

3-1-2011

# Complete Characterizations of Local Weak Sharp Minima With Applications to Semi-Infinite Optimization and Complementarity

Boris S. Mordukhovich  
*Wayne State University, boris@math.wayne.edu*

Naihua Xiu  
*Beijing Jiaotong University, China, nhxiu@bjtu.edu.cn*

Jinchuan Zhou  
*Shandong University of Technology, Zibo, China, jinchuanzhou@yahoo.com.cn*

---

## Recommended Citation

Mordukhovich, Boris S.; Xiu, Naihua; and Zhou, Jinchuan, "Complete Characterizations of Local Weak Sharp Minima With Applications to Semi-Infinite Optimization and Complementarity" (2011). *Mathematics Research Reports*. Paper 84.  
[http://digitalcommons.wayne.edu/math\\_reports/84](http://digitalcommons.wayne.edu/math_reports/84)

This Technical Report is brought to you for free and open access by the Mathematics at DigitalCommons@WayneState. It has been accepted for inclusion in Mathematics Research Reports by an authorized administrator of DigitalCommons@WayneState.

**COMPLETE CHARACTERIZATIONS OF LOCAL  
WEAK SHARP MINIMA WITH APPLICATIONS TO  
SEMI-INFINITE OPTIMIZATION AND  
COMPLEMENTARITY**

**BORIS S. MORDUKHOVICH, NAIHUA XIU, and JINCHUAN ZHOU**

**WAYNE STATE  
UNIVERSITY**

**Detroit, MI 48202**

**Department of Mathematics  
Research Report**

**2011 Series  
#3**

*This research was partly supported by the U.S.A. National Science Foundation*

# Complete Characterizations of Local Weak Sharp Minima with Applications to Semi-Infinite Optimization and Complementarity

Boris S. Mordukhovich\*    Naihua Xiu †    Jinchuan Zhou‡

February 2, 2011

## Abstract

In this paper we identify a favorable class of nonsmooth functions for which local weak sharp minima can be completely characterized in terms of normal cones and subdifferentials, or tangent cones and subderivatives, or their mixture in finite-dimensional spaces. The results obtained not only significantly extend previous ones in the literature, but also allow us to provide new types of criteria for local weak sharpness. Applications of the developed theory are given to semi-infinite programming and to semi-infinite complementarity problems.

**Key words.** local weak sharp minima, normal cones, tangent cones, subderivatives, subdifferentials, semi-infinite optimization, semi-infinite complementarity.

**AMS subject classifications.** 49J52, 65K10, 90C26.

---

\*Department of Mathematics, Wayne State University, Detroit, MI 48202, USA; boris@math.wayne.edu. Research of this author was partially supported by the U.S.A National Science Foundation under grant DMS-1007132 and by the Australian Research Council under grant DP-12092508.

†Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing 100044, P.R.China; nhxiu@bjtu.edu.cn. Research of this author was partly supported by the National Basic Research Program of China (2010CB732501).

‡Department of Mathematics, School of Science, Shandong University of Technology, Zibo 255049, P.R.China; jinchuanzhou@yahoo.com.cn. Research of this author was partly supported by Shandong Province Natural Science Foundation (ZR2010AQ026).

# 1 Introduction

This paper is mainly devoted to the study and complete characterizations of local weak sharp minima and their applications to problems of semi-infinite optimization and semi-infinite complementarity in finite-dimensional spaces.

Given an extended-real-valued function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$  and a point  $x \in \mathbb{R}^n$  with  $f(x) < \infty$ , recall that  $x$  is *local weak sharp minimum* of  $f$  if there exist positive scalars  $\eta$  and  $\delta$  such that

$$\eta \text{dist}(z, L_f(x)) \leq f(z) - f(x) \text{ for all } z \in B(x, \delta), \quad (1.1)$$

where  $B(x, \delta)$  is the closed ball with center  $x$  and radius  $\delta > 0$ , where

$$L_f(x) := \{z \in \mathbb{R}^n \mid f(z) = f(x)\},$$

is the level set of  $f$  at  $x$ , and where  $\text{dist}(x, A)$  is the distance function from  $x$  to a given set  $A \subset \mathbb{R}^n$  defined by

$$\text{dist}(x, A) := \inf_{y \in A} \|x - y\|.$$

Definition (1.1) clearly implies that  $x$  is a local minimum of  $f$ .

The notion of weak sharp minima was introduced by Ferris in [16] as a generalization of sharp minima due to Polyak [30] to include the possibility of non-unique solutions. During the last two decades the study of weak sharp minima has drawn much attention motivated by its importance in the treatment of sensitivity analysis (see, e.g., [1, 8]) and of convergence analysis for a wide range of optimization algorithms; we refer the reader to [6, 7, 9, 14, 17, ?, 23, 38] and the bibliographies therein. Roughly speaking, efficient conditions for weak sharp minima obtained in these papers via generalized differentiation can be classified into two types: *primal* conditions and *dual* conditions. The former involve tangent cones and directional derivatives, while the latter employ normal cones and subdifferentials.

Observe that necessary and sufficient conditions for local weak sharp minima were established in two special cases. The first case concerns the situation when  $x$  is a strict local minimum. Then definition (1.1) reduces, by shrinking  $\delta$  if necessary, to

$$\eta \|z - x\| \leq f(z) - f(x) \text{ for all } z \in B(x, \delta), \quad (1.2)$$

which is often referred to as local sharp minimum and is also called strongly unique local minimum; cf. [13, 30]). In this case it is not difficult to verify (see, e.g., [1, Chapter 3]) that (1.2) holds if and only if  $df(x)(w) > 0$  for all nonzero  $w \in \mathbb{R}^n$  via the subderivative of  $f$  defined in Section 2. Second, when the problem data are convex Burke and Ferris [6] provided several primal and dual characterizations of weak sharp minima and studied its impact to convex programming and convergence analysis in finite-dimensional setting; this was further extended by Burke and Deng [3] to infinite dimensions. Furthermore, close relationships between weak sharp minima, linear regularity, metric regularity, and error bound were exploited in [4, 5]. The recent paper [21] considers weak sharp minima for convex constrained optimization problems on Riemannian manifolds, containing also new characterizations for the case of conventional convex problems in finite-dimensional spaces.

In the general case, however, the nonconvexity of  $f$  and the non-uniqueness of solutions give rise to a lot of complications that invalidate classical techniques. To circumvent these difficulties, several approaches have been proposed. In particular, Wu and Ye [39] obtained dual sufficient conditions for global weak sharp minima in terms of an abstract subdifferential, a fairly general concept unifying most of specific subdifferentials useful in variational analysis. In [28], Ng and Zheng presented primal sufficient conditions for a proper lower semicontinuous function on a Banach space to have global weak sharp minima by using various kinds of lower generalized derivatives.

It is worth noting that the notion of weak sharp minima defined in (1.1) underlines a first-order growth of the objective function away from the level set  $L_f(x)$ . Meanwhile, weak sharp minima of higher order growth are also of interest in parametric optimization, because it can be used to establish Hölder continuity properties of solution mappings. In particular, weak sharp minima of order two was studied by Bonnans and Ioffe [2] in the case when  $f$  is a pointwise maximum of twice continuously differentiable convex functions. Sufficient conditions for weak sharp minima of order  $m \geq 1$  for nonconvex functions in finite dimensions were obtained by Studniarski and Ward [37] via the limiting normal cone by Mordukhovich and a certain extension of the regular tangent cone by Clarke.

Observe that, except for the two cases mentioned above and some particular situations, most of the conditions obtained for local weak sharp minima are either necessary or sufficient but not both. A natural and important question arises about the possibility to establish necessary and sufficient conditions for local weak sharp minima when  $f$  is not necessarily convex and  $x$  is not restricted to be a strict local minimum. An significant step in this direction was made by Zheng and Yang [40] who derived characterizations of local weak sharp minima for semi-infinite programming by exploiting the special structure of functions involved therein; see more details below.

The main purpose of this paper is to obtain efficient characterizations of local weak sharp minima in the general nonconvex framework of nonsmooth functions  $f$  in (1.1) and then to apply them to important classes of optimization-related problems. Our necessary and sufficient conditions are not only essentially extend the aforementioned ones to much broader classes of problems but also offer verifiable criteria of new types to characterize local weak sharp minima in both convex and nonconvex settings.

To achieve our goals, we introduce a new class of nonsmooth functions, called *inf-differentiable* functions, which are certainly of their independent interest. It is shown below that this class is sufficiently broad to cover a number of special classes of functions overwhelmingly encountered in variational analysis and optimization. Among those, besides the classical classes of smooth and convex functions, we particularly mention semidifferentiable functions, lower- $C^1$  functions, and functions given by parametric integrals with respect to finite measures over compact sets. The main results of this paper provide primal, dual, and mixed characterizations of weak local sharp minima for inf-differentiable functions in finite-dimensional spaces. These results enable us to fully characterize weak local sharp minima of semi-infinite programs in terms their initial data and to derive necessary and sufficient conditions for local error bounds of residuals in semi-infinite complementarity problems.

The rest of the paper is organized as follows. Section 2 collects some preliminaries from generalized differentiation widely used in the sequel. In Section 3 we introduce inf-differentiable functions and establish their relationships with other favorable classes of functions in variational analysis and optimization. Section 4 is devoted to characterizing local weak sharp minima for inf-differentiable functions. In Sections 5 and 6 we illustrate applications of the developed theory to important classes of problems in semi-infinite programming and semi-infinite complementarity, respectively. The final Section 7 presents concluding remarks and discussions on further research.

## 2 Preliminaries from Generalized Differentiation

In this section we briefly overview for the reader's convenience some constructions of generalized differentiation in variational analysis needed in what follows. We refer the reader to the monographs [25, 35] for more details, proofs, and notation.

Recall that the symbols  $\text{cl } A$ ,  $\text{co } A$ , and  $\text{cone } A$  stand for the closure, convex hull, and conic hull of a nonempty subset  $A \subset \mathbb{R}^n$ , respectively. Denote by  $B$  the closed unit ball in  $\mathbb{R}^n$  and by  $\mathbb{N} := \{1, 2, \dots\}$  the collections of natural numbers. The *support function*  $\sigma^*(\cdot|A): \mathbb{R}^n \rightarrow \overline{\mathbb{R}} := (-\infty, \infty]$  of  $A$  is defined by  $\sigma^*(w|A) := \sup_{a \in A} \langle a, w \rangle$  for all  $w \in \mathbb{R}^n$ . By

$$P_A(x) := \{y \in \text{cl } A \mid \|x - y\| = \text{dist}(x, A)\}$$

we denote the projection (the set of best approximations) of  $x$  onto  $A$ . The polar of  $A$  is  $A^0 := \{v \in \mathbb{R}^n \mid \langle v, a \rangle \leq 0 \text{ for all } a \in A\}$ .

Given a set-valued mapping  $S$  from  $\mathbb{R}^n$  into  $\mathbb{R}^m$ , define the *Painlevé–Kuratowski outer* and *inner limit* of  $S(z)$  as  $z \rightarrow x$  by, respectively,

$$\text{Lim sup}_{z \rightarrow x} S(z) := \{v \in \mathbb{R}^m \mid \exists z_k \rightarrow x \text{ and } v_k \rightarrow v \text{ with } v_k \in S(z_k)\},$$

$$\text{Lim inf}_{z \rightarrow x} S(z) := \{v \in \mathbb{R}^m \mid \forall z_k \rightarrow x, \exists v_l \rightarrow v \text{ with } v_l \in S(z_k)\}.$$

If no confusion arise, the symbols  $x' \xrightarrow{A} x$  and  $x' \xrightarrow{f} x$  mean that  $x' \rightarrow x$  with  $x' \in A$  and  $x' \rightarrow x$  with  $f(x') \rightarrow f(x)$ , respectively.

At any point  $x \in A$ , the *Bouligand-Severi tangent/contingent cone*  $T_A(x)$  is a closed cone defined via the outer limit

$$T_A(x) := \text{Lim sup}_{\tau \downarrow 0} \frac{A - x}{\tau}, \quad (2.1)$$

while the *Clarke/regular tangent cone* is defined by the inner limit

$$\widehat{T}_A(x) := \text{Lim inf}_{\substack{z \xrightarrow{A} x \\ \tau \downarrow 0}} \frac{A - z}{\tau}. \quad (2.2)$$

The *Fréchet/regular normal cone* (also known as the prenormal cone) is given by

$$\widehat{N}_A(x) := \left\{ v \in \mathbb{R}^n \mid \limsup_{z \xrightarrow{A} x} \frac{\langle v, z - x \rangle}{\|z - x\|} \leq 0 \right\} \quad (2.3)$$

and the *Mordukhovich/limiting/basic normal cone* can be equivalently defined by

$$N_A(x) := \text{Lim sup}_{z \xrightarrow{A} x} \widehat{N}_A(z) = \text{Lim sup}_{z \xrightarrow{A} x} [\text{cone}(z - P_A(z))]. \quad (2.4)$$

It follows from the above definitions that  $\widehat{T}_A(x) \subset T_A(x)$  and  $\widehat{N}_A(x) \subset N_A(x)$ . Furthermore, we have the polarity/duality relationships

$$\widehat{N}_A(x) = T_A(x)^0 \quad \text{and} \quad \widehat{T}_A(x) = N_A(x)^0. \quad (2.5)$$

However, the converse duality  $N_A(x) = \widehat{T}_A(x)^0$  is usually violated, since the limiting normal cone (2.4) is typically nonconvex in the absence of the normal regularity [25, 35], while polarity always generates convexity. The full duality is achieved by the *Clarke/convexified normal cone* defined by

$$\overline{N}_A(x) := \widehat{T}_A(x)^0 \quad \text{with} \quad \widehat{T}_A(x) = \overline{N}_A(x)^0, \quad (2.6)$$

which may be much larger than (2.4) satisfying the relationship  $\overline{N}_A(x) = \text{clco}N_A(x)$ . It follows from (2.6) implies that the regular tangent cone (2.2) is always convex—as well as the regular normal cone (2.3)—while the contingent cone (2.1) is usually not.

Given next an extended-real-valued function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  with the effective domain  $\text{dom } f := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$ , we define by

$$\widehat{\partial}f(x) := \{v \in \mathbb{R}^n \mid \liminf_{z \rightarrow x} \frac{f(z) - f(x) - \langle v, z - x \rangle}{\|z - x\|} \geq 0\}, \quad (2.7)$$

$$\partial f(x) := \text{Lim sup}_{z \xrightarrow{f} x} \widehat{\partial}f(z) \quad \text{and} \quad \partial^\infty f(x) := \text{Lim sup}_{\substack{z \xrightarrow{f} x \\ \lambda \downarrow 0}} \lambda \widehat{\partial}f(z), \quad (2.8)$$

$$\overline{\partial}f(x) := \text{clco}(\partial f(x) + \partial^\infty f(x)) \quad (2.9)$$

the *Fréchet subdifferential*, the *Mordukhovich basic/limiting subdifferential* and *singular subdifferential*, and the *Clarke subdifferential* of  $f$  at  $x \in \text{dom } f$ , respectively. It is well known that a lower semicontinuous (l.s.c) function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is locally Lipschitz continuous around  $x \in \text{dom } f$  if and only if  $\partial^\infty f(x) = \{0\}$ .

In what follows we also need two directional derivatives notions for a function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  at  $x \in \text{dom } f$ . The *subderivative* of  $f$  at  $x$  at the direction  $w \in \mathbb{R}^n$  is defined by

$$df(x)(w) := \liminf_{\substack{w' \rightarrow w \\ \tau \downarrow 0}} \frac{f(x + \tau w') - f(x)}{\tau}, \quad (2.10)$$

while the corresponding *Rockafellar/regular subderivative* is given by

$$\widehat{d}f(x)(w) := \lim_{\delta \downarrow 0} \left( \limsup_{\substack{x' \xrightarrow{f} x \\ \tau \downarrow 0}} \left[ \inf_{w' \in B(w, \delta)} \frac{f(x' + \tau w') - f(x')}{\tau} \right] \right). \quad (2.11)$$

Note that the Rockafellar subderivative  $\widehat{d}f(x)$  is always convex in directions and reduces to the generalized directional derivative of Clarke when  $f$  is locally Lipschitzian around the

reference point. Recall that a function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is *directionally regular* at  $x \in \text{dom } f$  if it is l.s.c. around this point and  $df(x) = \widehat{df}(x)$ . For locally Lipschitzian function the latter property is equivalent to the *subdifferential regularity*  $\widehat{\partial}f(\bar{x}) = \partial f(\bar{x})$  of  $f$  at  $\bar{x}$ .

The next result plays a significant role in our subsequent analysis.

**Lemma 2.1 (subderivative representation for regular functions).** *Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be locally Lipschitzian around  $x$  and directionally/subdifferentially regular at this point. Then we have*

$$df(x)(w) = \sigma^*(w | \partial f(x)) \text{ for all } w \in \mathbb{R}^n.$$

**Proof.** It follows from [35, Theorem 8.30] in which the requirement of  $\partial f(x) \neq \emptyset$  is satisfied by the local Lipschitz continuity of  $f$  around  $\bar{x}$ .  $\square$

### 3 Inf-Differentiable Functions

In this section we introduce a new class of nondifferentiable functions and discuss its relationships with other favorable classes of functions encountered in variational analysis, optimization, and their applications. The inf-differentiable functions introduced below play crucial roles in characterizing local weak sharp minima in Section 4 and the subsequent applications in Sections 5 and 6.

**Definition 3.1 (inf-differentiability and single inf-differentiability).** *Consider a function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  and a set  $S \subset \mathbb{R}^n$ . We say that  $f$  is INF-DIFFERENTIABLE at a given point  $x \in \mathbb{R}^n$  RELATIVE to  $S$  if*

$$\liminf_{\substack{z \rightarrow x \\ u \neq z, u \rightarrow x \\ S}} \frac{f(z) - f(u) - df(u)(z - u)}{\|z - u\|} = 0. \quad (3.1)$$

*In particular, if (3.1) holds with  $S = \mathbb{R}^n$  and with  $S = \{x\}$ , then  $f$  is called to be INF-DIFFERENTIABLE at  $x$  and SINGLE INF-DIFFERENTIABLE at  $x$ , respectively. Finally,  $f$  is inf-differentiable on a subset  $D \subset \mathbb{R}^n$  relative to  $S$  if (3.1) holds at every  $x \in D$ .*

It is easy to see that inf-differentiability of  $f$  at  $x$  implies its single inf-differentiability at this point, but not vice versa. We first consider the case of single inf-differentiability. In this case condition (3.1) clearly reduces to

$$\liminf_{w \rightarrow 0} \frac{f(x + w) - f(x) - df(x)(w)}{\|w\|} = 0. \quad (3.2)$$

Let us show that it holds for any local Lipschitzian function.

**Proposition 3.2 (single inf-differentiability from Lipschitz continuity).** *If  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is locally Lipschitzian around  $\bar{x}$ , it is single inf-differentiable at this point.*



**Proof.** We first prove the inequality

$$\liminf_{w \rightarrow 0} \frac{f(x+w) - f(x) - df(x)(w)}{\|w\|} \leq 0. \quad (3.3)$$

Assume on the contrary that (3.3) does not hold, i.e., there are  $\varepsilon, \delta > 0$  such that

$$f(x+w) - f(x) - df(x)(w) \geq \varepsilon\|w\| \text{ for all } w \neq 0 \text{ with } \|w\| \leq \delta.$$

It follows from the above inequality that

$$f(x+tw) - f(x) - tdf(x)(w) \geq \varepsilon t\|w\|$$

for all  $w \neq 0$  and  $t > 0$  sufficiently small. Dividing the latter by  $t > 0$  and taking the lower limit as  $t \downarrow 0$  give us the estimate

$$\liminf_{t \rightarrow 0} \frac{f(x+tw) - f(x)}{t} \geq df(x)(w) + \varepsilon\|w\|. \quad (3.4)$$

Furthermore, the local Lipschitzian continuity of  $f$  around  $x$  implies that

$$\begin{aligned} df(x)(w) &= \liminf_{\substack{w' \rightarrow w \\ t \downarrow 0}} \frac{f(x+tw') - f(x)}{t} \\ &= \liminf_{\substack{w' \rightarrow w \\ t \downarrow 0}} \left( \frac{f(x+tw') - f(x+tw)}{t} + \frac{f(x+tw) - f(x)}{t} \right) \\ &= \liminf_{t \rightarrow 0} \frac{f(x+tw) - f(x)}{t}. \end{aligned}$$

Combining this and (3.4) yields that

$$df(x)(w) \geq df(x)(w) + \varepsilon\|w\|, \quad w \in \mathbb{R}^n,$$

which is a contradiction that justifies (3.3).

To complete the proof of the proposition, it remains to show that the strict inequality

$$\liminf_{w \rightarrow 0} \frac{f(x+w) - f(x) - df(x)(w)}{\|w\|} < 0$$

does not hold in (3.3) Supposing the contrary, we find a sequence  $\{w_k\}$  such that

$$\lim_{k \rightarrow \infty} \frac{f(x+w_k) - f(x) - df(x)(w_k)}{\|w_k\|} < 0,$$

i.e., there exists a small positive number  $\varepsilon$  with

$$f(x+w_k) - f(x) < df(x)(w_k) - \varepsilon\|w_k\| \quad (3.5)$$

for all  $k \in \mathbb{N}$  sufficiently large. It follows from the assumed local Lipschitz continuity of  $f$  around  $x$  that the subderivative  $df(x)$  is globally Lipschitz continuous on  $\mathbb{R}^n$ ; see [35, Exercise 9.15]. Denoting  $t_k := \|w_k\|$  and  $v_k = \frac{w_k}{\|w_k\|}$  and using the boundedness of  $\{v_k\}$ , we

suppose without loss of generality that the whole sequence  $\{v_k\}$  converges to some limit  $v$  as  $k \rightarrow \infty$ . Dividing by  $\|w_k\|$  on both sides of (3.5) and taking the lower limit therein give us

$$\begin{aligned} df(x)(v) &\leq \liminf_{k \rightarrow \infty} \frac{f(x + t_k v_k) - f(x)}{t_k} \\ &\leq \lim_{k \rightarrow \infty} df(x)(v_k) - \varepsilon \\ &= df(x)(v) - \varepsilon, \end{aligned}$$

which is a clear contradiction that justifies the single inf-differentiability (3.2) and thus completes the proof of the proposition.  $\square$

Next we study the single inf-differentiability of functions from two remarkable classes widely used in variational analysis and optimization: semidifferentiable and  $B$ -differentiable functions. Being interrelated, these two classes are generally different from each other and contain functions that are not locally Lipschitzian.

Recall that a function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  finite at  $x$  is *semidifferentiable* at  $x$  [35] (known also as directionally differentiable in the Hadamard sense [1, Chapter 2]) if the limit

$$\lim_{\substack{w' \rightarrow w \\ t \downarrow 0}} \frac{f(x + tw') - f(x)}{t}$$

exists (may be infinite) for all  $w \in \mathbb{R}^n$ .

**Proposition 3.3 (single inf-differentiability from semidifferentiability).** *Any function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  semidifferentiable at  $x$  is single inf-differentiable at this point.*

**Proof.** It follows from [35, Theorem 7.21] that the semidifferentiability of  $f$  at  $x$  implies the representation via its subderivative:

$$f(x + w) = f(x) + df(x)(w) + o(\|w\|), \quad w \in \mathbb{R}^n.$$

The latter readily yields (3.2), which means the single inf-differentiability of  $f$  at  $x$ .  $\square$

Recall further that  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is  $B$ -differentiable (in the sense of Robinson [33]) at  $x \in \text{dom } f$  if it is directionally differentiable at this point and

$$\lim_{z \rightarrow x} \frac{f(z) - f(x) - f'(x; z - x)}{\|z - x\|} = 0. \quad (3.6)$$

Note that the class of  $B$ -differentiable functions and its semismooth subclass play an important role in many aspects of optimization, especially for designing and justifying nonsmooth Newton-type algorithms to solve nonsmooth equations; see, e.g., [15] and the references therein. For locally Lipschitzian functions the  $B$ -differentiability is equivalent to the classical directional differentiability [36]. The next result shows that the single inf-differentiability is weaker than the  $B$ -differentiability under some mild assumptions. Furthermore, we give an example to illustrate the failure of the converse implication.

**Proposition 3.4 (single inf-differentiability from  $B$ -differentiability).** *Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be  $B$ -differentiable at  $x \in \text{dom } f$ , and the condition  $df(x)(0) > -\infty$  be satisfied. Then  $f$  is single inf-differentiable at this point.*

**Proof.** We first claim that  $df(x)(w) > -\infty$  for all  $w \in \mathbb{R}^n$ . Indeed, assuming on the contrary that there is some vector  $w \in \mathbb{R}^n$  such that  $df(x)(w) = -\infty$  gives us

$$\begin{aligned} df(x)(0) &\leq \liminf_{t \downarrow 0} df(x)(tw) \\ &= \liminf_{t \downarrow 0} (tdf(x)(w)) \\ &= -\infty, \end{aligned}$$

where the inequality is due to the lower semicontinuity of  $df(x)$  by [35, Theorem 8.18] and the first equality is due to the positive homogeneity of  $df(x)$ . This contradicts the hypothesis  $df(x)(0) > -\infty$ . On the other hand, since  $df(x)(w) \leq f'(x; w)$  for all  $w$  by definition, the subderivative  $df(x)$  is finite everywhere. Hence we have

$$f(x+w) - f(x) - df(x)(w) \geq f(x+w) - f(x) - f'(x; w), \quad w \in \mathbb{R}^n,$$

which, together with the B-differentiability of  $f$ , implies that

$$\liminf_{w \rightarrow 0} \frac{f(x+w) - f(x) - df(x)(w)}{\|w\|} \geq \lim_{w \rightarrow 0} \frac{f(x+w) - f(x) - f'(x; w)}{\|w\|} = 0.$$

To complete the proof of the proposition, it remains to show that the strict inequality

$$\liminf_{w \rightarrow 0} \frac{f(x+w) - f(x) - df(x)(w)}{\|w\|} > 0 \tag{3.7}$$

does not hold. Suppose on the contrary that (3.7) holds and find  $\varepsilon, \delta > 0$  such that

$$f(x+w) - f(x) - df(x)(w) \geq \varepsilon \|w\| \tag{3.8}$$

whenever  $w$  is sufficiently close to zero. Given any nonzero  $w \in \mathbb{R}^n$  and  $w' \rightarrow w$ , it follows from (3.8) that for  $t$  sufficiently small we have

$$f(x+tw') - f(x) \geq tdf(x)(w') + t\varepsilon \|w'\|.$$

Taking into account the lower semicontinuity of  $df(x)$ , the latter inequality implies that

$$\begin{aligned} df(x)(w) &= \liminf_{\substack{w' \rightarrow w \\ t \downarrow 0}} \frac{f(x+tw') - f(x)}{t} \\ &\geq \liminf_{w' \rightarrow w} df(x)(w') + \varepsilon \|w\| \\ &\geq df(x)(w) + \varepsilon \|w\|, \end{aligned}$$

which is a contradiction that completes the proof.  $\square$

The following example shows that the converse statements to all the three Propositions 3.2–3.4 do not generally hold.

**Example 3.5 (single inf-differentiability of non-Lipschitzian, non-B-differentiable, and non-semidifferentiable functions).** Define a function  $f: \mathbb{R} \rightarrow \mathbb{R}$  by

$$f(x) := \begin{cases} \max\{0, x \sin \frac{1}{x}\} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

This function clearly fails to be locally Lipschitz continuous, B-differentiable, and semidifferentiable at the origin, since the directional derivative does not exist. Note that  $df(x)$  is positively homogeneous, and hence its values are completely determined by those at  $w = \pm 1$ . By the simple calculation we get  $df(0)(1) = df(0)(-1) = 0$ . This implies that

$$\begin{aligned} & \liminf_{t \rightarrow 0} \frac{f(t) - f(0) - df(0)(t)}{t} \\ &= \min \left\{ \liminf_{t \downarrow 0} \frac{f(t) - f(0) - df(0)(t)}{t}, \liminf_{t \uparrow 0} \frac{f(t) - f(0) - df(0)(t)}{t} \right\} \\ &= \min \left\{ \liminf_{t \downarrow 0} \left[ \max \left\{ 0, \sin \frac{1}{t} \right\} \right], 0 \right\} = 0 \end{aligned}$$

and thus establishes the single inf-differentiability of  $f$  at the origin.

The next example shows that the statements of Propositions 3.2–3.4 may fail for inf-differentiable functions in the sense of Definition 3.1.

**Example 3.6 (inf-differentiable versus single inf-differentiable functions).** Consider a function  $f: \mathbb{R} \rightarrow \mathbb{R}$  given by  $f(x) := -|x|$ . This function is obviously locally Lipschitzian, B-differentiable, and semidifferentiable; thus it is single inf-differentiable by Propositions 3.2–3.4. On the other hand, for  $z \downarrow 0$  and  $u = -z$ , we have

$$\liminf_{\substack{z \rightarrow 0 \\ u \neq z, u \rightarrow 0}} \frac{f(z) - f(u) - df(u)(z - u)}{\|z - u\|} \leq \liminf_{\substack{z \downarrow 0 \\ u = -z}} \frac{f(z) - f(u) - df(u)(z - u)}{\|z - u\|} = -1 < 0,$$

which shows that (3.1) does not hold for  $S = \mathbb{R}$ , and thus  $f$  is not inf-differentiable at  $x = 0$ .

Now we proceed with studying inf-differentiability. Let us first make a useful observation that makes it more convenient to check inf-differentiability.

**Proposition 3.7 (inf-differentiability relative to subsets).** *Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be locally Lipschitzian around  $x \in \mathbb{R}^n$ , and let  $S \subset \mathbb{R}^n$  be a given subset. If  $f$  is inf-differentiable at  $x$  relative to  $S$ , then  $f$  is inf-differentiable at  $x$  relative to every nonempty subset of  $S$ .*

**Proof.** Take any nonempty subset  $S_1 \subset S$ . Then we have

$$\begin{aligned} \liminf_{z \rightarrow x} \frac{f(z) - f(x) - df(x)(z - x)}{\|z - x\|} &\geq \liminf_{\substack{z \rightarrow x \\ u \neq z, u \rightarrow x \\ S_1}} \frac{f(z) - f(x) - df(z)(z - x)}{\|z - x\|} \\ &\geq \liminf_{\substack{z \rightarrow x \\ u \neq z, u \rightarrow x \\ S}} \frac{f(z) - f(x) - df(z)(z - x)}{\|z - x\|}. \end{aligned} \tag{3.9}$$

It follows from Proposition 3.2 that the value on the left-hand side of (3.9) is zero. At the same time the inf-differentiability of  $f$  relative to  $S$  yields that the value of the right side is zero as well. This completes the proof of the proposition.  $\square$

Recall that a function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is lower- $C^1$  around  $\bar{x} \in \text{dom } f$  if there is a neighborhood  $U$  of  $\bar{x}$  such that

$$f(x) = \max_{y \in Y} \varphi(x, y) \text{ for all } x \in U, \tag{3.10}$$

where the index set  $Y$  is compact, and where the function  $\varphi$  is of class  $C^1$  in the first variable with the continuous partial derivative on  $U \times Y$ ; see [35, Definition 10.29]. Note that, besides smooth functions, the lower- $C^1$  class—known in fact under different names—includes remarkable collections of functions well-recognized and applied in variational analysis and optimization; see, e.g., [26, pp. 135–136], [35, pp. 447–452], and the references therein.

The next theorem, which is the main result of this section, establishes the inf-differentiability of lower- $C^1$  and other favorable classes of functions used in what follows.

**Theorem 3.8 (inf-differentiability of favorable classes of functions).** *Given a function  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ , the following assertions hold:*

- (i) *If  $f$  is convex, then it is inf-differentiable on any closed and bounded subsets of the relative interior of its domain.*
- (ii) *If  $f$  is a lower- $C^1$  function around  $x \in \text{dom } f$ , then  $f$  is inf-differentiable at this point.*
- (iii) *Let  $f$  be an integral function given by*

$$f(x) := \int_{\Omega} \varphi(x, y) d\mu(y). \quad (3.11)$$

*Then it is inf-differentiable on  $\mathbb{R}^n$  provided validity of the following assumptions:*

- (a)  $\Omega$  is a compact set and  $\mathcal{F}$  is its Borel sigma-algebra;
- (b)  $\mu$  is a finite measure defined on a measurable space  $(\Omega, \mathcal{F})$ ;
- (c)  $\nabla_x \varphi$  is continuous on  $\mathbb{R}^n \times \Omega$ .

**Proof.** First we note that all the functions  $f$  in assertions (i)–(iii) are Lipschitz continuous on the corresponding sets under the assumptions made. Indeed, it follows from [34, Theorem 10.4] for (i), from [35, Theorem 10.31] for (ii), and from [12, Theorem 2.7.2] for (iii). According to Proposition 3.2 and the estimate in (3.9) of Proposition 3.7, the inf-differentiability of  $f$  in all the assertions (i)–(iii) follows from the inequality

$$\liminf_{\substack{z \rightarrow x \\ u \neq z, u \rightarrow x}} \frac{f(z) - f(u) - df(u)(z - u)}{\|z - u\|} \geq 0. \quad (3.12)$$

which we are going to verify in the cases under consideration.

To justify assertion (i), recall that any convex function is directionally differentiable on its domain and satisfies the estimates

$$df(u)(z - u) \leq f'(u; z - u) \leq f(z) - f(u),$$

which clearly yields inequality (3.12) in the convex case (i).

To prove assertion (ii), we use representation (3.10) of  $f$  due to its assumed lower- $C^1$  property. Given arbitrary positive scalars  $\varepsilon$  and  $\delta_1$ , the continuity of  $\nabla_x \varphi$  and the compactness of the set  $B(x, \delta_1) \times Y$  imply the existence of  $\delta_2 > 0$  with  $\delta_2 < \delta_1$  such that

$$\|\nabla_x \varphi(x_1, y) - \nabla_x \varphi(x_2, y)\| < \varepsilon \text{ for all } y \in Y \quad (3.13)$$

whenever  $x_1, x_2 \in B(x, \delta_1)$  satisfy  $\|x_1 - x_2\| < \delta_2$ . Let  $\delta := \frac{1}{2}\delta_2$ . Choosing  $z, u \in B(x, \delta)$  and  $y \in Y(u) := \arg \max_{y \in Y} \phi(u, y)$ , it follows from (3.13) that

$$\begin{aligned} \varphi(z, y) - f(u) &= \varphi(z, y) - \varphi(u, y) \\ &= \int_0^1 \langle \nabla_x \varphi(u + t(z - u), y), z - u \rangle dt \\ &= \int_0^1 \langle \nabla_x \varphi(u + t(z - u), y) - \nabla_x \varphi(u, y), z - u \rangle dt + \langle \nabla_x \varphi(u, y), z - u \rangle \\ &\geq \end{aligned}$$

Taking the pointwise supremum of  $y \in Y(u)$  on the both sides of the above inequality yields

$$\begin{aligned} f(z) - f(u) &\geq \max_{y \in Y(u)} \varphi(z, y) - f(u) \\ &\geq -\varepsilon \|z - u\| + \max_{y \in Y(u)} \langle \nabla_x \varphi(u, y), z - u \rangle \\ &= -\varepsilon \|z - u\| + df(u)(z - u), \end{aligned}$$

where the last equality follows from the subderivative calculus rule established in [35, Theorem 10.31]. This proves assertions (ii).

Assertion (iii) can be proved in the same vein. Indeed, according to [12, Theorem 2.7.2] and [31, Proposition 1] we have under the assumptions made in (a)–(c) that

$$df(u)(w) = f'(u; w) = \int_{\Omega} \langle \nabla_x \varphi(u, y), w \rangle d\mu(y) \text{ for all } w \in \mathbb{R}^n.$$

This clearly implies the relationships

$$\begin{aligned} &f(z) - f(u) - df(u)(z - u) \\ &= \int_{\Omega} [\varphi(z, y) - \varphi(u, y) - \langle \nabla_x \varphi(u, y), z - u \rangle] d\mu(y) \\ &= \int_{\Omega} \left\{ \int_0^1 [\langle \nabla_x \varphi(u + t(z - u), y), z - u \rangle - \langle \nabla_x \varphi(u, y), z - u \rangle] dt \right\} d\mu(y) \\ &\geq - \int_{\Omega} \left\{ \int_0^1 [\|\nabla_x \varphi(u + t(z - u), y) - \nabla_x \varphi(u, y)\| \|z - u\|] dt \right\} d\mu(y) \\ &\geq -\varepsilon \mu(\Omega) \|z - u\|, \end{aligned}$$

where the last inequality is due to (3.13). Since  $\mu$  is a finite measure and  $\Omega$  is compact, we get  $\mu(\Omega) < \infty$ , which justifies assertion (iii) due to the arbitrary choice of  $\varepsilon > 0$ . This completes the proof of the theorem.  $\square$

## 4 Characterizations of Local Weak Sharp Minima

One of the main features of inf-differentiable functions is that local weak sharp minima can be completely characterized via the primal constructions of tangent cones and subderivatives, as well as via the dual constructions of normal cones and subdifferentials, and also in terms of their mixture. This is the main contents of this section.

We start with the following *dual characterizations* of this important notion of minima in optimization and variational analysis.

**Theorem 4.1 (dual characterizations of local weak sharp minima).** *Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be locally Lipschitzian around  $x$  and subdifferentially regular at this point, and let  $f$  be inf-differentiable at  $x \in \text{dom } f$  relative to the level set  $L_f(x)$ . Then the following are equivalent:*

- (a) *The point  $x$  is a local weak sharp minimum of  $f$ .*
- (b) *There exist two positive scalars  $\eta$  and  $\delta$  such that*

$$\widehat{N}_{L_f(x)}(u) \cap \eta B \subset \partial f(u) \text{ for all } u \in L_f(x) \cap B(x, \delta).$$

- (c) *There exist two positive scalars  $\eta$  and  $\delta$  such that*

$$N_{L_f(x)}(u) \cap \eta B \subset \partial f(u) \text{ for all } u \in L_f(x) \cap B(x, \delta). \quad (4.1)$$

- (d) *There exist two positive scalars  $\eta$  and  $\delta$  such that*

$$\overline{N}_{L_f(x)}(u) \cap \eta B \subset \partial f(u) \text{ for all } u \in L_f(x) \cap B(x, \delta). \quad (4.2)$$

**Proof.** To justify (a)  $\implies$  (b), pick  $u \in L_f(x) \cap B(x, \frac{\delta}{2})$  and  $u^* \in \widehat{N}_{L_f(x)}(u) \cap B$ . Given any  $\varepsilon > 0$  and using definition (2.3) of the Fréchet normals, find  $\tau \in (0, \frac{\delta}{2})$  such that

$$\langle u^*, v - u \rangle \leq \varepsilon \|v - u\| \text{ for all } v \in L_f(x) \cap B(u, \tau).$$

Taking further  $z \in B(u, \frac{\tau}{2}) \subset B(x, \delta)$  and a projection  $v \in P_{L_f(x)}(z)$  (not necessarily unique due to the nonconvexity of  $L_f(x)$ ), we get  $\|v - u\| \leq \tau$ . Hence

$$\begin{aligned} \langle u^*, z - u \rangle &= \langle u^*, z - v \rangle + \langle u^*, v - u \rangle \\ &\leq \|z - v\| + \varepsilon \|v - u\| \\ &\leq (1 + \varepsilon) \|z - v\| + \varepsilon \|z - u\| \\ &= (1 + \varepsilon) \text{dist}(z, L_f(x)) + \varepsilon \|z - u\|. \end{aligned}$$

From definition (1.1) of local weak sharp minima and the obvious inclusion  $B(u, \frac{\tau}{2}) \subset B(x, \delta)$ , we have the estimate

$$\eta \langle u^*, z - u \rangle \leq (1 + \varepsilon)(f(z) - f(u)) + \eta \varepsilon \|z - u\| \text{ for all } z \in B(u, \frac{\tau}{2}),$$

which is equivalent to the inequality

$$0 \leq \eta \langle u^*, u - z \rangle + (1 + \varepsilon)(f(z) - f(u)) + \eta \varepsilon \|z - u\| \text{ for all } z \in B(u, \frac{\tau}{2}).$$

Noting that  $f(u) = f(x)$  by the definition of the level set  $L_f(x)$ , the above inequality implies that the function  $-\eta \langle u^*, z \rangle + (1 + \varepsilon)f(z) + \eta \varepsilon \|z - u\|$  attains its minimum at  $u$ . Invoking now the well-known necessary optimality condition in term in the limiting subdifferential (2.8) (see, e.g., [26, Proposition 5.3] and [35, Theorem 6.12]), we get

$$0 \in -\eta u^* + (1 + \varepsilon) \partial f(u) + \eta \varepsilon B,$$

which is equivalent to the inclusion

$$\eta u^* \in \partial f(u) + \varepsilon \partial f(u) + \eta \varepsilon B. \quad (4.3)$$

Noting that the local Lipschitz continuity of  $f$  around  $x$  ensures the local boundedness of  $\partial f$  around this point (see, e.g., [35, Theorem 9.13]) and then letting  $\varepsilon \downarrow 0$  in (4.3) yield that  $\eta u^* \in \partial f(u)$ , which justifies implication (ii).

To justify implication (ii)  $\implies$  (iii), pick arbitrary  $u \in L_f(x) \cap B(x, \frac{\delta}{2})$  and  $u^* \in N_{L_f(x)}(u) \cap \frac{\eta}{2}B$ . We have from definition (2.4) of the limiting normal cone that there are sequences  $\{u_k\} \subset L_f(x)$  and  $\{u_k^*\} \subset \widehat{N}_{L_f(x)}(u_k)$  such that  $u_k \rightarrow u$  and  $u_k^* \rightarrow u^*$  as  $k \rightarrow \infty$ . Since  $u_k \in B(x, \delta)$  and  $u_k^* \in \eta B$  for all  $k$  sufficiently large, assertion (ii) implies that  $u_k^* \in \partial f(u_k)$ . The desired result follows now by taking the limit as  $k \rightarrow \infty$  and using the well-known outer semicontinuity of the limiting subdifferential  $\partial f$ . Implication (iii)  $\implies$  (iv) is obtained in this way by taking the closed convex hull in the left-hand side of (4.1).

Let us finally justify implication (iv)  $\implies$  (i). Given an arbitrary  $\varepsilon \in (0, \frac{\eta}{2})$ , the inf-differentiability of  $f$  implies the existence of  $\delta_1 > 0$  such that

$$df(u)(z - u) \leq f(z) - f(u) + \varepsilon \|z - u\| \text{ for all } z \in B(x, \delta) \text{ and } u \in B(x, \delta) \cap L_f(x). \quad (4.4)$$

Denoting  $\delta' := \min\{\delta_1, \delta\}$  and taking  $z \in B(x, \frac{\delta'}{2}) \setminus L_f(x)$  and  $u \in P_{L_f(x)}(z)$ , it follows from [35, Example 6.16] that

$$\frac{\eta(z - u)}{\|z - u\|} \in \widehat{N}_{L_f(x)}(u) \cap \eta B \subset \overline{N}_{L_f(x)}(u) \cap \eta B,$$

which implies together with (4.2) that  $\frac{\eta(z - u)}{\|z - u\|} \in \partial f(u)$ . This gives us

$$\sigma^* \left( z - u \mid \frac{\eta(z - u)}{\|z - u\|} \right) \leq \sigma^*(z - u \mid \partial f(u))$$

and yields by (4.4) and Lemma 2.1 that

$$\eta \|z - u\| \leq df(u)(z - u) \leq f(z) - f(u) + \varepsilon \|z - u\|.$$

The latter implies in turn that

$$(\eta - \varepsilon) \|z - u\| \leq f(z) - f(u)$$

and thus completes the proof of the theorem.  $\square$

Note that in a number of papers [3, 6, 21, 28, 37, 40] contain either necessary or sufficient conditions of the dual-type for weak sharp minima given in terms of some normal cones and subdifferentials. In particular, the necessity part of (ii), and hence of (iii) and (iv), is proved in [27] for the general Banach space setting. The results of Theorem 4.1 show that the inf-differentiability allows us to justify also the sufficiency of the conditions above for weak sharp minima, i.e., to obtain full dual characterizations of this concept.

Next we derive *primal characterizations* of local weak sharp minima via tangent cones and subderivative constructions of Section 2.

**Theorem 4.2 (primal characterizations of local weak sharp minima).** *Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be locally Lipschitzian around  $x$  and subdifferentially regular at this point, and let  $f$  be inf-differentiable at  $x \in \text{dom } f$  relative to the level set  $L_f(x)$ . Then the following are equivalent:*



(i) The point  $x$  is a local weak sharp minimum of  $f$ .

(ii) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \operatorname{dist}(w, T_{L_f(x)}(u)) \leq df(u)(w) \text{ for all } u \in L_f(x) \cap B(x, \delta) \text{ and } w \in \mathbb{R}^n. \quad (4.5)$$

(iii) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \operatorname{dist}(w, \widehat{T}_{L_f(x)}(u)) \leq df(u)(w) \text{ for all } u \in L_f(x) \cap B(x, \delta) \text{ and } w \in \mathbb{R}^n.$$

(iv) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \|z - u\| \leq df(u)(z - u) \text{ for all } z \in B(x, \delta) \text{ and } u \in P_{L_f(x)}(z). \quad (4.6)$$

**Proof.** First we justify implication (i)  $\implies$  (iii). As showed in Theorem 4.1,  $x$  is a local weak sharp minimum of  $f$  if and only if there exist two positive scalars  $\eta$  and  $\delta$  such that  $\overline{N}_{L_f(x)}(u) \cap \eta B \subset \partial f(u)$  for all  $u \in L_f(x) \cap B(x, \delta)$ . Pick  $u \in L_f(x) \cap B(x, \delta)$  and  $w \in \mathbb{R}^n$ . Since the convergence  $u' \xrightarrow{L_f(x)} u$  clearly implies that of  $u' \xrightarrow{f} u$ , it follows from the subdifferential/directional regularity of  $f$  at  $x$  that

$$\begin{aligned} df(u)(w) &= \widehat{d}f(u)(w) \\ &= \lim_{\delta \downarrow 0} \left( \limsup_{\substack{u' \xrightarrow{f} u \\ \tau \downarrow 0}} \left[ \inf_{w' \in B(w, \delta)} \frac{f(u' + \tau w') - f(u')}{\tau} \right] \right) \\ &\geq \lim_{\delta \downarrow 0} \left( \limsup_{\substack{u' \xrightarrow{L_f(x)} u \\ \tau \downarrow 0}} \left[ \inf_{w' \in B(w, \delta)} \frac{f(u' + \tau w') - f(u')}{\tau} \right] \right) \geq 0, \end{aligned}$$

where the last inequality follows from (1.1) due to  $f(z) \geq f(x) = f(u')$  when  $z$  is sufficiently close to  $x$ . This readily implies the estimate

$$df(u)(w) \geq \eta \operatorname{dist}(w, \widehat{T}_{L_f(x)}(u)) \text{ for all } w \in \widehat{T}_{L_f(x)}(u).$$

Now we consider  $w \notin \widehat{T}_{L_f(x)}(u)$  and let  $\bar{w} = P_{\widehat{T}_{L_f(x)}(u)}(w)$ , where the uniqueness of projections comes from the convexity of the regular tangent cone  $\widehat{T}_{L_f(x)}(u)$  (2.2); see, e.g., [35, Theorem 6.26]. Furthermore, it follows from [35, Example 6.16] and the subdifferential regularity of  $f$  at  $x$  that  $w - \bar{w} \in \widehat{N}_{\widehat{T}_{L_f(x)}(u)}(\bar{w}) = N_{\widehat{T}_{L_f(x)}(u)}(\bar{w})$ , which in turn implies that

$$\langle w - \bar{w}, z - \bar{w} \rangle \leq 0 \text{ for all } z \in \widehat{T}_{L_f(x)}(u). \quad (4.7)$$

Since  $\widehat{T}_{L_f(x)}(u)$  is a convex cone, the relationship in (4.7) can be rewritten as

$$\langle w - \bar{w}, z \rangle \leq 0 \text{ and } \langle w - \bar{w}, \bar{w} \rangle = 0 \text{ for all } z \in \widehat{T}_{L_f(x)}(u),$$

which implies together with (2.6) that  $w - \bar{w} \in (\widehat{T}_{L_f(x)}(u))^0 = \overline{N}_{L_f(x)}(u)$ . Combining this with (4.2) gives us  $\frac{\eta(w - \bar{w})}{\|w - \bar{w}\|} \in \partial f(u)$ . By using Lemma 2.1 we get that

$$\begin{aligned} \eta \operatorname{dist}(w, \widehat{T}_{L_f(x)}(u)) &= \eta \|w - \bar{w}\| \\ &= \left\langle \frac{\eta(w - \bar{w})}{\|w - \bar{w}\|}, w - \bar{w} \right\rangle \\ &= \left\langle \frac{\eta(w - \bar{w})}{\|w - \bar{w}\|}, w \right\rangle \\ &\leq \sigma^*(w | \partial f(u)) = df(u)(w), \end{aligned}$$

which completes the proof of assertion (iii).

To justify implication (iii)  $\implies$  (ii), it suffices to show that

$$\operatorname{dist}(w, \widehat{T}_{L_f(x)}(u)) \geq \operatorname{dist}(w, T_{L_f(x)}(u)), \quad w \in \mathbb{R}^n,$$

since the inclusion  $\widehat{T}_{L_f(x)}(u) \subset T_{L_f(x)}(u)$  always holds by [35, Theorem 6.26].

Let us next prove implication (ii)  $\implies$  (iv). Pick  $z \in B(x, \frac{\delta}{2}) \setminus L_f(x)$  and  $u \in P_{L_f(x)}(z)$  and observe that

$$\frac{z - u}{\|z - u\|} \in \widehat{N}_{L_f(x)}(u) = (T_{L_f(x)}(u))^0,$$

which means that  $\left\langle \frac{z - u}{\|z - u\|}, \xi \right\rangle \leq 0$  for all  $\xi \in T_{L_f(x)}(u)$ . Using this and the classical Cauchy-Schwartz inequality, we get

$$\|z - u\| \leq \left\langle \frac{z - u}{\|z - u\|}, z - u - \xi \right\rangle \leq \|z - u - \xi\| \quad \text{for all } \xi \in T_{L_f(x)}(u),$$

which in turn implies that  $0 \in P_{T_{L_f(x)}(u)}(z - u)$  and show that  $\|z - u\| = \operatorname{dist}(z - u, T_{L_f(x)}(u))$ . The result now follows by replacing  $w$  by  $z - u$  in (4.5).

Finally, we prove implication (iv)  $\implies$  (i). Taking  $\varepsilon \in (0, \frac{\eta}{2})$  and using the inf-differentiability of  $f$  yields the existence of  $\delta_1 > 0$  that ensures the fulfillment of the estimate

$$df(u)(z - u) \leq f(z) - f(u) + \varepsilon \|z - u\| \quad \text{for all } z \in B(x, \delta_1) \text{ and } u \in B(x, \delta_1) \cap L_f(x). \quad (4.8)$$

Denoting  $\delta' := \frac{1}{2} \min\{\delta_1, \delta\}$ , we get from (4.6) and (4.8) that  $\eta \|z - u\| \leq f(z) - f(u) + \varepsilon \|z - u\|$ , i.e.,  $(\eta - \varepsilon) \|z - u\| \leq f(z) - f(u)$ . The latter justifies the validity of (1.1) with constants  $\frac{\eta}{2}$  and  $\delta'$  therein and thus completes the proof of the theorem.  $\square$

The next two theorems provide *mixed characterizations* of local weak sharp minima involving both primal and dual constructions of generalized differentiation.

**Theorem 4.3 (mixed characterizations of local weak sharp minima, I).** *Assume that  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is locally Lipschitzian around  $x$  and subdifferentially regular at this point, and that it is inf-differentiable at  $x \in \operatorname{dom} f$  relative to the level set  $L_f(x)$ . Then  $x$  is a local weak sharp minimum of  $f$  if and only if the following two conditions hold:*

- (a)  $T_{L_f(x)}(x) = \{w \in \mathbb{R}^n \mid df(x)(w) \leq 0\}$ .
- (b) *There exist two positive scalars  $\eta$  and  $\delta$  such that*

$$\overline{N}_{L_f(x)}(u) \cap \eta B \subset \partial f(u) \quad \text{for all } u \in L_f(x) \cap B(x, \delta) \setminus \{x\}.$$

**Proof.** First we justify the necessity part. According to criterion (ii) in Theorem 4.2 and criterion (iv) in Theorem 4.1 we only need to check the inclusion

$$T_{L_f(x)}(x) \subset \{w \in \mathbb{R}^n \mid df(x)(w) \leq 0\}. \quad (4.9)$$

To proceed, observe from definition (2.1) of the contingent cone that for any  $w \in T_{L_f(x)}(x)$  there exist sequences  $t_k \downarrow 0$  and  $w_k \rightarrow w$  as  $k \rightarrow \infty$  such that  $x_k = x + t_k w_k \in L_f(x)$ , i.e.,  $f(x_k) = f(x)$  for all  $k \in \mathbb{N}$  by the construction of the level set. The inf-differentiability of  $f$  implies the existence of  $\delta > 0$  for which

$$df(u)(z - u) \leq f(z) - f(u) + \varepsilon \|z - u\| \text{ for all } z \in B(x, \delta) \text{ and } u \in L_f(x) \cap B(x, \delta).$$

Thus for all  $k \in \mathbb{N}$  sufficiently large we have

$$df(x)(x_k - x) \leq f(x_k) - f(x) + \varepsilon \|x_k - x\| = \varepsilon t_k \|w_k\|.$$

The latter implies by the positive homogeneity of  $df(x)$  that

$$df(x)(w_k) \leq \varepsilon \|w_k\| \text{ for all large } k \in \mathbb{N}$$

Taking the lower limit in the above inequality as  $k \rightarrow \infty$  yields

$$df(x)(w) \leq \liminf_{k \rightarrow \infty} df(x)(w_k) \leq \varepsilon \|w\|,$$

where the first inequality is due to the lower semicontinuity of  $df(x)$  by [35, Theorem 18.18]. Since  $\varepsilon > 0$  was chosen arbitrarily, it follows from the above that  $df(x)(w) \leq 0$ , and hence we arrive at claimed inclusion (4.9).

To prove the sufficiency part, we first show that there is  $\eta_1 > 0$  for which

$$\eta_1 \|w\| \leq df(x)(w) \text{ whenever } w \in \widehat{N}_{L_f(x)}(x). \quad (4.10)$$

Suppose the contrary and find a sequence  $\{w_k\} \subset \widehat{N}_{L_f(x)}(x)$  with  $\|w_k\| = 1$  and

$$df(x)(w_k) < \frac{1}{k} \text{ for all } k \in \mathbb{N}. \quad (4.11)$$

Assume without loss of generality that  $w_k \rightarrow \bar{w}$  as  $k \rightarrow \infty$  for some  $\bar{w} \in \mathbb{R}^n$  of the unit norm. Hence  $\bar{w} \in \widehat{N}_{L_f(x)}(x)$  by the closedness of the latter cone. Taking the limit in (4.11) as  $k \rightarrow \infty$  and using the lower semicontinuity of  $df(x)$ , we get  $df(x)(\bar{w}) \leq 0$ , and thus  $\bar{w} \in T_{L_f(x)}(x)$  according to condition (a). It follows from the the first duality relation in (2.6) that  $\bar{w} \in \widehat{N}_{L_f(x)}(x) \cap T_{L_f(x)}(x) = (T_{L_f(x)}(x))^0 \cap T_{L_f(x)}(x)$ , which in turn obviously implies that  $\bar{w} = 0$ , a contradiction.

Further, let  $z \in B(x, \frac{\delta}{2}) \setminus L_f(x)$  and  $u \in P_{L_f(x)}(z)$ . Then  $u \in B(x, \delta)$  and

$$\frac{z - u}{\|z - u\|} \in \widehat{N}_{L_f(x)}(u) \subset \overline{N}_{L_f(x)}(u). \quad (4.12)$$

We now consider separately the two possible cases: (i)  $u = x$  and (ii)  $u \neq x$ . In case (i) it follows from (4.10) and (4.12) that

$$\eta_1 \|z - x\| \leq df(x)(z - x). \quad (4.13)$$

In case (ii) we get from condition (b) of the theorem that

$$\frac{\eta(z-u)}{\|z-u\|} \in \overline{N}_{L_f(x)}(u) \cap \eta B \subset \partial f(u).$$

Hence it follows from Lemma 2.1 that

$$\begin{aligned} \eta\|z-u\| &= \left\langle \frac{\eta(z-u)}{\|z-u\|}, z-u \right\rangle \\ &\leq \sigma^*(z-u | \partial f(u)) = df(u)(z-u). \end{aligned}$$

Combining the above inequality with (4.13), we arrive at characterization (iv) of local weak sharp minima in Theorem 4.2 and thus complete the proof of this theorem.  $\square$

Note that our Theorems 4.1–4.3 for the general class of inf-differentiable functions can be treated as far-going extensions of the results in [40, Theorems 3.1, 3.2 and Proposition 3.1] obtained for functions of the lower- $C^1$  type

$$f(x) = \max_{y \in Y} [\varphi(x, y)]_+, \quad (4.14)$$

where  $\varphi \in C^1$  and  $Y$  is compact. Although the function  $f$  in (4.14) is not given exactly in form (3.10) due to the additional nonsmooth operation  $[z]_+ := \max\{0, z\}$ , its inf-differentiability readily follows from the proof of Theorem 3.8(ii).

The last result of this section establishes mixed characterizations of local weak sharp minima of another type that extend, in particular, sufficient conditions for such minimizers developed in [19] in a special setting.

**Theorem 4.4 (mixed characterizations of local weak sharp minima, II).** *Let  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be locally Lipschitzian around  $x$  and subdifferentially regular at this point, and let  $f$  be inf-differentiable at  $x \in \text{dom } f$  relative to the level set  $L_f(x)$ . Then  $x$  is a local weak sharp minimum of  $f$  if and only if the following conditions hold:*

- (a)  $\widehat{N}_{L_f(x)}(x) = \text{cl cone}(\partial f(x))$ .
- (b) *There exist two positive scalars  $\eta$  and  $\delta$  such that*

$$\eta \text{dist}(w, \widehat{T}_{L_f(x)}(u)) \leq df(u)(w) \text{ for all } u \in L_f(x) \cap B(x, \delta) \setminus \{x\} \text{ and } w \in \mathbb{R}^n.$$

**Proof.** First we verify the equivalence between conditions (a) and (b) of this theorem and conditions (a) and (b) given in Theorem 4.3, respectively. It is easy to see that

$$\begin{aligned} T_{L_f(x)}(x) &= \{w \in \mathbb{R}^n \mid df(x)(w) \leq 0\} \\ &= \{w \in \mathbb{R}^n \mid \sigma^*(w | \partial f(x)) \leq 0\} \\ &= \{w \in \mathbb{R}^n \mid \langle w, z \rangle \leq 0 \text{ for all } z \in \partial f(x)\} \\ &= (\partial f(x))^0. \end{aligned} \quad (4.15)$$

By the first duality correspondence in (2.5) and [35, Theorem 6.28] we have

$$\widehat{N}_{L_f(x)}(x) = (T_{L_f(x)})^0 = ((\partial f(x))^0)^0 = \text{cl cone}(\partial f(x)), \quad (4.16)$$

where the last equality follows from [35, Corollary 6.21]. We have from (4.15) that the tangent cone  $T_{L_f(x)}(x)$  is convex under the assumptions made. This implies by (4.16) that

$$(\widehat{N}_{L_f(x)}(x))^0 = (T_{L_f(x)}(x)^0)^0 = T_{L_f(x)}(x).$$

Therefore the cones  $\widehat{N}_{L_f(x)}(x)$  and  $T_{L_f(x)}$  are polar to each other, which thus justifies the fulfillment of the claimed equivalence.

Let us next prove the equality

$$\eta \text{dist}(w, \widehat{T}_{L_f(x)}(u)) = \sigma^*(w | \overline{N}_{L_f(x)}(x) \cap \eta B) \text{ for all } w \in \mathbb{R}^n.$$

which is, by the second duality correspondence in (2.5), a particular case of the relationship

$$\eta \text{dist}(w, K) = \sigma^*(w | K^0 \cap \eta B) \quad (4.17)$$

held for any closed and convex cone  $K$ . To check the latter, consider the singleton  $w = P_K(w) + P_{K^0}(w)$  due the closedness and convexity of  $K$  and get by [35, Exercise 12.22] that

$$\begin{aligned} \sigma^*(w | K^0 \cap \eta B) &= \sigma^*(P_K(w) + P_{K^0}(w) | K^0 \cap \eta B) \\ &= \sigma^*(P_K(w) | K^0 \cap \eta B) + \sigma^*(P_{K^0}(w) | K^0 \cap \eta B) \\ &= \eta \|P_{K^0}(w)\| = \eta \|w - P_K(w)\| = \eta \text{dist}(w, K). \end{aligned}$$

Specifying (4.17) for  $K = \widehat{T}_{L_f(x)}(u)$ , we arrive at the equivalences

$$\begin{aligned} &\overline{N}_{L_f(x)}(u) \cap \eta B \subset \partial f(u), && u \in L_f(x) \cap B(x, \delta) \setminus \{x\}, \\ \iff &\sigma^*(w | \overline{N}_{L_f(x)}(u) \cap \eta B) \leq \sigma^*(w | \partial f(u)), && u \in L_f(x) \cap B(x, \delta) \setminus \{x\}, w \in \mathbb{R}^n, \\ \iff &\eta \text{dist}(w, \widehat{T}_{L_f(x)}(u)) \leq df(u)(w), && u \in L_f(x) \cap B(x, \delta) \setminus \{x\}, w \in \mathbb{R}^n, \text{inequ} \end{aligned}$$

where the first equivalence is due to [34, Corollary 13.1.1]. This completes the proof.  $\square$

## 5 Applications to Semi-Infinite Programming

In this section we develop some applications of the newly obtained characterizations of local weak sharp minima to problems of semi-infinite programming (SIP).

Consider the following canonical SIP problem with inequality constraints:

$$\text{minimize } f(x) \text{ subject to } g(x, s) \leq 0 \text{ for all } s \in \Omega, \quad (5.1)$$

where  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  and  $g : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$  are continuously differentiable functions, and where  $\Omega \subset \mathbb{R}^m$  is a compact set. Problems of this type arise in various fields of mathematics, engineering, and applied science. Among such areas we mention approximation theory, optimal control, resource allocation in decentralized systems, decision making under competition, optimum filter design in signal processing, control of water resources, etc. For more details and discussions we refer the reader to, e.g., [10, 11, 18, 20, 32] and the bibliographies therein.

Denote by  $X$  the set of feasible solutions to the SIP problem (5.1). Adopting the general concept (1.1) of weak sharp minimizers for unconstrained optimization problems and talking

into account the specific structure of the constraints in (5.1), we say that  $x$  is a *SIP local weak sharp minimum* for (5.1) if  $x \in X$  and there are  $\eta, \delta > 0$  such that

$$\eta \operatorname{dist}(z, L_f(x) \cap X) \leq f(z) - f(x) + \int_{\Omega} [g(z, s)]_+ d\mu(s) \text{ for all } z \in B(x, \delta), \quad (5.2)$$

where  $\mathcal{F}$  is the Borel  $\sigma$ -algebra on  $\Omega$ , and where  $\mu$  is a finite measure defined on a measurable space  $(\Omega, \mathcal{F})$  and satisfying the support condition

$$\Omega = \operatorname{supp}(\Omega) := \left\{ w \in \Omega \mid \int_{B(w, \delta) \cap \Omega} d\mu(s) > 0 \text{ for all } \delta > 0 \right\}.$$

Observe that our approach to handle constraints in the SIP framework (5.2) of local weak sharp minima seems to be natural and convenient for the subsequent analysis being somewhat different from the one in [40], where the function

$$\psi(x) := \int_{\Omega} \phi_s(x) d\mu(s) \text{ with } \phi_s(x) := [g(x, s)]_+, \quad x \in \mathbb{R}^n \quad (5.3)$$

in (5.2) is replaced by  $\max_{s \in \Omega} [g(x, s)]_+$ .

The following result establishes all the properties of the function  $\psi$  in (5.3) needed for applications of our characterizations of local weak sharp minima obtained in Section 4.

**Lemma 5.1 (inf-differentiability and subdifferential regularity of integral functions associated with SIP).** *Under the assumptions made the integral function  $\psi$  defined in (5.3) is Lipschitz continuous, directionally differentiable, subdifferentially regular, and inf-differentiable on the feasible set  $X$ .*

**Proof.** The Lipschitz continuity and directional differentiability of  $\psi$  at every  $x \in X$  follow from [1, Proposition 5.108], where the formula

$$\psi'(x; w) = \int_{\Omega} \phi'_s(x; w) d\mu(s), \quad w \in \mathbb{R}^n, \quad (5.4)$$

is justified. Furthermore, we derive from [12, Theorem 2.7.3] that  $\phi$  and then  $\psi$  in (5.3) are directionally differentiable at each  $x \in X$ . Thus  $\psi$  is subdifferentially regular at  $x$  due its Lipschitz continuity. It is easy to observe from the construction of  $\phi$  in (5.3) that the directional derivative of  $\phi$  is computed by

$$\phi'_s(x; w) = \begin{cases} [\langle \nabla_x g(x, s), w \rangle]_+ & \text{if } s \in \Omega_0(x), \\ 0 & \text{if } s \in \Omega_-(x), \\ \langle \nabla_x g(x, s), w \rangle & \text{if } s \in \Omega_+(x), \end{cases} \quad (5.5)$$

where the sets  $\Omega_0(x)$ ,  $\Omega_-(x)$ , and  $\Omega_+(x)$  are given by

$$\begin{aligned} \Omega_0(x) &:= \{s \in \Omega \mid g(x, s) = 0\}, \\ \Omega_-(x) &:= \{s \in \Omega \mid g(x, s) < 0\}, \\ \Omega_+(x) &:= \{s \in \Omega \mid g(x, s) > 0\}, \end{aligned}$$

respectively. Defining further a nonnegative number  $C$  by the maximum

$$C := \max_{(x',s) \in B(x,\delta) \times \Omega} \|\nabla_x g(x', s)\|$$

over the compact set  $B(x; \delta) \times \Omega$ , we get the estimates

$$\begin{aligned} \|\phi_s(z) - \phi_s(u)\| &= \|[g(z, s)]_+ - [g(u, s)]_+\| \leq \|g(z, s) - g(u, s)\| \\ &\leq \int_0^1 \|\nabla_x g(u + t(z-u), s)\| \cdot \|z - u\| dt \\ &\leq KC\|z - u\| \text{ for any } z, u \in B(x, \delta), \end{aligned} \quad (5.6)$$

where the first inequality is due to the nonexpansivity of the projection mapping by [35, Corollary 12.20]. It follows from (5.3)-(5.5) that  $\|\phi'_s(x; w)\| \leq \|\nabla_x g(x, s)\| \cdot \|w\|$ , which together with (5.6) yields the estimates that

$$\begin{aligned} \phi_s(z) - \phi_s(u) - \phi'_s(u; z - u) &\geq -\|\phi_s(z) - \phi_s(u)\| - \|\phi'_s(u; z - u)\| \\ &\geq -2C\|z - u\|. \end{aligned}$$

Now applying the classical Fatou theorem gives us the inequality

$$\liminf_{\substack{z \rightarrow x \\ z \neq u, u \rightarrow x}} \int_{\Omega} \frac{\phi_s(z) - \phi_s(u) - \phi'_s(u; z - u)}{\|z - u\|} d\mu(s) \geq \int_{\Omega} \liminf_{\substack{z \rightarrow x \\ z \neq u, u \rightarrow x}} \frac{\phi_s(z) - \phi_s(u) - \phi'_s(u; z - u)}{\|z - u\|} d\mu(s).$$

Consequently we have the relationships

$$\begin{aligned} \liminf_{\substack{z \rightarrow x \\ z \neq u, u \rightarrow x}} \frac{\psi(z) - \psi(u) - \psi'(u; z - u)}{\|z - u\|} &= \liminf_{z \neq u, u \rightarrow x} \int_{\Omega} \frac{\phi_s(z) - \phi_s(u) - \phi'_s(u; z - u)}{\|z - u\|} d\mu(s) \\ &\geq \int_{\Omega} \liminf_{\substack{z \rightarrow x \\ z \neq u, u \rightarrow x}} \frac{\phi_s(z) - \phi_s(u) - \phi'_s(u; z - u)}{\|z - u\|} d\mu(s) \\ &\geq 0, \end{aligned} \quad (5.7)$$

where the last inequality follows from the inf-differentiability of  $\phi_s$  by Theorem 3.8(ii). On the other hand, it follows from Theorem 3.2 and the Lipschitz continuity of  $\psi$  that

$$\liminf_{\substack{z \rightarrow x \\ z \neq u, u \rightarrow x}} \frac{\psi(z) - \psi(u) - \psi'(u; z - u)}{\|z - u\|} \leq \liminf_{z \rightarrow x} \frac{\psi(z) - \psi(x) - \psi'(x; z - x)}{\|z - x\|} = 0,$$

which together with (5.7) ensures the inf-differentiability of  $\psi$  at each  $x \in X$  and thus completes the proof of the lemma.  $\square$

To apply the results developed in the previous section to SIP (5.1), we first need to obtain explicit descriptions of the subderivative  $d\psi$  and the subdifferential  $\partial\psi$  of the integral function (5.3) at feasible points. Since we know that  $\psi$  is directionally regular, it suffices to obtain a formula for its directional derivative that implies the corresponding subdifferential representation.

**Lemma 5.2 (computing directional derivatives and subdifferentials of integral functions).** *Let  $\psi$  be defined by (5.3). Then we have the following representations:*

$$\psi'(x; w) = \int_{\Omega_+(x)} \langle \nabla_x g(x, s), w \rangle d\mu(s) + \int_{\Omega_0(x)} [\langle \nabla_x g(x, s), w \rangle]_+ d\mu(s), \quad w \in \mathbb{R}^n, \quad (5.8)$$

$$\partial\psi(x) = \left\{ \int_{\Omega_+(x)} \nabla_x g(x, s) d\mu(s) + \int_{\Omega_0(x)} \lambda(s) \nabla_x g(x, s) d\mu(s) \mid \lambda \in \wedge \right\}, \quad (5.9)$$

where  $\wedge$  denotes the set of all summable mappings from  $\Omega_0(x)$  to  $[0, 1]$ .

**Proof.** Substituting (5.5) into (5.4) yields the claimed expression (5.8) of the directional derivative. Invoking further Lemma 2.1, we get the equality

$$\psi'(x; w) = \sigma^*(w \mid \partial\psi(x)) \text{ for all } w \in \mathbb{R}^n,$$

which together with (5.8) implies the subdifferential formula (5.9).  $\square$

Now we are ready to establish primal, dual, and mixed characterizations of SIP local weak sharp minima (5.2) or semi-infinite programs of type (5.1).

**Theorem 5.3 (characterizations of SIP local weak sharp minima).** *Under the standing assumptions of this section the following statements are equivalent for every  $x \in X$ :*

(i)  $x$  is a SIP local weak sharp minimum of (5.1).

(ii) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$N_{L_f(x) \cap X}(u) \cap \eta B \subset \nabla f(u) + \left\{ \int_{\Omega_+(x)} \nabla_x g(x, s) d\mu(s) + \int_{\Omega_0(x)} \lambda(s) \nabla_x g(x, s) d\mu(s) \mid \lambda \in \wedge \right\}$$

for all  $u \in L_f(x) \cap X \cap B(x, \delta)$ .

(iii) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \text{dist}(w, T_{L_f(x) \cap X}(u)) \leq \langle \nabla f(u), w \rangle + \int_{\Omega_+(x)} \langle \nabla_x g(x, s), w \rangle d\mu(s) + \int_{\Omega_0(x)} [\langle \nabla_x g(x, s), w \rangle]_+ d\mu(s)$$

for all  $u \in L_f(x) \cap X \cap B(x, \delta)$ .

(iv) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\bar{N}_{L_f(x) \cap X}(u) \cap \eta B \subset \nabla f(u) + \left\{ \int_{\Omega_+(x)} \nabla_x g(x, s) d\mu(s) + \int_{\Omega_0(x)} \lambda(s) \nabla_x g(x, s) d\mu(s) \mid \lambda \in \wedge \right\}$$

for all  $u \in L_f(x) \cap X \cap B(x, \delta) \setminus \{x\}$  and

$$T_{L_f(x) \cap X}(x) = \left\{ w \mid \langle \nabla f(x), w \rangle + \int_{\Omega_+(x)} \langle \nabla_x g(x, s), w \rangle d\mu(s) + \int_{\Omega_0(x)} [\langle \nabla_x g(x, s), w \rangle]_+ d\mu(s) = 0 \right\}.$$

(v) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \text{dist}(w, \widehat{T}_{L_f(x) \cap X}(u)) \leq \langle \nabla f(u), w \rangle + \int_{\Omega_+(x)} \langle \nabla_x g(x, s), w \rangle d\mu(s) + \int_{\Omega_0(x)} [\langle \nabla_x g(x, s), w \rangle]_+ d\mu(s)$$

for all  $u \in L_f(x) \cap X \cap B(x, \delta) \setminus \{x\}$  and

$$\widehat{N}_{L_f(x) \cap X}(x) = \text{cl cone} \left( \nabla f(x) + \left\{ \int_{\Omega_+(x)} \nabla_x g(x, s) d\mu(s) + \int_{\Omega_0(x)} \lambda(s) \nabla_x g(x, s) d\mu(s) \mid \lambda \in \wedge \right\} \right).$$



**Proof.** The smoothness of  $f$  and the properties of  $\psi$  established in Lemma 5.1 ensure that  $f + \psi$  is Lipschitz continuous, subdifferentially regular, and inf-differentiable on  $X$ . By the calculus rules of [25, Proposition 1.107(ii)] and [35, Corollary 10.9], we have

$$\partial(f + \psi)(x) = \nabla f(x) + \partial\psi(x),$$

$$d(f + \psi)(x)(w) = df(x)(w) + d\psi(x)(w) = \langle \nabla f(x), w \rangle + \psi'(x; w) \text{ for all } w \in \mathbb{R}^n.$$

Applying finally Theorems 4.1–4.4 to the sum  $f + \psi$  yields the desired results and thus completes the proof of the theorem.  $\square$

## 6 Applications to Semi-Infinite Complementarity Problems

In this section we apply our general characterizations of local weak sharp minima obtained in Section 4 to the following *semi-infinite complementarity problems* (SICP) defined as follows: find a vector  $x \in \mathbb{R}^n$  such that

$$x \geq 0, \quad F(x, s) \geq 0, \quad x^T F(x, s) = 0 \text{ for all } s \in \Omega, \quad (6.1)$$

where  $\Omega \subset \mathbb{R}^m$ ,  $F: \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}^m$ , and “ $T$ ” stands for transposition. In contrast to classical complementarity problems, the number of complementarity constraints in (6.1) may be infinite while the decision vector  $x$  is finite-dimensional.

Recall that a function  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$  is an *NCP function* if it has the property

$$\phi(a, b) = 0 \iff [a \geq 0, b \geq 0, ab = 0].$$

For recent years NCP functions have been used as a powerful tool of dealing with classical complementarity problems since they allow us to reformulate complementarity problems as either equations or minimization problems. Such formulations are very beneficial for both analytical and computational purposes. Indeed, powerful developments from classical analysis of systems of equations can be applied to treat classical complementarity problems for justifying the existence of solutions and for analyzing these solution properties. Furthermore, efficient algorithms for solving equations and optimization problems can be applied and extended to solve classical complementarity problems; see, e.g., [15].

Similar to classical complementarity problems, we obtain the following equivalent reformulation of SICP as a system of equations:

$$x \in S \iff \Phi(x, s) = 0 \text{ for all } s \in \Omega, \quad (6.2)$$

where the vector function  $\Phi: \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}^n$  is defined by

$$\Phi(x, s) := \begin{pmatrix} \phi(x_1, F_1(x, s)) \\ \vdots \\ \phi(x_n, F_n(x, s)) \end{pmatrix}.$$

Using the SICP formulation (6.2), it is natural to choose a *residual function* as

$$r(x) := \frac{1}{2} \max_{s \in \Omega} \|\Phi(x, s)\|^2. \quad (6.3)$$

Indeed, to solve the semi-infinite complementarity problem clearly means to find a root of the equation  $r(x) = 0$ , or equivalently, to find an optimal solution of the following minimization problem with optimal objective value equal to *zero*:

$$\min_{x \in \mathbb{R}^n} r(x) = \frac{1}{2} \max_{s \in \Omega} \|\Phi(x, s)\|^2. \quad (6.4)$$

Noting that the latter minimization problem is a typical semi-infinite minimax programming problem [29]; it offers another explanation for labeling problem (6.1) as a semi-infinite complementarity problem in agreement with SIP.

Denote by  $S$  the solution set to (6.1). We say that a residual function  $r$  has a *local error bound* at  $x \in S$  if there exist two positive scalars  $\eta$  and  $\delta$  such

$$\eta \text{dist}(z, S) \leq r(z) \text{ for all } z \in B(x, \delta), \quad (6.5)$$

which is equivalent to saying that  $x \in S$  is a local weak sharp minimum of the residual function  $r$  from (6.3) since the corresponding level set is

$$L_r(x) = \{z \in \mathbb{R}^n \mid r(z) = r(x) = 0\} = S.$$

From now on we concentrate in the above scheme on the *Fischer-Burmeister function*

$$\phi_{\text{FB}}(a, b) := \sqrt{a^2 + b^2} - (a + b) \quad (6.6)$$

and denote the corresponding functions  $\Phi$  in (6.2) and  $r$  in (6.3) by  $\Phi_{\text{FB}}$  and  $r_{\text{FB}}$ , respectively. In fact, similar results can be obtained via other residual functions whenever the square  $\phi^2$  is continuously differentiable. As usual,  $J_x F$  stands for the classical Jacobian matrix of  $F$  with respect to  $x$ .

The next lemma summarizes basic properties of the auxiliary functions associated with (6.1) and (6.6) needed in what follows.

**Lemma 6.1 (properties of residual functions).** *Let  $\Omega$  be compact, and let  $F$  be continuously differentiable on  $\mathbb{R}^n \times \Omega$  in the framework of SICP in (6.1). Then we have the following assertions fulfilled:*

- (i) *The function  $\frac{1}{2} \|\Phi_{\text{FB}}(x, s)\|^2$  is continuous differentiable on  $\mathbb{R}^n \times \Omega$  and its partial gradient with respect to  $x$  is equal to  $H^T \Phi_{\text{FB}}(x, s)$  for every  $H \in \partial_x \Phi_{\text{FB}}(x, s)$ , where  $\partial_x \Phi_{\text{FB}}$  is the subdifferential of  $\Phi_{\text{FB}}$  with respect to the first variable satisfying*

$$\partial_x \Phi_{\text{FB}}(x, s) \subset \mathcal{D}_a(x, s) + \mathcal{D}_b(x, s) J_x F(x, s),$$

*and where  $\mathcal{D}_a(x, s)$  and  $\mathcal{D}_b(x, s)$  are the sets of  $n \times n$  diagonal matrices  $\text{diag}(a_1(x, s), \dots, a_n(x, s))$  and  $\text{diag}(b_1(x, s), \dots, b_n(x, s))$ , respectively, with*

$$(a_i(x, s), b_i(x, s)) \begin{cases} = \frac{(x_i, F_i(x, s))}{\sqrt{x_i^2 + F_i^2(x, s)}} - (1, 1) & \text{if } (x_i, F_i(x, s)) \neq 0, \\ \in B(0, 1) - (1, 1) & \text{if } (x_i, F_i(x, s)) = 0. \end{cases}$$

(ii) The residual function  $r_{\text{FB}}(x)$  in (6.3) associated with (6.6) is Lipschitz continuous, semismooth, and inf-differentiable on  $\mathbb{R}^n$ .

(iii) The subderivative and subdifferential of  $r_{\text{FB}}$  at  $x$  are computed by

$$dr_{\text{FB}}(x)(w) = \max_{\substack{s \in \Omega(x) \\ H \in \partial_x \Phi_{\text{FB}}(u, s)}} \langle H^T \Phi_{\text{FB}}(x, s), w \rangle \text{ for all } w \in \mathbb{R}^n,$$

$$\partial r_{\text{FB}}(x) = \text{co}\{H^T \Phi_{\text{FB}}(x, s) \mid s \in \Omega(x) \text{ and } H \in \partial_x \Phi_{\text{FB}}(u, s)\},$$

$$\text{with } \Omega(x) := \{s \in \Omega \mid \|\Phi_{\text{FB}}(x, s)\|^2 = r_{\text{FB}}(x)\}.$$

**Proof.** Assertion (i) follows from [15, Proposition 9.1.4]. In assertion (ii) the Lipschitz continuity comes from [35, Proposition 9.10], the semismoothness is due to [24, Theorem 2], and the inf-differentiability follows from Theorem 3.8(ii) above since the residual  $r$  is the pointwise supremum of a family of smooth functions over a compact index set. Finally, assertion (iii) follows from [35, Theorem 10.31].  $\square$

Based on the results obtained in Section 4 and the properties of Lemma 6.1, we are now ready to efficiently characterize semi-infinite complementary problems whose residuals possess a local error bound.

**Theorem 6.2 (local error bounds for SICP).** *Consider a SICP (6.1) with a compact set  $\Omega \subset \mathbb{R}^n$  and a continuously differentiable function  $F: \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}^m$ . Let  $x \in S$  be a solution to (6.1). Then the following assertions are equivalent:*

(i) The residual  $r_{\text{FB}}$  has a local error bound at  $x$  in the sense of (6.5).

(ii) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$N_S(u) \cap \eta B \subset \text{co}\{H^T \Phi_{\text{FB}}(u, s) \mid s \in \Omega(u) \text{ and } H \in \partial_x \Phi_{\text{FB}}(u, s)\} \text{ as } u \in S \cap B(x, \delta).$$

(iii) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \text{dist}(w, T_S(u)) \leq \max_{\substack{s \in \Omega(u) \\ H \in \partial_x \Phi_{\text{FB}}(u, s)}} \langle H^T \Phi_{\text{FB}}(u, s), w \rangle \text{ for all } u \in S \cap B(x, \delta), w \in \mathbb{R}^n.$$

(iv) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\overline{N}_S(u) \cap \eta B \subset \text{co}\{H^T \Phi_{\text{FB}}(u, s) \mid s \in \Omega(u) \text{ and } H \in \partial_x \Phi_{\text{FB}}(u, s)\}$$

for all  $u \in S \cap B(x, \delta) \setminus \{x\}$  and

$$T_S(x) = \{w \in \mathbb{R}^n \mid \max_{\substack{s \in \Omega(x) \\ H \in \partial_x \Phi_{\text{FB}}(x, s)}} \langle H^T \Phi_{\text{FB}}(x, s), w \rangle \leq 0\}.$$

(v) There exist two positive scalars  $\eta$  and  $\delta$  such that

$$\eta \text{dist}(w, T_S(u)) \leq \max_{\substack{s \in \Omega(u) \\ H \in \partial_x \Phi_{\text{FB}}(u, s)}} \langle H^T \Phi_{\text{FB}}(u, s), w \rangle$$

for all  $u \in S \cap B(x, \delta) \setminus \{x\}$  and  $w \in \mathbb{R}^n$ , and we have

$$\widehat{N}_S(x) = \text{co}\{H^T \Phi_{\text{FB}}(x, s) \mid s \in \Omega(x), H \in \partial_x \Phi_{\text{FB}}(u, s)\}.$$

**Proof.** We know from the previous discussions that the residual  $r$  has a local error bound at  $x \in S$  if and only if  $x$  is a local weak sharp minima of  $r$ . Thus applying Theorems 4.1–4.4 to  $r_{FB}$  and employing the explicit structures of the subderivative and the subdifferential obtained in Lemma 6.1, we justify all the assertions of the theorem.  $\square$

## 7 Concluding Remarks

In this paper we introduce a new class of nonsmooth functions under the name of *inf-differentiable* functions, which is sufficiently broad to include many remarkable collections of nonsmooth functions important in variational analysis, optimization, and their numerous applications. One of the most significant applications of inf-differential functions provided in the paper is that local weak sharp minima can be completely characterized for them via primal and dual constructions of generalized differentiation. Among the main purposes of our future research are developing calculus results for this remarkable class of functions and their further applications to various optimization and optimization-related problems.

## References

- [1] J. F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
- [2] J. F. Bonnans, A. D. Ioffe, *Quadratic growth and stability in convex programming problems with multiple solutions*, J.Convex Anal., 2:41–57, 1995.
- [3] J. V. Burke, S. Deng, *Weak sharp minima revisited, I: Basic theory*, Control and Cybernetics, 31:439–469, 2002.
- [4] J. V. Burke, S. Deng, *Weak sharp minima revisited, II: Application to linear regularity and error bounds*, Math. Program., 104:235–261, 2005.
- [5] J. V. Burke, S. Deng, *Weak sharp minima revisited, III: Error bounds for differentiable convex inclusions*, Math. Program., 116:37–56, 2009.
- [6] J. V. Burke, M. C. Ferris, *Weak sharp minima in mathematical programming*, SIAM J. Control Optim., 31:1340–1359, 1993.
- [7] J. V. Burke, M. C. Ferris, *A Gauss-Newton method for convex composite optimization*, Math. Program., 71:179–194, 1995.
- [8] J. V. Burke, A. S. Lewis, M. L. Overton, *Optimizing matrix stability*, Proc. Amer. Math. Soc., 129: 1635–1642, 2000.
- [9] J. V. Burke, J. J. Moré, *On the identification of active constraints*, SIAM J. Numer. Anal., 25:1197–1211.
- [10] M. J. Cánovas, M. A. López, B. S. Mordukhovich, J. Parra, *Variational analysis in semi-infinite and infinite programming, I: Stability of linear inequality systems of feasible solutions*, SIAM J. Optim., 20:1504–1526, 2009.

- [11] M. J. Cánovas, M. A. López, B. S. Mordukhovich, J. Parra, *Variational analysis in semi-infinite and infinite programming, II: Necessary optimality conditions*, SIAM J. Optim., 20:2788–2806, 2010.
- [12] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
- [13] L. Cromme, *Strong uniqueness: A far reaching criterion for the convergence of iterative procedures*, Numer. Math., 29:179–193, 1978.
- [14] S. Deng, *Some remarks on finite termination of descent methods*, Pacif. J. Optim., 1:31–37, 2005.
- [15] F. Facchinei, J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, I,II*, Springer, New York, 2003.
- [16] M. C. Ferris, *Weak Sharp Minima and Penalty Functions in Mathematical Programming*, PhD Dissertation, University of Cambridge, Cambridge, UK, 1988.
- [17] M. C. Ferris, *Finite termination of the proximal point algorithm*. Math. Program, 50:359–366, 1991.
- [18] M. A. Goberna, M. A. López, *Linear Semi-Infinite Optimization*, Wiley, Chichester, 1998.
- [19] R. Henrion, J. Outrata, *Calmness of constraint systems with applications*, Math. Program., 104:437–464, 2005.
- [20] R. Hettich, K. O. Kortanek, *Semi-infinite programming: Theory, methods, and applications*, SIAM Review, 35:380–429, 1993.
- [21] C. Li, B. S. Mordukhovich, J. Wang, J. C. Yao, *Weak sharp minima on Riemannian manifolds*, to appear in SIAM J. Optim., 2011.
- [22] C. Ling, L. Qi, G. Zhou, L. Caccetta, *The  $SC^1$  property of an expected residual function arising from stochastic complementarity problems*, Oper. Res. Lett., 36:456–460, 2008.
- [23] P. Marcotte, D. L. Zhu, *Weak sharp solutions of variational inequalities*, SIAM J. Optim., 9:179–189, 1998.
- [24] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., 15:957–972, 1977.
- [25] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, 2006.
- [26] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, II: Applications*, Springer, 2006.
- [27] B. S. Mordukhovich, N. M. Nam, N. D. Yen, *Frechet subdifferential calculus and optimality conditions in nondifferentiable programming*, Optimization, 55:685–708, 2006.

- [28] K. F. Ng, X. Y. Zheng, *Global weak sharp minima on Banach spaces*, SIAM J. Control Optim., 41:1868–1885, 2003.
- [29] E. Polak, *Optimization: Algorithms and Consistent Approximation*, Springer, New York, 1997.
- [30] B. T. Polyak, *Introduction to Optimization*, Optimization Software, New York, 1987.
- [31] L. Qi, A. Shapiro, C. Ling, *Differentiability and semismoothness properties of integral functions and their applications*, Math. Program., 102:223–248, 2005.
- [32] R. Reemtsen, J. J. Rückmann (eds.), *Semi-Infinite Programming*, John Wiley, New York, 1998.
- [33] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity*, Math. Program., 30:45–66, 1987.
- [34] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
- [35] R. T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
- [36] A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl., 66:477–487, 1990.
- [37] M. Studniarski, D. E. Ward, *Weak sharp minima: Characterizations and sufficient conditions*, SIAM J. Control Optim., 38:219–236, 1999.
- [38] Z. L. Wu, S. Y. Wu, *Weak sharp solutions of variational inequalities in Hilbert spaces*, SIAM J. Optim., 14:1011–1027, 2004.
- [39] Z. L. Wu, J. J. Ye, *Sufficient conditions for error bounds*, SIAM J. Optim., 12:421–435, 2001.
- [40] X. Y. Zheng, X. Q. Yang, *Weak sharp minima for semi-infinite optimization problems with applications*, SIAM J. Optim., 18:573–588, 2007.