Math. Program., Ser. A (2007) 110:405–429
DOI 10.1007/s10107-006-0007-2

FULL LENGTH PAPER

Error bounds of regularized gap functions for nonsmooth variational inequality problems

Kung Fu Ng · Lu Lin Tan

Received: 19 April 2005 / Accepted: 30 March 2006 /
Published online: 22 June 2006

Abstract We study the Clarke–Rockafellar directional derivatives of the regularized gap functions (and of some modified ones) for the variational inequality problem (VIP) defined by a locally Lipschitz but not necessarily differentiable function on a closed convex set in an Euclidean space. As applications we show that, under the strong monotonicity assumption, the regularized gap functions have fractional exponent error bounds and consequently that the sequences provided by an algorithm of Armijo type converge to the solution of the (VIP).

Keywords Nonsmooth variational inequality problem · Merit function · Regularized gap function · Error bound

Mathematics Subject Classification (2000) 49J52 · 49J40 · 90C30

1 Introduction

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

The research of this author was supported by an Earmarked Grant from the Research Council of Hong Kong.

K. F. Ng · L. L. Tan (✉)
Department of Mathematics,
The Chinese University of Hong Kong,
Hong Kong, People's Republic of China
e-mail: lltan@math.cuhk.edu.hk

K. F. Ng
e-mail: kfng@math.cuhk.edu.hk

$$\langle F(x^*),\ x - x^*\rangle \geq 0 \quad \forall x \in P. \tag{1}$$

When P is the nonnegative orthant in R^n, VIP(F, P) reduces to the nonlinear complementarity problem NCP(F, P). The variational inequality problems have many applications in different fields (including mathematical programming problems and some equilibrium problems), we refer the reader to the very informative recent boobook [3] by Facchinei and Pang for the background information and motivations of the variational inequality problems covering both smooth and nonsmooth functions (in fact nonsmooth variational problems are quite abundant, see [7, 17, B.F. Hobbs and J.S. Pang, submitted]for recent developments). Many authors have studied these problems (with various degree of generality such as for smooth or nonsmooth F) using various methods pertinent to various issues or aspects of the problem. See [2, 3, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 18, 24, 26, 27]. One approach is by merit functions such as the regularized gap function f_γ [5] defined by

$$f_\gamma(\tau) := -\inf_{x\in P}\left\{\langle F(\tau),\ x-\tau\rangle + \frac{\gamma}{2}\|x-\tau\|^2\right\}, \quad \tau \in P,\ \gamma > 0. \tag{2}$$

By [26, Proposition 3.3 and Theorem 3.1] x^* solves VIP(F, P) if and only if $f_\gamma(x^*) = 0$ and x^* solves the constrained minimization problem

$$\min f_\gamma(\tau) \quad \text{subject to} \quad \tau \in P. \tag{3}$$

In the case when F is continuously differentiable on P, f_γ is also continuously differentiable, and in fact one has [26]

$$\nabla f_\gamma(\tau) = -\nabla F(\tau)\left(\pi_\gamma(\tau) - \tau\right) + F(\tau) + \gamma\left(\pi_\gamma(\tau) - \tau\right), \tag{4}$$

where $\pi_\gamma(\tau)$ denotes the best approximation to $\tau - \frac{F(\tau)}{\gamma}$ from P, namely

$$\pi_\gamma(\tau) = \mathrm{Proj}_P\left(\tau - \frac{F(\tau)}{\gamma}\right). \tag{5}$$

An advantage of the regularized gap function is its differentiability when F is smooth and this helps to develop descent-based algorithms to solve the original VIP. In particular, by virtue of the consideration of differentials, Wu et al. [26], Yamashita et al. [27], and Huang et al. [8] have addressed the error bound issues for f_γ when F is assumed to be strongly monotone, and thereby established convergence results of sequences obtained by an algorithm of Armijo type. The above results are extended here to cover the case that F is not necessarily smooth. We show in particular that the regularized gap function f_γ is locally Lipshitz on P when F is, and hence the corresponding Clarke–Rockafellar directional derivative can play a similar role as that played by the directional derivative for the smooth case with the above mentioned algorithm of Armijo type. The organization of the rest of this paper as follows. In the next section we set our notations (which are standard) and preliminaries. In particular three basic lemmas regarding the Clarke–Rockafellar directional derivatives are given. These results are more or less folklore (cf. [1, 19, 20, 21, 25, 29]). But as, in the literature, we cannot find results presented in the forms required for our need we include their proofs here even though the proofs are quite technical. Although f_γ is not necessarily smooth at certain points, we show in sect. 4 that the Clarke–Rockafellar directional derivatives of f_γ (and of $\overline{f}_\gamma$ which agrees with f_γ on P but takes value $+\infty$ on $R^n\backslash P$) at these points can be explicitly represented in a similar spirit as that given in [1, Theorem 2.8.6]. This is established via a more general result in Sect. 3 regarding max-functions. Applications are given in the last two sections: we show in Sect. 5 that $\sqrt{f_\gamma}$ has an error bound on P and in Sect. 6 we present a convergence result for a descent method of Armijo type leading to the solution of the corresponding VIP. Our treatment here follows the earlier ones given by Yamashita et al. [27] and Huang et al. [8] except that the Armijo-type line search is justified now by the study of the Clarke–Rockafellar directional derivative of the regularized gap function (in place of the directional derivative in the smooth case).

2 Notation and preliminary results

For a proper lower semicontinuous function $h : R^n \to R \cup \{+\infty\}$, let dom $h :=$ $\{x \in R^n : h(x) < +\infty\}$. For any $x \in$ dom h and $v \in R^n$, we denote upper and lower Dini-directional derivatives, respectively, by

$$\overline{d}^+ h(x)(v) := \limsup_{t\downarrow 0} \frac{h(x+tv)-h(x)}{t}$$

and

$$\underline{d}^+ h(x)(v) := \liminf_{t\downarrow 0} \frac{h(x+tv)-h(x)}{t}.$$

The Clarke–Rockafellar directional derivative of h at x in the direction v is denoted by $h^\uparrow(x, v)$ and defined by (cf. [29])

$$h^\uparrow(x,v) = \lim_{\varepsilon\downarrow 0} \limsup_{y\to_h x, t\downarrow 0} \inf_{\|u-v\|\le\varepsilon} \frac{h(y+tu)-h(y)}{t}, \tag{6}$$

where $y \to_h x$ means that $y \to x$ and $h(y) \to h(x)$. The Clarke subdifferential of h at x is defined by (cf. [29])

$$\partial_C h(x) := \{\xi \in R^n : \langle \xi,\ v\rangle \le h^\uparrow(x,v) \quad \text{for all } v \text{ in } R^n\}.$$

It is well known (cf. [29, Proposition 3.2.2]) that the epigraph of $h^{\uparrow}(x,\cdot)$ equals the Clarke tangent cone of the epigraph of h at $(x,h(x))$, that is

$$\text{epi } h^{\uparrow}(x,\cdot) = T_C\left(\text{epi } h, (x,h(x))\right). \tag{7}$$

Throughout we use B to denote the open unit ball of R^n. The following three lemmas regarding the directional derivatives will be useful for us.

Lemma 1 *Let $h : R^n \to R \cup \{+\infty\}$ be a proper lower semicontinuous function, $x \in$ dom h, and $v \in R^n$. Then the following assertions hold:*

(i) *For any function ξ with the property that $\xi(y) \to v$ when $y \to_h x$, one has*

$$h^{\uparrow}(x,v) \le \limsup_{y\to_h x, t\downarrow 0} \frac{h(y+t\xi(y)) - h(y)}{t}.$$

(ii) *For any sequences (x_k) and (t_k) with $(x_k) \to_h x$ and $(t_k) \downarrow 0$, there exists a sequence (v_k) such that $(v_k) \to v$ and*

$$\limsup_{k\to+\infty} \frac{h(x_k + t_k v_k) - h(x_k)}{t_k} \le h^{\uparrow}(x,v). \tag{8}$$

(iii) *If h is assumed to be Lipschitz around x, then there exist sequences (x_k), (t_k) and (v_k) such that $(x_k) \to_h x$, $(t_k) \downarrow 0$, $(v_k) \to v$ and*

$$\lim_{k\to+\infty} \frac{h(x_k + t_k v_k) - h(x_k)}{t_k} = h^{\uparrow}(x,v). \tag{9}$$

Proof Let $\varepsilon > 0$. Then there exists $\delta_0 > 0$ such that $\|\xi(y) - v\| < \varepsilon$ whenever $\|y - x\| < \delta_0$ and $|h(y) - h(x)| < \delta_0$. Thus, for any $\delta \in (0,\delta_0)$

$$\sup_{\substack{\|y-x\|\le\delta, 0<t\le\delta \\ |h(y)-h(x)|\le\delta}} \inf_{\|u-v\|\le\varepsilon} \frac{h(y+tu)-h(y)}{t} \le \sup_{\substack{\|y-x\|\le\delta, 0<t\le\delta \\ |h(y)-h(x)|\le\delta}} \frac{h(y+t\xi(y))-h(y)}{t}.$$

Taking infima on both sides over all $\delta \in (0,\delta_0)$, it follows that

$$\limsup_{\substack{y\to_h x \\ t\downarrow 0}} \inf_{\|u-v\|\le\varepsilon} \frac{h(y+tu)-h(y)}{t} \le \limsup_{y\to_h x, t\downarrow 0} \frac{h(y+t\xi(y))-h(y)}{t}.$$

Since this is true for arbitrary $\varepsilon > 0$, (i) follows. To prove (ii), let $(t_k) \downarrow 0$ and $(x_k) \to_h x$, that is $(x_k, h(x_k)) \to (x, h(x))$ in epi h. Since $(v, h^{\uparrow}(x,v))$ belongs to $T_C\left(\text{epi } h, (x,h(x))\right)$ by (7), it follows that there exists a sequence $(v_k, s_k) \to (v, h^{\uparrow}(x,v))$ such that $(x_k, h(x_k)) + t_k(v_k, s_k) \in$ epi h, i.e., $h(x_k + t_k v_k) \le h(x_k) + t_k s_k$ for each k. Thus,

$$\frac{h(x_k + t_k v_k) - h(x_k)}{t_k} \le s_k.$$

Letting $k \to +\infty$, we see that (ii) holds. Finally by (i), we have

$$h^{\uparrow}(x,v) \leq \limsup_{y \to_h x, t \downarrow 0} \frac{h\,(y + t(v + x - y)) - h(y)}{t},$$

and hence there exist sequences (x_k), (t_k) such that $(x_k) \to_h x$, $(t_k) \downarrow 0$ and

$$h^{\uparrow}(x,v) \leq \lim_{k \to +\infty} \frac{h\,(x_k + t_k(v + x - x_k)) - h(x_k)}{t_k}. \tag{10}$$

Let (v_k) be a sequence with the properties as that stated in (8) of (ii). Since h is assumed to be Lipschitz around x, (9) follows by combining (8) and (10). Thus (iii) is valid and the proof is complete. □

Recall that the cone of feasible directions of a convex set $C \subset R^n$ at a point $x \in C$ is, by definition, the set

$$\mathcal{F}_C(x) := \{v \in R^n : x + tv \in \ C \quad \text{for some} \ \ t > 0\}, \tag{11}$$

that is $\mathcal{F}_C(x)$ is the cone generated by $C - x$.

Lemma 2 *Let $h : R^n \to R \cup \{+\infty\}$ be a proper lower semicontinuous function. Suppose that* dom h *is a convex set and that the restriction of h to* dom h *is locally Lipschitz. Let $x \in$* dom h *and $v \in \mathcal{F}_{dom\,h}(x)$. Then*

$$\overline{d}^{+} h(x)(v) \leq h^{\uparrow}(x,v). \tag{12}$$

Moreover, if $0 < h(x) < +\infty$ *then*

$$\overline{d}^{+} \sqrt{h}(x)(v) \leq \frac{h^{\uparrow}(x,v)}{2\sqrt{h(x)}}. \tag{13}$$

Proof By assumptions, there exist $r > 0$ and $K_x > 0$ such that

$$|h(x_1) - h(x_2)| \leq K_x \|x_1 - x_2\| \quad \forall x_1, x_2 \in (x + rB) \cap \text{dom}\, h, \tag{14}$$

and there exists $t_0 > 0$ such that

$$x + tv \in \text{dom}\, h \quad \text{for} \quad \text{all} \ \ 0 < t \leq t_0. \tag{15}$$

Let $\varepsilon > 0$, $\delta := \min\{r/\|v\| + \varepsilon, t_0\}$ and let $U := \{u \in v + \varepsilon B : x + t_0 u \in \text{dom}\, h\}$. Then $x + tu \in (x + rB) \cap \text{dom}\, h$ for all $t \in (0, \delta)$ and $u \in U$. It follows from (14) and (15) that

$$|h(x + tu) - h(x + tv)| \leq K_x t\, \|v - u\| \leq K_x t \varepsilon, \quad \text{for all} \ \ t \in (0, \delta), u \in U.$$

Hence

$$\inf_{\|u-v\|\le\varepsilon} \frac{h(x+tu)-h(x)}{t} \ge -K_x\varepsilon + \frac{h(x+tv)-h(x)}{t}.$$

Taking upper limits on both sides, we have

$$\begin{aligned}
&\limsup_{y\to_h x, t\downarrow 0} \inf_{\|u-v\|\le\varepsilon} \frac{h(y+tu)-h(y)}{t} \\
&\quad \ge \limsup_{t\downarrow 0} \inf_{\|u-v\|\le\varepsilon} \frac{h(x+tu)-h(x)}{t} \\
&\quad \ge -K_x\varepsilon + \limsup_{t\downarrow 0} \frac{h(x+tv)-h(x)}{t} \\
&\quad = -K_x\varepsilon + \overline{d}^+ h(x)(v),
\end{aligned}$$

and (12) follows by taking limits as $\varepsilon \to 0$.

Finally, suppose $h(x) > 0$. Then $h(x+tv) > 0$ when $t > 0$ is small enough (since h is continuous on dom h). Passing to the upper limits in

$$\frac{\sqrt{h(x+tv)} - \sqrt{h(x)}}{t} = \frac{h(x+tv)-h(x)}{t} \bullet \frac{1}{\sqrt{h(x+tv)} + \sqrt{h(x)}},$$

one has

$$\overline{d}^+ \sqrt{h}(x)(v) = \frac{\overline{d}^+ h(x)(v)}{2\sqrt{h(x)}}$$

and so (13) follows from (12).

Lemma 3 *Let h, x and v be as in Lemma 2. Then*

$$\limsup_{\substack{y\to_h x \\ q\to v, q\in\mathcal{F}_{dom\,h}(y)}} h^{\uparrow}(y,q) \le h^{\uparrow}(x,v),$$

where

$$\limsup_{\substack{y\to_h x \\ q\to v, q\in\mathcal{F}_{dom\,h}(y)}} h^{\uparrow}(y,q) := \inf_{\delta>0} \sup_{\substack{\|y-x\|\le\delta, |h(y)-h(x)|\le\delta \\ \|q-v\|\le\delta, q\in\mathcal{F}_{dom\,h}(y)}} h^{\uparrow}(y,q).$$

Proof We assume $h^{\uparrow}(x,v) \ne +\infty$. Let (y_k) and (v_k) be two arbitrary sequences such that $(y_k) \to_h x$, $(v_k) \to v$ and $v_k \in \mathcal{F}_{\mathrm{dom}\,h}(y_k)$ with the associate $\lambda_k > 0$ (that is the line-segment $[y_k, y_k + \lambda_k v_k]$ is contained in dom h) for each k. It suffices to show that

$$\limsup_{k\to+\infty} h^{\uparrow}(y_k, v_k) \le h^{\uparrow}(x, v). \tag{16}$$

For each k, by the definition, we have

$$\begin{aligned} h^{\uparrow}(y_k, v_k) &= \sup_{\varepsilon>0} \limsup_{y\to_h y_k, t\downarrow 0} \inf \left\{ \tfrac{h(y+tu)-h(y)}{t} : \|u - v_k\| \le \varepsilon \right\} \\ &\le \sup_{\varepsilon>0} \limsup_{y\to_h y_k, t\downarrow 0} \inf \left\{ \tfrac{h(y+tu)-h(y)}{t} : u \in \mathcal{F}_{\mathrm{dom}\, h}(y) \text{ and } \|u - v_k\| \le \varepsilon \right\} \end{aligned}$$

(note that the set $\Delta(y) := \{u : u \in \mathcal{F}_{\mathrm{dom}\, h}(y) \text{ and } \|u - v_k\| \le \varepsilon\}$ is not empty for any $y \in y_k + \lambda_k \varepsilon B$; for instance it contains $v_k + y_k - y/\lambda_k$).
Let (m_k) be a sequence with $m_k < h^{\uparrow}(y_k, v_k)$ for each k. Then there exists a sequence $(\varepsilon_k) \downarrow 0$ such that $\varepsilon_k \lambda_k < 1/k$ for each k and

$$m_k < \limsup_{y\to_h y_k, t\downarrow 0} \inf \left\{ \frac{h(y+tu) - h(y)}{t} : u \in \mathcal{F}_{\mathrm{dom}\, h}(y) \text{ and } \|u - v_k\| \le \varepsilon_k \right\}.$$

Hence, for each k, there exist z_k and $t_k > 0$ such that

$$\|z_k - y_k\| \le \varepsilon_k \lambda_k, \qquad 0 < t_k \le \varepsilon_k \lambda_k, \qquad \|h(z_k) - h(y_k)\| \le \varepsilon_k \lambda_k \tag{17}$$

and

$$m_k < \inf \left\{ \frac{h(z_k + t_k u) - h(z_k)}{t_k} : u \in \mathcal{F}_{\mathrm{dom}\, h}(z_k) \text{ and } \|u - v_k\| \le \varepsilon_k \right\}. \tag{18}$$

Let $u_k := v_k + y_k - z_k/\lambda_k$. Then $\|u_k - v_k\| = \|y_k - z_k\|/\lambda_k \le \varepsilon_k$ and $u_k \in \mathcal{F}_{\mathrm{dom}\, h}(z_k)$ because $z_k + \lambda_k u_k = y_k + \lambda_k v_k \in \mathrm{dom}\, h$. Hence by (18), one has

$$m_k < \frac{h(z_k + t_k u_k) - h(z_k)}{t_k}. \tag{19}$$

Since $(y_k) \to_h x$ and by (17), we have $(z_k) \to_h x$. Hence one can apply Lemma 1 (ii) to find (u'_k) such that $(u'_k) \to v$ and

$$\limsup_{k\to+\infty} \frac{h\left(z_k + t_k u'_k\right) - h(z_k)}{t_k} \le h^{\uparrow}(x, v). \tag{20}$$

Moreover we assume without loss of generality that $u'_k \in \mathcal{F}_{\mathrm{dom}\, h}(z_k)$ with $z_k + t_k u'_k \in \mathrm{dom}\, h$ for each k thanks to the fact that $h^{\uparrow}(x, v) < +\infty$. Therefore, by (19),

$$m_k \le \frac{h\left(z_k + t_k u'_k\right) - h(z_k)}{t_k} + \frac{h\left(z_k + t_k u_k\right) - h\left(z_k + t_k u'_k\right)}{t_k}, \tag{21}$$

where the second term converges to zero as $k \to +\infty$ because the restriction of h to dom h is assumed to be locally Lipschitz. Thus, by (20) and (21), one has

$$\limsup_{k \to +\infty} m_k \leq h^{\uparrow}(x, v).$$

Therefore (16) is true since each m_k is any (arbitrary) number such that $m_k < h^{\uparrow}(y_k, v_k)$. □

We end this section with two propositions from the measure theory. The first is an extended version of Rademacher's Theorem (see [4]). The proof for the second result will be omitted as it is standard (via the polar coordinates).

Proposition 1 *Let $h : E \to R$ be a real-valued measurable function defined on a bounded measurable set E in R^n. Then h is differentiable almost everywhere on E if and only if the difference quotients of h are locally bounded almost everywhere on E.*

Proposition 2 *Let S be a set in R^n of measure zero, and let $z_0 \in R^n$. Then, for almost every $y \in R^n$, the intersection of S with the line-segment $L(y)$ with the end points y and z_0 is of zero measure with respect to the 1-dimensional Lebesgue measure of $L(y)$.*

3 Max-functions

This section is devoted to study the max-function G of the following type

$$G(\tau) = \max_{i \in I} G_i(\tau), \quad \text{for each } \tau \in C, \tag{22}$$

where I is an index-set, C is a nonempty closed convex subset of R^n and each G_i is a real-valued function on C. The investigation here follows closely to that of Clarke given in [1, Theorem 2.8.6], but our arguments are more complicated because of the presence of the constraint set C. For simplicity, assume that C affinely spans R^n. (Hence C is of positive measure in R^n.) Further, we make the following blanket assumptions which are in force throughout this section.

Assumption 1

(I) *For each $\tau \in C$, assume the active index subset $I(\tau)$ for τ is nonempty, that is*

$$I(\tau) := \{i \in I : G_i(\tau) = G(\tau)\} \neq \emptyset. \tag{23}$$

(II) *For each $\tau \in C$, there exist positive real numbers δ_τ, L_τ such that for each $z \in (\tau + \delta_\tau B) \cap C$ and each $i \in I(z)$, G_i is Lipschitz on $(\tau + \delta_\tau B) \cap C$ with modulus L_τ, that is*

$$|G_i(\tau') - G_i(\tau'')| \leq L_\tau \|\tau' - \tau''\|, \quad \forall \tau', \tau'' \in (\tau + \delta_\tau B) \cap C. \tag{24}$$

Remark 1

(a) By (24), G is also locally Lipschitz around τ with modulus L_τ, that is,

$$|G(\tau') - G(\tau'')| \le L_\tau \|\tau' - \tau''\|, \quad \forall \tau', \tau'' \in (\tau + \delta_\tau B) \cap C. \tag{25}$$

In fact, for $i \in I(\tau')$, one has

$$G(\tau') - G(\tau'') = G_i(\tau') - G(\tau'') \le G_i(\tau') - G_i(\tau'') \le L_\tau \|\tau' - \tau''\|,$$

where $\tau', \tau'' \in (\tau + \delta_\tau B) \cap C$; thus (25) holds by symmetry.

(b) By (a) and the Rademacher Theorem (see Proposition 1), G and each G_i are differentiable almost everywhere on C. Note that, if $z \in (\tau + \delta_\tau B) \cap \operatorname{int} C$ and if G (resp.G_i) is differentiable at z, then

$$\|\nabla G(z)\| \le L_\tau \quad (\text{resp. } \|\nabla G_i(z)\| \le L_\tau). \tag{26}$$

For our convenience and for our subsequent use, we define a "modified function" $\overline{G}$ of G by

$$\overline{G}(\tau) = \begin{cases} G(\tau), & \tau \in C; \\ +\infty, & \text{otherwise.} \end{cases}$$

Clearly, $\overline{G}$ satisfies the conditions assumed in Lemma 2 (stated for the function h). Modified from [1, Theorem 2.8.6], we have the following proposition that provides an upper estimate for the directional derivative $\overline{G}^{\uparrow}(\tau, v)$ at $\tau \in C$ along a feasible direction v (see (28)).

Proposition 3 *Let $\tau \in C$ and $v \in \mathcal{F}_C(\tau)$ with the associate $\lambda_0 > 0$ such that the line-segment*

$$[\tau, \tau + \lambda_0 v] \subset C. \tag{27}$$

Let S be a subset of C with measure zero and let $A(v)$ be defined by

$$A(v) := \left\{ \lim_{k\to+\infty} \langle \nabla G_{i_k}(\tau_k),\ v\rangle : \tau_k, z_k \to \tau \text{ with } i_k \in I(z_k), \right.$$
$$\left. \tau_k \in \operatorname{int} C \backslash S \text{ and } z_k \in C \text{ for each } k \right\},$$
$$A_1(v) := \left\{ \lim_{k\to+\infty} \langle \nabla G_{i_k}(\tau_k),\ v_k\rangle : \tau_k, z_k \to \tau \text{ with } i_k \in I(z_k), \right.$$
$$\left. \tau_k \in \operatorname{int} C \backslash S \text{ and } z_k \in C \text{ for each } k \right\},$$

where (v_k) is any sequence convergent to v; for example, corresponding to $(\tau_k) \to \tau$, let $v_k := v + (\tau - \tau_k/\lambda_0)$.

Then the following assertions hold:

(i) $A_1(v) = A(v)$.
(ii) $A(v)$ *is nonempty and compact.*
(iii) *A real number r belongs to* $A(v)$ *if and only if for any each* $\varepsilon > 0$, *there exist* $\tau^\varepsilon \in (\tau + \varepsilon B)/S$, $z^\varepsilon \in \tau + \varepsilon B$ *and* $t^\varepsilon \in I(z^\varepsilon)$ *such that* G_{t^ε} *is differentiable at* τ^ε *and*

$$|r - \langle \nabla G_{t^\varepsilon}\left(\tau^\varepsilon\right), v\rangle| < \varepsilon.$$

(iv) *The Clarke–Rockafellar directional derivative of* $\overline{G}$ *at* τ *along the direction* v *satisfies the inequality*

$$\overline{G}^{\uparrow}(\tau, v) \leq \max A(v) := \max_{\xi \in A(v)} \xi. \tag{28}$$

Proof Take a sequence $(z_k) \to \tau$ with $z_k \in C$ for each k. By (I) of Assumption 1, there exists a sequence (i_k) such that $i_k \in I(z_k)$ for any k. Let $\Omega = \cup_{k=1}^{+\infty}\Omega_{ik}$ where Ω_{ik} denotes the set of points in $(\tau + \delta_\tau B) \cap C$ at which G_{i_k} fails to be differentiable. Then Ω is of measure zero by the Rademacher Theorem. Since int C is dense in C (as aff $C = R^n$ and C is convex), it follows that $((\tau + \delta_\tau B) \cap \text{int } C) / (S \cup \Omega)$ is dense in $(\tau + \delta_\tau B) \cap C$. Thus there exists a sequences (τ_k) in $(\tau + \delta_\tau B) \cap (\text{int } C \backslash S)$ convergent to τ such that for each k, $\nabla G_{i_k}(\tau_k)$ exists. Note further that

$$\|\nabla G_{i_k}(\tau_k)\| \leq L_\tau \quad \text{for each } k.$$

It is now clear that $A(v)$ is a nonempty bounded set. Moreover, for any (v_k) as in the definition for set $A_1(v)$, one has

$$\|\langle \nabla G_{i_k}(\tau_k), v\rangle - \langle \nabla G_{i_k}(\tau_k), v_k\rangle\| \leq L_\tau \|v - v_k\|,$$

and it follows that $A_1(v) = A(v)$. The verification for (iii) is routine, as well as the verification that $A(v)$ is closed thanks to (iii). Therefore $A(v)$ is compact and hence has a maximal element. Recalling that $\mathcal{F}_C(\tau)$ is the cone generated by $C - \tau$, to prove (28) it suffices to consider the case when $v \in C - \tau$ (then $\lambda_0 = 1$ in (27)). Denote m for $\max A(v)$. Then for any $\varepsilon > 0$, the definitions of m and $A_1(v)$ imply that there exists $\delta \in (0, \delta_\tau/2)$ such that if $y, z \in (\tau + 2\delta B) \cap C$ and $i \in I(z)$ satisfy

$$y \in \text{int } C\backslash S, \quad \nabla G_i(y) \text{ exists},$$

then one has

$$\langle \nabla G_i(y), v + (\tau - y)\rangle < m + \varepsilon. \tag{29}$$

Let x be any point in $(\tau + \delta B) \cap C$. Let $\lambda_1 = \min\{(1/2), (\delta/2\|v\|)\}$ and s be any number in $(0, \lambda_1]$. Let $z = x + s(v + \tau - x)$ and $i \in I(z)$. Let Ω_i be the set of points in $(\tau + \delta_\tau B) \cap C$ at which G_i fails to be differentiable. Then Ω_i is of measure zero. For all $y \in (\tau + \delta B) \cap C$, one has

$$[y, y + s(v + \tau - y)] \subset [y, y + (v + \tau - y)] = [y, \tau + v],$$

$$\|y + \lambda_1(v + \tau - y) - y\| \leq \lambda_1\|v\| + \lambda_1\|\tau - y\| < \delta,$$

and so, by (27)

$$[y, y + s(v + \tau - y)] \subset (\tau + 2\delta B) \cap C,$$

(in particular, replacing y by x one has $z \in (\tau + 2\delta B) \cap C$). On the other hand, $S \cup \text{bdry } C \cup \Omega_i$ is of measure zero since int C is dense in C. Thus, it follows from Proposition 2 (applied to $v + \tau$ in place of z_0) that for almost every $y \in R^n$, the intersection of $S \cup \text{bdry } C \cup \Omega_i$ with the line-segment $[y, v + \tau]$ is of measure zero with respect to the measure in that line-segment; for our convenience let Y denote the set of all $y \in R^n$ with the above property. Then $(\tau + \delta B) \cap C \cap Y$ is dense in $(\tau + \delta B) \cap C$. Moreover, let $y \in (\tau + \delta B) \cap C \cap Y$. Note that the set $\{\mu \in [0, s] : y + \mu(v + \tau - y) \in S \cup \text{bdry } C \cup \Omega_i\} = \{\mu \in [0, s] : (1 - \mu)y + \mu(v + \tau) \in S \cup \text{bdry } C \cup \Omega_i\}$ is of measure zero in $[0, s]$. Consequently, it follows from (29) that for almost every $\mu \in [0, s]$

$$\Big\langle \nabla G_i\,(y + \mu\,(v + \tau - y))\,,\; v + \tau - (y + \mu\,(v + \tau - y))\Big\rangle < m + \varepsilon,$$

that is

$$\Big\langle \nabla G_i\,(y + \mu\,(v + \tau - y))\,,\; v + \tau - y\Big\rangle < \frac{1}{1 - \mu}\,(m + \varepsilon)\,.$$

By integration over $[0, s]$ with respect to μ, this gives

$$\begin{aligned} G_i\,(y + s\,(v + \tau - y)) - G_i(y) &\leq (m + \varepsilon)\int_0^s \frac{1}{1 - \mu}\mathrm{d}\mu \\ &= (m + \varepsilon)\ln\frac{1}{1 - s}. \end{aligned} \tag{30}$$

By the continuity of G_i, it follows that (30) holds for every y in $(\tau + \delta B) \cap C$. Taking $y = x$, we have

$$\begin{aligned} G(z) - G(x) &\leq G_i(z) - G_i(x) \\ &= G_i\,(x + s\,(v + \tau - x)) - G_i(x) \\ &\leq (m + \varepsilon)\ln\tfrac{1}{1 - s}. \end{aligned}$$

That is

$$\frac{1}{s}\Big[\overline{G}\left(x+s\left(v+\tau-x\right)\right)-\overline{G}(x)\Big]=\frac{1}{s}\Big[G\left(x+s\left(v+\tau-x\right)\right)-G(x)\Big]$$
$$\leq\frac{1}{s}\left(m+\varepsilon\right)\ln\frac{1}{1-s}.$$

Passing to the limits as $x\rightarrow_C \tau$ and $s\downarrow 0$, this implies that

$$\limsup_{x\rightarrow_C\tau,s\downarrow 0}\frac{\overline{G}\left(x+s\left(v+\tau-x\right)\right)-\overline{G}(x)}{s}\leq m+\varepsilon,$$

and it follows from Lemma 1 (i) that

$$\overline{G}^{\uparrow}(\tau,v)\leq m+\varepsilon.$$

Then (28) holds as $\varepsilon>0$ is arbitrary.

4 Regularized gap functions

Without loss of generality, we assume throughout that P affinely spans R^n and that P contains the origin (by a translation argument if needed). We always assume that $F:P\rightarrow R^n$ is a locally Lipschitz map. Let $\gamma>0$ and let f_γ be defined as (2), namely

$$f_\gamma(\tau)=\max_{x\in P}\Psi(x,\tau)=\Psi\left(\pi_\gamma(\tau),\tau\right), \tag{31}$$

where $\pi_\gamma(\tau)$ is defined as in (5) and

$$\begin{aligned}\Psi(x,\tau)&:=\langle F(\tau),\ \tau-x\rangle-\tfrac{\gamma}{2}\|x-\tau\|^2,\\ &=-\tfrac{\gamma}{2}\|\tau-\tfrac{F(\tau)}{\gamma}-x\|^2+\tfrac{\|F(\tau)\|^2}{2\gamma}\qquad\forall(x,\tau)\in P\times P.\end{aligned} \tag{32}$$

It is not difficult to verify that π_γ and f_γ are locally Lipschitz on P.
Define $\overline{f}_\gamma:R^n\rightarrow R\cup\{+\infty\}$ by

$$\overline{f}_\gamma(\tau)=\begin{cases}f_\gamma(\tau), & \tau\in P;\\ +\infty, & \text{otherwise.}\end{cases} \tag{33}$$

Then $\operatorname{dom}\overline{f}_\gamma=P$ and $\overline{f}_\gamma$ satisfies the condition assumed in Lemma 2 stated for h. Hence by (13), for each $\tau\in P$ with $f_\gamma(\tau)>0$ and $v\in\mathcal{F}_P(\tau)$,

$$\underline{d}^+\sqrt{\overline{f}_\gamma}(\tau)(v)\leq\overline{d}^+\sqrt{\overline{f}_\gamma}(\tau)(v)\leq\frac{\overline{f}_\gamma^{\uparrow}(\tau,v)}{2\sqrt{f_\gamma(\tau)}}. \tag{34}$$

Recall from (5), (31) and (32) that

$$f_\gamma(\tau) = \max_{x \in P} \Psi(x, \tau),$$

the maximum being attained exactly at one point $x = \pi_\gamma(\tau)$. Together with the following lemma, $(f_\gamma,\ \Psi(x,\cdot),\ P,\ x \in P)$ satisfies Assumption 1 stated for $(G,\ G_i,\ C,\ i \in I)$ and hence, by Proposition 3,

$$\overline{f}_\gamma^{\uparrow}(\tau, v) \quad \leq \max A(v) \quad \text{for all } \tau \in P \text{ and } v \in \mathcal{F}_P(\tau), \tag{35}$$

where

$$A(v) = \left\{ \lim_{k \to +\infty} \langle \nabla_2 \Psi\left(\pi_\gamma(z_k), \tau_k\right),\ v \rangle : \tau_k, z_k \to \tau \text{ with each } \tau_k \in \operatorname{int} P \backslash \Omega_F \right\}, \tag{36}$$

$\nabla_2 \Psi\left(\pi_\gamma(z_k), \tau_k\right)$ denotes the derivative of the function $\Psi\left(\pi_\gamma(z_k), \cdot\right)$ at τ_k, $U_\tau := (\tau + \delta_\tau B) \cap P$ with some $\delta_\tau > 0$ is a neighborhood of τ on which F is Lipschitz, and

$$\Omega_F := \{y \in U_\tau : F \text{ fails to be differentiable at } y\}$$

(Ω_F is of measure zero by the Rademacher Theorem and $\operatorname{int} P \neq \emptyset$ since P is convex and $\operatorname{aff} P = R^n$).

Lemma 4 *Let $\tau \in P$. Let U_τ and Ω_F be as explained before the statement of the lemma. Then there exists a constant $L_\tau > 0$ such that each function in the family*

$$\{\Psi\left(\pi_\gamma(z), \cdot\right) : z \in U_\tau\}$$

is Lipschitz on U_τ with modulus L_τ, that is for each $z \in U_\tau$

$$|\Psi\left(\pi_\gamma(z), \tau'\right) - \Psi\left(\pi_\gamma(z), \tau''\right)| \leq L_\tau \|\tau' - \tau''\| \quad \text{for each } \tau', \tau'' \in U_\tau. \tag{37}$$

Consequently, f_γ is also Lipschitz with modulus L_τ on U_τ.

Proof Let $M_\tau > 0$ be a Lipschitz constant for F on U_τ. Then there exists a constant $C_1 > 0$ such that $\|F\left(\tau'\right)\| \leq C_1$ for all $\tau' \in U_\tau$ (e.g., take $C_1 := \|F(\tau)\| + M_\tau \delta_\tau$). Similarly, by (5) and since the projection is non-expansive, there exist constants $M_2, C_2 > 0$ such that

$$\|\pi_\gamma\left(\tau'\right) - \pi_\gamma\left(\tau''\right)\| \leq M_2 \|\tau' - \tau''\| \quad \forall \tau', \tau'' \in U_\tau$$

and

$$\|\pi_\gamma\left(\tau'\right)\| \leq C_2 \quad \forall \tau' \in U_\tau.$$

Now, in view of (32), we write for all $z, \tau' \in U_\tau$,

$$\Psi\left(\pi_\gamma(z), \tau'\right) = -\Psi_1^2\left(\pi_\gamma(z), \tau'\right) + \Psi_2^2\left(\pi_\gamma(z), \tau'\right)$$

where

$$\Psi_1\left(\pi_\gamma(z), \tau'\right) := \left(\frac{\gamma}{2}\right)^{\frac{1}{2}} \left\| \tau' - \frac{F\left(\tau'\right)}{\gamma} - \pi_\gamma(z) \right\|$$

$$\Psi_2\left(\pi_\gamma(z), \tau'\right) := \left(\frac{1}{2\gamma}\right)^{\frac{1}{2}} \|F\left(\tau'\right)\|.$$

Then there exist $M_3, C_3 > 0$ such that for $z \in U_\tau$, $\Psi_1\left(\pi_\gamma(z), \cdot\right)$ is Lipschitz on U_τ with modulus M_3 and

$$|\Psi_1\left(\pi_\gamma(z), \cdot\right)| \leq C_3 \text{ on } U_\tau.$$

Similar constants M_4, C_4 are for Ψ_2. Note that, for all $z, \tau', \tau'' \in U_\tau$,

$$\begin{aligned} &|\Psi_1^2\left(\pi_\gamma(z), \tau'\right) - \Psi_1^2\left(\pi_\gamma(z), \tau''\right)| \\ &\quad \leq 2C_3 \|\Psi_1\left(\pi_\gamma(z), \tau'\right) - \Psi_1\left(\pi_\gamma(z), \tau''\right)\| \\ &\quad \leq 2C_3 M_3 \|\tau' - \tau''\| \end{aligned}$$

and

$$|\Psi_2^2\left(\pi_\gamma(z), \tau'\right) - \Psi_2^2\left(\pi_\gamma(z), \tau''\right)| \leq 2C_4 M_4 \|\tau' - \tau''\|.$$

Thus, (37) holds with $L_\tau := 2C_3M_3 + 2C_4M_4$. Consequently, it follows from Remark 1 (a) that f_γ is also Lipschitz with modulus L_τ on U_τ because $f_\gamma(\tau) = \max_{x \in P} \Psi(x, \tau)$ and $\left(f_\gamma,\ \Psi(x, \cdot),\ P,\ x \in P,\ \pi_\gamma(z)\right)$ satisfies Assumption 1 stated for $(G,\ G_i,\ C,\ i \in I,\ I(z))$.

Theorem 1 *Let $\tau \in P$ and $\omega = \pi_\gamma(\tau) - \tau$. Let $v \in \mathcal{F}_P(\tau)$. Then*

$$\overline{f}_\gamma^{\uparrow}(\tau, v) = \max\{\langle \xi, -v \rangle : \xi \in D(\omega)\} + \langle F(\tau),\ v \rangle + \gamma \langle \omega,\ v \rangle \tag{38}$$

where

$$D(\omega) := \left\{ \lim_{k \to +\infty} \nabla F(\tau_k)\omega : (\tau_k) \to \tau \text{ and } \tau_k \in \mathit{int}\ P \backslash \Omega_F \text{ for each } k \right\}. \tag{39}$$

Proof Let $M := \max A(v)$, where $A(v)$ is as in (36). Then $\overline{f}_\gamma^{\uparrow}(\tau, v) \le M$ as in (35) and there exist some sequences $(\tau_k), (z_k)$ in U_τ convergent to τ such that each $\tau_k \in \operatorname{int} P \backslash \Omega_F$ and

$$M = \lim_{k \to +\infty} \left\langle \nabla_2 \Psi \left(\pi_\gamma(z_k), \tau_k \right), \ v \right\rangle.$$

By (32), we note (similar as in (4)) that

$$\nabla_2 \Psi \left(\pi_\gamma(z_k), \tau_k \right) = \nabla F(\tau_k) \left(\tau_k - \pi_\gamma(z_k) \right) + F(\tau_k) + \gamma \left(\pi_\gamma(z_k) - \tau_k \right).$$

Since F is Lipschitz around τ and $\tau_k \in \operatorname{int} P$ for all k, we assume without loss of generality that $\lim_{k \to +\infty} \nabla F(\tau_k)$ exists, and it follows that

$$\lim_{k \to +\infty} \nabla_2 \Psi \left(\pi_\gamma(z_k), \tau_k \right) = \lim_{k \to +\infty} \nabla F(\tau_k)(-\omega) + F(\tau) + \gamma\omega = -\overline{\xi} + F(\tau) + \gamma\omega,$$

where $\overline{\xi} := \lim_{k \to +\infty} \nabla F(\tau_k)\omega$. Note that $\overline{\xi} \in D(\omega)$ by (39) and

$$\overline{f}_\gamma^{\uparrow}(\tau, v) \le M = \left\langle -\overline{\xi} + F(\tau) + \gamma\omega, \ v \right\rangle.$$

Therefore, to prove (38), it suffices to show that

$$\overline{f}_\gamma^{\uparrow}(\tau, v) \ge \langle \xi, -v \rangle + \langle F(\tau), \ v \rangle + \gamma \langle \omega, \ v \rangle \qquad \text{for each } \xi \in D(\omega). \tag{40}$$

To do this, let $\xi \in D(\omega)$. Then $\xi = \theta \cdot \omega$ where $\theta = \lim\limits_{k \to +\infty} \nabla F(\tau_k)$ for some sequence $(\tau_k) \to \tau$ such that $\tau_k \in \operatorname{int} P \backslash \Omega_F$ for all k. Note that

$$\begin{aligned} \langle \xi, -v \rangle &= \langle \theta \cdot \omega, \ -v \rangle \\ &= \left\langle -\omega, \ \lim_{k \to +\infty} \nabla F(\tau_k)^T v \right\rangle \\ &= \left\langle \lim_{k \to +\infty} \lim_{t \downarrow 0} \tfrac{F(\tau_k + tv) - F(\tau_k)}{t}, \ -\omega \right\rangle. \end{aligned}$$

Consequently there exist a subsequence $\left(\tau_{k_i}\right)$ of (τ_k) and a sequence $(t_i) \downarrow 0$ such that

$$\langle \xi, -v \rangle = \left\langle \lim_{i \to +\infty} \frac{F\left(\tau_{k_i} + t_i v\right) - F\left(\tau_{k_i}\right)}{t_i}, \ -\omega \right\rangle.$$

For simplicity of notations, we henceforth assume that the above $\left(\tau_{k_i}\right)$ is (τ_k) itself, that is

$$\langle \xi, -v \rangle = \left\langle \lim_{k \to +\infty} \frac{F\left(\tau_k + t_k v\right) - F\left(\tau_k\right)}{t_k}, \ -\omega \right\rangle. \tag{41}$$

On the other hand, by Lemma 1 (ii), there exists a sequence $(v_k) \to v$ such that

$$\limsup_{k\to+\infty} \frac{\overline{f}_\gamma(\tau_k + t_k v_k) - \overline{f}_\gamma(\tau_k)}{t_k} \leq \overline{f}_\gamma^{\uparrow}(\tau, v) \ (\leq M < +\infty). \tag{42}$$

Then we can assume that $\overline{f}_\gamma(\tau_k + t_k v_k) < \infty$ for each k, and $\overline{f}_\gamma$ can be replaced by f_γ in the left-hand side of (42). We note that

$$\begin{aligned}
&f_\gamma(\tau_k + t_k v_k) - f_\gamma(\tau_k)\\
&\quad = \Psi\left(\pi_\gamma(\tau_k + t_k v_k), \tau_k + t_k v_k\right) - \Psi\left(\pi_\gamma(\tau_k), \tau_k\right)\\
&\quad \geq \Psi\left(\pi_\gamma(\tau_k), \tau_k + t_k v_k\right) - \Psi\left(\pi_\gamma(\tau_k), \tau_k\right)\\
&\quad = \langle F(\tau_k + t_k v_k),\ \tau_k + t_k v_k - \pi_\gamma(\tau_k)\rangle - \frac{\gamma}{2}\|\pi_\gamma(\tau_k) - (\tau_k + t_k v_k)\|^2\\
&\qquad -\langle F(\tau_k),\ \tau_k - \pi_\gamma(\tau_k)\rangle + \frac{\gamma}{2}\|\pi_\gamma(\tau_k) - \tau_k\|^2\\
&\quad = \langle F(\tau_k + t_k v_k) - F(\tau_k),\ \tau_k - \pi_\gamma(\tau_k)\rangle + \langle F(\tau_k + t_k v_k),\ t_k v_k\rangle\\
&\qquad -\frac{\gamma}{2}\left(\|\pi_\gamma(\tau_k) - (\tau_k + t_k v_k)\|^2 - \|\pi_\gamma(\tau_k) - \tau_k\|^2\right)
\end{aligned}$$

and hence that

$$\begin{aligned}
&\frac{1}{t_k}\left[f_\gamma(\tau_k + t_k v_k) - f_\gamma(\tau_k)\right]\\
&\quad \geq \left\langle \frac{F(\tau_k + t_k v_k) - F(\tau_k)}{t_k},\ -\left(\pi_\gamma(\tau_k) - \tau_k\right)\right\rangle + \langle F(\tau_k + t_k v_k),\ v_k\rangle\\
&\qquad -\frac{\gamma}{2}\frac{\|\pi_\gamma(\tau_k) - (\tau_k + t_k v_k)\|^2 - \|\pi_\gamma(\tau_k) - \tau_k\|^2}{t_k}\\
&\quad \to \langle \xi, -v\rangle + \langle F(\tau),\ v\rangle + \gamma\langle \omega,\ v\rangle \ \text{ as } k \to +\infty.
\end{aligned}$$

Here we have made use of (41) as well as the facts that $\left(\pi_\gamma(\tau_k) - \tau_k\right) \to \omega$, $v_k \to v$ and that F is Lipschitz on U_τ (and so $\{F(\tau_k + t_k v_k) - F(\tau_k)/t_k : k \in N\}$ is bounded). Consequently it follows from (42) that

$$\overline{f}_\gamma^{\uparrow}(\tau, v) \geq \langle \xi, -v\rangle + \langle F(\tau),\ v\rangle + \gamma\langle \omega,\ v\rangle,$$

i.e., (40) holds. □

5 Error bounds results

For the remainder of this paper, let F, P, γ, f_γ and Ψ be as at the beginning of the Sect. 4 and we assume that F is strongly monotone with modulus $\lambda > 0$ namely

$$\langle F(x') - F(x),\ x' - x\rangle \geq \lambda \parallel x' - x \parallel^2 \quad \forall x, x' \in P. \tag{43}$$

Under this assumption, VIP*(F, P)* is known to have a unique solution (cf. [3, Theorem 2.3.3]). We use x^* to denote the unique solution of VIP*(F, P)*. Thanks to the assumption (43), the following result is known (cf. [27, Lemma 3.1] and [26, Theorem 3.1]): For any $\tau \in P$, $f_\gamma(\tau) \geq 0$ and

$$f_\gamma(\tau) = 0 \Longleftrightarrow \pi_\gamma(\tau) = \tau \Longleftrightarrow \tau \text{ solves VIP}(F, P). \tag{44}$$

This section is devoted to show that the function $\sqrt{f_\gamma}$ has an error bound. Recall that (see e.g. [8, 15] and references therein) a proper function $h : R^n \to R \cup \{+\infty\}$ is said to have an error bound $\delta > 0$ on P if

$$\delta \ \mathrm{dist}(L_h, x) \leq h(x) \quad \text{for each } x \in P$$

where $L_h := \{z \in P : h(z) \leq 0\}$ and $\mathrm{dist}(L_h, x)$ denotes the distance from x to L_h. Recall from the beginning of this section that x^* denotes the unique solution of VIP (F, P). By (44), we have

$$0 = \inf_{\tau \in P} \sqrt{f_\gamma}(\tau) = \sqrt{f_\gamma}(x^*) \quad \text{and} \quad L_{\sqrt{f_\gamma}} = \{x^*\}.$$

Lemma 5 *Let τ, ω and $D(\omega)$ be as in Theorem 1. Then*

$$\overline{f}_\gamma^{\uparrow}(\tau, \omega) = \max\{\langle \xi, -\omega\rangle : \xi \in D(\omega)\} + \langle F(\tau),\ \omega\rangle + \gamma \|\omega\|^2. \tag{45}$$

Moreover,

$$\underline{d}^+ \sqrt{f_\gamma}(\tau)(\omega) \leq \overline{d}^+ \sqrt{f_\gamma}(\tau)(\omega) \leq \frac{\overline{f}_\gamma^{\uparrow}(\tau, \omega)}{2\sqrt{f_\gamma(\tau)}}. \tag{46}$$

Proof Since P is convex, it is easy to verify that $\omega \in \mathcal{F}_P(\tau)$. Thus (45) and (46) follow from (38) and (34) respectively. □

Definition 1 *Let $\lambda > 0$. Let $\iota_\lambda : (0, +\infty) \to (0, +\infty)$ be defined by*

$$\iota_\lambda(t) := min\left\{\frac{\sqrt{\lambda}}{2}, \frac{\lambda}{2\sqrt{t}}\right\} = \begin{cases} \frac{\sqrt{\lambda}}{2}, & 0 < t \leq \lambda \\ \frac{\lambda}{2\sqrt{t}}, & t > \lambda. \end{cases}$$

Theorem 2 *Let F, λ satisfy* (43) *and let $\gamma > 0$. Then for any $\tau \in P \backslash \{x^*\}$, one has*

$$\overline{f}_\gamma^{\uparrow}\left(\tau, \frac{\pi_\gamma(\tau) - \tau}{\|\pi_\gamma(\tau) - \tau\|}\right) \leq -2\iota_\lambda\left(\frac{\gamma}{2}\right)\sqrt{f_\gamma(\tau)} \tag{47}$$

and

$$\iota_\lambda\left(\frac{\gamma}{2}\right)\|\tau - x^*\| \le \sqrt{f_\gamma(\tau)} \quad \text{for each } \tau \in P, \tag{48}$$

where $f_\gamma, \overline{f}_\gamma$ *are defined by* (31) *and* (33).

Proof Let $\tau \in P\backslash\{x^*\}$. Then $f_\gamma(\tau) > 0$ and (33) shows that $\overline{f}_\gamma(\tau) = f_\gamma(\tau)$. For brevity, we denote $\omega := \pi_\gamma(\tau) - \tau$ as in Theorem 1. If (47) is valid, then

$$\underline{d}^+\sqrt{\overline{f}_\gamma}(\tau)\left(\frac{\omega}{\|\omega\|}\right) \le \overline{f}_\gamma^\uparrow\left(\tau, \frac{\omega}{\|\omega\|}\right) \Big/ \left(2\sqrt{f_\gamma(\tau)}\right) \le -\iota_\lambda\left(\frac{\gamma}{2}\right), \tag{49}$$

thanks to (46). Therefore,(48) follows from (47) and [15, Corollary 2.6]. We claim that

$$\overline{f}_\gamma^\uparrow(\tau,\omega) \le -\left(\lambda - \frac{\gamma}{2}\right)\|\omega\|^2 - f_\gamma(\tau). \tag{50}$$

To prove (50), note first that Lemma 5 shows

$$\overline{f}_\gamma^\uparrow(\tau,\omega) = \max\{\langle\xi, -\omega\rangle : \xi \in D(\omega)\} + \langle F(\tau),\ \omega\rangle + \gamma\|\omega\|^2. \tag{51}$$

By (39), each ξ in $D(\omega)$ can be expressed in the form $\xi = \lim_{k\to+\infty} \nabla F(\tau_k)\omega$ for some sequence $(\tau_k) \to \tau$ such that $\tau_k \in \operatorname{int} P\backslash\Omega_F$for each k. Since by (43), $\langle\nabla F(\tau_k)\omega,\ \omega\rangle \ge \lambda\|\omega\|^2$, it follows that

$$\langle\xi, -\omega\rangle \le -\lambda\|\omega\|^2. \tag{52}$$

On the other hand, since $\pi_\gamma(\tau)$ is the maximizer of the function $\Psi(\cdot,\tau)$ on P, the first order optimality condition implies that $\langle\nabla_1\Psi\left(\pi_\gamma(\tau),\tau\right),\ \tau' - \pi_\gamma(\tau)\rangle \le 0$ for any $\tau' \in P$. Letting $\tau' = \tau$ and noting $\nabla_1\Psi\left(\pi_\gamma(\tau),\tau\right) = -F(\tau) - \gamma\left(\pi_\gamma(\tau) - \tau\right)$, we have

$$\langle F(\tau) + \gamma\left(\pi_\gamma(\tau) - \tau\right),\ \pi_\gamma(\tau) - \tau\rangle \le 0,$$

that is

$$\langle F(\tau),\ \omega\rangle + \gamma\|\omega\|^2 \le 0.$$

Moreover, noting that $f_\gamma(\tau) = \langle F(\tau),\ -\omega\rangle - \frac{\gamma}{2}\|\omega\|^2$, the above inequality shows that

$$\frac{\|\omega\|}{\sqrt{f_\gamma(\tau)}} \le \sqrt{\frac{2}{\gamma}} \tag{53}$$

and (51), (52) also imply that

$$\overline{f}_\gamma^{\uparrow}(\tau,\omega)+f_\gamma(\tau)\le -\lambda\|\omega\|^2+\frac{\gamma}{2}\|\omega\|^2.$$

So, we have (50). The verification for (47) is now divided into three cases.
(a) $\gamma/2\le\lambda$ and $\|\omega\|/\sqrt{f_\gamma(\tau)}<1/\sqrt{\lambda}$.
(b) $\gamma/2\le\lambda$ and $\|\omega\|/\sqrt{f_\gamma(\tau)}\ge 1/\sqrt{\lambda}$.
(c) $\gamma/2>\lambda$.
In case (a), we have by (50) that

$$\begin{aligned}\overline{f}_\gamma^{\uparrow}\left(\tau,\tfrac{\omega}{\|\omega\|}\right)&\le-\left(\lambda-\tfrac{\gamma}{2}\right)\|\omega\|-\tfrac{f_\gamma(\tau)}{\|\omega\|}\\&\le-\tfrac{f_\gamma(\tau)}{\|\omega\|}\\&<-\sqrt{\lambda}\sqrt{f_\gamma(\tau)}\\&=-2\iota_\lambda\left(\tfrac{\gamma}{2}\right)\sqrt{f_\gamma(\tau)}.\end{aligned}$$

where the last equality holds by Definition 1.
In case (b), we have by (50) and (53) that

$$\overline{f}_\gamma^{\uparrow}\left(\tau,\frac{\omega}{\|\omega\|}\right)\le-\lambda\|\omega\|\le-\lambda\cdot\frac{1}{\sqrt{\lambda}}\sqrt{f_\gamma(\tau)}=-\sqrt{\lambda}\sqrt{f_\gamma(\tau)}=-2\iota_\lambda\left(\frac{\gamma}{2}\right)\sqrt{f_\gamma(\tau)}.$$

Finally, in case (c), we have by (50) and (53) that

$$\begin{aligned}\overline{f}_\gamma^{\uparrow}\left(\tau,\tfrac{\omega}{\|\omega\|}\right)&\le\left(\tfrac{\gamma}{2}-\lambda\right)\|\omega\|-\tfrac{f_\gamma(\tau)}{\|\omega\|}\\&\le\left(\tfrac{\gamma}{2}-\lambda\right)\sqrt{\tfrac{2}{\gamma}}\sqrt{f_\gamma(\tau)}-\sqrt{\tfrac{\gamma}{2}}\sqrt{f_\gamma(\tau)}\\&=-\lambda\sqrt{\tfrac{2}{\gamma}}\sqrt{f_\gamma(\tau)}\\&=-2\iota_\lambda\left(\tfrac{\gamma}{2}\right)\sqrt{f_\gamma(\tau)}.\end{aligned}$$

Therefore, (47) holds in all cases. □

Remark 2 One can consider more general type of regularized gap functions such as the one defined by (55) below, where one replaces the term $(1/2)\|x-\tau\|^2$ in (32) by a general function θ with the property (54). The following result not only extends [8, Theorem 2.1] (to the nonsmooth setting), but also provides an error bound constant which is defined by a function of one variable rather than by that of two variables as done in [8]. See Lemma 6 for the relation of these two functions.

Theorem 3 *Let F satisfy* (43). *Let $\theta : P\times P\to[0,\infty)$ be a function and $A>0$ such that*

$$\theta(x,\tau)\le A\|x-\tau\|^2,\quad \text{for all } x,\tau\in P. \tag{54}$$

Let $\gamma > 0$ *and* $\gamma' = \gamma A$. *Let* f_γ^θ *be defined by*

$$f_\gamma^\theta(\tau) := -\inf_{x\in P}\{F(\tau)(x-\tau) + \gamma\theta(x,\tau)\}, \quad \text{for each } \tau \in P. \tag{55}$$

Then, for any $\tau \in P$,

$$\sqrt{f_\gamma^\theta(\tau)} \geq \iota_\lambda(\gamma')\|\tau - x^*\|. \tag{56}$$

Proof By (2), we have

$$f_{2\gamma'}(\tau) := -\inf_{x\in P}\{\langle F(\tau),\ x-\tau\rangle + \gamma'\|x-\tau\|^2\} \quad \text{for each } \tau \in P.$$

Then, by Theorem 2,

$$\iota_\lambda(\gamma')\|\tau - x^*\| \leq \sqrt{f_{2\gamma'}(\tau)}. \tag{57}$$

By (54) , we have

$$-\gamma\theta(x,\tau) \geq -\gamma'\|x-\tau\|^2,$$

which implies that $f_\gamma^\theta(\tau) \geq f_{2\gamma'}(\tau)$ for each $\tau \in P$. Therefore, the result follows from (57).

Let $\lambda > 0$. For any $\gamma > 0$, following [8], we define $\delta_\gamma : (0,+\infty)\times(0,+\infty) \to [0,+\infty)$ by

$$\delta_\gamma(\sigma,\eta) := \begin{cases} \min\left\{\lambda\sigma, \frac{1}{4\sigma}\right\}, & \text{if } 0 < \gamma \leq \lambda \\ \min\{\lambda\sigma, \eta\sigma\}, & \text{if } \gamma > \lambda \quad \text{and} \quad 0 < \sigma \leq \frac{1}{2\sqrt{\gamma-\lambda+\eta}} \\ 0, & \text{otherwise.} \end{cases} \tag{58}$$

For the case when F is assumed to be smooth, it was shown in [8] that each non-zero value $\delta_{\gamma'}(\sigma,\eta)$($\gamma'$ is defined as in Theorem 3) is an error bound for the function f_γ^θ (see (55)).

Lemma 6 *Let* γ' *be defined as in Theorem 3. Then*

$$\iota_\lambda(\gamma') = \max_{\sigma,\eta>0} \delta_{\gamma'}(\sigma,\eta).$$

Proof First, we claim that

$$\iota_\lambda(\gamma') \geq \delta_{\gamma'}(\sigma,\eta) \quad \text{for all } \sigma,\eta > 0. \tag{59}$$

If $\delta_{\gamma'}(\sigma,\eta) = 0$, (59) is trivial. We suppose henceforth that $\delta_{\gamma'}(\sigma,\eta) > 0$. If $0 < \gamma' \le \lambda$, then by Definition 1, $\iota_\lambda(\gamma') = \sqrt{\lambda}/2 \ge \delta_{\gamma'}(\sigma,\eta)$ because

$$\delta_{\gamma'}(\sigma,\eta) = \min\left\{\lambda\sigma, \frac{1}{4\sigma}\right\} = \begin{cases} \lambda\sigma \le \frac{\sqrt{\lambda}}{2}, & \text{if } \sigma \le \frac{1}{2\sqrt{\lambda}} \\ \frac{1}{4\sigma} \le \frac{\sqrt{\lambda}}{2}, & \text{if } \sigma \ge \frac{1}{2\sqrt{\lambda}}. \end{cases}$$

If $\gamma' > \lambda$, Definition 1 shows that

$$\iota_\lambda(\gamma') = \frac{\lambda}{2\sqrt{\gamma'}} = \frac{\lambda}{2\sqrt{\gamma' - \lambda + \lambda}}. \tag{60}$$

In view of (58), we may suppose also that $0 < \sigma \le \frac{1}{2\sqrt{\gamma'-\lambda+\eta}}$. Then

$$\lambda\sigma \le \frac{\lambda}{2\sqrt{\gamma' - \lambda + \eta}} \qquad \text{and} \qquad \eta\sigma \le \frac{\eta}{2\sqrt{\gamma' - \lambda + \eta}}. \tag{61}$$

Since $\frac{\lambda}{2\sqrt{\gamma'-\lambda+\lambda}}$ dominates $\frac{\lambda}{2\sqrt{\gamma'-\lambda+\eta}}$ if $\lambda \le \eta$ and $\frac{\eta}{2\sqrt{\gamma'-\lambda+\eta}}$ if $\lambda > \eta$, it follows from (60) and (61) that

$$\iota_\lambda(\gamma') \ge \min\{\lambda\sigma, \eta\sigma\} = \delta_{\gamma'}(\sigma,\eta).$$

Therefore, by (58), (59) holds in all cases.

It remains to show that there exist $\sigma_0, \eta_0 > 0$ such that $\delta_{\gamma'}(\sigma_0,\eta_0) = \iota_\lambda(\gamma')$. Indeed, if $0 < \gamma' \le \lambda$, then letting $\sigma_0 = 1/2\sqrt{\lambda}$, it follows from the Definitions that $\delta_{\gamma'}(\sigma_0,\eta) = \iota_\lambda(\gamma') = \sqrt{\lambda}/2$ for any $\eta > 0$. Otherwise, $\gamma' \ge \lambda$, take $\sigma_0 = 1/2\sqrt{\gamma'}$ and $\eta_0 = \lambda$. Then $\delta_{\gamma'}(\sigma_0,\eta_0) = \iota_\lambda(\gamma') = \lambda/2\sqrt{\gamma'}$.

6 A descent method

Let $\gamma > 0$. Then by Theorem 2 one has for each $\tau \in P\backslash\{x^*\}$ and $\omega := \pi_\gamma(\tau) - \tau$,

$$\bar{f}_\gamma^\uparrow\left(\tau, \frac{\omega}{\|\omega\|}\right) \le -2\iota_\lambda\left(\frac{\gamma}{2}\right)\sqrt{f_\gamma(\tau)} < 0 \tag{62}$$

and

$$0 < \iota_\lambda(\frac{\gamma}{2})\|\tau - x^*\| \le \sqrt{f_\gamma(\tau)}. \tag{63}$$

Furthermore, it follows from (46) that

$$\overline{d}^{+}\sqrt{\overline{f}_{\gamma}}(\tau)(\omega) \leq \frac{\overline{f}_{\gamma}^{\uparrow}(\tau,\omega)}{2\sqrt{\overline{f}_{\gamma}}(\tau)} \leq -\iota_{\lambda}\left(\frac{\gamma}{2}\right)\|\omega\| < 0. \tag{64}$$

Hence

$$\sqrt{\overline{f}_{\gamma}(\tau+t\omega)} - \sqrt{\overline{f}_{\gamma}(\tau)} < -\frac{\iota_{\lambda}(\frac{\gamma}{2})}{2}t\|\omega\|, \quad \tau \in P\backslash\{x^*\} \tag{65}$$

for all sufficiently small $t > 0$. Moreover, $\overline{f}_{\gamma}$ can be replaced by f_{γ} in (65), because P is convex and $f = \overline{f}$ on P. Below we consider an algorithm of Armijo type.

Algorithm

Step 1. Let $\rho \in (0,1)$. Let τ_0 be a given vector in P. Set $k = 0$.
Step 2. If $f_{\gamma}(\tau_k) = 0$ then stop. If not then go to step 3.
Step 3. Let $\omega_k := \pi_{\gamma}(\tau_k) - \tau_k$.
Step 4. Let m_k be the smallest nonnegative integer such that

$$\sqrt{f_{\gamma}(\tau_k+\rho^{m_k}\omega_k)} - \sqrt{f_{\gamma}(\tau_k)} \leq -\frac{\iota_{\lambda}(\gamma/2)}{2}\rho^{m_k}\|\omega_k\| \tag{66}$$

and set $\tau_{k+1} = \tau_k + \rho^{m_k}\omega_k$. Return to step 2 with k replace by $k+1$.

Remark 3 By (65) and since $\rho \in (0,1)$, m_k in (66) dose exist. Moreover $\tau_{k+1} \in P$ because P is convex.

The following result is known as the Zagrodny Mean-valued theorem (see [28]), and we state it in a version that is convenient to us.

Lemma 7 *Let h be a lower semicontinuous function on R^n, $a, b \in$ dom h and $a \neq b$. Let $r \in R$ with $r \leq h(b)$. Then there exist sequences (x^k), $(x^k)^*$ in R^n and a point $c \in [a,b)$ such that $(x^k) \to_h c$ and $(x^k)^* \in \partial_C h(x^k)$ for each k such that*

$$r - h(a) \leq \liminf_{k\to+\infty}\langle (x^k)^*,\ b-a\rangle.$$

Theorem 4 *Let $\gamma > 0$. Suppose that F is strongly monotone and locally Lipschitz on P. Then the sequence (τ_k) generated by the above algorithm converges to the unique solution of VIP(F, P).*

Proof If $f_{\gamma}(\tau_k) = 0$ then $\tau_k = x^*$ by (44). Suppose therefore that $f_{\gamma}(\tau_k) > 0$ for each k. It follows from (66) that the sequence $(f_{\gamma}(\tau_k))$ is decreasing and hence converges to a nonnegative real number. Noting that the number of the right-hand side of (66) is negative, it follows that

$$\lim_{k\to+\infty} \rho^{m_k}\|\omega_k\| = 0. \tag{67}$$

Moreover the monotonicity of $\left(\sqrt{f_\gamma(\tau_k)}\right)$ also implies that

$$\tau_k \in P \text{ and } f_\gamma(\tau_k) \le f_\gamma(\tau_0) \quad \text{for each } k \tag{68}$$

and we deduce from (63) that

$$\iota_\lambda\left(\frac{\gamma}{2}\right)\|\tau_k - x^*\| \le \sqrt{f_\gamma(\tau_k)} \le \sqrt{f_\gamma(\tau_0)} \quad \text{for each } k.$$

In particular the sequence (τ_k) is bounded. Suppose that (τ_{k_i}) is a subsequence of $\{\tau_k\}$ such that $\lim_{(k_i\to+\infty)} \tau_{k_i} = x_*$ for some x_*. If $\pi_\gamma(x_*) = x_*$, then x_* is the solution of VIP*(F, P)* by (44). Now we assume that $\pi_\gamma(x_*) \neq x_*$. Since $\omega_{k_i} = \pi_\gamma(\tau_{k_i}) - \tau_{k_i} \to \pi_\gamma(x_*) - x_* \neq 0$, and by considering subsequences if necessary we suppose without loss of generality that $\{\|\omega_{k_i}\|\}_{i=1}^{+\infty}$ is bounded away from zero. Thus, (67) implies that $\lim_{k_i\to+\infty}\rho^{m_{k_i}} = 0$ (and so $m_{k_i} \to +\infty$). Note that, by continuity

$$\frac{\omega_{k_i}}{\|\omega_{k_i}\|} = \frac{\pi_\gamma(\tau_{k_i}) - \tau_{k_i}}{\|\pi_\gamma(\tau_{k_i}) - \tau_{k_i}\|} \to \frac{\pi_\gamma(x_*) - x_*}{\|\pi_\gamma(x_*) - x_*\|}.$$

Below let us consider an arbitrary i and keep it fixed. By Lemma 7 (applied to $\overline{f}_\gamma$, τ_{k_i}, $\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i}$ and $\overline{f}_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i})$ in place of h, a, b and r), there exist a point $c_{k_i} \in [\tau_{k_i}, \tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i})$ and sequences $(x_{k_i}^j)$ and $(x_{k_i}^j)^*$ with $(x_{k_i}^j) \to_{\overline{f}_\gamma} c_{k_i}$ and $(x_{k_i}^j)^* \in \partial_C\overline{f}_\gamma(x_{k_i}^j)$ for every natural number j such that

$$\frac{f_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i}) - f_\gamma(\tau_{k_i})}{\rho^{m_{k_i}-1}\|\omega_{k_i}\|} \le \liminf_{j\to+\infty}\left\langle (x_{k_i}^j)^*, \quad \frac{\omega_{k_i}}{\|\omega_{k_i}\|}\right\rangle \le \liminf_{j\to+\infty}\overline{f}_\gamma^\uparrow\left(x_{k_i}^j, \frac{\omega_{k_i}}{\|\omega_{k_i}\|}\right).$$

Since $(x_{k_i}^j) \to c_{k_i}$, $(\overline{f}_\gamma(x_{k_i}^j)) \to \overline{f}_\gamma(c_{k_i})$ and $\frac{\omega_{k_i}}{\|\omega_{k_i}\|} \in \mathcal{F}_P(c_{k_i})$, one can apply Lemma 3 to conclude that

$$\begin{aligned}\frac{f_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i}) - f_\gamma(\tau_{k_i})}{\rho^{m_{k_i}-1}\|\omega_{k_i}\|} &\le \limsup_{j\to+\infty}\overline{f}_\gamma^\uparrow\left(x_{k_i}^j, \frac{\omega_{k_i}}{\|\omega_{k_i}\|}\right)\\ &\le \overline{f}_\gamma^\uparrow\left(c_{k_i}, \frac{\omega_{k_i}}{\|\omega_{k_i}\|}\right).\end{aligned} \tag{69}$$

As this is shown to be valid for an arbitrary i, in passing to the limits as $i \to +\infty$, it follows that

$$\begin{aligned}\limsup_{i\to+\infty}\frac{f_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i}) - f_\gamma(\tau_{k_i})}{\rho^{m_{k_i}-1}\|\omega_{k_i}\|} &\le \limsup_{i\to+\infty}\overline{f}_\gamma^\uparrow\left(c_{k_i}, \frac{\omega_{k_i}}{\|\omega_{k_i}\|}\right)\\ &\le \overline{f}_\gamma^\uparrow\left(x_*, \frac{\pi_\gamma(x_*) - x_*}{\|\pi_\gamma(x_*) - x_*\|}\right),\end{aligned} \tag{70}$$

where the last inequality holds by Lemma 3 as $\pi_\gamma(x_*) - x_*/\|\pi_\gamma(x_*) - x_*\| \in \mathcal{F}_P(x_*)$ and $c_{k_i} \to_{\overline{f}_\gamma} x_*$ because $\overline{f}_\gamma$ is continuous on P and $(c_{k_i}) \subset P$. Since the line search rule (step 4) ensures

$$\frac{\sqrt{f_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i})} - \sqrt{f_\gamma(\tau_{k_i})}}{\rho^{m_{k_i}-1}\|\omega_{k_i}\|} > -\frac{\iota_\lambda(\frac{\gamma}{2})}{2} \quad \text{for each } i, \tag{71}$$

it follows from (70) that

$$\begin{aligned}
&\overline{f_\gamma}^{\uparrow}\left(x_*, \tfrac{\pi_\gamma(x_*)-x_*}{\|\pi_\gamma(x_*)-x_*\|}\right) \\
&\quad \geq \limsup_{k_i \to +\infty} \tfrac{f_\gamma(\tau_{k_i}+\rho^{m_{k_i}-1}\omega_{k_i}) - f_\gamma(\tau_{k_i})}{\rho^{m_{k_i}-1}\|\omega_{k_i}\|} \\
&\quad = \limsup_{k_i \to +\infty} \tfrac{\sqrt{f_\gamma(\tau_{k_i}+\rho^{m_{k_i}-1}\omega_{k_i})} - \sqrt{f_\gamma(\tau_{k_i})}}{\rho^{m_{k_i}-1}\|\omega_{k_i}\|} \\
&\quad\quad \times \lim_{k_i \to +\infty} \left(\sqrt{f_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i})} + \sqrt{f_\gamma(\tau_{k_i})}\right) \\
&\quad \geq -\tfrac{\iota_\lambda(\frac{\gamma}{2})}{2} \lim_{k_i \to +\infty} \left(\sqrt{f_\gamma(\tau_{k_i} + \rho^{m_{k_i}-1}\omega_{k_i})} + \sqrt{f_\gamma(\tau_{k_i})}\right) \\
&\quad = -\iota_\lambda(\tfrac{\gamma}{2})\sqrt{f_\gamma(x_*)}.
\end{aligned}$$

This contradicts (62) unless $x_* = x^*$. Consequently, (τ_k) must also converge to x^*. □

Remark 4 There already exist projection-type methods providing iterative sequences that converge to a solution assuming only F is monotone and continuous (e.g., [22, 23]). Our present approach (which requires the stronger assumption that F is strongly monotone and locally Lipschitz) is based on the consideration of error bounds of the merit function f_γ and hence we not only have the convergence result Theorem 4 but also know by (63) that how near to the solution from the kth point of the iteration.

Acknowledgements We are indebted to Professor J. S. Pang for helpful suggestions and for the references [7, 17, B.F.Hobbs and J.S.Pang sumbited]. We also wish to thank the referees and the Associate Editor for their valuable suggestions and for the references [22, 23].

References

1. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
2. Cottle, R.W.: Nolinear programs with positively bounded Jacobians. SIAM J. Appl. Math. **14**, 147–158 (1966)
3. Facchinei, F., Pang, J.S.: Finte-Dimensional Variational Inequalities and Complementarity Problems. Springer, Berlin Helidelberg, New York (2003)
4. Fadell, A.G.: A generalization of Rademacher's theorem on complete differential. Colloq. Math. **27**, 126–131 (1973)
5. Fukushima, M.: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Math. Program. **53**, 99–110 (1992)

6. 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)
7. Harrington, J.E., Hobbs, B.F., Pang, J.S., Liu, A., Roch, G.: Collusive game solutions via optimization. Math. Program. **104**, 407–435 (2005)
8. Huang, L.R., Ng, K.F.: Equivalent Optimization Formulations and Error Bounds for Variational Inequality Problems. J. Optim. Theory Appl. **125**, 299–314 (2005)
9. Jiang, H., Qi, L.: Local uniqueness and convergence of iterative methods for nonsmooth variational inequalities. J. Math. Anal. Appl. **196**, 314–331 (1995)
10. King, A.J., Rockafellar, R.T.: Sensitivity analysis for nonsmooth generalized equations. Math. Program. **55**, 193–202 (1992)
11. Kyparisis, J.: Uniqueness and differentiability of solutions of parametric nonlinear complementarity problems. Math. Program. **36**, 105–113 (1986)
12. Luc, D.T.: Existence results for densely pseudomonotone variational inequalities. J. Math. Anal. Appl. **254**, 291–308 (2001)
13. Luc, D.T.: The Frechet approximate jacobian and local uniqueness in variational inequalities. J. Math. Anal. Appl. **268**, 629–646 (2002)
14. Mordukhovich, B.S.: Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis. Trans. Am. Math. Soc. **343**, 609–658 (1994)
15. Ng, K.F., Zheng, X.Y.: Error bound for lower semicontinuous functions in normed space. SIAM J. Optim. **12**, 1–17 (2001)
16. Pang, J.S.: Complementarity problems. In: Horst, R., Pardalos, P. (eds) Handbook on Global Optimization, Kluwer, Boston (1994)
17. Pang, J.S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. Comput. Manag. Sci. **1**, 21–56 (2005)
18. Robinson, S.M.: Generalized equations. In: Bachem, A., Grötschel, M., Korte, B. (eds) Mathematical Programming: The State of the Art, pp 346–367, Springer,Berlin Helidelberg, New York (1983)
19. Rockafellar, R.T: Directionally Lipschitzian functions and subdifferential calculus. Proc. Lond Math. Soc. **39**, 331–355 (1979)
20. Rockafellar, R.T: Generalized directional derivatives and subgradients of nonconvex functions. Can. J. Math. **32**, 257–280 (1980)
21. Rockafellar, R.T., Wets, R.J-B.: Variational Analysis. Springer, Berlin Helidelberg, New York (1998)
22. Solodov, M.V., Tseng, P.: Modified projection-type methods for monotone variational inequalities. SIAM J. Control Optim. **34**, 1814–1830 (1996)
23. Solodov, M.V., Svaiter, B.F.: A new projection method for variational inequlity problems. SIAM J. Control Optim. **37**, 765–776 (1999)
24. Stampacchia, G.: Formes bilinaires coercitives sur les ensembles convexes. C. R. Acad. Sci. Paris **258**, 4413–4416 (1964)
25. Ward, D.E., Borwein, J.M.: Nonsmooth calculus in finite dimensions. SIAM J. Control Optim. **25**, 1312–1340 (1987)
26. Wu, J.H., Florian, M., Marcotte, P.: A general descent framework for the monotone variational inequality problem. Math. Program. **61**, 281–300 (1993)
27. Yamashita, N., Taji, K., Fukushima, M.: Unconstrained optimization reformulations of variational inequality problems. J. Optim. Theory Appl. **92**, 439–456 (1997)
28. Zagrodny, D.: Approximate mean value theorem for upper subderivatives. Nonlinear Anal. **12**, 1413–1428 (1988)
29. Zalinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)