x - \bar{x}\|) \quad \text{for all } x \in C.$$

Another way of defining the regular normal cone is via the equality $\widehat{N}_C(x) = T_C(x)^*$.

(iii) The normal cone to C at x is denoted by $N_C(x)$, i.e., $v \in N_C(x)$ if there exist sequences $x_k \to x$ and $v_k \to v$ with $x_k \in C$ and $v_k \in \widehat{N}_C(x_k)$ for all k.

(iv) C is said to be regular at x in the sense of Clarke if it is locally closed at x (i.e., $C \cap U$ is closed for some closed neighborhood U of x) and $\widehat{N}_C(x) = N_C(x)$.

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ be an extended real-valued function. We denote the epigraph of f by $\operatorname{epi} f := \{(x, \alpha) \mid f(x) \leq \alpha\}$. The lower level set with a level of α is defined and denoted by $[f \leq \alpha] := \{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}$. In a similar way, we define $[f < \alpha] := \{x \in \mathbb{R}^n \mid f(x) < \alpha\}$ and $[\alpha < f < \beta] := \{x \in \mathbb{R}^n \mid \alpha < f(x) < \beta\}$.

Definition 2.2 Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be an extended real-valued function and let $\bar{x}$ be a point with $f(\bar{x})$ finite.

(i) The vector $v \in \mathbb{R}^n$ is a regular/Fréchet subgradient of f at $\bar{x}$, written $v \in \widehat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle + o(||x - \bar{x}||).$$

(ii) The vector $v \in \mathbb{R}^n$ is a general/limiting subgradient of f at $\bar{x}$, written $v \in \partial f(\bar{x})$, if there exist sequences $x_k \to \bar{x}$ and $v_k \to v$ with $f(x_k) \to f(\bar{x})$ and $v_k \in \widehat{\partial} f(x_k)$.

(iii) The function f is said to be (subdifferentially) regular at $\bar{x}$ if $\operatorname{epi} f$ is regular in the sense of Clarke at $(\bar{x}, f(\bar{x}))$ as a subset of $\mathbb{R}^n \times \mathbb{R}$.

(iv) The subderivative $df(\bar{x}) : \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined by

$$df(\bar{x})(w) := \liminf_{t \to 0_+, w' \to w} \frac{f(\bar{x} + tw') - f(\bar{x})}{t}.$$

(v) The set of Clarke subgradients of f at $\bar{x}$ is defined by

$$\overline{\partial} f(\bar{x}) := \{v | (v, -1) \in \operatorname{cl} \operatorname{conv} N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))\},$$

where $\operatorname{cl} \operatorname{conv} N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$ denotes the closed and convex hull of $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$.

Remark 2.1 The regular subgradients can be derived from the subderivative as follows [37, Exercise 8.4]:

$$\widehat{\partial} f(\bar{x}) = \{v \in \mathbb{R}^n | \langle v, w \rangle \leq df(\bar{x})(w) \ \forall w \in \mathbb{R}^n\}.$$

Following [4, 7, 8, 23], we introduce the notion of the Kurdyka-Łojasiewicz (KL, for short) inequality.

Definition 2.3 For a proper lower semicontinuous function $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, a point $\bar{x} \in \mathbb{R}^n$ with $\partial f(\bar{x}) \neq \emptyset$, and some $\alpha \in [0, 1)$, we say that f satisfies the KL inequality at $\bar{x}$ with an exponent of α, if there exist $\mu, \epsilon > 0$ and $\nu \in (0, +\infty]$ so that

$$d(0, \partial f(x)) \geq \mu (f(x) - f(\bar{x}))^{\alpha}$$

whenever $\|x - \bar{x}\| \leq \epsilon$ and $f(\bar{x}) < f(x) < f(\bar{x}) + \nu$. If f satisfies the KL inequality at every $x \in \mathbb{R}^n$ with $\partial f(x) \neq \emptyset$ and with the same exponent α, we say that f is a KL function with an exponent of α.

Following [12], we introduce the notion of local and global error bounds as follows.

Definition 2.4 For a proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and a set $C \subset \mathbb{R}^n$, we say that f has a local error bound on C if there exist two positive constants τ and ϵ such that for all $x \in [f \leq \epsilon] \cap C$

$$d(x, [f \leq 0] \cap C) \leq \tau \max\{f(x), 0\}.$$

Furthermore, we say that f has a global error bound on C if there exists a constant $\tau > 0$ such that the above inequality holds for all $x \in C$.

Definition 2.5 Let $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping and $(\bar{x}, \bar{u}) \in \operatorname{gph} S := \{(x, u) \mid u \in S(x)\}$.

(i) The graphical derivative of S at $\bar{x}$ for $\bar{u}$ is the mapping $DS(\bar{x} \mid \bar{u}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined by

$$z \in DS(\bar{x} \mid \bar{u})(w) \Longleftrightarrow (w, z) \in T_{\text{gph}\, S}(\bar{x}, \bar{u}).$$

(ii) The regular coderivative of S at $\bar{x}$ for $\bar{u}$ is the mapping $\widehat{D}^* S(\bar{x} \mid \bar{u}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$x^* \in \widehat{D}^* S(\bar{x} \mid \bar{u})(u^*) \Longleftrightarrow (x^*, -u^*) \in \widehat{N}_{\text{gph}\, S}(\bar{x}, \bar{u}).$$

(iii) The coderivative of S at $\bar{x}$ for $\bar{u}$ is the mapping $D^* S(\bar{x} \mid \bar{u}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$x^* \in D^* S(\bar{x} \mid \bar{u})(u^*) \Longleftrightarrow (x^*, -u^*) \in N_{\text{gph}\, S}(\bar{x}, \bar{u}).$$

Here the notation $DS(\bar{x} \mid \bar{u})$, $D^* S(\bar{x} \mid \bar{u})$ and $\widehat{D}^* S(\bar{x} \mid \bar{u})$ is simplified to $DS(\bar{x})$, $D^* S(\bar{x})$ and $\widehat{D}^* S(\bar{x})$ when S is single-valued at $\bar{x}$, i.e., $S(\bar{x}) = \{\bar{u}\}$.

Definition 2.6 Let F be a single-valued mapping defined on $\mathbb{R}^n$, with values in $\mathbb{R}^m$.

(i) F is globally Lipschitz continuous if there exists $\kappa \in \mathbb{R}_+ := [0, \infty)$ with

$$\|F(x') - F(x)\| \leq \kappa \|x' - x\| \quad \forall x, x' \in \mathbb{R}^n.$$

Then κ is called a Lipschitz constant for F.

(ii) F is locally Lipschitz continuous at a point $\bar{x} \in \mathbb{R}^n$ if the value

$$\text{lip}\, F(\bar{x}) := \limsup_{x, x' \to \bar{x}, x \neq x'} \frac{\|F(x') - F(x)\|}{\|x' - x\|}$$

is finite. Here $\text{lip}\, F(\bar{x})$ is the Lipschitz modulus of F at $\bar{x}$.

(iii) F is locally Lipschitz continuous if F is locally Lipschitz continuous at every $\bar{x} \in \mathbb{R}^n$.

Lemma 2.1 *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be an extended real-valued function and let $\bar{x}$ be a point with $f(\bar{x})$ finite. Assume that f is locally Lipschitz continuous at $\bar{x}$. The following properties hold:*

(a) *$\partial f(\bar{x})$ is nonempty and compact.*

(b) $df(\bar{x})(w) = \liminf_{t \to 0_+} \dfrac{f(\bar{x} + tw) - f(\bar{x})}{t}$.

(c) $\overline{\partial} f(\bar{x}) = \text{conv}(\partial f(\bar{x}))$.

Proof **(a-c)** can be found in [37, Theorem 9.13, Exercise 9.15, Theorem 9.61], respectively. □

Lemma 2.2 *Assume that $F : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz continuous at a point $\bar{x} \in \mathbb{R}^n$. The following properties hold:*

(a) $D^* F(\bar{x})(0) = \{0\}$*, which is also sufficient for F being locally Lipschitz continuous at $\bar{x}$.*

(b) *The mappings $DF(\bar{x})$ and $D^* F(\bar{x})$ are nonempty-valued and locally bounded.*

(c) $||z|| \leq (\text{lip}\, F(\bar{x}))\, ||w||$ *holds for all* $(w, z) \in \text{gph}(DF(\bar{x}))$.

(d) $||x^*|| \leq (\text{lip}\, F(\bar{x}))\, ||u^*||$ *holds for all* $(u^*, x^*) \in \text{gph}(D^* F(\bar{x}))$.

(e) $z \in DF(\bar{x})(w)$ *if and only if there is some* $\tau^\nu \to 0_+$ *such that* $\frac{F(\bar{x} + \tau^\nu w) - F(\bar{x})}{\tau^\nu} \to z$.

Proof **(a)** follows directly from the Mordukhovich criterion [37, Theorem 9.40]. **(b-d)** follow from [37, Proposition 9.24]. **(e)** follows from the definitions of the graphical derivative and the local Lipschitzian continuity. □

Assume now that $F : \mathbb{R}^n \to \mathbb{R}^m$ is a locally Lipschitz continuous function and let D be the subset of $\mathbb{R}^n$ consisting of the points where F is differentiable. By the Rademacher Theorem [37, Theorem 9.60], F is differentiable almost everywhere with $\mathbb{R}^n \backslash D$ being negligible. For each $\bar{x} \in \mathbb{R}^n$, define

$$\overline{\nabla} F(\bar{x}) := \{A \in \mathbb{R}^{m \times n} \mid \exists x^\nu \to \bar{x} \text{ with } x^\nu \in D,\ \nabla F(x^\nu) \to A\}, \tag{2}$$

in terms of which, the generalized Jacobian $\overline{\partial} F(x)$ [10, Definition 2.6.1] of F at $\bar{x}$ can be written as

$$\overline{\partial} F(\bar{x}) := \operatorname{conv} \overline{\nabla} F(\bar{x}). \tag{3}$$

According to [37, Theorem 9.62], $\overline{\nabla} F(\bar{x})$ is a nonempty, compact set of matrices, and for every $w \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ one has

$$\operatorname{conv} D^* F(\bar{x})(y) = \operatorname{conv}\{A^T y \mid A \in \overline{\nabla} F(\bar{x})\} = \{A^T y \mid A \in \operatorname{conv} \overline{\nabla} F(\bar{x})\} \tag{4}$$

and

$$\operatorname{conv} D_* F(\bar{x})(w) = \operatorname{conv}\{A w \mid A \in \overline{\nabla} F(\bar{x})\} = \{A w \mid A \in \operatorname{conv} \overline{\nabla} F(\bar{x})\}, \tag{5}$$

where $D_* F(\bar{x})$ stands for the strict derivative mapping of F at $\bar{x}$ [37, Definition 9.53], and has the following definition by taking into account that F is locally Lipschitz continuous:

$$D_* F(\bar{x})(w) := \{z \mid \exists \tau^\nu \to 0_+, x^\nu \to \bar{x} \text{ with } (F(x^\nu + \tau^\nu w) - F(x^\nu))/\tau^\nu \to z\}. \tag{6}$$

Note that $D_* F(\bar{x})$ is also known as the Thibault's strict derivative (cf. [41]), and that by definition

$$\operatorname{gph} DF(\bar{x}) \subset \operatorname{gph} D_* F(\bar{x}). \tag{7}$$

Definition 2.7 [12] Let C be a subset of $\mathbb{R}^n$, and let F be a single-valued mapping defined on $\mathbb{R}^n$, with values in $\mathbb{R}^n$. F is said to be coercive on C if

$$\lim_{x \in C,\ \|x\| \to \infty} \frac{\langle F(x), x - y\rangle}{\|x\|} = +\infty$$

holds for all $y \in C$ (if C is bounded, then F is by convention coercive on C); and F is said to be strongly monotone on C (with modulus $\mu > 0$) if $\langle F(x) - F(y),\ x - y\rangle \geq \mu \|x - y\|^2$ holds for all $x, y \in C$.

3 Subderivatives and subgradients of gap functions

In the remaining part of this paper, we impose the following general assumptions on the problem data and certain constants. For simplicity's sake, we will not mention them when stating a result.

- $K \subset \mathbb{R}^n$ is a nonempty closed and convex set.
- $F : \mathbb{R}^n \to \mathbb{R}^n$ is a locally Lipschitz continuous function.
- a, b, c are fixed positive numbers with $a < b$.

The aim of this section is to investigate the subderivatives and subgradients of f_{ab} and f_c at a certain point $\bar{x}$ by utilizing the graphical derivative $DF(\bar{x})$, as well as the coderivatives $D^* F(\bar{x})$ and $\widehat{D}^* F(\bar{x})$. Moreover, we will often make use of the following projection operator π_c which is associated with F and K as follows:

$$\pi_c(x) := P_K\left(x - \frac{F(x)}{c}\right).$$

The projection operators π_a and π_b are defined in the same way.

To begin with, we summarize some basic properties of the regularized gap function f_c and the D-gap function f_{ab} below. Most of these properties can be found in the literature and are beneficial for the further development in the sequel.

Lemma 3.1 *The following properties hold:*

(a) $\frac{b-a}{2}||x-\pi_b(x)||^2+\frac{a}{2}||\pi_b(x)-\pi_a(x)||^2 \leq f_{ab}(x) \leq \frac{b-a}{2}||x-\pi_a(x)||^2-\frac{b}{2}||\pi_b(x)-\pi_a(x)||^2$.

(b) $||\pi_b(x)-\pi_a(x)|| \leq \frac{b-a}{a}||x-\pi_a(x)||$ *and* $||x-\pi_b(x)|| \leq ||x-\pi_a(x)|| \leq \frac{b}{a}||x-\pi_b(x)||$.

(c) $x \in \mathbb{R}^n$ *solves (VIP)* $\Leftrightarrow x=\pi_c(x)$ *for any* $c>0 \Leftrightarrow f_{ab}(y) \geq f_{ab}(x)=0$ *for all* $y \in \mathbb{R}^n$ $\Leftrightarrow x \in K$ *and* $f_c(y) \geq f_c(x)=0$ *for all* $y \in K$.

(d) $\langle a(x-\pi_a(x))-b(x-\pi_b(x)), \pi_a(x)-\pi_b(x)\rangle \geq 0$.

(e) $\pi_a(x)-\pi_b(x) \in T_{ab}(x,F,K) := T_K(\pi_b(x)) \cap (-T_K(\pi_a(x))) \cap (F(x))^*$.

(f) π_a, π_b, π_c, f_c *and* f_{ab} *are locally Lipschitz continuous. If* F *is globally Lipschitz continuous, then* π_a, π_b, π_c, f_c *and* f_{ab} *are also globally Lipschitz continuous.*

(g) *The following hold:*

$$\begin{aligned}&\arg\max_{y\in K}\left\{\langle F(x), x-y\rangle-\tfrac{c}{2}||y-x||^2\right\}=\{\pi_c(x)\},\\ &f_c(x)=\langle F(x), x-\pi_c(x)\rangle-\tfrac{c}{2}||x-\pi_c(x)||^2,\\ &f_{ab}(x)=\langle F(x), \pi_b(x)-\pi_a(x)\rangle-\tfrac{a}{2}||x-\pi_a(x)||^2+\tfrac{b}{2}||x-\pi_b(x)||^2.\end{aligned}$$

Proof **(a)** and **(b)** can be found in [38, Lemma 1] and [28], respectively. **(c)** can be found in [13] and [42]. **(d)** and **(e)** can be found in [21, Lemma 4.4] or in [12, Theorem 10.3.4]. **(f)** can be found in [22, Lemma 3.1]. **(g)** can be found in [42] or deduced from the standard optimality condition for convex programs. This completes the proof. □

3.1 Subderivatives and subgradients of f_c

We first present the formulas for the subderivative, the regular subdifferential set and the limiting subdifferential set of f_c at a point $\bar{x}$.

Proposition 3.1 *Let* $\bar{x} \in \mathbb{R}^n$ *and let* $w \in \mathbb{R}^n$. *We have the following formulas:*

$$\begin{aligned}&df_c(\bar{x})(w)=\langle F(\bar{x}),\ w\rangle+\min\langle(DF(\bar{x})-cI)\,w,\ \bar{x}-\pi_c(\bar{x})\rangle,\\ &\widehat{\partial} f_c(\bar{x})=\left(\widehat{D}^*F(\bar{x})-cI\right)(\bar{x}-\pi_c(\bar{x}))+F(\bar{x}),\\ &\partial f_c(\bar{x})=\left(D^*F(\bar{x})-cI\right)(\bar{x}-\pi_c(\bar{x}))+F(\bar{x}),\end{aligned}$$

where

$$\min\langle(DF(\bar{x})-cI)\,w,\ \bar{x}-\pi_c(\bar{x})\rangle := \min_{v\in DF(\bar{x})(w)}\langle v-cw,\ \bar{x}-\pi_c(\bar{x})\rangle.$$

Proof Let $w \in \mathbb{R}^n$ be fixed. Since F is locally Lipschitz continuous, it follows from Lemma 2.2 **(b)** and **(e)** that for any continuous function $M:\mathbb{R}\to\mathbb{R}^n$,

$$\liminf_{t\to 0_+}\langle\frac{F(\bar{x}+tw)-F(\bar{x})}{t},\ M(t)\rangle=\min_{v\in DF(\bar{x})(w)}\langle v,\ M(0)\rangle. \tag{8}$$

By Lemma 3.1 **(f)**, f_c is a locally Lipschitz continuous function. This implies, by Lemma 2.1 **(b)**, that

$$df_c(\bar{x})(w)=\liminf_{t\to 0_+}\frac{f_c(\bar{x}+tw)-f_c(\bar{x})}{t}.$$

In view of Lemma 3.1 **(g)**, we have for all t that

$$f_c(\bar{x}) \geq \langle F(\bar{x}), \bar{x} - \pi_c(\bar{x} + tw)\rangle - \frac{c}{2}||\bar{x} - \pi_c(\bar{x} + tw)||^2,$$

and

$$f_c(\bar{x} + tw) = \langle F(\bar{x} + tw), \bar{x} + tw - \pi_c(\bar{x} + tw)\rangle - \frac{c}{2}||\bar{x} + tw - \pi_c(\bar{x} + tw)||^2.$$

This, together with (8) and the fact that π_c is locally Lipschitz continuous (cf. Lemma 3.1 **(f)**), implies that

$$\begin{aligned} df_c(\bar{x})(w) &\leq \liminf_{t\to 0_+} \langle \frac{F(\bar{x} + tw) - F(\bar{x})}{t}, \ \bar{x} - \pi_c(\bar{x} + tw)\rangle + \lim_{t\to 0_+} \langle F(\bar{x} + tw), \ w\rangle \\ &\quad + \lim_{t\to 0_+} \frac{c}{2}\langle 2(\bar{x} - \pi_c(\bar{x} + tw)) + tw, \ -w\rangle \\ &= \min_{v\in DF(\bar{x})(w)} \langle v, \bar{x} - \pi_c(\bar{x})\rangle + \langle F(\bar{x}), \ w\rangle - c\langle \bar{x} - \pi_c(\bar{x}), \ w\rangle \\ &= \min\langle (DF(\bar{x}) - cI)\, w, \ \bar{x} - \pi_c(\bar{x})\rangle + \langle F(\bar{x}), \ w\rangle. \end{aligned}$$

To prove the inequality in the other direction, we can simply follow a similar approach by observing, from Lemma 3.1 **(g)**, that for all t,

$$f_c(\bar{x}) = \langle F(\bar{x}), \bar{x} - \pi_c(\bar{x})\rangle - \frac{c}{2}||\bar{x} - \pi_c(\bar{x})||^2,$$

and

$$f_c(\bar{x} + tw) \geq \langle F(\bar{x} + tw), \bar{x} + tw - \pi_c(\bar{x})\rangle - \frac{c}{2}||\bar{x} + tw - \pi_c(\bar{x})||^2.$$

To get the formula for $\widehat{\partial} f_c(\bar{x})$, we resort to the formula for $df_c(\bar{x})$ and the equality in Remark 2.1. Specifically, by defining $\bar{v} := F(\bar{x}) - c(\bar{x} - \pi_c(\bar{x}))$, we have

$$\begin{aligned} & v \in \widehat{\partial} f_c(\bar{x}) \\ \iff & \langle v, w\rangle \leq \langle \bar{v}, w\rangle + \min\langle DF(\bar{x})(w), \bar{x} - \pi_c(\bar{x})\rangle \quad \forall w \in \mathbb{R}^n, \\ \iff & \langle v - \bar{v}, w\rangle \leq \langle z, \bar{x} - \pi_c(\bar{x})\rangle \quad \forall (w, z) \in \mathrm{gph}(DF(\bar{x})) = T_{\mathrm{gph}\, F}(\bar{x}, F(\bar{x})), \\ \iff & (v - \bar{v}, -\bar{x} + \pi_c(\bar{x})) \in (T_{\mathrm{gph}\, F}(\bar{x}, F(\bar{x})))^* = \widehat{N}_{\mathrm{gph} F}(\bar{x}, F(\bar{x})), \\ \iff & v - \bar{v} \in \widehat{D}^* F(\bar{x})(\bar{x} - \pi_c(\bar{x})). \end{aligned}$$

This provides us with the formula for $\widehat{\partial} f_c(\bar{x})$.

To show $\partial f_c(\bar{x}) \subset U := \big(D^* F(\bar{x}) - cI\big)\,(\bar{x} - \pi_c(\bar{x})) + F(\bar{x})$, let $v \in \partial f_c(\bar{x})$. By the formula for $\widehat{\partial} f_c(x_k)$, there are sequences $x_k \to \bar{x}$ and $v_k \to v$ such that

$$(v_k - \bar{v}_k, \pi_c(x_k) - x_k) \in \widehat{N}_{\mathrm{gph}\, F}(x_k, F(x_k)) \quad \forall k,$$

where $\bar{v}_k := F(x_k) - c(x_k - \pi_c(x_k))$. Since F and π_c are locally Lipschitz continuous functions (cf. Lemma 3.1 **(f)**), we have $\bar{v}_k \to F(\bar{x}) - c(\bar{x} - \pi_c(\bar{x})), x_k - \pi_c(x_k) \to \bar{x} - \pi_c(\bar{x})$. Consequently,

$$(v - F(\bar{x}) + c(\bar{x} - \pi_c(\bar{x})), \pi_c(\bar{x}) - \bar{x}) \in N_{\mathrm{gph}\, F}(\bar{x}, F(\bar{x})),$$

or equivalently,

$$v - F(\bar{x}) + c(\bar{x} - \pi_c(\bar{x})) \in D^* F(\bar{x})(\bar{x} - \pi_c(\bar{x})).$$

This verifies that $v \in U$ and hence that $\partial f_c(\bar{x}) \subset U$.

To show $U \subset \partial f_c(\bar{x})$, let $v \in \left(D^*F(\bar{x}) - cI\right)(\bar{x} - \pi_c(\bar{x})) + F(\bar{x})$. Then we have

$$z{:=}v + c(\bar{x}-\pi_c(\bar{x}))-F(\bar{x}) \in D^*F(\bar{x})(\bar{x}-\pi_c(\bar{x})) \iff (z, -\bar{x}+\pi_c(\bar{x}))\in N_{\text{gph}\,F}(\bar{x}, F(\bar{x})).$$

According to the definitions of normal cone (cf. Definition 2.1) and the regular coderivative (cf. Definition 2.5), there exist sequences $x_k \to \bar{x}$, $z_k \to z$ and $w_k \to \bar{x} - \pi_c(\bar{x})$ such that for all k,

$$(z_k, -w_k) \in \widehat{N}_{\text{gph}\,F}(x_k, F(x_k)) \iff (z_k, -w_k) \in (\text{gph}\, DF(x_k))^*,$$

or explicitly,

$$\langle z_k, w\rangle - \langle x_k - \pi_c(x_k), z\rangle \le \langle w_k - x_k + \pi_c(x_k), z\rangle \quad \forall z \in DF(x_k)(w). \tag{9}$$

By the Cauchy-Schwarz inequality and Lemma 2.2 **(c)**, we have for all k,

$$\langle w_k - x_k + \pi_c(x_k), z\rangle \le \epsilon_k \|w\| \quad \forall z \in DF(x_k)(w),$$

where $\epsilon_k := \text{lip}F(x_k)\|w_k - x_k + \pi_c(x_k)\|$. It then follows from (9) that for all k,

$$\langle z_k, w\rangle \le \min\langle DF(x_k)(w), x_k - \pi_c(x_k)\rangle + \epsilon_k\|w\| \quad \forall w \in \mathbb{R}^n.$$

By the formula for the subderivative $df_c(x_k)(w)$, we have for all k,

$$\langle z_k - c(x_k - \pi_c(x_k)) + F(x_k),\ w\rangle \le df_c(x_k)(w) + \epsilon_k\|w\| \quad \forall w \in \mathbb{R}^n. \tag{10}$$

Since F and π_c are locally Lipschitz continuous functions (cf. Lemma 3.1 **(f)**), taking the limit as $k \to +\infty$ yields

$$z_k - c(x_k - \pi_c(x_k)) + F(x_k) \to z - c(\bar{x} - \pi_c(\bar{x})) + F(\bar{x}) = v,$$

and $\epsilon_k \to 0$ (due to the upper semicontinuity of lip $F(\cdot)$ [37, Theorem 9.2] and $w_k - x_k + \pi_c(x_k) \to 0$). Then by [37, Proposition 10.46] and (10), we have $v \in \partial f_c(\bar{x})$. This completes the proof. □

By virtue of the formula for the limiting subdifferential set $\partial f_c(\bar{x})$ in Proposition 3.1, we readily derive the formula for the Clarke subdifferential set $\overline{\partial} f_c(\bar{x})$, initially obtained in [45, Lemma 3.2].

Corollary 3.1 *Let $\bar{x} \in \mathbb{R}^n$. We have*

$$\overline{\partial} f_c(\bar{x}) = \left(\overline{\partial} F(\bar{x})^T - cI\right)(\bar{x} - \pi_c(\bar{x})) + F(\bar{x}),$$

where $\overline{\partial} F(\bar{x})$ denotes the generalized Jacobian of F at $\bar{x}$ (cf. (3)).

Proof By Lemma 3.1 **(f)** and Lemma 2.1 **(c)**, the function f_c is locally Lipschitz continuous. Consequently, we have $\overline{\partial} f_c(\bar{x}) = \text{co}(\partial f_c(\bar{x}))$. The formula for $\overline{\partial} f_c(\bar{x})$ then follows directly from Proposition 3.1 and the coderivative duality given by (4). This completes the proof. □

3.2 Subderivatives and subgradients of f_{ab}

Following a parallel approach to Subsection 3.1, we present some differential properties of the D-gap function f_{ab} in this subsection. Most proofs are omitted as they closely mirror those in Subsection 3.1.

Proposition 3.2 *Let $\bar{x} \in \mathbb{R}^n$ and $w \in \mathbb{R}^n$. We have the following formulas:*

$$\begin{aligned} df_{ab}(\bar{x})(w) &= (b-a)\langle \bar{x} - \pi_a(\bar{x}),\ w\rangle + \min\langle (DF(\bar{x}) - bI)\, w,\ \pi_b(\bar{x}) - \pi_a(\bar{x})\rangle, \\ \widehat{\partial} f_{ab}(\bar{x}) &= \left(\widehat{D}^* F(\bar{x}) - bI\right)(\pi_b(\bar{x}) - \pi_a(\bar{x})) + (b-a)(\bar{x} - \pi_a(\bar{x})), \\ \partial f_{ab}(\bar{x}) &= (D^* F(\bar{x}) - bI)\,(\pi_b(\bar{x}) - \pi_a(\bar{x})) + (b-a)(\bar{x} - \pi_a(\bar{x})), \end{aligned}$$

where

$$\min\langle (DF(\bar{x}) - bI)\, w,\ \pi_b(\bar{x}) - \pi_a(\bar{x})\rangle := \min_{v \in DF(\bar{x})(w)} \langle (v - bw)\,,\ \pi_b(\bar{x}) - \pi_a(\bar{x})\rangle.$$

Proof Since $f_{ab} = f_a - f_b$ is a locally Lipschitz continuous function, we have

$$df_{ab}(\bar{x})(w) = \liminf_{t \to 0_+} \left[\frac{f_a(\bar{x} + tw) - f_a(\bar{x})}{t} - \frac{f_b(\bar{x} + tw) - f_b(\bar{x})}{t} \right].$$

By Lemma 3.1 **(g)**, the following inequalities hold for all t:

$$f_a(\bar{x}) \geq \langle F(\bar{x}), \bar{x} - \pi_a(\bar{x} + tw)\rangle - \frac{a}{2}||\bar{x} - \pi_a(\bar{x} + tw)||^2$$

and

$$f_b(\bar{x} + tw) \geq \langle F(\bar{x} + tw), \bar{x} + tw - \pi_b(\bar{x})\rangle - \frac{b}{2}||\bar{x} + tw - \pi_b(\bar{x})||^2.$$

Combining these with (8) and the local Lipschitz continuity of π_a and π_b (see Lemma 3.1 **(f)**), we derive

$$\begin{aligned} df_{ab}(\bar{x})(w) \leq\ & \liminf_{t \to 0_+} \langle \frac{F(\bar{x} + tw) - F(\bar{x})}{t}, \pi_b(\bar{x}) - \pi_a(\bar{x} + tw)\rangle \\ & - \lim_{t \to 0_+} \frac{a}{2} \frac{||\bar{x} + tw - \pi_a(\bar{x} + tw)||^2 - ||\bar{x} - \pi_a(\bar{x} + tw)||^2}{t} \\ & + \lim_{t \to 0_+} \frac{b}{2} \frac{||\bar{x} + tw - \pi_b(\bar{x})||^2 - ||\bar{x} - \pi_b(\bar{x})||^2}{t} \\ =\ & \min_{v \in DF(\bar{x})(w)} \langle v, \pi_b(\bar{x}) - \pi_a(\bar{x})\rangle + \langle b(\bar{x} - \pi_b(\bar{x})) - a(\bar{x} - \pi_a(\bar{x})), w\rangle. \end{aligned}$$

To establish the reverse inequality, we use a similar approach. By Lemma 3.1 **(g)**, for all t, we have

$$f_a(\bar{x} + tv) \geq \langle F(\bar{x} + tv), \bar{x} + tv - \pi_a(\bar{x})\rangle - \frac{a}{2}||\bar{x} + tv - \pi_a(\bar{x})||^2$$

and

$$f_b(\bar{x}) \geq \langle F(\bar{x}), \bar{x} - \pi_b(\bar{x} + tv)\rangle - \frac{b}{2}||\bar{x} - \pi_b(\bar{x} + tv)||^2.$$

This completes the proof of the formula for $df_{ab}(\bar{x})(w)$. The remaining two formulas can be derived analogously to Proposition 3.1. □

Corollary 3.2 *Let $\bar{x} \in \mathbb{R}^n$. The following properties hold:*

(a) *We have the formula for the Clarke subdifferential set of f_{ab} at $\bar{x}$ as follows:*

$$\overline{\partial} f_{ab}(\bar{x}) = \left(\overline{\partial} F(\bar{x})^T - bI\right)(\pi_b(\bar{x}) - \pi_a(\bar{x})) + (b-a)(\bar{x} - \pi_a(\bar{x})).$$

(b) *$\bar{x}$ solves (VIP) if and only if $0 \in \partial f_{ab}(\bar{x})$ and $\pi_a(\bar{x}) = \pi_b(\bar{x})$.*

Remark 3.1 The formula for $\overline{\partial} f_{ab}(\bar{x})$ was first obtained in [45, Lemma 3.3]. Subsequently, it was also derived in [28, Theorem 4.1] and [21, Theorem 3.1] for some generalized D-gap functions. According to the generalized Fermat's rule [37, Theorem 10.1], the condition

$$0 \in \partial f_{ab}(\bar{x}) \tag{11}$$

is necessary for $\bar{x}$ to be locally optimal for the optimization problem

$$\min_{x \in \mathbb{R}^n} f_{ab}(x),$$

and thus it is necessary for $\bar{x}$ to be a solution of (VIP) (cf. Lemma 3.1 **(c)**). Another necessary condition for $\bar{x}$ to be a solution of (VIP), by Lemma 3.1 **(c)**, is the equality

$$\pi_a(\bar{x}) = \pi_b(\bar{x}). \tag{12}$$

Although these two necessary conditions together are sufficient for $\bar{x}$ to be a solution of (VIP), it is interesting to note that neither of them alone is sufficient.

It was shown in [21, Theorem 4.3] that $\bar{x}$ solves (VIP) if and only if $0 \in \overline{\partial} f_{ab}(\bar{x})$ and

$$\left.\begin{array}{c} w \in T_{ab}(x, F, K), \quad Z \in \overline{\partial} F(x) \\ Z^T w \in T_{ab}(x, F, K)^* \end{array}\right\} \Rightarrow F(x)^T w = 0, \tag{13}$$

where $T_{ab}(x, F, K)$ is a cone defined as in Lemma 3.1 **(e)**. However, by resorting to Corollary 3.2 **(b)** and noting that $\overline{\partial} f_{ab}(\bar{x}) = \partial f_{ab}(\bar{x})$ in the presence of (12), we can refine [21, Theorem 4.3] as follows: $\bar{x}$ solves (VIP) if and only if $0 \in \overline{\partial} f_{ab}(\bar{x})$ and (12) holds. Note that $\pi_a(\bar{x})$ and $\pi_b(\bar{x})$ are involved in the definition of $T_{ab}(x, F, K)$. So in contrast to the verification of (13), it is much easier to verify (12).

It is also worth noting that (12) is implied by (11) whenever the inequality

$$d(0, \partial f_{ab}(\bar{x})) \geq \mu |\pi_b(\bar{x}) - \pi_a(\bar{x})| \tag{14}$$

holds for some $\mu > 0$. Inequalities in the form of (14) will play a crucial role in the next section.

4 The Kurdyka-Łojasiewicz inequality and error bounds of f_{ab}

In this section, we investigate the KL inequality and error bounds for the D-gap function f_{ab} by utilizing the formula for the limiting subdifferential sets $\partial f_{ab}(x)$ presented in the previous section. Before stating our main results in Theorem 4.1, we provide in Lemmas 4.1-4.4 several results regarding the necessary and sufficient conditions for the following inequality:

$$d(0, \partial f_{ab}(x)) \geq \mu \|\pi_b(x) - \pi_a(x)\| \quad \forall x \in V,$$

where V is some open set in $\mathbb{R}^n$.

Lemma 4.1 *Let $x \in \mathbb{R}^n$ and let $\mu > 0$. If $d(0, \partial f_{ab}(x)) \geq \mu \|\pi_b(x) - \pi_a(x)\|$, then*

$$d(0, \partial f_{ab}(x)) \geq \frac{\mu(b-a)}{\mu + b + \operatorname{lip} F(x)} \|x - \pi_a(x)\|. \tag{15}$$

Proof Let $w := \pi_b(x) - \pi_a(x)$ and let $u := x - \pi_a(x)$. By applying the formula for $\partial f_{ab}(x)$ in Proposition 3.2, we can find some $z^* \in D^*F(x)(w)$ such that $d(0, \partial f_{ab}(x)) = \|z^* - bw + (b-a)u\|$. Then we get (15) as follows:

$$\begin{aligned} d(0, \partial f_{ab}(x)) &\geq -\|z^*\| - b\|w\| + (b-a)\|u\| \\ &\geq -(b + \operatorname{lip} F(x))\|w\| + (b-a)\|u\| \\ &\geq -\tfrac{b+\operatorname{lip} F(x)}{\mu} d(0, \partial f_{ab}(x)) + (b-a)\|u\|, \end{aligned}$$

where the first inequality follows from the triangle inequality, the second one from Lemma 2.2 **(d)**, and the last one from the assumption that $d(0, \partial f_{ab}(x)) \geq \mu\|w\|$. This completes the proof. □

Lemma 4.2 *Assume that* lip $F(x)$ *is bounded from above on a nonempty subset* V *of* $\mathbb{R}^n$*, as is true in particular when* V *is bounded. Then the following properties are equivalent:*

(a) *There is some* $\mu > 0$ *such that* $d(0, \partial f_{ab}(x)) \geq \mu\sqrt{f_{ab}(x)} \quad \forall x \in V$.
(b) *There is some* $\mu > 0$ *such that* $d(0, \partial f_{ab}(x)) \geq \mu\|x - \pi_a(x)\| \quad \forall x \in V$.
(c) *There is some* $\mu > 0$ *such that* $d(0, \partial f_{ab}(x)) \geq \mu\|\pi_b(x) - \pi_a(x)\| \quad \forall x \in V$.

Therefore, f_{ab} *satisfies the KL inequality at any solution* $\bar{x}$ *of (VIP) with an exponent of* $\frac{1}{2}$ *if and only if any of* **(a)**, **(b)** *and* **(c)** *holds with* V *being some neighborhood of* $\bar{x}$.

Proof The relations **(a)** $\iff$ **(b)** $\implies$ **(c)** follow directly from Lemma 3.1 **(a)**. Since lip $F(x)$ is upper semicontinuous ([37, Theorem 9.2]), it follows from [37, Corollary 1.10] that lip $F(x)$ is bounded from above on each bounded subset of $\mathbb{R}^n$. We now show **(c)** $\implies$ **(b)** by assuming that **(c)** holds for some $\mu > 0$ and that there is some $L > 0$ such that lip $F(x) \leq L \; \forall x \in V$. By Lemma 4.1, we get **(b)** as follows:

$$d(0, \partial f_{ab}(x)) \geq \frac{\mu(b-a)}{\mu + b + \operatorname{lip} F(x)}\|x - \pi_a(x)\| \geq \frac{\mu(b-a)}{\mu + b + L}\|x - \pi_a(x)\| \quad \forall x \in V.$$

Let $\bar{x}$ be a solution of (VIP). Firstly, we note that f_{ab} is locally Lipschitz continuous with $f_{ab} \geq 0$ and $f_{ab}(\bar{x}) = 0$ (cf. Lemma 3.1 **(c)**). Then f_{ab} satisfies the KL inequality at $\bar{x}$ with an exponent of $\frac{1}{2}$ provided that, according to Definition 2.3, **(a)** holds with V being some bounded neighborhood of $\bar{x}$. By the previous argument, **(a)**, **(b)** and **(c)** are equivalent whenever V is bounded. Hence, the last assertion is true. This completes the proof. □

Lemma 4.3 *Assume that the solution set of (VIP) is nonempty. If there are some* $\mu \in (0, +\infty)$ *and* $\varepsilon \in (0, +\infty]$ *such that*

$$d\left(0, \partial f_{ab}(x)\right) \geq \mu\|\pi_b(x) - \pi_a(x)\| \quad \forall x \in [f_{ab} < \varepsilon], \tag{16}$$

and

$$L := \sup_{x \in [0 < f_{ab} < \varepsilon]} \operatorname{lip} F(x) < +\infty, \tag{17}$$

then

$$\sqrt{\frac{b-a}{2}}\frac{\mu}{\mu + b + L}\, d\left(x, [f_{ab} \leq \theta]\right) \leq \left(\sqrt{f_{ab}(x)} - \sqrt{\theta}\right)_+ \quad \forall \theta \in [0, \varepsilon),\; \forall x \in [f_{ab} < \varepsilon], \tag{18}$$

which, in particular, implies the following error bound property:

$$\sqrt{\frac{b-a}{2}}\frac{\mu}{\mu + b + L}\, d\left(x, [f_{ab} \leq 0]\right) \leq \sqrt{f_{ab}(x)} \quad \forall x \in [f_{ab} \leq \varepsilon].$$

Proof It suffices to show (18) by assuming (16) and (17) for some given $\mu \in (0, +\infty)$ and $\varepsilon \in (0, +\infty]$. Since the solution set of (VIP) is nonempty, we infer from Lemma 3.1 **(c)** that $[f_{ab} \leq 0] \neq \emptyset$. In the following, we assume that $[0 < f_{ab} < \varepsilon]$ is nonempty; otherwise, (18) holds trivially. Fix an $x \in [0 < f_{ab} < \varepsilon]$. In light of (16) and (17), it follows from Lemma 4.1 that

$$d(0, \partial f_{ab}(x)) \geq \frac{\mu(b-a)}{\mu+b+L}\|x - \pi_a(x)\|.$$

Then by Lemma 3.1 **(a)**, we have

$$d(0, \partial f_{ab}(x)) \geq \frac{\mu\sqrt{2(b-a)}}{\mu+b+L}\sqrt{f_{ab}(x)}.$$

Through some straightforward calculation, we have $\partial\sqrt{f_{ab}}(x) = \dfrac{\partial f_{ab}(x)}{2\sqrt{f_{ab}}(x)}$ and thus

$$d\left(0, \partial\sqrt{f_{ab}}(x)\right) \geq \sqrt{\frac{b-a}{2}}\frac{\mu}{\mu+b+L}.$$

Then by [24, Lemma 2.1 **(ii')**], we have

$$|\nabla\sqrt{f_{ab}}|(x) \geq \sqrt{\frac{b-a}{2}}\frac{\mu}{\mu+b+L},$$

where for a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $\bar{y} \in \mathbb{R}^n$,

$$|\nabla f|(\bar{y}) := \limsup_{y\to\bar{y},\ y\neq\bar{y}} \frac{(f(\bar{y}) - f(y))_+}{\|y-\bar{y}\|}$$

denotes the the strong slope of f at $\bar{y}$, introduced by De Giorgi et al. [14]. Since $x \in [0 < f_{ab} < \varepsilon]$ is chosen arbitrarily, we can apply [6, Theorem 2.1] to deduce that

$$\begin{aligned}\inf_{0\leq\sqrt{\theta}<\sqrt{\varepsilon}}\ \inf_{x\in[\sqrt{\theta}<\sqrt{f_{ab}}<\sqrt{\varepsilon}]} \frac{\sqrt{f_{ab}(x)} - \sqrt{\theta}}{d\left(x, \left[\sqrt{f_{ab}} \leq \sqrt{\theta}\right]\right)} &= \inf_{x\in[0<\sqrt{f_{ab}}<\sqrt{\varepsilon}]} |\nabla\sqrt{f_{ab}}|(x)\\ &\geq \sqrt{\frac{b-a}{2}}\frac{\mu}{\mu+b+L},\end{aligned}$$

from which, (18) follows readily. This completes the proof. □

There are numerous existing conditions in the literature that are sufficient for Lemma 4.2 **(c)** or (16). This can be observed from the following lemma, in which we also present a new sufficient condition that can be regarded as a form of restricted strong monotonicity.

Lemma 4.4 *Let $\mu > 0$ and let $V \subset \mathbb{R}^n$ be open. Consider the following properties:*

(a) *F is strongly monotone on V with modulus μ, which holds in the case of V being convex if and only if the following inequality holds for all $x \in V$ where F is differentiable:*

$$\langle\nabla F(x)w,\ w\rangle \geq \mu||w||^2 \quad \forall w \in \mathbb{R}^n. \tag{19}$$

(b) *The following holds for all $x \in V$ where F is differentiable:*

$$\langle\nabla F(x)w,\ w\rangle \geq \mu||w||^2 \quad \forall w \in T_{ab}(x, F, K).$$

(c) *The following holds for all $x \in V$ where F is differentiable:*

$$\langle \nabla F(x)(\pi_a(x) - \pi_b(x)),\ \pi_a(x) - \pi_b(x)\rangle \geq \mu||\pi_a(x) - \pi_b(x)||^2.$$

(d) $d(0, \partial f_{ab}(x)) \geq \mu\|\pi_b(x) - \pi_a(x)\| \quad \forall x \in V.$

We have **(a)** $\Longrightarrow$ **(b)** $\Longrightarrow$ **(c)** $\Longrightarrow$ **(d)**.

Proof According to [16, Proposition 2.3 **(b)**], the following inequality holds for all $x \in V$:

$$\langle Zw,\ w\rangle \geq \mu||w||^2 \quad \forall Z \in \overline{\nabla} F(x),\ \forall w \in \mathbb{R}^n, \tag{20}$$

if F is strongly monotone on V with modulus μ. Moreover, the converse is true whenever V is convex. Since $\nabla F(x) \in \overline{\nabla} F(x)$ when F is differentiable at x, (19) is implied by (20). To prove that (20) is implied by (19), let $x \in V$ and let $Z \in \overline{\nabla} F(x)$. By the definition of $\overline{\nabla} F(x)$ (cf. (2)), there is a sequence $x_k \to x$ such that F is differentiable at x_k for all k and $\nabla F(x_k) \to Z$. Then by (19), we have for all sufficiently large k that

$$\langle \nabla F(x_k)w,\ w\rangle \geq \mu||w||^2 \quad \forall w \in \mathbb{R}^n,$$

which implies (20) by taking the limit as $k \to \infty$.

By the previous argument, we get **(b)** from **(a)** in a straightforward way. To get **(c)** from **(b)**, it suffices to note the following facts: (1) $\pi_a(x) - \pi_b(x) \in T_{ab}(x, F, K)$ (cf. Lemma 3.1 **(e)**); (2) $\pi_a(x) = \pi_b(x)$ whenever $f_{ab}(x) = 0$ (cf. Lemma 3.1 **(c)**).

We now show **(c)** $\Longrightarrow$ **(d)**. Let $x \in V$. Set $w := \pi_b(x) - \pi_a(x)$ and $u := x - \pi_a(x)$. We first claim that the following inequality holds for all $z^* \in \operatorname{conv} D^*F(x)(w)$:

$$\langle z^*, w\rangle \geqslant \mu||w||^2. \tag{21}$$

By the coderivative duality (4) for a locally Lipschitz continuous mapping, we have $z^* \in \{A^T w \mid A \in \operatorname{conv} \overline{\nabla} F(x)\}$. Then there exist a positive integer r and some $A^i \in \overline{\nabla} F(x)$ such that

$$z^* = \left(\sum_{i=1}^{r} \lambda^i A^i\right)^T w = \sum_{i=1}^{r} \lambda^i \left(A^i\right)^T w, \tag{22}$$

where $\lambda^i \geq 0$ for all i and $\sum_{i=1}^r \lambda^i = 1$. For each $A^i \in \overline{\nabla} F(x)$, by its definition, there exists some sequence $\{x_k^i\}$ such that F is differentiable at x_k^i for all k, $x_k^i \to x$ and $\nabla F(x_k^i) \to A^i$ as $k \to \infty$. Then by **(c)**, for all k large enough, we have

$$\langle \nabla F(x_k^i)(\pi_a(x_k^i) - \pi_b(x_k^i)), \pi_a(x_k^i) - \pi_b(x_k^i)\rangle \geqslant \mu||\pi_b(x_k^i) - \pi_a(x_k^i)||^2.$$

Thus, by taking into account that π_a and π_b are locally Lipschitz continuous and letting $k \to \infty$, we get

$$\langle A^i(\pi_a(x) - \pi_b(x)), \pi_a(x) - \pi_b(x)\rangle \geqslant \mu||\pi_b(x) - \pi_a(x)||^2.$$

In terms of w, this is equivalent to $\langle (A^i)^T w, w\rangle \geq \mu\|w\|^2$. Combined with (22), this leads to (21).

By applying the formula for $\partial f_{ab}(x)$ in Proposition 3.2, we can find some $\bar{z}^* \in D^*F(x)(w) \subset \operatorname{conv} D^*F(x)(w)$ such that $d(0, \partial f_{ab}(x)) = \|\bar{z}^* - bw + (b-a)u\|$. Then we get **(d)**, since we have

$$d(0, \partial f_{ab}(x))\,\|w\| \geq \langle \bar{z}^* - bw + (b-a)u,\ w\rangle \geq \langle \bar{z}^*, w\rangle \geq \mu\|w\|^2,$$

where the first inequality follows from the Cauchy-Schwarz inequality, the second one from Lemma 3.1 **(d)**, and the last one from (21). This completes the proof. ◻

Remark 4.1 Since $\nabla F(x) \in \overline{\nabla} F(x) \subset \overline{\partial} F(x)$ when F is differentiable at x, Lemma 4.4 **(b)** holds if the following inequality holds for all $x \in V$:

$$\langle Z^T w, w \rangle \geq \mu ||w||^2 \quad \forall Z \in \overline{\partial} F(x), \ \forall w \in T_{ab}(x, F, K). \tag{23}$$

The supremum of all possible positive μ satisfying (23) for all $x \in \mathbb{R}^n$ with $f_{ab}(x) > 0$ can be reformulated as

$$\mu_{ab} := \inf\{w^T Z w \mid Z \in \overline{\partial} F(x), \ w \in T_{ab}(x, F, K), \ \|w\| = 1, \ f_{ab}(x) > 0\}. \tag{24}$$

The quantity μ_{ab} was first introduced for a general case in [21, Theorem 4.2], where the condition $\mu_{ab} > 0$ was utilized to study the local error bounds for f_{ab}.

Remark 4.2 Lemma 4.4 **(c)** can be easily reformulated as

$$\langle z^*, \pi_b(x) - \pi_a(x) \rangle \geq \mu ||\pi_a(x) - \pi_b(x)||^2 \quad \forall x \in V, \ z^* \in \text{conv } D^* F(x)(\pi_b(x) - \pi_a(x)),$$

or

$$\langle z, \pi_a(x) - \pi_b(x) \rangle \geq \mu ||\pi_a(x) - \pi_b(x)||^2 \quad \forall x \in V, \ z \in \text{conv } D_* F(x)(\pi_a(x) - \pi_b(x)),$$

where $D_* F(x)$ stands for the strict derivative mapping of F at x (cf. (6)).

Moreover, it should be noticed that $\pi_a(x) - \pi_b(x) \in K - K \subset L$ for all $x \in \mathbb{R}^n$, where L is the parallel linear subspace of K. This implies that Lemma 4.4 **(c)** holds if the following stronger property holds for all $x \in V$ where F is differentiable:

$$\langle \nabla F(x) w, \ w \rangle \geq \mu ||w||^2 \quad \forall w \in L.$$

Evidently, verifying the aforementioned property is much simpler than verifying Lemma 4.4 **(c)**. Whenever L is a proper subset of $\mathbb{R}^n$ (i.e., $L \subsetneq \mathbb{R}^n$), the aforementioned property can still be less restricted than the strong monotonicity stated in Lemma 4.4 **(a)**. The following example serves to illustrate this case.

Example 4.1 Consider an affine variational inequality problem (VIP) defined by the set $K := \{x \in \mathbb{R}^n | Cx \leq d\}$ and the mapping $F(x) := Ax + q$, where $C \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$, $A \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$. For this problem, verifying the strong monotonicity of F (i.e., Lemma 4.4 **(a)**) requires showing that A is positive-definite on the entire space $\mathbb{R}^n$. Assume that no row $C_1, \ldots, C_m$ of C is a zero vector. Let $\mathcal{J} := \{j \in \{1, \ldots, m\} | C_j y = d_j\}$, where $y \in \mathbb{R}^n$ is an arbitrary relative interior point of K. The linear parallel subspace L of K is then the kernel $\ker C_{\mathcal{J}}$ of the matrix $C_{\mathcal{J}}$, where $C_{\mathcal{J}}$ denotes the submatrix of C consisting of the rows C_j with $j \in \mathcal{J}$, and $\ker C_{\mathcal{J}}$ is set by convention to be $\mathbb{R}^n$ when $\mathcal{J} = \emptyset$. As explained in Remark 4.2, to verify Lemma 4.4 **(c)**, it suffices to show that A is positive-definite on $\ker C_{\mathcal{J}}$. This condition is strictly less restrictive than the strong monotonicity of F when $\mathcal{J} \neq \emptyset$ or equivalently the Slater condition for the linear system $Cx \leq d$ fails to hold.

In the case where F is an affine mapping and K is the entire space $\mathbb{R}^n$, we can provide details regarding how the properties in Lemma 4.4 differ from one another in the following example.

Example 4.2 Let $A \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$ be such that $q + \text{rge } A \neq \{0\}$, where rge A denotes the range space of A. Consider an instance of (VIP) with $K = \mathbb{R}^n$ and $F(x) = Ax + q$. In this case, finding a solution of (VIP) amounts to finding a solution to the linear equations $Ax + q = 0$, which exists if and only if $q \in \text{rge } A$. Clearly, F is continuously differentiable on $\mathbb{R}^n$ with $\nabla F(\cdot) = A$, which implies that f_{ab} is continuously differentiable on $\mathbb{R}^n$. Through some direct computation, we have

$$\pi_b(x) - \pi_a(x) = \frac{b-a}{ab}(Ax+q), \quad f_{ab}(x) = \frac{b-a}{2ab}\|Ax+q\|^2,$$

and

$$\nabla f_{ab}(x) = \frac{b-a}{ab}A^T(Ax+q), \quad T_{ab}(x, F, K) = \{w \mid \langle Ax+q, w\rangle \le 0\}.$$

Then, in the case of $V = \mathbb{R}^n$, Lemma 4.4 **(a)-(d)** can be reduced respectively to the following:

(a) $A - \mu I$ is positive-semidefinite on $\mathbb{R}^n$.

(b) $A - \mu I$ is positive-semidefinite on at least one closed-half space containing the origin and hence on the whole space $\mathbb{R}^n$. (Therefore, **(a)** and **(b)** coincide, both of which implies that A is positive-definite on $\mathbb{R}^n$ and that the linear equatios $Ax + q = 0$ has a unique solution.)

(c) $A - \mu I$ is positive-semidefinite on the linear subspace $\mathbb{R}\{q\} + \operatorname{rge} A$, which entails positive-semidefiniteness of $A^TAA - \mu A^TA$ on $\mathbb{R}^n$ and is equivalent to it when $q \in \operatorname{rge} A$. (The latter property can be fulfilled for a symmetric matrix A if and only if A is positive-semidefinite and $0 < \mu < \lambda_i$, where λ_i is any positive eigenvalue of A.)

(d) $AA^T - \mu^2 I$ is positive-semidefinite on the linear subspace $\mathbb{R}\{q\} + \operatorname{rge} A$, which entails positive-semidefiniteness of $(A^TA)^2 - \mu^2 A^TA$ on $\mathbb{R}^n$ and is equivalent to it when $q \in \operatorname{rge} A$. (The latter property can be fulfilled as long as $0 < \mu \le \sqrt{\lambda_i}$, where λ_i is any positive eigenvalue of A^TA.)

Therefore, in the case where $q \in \operatorname{rge} A$ and A is symmetric and positive-semidefinite (but not positive-definite), Lemma 4.4 **(a)-(b)** cannot hold. However, Lemma 4.4 **(c)** can hold as long as $0 < \mu < \lambda_i$, where λ_i is any positive eigenvalue of A. This demonstrates that Lemma 4.4 **(c)** can be strictly weaker than Lemma 4.4 **(a)-(b)**. On the other hand, in the case where $q \in \operatorname{rge} A$ and A is symmetric but not positive-semidefinite, Lemma 4.4 **(c)** cannot hold. Nevertheless, Lemma 4.4 **(d)** can hold as long as μ is less than or equal to the square root of the smallest positive eigenvalue of A^TA. This demonstrates that Lemma 4.4 **(d)** can be strictly weaker than Lemma 4.4 **(c)**.

Theorem 4.1 *Assume that any of* **(a)-(d)** *in Lemma 4.4 holds with some $\mu > 0$ and $V = \mathbb{R}^n$. Then the following properties hold:*

(a) *f_{ab} is a KL function with an exponent of $\frac{1}{2}$.*

(b) *If F is coercive on $\mathbb{R}^n$, then the solution set of (VIP) is nonempty and compact, and $\sqrt{f_{ab}}$ has a local error bound on $\mathbb{R}^n$, i.e., the following holds for any given $\varepsilon > 0$:*

$$\sqrt{\frac{b-a}{2}}\frac{\mu}{\mu+b+L}\, d\left(x, [f_{ab} \le 0]\right) \le \sqrt{f_{ab}(x)} \quad \forall x \in [f_{ab} \le \varepsilon].$$

where L is any number such that $L \ge \operatorname{lip} F(x)$ for all $x \in [0 < f_{ab} < \varepsilon]$.

(c) *If the solution set of (VIP) is nonempty and F is globally Lipschitz continuous with a constant $L > 0$, then $\sqrt{f_{ab}}$ has a global error bound on $\mathbb{R}^n$, i.e., the following holds:*

$$\sqrt{\frac{b-a}{2}}\frac{\mu}{\mu+b+L}\, d\left(x, [f_{ab} \le 0]\right) \le \sqrt{f_{ab}(x)} \quad \forall x \in \mathbb{R}^n.$$

Proof For each x that is a solution of (VIP), it follows from Lemma 4.2 that f_{ab} is a KL function at x with an exponent of $\frac{1}{2}$. For each x that is not a solution of (VIP), we assert that $0 \notin \partial f_{ab}(x)$, and thus f_{ab} is a KL function at x with an exponent of 0. Otherwise, if $0 \in \partial f_{ab}(x)$, together with the equality $\pi_a(x) = \pi_b(x)$ which can be guaranteed by Lemma 4.4 **(d)**, it would imply that x is a solution of (VIP) (cf. Corollary 3.2 **(b)**). Overall, f_{ab} is indeed a KL function with an exponent of $\frac{1}{2}$. This verifies **(a)**.

To show **(b)**, fix any $\varepsilon > 0$ and let $\bar{L} := \sup_{x\in[0<f_{ab}<\varepsilon]} \operatorname{lip} F(x)$. By the coerciveness of F on $\mathbb{R}^n$ (hence on K), the solution set of (VIP) is nonempty and compact (cf. [12, Proposition 2.2.7]), and the level set $[f_{ab} \le \varepsilon]$ is bounded (cf. [21, Lemma 4.1]). Since $\operatorname{lip} F(x)$ is upper semicontinuous (cf. [37, Theorem 9.2]), it follows from [37, Corollary 1.10] that $\operatorname{lip} F(x)$ is bounded from above on each bounded subset of $\mathbb{R}^n$. So we have $\bar{L} < +\infty$. Then by Lemma 4.3, we get **(b)** in a straightforward manner.

To show **(c)**, we apply Lemma 4.3 again by noting that

$$\sup_{x\in[0<f_{ab}<+\infty]} \operatorname{lip} F(x) \le L.$$

This completes the proof. □

Remark 4.3 In the case where Lemma 4.4 **(a)** holds for some $\mu > 0$ and $V = \mathbb{R}^n$ (i.e., F is strongly monotone on $\mathbb{R}^n$ with modulus μ), it was pointed out by [21, Remark 2.1 (ii)] that F is coercive on $\mathbb{R}^n$. In this situation, Theorem 4.1 **(b)** holds without explicitly assuming coerciveness. On the other hand, when Lemma 4.4 **(b)** holds with $V = \mathbb{R}^n$ and some $\mu > 0$, Theorem 4.1 **(b)** can be deduced from [21, Theorem 4.2](cf. Remark 4.1). To the best of our knowledge, all the results in Theorem 4.1, except for the ones mentioned above, are new.

Example 4.3 ([21], Example 4.4) Consider an instance of (VIP) with $K = \mathbb{R}^2_+$ and $F : \mathbb{R}^2 \to \mathbb{R}^2$ being given by $F(x) = \big(x_1 + (x_1)_+(x_2)_+,\quad x_2 + \frac{3}{2}(x_1)_+\big)^T$. Clearly, F is differentiable at $x \in \mathbb{R}^2$ if and only if $x_1x_2 \ne 0$, and moreover,

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

Let $a \in (0, 1)$ and $b = 1$. According to [21, Example 4.4], F is coercive and not monotone on $\mathbb{R}^2$, and $\sqrt{f_{ab}}$ has a local error bound on $\mathbb{R}^2$ (with some error bound modulus expressed in an abstract way), and $\mu_{ab} \ge 1$, where μ_{ab} is defined by (24).

In what follows, by virtue of Lemma 4.4 **(c)**, we can show that $\mu_{ab} = 1$ and that some error bound modulus expressed in an explicit way can be provided. First, by some direct calculation, we have $\pi_b(x) = (0, 0)^T$ for all $x \in \mathbb{R}^2$ and

$$\pi_a(x) - \pi_b(x) = \begin{cases} \left(\dfrac{a-1}{a}x_1,\ 0\right)^T & \text{if } x_1 \le 0, x_2 \ge 0, \\ \left(\dfrac{a-1}{a}x_1,\ \dfrac{a-1}{a}x_2\right)^T & \text{if } x_1 \le 0, x_2 \le 0, \\ \left(0,\ \dfrac{a-1}{a}x_2 - \dfrac{3}{2a}x_1\right)^T & \text{if } 0 \le x_1 \le \frac{2(a-1)}{3}x_2, \\ (0,\ 0)^T & \text{otherwise.} \end{cases}$$

Then it is straightforward to verify that the inequality

$$\langle \nabla F(x)(\pi_a(x) - \pi_b(x)),\ \pi_a(x) - \pi_b(x)\rangle \ge \mu||\pi_a(x) - \pi_b(x)||^2$$

holds for all $x \in \mathbb{R}^2$ with $x_1x_2 \neq 0$ if and only if $0 < \mu \leq 1$. That is, Lemma 4.4 **(c)** holds with $V = \mathbb{R}^2$ if and only if $0 < \mu \leq 1$. As Lemma 4.4 **(c)** is implied by Lemma 4.4 **(b)**, we deduce that Lemma 4.4 **(b)** cannot hold with $V = \mathbb{R}^2$ and $\mu > 1$, which implies that μ_{ab} cannot be greater than 1 (cf. Remark 4.1). Therefore, we confirm that $\mu_{ab} = 1$. Furthermore, we can apply Theorem 4.1 to get the following: (i) f_{ab} is a KL function with an exponent of $\frac{1}{2}$; (ii) $\sqrt{f_{ab}}$ has a local error bound on $\mathbb{R}^2$, i.e., for any given $\varepsilon > 0$,

$$\sqrt{\frac{b-a}{2}}\frac{1}{1+b+L}\, d\left(x, [f_{ab} \leq 0]\right) \leq \sqrt{f_{ab}(x)} \quad \forall x \in [f_{ab} \leq \varepsilon],$$

where L is any number such that $L \geq \sup_{x \in [0 < f_{ab} < \varepsilon]} \operatorname{lip} F(x)$.

5 A derivative-free descent method for (VIP)

In this section, we investigate the convergence behavior of the following descent algorithm with an Armijo line search. This algorithm is essentially identical to those studied in [17, 21, 34, 35, 45, 47], particularly in terms of the manner in which descent directions are chosen.

Algorithm

Step 1. Set $0 < a < b$ and $0 < \rho < 1$. Choose three positive constants α, β, τ such that β and τ are small and that α is close to $b - a$. Select a start point $x_0 \in \mathbb{R}^n$, and set $n = 0$.

Step 2. If $f_{ab}(x_n) = 0$, stop. Otherwise, go to Step 3.

Step 3. Let $u_n = \pi_a(x_n) - x_n$ and $w_n = \pi_a(x_n) - \pi_b(x_n)$. If $\beta||u_n|| < ||w_n||$, set $d_n = w_n$ and select m_n as the smallest nonnegative integer m such that

$$f_{ab}(x_n + \rho^m d_n) - f_{ab}(x_n) \leq -\tau\rho^m ||d_n||^2. \tag{25}$$

Otherwise, set $d_n = u_n$ and select m_n as the smallest nonnegative integer m such that

$$f_{ab}(x_n + \rho^m d_n) - f_{ab}(x_n) \leq -\left(b - a - \alpha\right)\rho^m ||d_n||^2. \tag{26}$$

Step 4. Set $t_n = \rho^{m_n}$, $x_{n+1} = x_n + t_n d_n$ and $n = n + 1$, and go to Step 2.

In the following, we will make the following assumptions.

Assumption (i) F is globally Lipschitz continuous with a constant $L > 0$ (implying that π_a and π_b are globally Lipschitz continuous).

Assumption (ii) There exists some $\mu^* > 0$ such that the inequality

$$\langle \nabla F(x)(\pi_a(x) - \pi_b(x)),\ \pi_a(x) - \pi_b(x)\rangle \geq \mu^* ||\pi_a(x) - \pi_b(x)||^2$$

holds for all $x \in \mathbb{R}^n$ where F is differentiable. This implies by Theorem 4.1 that f is a KL function with an exponent of $\frac{1}{2}$, and by Remark 4.2 and (7) that

$$\min_{z \in DF(x)(\pi_a(x) - \pi_b(x))} \langle z, \pi_a(x) - \pi_b(x)\rangle \geq \mu^* ||\pi_a(x) - \pi_b(x)||^2 \quad \forall x \in \mathbb{R}^n.$$

Assumption (iii) The parameters α, β, τ in the Algorithm are chosen such that

$$0 < \beta < \frac{b-a}{b+L}, \quad (b+L)\beta < \alpha < b - a, \quad 0 < \tau < \mu^*.$$

Firstly, we present two technical lemmas that are beneficial for our subsequent analysis.

Lemma 5.1 *Under* **Assumption (i)***, we have the following for all $x \in \mathbb{R}^n$ and $v \in \partial f_{ab}(x)$:*

$$\begin{aligned}\|v\| &\le (b + \operatorname{lip} F(x))\|\pi_b(x) - \pi_a(x)\| + (b-a)\|x - \pi_a(x)\| \\ &\le (b + L)\|\pi_b(x) - \pi_a(x)\| + (b-a)\|x - \pi_a(x)\|.\end{aligned}$$

Proof In view of Lemma 2.2 **(d)** and the formula for $\partial f_{ab}(x)$ presented in Proposition 3.2, we get the first inequality. The second inequality follows directly from **Assumption (i)**. □

Lemma 5.2 *Consider a locally Lipschitz continuous function $g : \mathbb{R}^n \to \mathbb{R}$. For some $x \in \mathbb{R}^n$ and $w \in \mathbb{R}^n \setminus \{0\}$, assume that there are some $\sigma > 0$ and $0 < t_0 < t_1$ such that*

$$g(x + t_0 w) - g(x) \le -\sigma t_0 ||w||^2 \quad \textit{and} \quad g(x + t_1 w) - g(x) > -\sigma t_1 ||w||^2.$$

Then there exist some $\theta^* \in (0,1)$ and $v^* \in \partial g(x + \theta^* t_1 w)$ such that

$$g(x + t_1 w) - g(x) = t_1 \langle v^*, w \rangle.$$

Proof Define $\varphi : \mathbb{R} \to \mathbb{R}$ by $\varphi(\theta) := g(x + \theta t_1 w) - g(x) + \theta[g(x) - g(x + t_1 w)]$. Evidently, φ is locally Lipschitz continuous, and $\varphi(0) = \varphi(1) = 0$. Moreover, it follows from the assumption that

$$\varphi(t_0/t_1) = g(x + t_0 w) - g(x) + (t_0/t_1)[g(x) - g(x + t_1 w)] < 0.$$

This implies the existence of at least one $\theta^* \in (0,1)$ such that φ attains its minimum over $[0,1]$ at θ^*. By the Fermat's rule, this further implies that $0 \in \partial\varphi(\theta^*)$. In light of the local Lipschitz continuity of g, we get from the calculus rules [37, Exercise 8.8 and Theorem 10.6] that

$$\partial\varphi(\theta^*) \subset g(x) - g(x + t_1 w) + \left\{ t_1 \langle v, w \rangle \mid v \in \partial g(x + \theta^* t_1 w) \right\}.$$

This completes the proof. □

Proposition 5.1 *Under* **Assumptions (i)-(iii)***, Step 3 of the Algorithm is well defined.*

Proof To prove that Step 3 in the Algorithm is well defined, it suffices to show that if $\beta||u_n|| < ||w_n||$, $-d(-f_{ab})(x_n)(w_n) < -\tau\|w_n\|^2$, and if $\beta||u_n|| \ge ||w_n||$, $-d(-f_{ab})(x_n)(u_n) < -(b - a - \alpha)\|u_n\|^2$. Based on the proof of the formula for $df_{ab}(\bar{x})(w)$ in Proposition 3.2, we get the formula for the subderivative of $-f_{ab}$ at a point $\bar{x} \in \mathbb{R}^n$ as follows:

$$-d(-f_{ab})(\bar{x})(w) = (b-a)\langle \bar{x} - \pi_a(\bar{x}),\ w\rangle - \min\langle (DF(\bar{x}) - bI)\, w,\ -\pi_b(\bar{x}) + \pi_a(\bar{x})\rangle.$$

In the case where $\beta||u_n|| < ||w_n||$, we have

$$\begin{aligned}&-d(-f_{ab})(x_n)(w_n) \\ &= \langle b(x_n - \pi_b(x_n)) - a(x_n - \pi_a(x_n)), w_n\rangle - \min\nolimits_{z \in DF(x_n)(w_n)} \langle z, w_n\rangle \\ &\le -\min\nolimits_{z \in DF(x_n)(w_n)} \langle z, w_n\rangle \\ &\le -\mu^* ||w_n||^2 \\ &< -\tau ||w_n||^2,\end{aligned}$$

where the first inequality follows from Lemma 3.1 **(d)**, the second inequality follows from **Assumption (ii)**, and the third inequality follows from **Assumption (iii)**. In the case where

$\beta||u_n|| \geq ||w_n||$, we have

$$\begin{aligned}
&-d(-f_{ab})(x_n)(u_n)\\
&= \langle b(x_n - \pi_b(x_n)) - a(x_n - \pi_a(x_n)), u_n\rangle - \min_{z\in DF(x_n)(u_n)}\langle z, w_n\rangle\\
&= -(b-a)||u_n||^2 + b\langle \pi_a(x_n) - \pi_b(x_n), u_n\rangle + \max_{z\in DF(x_n)(u_n)}\langle z, -w_n\rangle\\
&\leq -[(b-a) - b\beta]||u_n||^2 + \max_{z\in DF(x_n)(u_n)}\langle z, -w_n\rangle\\
&\leq -[(b-a) - b\beta]||u_n||^2 + L||u_n||\cdot||w_n||\\
&\leq -[(b-a) - (b+L)\beta]||u_n||^2\\
&< -[(b-a) - \alpha]||u_n||^2,
\end{aligned}$$

where the first inequality follows by using the Cauchy-Schwarz inequality and the inequality $\beta||u_n|| \geq ||w_n||$, the second inequality follows from Lemma 2.2 **(c)** and **Assumption (i)**, the third inequality follows from the inequality $\beta||u_n|| \geq ||w_n||$, and the last inequality follows from **Assumption (iii)**. This completes the proof. □

Proposition 5.2 *Assume that the sequence $\{x_n\}$ generated by the Algorithm satisfies $f_{ab}(x_n) > 0$ for all n. Under* **Assumptions (i)-(iii)***, there is some $t^* > 0$ such that $t_n \geq t^*$ for all n, i.e., the step size sequence $\{t_n\}$ generated by the Algorithm has a positive lower bound.*

Proof Recall that in Step 3 of the Algorithm, we set $u_n := \pi_a(x_n) - x_n$, $w_n := \pi_a(x_n) - \pi_b(x_n)$, and $d_n := u_n$ if $\beta\|u_n\| \geq \|w_n\|$, and $d_n := w_n$ if $\beta\|u_n\| < \|w_n\|$. Considering the setting for d_n and our assumption that $f_{ab}(x_n) > 0$ for all n, we can infer from Lemma 3.1 **(c)** that $d_n \neq 0$ for all n.

Suppose, for the sake of contradiction, that the step length sequence $\{t_n\}$ does not have a positive lower bound. That is, by taking a subsequence if necessary we assume that $t_n \to 0+$ as $n \to +\infty$. Due to $t_n = \rho^{m_n}$, we have $m_n \to +\infty$ as $n \to +\infty$. Without loss of generality, we may assume that $m_n \geq 1$ for all n. In light of the line search strategy in Step 3 of the Algorithm, we apply Lemma 5.2 to get

$$f_{ab}(x_n + \rho^{m_n-1}d_n) - f_{ab}(x_n) = \rho^{m_n-1}\langle v_n, d_n\rangle \quad \forall n, \tag{27}$$

where $v_n \in \partial f_{ab}(y_n)$ with $y_n := x_n + \theta_n^*\rho^{m_n-1}d_n$ and $\theta_n^* \in (0, 1)$. By the formula for $\partial f_{ab}(y_n)$ in Proposition 3.2, there exists some $z_n^* \in D^*F(\pi_b(y_n) - \pi_a(y_n))$ such that

$$v_n = z_n^* + b(y_n - \pi_b(y_n)) - a(y_n - \pi_a(y_n)). \tag{28}$$

By Lemma 2.2 **(d)** and **Assumption (i)**, we have

$$||z_n^*|| \leq L||\pi_b(y_n) - \pi_a(y_n)||. \tag{29}$$

First, we consider the case where $\beta\|u_n\| \geq \|w_n\|$ in Step 3. In this case, we have $d_n = u_n = \pi_a(x_n) - x_n$ and $y_n := x_n + \theta_n^*\rho^{m_n-1}u_n$. Thanks to the line search strategy proposed in the Algorithm, we have $f_{ab}(x_n + \rho^{m_n-1}u_n) - f_{ab}(x_n) > -(b-a-\alpha)\rho^{m_n-1}\|u_n\|^2$. This, together with (27), (28) and (29), implies that

$$\begin{aligned}
-(b-a-\alpha)||u_n||^2 &< \langle v_n, u_n\rangle\\
&= \langle z_n^*, u_n\rangle + \langle b(y_n - \pi_b(y_n)) - a(y_n - \pi_a(y_n)), u_n\rangle\\
&= \langle z_n^*, u_n\rangle + b\langle \pi_a(y_n) - \pi_b(y_n), u_n\rangle + (b-a)\langle y_n - \pi_a(y_n), u_n\rangle\\
&\leq (L+b)||\pi_b(y_n) - \pi_a(y_n)||\cdot||u_n||\\
&\quad -(b-a)\langle \pi_a(y_n) - y_n, u_n\rangle.
\end{aligned} \tag{30}$$

Moreover, by **Assumption (i)**, we have

$$\begin{aligned}||\pi_a(y_n) - \pi_b(y_n)|| &\le ||w_n|| + ||\pi_a(y_n) - \pi_b(y_n) - w_n|| \\ &\le \|w_n\| + ||\pi_a(y_n) - \pi_a(x_n)\| + \|\pi_b(y_n) - \pi_b(x_n)|| \\ &\le \beta\|u_n\| + (1 + \frac{L}{a})\|y_n - x_n\| + (1 + \frac{L}{b})\|y_n - x_n\| \\ &= [\beta + (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n-1}]||u_n||, \end{aligned} \tag{31}$$

and

$$\begin{aligned}||\pi_a(y_n) - y_n - u_n|| &= ||\pi_a(y_n) - y_n - \pi_a(x_n) + x_n|| \\ &\le \|\pi_a(y_n) - \pi_a(x_n)\| + \|y_n - x_n\| \\ &\le (2 + \tfrac{L}{a})||y_n - x_n|| = (2 + \tfrac{L}{a})\theta_n^* \rho^{m_n-1}||u_n||.\end{aligned}$$

The latter condition entails that

$$\langle \pi_a(y_n) - y_n, u_n\rangle = ||u_n||^2 + (2 + \frac{L}{a})\theta_n^* \rho^{m_n-1}||u_n||^2 \langle c_n, \frac{u_n}{\|u_n\|}\rangle, \tag{32}$$

where $c_n := \frac{\pi_a(y_n)-y_n-u_n}{(2+\frac{L}{a})\theta_n^* \rho^{m_n-1}||u_n||}$ having the property that $\|c_n\| \le 1$. Combining (30-32), we have

$$\begin{aligned}-(b - a - \alpha) < &(L + b)[\beta + (2 + \tfrac{L}{a} + \tfrac{L}{b})\theta_n^* \rho^{m_n-1}] \\ &-(b - a)[1 + (2 + \tfrac{L}{a})\theta_n^* \rho^{m_n-1}\langle c_n, \tfrac{u_n}{||u_n||}\rangle].\end{aligned} \tag{33}$$

Next, we consider the case where $\beta\|u_n\| < \|w_n\|$ in Step 3. In this case, we have $d_n = w_n = \pi_a(x_n) - \pi_b(x_n)$ and $y_n := x_n + \theta_n^* \rho^{m_n-1} w_n$. Due to the line search strategy proposed in the Algorithm, we have $f_{ab}(x_n + \rho^{m_n-1} w_n) - f_{ab}(x_n) > -\tau\rho^{m_n-1}\|w_n\|^2$, which, together with (27), (28) and (29), implies that

$$\begin{aligned}&-\tau\|w_n\|^2 \\ &< \langle v_n,\ w_n\rangle \\ &= \langle z_n^* + b(y_n - \pi_b(y_n)) - a(y_n - \pi_a(y_n)), w_n\rangle \\ &\le \langle z_n^*,\ \pi_a(y_n) - \pi_b(y_n)\rangle + \langle z_n^*, w_n - (\pi_a(y_n) - \pi_b(y_n))\rangle \\ &\quad +\langle b(y_n - \pi_b(y_n)) - a(y_n - \pi_a(y_n)), w_n - (\pi_a(y_n) - \pi_b(y_n))\rangle \\ &\le -\mu^*||\pi_a(y_n) - \pi_b(y_n)||^2 + \langle z_n^*, w_n - (\pi_a(y_n) - \pi_b(y_n))\rangle \\ &\quad +\langle b(y_n - \pi_b(y_n)) - a(y_n - \pi_a(y_n)), w_n - (\pi_a(y_n) - \pi_b(y_n))\rangle \\ &\le -\mu^*||\pi_a(y_n) - \pi_b(y_n)||^2 + L||\pi_a(y_n) - \pi_b(y_n)|| \cdot ||w_n - (\pi_a(y_n) - \pi_b(y_n))|| \\ &\quad +[(b - a)||\pi_a(y_n) - y_n|| + b||\pi_a(y_n) - \pi_b(y_n)||]||w_n - (\pi_a(y_n) - \pi_b(y_n))||,\end{aligned} \tag{34}$$

where the second inequality follows from Lemma 3.1 **(d)**, the third one follows from **Assumption (ii)**, the last one follows from Cauchy-Schwarz inequality. Moreover, by **Assumption (i)**, we have

$$||\pi_a(y_n) - \pi_b(y_n) - w_n|| \le (2 + \frac{L}{a} + \frac{L}{b})||y_n - x_n|| = (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n-1}||w_n||, \tag{35}$$

$$||\pi_a(y_n) - \pi_b(y_n)|| \le [1 + (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n-1}]||w_n||, \tag{36}$$

$$
\begin{aligned}
||\pi_a(y_n) - y_n|| &\le \|u_n\| + ||\pi_a(y_n) - y_n - u_n|| \\
&\le ||u_n|| + (2 + \frac{L}{a})\theta_n^* \rho^{m_n - 1} ||w_n|| \\
&\le [\frac{1}{\beta} + (2 + \frac{L}{a})\theta_n^* \rho^{m_n - 1}]||w_n|| \qquad (37)
\end{aligned}
$$

and then there exists b_n with $||b_n|| \le 1$ such that

$$
\pi_a(y_n) - \pi_b(y_n) = w_n + (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1} ||w_n|| b_n. \qquad (38)
$$

Combining (34-38), we obtain

$$
\begin{aligned}
-\tau < &-\mu^*[1 + 2\langle \frac{w_n}{||w_n||}, (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1} b_n\rangle + (2 + \frac{L}{a} + \frac{L}{b})^2 (\theta_n^* \rho^{m_n - 1})^2 ||b_n||^2] \\
&+L[1 + (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1}](2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1} \\
&+(b - a)[\frac{1}{\beta} + (2 + \frac{L}{a})\theta_n^* \rho^{m_n - 1}](2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1} \\
&+b[1 + (2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1}](2 + \frac{L}{a} + \frac{L}{b})\theta_n^* \rho^{m_n - 1}. \qquad (39)
\end{aligned}
$$

Our assumption that $f_{ab}(x_n) > 0$ for all n implies that there are infinitely many positive integers n such that either $\beta\|u_n\| \ge \|w_n\|$ or $\beta\|u_n\| < \|w_n\|$. This, in turn, means that there are infinitely many positive integers n for which either the inequality (33) or (39) holds. Given that $\rho^{m_n - 1} \to 0+$, we have consequently either $-(b - a - \alpha) \le (L + b)\beta - (b - a)$ or $-\tau \le -\mu^*$, both contradicting to **Assumption (iii)**. This contradiction demonstrates that the step length sequence $\{t_n\}$ generated by the Algorithm has a positive lower bound. This completes the proof. □

Proposition 5.3 *Assume that the sequence $\{x_n\}$ generated by the Algorithm satisfies $f_{ab}(x_n) > 0$ for all n. Under* **Assumptions (i)-(iii)**, *the following inequalities hold for all n:*

$$
f_{ab}(x_{n+1}) - f_{ab}(x_n) \le -M_1 ||x_{n+1} - x_n||^2 \qquad (40)
$$

and

$$
d(0, \partial f_{ab}(x_n)) \le \frac{M_2}{t^*} ||x_{n+1} - x_n||, \qquad (41)
$$

where $M_1 := \min\{b - a - \alpha, \tau\}$, $M_2 := L + b + \frac{b-a}{\beta}$ and t^* is a positive lower bound of $\{t_n\}$.

Proof By Steps 3 and 4 of the Algorithm, we have $0 < t_n \le 1$, $x_{n+1} = x_n + t_n d_n$ and $f_{ab}(x_{n+1}) - f_{ab}(x_n) \le -M_1 t_n \|d_n\|^2$ for all n. From this, we can immediately derive (40) . By Lemma 5.1, we have

$$
d(0, \partial f_{ab}(x_n)) \le (L + b)||w_n|| + (b - a)||u_n||,
$$

where L is given as in **Assumption (i)**, and $w_n = \pi_a(x_n) - \pi_b(x_n)$ and $u_n = \pi_a(x_n) - x_n$ are defined as in Step 3. If $\beta||u_n|| < ||w_n||$, from Steps 3 and 4 of the Algorithm, we get $\|x_{n+1} - x_n\| = t_n \|w_n\|$ and hence that

$$
(L + b)||w_n|| + (b - a)||u_n|| < (L + b + \frac{b - a}{\beta})||w_n|| = \frac{M_2}{t_n} ||x_{n+1} - x_n||.
$$

Alternatively, if $\beta||u_n|| \geq ||w_n||$, from Steps 3 and 4 of the Algorithm, we have $\|x_{n+1}-x_n\| = t_n\|u_n\|$ and hence that

$$(L+b)||w_n|| + (b-a)||u_n|| \leq \beta(L+b+\frac{b-a}{\beta})||u_n|| \leq \frac{M_2}{t_n}||x_{n+1}-x_n||,$$

where the second inequality follows from the fact that $0 < \beta < \frac{b-a}{b+L} < 1$ according to **Assumption (iii)**. In both cases, we get (41) by noting that the existence of a positive lower bound t^* of $\{t_n\}$ is guaranteed by Proposition 5.2. This completes the proof. □

Remark 5.1 In light of (40) and (41) in Proposition 5.3, we confirm that the sequence $\{x_n\}$ satisfies **Assumption (H1)** and a variant of **Assumption (H2)** in [5]. Notably, **Assumption (H2)** in [5] requires an upper estimate of the form (41) for $\partial f(x_{n+1})$ (instead of $\partial f(x_n)$). However, it remains unclear whether the sequence $\{x_n\}$ possesses a convergent subsequence (i.e., whether **Assumption (H3)** in [5] holds). Consequently, [5, Theorem 2.9] cannot be directly applied unless we assume either that $\{x_n\}$ has a convergent subsequence or that the level set $[f_{ab} \leq f_{ab}(x_0)]$ is bounded. Nevertheless, the following theorem and its corollary demonstrate that **Assumptions (i)-(iii)** collectively ensure the convergence of $\{x_n\}$ and the nonemptiness of the solution set of (VIP).

Theorem 5.1 *Assume that the sequence $\{x_n\}$ generated by the Algorithm satisfies $f_{ab}(x_n) > 0$ for all n. Under* **Assumptions (i)-(iii)***, the following assertions hold:*

(a) *The sequence $f_{ab}(x_n)$ converges* Q*-linearly to 0.*
(b) *The sequence x_n has a finite length, i.e., $\sum_{n=0}^{+\infty}||x_{n+1}-x_n|| < +\infty$.*
(c) *The sequence x_n converges* R*-linearly to a solution $\bar{x}$ of (VIP).*

Proof According to Proposition 5.3, (40) and (41) hold with $M_1 := \min\{\tau, b-a-\alpha\}$, $M_2 := L+b+\frac{b-a}{\beta}$ and t^* being some positive lower bound of $\{t_n\}$.

We first show **(a)**. By the line search strategy in Step 3 of the Algorithm, the following hold for all n:

$$\|d_n\| \geq \beta\|x_n - \pi_a(x_n)\|, \tag{42}$$

and

$$\begin{aligned} f_{ab}(x_{n+1}) - f_{ab}(x_n) &\leq -\min\{\tau, b-a-\alpha\}t_n\|d_n\|^2 \\ &\leq -\min\{\tau, b-a-\alpha\}t^*\|d_n\|^2 \\ &< 0. \end{aligned} \tag{43}$$

In light of (42), we get from Lemma 3.1 **(a)** that $\|d_n\|^2 \geq \frac{2\beta^2}{b-a}f_{ab}(x_n)$. This, together with (43) and the definition of M_1, implies that

$$f_{ab}(x_{n+1}) \leq -M_1t^*||d_n||^2 + f_{ab}(x_n) \leq (1-\frac{2\beta^2M_1t^*}{b-a})f_{ab}(x_n),$$

and hence that,

$$\frac{f_{ab}(x_{n+1})}{f_{ab}(x_n)} \leq 1 - \frac{2\beta^2M_1t^*}{b-a} =: \eta. \tag{44}$$

Clearly, we have $0 < \eta < 1$. Then, considering the fact that $f_{ab} \geq 0$, and by definition [29, pp.619-620], the sequence $f_{ab}(x_n)$ converges Q-linearly to 0.

To prove **(b)**, we first observe that the following inequality holds for all n:

$$f_{ab}(x_n) + f_{ab}(x_{n+1}) \geq 2\sqrt{f_{ab}(x_n)f_{ab}(x_{n+1})},$$

or equivalently,

$$\sqrt{f_{ab}(x_n)} - \sqrt{f_{ab}(x_{n+1})} \geq \frac{1}{2\sqrt{f_{ab}(x_n)}}(f_{ab}(x_n) - f_{ab}(x_{n+1})). \tag{45}$$

In view of Assumptions **(i)** and **(ii)**, we get from Lemma 4.2 that for all n,

$$d(0, \partial f_{ab}(x_n)) \geq \mu^* \sqrt{f_{ab}(x_n)}. \tag{46}$$

Combing (40-41) and (45-46), we have

$$\sqrt{f_{ab}(x_n)} - \sqrt{f_{ab}(x_{n+1})} \geq \frac{M_1 \mu^* t^*}{2M_2} \|x_{n+1} - x_n\| \quad \forall n,$$

and hence

$$S_m := \sum_{n=0}^{m} \|x_{n+1} - x_n\| \leq \frac{2M_2}{M_1 \mu^* t^*}(\sqrt{f_{ab}(x_0)} - \sqrt{f_{ab}(x_{m+1})}) \leq \frac{2M_2 \sqrt{f_{ab}(x_0)}}{M_1 \mu^* t^*} \quad \forall m.$$

Since the positive sequence $\{S_m\}$ is increasing and bounded above, it must be convergent. This establishes**(b)**.

To prove **(c)**, we first demonstrate that $\{x_n\}$ is a convergence sequence. Let $\varepsilon > 0$ be arbitrary. Since $\{S_m\}$ is a convergent sequence, there exists some positive integer N such that for any $p, q > N$, we have $|S_p - S_q| \leq \varepsilon$. Given that $\|x_{p+1} - x_{q+1}\| \leq |S_p - S_q|$, the sequence $\{x_n\}$ is clearly a Cauchy sequence, and thus converges to some $\bar{x} \in \mathbb{R}^n$. Since f_{ab} is continuous and $f_{ab}(x_n)$ converges to 0, we have $f_{ab}(\bar{x}) = 0$, i.e., $\bar{x}$ is a solution of (VIP). By the triangle inequality, the following inequality holds for all positive integers n and m with $m > n$:

$$\|x_n - \bar{x}\| \leq \sum_{k=n}^{m} \|x_{k+1} - x_k\| + \|x_{m+1} - \bar{x}\|.$$

In view of **(b)** and the fact that $\|x_{m+1} - \bar{x}\| \to 0$ as $m \to \infty$, we have

$$\sum_{k=n}^{m} \|x_{k+1} - x_k\| + \|x_{m+1} - \bar{x}\| \to \sum_{k=n}^{\infty} \|x_{k+1} - x_k\| \text{ as } m \to \infty,$$

and hence

$$\|x_n - \bar{x}\| \leq \sum_{k=n}^{\infty} \|x_{k+1} - x_k\|.$$

Using (40) and (44), we further derive

$$\|x_n - \bar{x}\| \leq \sum_{k=n}^{\infty} \sqrt{\frac{f_{ab}(x_k)}{M_1}} \leq \sqrt{\frac{f_{ab}(x_n)}{M_1}} \sum_{k=0}^{\infty} \sqrt{\eta^k} = \sqrt{\frac{f_{ab}(x_n)}{M_1}} \frac{1}{1 - \sqrt{\eta}} =: \zeta_n,$$

and

$$\frac{\zeta_{n+1}}{\zeta_n} = \sqrt{\frac{f_{ab}(x_{n+1})}{f_{ab}(x_n)}} \leq \sqrt{\eta}.$$

Since $0 < \eta < 1$, we have $0 < \sqrt{\eta} < 1$. Then by definition [29, pp.619-620], ζ_n converges Q-linearly to 0, and x_n converges R-linearly to $\bar{x}$. This completes the proof. □

By noting that **Assumption (iii)** can be ensured by appropriately setting the parameters α, β, and τ under **Assumptions (i)** and **(ii)**, the above convergence theorem implies that **Assumptions (i)** and **(ii)** together guarantee the nonemptiness of the solution set of (VIP).

Corollary 5.1 *Consider the (VIP) with F being globally Lipschitz continuous. If there exists some $\mu > 0$ such that the inequality*

$$\langle \nabla F(x)(\pi_a(x) - \pi_b(x)),\ \pi_a(x) - \pi_b(x)\rangle \geq \mu||\pi_a(x) - \pi_b(x)||^2$$

holds for all $x \in \mathbb{R}^n$ where F is differentiable, then the (VIP) has a solution.

6 Concluding remarks

In this paper, we studied the D-gap function for a nonsmooth and nonmonotone variational inequality problem. By deriving exact formulas for the subderivative and the regular (limiting) subdifferential set of the D-gap function, we established necessary and sufficient conditions for the Kurdyka-Łojasiewicz (KL) inequality property and the error bound property of the D-gap function. Moreover, we demonstrated how these results can be applied to analyze the linear convergence of certain derivative-free descent algorithms with an inexact line search.

For future research, we plan to extend our techniques and subdifferential formulas to other gap functions (see [12]) that rely less on projections. Furthermore, our objective is to develop more efficient algorithms that involve minimal projections when determining step sizes by (25) and (26).

Acknowledgements We sincerely thank the two anonymous referees for their constructive comments and suggestions, which have significantly improved the quality of this manuscript. Li's work was supported in part by the National Natural Science Foundation of China (No.12271072), by the Science and Technology Research Program of Chongqing Municipal Education Commission (No.KJZD-M202201303) and by Natural Science Foundation of Chongqing Municipal Science and Technology Commission (No.CSTB2022NSCQ-MSX0406). Meng's work was supported in part by the National Natural Science Foundation of China (No. 11671329). Yang's work was supported in part by Research Grants Council of Hong Kong (15205223 and N_PolyU507/22).

Funding Open access funding provided by The Hong Kong Polytechnic University

Data Availability Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. Auslender, A.: Differential stability in nonconvex and nondifferentiable programming. Math. Program. Study **10**, 29–41 (1979)
2. Auchmuty, G.: Variational principles for variational inequalities. Numer. Funct. Anal. Optim. **10**, 863–874 (1989)

3. Auslender, A.: Optimization, Méthodes numériques, Masson, Paris, France (1976)
4. Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality. Math. Oper. Res. **35**, 438–457 (2010)
5. Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Program. **137**, 91–129 (2013)
6. Azé, D., Corvellec, J.-N.: Characterizations of error bounds for lower semicontinuous functions on metric spaces. ESAIM: Control, Optim. Calc. Var. **10**, 409–425 (2004)
7. Bolte, J., Daniilidis, A., Lewis, A.: The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. **17**, 1205–1223 (2007)
8. Bolte, J., Daniilidis, A., Ley, O., Mazet, L.: Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. Trans. Amer. Math. Soc. **362**, 3319–3363 (2010)
9. Chen, G.Y., Goh, C.J., Yang, X.Q.: On gap functions and duality of variational inequality problems. J. Math. Anal. Appl. **214**, 658–673 (1997)
10. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
11. Dinh, B.V., Pham, T.S.: Error bounds of regularized gap functions for polynomial variational inequalities. J. Optim. Theory Appl. **192**, 226–247 (2022)
12. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I–II. Springer, Berlin Helidelberg, New York (2003)
13. Fukushima, M.: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Math. Program. **53**, 99–110 (1992)
14. De Giorgi, E., Marino, A., Tosques, M.: Problemi di evoluzione in spazi metrici e curve di massima pendenza (evolution problems in metric spaces and curves of maximal slope). Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **68**, 180–187 (1980)
15. 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)
16. 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)
17. Kanzow, C., Fukushima, M.: Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities. Math. Program. **83**, 55–87 (1998)
18. Konnov, I.: Descent methods for nonsmooth variational inequalities. Comput. Math. Math. Phys. **46**, 1186–1192 (2006)
19. Hogan, W.W.: Directional derivatives for extremal-value functions with applications to the completely convex case. Oper. Res. **21**, 188–209 (1973)
20. Larsson, T., Patriksson, M.: A class of gap functions for variational inequalities. Math. Program. **64**, 53–79 (1994)
21. Li, G.Y., Ng, K.F.: Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems. SIAM J. Optim. **20**, 667–690 (2009)
22. Li, G.Y., Tang, C., Wei, Z.: Error bound results for generalized D-gap functions of nonsmooth variational inequality problems. J. Comput. Appl. Math. **233**, 2795–2806 (2010)
23. Li, G.Y., Pong, T.K.: Calculus of the exponent of Kurdyka-Łojasiewicz inequality and its applications to linear convergence of first-order methods. Found. Comput. Math. **18**, 1199–1232 (2018)
24. Meng, K.W., Yang, X.Q.: Equivalent conditions for local error bounds. Set-Valued Var. Anal. **20**, 617–636 (2012)
25. Minchenko, L.I., Sakolchik, P.P.: Hölder behavior of optimal solutions and directional differentiability of marginal functions in nonlinear programming. J. Optim. Theory Appl. **90**, 555–580 (1996)
26. Mordukhovich, B.S., Nam, N.M., Yen, N.D.: Subgradients of marginal functions in parametric mathematical programming. Math. Program. **116**, 369–396 (2009)
27. Ng, K.F., Tan, L.L.: Error bounds of regularized gap functions for nonsmooth variational inequality problems. Math. Program. **110**, 405–429 (2007)
28. Ng, K.F., Tan, L.L.: D-gap functions for nonsmooth variational inequality problems. J. Optim. Theory Appl. **133**, 77–97 (2007)
29. Nocedal, J., Wright, S.: Numerical Optimization, 2nd edn. Springer-Verlag, Berlin, New York (2006)
30. Pang, J.-S.: A posteriori error bounds for the linearly-constrained variational inequality problem. Math. Oper. Res. **12**, 474–484 (1987)
31. Panicucci, B., Pappalardo, M., Passacantando, M.: A globally convergent descent method for nonsmooth variational inequalities. Comput. Optim. Appl. **43**, 197–211 (2009)
32. Pappalardo, M., Mastroeni, G., Passacantando, M.: Merit functions: a bridge between optimization and equilibria. 4OR **12**, 1–33 (2014)

33. Peng, J.M.: Equivalence of variational inequality problems to unconstrained optimization. Math. Program. **78**, 347–355 (1997)
34. 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)
35. Qu, B., Wang, C.Y., Zhang, J.Z.: Convergence and error bound of a method for solving variational inequality problems via the generalized D-gap function. J. Optim. Theory Appl. **119**, 535–552 (2003)
36. Rockafellar, R.T.: Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming. Math. Program. Study **17**, 28–66 (1982)
37. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
38. Solodov, M.V., Tseng, P.: Some methods based on the D-gap function for solving monotone variational inequalities. Comput. Optim. Appl. **17**, 255–277 (2000)
39. Solodov, M.V.: Merit functions and error bounds for generalized variational inequalities. J. Math. Anal. Appl. **287**, 405–414 (2003)
40. Tan, L.L.: Regularized gap functions for nonsmooth variational inequality problems. J. Math. Anal. Appl. **334**, 1022–1038 (2007)
41. Thibault, L.: Tangent cones and quasi-interiorly tangent cones to multifunctions. Trans. Amer. Math. Soc. **277**, 601–621 (1983)
42. Wu, J.H., Florian, M., Marcotte, P.: A general descent framework for the monotone variational inequality problem. Math. Program. **61**, 281–300 (1993)
43. Wu, Z.L., Wu, S.Y.: Gâteaux differentiability of the dual gap function of a variational inequality. Eur. J. Oper. Res. **190**, 328–344 (2008)
44. Wu, Z.L.: Characterizations of weakly sharp solutions for a variational inequality with a pseudomonotone mapping. Eur. J. Oper. Res. **265**, 448–453 (2018)
45. Xu, H.F.: Regularized gap functions and D-gap functions for nonsmooth variational inequalities, Optimization and Related Topics, A. Rubinov and B. Glover eds., Kluwer Academic Publishers (2001)
46. Yamashita, N., Fukushima, M.: Equivalent unconstrained minimization and global error bounds for variational inequality problems. SIAM J. Control. Optim. **35**, 273–284 (1997)
47. Yamashita, N., Taji, K., Fukushima, M.: Unconstrained optimization reformulations of variational inequality problems. J. Optim. Theory Appl. **92**, 439–456 (1997)
48. Zhu, D.L., Marcotte, P.: Modified descent methods for solving the monotone variational inequality problem. Oper. Res. Lett. **14**, 111–120 (1993)
49. Zhu, D.L., Marcotte, P.: An extended descent framework for variational inequalities. J. Optim. Theory Appl. **80**, 349–366 (1994)