Math. Program., Ser. A (2012) 134:365–393
DOI 10.1007/s10107-011-0440-8

FULL LENGTH PAPER

Subsmooth semi-infinite and infinite optimization problems

Xi Yin Zheng · Kung Fu Ng

Received: 5 March 2010 / Accepted: 23 December 2010 / Published online: 19 February 2011

Abstract We first consider subsmoothness for a function family and provide formulas of the subdifferential of the pointwise supremum of a family of subsmooth functions. Next, we consider subsmooth infinite and semi-infinite optimization problems. In particular, we provide several dual and primal characterizations for a point to be a sharp minimum or a weak sharp minimum for such optimization problems.

Keywords Subsmoothness · Infinite optimization · Semi-infinite optimization · Sharp minima · Weak sharp minima

Mathematics Subject Classification (2000) 90C30 · 90C34 · 49J52 · 65K10

1 Introduction

Among many operations in convex analysis and variational analysis, an important one is the classical operation of taking the pointwise supremum

$$\Phi(x) := \sup\{\phi_y(x) : \ y \in Y\} \quad \forall x \in X \tag{1.1}$$

This research was supported by an Earmarked Grant (GRF) from the Research Grant Council of Hong Kong and the National Natural Science Foundation of P. R. China (Grant No. 11061038).

X. Y. Zheng (✉)
Department of Mathematics, Yunnan University, Kunming 650091, People's Republic of China
e-mail: xyzheng@ynu.edu.cn

K. F. Ng
Department of Mathematics (IMS), The Chinese University of Hong Kong,
Shatin, New Territory, Hong Kong
e-mail: kfng@math.cuhk.edu.hk

of an arbitrarily indexed family of proper lower semicontinuous functions ϕ_y on a Banach space X with the index set Y. The objective of this paper is twofold. First we study the issue of representing the subdifferential $\partial\Phi(x)$ at $x \in X$ in terms of the subdifferentials $\partial\phi_y(x)$ of the functions ϕ_y. Second we consider the optimization problem with inequality constraint defined by $\{\phi_y : y \in Y\}$

$$\min f(x) \quad \text{subject to } \phi_y(x) \le 0 \quad \forall y \in Y \tag{1.2}$$

or, more generally,

$$\min f(x) \quad \text{subject to } \phi_y(x) \le 0 \quad \forall y \in Y \text{ and } x \in A \tag{OP}$$

where f is an extended-real valued function and A is a subset of X.

Throughout we make the following assumptions:

- X is a Banach space (with the topological dual denoted by X^*, the closed unit ball denoted by B_X, while $B(x, r)$ denotes the open ball with center x and radius r);
- the index set Y is a compact topological space;
- $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$;
- $f : X \to \overline{\mathbb{R}}$ is proper (so $\mathrm{dom}(f) := \{x \in X : f(x) < +\infty\}$ is nonempty) and lower semicontinuous;
- the function $(x, y) \mapsto \phi_y(x)$ is continuous on $X \times Y$.

When X is infinite dimensional, problem (1.2) is usually called an infinite optimization problem (cf. [32]). When X is finite dimensional, (1.2) is well studied as a semi-infinite optimization problem and has many important and interesting applications in engineering design, control of robots, mechanical stress of materials and social sciences; see the survey paper [15] and the books [3,11,28]. In the last three decades, semi-infinite optimization and its broad range of applications have been an active study area in mathematical programming (see [1,12,18,23,30] and references therein). In particular, many authors have studied first order optimality conditions of semi-infinite optimization problems with linear, convex or smooth data (cf. [17,20,33,38] and references).

The notion of a sharp minimum (namely a strong isolated minimum or strong unique local minimum) of real-valued functions plays an important role in the convergence analysis of numerical algorithms in mathematical programming problems (see [9,16,24,26]). As a generalization of sharp minima, the notion of weak sharp minima for real-valued functions was introduced and studied in [10]. Extensive study of weak sharp minima for real-valued convex functions has been done in the literature (cf. [4,5,31,34,35]). It has been found that the weak sharp minimum is closely related to the error bound in convex programming, a notion that has received much attention and has produced a vast number of publications (see [14,19,25,35,36] and references therein). Zheng and Yang [39,40] studied weak sharp minima for a semi-infinite optimization problem for both smooth and convex cases.

Covering both smooth and convex cases as well as the prox-regularity introduced by Poliquin and Rockafellar [27], a valuable extension is the notion of subsmoothness introduced and well studied by Aussel et al. [2]. Motivated by [2], Definition 3.1

introduces the notion of subsmoothness for a function family. For a subsmooth function family $\{\phi_y : y \in Y\}$ and under suitable Lipschitz conditions, we establish the following representation for the subdifferential of the supremum function Φ at $a \in X$:

$$\partial\Phi(a) = \overline{\text{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right) \tag{1.3}$$

and if X is finite dimensional,

$$\partial\Phi(a) = \text{co}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right) \tag{1.4}$$

where

$$Y(a) := \{y \in Y : \ \Phi(a) = \phi_y(a)\}$$

and the notations co and $\overline{\text{co}}^{w^*}$ (the weak*-closed convex hull) are standard. Results of types (1.3) and (1.4) have been established by several researchers under various degrees of generality and they have played a major role in establishing optimality conditions (see [6,19,25,33] and references therein). In Sect. 4 of this paper, (1.3) and (1.4) are applied to provide necessary/sufficient conditions (of Lagrangian type) for sharp/weak sharp minima of (OP) under appropriate subsmooth and Lipschitz assumptions on f, A and $\{\phi_y : \ y \in Y\}$. The last section is devoted to the finite dimensional case (with $\dim(X) = m - 1$ for some $m \geq 2$). Extending the well known results on smooth and convex semi-infinite optimization problems, we show in particular that (under the subsmooth and appropriate Lipschitz assumptions on the given data) if a feasible point $\bar{x}$ is a local solution of (OP) then there exist active indices y_i and $\lambda_i \in [0, \ +\infty)$ not all zero such that

$$0 \in \lambda_0\partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i\partial\phi_{y_i}(\bar{x}) + N(A, \bar{x})$$

(and $\lambda_0 \neq 0$ under a constraint qualification). In the same spirit we also provide a characterization for $\bar{x}$ to be a sharp/weak sharp minimum of (OP) under the subsmooth setting.

2 Preliminaries

Let A be a closed subset of X and $a \in A$. We denote by $T_c(A, a)$ and $T(A, a)$ the Clarke tangent cone and the contingent cone of A at a which are defined, respectively, by

$$T_c(A,a)=\liminf_{x\xrightarrow{A}a,t\to 0^+}\frac{A-x}{t}\quad\text{and}\quad T(A,a)=\limsup_{t\to 0^+}\frac{A-a}{t},$$

where $x \xrightarrow{A} a$ means that $x \to a$ with $x \in A$. Thus, $v \in T_c(A,a)$ if and only if, for each sequence $\{a_n\}$ in A converging to a and each sequence $\{t_n\}$ in $(0, \infty)$ decreasing to 0, there exists a sequence $\{v_n\}$ in X converging to v such that $a_n + t_n v_n \in A$ for all n in the set $\mathbb{N}$ of all natural numbers, while $v \in T(A,a)$ if and only if there exist a sequence $\{v_n\}$ converging to v and a sequence $\{t_n\}$ in $(0,\infty)$ decreasing to 0 such that $a + t_n v_n \in A$ for all $n \in \mathbb{N}$. We denote by $N(A,a)$ the Clarke normal cone of A at a, that is,

$$N(A,a) := \{x^* \in X^* : \langle x^*, h\rangle \le 0 \quad \text{for all } h \in T_c(A,a)\}.$$

For $\varepsilon \ge 0$ and $a \in A$, the nonempty set

$$\hat{N}_\varepsilon(A,a) := \left\{x^* \in X^* : \limsup_{x\xrightarrow{A}a}\frac{\langle x^*, x-a\rangle}{\|x-a\|} \le \varepsilon\right\}$$

is called the set of Fréchet ε-normals of A at a. When $\varepsilon = 0$, $\hat{N}_\varepsilon(A,a)$ is a convex cone which is called the Fréchet normal cone of A at a and is denoted by $\hat{N}(A,a)$. Let $N_M(A,a)$ denote the limiting normal cone of A at a in the Mordukhovich sense, that is,

$$N_M(A,a) := \limsup_{x\xrightarrow{A}a,\varepsilon\to 0^+}\hat{N}_\varepsilon(A,x).$$

Thus, $x^* \in N_M(A,a)$ if and only if there exists a sequence $\{(x_n, \varepsilon_n, x_n^*)\}$ in $A \times \mathbb{R}_+ \times X^*$ such that $(x_n, \varepsilon_n) \to (a, 0)$, $x_n^* \xrightarrow{w^*} x^*$ and $x_n^* \in \hat{N}_{\varepsilon_n}(A, x_n)$ for each n. It is known that

$$\hat{N}(A,a) \subset N_M(A,a) \subset N(A,a)$$

(cf. [22]). If A is convex, then $T_c(A,a) = T(A,a)$ and

$$N(A,a) = \hat{N}(A,a) = \{x^* \in X^* : \langle x^*, x\rangle \le \langle x^*, a\rangle \ \text{ for all } x \in A\}.$$

Let $f : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function. For $x \in \mathrm{dom}(f) := \{y \in X : f(y) < +\infty\}$ and $h \in X$, the generalized Rockafellar directional derivative of f at x along the direction h is defined by (see [6,29])

$$f^\circ(x;h) := \lim_{\varepsilon\downarrow 0}\limsup_{z\xrightarrow{f}x,\,t\downarrow 0}\inf_{w\in h+\varepsilon B_X}\frac{f(z+tw)-f(z)}{t},$$

where $z \xrightarrow{f} x$ means that $z \to x$ and $f(z) \to f(x)$. When f is locally Lipschitz at x, it is known that the generalized Rockafellar directional derivative reduces to the Clarke directional derivative, that is,

$$f^\circ(x, h) = \limsup_{y \to x,\, t \to 0^+} \frac{f(y + th) - f(y)}{t}.$$

Let $\partial f(x)$ denote the Clarke subdifferential of f at x, that is,

$$\partial f(x) := \{x^* \in X^* : \langle x^*, h\rangle \le f^\circ(x; h)\ \forall h \in X\}.$$

It is well known that

$$\partial f(x) = \{x^* \in X^* :\ (x^*, -1) \in N(\text{epi}(f), (x, f(x)))\}.$$

The Fréchet subdifferential and limiting(basic/Mordukhovich) subdifferential of f at x are denoted by $\hat{\partial} f(x)$ and $\partial_M f(x)$, respectively, that is,

$$\hat{\partial} f(x) := \{x^* \in X^* :\ (x^*, -1) \in \hat{N}(\text{epi}(f), (x, f(x)))\}$$

and

$$\partial_M f(x) := \{x^* \in X^* :\ (x^*, -1) \in N_M(\text{epi}(f), (x, f(x)))\}.$$

It is well known that

$$\hat{\partial} f(x) = \left\{x^* \in X^* : \liminf_{y \to x} \frac{f(y) - f(x) - \langle x^*, y - x\rangle}{\|y - x\|} \ge 0\right\}.$$

Recall that a Banach space X is called an Asplund space if every continuous convex function on X is Fréchet differentiable at each point of a dense subset of X. It is well known (cf. [22]) that X is an Asplund space if and only if every separable subspace of X has a separable dual space. In particular, every reflexive Banach space is an Asplund space. When X is an Asplund space, it is known (cf. [22]) that

$$N(A, a) = \overline{\text{co}}^{w^*}(N_M(A, a)), \quad N_M(A, a) = \limsup_{x \xrightarrow{A} a} \hat{N}(A, x), \tag{2.1}$$

$$\partial_M f(x) = \limsup_{u \xrightarrow{f} x} \hat{\partial} f(u) \quad \text{and} \quad \partial f(x) = \overline{\text{co}}^{w^*}(\partial_M f(x) + \partial_M^\infty f(x)), \tag{2.2}$$

where $\partial_M^\infty f(x) := \{x^* \in X^* :\ (x^*, 0) \in N_M(\text{epi}(f), (x, f(x)))\}$.

The following three lemmas can be found in [6] and are useful in the proofs of main results.

Lemma 2.1 *Let $x_1, x_2 \in X$ and suppose that f is a Lipschitz function on an open set containing the line segment $[x_1,\ x_2]$. Then there exists $u \in (x_1,\ x_2)$ and $u^* \in \partial f(u)$ such that*

$$f(x_2) - f(x_1) = \langle u^*, x_2 - x_1\rangle.$$

Lemma 2.2 *Let $f_1, f_2 : X \to \overline{\mathbb{R}}$ be proper lower semicontinuous functions. Let $\bar{x} \in \mathrm{dom}(f_1)$ and suppose that f_2 is locally Lipschitz at $\bar{x}$. Then*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

Lemma 2.3 *Let X, W be Banach spaces, $g : X \to W$ be a smooth function and $\psi : W \to \overline{\mathbb{R}}$ be a lower semicontinuous convex function. Let $\bar{x} \in X$ be such that $g(\bar{x}) \in \mathrm{dom}(\psi)$. Then*

$$\partial(\psi \circ g)(\bar{x}) = g'(\bar{x})^*(\partial\psi(g(\bar{x}))),$$

where $g'(\bar{x})^*$ denotes the conjugate operator of the derivative $g'(\bar{x})$.

We will also need the following approximate projection result (cf. [37, Theorem 3.1]).

Lemma 2.4 *Let X be a Banach space and A be a closed nonempty subset of X. Let $\gamma \in (0, 1)$. Then for any $x \notin A$ there exist a boundary point a of A and $a^* \in N(A, a)$ with $\|a^*\| = 1$ such that*

$$\gamma\|x - a\| < \min\left\{d(x, A), \langle a^*, x - a\rangle\right\},$$

where $d(x, A) := \inf\{\|x - u\| : u \in A\}$.

3 Subsmoothness for a function family

As an extension of convexity, prox-regularity expresses a variational behavior of "order two" and plays an important role in variational analysis and optimization (see [7,8,27,29]). As a generalization of the prox-regularity, Aussel et al. [2] introduced and studied the subsmoothness. A closed set A in X is said to be subsmooth at $a \in A$ if for any $\varepsilon > 0$ there exists $r > 0$ such that

$$\langle x^* - u^*, x - u\rangle \geq -\varepsilon\|x - u\|$$

whenever $x, u \in A \cap B(a, r)$, $x^* \in N(A, x) \cap B_{X^*}$ and $u^* \in N(A, u) \cap B_{X^*}$.

It is known (cf. [37]) that A is subsmooth at $a \in A$ if and only if for any $\varepsilon > 0$ there exists $r > 0$ such that

$$\langle u^*, x - u\rangle \leq \varepsilon\|x - u\| \quad \forall x, u \in A \cap B(a, r) \text{ and } u^* \in N(A, u) \cap B_{X^*}.$$

The following known lemma (cf. [37, Proposition 2.1]) is useful for us.

Lemma 3.1 *Let A be subsmooth at $a \in A$. Then, for any $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\langle u^*, x - u\rangle \le d(x, A) + \varepsilon\|x - u\| \;\; \forall x \in B(a, \delta)$$

whenever $u \in A \cap B(a, \delta)$ and $u^* \in N(A, u) \cap B_{X^*}$.

From this, it is easy to verify the following proposition.

Proposition 3.1 *Let $f : X \to \overline{\mathbb{R}}$ be a proper lower semicontinuous function and suppose that f is locally Lipschitz at $a \in \mathrm{dom}(f)$. Then $\mathrm{epi}(f)$ is subsmooth at $(a, f(a))$ if and only if for any $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\langle u^*, x - u\rangle \le f(x) - f(u) + \varepsilon\|x - u\| \;\; \forall x, u \in B(a, \delta) \text{ and } \forall u^* \in \partial f(u). \quad (*)$$

In view of Proposition 3.1, we say that a proper lower semicontinuous function $f : X \to \overline{\mathbb{R}}$ is subsmooth at a if for any $\varepsilon > 0$ there exists $\delta > 0$ such that the above (*) holds.

In the same line we can define the subsmoothness for a function family $\{\phi_y : y \in Y\}$ as follows.

Definition 3.1 We say that a function family $\{\phi_y : y \in Y\}$ is subsmooth at $a \in X$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\langle u^*, x - u\rangle \le \phi_y(x) - \phi_y(u) + \varepsilon\|x - u\| \quad (3.1)$$

whenever $(x, y), (u, y) \in B(a, \delta) \times Y$ and $u^* \in \partial\phi_y(u)$.

Further, we say that the family $\{\phi_y : y \in Y\}$ is subsmooth around a if there exists $\delta > 0$ such that it is subsmooth at each $x \in B(a, \delta)$.

The following proposition shows that the smooth assumption on the family $\{\phi_y : y \in Y\}$ (often considered in the literature on semi-infinite optimization problem (1.2)) implies the subsmoothness.

Proposition 3.2 *Suppose that ϕ_y is smooth for each $y \in Y$ and that the function $(u, y) \mapsto \phi'_y(u)$ is continuous on $X \times Y$, where $\phi'_y(u)$ denotes the derivative of ϕ_y at u. Then $\{\phi_y : y \in Y\}$ is subsmooth at each $a \in X$.*

Proof Let $a \in X$. We claim that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|\phi'_y(x_1) - \phi'_y(x_2)\| < \varepsilon \quad \forall (x_1, y), (x_2, y) \in B(a, \delta) \times Y. \quad (3.2)$$

Granting this and noting that $\partial_y\phi(u) = \{\phi'_y(u)\}$ and

$$\phi_y(x) - \phi_y(u) - \langle\phi'_y(u), x - u\rangle = \langle\phi'_y(u + \theta(x - u)) - \phi'_y(u), x - u\rangle$$

for all $(x, y), (u, y) \in X \times Y$ with corresponding $\theta \in (0,\ 1)$, it is easy to verify the desired assertion that $\{\phi_y : y \in Y\}$ is subsmooth at a. To prove (3.2), suppose to the contrary that there exist $\varepsilon_0 > 0$ and a sequence $\{(x_n, u_n, y_n)\}$ in $X \times X \times Y$ such that $(x_n, u_n) \to (a, a)$ and

$$\|\phi'_{y_n}(x_n) - \phi'_{y_n}(u_n)\| \geq \varepsilon_0 \quad \forall n \in \mathbb{N}, \tag{3.3}$$

where $\mathbb{N}$ denotes the set of all natural numbers. Since Y is compact, we can assume without loss of generality that $\{y_n\}$ converges to some $y_0 \in Y$ (passing to a *generalized* subsequence if necessary). Since $(u, y) \to \phi'_y(u)$ is continuous, it follows that $\phi'_{y_n}(x_n) \to \phi'_{y_0}(a)$ and $\phi'_{y_n}(u_n) \to \phi'_{y_0}(a)$. This contradicts (3.3). Hence, for any $\varepsilon > 0$ there exists $\delta > 0$ such that (3.2) holds. The proof is completed.

For several results later, let us introduce the following notion: a family $\{\phi_y : y \in Y\}$ is said to be locally Lipschitz at $a \in X$ if for any $v \in Y$ there exist $L_v, r_v \in (0, +\infty)$ and a neighborhood U_v of v such that

$$|\phi_y(x_1) - \phi_y(x_2)| \leq L_v\|x_1 - x_2\| \quad \forall (x_1, y), (x_2, y) \in B(a, r_v) \times U_v. \tag{3.4}$$

The following simple lemma is useful for our analysis later. Recall that Φ denotes the pointwise maximum of $\{\phi_y : y \in Y\}$, that is,

$$\Phi(x) = \sup_{y \in Y} \phi_y(x) \quad \forall x \in X.$$

Lemma 3.2 *Let $\{\phi_y : y \in Y\}$ be locally Lipschitz at a. Then there exist $L, r \in (0, +\infty)$ such that*

$$|\phi_y(x_1) - \phi_y(x_2)| \leq L\|x_1 - x_2\| \quad \forall x_1, x_2 \in B(a, r) \text{ and } y \in Y \tag{3.5}$$

and

$$|\Phi(x_1) - \Phi(x_2)| \leq L\|x_1 - x_2\| \quad \forall x_1, x_2 \in B(a, r). \tag{3.6}$$

Proof As it is easy to verify that (3.5) implies (3.6), we only need to show that (3.5) holds for some $L, r \in (0, +\infty)$. By the assumption, for each $v \in Y$ there exist $L_v, r_v \in (0, +\infty)$ and a neighborhood U_v of v such that (3.4) holds. Hence $\{U_v : v \in Y\}$ is an open cover of Y, and it follows from the compactness of Y that there exist $v_1, \ldots, v_k \in Y$ such that $Y = \bigcup_{i=1}^k U_{v_i}$. Letting $L := \max_{1 \leq i \leq k} L_{v_i}$ and $r := \min_{1 \leq i \leq k} r_{v_i}$, it follows from (3.4) that (3.5) holds.

An important class of subsmooth families is the composite-convex one.

Proposition 3.3 *Let X, W be Banach spaces and Y be a compact topological space. Let $\psi : W \times Y \to \mathbb{R}$ be a continuous function such that the function $z \mapsto \psi(z, y)$ is convex for each $y \in Y$ and let $g : X \to W$ be a smooth function. Let*

$$\phi_y(x) = \psi(g(x), y) \quad \forall (x, y) \in X \times Y.$$

Then $\{\phi_y : y \in Y\}$ is subsmooth and locally Lipschitz at each $a \in X$.

Proof Let $a \in X$. We first show that the family $\{\psi(\cdot, y) : \ y \in Y\}$ is locally Lipschitz at $g(a)$. Let $v \in Y$. Then there exist $M, \delta \in (0, \ +\infty)$ and a neighborhood U of v such that

$$|\psi(x, y)| \leq M \quad \forall (x, y) \in B(g(a), 2\delta) \times U \tag{3.7}$$

(thanks to the continuity of ψ). Let $y \in U, z_1, z_2 \in B(g(a), \delta)$ with $z_1 \neq z_2$, and let $z := z_2 + \frac{\delta(z_2 - z_1)}{\|z_2 - z_1\|}$. Then $z \in B(g(a), 2\delta)$ and $z_2 = \frac{z}{1+t} + \frac{tz_1}{1+t}$, where $t = \frac{\delta}{\|z_2 - z_1\|}$. It follows from the convexity assumption and (3.7) that

$$\begin{aligned} \psi(z_2, y) - \psi(z_1, y) &\leq \frac{1}{1+t}(\psi(z, y) - \psi(z_1, y)) \\ &\leq \frac{2M}{1+t} = \frac{2M\|z_2 - z_1\|}{\delta} \end{aligned}$$

Exchanging z_1 for z_2, it follows that $|\psi(z_2, y) - \psi(z_1, y)| \leq \frac{2M\|z_2 - z_1\|}{\delta}$. This shows that $\{\psi(\cdot, y) : \ y \in Y\}$ is locally Lipschitz at $g(a)$. By Lemma 3.2, there exist $L, r \in (0, \ +\infty)$ such that

$$|\psi(z_1, y) - \psi(z_2, y)| \leq L\|z_1 - z_2\| \quad \forall (z_1, y), (z_2, y) \in B(g(a), r) \times Y. \tag{3.8}$$

It follows that

$$\sup\{\|z^*\| : \ z^* \in \partial\psi(\cdot, y)(B(g(a), r))\} \leq L \quad \forall y \in Y. \tag{3.9}$$

Let $\varepsilon > 0$. Since g is smooth, there exist $L_1, \delta > 0$ such that

$$g(x) \in B(g(a), r), \ \|g(x) - g(u)\| \leq L_1\|x - u\|$$

and

$$\|g(x) - g(u) - g'(u)(x - u)\| \leq \frac{\varepsilon\|x - u\|}{L} \tag{3.10}$$

for all $x, u \in B(a, \delta)$. It follows from (3.8) that

$$|\phi_y(x) - \phi_y(u)| \leq LL_1\|x - u\| \quad \forall x, u \in B(a, \delta) \text{ and } y \in Y.$$

This shows that $\{\phi_y : \ y \in Y\}$ is locally Lipschitz at a.

On the other hand, by the convexity and smoothness assumptions, Lemma 2.3 implies that

$$\partial_y\phi(u) = g'(u)^*(\partial\psi(\cdot, y)(u)) \quad \forall (u, y) \in X \times Y.$$

Let $x, u \in B(a,\delta)$, $y \in Y$ and $z^* \in \partial\psi(\cdot, y)(u)$. Then, by (3.10), (3.9) and the convexity assumption, one has

$$\begin{aligned}\langle g'(u)^*(z^*), x-u\rangle &= \langle z^*, g'(u)(x-u)\rangle \\ &\le \langle z^*, g(x)-g(u)\rangle + \|z^*\|\|g(x)-g(u)-g'(u)(x-u)\| \\ &\le \psi(g(x),y) - \psi(g(u),y) + \varepsilon\|x-u\| \\ &= \phi_y(x) - \phi_y(u) + \varepsilon\|x-u\|.\end{aligned}$$

This shows that $\{\phi_y : y \in Y\}$ is subsmooth at a. The proof is completed.

The following theorem is a key of the proofs of the main results in this paper. Recall that

$$Y(x) = \{y \in Y : \phi_y(y) = \Phi(x)\} \quad \forall y \in Y.$$

Since the index set Y is compact and the function $(x, y) \mapsto \phi_y(x)$ is continuous, $Y(x)$ is nonempty.

Theorem 3.1 *Suppose that* $\{\phi_y : y \in Y\}$ *is subsmooth at* $a \in X$. *Then*

$$\partial\Phi(a) \supset \overline{\text{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right) \tag{3.11}$$

and for any $\varepsilon > 0$ *there exists* $\delta > 0$ *such that*

$$\langle x^*, x-a\rangle \le \Phi(x) - \Phi(a) + \varepsilon\|x-a\| \tag{3.12}$$

whenever $x \in B(a,\delta)$ *and* $x^* \in \overline{\text{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right)$. *If, in addition,* $\{\phi_y : y \in Y\}$ *is locally Lipschitz at a, then*

$$\partial\Phi(a) = \overline{\text{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right). \tag{3.13}$$

Proof Let $\varepsilon > 0$. By the subsmoothness assumption, there exists $\delta > 0$ such that (3.1) holds for all $(x, y), (u, y) \in B(a,\delta) \times Y$ and $u^* \in \partial\phi_y(u)$. Let $x^* \in \overline{\text{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right)$. Then there exists a *generalized* sequence $\{x^*_\alpha\}_{\alpha\in\Lambda}$ in $\text{co}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right)$ such that $x^*_\alpha \xrightarrow{w^*} x^*$. For each $\alpha \in \Lambda$, take a finite subset I_α of $Y(a)$, $t_i \ge 0$ and $x^*_i \in \partial\phi_i(a)$ ($i \in I_\alpha$) such that

$$\sum_{i\in I_\alpha} t_i = 1 \quad \text{and} \quad x^*_\alpha = \sum_{i\in I_\alpha} t_i x^*_i.$$

Noting that $\phi_{y'}(a) = \Phi(a)$ for all $y' \in Y(a)$, it follows from (3.1) that

$$\begin{aligned}\langle x_\alpha^*, x - a\rangle &= \sum_{i\in I_\alpha} t_i \langle x_i^*, x - a\rangle \\ &\le \sum_{i\in I_\alpha} t_i(\phi_i(x) - \phi_i(a) + \varepsilon\|x - a\|) \\ &\le \Phi(x) - \Phi(a) + \varepsilon\|x - a\|\end{aligned}$$

for all $x \in B(a, \delta)$. This implies that (3.12) holds. Hence $x^* \in \hat{\partial}\Phi(a) \subset \partial\Phi(a)$ and so (3.11) holds. Next suppose that the Lipschitz assumption holds. To prove (3.13), by (3.11) we only need to show that

$$\partial\Phi(a) \subset \overline{\text{co}}^* \left(\bigcup_{y\in Y(a)} \partial\phi_y(a) \right).$$

To do this, suppose to the contrary that there exists

$$x_0^* \in \partial\Phi(a) \backslash \overline{\text{co}}^{w^*} \left(\bigcup_{y\in Y(a)} \partial\phi_y(a) \right).$$

Noting that the weak*-closed convex set $\overline{\text{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right)$ is nonempty (because $Y(a) \neq \emptyset$ and $\partial\phi_y(a) \neq \emptyset$ for all $y \in Y(a)$), it follows from the separation theorem that there exists $h \in X\backslash\{0\}$ such that

$$\langle x_0^*, h\rangle > \sup\{\langle x^*, h\rangle :\ x^* \in \bigcup_{y\in Y(a)} \partial\phi_y(a)\}. \tag{3.14}$$

By the local Lipschitz assumption and Lemma 3.2, Φ is locally Lipschitz at a. Hence there exists a sequence $\{(x_n, t_n)\}$ in $X \times (0, +\infty)$ such that $(x_n, t_n) \to (a, 0)$ and

$$\lim_{n\to\infty} \frac{\Phi(x_n + t_n h) - \Phi(x_n)}{t_n} = \Phi^\circ(a, h).$$

Noting that $\langle x_0^*, h\rangle \le \Phi^\circ(a, h)$, it follows that

$$\langle x_0^*, h\rangle \le \lim_{n\to\infty} \frac{\Phi(x_n + t_n h) - \Phi(x_n)}{t_n}. \tag{3.15}$$

For each $n \in \mathbb{N}$, take $y_n \in Y(x_n + t_n h)$. Then

$$\phi_{y_n}(x_n + t_n h) - \phi_{y_n}(x_n) \ge \Phi(x_n + t_n h) - \Phi(x_n) \ \ \forall n \in \mathbb{N} \tag{3.16}$$

and

$$\phi_y(x_n + t_n h) \leq \phi_{y_n}(x_n + t_n h) \quad \forall (y, n) \in Y \times \mathbb{N}. \tag{3.17}$$

Since Y is compact, we can assume without loss of generality that $y_n \to y_0 \in Y$ (taking a *generalized* subsequence if necessary). Noting that $x_n + t_n h \to a$ and the function $(x, y) \mapsto \phi_y(x)$ is continuous, it follows from (3.17) that $\phi_y(a) \leq \phi_{y_0}(a)$ for all $y \in Y$, that is, $y_0 \in Y(a)$. By the Lipschitz assumption and Lemma 3.2, there exist $L \in (0, +\infty)$, $r \in (0, \delta)$ such that (3.5) holds. Since $(x_n, t_n) \to (a, 0)$, we can assume without loss of generality that $x_n, x_n + t_n h \in B(a, r)$ for all $n \in \mathbb{N}$. By (3.5) and Lemma 2.1, there exist $\theta_n \in (0, 1)$ and $x_n^* \in \partial\phi_{y_n}(x_n + \theta_n t_n h)$ such that $\|x_n^*\| \leq L$ and

$$\phi_{y_n}(x_n + t_n h) - \phi_{y_n}(x_n) = \langle x_n^*, t_n h\rangle.$$

Since B_{X^*} is compact with respect to the weak* topology, we can assume that $x_n^* \stackrel{w^*}{\to} a^*$ (passing to a *generalized* subsequence if necessary). Hence

$$\limsup_{n\to\infty} \frac{\phi_{y_n}(x_n + t_n h) - \phi_{y_n}(x_n)}{t_n} = \langle a^*, h\rangle.$$

It follows from (3.15) and (3.16) that

$$\langle x_0^*, h\rangle \leq \langle a^*, h\rangle. \tag{3.18}$$

On the other hand, by (3.1) and $r \in (0, \delta)$, one has

$$\langle x_n^*, x - (x_n + t_n h)\rangle \leq \phi_{y_n}(x) - \phi_{y_n}(x_n + t_n h) + \varepsilon\|x - (x_n + t_n h)\|$$

for all $x \in B(a, \delta)$ and $n \in \mathbb{N}$. It follows that $\langle a^*, x - a\rangle \leq \phi_{y_0}(x) - \phi_{y_0}(a)$ for all $x \in B(a, \delta)$. This implies that $a^* \in \hat{\partial}\phi_{y_0}(a) \subset \partial\phi_{y_0}(a)$, contradicting (3.14) and (3.18). The proof is completed.

In the finite dimensional case, we have the following sharper result.

Theorem 3.2 *Suppose that $\{\phi_y : y \in Y\}$ is subsmooth at $a \in X$ and locally Lipschitz at a. Further suppose that X is finite dimensional. Then*

$$\partial\Phi(a) = \mathrm{co}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right). \tag{3.19}$$

Proof Let $x^* \in \overline{\mathrm{co}}^{w^*}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right)$. By Theorem 3.1, it suffices to show that

$$x^* \in \mathrm{co}\left(\bigcup_{y\in Y(a)} \partial\phi_y(a)\right).$$

Take a *generalized* sequence $\{x_\alpha^*\}_{\alpha\in\Lambda}$ in co $\left(\bigcup_{y\in Y(a)}\partial\phi_y(a)\right)$ such that $x_\alpha^* \stackrel{w^*}{\to} x^*$. Let $m := \dim(X)+1$, where $\dim(X)$ denotes the dimension of X. Then, by the Carathéodory theorem, for each $\alpha\in\Lambda$ there exist $y_\alpha(k)\in Y(a)$, $x_\alpha^*(k)\in\partial\phi_{y_\alpha(k)}(a)$ and $t_\alpha(k)\in[0,\ 1]$ $(k=1,\dots,m)$ such that

$$\sum_{k=1}^m t_\alpha(k)=1 \ \text{ and } \ x_\alpha^*=\sum_{k=1}^m t_\alpha(k)x_\alpha^*(k)\stackrel{w^*}{\to} x^*. \tag{3.20}$$

Since $Y(a)$ is a closed subset of the compact topological space Y, without loss of generality, we assume that

$$t_\alpha(k)\to t_k \quad \text{and} \quad y_\alpha(k)\to y_k\in Y(a),\quad k=1,\dots,m. \tag{3.21}$$

By the Lipschitz assumption and Lemma 3.2, there exist $L,r\in(0,\ +\infty)$ such that (3.5) holds. It follows that $\|x_\alpha(k)^*\|\le L$ for all $\alpha\in\Lambda$ and $k=1,\dots,m$. Without loss of generality, we can assume that

$$x_\alpha^*(k)\stackrel{w^*}{\to} x_k^*,\quad k=1,\dots,m \tag{3.22}$$

It follows from (3.20) and (3.21) that

$$\sum_{k=1}^m t_k=1 \quad \text{and} \quad \sum_{k=1}^m t_k x_k^*=x^*. \tag{3.23}$$

Let $\varepsilon>0$. By the subsmoothness assumption, there exists $\delta>0$ such that

$$\langle x_\alpha^*(k),x-a\rangle\le\phi_{y_\alpha(k)}(x)-\phi_{y_\alpha(k)}(a)+\varepsilon\|x-a\|$$

for all $x\in B(a,\delta)$, $\alpha\in\Lambda$ and $k=1,\dots,m$. Since the function $(x,y)\mapsto\phi_y(x)$ is continuous, it follows from (3.21) and (3.22) that

$$\langle x_k^*,x-a\rangle\le\phi_{y_k}(x)-\phi_{y_k}(a)+\varepsilon\|x-a\|\ \ \forall x\in B(a,\delta)\text{ and }k=1,\dots,m.$$

Hence $x_k^*\in\hat\partial\phi_{y_k}(a)\subset\partial\phi_{y_k}(a)$ for each k. This and (3.23) imply that $x^*\in$ co $\left(\bigcup_{y\in Y(a)}\partial\phi_y(a)\right)$. The proof is completed.

Remark The subdifferential formula of a pointwise maximum function is important in both theory and application. The following results can be found in [35] and [6].

Theorem I *Suppose that ϕ_y is convex for each $y\in Y$. Then (3.13) holds.*

Theorem II *Suppose that Y is a compact metric space and that there exist $L,r\in(0,\ +\infty)$ such that*

$$|\phi_y(x_1)-\phi_y(x_2)|\le L\|x_1-x_2\|\ \ \forall(x_1,y),(x_2,y)\in B(a,r)\times Y.$$

Then

$$\partial\Phi(a) \subset \left\{ \int_Y \partial^{[Y]}\phi_y(a)d\mu : \ \mu \in \mathcal{M}(a) \right\},$$

where $\mathcal{M}(a)$ denotes the set of all Radon probability measures whose supports are contained in $Y(a)$ and $\partial^{[Y]}\phi_y(a)$ denotes the set

$$\overline{\text{co}}^{w^*}\left\{x^* \in X^* : \ x_n^* \in \partial\phi_{y_n}(x_n), \ x_n \to x, \ y_n \to y, \ x^* \text{ is a weak}^* \text{ cluster of } \{x_n^*\}\right\}.$$

Theorem I is well known as Ioffe and Tikhomirov theorem. Recently, under the convexity assumption, Hantoute et al. [13] and Lopez and Volle [21] further provided some formulas for the subdifferential of pointwise supremum functions.

Remark Under the subsmoothness assumption, we can prove that $(x, y) \to \partial_{[Y]}\phi_y(x)$ is weak* closed. In the case when the index set Y is a compact metric space, Theorem 3.1 can be proved in virtue of Theorem II.

By Proposition 3.3, Theorem 3.1 clearly extends Theorem I and can be regarded as a supplement of Theorem II.

4 Subsmooth infinite optimization problem

In this section, we consider the case when X is a general Banach space. Let Z denote the feasible set of (OP), that is,

$$Z = \{x \in A : \ \phi_y(x) \leq 0 \ \ \forall y \in Y\}.$$

In the remainder of this paper, we always assume that $\bar{x}$ *is a fixed feasible point* ($\bar{x} \in Z$) and

$$S_{\bar{x}} := \{x \in Z : \ f(x) = f(\bar{x})\};$$

we will often use the following condition:

Condition S *f, $\{\phi_y : \ y \in Y\}$ and A are subsmooth at $\bar{x}$.*

Needless to say, this condition is weaker than the following one:

Condition S$^+$ *f and A are subsmooth at $\bar{x}$ and $\{\phi_y : \ y \in Y\}$ is subsmooth around $\bar{x}$ and locally Lipschitz at $\bar{x}$.*

As in [39,40], we say that $\bar{x}$ is a sharp minimum of (OP) if there exist $\eta, \delta \in (0, +\infty)$ such that

$$\eta\|x - \bar{x}\| \leq f(x) - f(\bar{x}) + \sup_{y\in Y}[\phi_y(x)]_+ + d(x, A) \ \ \forall x \in B(\bar{x}, \delta) \qquad \text{(SM)}$$

and that $\bar{x}$ is a weak sharp minimum of (OP) if there exist $\eta, \delta \in (0, +\infty)$ such that

$$\eta d(x, S_{\bar{x}}) \leq f(x) - f(\bar{x}) + \sup_{y \in Y}[\phi_y(x)]_+ + d(x, A) \;\; \forall x \in B(\bar{x}, \delta), \tag{WM}$$

where $[\phi_y(x)]_+ := \max\{\phi_y(x), 0\}$.

Clearly, (WM) implies that

$$\eta d(x, S_{\bar{x}}) \leq f(x) - f(\bar{x}) \quad \forall x \in Z \cap B(\bar{x}, \delta)$$

and so $\bar{x}$ is a local solution of (OP). It is clear that (SM) implies that $\bar{x}$ is a local solution of (OP) and $B(\bar{x}, \delta) \cap S_{\bar{x}} = \{\bar{x}\}$, which means that $\bar{x}$ is an isolated solution of (OP).

For $u \in Z$, let

$$Y_0(u) := \{y \in Y : \; \phi_y(u) = 0\}.$$

It is clear that if $u \in Z$ and $Y_0(u) \neq \emptyset$ then

$$Y(u) = Y_0(u) \;\text{ and }\; \Phi(u) = 0.$$

For a set Ω, we adopt the following convention

$$[0, \; 1]\Omega = \begin{cases} \{t\omega : \; t \in [0, \; 1] \text{ and } \omega \in \Omega\}, & \text{if } \Omega \neq \emptyset \\ \{0\}, & \text{if } \Omega = \emptyset \end{cases}$$

First we provide a dual sufficient condition for a feasible point to be a weak sharp minimum of optimization problem (OP).

Theorem 4.1 *Suppose that Condition* S *is satisfied and that there exist* $\eta, r \in (0, +\infty)$ *such that*

$$N(S_{\bar{x}}, u) \cap \eta B_{X^*} \subset \partial f(u) + [0, \; 1]\overline{\text{co}}^{w^*} \bigcup_{y \in Y_0(u)} \partial \phi_y(u) + N(A, u) \cap B_{X^*} \tag{4.1}$$

whenever $u \in S_{\bar{x}} \cap B(\bar{x}, r)$*. Then* $\bar{x}$ *is a weak sharp minimum of* (OP).

Proof Let $\varepsilon \in \left(0, \; \frac{\eta}{3}\right)$. By Condition S and Lemma 3.1, there exists $\delta \in (0, \; r)$ such that

$$\langle u_1^*, x - u \rangle \leq f(x) - f(u) + \varepsilon \|x - u\|, \tag{4.2}$$

$$\langle u_2^*, x - u \rangle \leq \phi_y(x) - \phi_y(u) + \varepsilon \|x - u\| \tag{4.3}$$

and

$$\langle u_3^*, x - a \rangle \leq d(x, A) + \varepsilon \|x - a\| \tag{4.4}$$

whenever $x, u \in B(\bar{x}, \delta), a \in A \cap B(\bar{x}, \delta), y \in Y, u_1^* \in \partial f(u), u_2^* \in \partial\phi_y(u)$ and $u_3^* \in N(A, a) \cap B_{X^*}$. Since $\Phi(u) = \phi_y(u)$ for all $y \in Y(u)$, it is easy from (4.3) to verify that

$$\langle u_4^*, x - u\rangle \le \Phi(x) - \Phi(u) + \varepsilon\|x - u\| \tag{4.5}$$

for all $x, u \in B(\bar{x}, \delta)$ and $u_4^* \in \overline{\mathrm{co}}^{w^*} \bigcup_{y\in Y(u)} \partial\phi_y(u)$. Let $x \in B(\bar{x}, \frac{\delta}{2})\backslash S_{\bar{x}}$ and $\gamma \in (\max\{\frac{3\varepsilon}{\eta}, \frac{2d(x,S_{\bar{x}})}{\delta}\},\ 1)$. By Lemma 2.4, there exist $u \in S_{\bar{x}}$ and $u^* \in N(S_{\bar{x}}, u)$ such that $\|u^*\| = 1$ and

$$\gamma\|x - u\| \le \min\{\langle u^*, x - u\rangle, d(x, S_{\bar{x}})\}. \tag{4.6}$$

Hence $\|x - u\| \le \frac{d(x,S_{\bar{x}})}{\gamma} < \frac{\delta}{2}$, and so $\|u - \bar{x}\| \le \|u - x\| + \|x - \bar{x}\| < \delta < r$. It follows from (4.1) that there exist $u_1^* \in \partial f(u), u_2^* \in [0,\ 1]\overline{\mathrm{co}}^{w^*} \bigcup_{y\in Y_0(u)} \partial\phi_y(u)$ and $u_3^* \in N(A, u) \cap B_{X^*}$ such that $\eta u^* = u_1^* + u_2^* + u_3^*$. We divide into two cases: (C1) $\bigcup_{y\in Y_0(u)} \partial\phi_y(u) \ne \emptyset$ and (C2) $\bigcup_{y\in Y_0(u)} \partial\phi_y(u) = \emptyset$. When (C1) holds, $Y_0(u) \ne \emptyset$. Since $u \in S_{\bar{x}} \subset Z$, this implies that $Y_0(u) = Y(u)$. Hence $\Phi(u) = 0$ and there exist $t \in [0,\ 1]$ and $u_4^* \in \overline{\mathrm{co}}^{w^*} \bigcup_{y\in Y(u)} \partial\phi_y(u)$ such that $u_2^* = tu_4^*$ and so $\eta u^* = u_1^* + tu_4^* + u_3^*$. By (4.2) and (4.4)–(4.6), this implies that

$$\gamma\eta\|x - u\| \le f(x) - f(u) + t\Phi(x) + d(x, A) + 3\varepsilon\|x - u\|.$$

Noting that $t\Phi(x) \le \sup_{y\in Y}[\phi_y(x)]_+$ and $f(u) = f(\bar{x})$, it follows that

$$(\gamma\eta - 3\varepsilon)\|x - u\| \le f(x) - f(\bar{x}) + \sup_{y\in Y}[\phi_y(x)]_+ + d(x, A). \tag{4.7}$$

When (C2) holds, $u_2^* = 0$ and $\eta u^* = u_1^* + u_3^*$. It follows from (4.2), (4.4) and (4.6) that

$$\begin{aligned}\gamma\eta\|x - u\| &\le f(x) - f(u) + d(x, A) + 2\varepsilon\|x - u\|\\ &= f(x) - f(\bar{x}) + d(x, A) + 2\varepsilon\|x - u\|\\ &\le f(x) - f(\bar{x}) + \sup_{y\in Y}[\phi_y(x)]_+ + d(x, A) + 2\varepsilon\|x - u\|.\end{aligned}$$

It follows that (4.7) also holds in this case. Since $u \in S_{\bar{x}}$, (4.7) implies that

$$(\gamma\eta - 3\varepsilon)d(x, S_{\bar{x}}) \le f(x) - f(\bar{x}) + \sup_{y\in Y}[\phi_y(x)]_+ + d(x, A).$$

Letting $\gamma \to 1^-$, one has

$$(\eta - 3\varepsilon)d(x, S_{\bar{x}}) \le f(x) - f(\bar{x}) + \sup_{y\in Y}[\phi_y(x)]_+ + d(x, A).$$

Since x is arbitrary in $B\left(a, \frac{\delta}{2}\right)\backslash S_{\bar{x}}$, this implies that $\bar{x}$ is a weak sharp minimum. The proof is completed.

Next we provide a necessary condition for a feasible point to be a weak sharp minimum of (OP).

Theorem 4.2 *Let $\bar{x}$ be a weak sharp minimum of* (OP) *and suppose that Condition* S^+ *is satisfied. Then there exist $\eta, \delta \in (0,\ +\infty)$ such that*

$$\hat{N}(S_{\bar{x}}, u) \cap \eta B_{X^*} \subset \partial f(u) + [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y \in Y_0(u)} \partial \phi_y(u) + N(A, u) \cap B_{X^*} \quad (4.8)$$

whenever $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$.

Proof Thanks to Condition S^+, Theorem 3.1 and Lemma 3.2 can be applied and there exist $L, r \in (0,\ +\infty)$ such that

$$\partial \Phi(u) = \overline{\text{co}}^{w^*} \bigcup_{y \in Y(u)} \partial \phi_y(u) \quad \forall u \in B(\bar{x}, r), \quad (4.9)$$

$$|\phi_y(x_2) - \phi_y(x_1)| \le L\|x_2 - x_1\| \quad \forall x_1, x_2 \in B(\bar{x}, r) \text{ and } y \in Y \quad (4.10)$$

and

$$|\Phi(x_2) - \Phi(x_1)| \le L\|x_2 - x_1\| \quad \forall x_1, x_2 \in B(\bar{x}, r). \quad (4.11)$$

Note further that

$$\partial[\Phi]_+(u) \subset [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y \in Y_0(u)} \partial \phi_y(u) \quad \forall u \in Z \cap B(\bar{x}, r), \quad (4.12)$$

where the function Φ_+ is defined by $[\Phi]_+(x) = \max\{\Phi(x), 0\}$ for all $x \in X$. Indeed, let $u \in Z \cap B(\bar{x}, r)$. Then $\Phi(u) \le 0$. If $\Phi(u) < 0$, then (4.11) implies that $[\Phi]_+$ is identically 0 on some neighborhood of u. Hence $\partial[\Phi]_+(u) = \{0\}$ and so (4.12) holds in this case. Suppose next that $\Phi(u) = 0$. Then $Y_0(u) = Y(u)$ and

$$\partial[\Phi]_+(u) \subset \text{co}\,(\partial\Phi(u) \cup \{0\}) = [0, 1]\partial\Phi(u).$$

Thus (4.9) entails (4.12). Therefore (4.12) is true.

Now by the assumption that $\bar{x}$ is a weak sharp minimum of (OP), there exist $\eta > 0$ and $\delta \in (0,\ r)$ such that (WM) holds. Let $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$ and $u^* \in \hat{N}(S_{\bar{x}}, u) \cap \eta B_{X^*}$. Then $f(u) = f(\bar{x})$ and $u^* \in \eta\hat{\partial} d(\cdot, S_{\bar{x}})(u)$. Hence, for each $n \in \mathbb{N}$ there exists $\delta_n > 0$ such that $B(u, \delta_n) \subset B(\bar{x}, \delta)$ and

$$\langle u^*, x - u \rangle \le \eta d(x, S_{\bar{x}}) + \frac{1}{n}\|x - u\| \quad \forall x \in B(u, \delta_n).$$

Noting that $[\Phi(x)]_+ = \sup_{y \in Y}[\phi_y(x)]_+$ for all $x \in X$, this and (WM) imply that

$$\langle u^*, x - u \rangle \le f(x) - f(u) + [\Phi(x)]_+ + d(x, A) + \frac{1}{n}\|x - u\| \quad \forall x \in B(u, \delta_n).$$

Letting

$$g(x) := -\langle u^*, x-u\rangle + f(x) - f(u) + [\Phi(x)]_+ + d(x, A) + \frac{1}{n}\|x-u\| \quad \forall x \in X,$$

it follows that u is a local minimizer of g. Hence

$$0 \in \partial g(u) \subset -u^* + \partial f(u) + \partial[\Phi]_+(u) + \partial d(\cdot, A)(u) + \frac{1}{n}B_{X^*}.$$

Noting that $\partial d(\cdot, A)(u) \subset N(A, u) \cap B_{X^*}$, this and (4.12) imply that there exist $u_n^* \in \partial f(u)$, $v_n^* \in [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y\in Y_0(u)} \partial\phi_y(u)$ and $w_n^* \in N(A, u) \cap B_{X^*}$ such that

$$\|u_n^* + v_n^* + w_n^* - u^*\| \le \frac{1}{n}.$$

By (4.10), one has $\|v_n^*\| \le L$. Since $\partial f(u)$, $[0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y\in Y_0(u)} \partial\phi_y(u)$ and $N(A, u)$ are weak*-closed and B_{X^*} is weak*-compact, we can assume without loss of generality that

$$v_n^* \stackrel{w^*}{\to} v^* \in [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y\in Y_0(u)} \partial_x\phi(u, y) \ \text{ and } \ w_n^* \stackrel{w^*}{\to} w^* \in N(A, u) \cap B_{X^*}$$

and so $u_n^* \stackrel{w^*}{\to} u^* - v^* - w^* \in \partial f(u)$. It follows that

$$u^* \in \partial f(u) + [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y\in Y_0(u)} \partial\phi_y(u) + N(A, u) \cap B_{X^*}.$$

This shows that (4.8) holds. The proof is completed.

Next we provide a characterization for a sharp minimum of (OP).

Theorem 4.3 *Suppose that Condition* S^+ *is satisfied. Then* $\bar{x}$ *is a sharp minimum of* (OP) *if and only if there exists* $\eta \in (0,\ +\infty)$ *such that*

$$\eta B_{X^*} \subset \partial f(\bar{x}) + [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y\in Y_0(\bar{x})} \partial\phi_y(\bar{x}) + N(A, \bar{x}) \cap B_{X^*}. \tag{4.13}$$

Proof Suppose that $\bar{x}$ is a sharp minimum. Then there exists $\delta > 0$ such that $S_{\bar{x}} \cap B(\bar{x}, \delta) = \{\bar{x}\}$ and so $\hat{N}(S_{\bar{x}}, \bar{x}) = X^*$. Thus the necessity part is clear by Theorem 4.2. For the sufficiency part, by Theorem 4.1, we only need to show that (4.13) implies that $S_{\bar{x}} \cap B(\bar{x}, r) = \{\bar{x}\}$ for some $r > 0$. Suppose to the contrary that there exists a sequence $\{x_n\}$ in $S_{\bar{x}}\backslash\{\bar{x}\}$ such that $x_n \to \bar{x}$. Take $x_n^* \in \eta B_{X^*}$ such that

$$\langle x_n^*, x_n - \bar{x}\rangle = \eta\|x_n - \bar{x}\|. \tag{4.14}$$

By (4.13), there exist $u_n^* \in \partial f(\bar{x})$, $v_n^* \in [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y \in Y_0(\bar{x})} \partial \phi_y(\bar{x})$ and $w_n^* \in N(A, \bar{x}) \cap B_{X^*}$ such that

$$x_n^* = u_n^* + v_n^* + w_n^*. \tag{4.15}$$

Since f and A are subsmooth at $\bar{x}$, there exists $\delta > 0$ such that

$$\langle x^*, x - \bar{x}\rangle \leq f(x) - f(\bar{x}) + \frac{\eta\|x - \bar{x}\|}{6} \quad \forall x \in B(\bar{x}, \delta) \text{ and } x^* \in \partial f(\bar{x}) \tag{4.16}$$

and

$$\langle x^*, x - \bar{x}\rangle \leq \frac{\eta\|x - \bar{x}\|}{6} \quad \forall x \in A \cap B(\bar{x}, \delta) \quad \text{and} \quad x^* \in N(A, \bar{x}) \cap B_{X^*}. \tag{4.17}$$

When $Y_0(\bar{x}) \neq \emptyset$, we have $Y_0(\bar{x}) = Y(\bar{x})$ and $\Phi(\bar{x}) = 0$; thus, taking a smaller δ if necessary, it is easy from Theorem 3.1 to verify that

$$\langle x^*, x - \bar{x}\rangle \leq [\Phi(x)]_+ + \frac{\eta\|x - \bar{x}\|}{6} \quad \forall x \in B(\bar{x}, \delta) \quad \text{and} \quad x^* \in [0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y \in Y_0(\bar{x})} \partial \phi_y(\bar{x})$$

(when $Y_0(\bar{x}) = \emptyset$ this inequality trivially holds because $[0,\ 1]\overline{\text{co}}^{w^*} \bigcup_{y \in Y_0(\bar{x})} \partial \phi_y(\bar{x}) = \{0\}$ in this case). Noting that $S_{\bar{x}} \subset Z \subset A$, $f(x) = f(\bar{x})$ for all $x \in S_{\bar{x}}$ and $x_n \to \bar{x}$, it follows from (4.15)–(4.17) that $\langle x_n^*, x_n - \bar{x}\rangle \leq \frac{\eta}{2}\|x_n - \bar{x}\|$ for all sufficiently large n. This contradicts (4.14). The proof is completed.

The following corollary is immediate from Proposition 3.3 and Theorems 4.1–4.3.

Corollary 4.1 *Let W be a Banach space, $\psi : W \times Y \to \mathbb{R}$ be a continuous function such that the function $z \mapsto \psi(z, y)$ is convex for each $y \in Y$ and let $g : X \to W$ be a smooth function. Let*

$$\phi_y(x) = \psi(g(x), y) \quad \forall (x, y) \in X \times Y$$

and consider the following statements:

(i) *$\bar{x}$ is a sharp minimum of* (OP).
(ii) *There exist $\eta, \delta \in (0,\ +\infty)$ such that* (4.13) *holds.*
(iii) *$\bar{x}$ is a weak sharp minimum of* (OP).
(iv) *There exist $\eta, \delta \in (0,\ +\infty)$ such that* (4.1) *holds.*
(v) *There exist $\eta, \delta \in (0,\ +\infty)$ such that* (4.8) *holds.*

Then $(i) \Leftrightarrow (ii)$ and $(iv) \Rightarrow (iii) \Rightarrow (v)$.

5 Subsmooth semi-infinite optimization problem

In this section, we assume that X is a finite dimensional Euclidean space, and the corresponding (OP) is to be referred as a generalized semi-infinite optimization problem ((GSOP) in brief). In the remainder, let $\dim(X)$ denote the dimension of X and

$$m := \dim(X) + 1.$$

Recall that a function $g : X \to \mathbb{R}$ is directionally differentiable at $\bar{x} \in X$ in $h \in X$ if the limit

$$d^+g(\bar{x}, h) := \lim_{t\to 0^+} \frac{g(\bar{x} + th) - g(\bar{x})}{t}$$

exists.

We need the following lemma.

Lemma 5.1 *Suppose that f is subsmooth at $\bar{x}$ and locally Lipschitz at $\bar{x}$. Then f is directionally differentiable at $\bar{x}$ in each $h \in X$ and*

$$d^+f(\bar{x}, h) = f^\circ(\bar{x}, h) \quad \forall h \in X. \tag{5.1}$$

Proof From the subsmoothness, it is easy to verify that

$$\langle x^*, h\rangle \le \liminf_{t\to 0^+} \frac{f(\bar{x} + th) - f(\bar{x})}{t} \quad \forall (x^*, h) \in \partial f(\bar{x}) \times X. \tag{5.2}$$

Since f is locally Lipschitz,

$$f^\circ(\bar{x}, h) = \max\{\langle x^*, h\rangle : x^* \in \partial f(\bar{x})\} \text{ and } \limsup_{t\to 0^+} \frac{f(\bar{x} + th) - f(\bar{x})}{t} \le f^\circ(\bar{x}, h)$$

for all $h \in X$. Thus the result is clear.

Lemma 5.2 *Suppose that f and $\{\phi_y : y \in Y\}$ are subsmooth around $\bar{x}$. Further suppose that f and $\{\phi_y : y \in Y\}$ are locally Lipschitz at $\bar{x}$. Then there exists $\delta > 0$ such that*

$$d^+f(u, h) = 0 \quad \text{and} \quad d^+\phi_y(u, h) \le 0 \tag{5.3}$$

for all $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$, $h \in T(S_{\bar{x}}, u)$ and all $y \in Y_0(u)$.

Proof By the assumptions, there exist $L, r \in (0, +\infty)$ satisfying (4.10) such that f and ϕ are subsmooth at each $u \in S_{\bar{x}} \cap B(\bar{x}, r)$, and

$$|f(x_1) - f(x_2)| \le L\|x_1 - x_2\| \quad \forall x_1, x_2 \in B(\bar{x}, r). \tag{5.4}$$

Let $u \in S_{\bar{x}} \cap B(\bar{x}, r)$ and $h \in T(S_{\bar{x}}, u)$ and $y \in Y_0(u)$. Then $\phi_y(u) = 0$ and there exist $t_n \to 0^+$ and $h_n \to h$ such that $u + t_n h_n \in S_{\bar{x}}$ for all $n \in \mathbb{N}$. Hence

$$f(u + t_n h_n) = f(u) = f(\bar{x}) \text{ and } \phi_y(u + t_n h_n) \leq 0 \quad \forall n \in \mathbb{N}.$$

It follows from (5.4) and (4.10) that

$$|f(u + t_n h) - f(u)| \leq L t_n \|h_n - h\|$$

and

$$\phi_y(u + t_n h) - \phi_y(u) \leq \phi_y(u + t_n h) - \phi_y(u + t_n h_n) \leq L t_n \|h_n - h\|$$

for all sufficiently large n. This and Lemma 5.1 imply that (5.3) holds. The proof is completed.

We first provide necessity conditions.

Theorem 5.1 *Let $\bar{x}$ be a local solution of (GSOP). Suppose that $\{\phi_y : y \in Y\}$ is subsmooth at $\bar{x}$ and that f and $\{\phi_y : y \in Y\}$ are locally Lipschitz at $\bar{x}$. Then there exist $y_1, \ldots, y_m \in Y$ and $\lambda_0, \lambda_1, \ldots, \lambda_m \in \mathbb{R}_+$ such that*

$$\sum_{i=0}^{m} \lambda_i = 1, \; \lambda_i \phi_{y_i}(\bar{x}) = 0 \; (1 \leq i \leq m) \tag{5.5}$$

and

$$0 \in \lambda_0 \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial \phi_{y_i}(\bar{x}) + N(A, \bar{x}). \tag{5.6}$$

If, in addition, there exists $h \in T(A, \bar{x})$ such that

$$d^+ \phi_y(\bar{x}, h) < 0 \quad \forall y \in Y_0(\bar{x}), \tag{5.7}$$

then there exist $y_1, \ldots, y_m \in Y$ and $\lambda_1, \ldots, \lambda_m \in \mathbb{R}_+$ such that

$$\lambda_i \phi_{y_i}(\bar{x}) = 0 \; (1 \leq i \leq m)$$

and

$$0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial \phi_{y_i}(\bar{x}) + N(A, \bar{x}).$$

Proof By Lemma 3.2, Φ is locally Lipschitz at $\bar{x}$. Since $\bar{x}$ is a local solution of (GSOP), it is easy to verify that $\bar{x}$ is a local solution of the following optimization problem:

$$\min f(x) \ \text{ subject to } \Phi(x) \le 0 \text{ and } x \in A.$$

It follows from [6, Theorem 6.1.1] that there exist $\lambda_0, \bar{\lambda} \in \mathbb{R}_+$ such that

$$\lambda_0 + \bar{\lambda} = 1,\ \bar{\lambda}\Phi(\bar{x}) = 0 \ \text{ and } \ 0 \in \lambda_0 \partial f(\bar{x}) + \bar{\lambda}\partial\Phi(\bar{x}) + N(A, \bar{x}). \tag{5.8}$$

We assume that $\bar{\lambda} \neq 0$ (otherwise the conclusion trivially holds). Thus, $\Phi(\bar{x}) = 0$ and so $Y_0(\bar{x}) = Y(\bar{x})$. By Theorem 3.2, one has

$$\partial\Phi(\bar{x}) = \text{co}\left(\bigcup_{y \in Y_0(\bar{x})} \partial\phi_y(\bar{x})\right).$$

It follows from (5.8) and the Carathéodory theorem that there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}_+$ and $y_1, \ldots, y_m \in Y$ such that (5.5) and (5.6) hold. Finally we consider the case when there exists $h \in T(A, \bar{x})$ such that (5.7) holds. We only need to show that $\lambda_0 \neq 0$ (the result is then clear as λ_i's can be replaced by suitable multiples if necessary). Suppose to the contrary that $\lambda_0 = 0$. Then (5.6) reduces to

$$0 \in \sum_{i=1}^{m} \lambda_i \partial\phi_{y_i}(\bar{x}) + N(A, \bar{x})$$

and so there exist $x_i^* \in \partial\phi_{y_i}(\bar{x})$ $(i = 1, \ldots, m)$ such that $-\sum_{i=1}^m \lambda_i x_i^* \in N(A, \bar{x})$. It follows that $\sum_{i=1}^m \lambda_i \phi_{y_i}^\circ(\bar{x}, h) \ge \sum_{i=1}^m \lambda_i \langle x_i^*, h\rangle \ge 0$. This and Lemma 5.1 imply that $\sum_{i=1}^m \lambda_i d^+\phi_{y_i}(\bar{x}, h) \ge 0$. Since $\sum_{i=1}^m \lambda_i = 1$, this contradicts (5.7). The proof is completed.

In the line of Theorem 5.2, the following theorems establishes a dual characterization for a sharp minimum of (GSOP) and is immediate from Theorems 3.1 and 4.3 together with the Carathéodory theorem.

Theorem 5.2 *Suppose that Condition* S^+ *is satisfied. Then* $\bar{x}$ *is a sharp minimum of (GSOP) if and only if there exists* $\eta > 0$ *such that for each* $x^* \in \eta B_{X^*}$ *there exist* $y_1, \ldots, y_m \in Y_0(\bar{x})$ *and* $\lambda_1, \ldots, \lambda_m \in \mathbb{R}_+$ *satisfying*

$$\sum_{i=1}^{m} \lambda_i \le 1 \ \text{ and } \ x^* \in \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial\phi_{y_i}(\bar{x}) + N(A, \bar{x}) \cap B_{X^*}.$$

Next we provide dual characterizations for a feasible point to be a weak sharp minimum of (GSOP).

Theorem 5.3 *Suppose that* f, $\{\phi_y :\ y \in Y\}$ *and* A *are subsmooth around* $\bar{x}$ *and that* $\{\phi_y :\ y \in Y\}$ *is locally Lipsctitz at* $\bar{x}$. *Then the following statements are equivalent.*

(i) *$\bar{x}$ is a weak sharp minimum of (GSOP).*

(ii) *There exist $\eta, r \in (0, +\infty)$ such that for each $u \in S_{\bar{x}} \cap B(\bar{x}, r)$ and each $u^* \in \hat{N}(S_{\bar{x}}, u) \cap \eta B_{X^*}$ there exist $y_1, \ldots, y_m \in Y_0(u)$ and $\lambda_1, \ldots, \lambda_m \in \mathbb{R}_+$ satisfying*

$$\sum_{i=1}^{m} \lambda_i \leq 1 \text{ and } u^* \in \partial f(u) + \sum_{i=1}^{m} \lambda_i \partial \phi_{y_i}(u) + N(A, u) \cap B_{X^*}. \quad (5.9)$$

(iii) *Same as (ii) but $\hat{N}(S_{\bar{x}}, u)$ is replaced by $N_M(S_{\bar{x}}, u)$.*

(iv) *Same as (ii) but $\hat{N}(S_{\bar{x}}, u)$ is replaced by $N(S_{\bar{x}}, u)$.*

Proof Thanks to the assumption and by Lemma 3.2, we take $L, r \in (0, +\infty)$ satisfying (4.10) and (4.11) such that f, $\{\phi_y : y \in Y\}$ and A are subsmooth at each $u \in B(\bar{x}, r)$. By Theorem 3.2, we have

$$\partial \Phi(u) = \text{co}\left(\bigcup_{y \in Y(u)} \partial \phi_y(u)\right) \quad \forall u \in B(\bar{x}, r). \quad (5.10)$$

Thus, by Theorems 4.1 and 4.2 together with the Carathéodory theorem, we have $(i) \Rightarrow (ii)$ and $(iv) \Rightarrow (i)$. It remains to show $(ii) \Rightarrow (iii)$ and $(iii) \Rightarrow (iv)$.

$(ii) \Rightarrow (iii)$ By (ii) we can assume without loss of generality that the above r together with some $\eta > 0$ has the property stated as in (ii). Let $u \in S_{\bar{x}} \cap B(\bar{x}, r)$ and $u^* \in N_M(S_{\bar{x}}, u) \cap \eta B_{X^*}$. Then there exist sequences $\{u_n\}$ in $S_{\bar{x}} \cap B(\bar{x}, r)$ and $\{u_n^*\}$ such that

$$u_n \to u, \; u_n^* \to u^* \text{ and } u_n^* \in \hat{N}(S_{\bar{x}}, u_n) \cap \eta B_{X^*} \;\; \forall n \in \mathbb{N}.$$

By (ii), for each $n \in \mathbb{N}$ there exist $x_n^* \in \partial f(u)$, $y_i(n) \in Y_0(u_n)$, $y_i^*(n) \in \partial \phi_{y_i(n)}(u_n)$, $\lambda_i(n) \in \mathbb{R}_+$ $(i = 1, \ldots, m)$ and $z_n^* \in N(A, u_n) \cap B_{X^*}$ such that

$$\sum_{i=1}^{m} \lambda_i(n) \leq 1 \text{ and } u_n^* = x_n^* + \sum_{i=1}^{m} \lambda_i(n) y_i^*(n) + z_n^*. \quad (5.11)$$

This and (4.10) imply that $\{x_n^*\}$ and $\{y_i^*(n)\}$ are bounded. By the compactness of Y, we assume without loss of generality that

$$x_n^* \to x^*, \; \lambda_i(n) \to \lambda_i, \; y_i^*(n) \to y_i^*, \; y_i(n) \to y_i \text{ and } z_n^* \to z^* \text{ as } n \to \infty.$$

It follows from (5.11) and the continuity of the function $(x, y) \mapsto \phi_y(x)$ that

$$\sum_{i=1}^{m} \lambda_i \leq 1, \; u^* = x^* + \sum_{i=1}^{m} \lambda_i y_i^* + z^* \text{ and } y_i \in Y_0(u) \; (1 \leq i \leq m).$$

Thus, to prove (iii), we only need to show that

$$x^* \in \partial f(u),\ y_i^* \in \partial \phi_{y_i}(u)\ (1 \le i \le m) \text{ and } z^* \in N(A, u). \tag{5.12}$$

Let $\varepsilon > 0$. By the subsmoothness, there exists $\delta > 0$ such that

$$\langle x_n^*, x - u_n\rangle \le f(x) - f(u_n) + \varepsilon\|x - u_n\|,\quad \langle z_n^*, z - u_n\rangle \le \varepsilon\|z - u_n\|$$

and

$$\langle y_i^*(n), x - u_n\rangle \le \phi_{y_i(n)}(x) - \phi_{y_i(n)}(u_n) + \varepsilon\|x - u_n\|\ \ (1 \le i \le m)$$

for any $x \in B(u, \delta)$, $z \in A \cap B(u, \delta)$ and all sufficiently large n. It follows that

$$\langle x^*, x - u\rangle \le f(x) - f(u) + \varepsilon\|x - u\|,\quad \langle z^*, z - u\rangle \le \varepsilon\|z - u\|$$

and

$$\langle y_i^*, x - u\rangle \le \phi_{y_i}(x) - \phi_{y_i}(u) + \varepsilon\|x - u\|\ \ (1 \le i \le m)$$

for any $x \in B(u, \delta)$ and $z \in A \cap B(u, \delta)$. This implies that (5.12) holds and so does (iii).

(iii)$\Rightarrow$ (iv) Let $u \in S_{\bar{x}} \cap B(\bar{x}, r)$. Since

$$[0,\ 1]\mathrm{co} \bigcup_{y\in Y_0(u)} \partial\phi_y(u) = \begin{cases} \{0\} & \text{if } Y_0(u) = \emptyset \\ [0,\ 1]\partial\Phi(u) & \text{if } Y_0(u) \ne \emptyset \end{cases}$$

and $\partial\Phi(u)$ is weak*-compact (by (4.11)), $[0,\ 1]\mathrm{co}\bigcup_{y\in Y_0(u)} \partial\phi_y(u)$ is weak*-compact. Noting that $\partial f(u)$ is a weak*-closed convex set and $N(A, u) \cap B_{X^*}$ is a weak*-compact convex set, it follows that $\partial f(u) + [0,\ 1]\mathrm{co}\bigcup_{y\in Y_0(u)} \partial\phi_y(u) + N(A, u) \cap B_{X^*}$ is weak*-closed and convex. Since (iii) means

$$N_M(S_{\bar{x}}, u) \cap \eta B_{X^*} \subset \partial f(u) + [0,\ 1]\mathrm{co} \bigcup_{y\in Y_0(u)} \partial\phi_y(u) + N(A, u) \cap B_{X^*},$$

it follows that

$$\overline{\mathrm{co}}^{w^*}(N_M(S_{\bar{x}}, u) \cap \eta B_{X^*}) \subset \partial f(u) + [0,\ 1]\mathrm{co} \bigcup_{y\in Y_0(u)} \partial\phi_y(u) + N(A, u) \cap B_{X^*}.$$

Since every finite dimensional space is an Asplund space, (2.1) implies that

$$N(S_{\bar{x}}, u) \cap \eta B_{X^*} = \overline{\mathrm{co}}^{w^*}(N_M(S_{\bar{x}}, u) \cap \eta B_{X^*}).$$

Hence

$$N(S_{\bar{x}}, u) \cap \eta B_{X^*} \subset \partial f(u) + [0,\ 1]\mathrm{co} \bigcup_{y \in Y_0(u)} \partial \phi_y(u) + N(A, u) \cap B_{X^*}.$$

By the Carathéodory theorem, one can see that (iv) holds. The proof is completed.

Next we provide primal characterizations for $\bar{x}$ to be a local weak sharp minimum of (GSOP). In what follows, for $u \in S_{\bar{x}}$ and $h \in X$, let us adopt the convention that

$$\sup_{y \in Y_0(u)} [d^+\phi_y(u, h)]_+ := 0 \quad \text{if } Y_0(u) = \emptyset.$$

For a closed subset Ω of X and $x \in X$, let $P_\Omega(x)$ denote the set of all projections of x to Ω, that is,

$$P_\Omega(x) := \{\omega \in \Omega :\ \|x - \omega\| = d(x, \Omega)\}.$$

To establish primal characterization, we need the following lemma, which should be known. Since we cannot find a reference on this lemma, we provide its proof for completeness.

Lemma 5.3 *Let K be a closed convex cone of a Banach space X and $x \in X \backslash K$. Then*

$$d(x, K) = \max\{\langle x^*, x\rangle :\ x^* \in N(K, 0) \cap B_{X^*}\}.$$

Proof Let $r := d(x, K)$. Then, $B(x, r) \cap K = \emptyset$ and it follows from the separation theorem that there exists $x^* \in X^*$ with $\|x^*\| = 1$ such that

$$\langle x^*, x\rangle - r = \inf\{\langle x^*, u\rangle :\ u \in B(x, r)\} \geq \sup\{\langle x^*, u\rangle :\ u \in K\}.$$

Since K is a cone, this implies that $\sup\{\langle x^*, u\rangle :\ u \in K\} = 0$. Hence $\langle x^*, x\rangle \geq r$ and $x^* \in N(K, 0) \cap B_{X^*}$. We need only show that

$$\max\{\langle x^*, x\rangle :\ x^* \in N(K, 0) \cap B_{X^*}\} \leq d(x, K). \tag{5.13}$$

Let $x^* \in N(K, 0) \cap B_{X^*}$ and $u \in K$. Then $\langle x^*, x\rangle \leq \langle x^*, x - u\rangle \leq \|x - u\|$. It follows that (5.13) holds. The proof is completed.

Theorem 5.4 *Let f, $\{\phi_y :\ y \in Y\}$ and A be as in Theorem 5.3 and further suppose that f is locally Lipschitz at $\bar{x}$. Then the following statements are equivalent.*

(i) *$\bar{x}$ is a local weak sharp minimum of (GSOP).*
(ii) *There exist $\eta, \gamma \in (0,\ +\infty)$ such that*

$$\eta d(h, T_c(S_{\bar{x}}, u)) \leq d^+ f(u, h) + \sup_{y \in Y_0(u)} [d^+\phi_y(u, h)]_+ + d(h, T(A, u)) \tag{5.14}$$

for all $u \in S_{\bar{x}} \cap B(\bar{x}, \gamma)$ and $h \in X$.

(iii) *Same as (ii) but* $T_c(S_{\bar{x}}, u)$ *is replaced by* $T(S_{\bar{x}}, u)$.
(iv) *There exist* $\eta, \gamma \in (0, +\infty)$ *such that*

$$\eta\|x-u\| \leq d^+ f(u, x-u) + \sup_{y\in Y_0(u)} [d^+\phi_y(u, x-u)]_+ + d(x-u, T(A, u)) \tag{5.15}$$

for any $x \in B(\bar{x}, \gamma)$ *and* $u \in P_{S_{\bar{x}}}(x)$.

Proof Take $L, r \in (0, +\infty)$ satisfying (4.10) and (5.4) such that f, $\{\phi_y : y \in Y\}$ and A are subsmooth at each $u \in S_{\bar{x}} \cap B(\bar{x}, r)$. Hence A is regular at each $u \in S_{\bar{x}} \cap B(\bar{x}, r)$ in the Clarke sense, namely

$$T(A, u) = T_c(A, u) \quad \forall u \in S_{\bar{x}} \cap B(\bar{x}, r). \tag{5.16}$$

$(i) \Rightarrow (ii)$. Suppose that (i) holds. Then, by Theorem 5.3 there exist $\eta > 0$ and $\gamma \in (0, r)$ such that (iv) of Theorem 5.3 holds. Let $u \in S_{\bar{x}} \cap B(\bar{x}, \gamma)$ and $h \in X$. By Lemma 5.2, one sees that (5.14) holds if $h \in T_c(S_{\bar{x}}, u)$. Now we assume that $h \notin T_c(S_{\bar{x}}, u)$. Since $T_c(S_{\bar{x}}, u)$ is a closed and convex cone, the projection theorem implies that there exists

$$h_0 \in P_{T_c(S_{\bar{x}}, u)}(h) \text{ and } \langle h - h_0, z - h_0\rangle \leq 0 \quad \forall z \in T_c(S_{\bar{x}}, u).$$

It follows that

$$\langle h - h_0, h_0\rangle = 0 \text{ and } \langle h - h_0, z\rangle \leq 0 \quad \forall z \in T_c(S_{\bar{x}}, u),$$

and so $\frac{\eta(h-h_0)}{\|h-h_0\|} \in N(S_{\bar{x}}, u) \cap \eta B_{X^*}$. Thus, by (iv) of Theorem 5.3, there exist $\lambda_i \in \mathbb{R}_+$ and $y_i \in Y_0(u)$ $(1 \leq i \leq m)$ such that

$$\frac{\eta(h-h_0)}{\|h-h_0\|} \in \partial f(u) + \sum_{i=1}^{m} \lambda_i \partial\phi_{y_i}(u) + N(A, u) \cap B_{X^*} \quad \text{and} \quad \sum_{i=1}^{m} \lambda_i \leq 1.$$

Noting that

$$f^\circ(u, h) = \max_{x^*\in\partial f(u)} \langle x^*, h\rangle, \quad \phi_{y_i}^\circ(u, h) = \max_{x^*\in\partial\phi_{y_i}(u)} \langle x^*, h\rangle,$$
$$d(h, T(A, u)) = d(h, T_c(A, u)) \text{ (by(5.16))} \quad \text{and} \quad N(A, u) = N(T_c(A, u), 0),$$

it follows from Lemmas 5.1 and 5.3 that

$$\begin{aligned}\eta d(h, T_c(S_{\bar{x}}, u)) &= \left\langle \frac{\eta(h-h_0)}{\|h-h_0\|}, h-h_0 \right\rangle \\ &= \left\langle \frac{\eta(h-h_0)}{\|h-h_0\|}, h \right\rangle \\ &\le d^+ f(u,h) + \sum_{i=1}^{m} \lambda_i d^+ \phi_{y_i}(u,h) + d(h, T(A,u)) \\ &\le d^+ f(u,h) + \sup_{y\in Y_0(u)} [d^+\phi_y(u,h)]_+ + d(h, T(A,u)).\end{aligned}$$

This shows that (ii) holds.

Since $T_c(S_{\bar{x}}, u) \subset T(S_{\bar{x}}, u)$ for any $u \in S_{\bar{x}}$, the implication $(ii) \Rightarrow (iii)$ is trivial. Let $x \in B(\bar{x}, \frac{\gamma}{2})\backslash S_{\bar{x}}$ and take $u \in P_{S_{\bar{x}}}(x)$. Then $u \in S_{\bar{x}} \cap B(\bar{x}, \gamma)$ and $\frac{x-u}{\|x-u\|} \in \hat{N}(S_{\bar{x}}, u)$ (cf. [29, Example 6.16]). This and [29, Proposition 6.5] imply that

$$\left\langle \frac{x-u}{\|x-u\|}, z \right\rangle \le 0 \quad \forall z \in T(S_{\bar{x}}, u).$$

It follows that

$$\|x-u\| \le \left\langle \frac{x-u}{\|x-u\|}, x-u-z \right\rangle \le \|x-u-z\| \quad \forall z \in T(S_{\bar{x}}, u).$$

Hence $\|x-u\| = d(x-u, T(S_{\bar{x}}, u))$. By (iii) (applied to $h = x-u$), one has that (5.15) holds. This shows that $(iii) \Rightarrow (iv)$ holds.

Suppose that (iv) holds with $\eta > 0$ and $\gamma \in (0,\ r)$. Let $\varepsilon \in (0,\ \frac{\eta}{3})$. By the subsmoothness, it is easy from Lemmas 5.1 and 3.1 to verify that there exists $\delta \in (0,\ \gamma)$ such that

$$d^+ f(u, x-u) = \max_{x^*\in\partial f(u)} \langle x^*, x-u\rangle \le f(x) - f(\bar{x}) + \varepsilon\|x-u\|,$$

$$d^+\phi_y(u, x-u) \le \phi_y(x) + \varepsilon\|x-u\|$$

and

$$\begin{aligned}d(x-u, T(A,u)) &= d(x-u, T_c(A,u)) \\ &= \max_{x^*\in N(A,u)\cap B_{X^*}} \langle x^*, x-u\rangle \le d(x,A) + \varepsilon\|x-u\|\end{aligned}$$

for all $x \in B(\bar{x}, \delta)$, $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$ and $y \in Y_0(u)$. Let $x \in B(\bar{x}, \frac{\delta}{2})\backslash S_{\bar{x}}$ and take $u \in P_{S_{\bar{x}}}(x)$. Then $u \in B(\bar{x}, \delta)$. Hence (5.15) holds for such x and u. It follows from the earlier estimates that

$$\eta\|x-u\| \le f(x) - f(\bar{x}) + \sup_{y\in Y}[\phi_y(x)]_+ + d(x,A) + 3\varepsilon\|x-u\|,$$

that is,

$$(\eta - 3\varepsilon)d(x, S_{\bar{x}}) = (\eta - 3\varepsilon)\|x - u\| \leq f(x) - f(\bar{x}) + \sup_{y \in Y}[\phi_y(x)]_+ + d(x, A).$$

This shows that (i) holds. The proof is completed.

Acknowledgments We thank the referees for their helpful comments and for [13,21].

References

1. Auslander, A., Goberna, M., López, M.A.: Penalty and smoothning methods for convex semi-infinite programming. Math. Oper. Res. **34**, 303–319 (2009)
2. Aussel, D., Daniilidis, A., Thibault, L.: Subsmooth sets: functional characterizations and related concepts. Trans. Am. Math. Soc. **357**, 1275–1301 (2005)
3. Brosowski, B.: Parametric Semi-Infinite Optimization. Verlag Peter Lang, Frankfurt (1982)
4. Burke, J.V., Deng, S.: Weak sharp minima revisited. Part I: basic theory. Control Cybern. **31**, 439–469 (2002)
5. Burke, J.V., Ferris, M.C.: Weak sharp minima in mathematical programming. SIAM J. Control Optim. **31**, 1340–1359 (1993)
6. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
7. Clarke, F.H., Stern, R., Wolenski, P.: Proximal smoothness and the lower-C^2 property. J. Convex Anal. **2**, 117–144 (1995)
8. Clarke, F.H., Ledyaev, Y., Stern, R., Wolenski, P.: Nonsmmoth Analysis and Control Theory. Springer, New York (1998)
9. Cromme, L.: Strong uniqueness, a far-reaching criterion for the convergence analysis of iterative procedures. Numer. Math. **29**, 179–193 (1978)
10. Ferris, M.C.: Weak Sharp Minima and Penalty Fucntions in Mathematical Programming. Ph.D. Thesis, University of Cambridge, Cambridge (1988)
11. Goberna, A., López, M.A.: Linear Semi-Infinite Optimization. Wiley, Chichster (1998)
12. Goberna, A., López, M.A. (eds.): Semi-Infinite Programming—Recent Advances. Kluwer, Boston (2001)
13. Hantoute, A., López, M.A., Zalinescu, C.: Subdifferential calculus rules in convex analysis: a unifying approach via pointwise supremum functions. SIAM J. Optim. **19**, 863–882 (2008)
14. Henrion, R., Outrata, J.: Calmness of constraint systems with applications. Math. Program. **104**, 437–464 (2005)
15. Hettich, R., Kortanek, K.O.: Semi-infinite programming: theory, methods, and applications. SIAM Rev. **35**, 380–429 (1993)
16. Jittorntrum, K., Osborne, M.R.: Strong uniqueness and second order convergence in nonlinear discrete approximation. Numer. Math. **34**, 439–455 (1980)
17. Jongen, H.T., Ruuckmann, J.-J., Stein, O.: Generalized semi-infinite optimization: a first order optimality condition and examples. Math. Program. **83**, 145–158 (1998)
18. Klatte, D., Henrion, R.: Regularity and stability in nonlinear semi-infinite optimization. Semi Infin Program Nonconvex Optim. Appl. **25**, 69–102 (1998)
19. Lewis, A., Pang, J.S.: Error bounds for convex inequality systems. In: Crouzeix, J.-P., Martinez-Legaz, J.-E., Volle, M. (eds.) Generalized Convexity, Generalized Monotonicity: Recent Results, Proceedings of the 5th Symposium on Generalized Convexity, Luminy, 1996, pp. 75-0110. Kluwer, Dordrecht (1997)
20. López, M.A., Vercher, E.: Optimality conditions for nondifferentiable convex semi-infinite programming. Math. Program. **27**, 307–319 (1983)
21. López, M.A., Volle, M.: A formula for the set of optimal solutions of a relaxed minimization problem, applications to subdifferential calculus. J. Convx Anal. **17**, 1057–1075 (2010)
22. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, vol. I/II, Basic Theory. Springer, Berlin (2006)

23. Nurnberger, G.: Global unicity in semi-infinite optimization. Numer. Funct. Anal. Optim. **8**, 173–191 (1985)
24. Osborne, M.R., Womersley, R.S.: Strong uniqueness in sequential linear programming. J. Austral. Math. Soc. Ser. B **31**, 379–384 (1990)
25. Pang, J.S.: Error bounds in mathematical programming. Math. Program. **79**, 299–332 (1997)
26. Polak, E.: Optimization. Springer, New York (1997)
27. Poliquin, R., Rockafellar, R.T.: Prox-regular functions in variational analysis. Trans. Am. Math. Soc. **348**(5), 1805–1838 (1996)
28. Reemtsen, R., Ruckmann, J.-J. (eds.): Semi-Infinite Programming. Kluwer, Boston (1998)
29. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
30. Stein, O.: Bi-Level Strategies in Semi-Infinite Programming. Kluwer, Boston (2003)
31. Studniarski, M., Ward, D.E.: Weak sharp minima: characterizations and sufficient conditions. SIAM J. Control Optim. **38**, 219–236 (1999)
32. Tapia, R.A., Trosset, M.W.: An extension of the Karush–Kuhn–Tucker necessity conditions to infinite programming. SIAM Rev. **36**, 1–17 (1994)
33. Vazquez, F.G., Ruckmann, J.J.: Extensions of the Kuhn–Tucker constraint qualification to generalized semi-infinite programming. SIAM J. Optim. **15**, 926–937 (2005)
34. Zalinescu, C.: Weak sharp minima, well-behaving functions and global error bounds for convex inequalities in Banach spaces. In: Proceedings of 12th Baikal International Conference on Optimization Methods and Their Applications, Irkutsk, pp. 272–284 (2001)
35. Zalinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)
36. Zheng, X.Y., Ng, K.F.: Metric regularity and constraint qualifications for convex inequalities on Banach spaces. SIAM J. Optim. **14**, 757–772 (2003)
37. Zheng, X.Y., Ng, K.F.: Linear regularity for a collection of subsmooth sets in Banach spaces. SIAM J. Optim. **19**, 62–76 (2008)
38. Zheng, X.Y., Yang, X.Q.: Lagrange multipliers in nonsmooth semi-infinite optimization problems. Math. Oper. Res. **32**, 168–181 (2007)
39. Zheng, X.Y., Yang, X.Q.: Weak sharp minima for semi-infinite optimization problems with applications. SIAM J. Optim. **18**, 573–588 (2007)
40. Zheng, X.Y., Yang, X.Q.: Global weak sharp minima for convex (semi-)infinite optimization problems. J. Math. Anal. Appl. **248**, 1021–1028 (2008)