Mathematical Programming (2021) 189:217–247
https://doi.org/10.1007/s10107-020-01592-0

FULL LENGTH PAPER

Series B

Subdifferential of the supremum function: moving back and forth between continuous and non-continuous settings

R. Correa[1,2] · A. Hantoute[3,4] · M. A. López[4,5]

Received: 23 January 2020 / Accepted: 3 November 2020 / Published online: 27 November 2020

Abstract
In this paper we establish general formulas for the subdifferential of the pointwise supremum of convex functions, which cover and unify both the compact continuous and the non-compact non-continuous settings. From the non-continuous to the continuous setting, we proceed by a compactification-based approach which leads us to problems having compact index sets and upper semi-continuously indexed mappings, giving rise to new characterizations of the subdifferential of the supremum by means of upper semicontinuous regularized functions and an enlarged compact index set. In the opposite sense, we rewrite the subdifferential of these new regularized functions by using the original data, also leading us to new results on the subdifferential of the supremum. We give two applications in the last section, the first one concerning the nonconvex Fenchel duality, and the second one establishing Fritz-John and KKT conditions in convex semi-infinite programming.

Keywords Supremum of convex functions · Subdifferentials · Stone–Čech compactification · Convex semi-infinite programming · Fritz-John and KKT optimality conditions

Mathematics Subject Classification 46N10 · 52A41 · 90C25

1 Introduction

In this paper we deal with the characterization of the subdifferential of the pointwise supremum $f := \sup_{t \in T} f_t$ of a family of convex functions $f_t : X \to \mathbb{R} \cup \{\pm\infty\}$,

---

Dedicated by his co-authors to Prof. Macro A. López on his 70th birthday

Research supported by CONICYT (Fondecyt 1190012 and 1190110), Proyecto/Grant PIA AFB-170001, MICIU of Spain and Universidad de Alicante (Grant Beatriz Galindo BEAGAL 18/00205), and Research Project PGC2018-097960-B-C21 from MICINN, Spain. The research of the third author is also supported by the Australian ARC - Discovery Projects DP 180100602.

Extended author information available on the last page of the article

$t \in T$, with T being an arbitrary nonempty set, defined on a separated locally convex space X. We obtain new characterizations which allow us to unify both the *compact continuous* and the *non-compact non-continuous* setting ([8,9,27,30], etc.). The first setting relies on the following standard conditions in the literature of convex analysis and non-differentiable semi-infinite programming:

$$T \text{ is compact and the mappings } f_{(\cdot)}(z),\ z \in X, \text{ are upper semi-continuous.}$$

In the other framework, called the *non-compact non-continuous* setting, we do not assume the above conditions. In other words (see, e.g., [14,15,18,21,29–31], etc.):

$$T \text{ is an arbitrary set, possibly infinite and without any prescribed topology,}$$
$$\text{and no requirement is imposed on the mappings } f_{(\cdot)}(z).$$

Going from the non-continuous to the continuous setting, we follow an approach based on the Stone–Čech compactification of the index set T. At the same time, we build an appropriate enlargement of the original family $f_t,\ t \in T$, which ensures the fulfillment of the upper semi-continuity property required in the compact setting. Since the new setting is naturally compact, by applying the results in [8,9], we obtain new characterizations given in terms of the exact subdifferential at the reference point of the new functions and the extended active set. In this way, we succeed in unifying both settings. In [10], we gave the first steps in this direction, using compactification arguments, but in the current paper we go further into the subject with some enhanced formulas.

To move in the other direction, we rewrite the subdifferential of these new regularizing functions in terms of the original data, and this also leads us to new results on the subdifferential of the supremum. In this last case, the characterizations are given upon limit processes on the ε-subdifferentials at the reference point of the almost-active original functions. These limit processes also involve approximations by finite-dimensional sections of the domain of the supremum function.

The main results of this paper are applied to derive formulas for the subdifferential of the conjugate function [3–5,23]. Our approach permits simple proofs of these results, with the aim of relating the solution set of a nonconvex optimization problem and its convexified relaxation. Additionally, our results give rise to new Fritz-John and KKT conditions in convex semi-infinite programming.

The paper is organized as follows. After a short section introducing the notation, in Sect. 3 we present some preliminary results in the continuous setting. In Sect. 4 we apply our compactification approach to obtain, in Theorem 4, a first characterization of the subdifferential of the supremum. Such a theorem constitutes an improved version of the main result in [10], as the requirement of equipping T with a completely regular topology is eliminated. Theorem 4 is enhanced in Sect. 5, allowing for a more natural interpretation of the regularized functions. The main result in Sect. 6 is Theorem 12, involving only the ε-subdifferentials of the original data functions. This theorem, whose proof is based on Lemmas 10 and 11, is crucial in the proposed approach to move from the continuous to the non-continuous setting. Finally, in Sect. 7, we give two applications. The first one addresses the extension of the classical Fenchel duality to nonconvex functions, and the second one establishes Fritz-John and KKT optimality conditions for convex semi-infinite optimization.

2 Notation

Let X be a (real) separated locally convex space, with its topological dual X^* endowed with the w^*-topology. By $\mathcal{N}_X$ ($\mathcal{N}_{X^*}$) we denote the family of closed, convex, and balanced neighborhoods of the origin in X (X^*), also called θ-neighborhoods. The spaces X and X^* are paired in duality by the bilinear form $(x^*, x) \in X^* \times X \mapsto \langle x^*, x\rangle := \langle x, x^*\rangle := x^*(x)$. The zero vectors in X and X^* are both denoted by θ. We use the notation $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ and $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, and adopt the convention $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$.

Given two nonempty sets A and B in X (or in X^*), we define the *algebraic* (or *Minkowski*) *sum* by

$$A + B := \{a + b : \ a \in A, \ b \in B\}, \quad A + \emptyset = \emptyset + A = \emptyset. \tag{1}$$

By $\operatorname{co}(A)$, $\operatorname{cone}(A)$, and $\operatorname{aff}(A)$, we denote the *convex,* the *conical convex,* and the *affine hulls* of A, respectively. Moreover, $\operatorname{int}(A)$ is the *interior* of A, and $\operatorname{cl} A$ and $\overline{A}$ are indistinctly used for denoting the *closure* of A (unless otherwise specified, the topology considered on X^* is the weak* topology). We use $\operatorname{ri}(A)$ to denote the (topological) *relative interior* of A (i.e., the interior of A in the topology relative to $\operatorname{aff}(A)$ if $\operatorname{aff}(A)$ is closed, and the empty set otherwise).

Associated with $A \neq \emptyset$ we consider the *polar set* and the *orthogonal subspace* given respectively by

$$A^\circ := \left\{x^* \in X^* : \ \langle x^*, x\rangle \leq 1 \text{ for all } x \in A\right\},$$

and

$$A^\perp := \left\{x^* \in X^* : \ \langle x^*, x\rangle = 0 \text{ for all } x \in A\right\}.$$

If $A \subset X^*$, then the following relation holds

$$\bigcap\nolimits_{L\in\mathcal{F}}(A + L^\perp) \subset \operatorname{cl} A, \tag{2}$$

where $\mathcal{F}$ is the family of finite-dimensional linear subspaces in X.

If $A \subset X$ is convex and $x \in X$, we define the *normal cone* to A at x as

$$\mathrm{N}_A(x) := \left\{x^* \in X^* : \ \langle x^*, z - x\rangle \leq 0 \text{ for all } z \in A\right\},$$

if $x \in A$, and the empty set otherwise.

The basic concepts in this paper are traced from [25,28]. Given a function $f : X \longrightarrow \overline{\mathbb{R}}$, its *(effective) domain* and *epigraph* are, respectively,

$$\operatorname{dom} f := \{x \in X : \; f(x) < +\infty\} \text{ and } \operatorname{epi} f := \{(x, \lambda) \in X \times \mathbb{R} : \; f(x) \leq \lambda\}.$$

We say that f is *proper* when $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$. By $\operatorname{cl} f$ and $\overline{\operatorname{co}} f$ we respectively denote the *closed* and the *closed convex hulls* of f, which are the functions such that $\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f)$ and $\operatorname{epi}(\overline{\operatorname{co}} f) = \overline{\operatorname{co}}(\operatorname{epi} f)$. We say that f is lower semicontinuous (lsc, for short) at x if $(\operatorname{cl} f)(x) = f(x)$, and lsc if $\operatorname{cl} f = f$.

Given $x \in X$ and $\varepsilon \geq 0$, the ε *-subdifferential* of f at x is

$$\partial_\varepsilon f(x) = \{x^* \in X^* : \; f(y) \geq f(x) + \langle x^*, y - x\rangle - \varepsilon \text{ for all } y \in X\},$$

when $x \in \operatorname{dom} f$, and $\partial_\varepsilon f(x) := \emptyset$ when $f(x) \notin \mathbb{R}$. The elements of $\partial_\varepsilon f(x)$ are called ε*-subgradients* of f at x. The *subdifferential* of f at x is $\partial f(x) := \partial_0 f(x)$, and its elements are called *subgradients* of f at x. If f and g are convex functions such that one of them is finite and continuous at a point of the domain of the other one, then Moreau-Rockafellar's theorem says that

$$\partial(f + g) = \partial f + \partial g. \tag{3}$$

Given a function $f : X \to \overline{\mathbb{R}}$, the *(Fenchel) conjugate* of f is the function $f^* : X^* \to \overline{\mathbb{R}}$ defined as

$$f^*(x^*) := \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}.$$

The *indicator* and the *support* functions of $A \subset X$ are respectively defined as

$$\mathrm{I}_A(x) := \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \in X \setminus A, \end{cases}$$

and

$$\sigma_A := \mathrm{I}_A^*.$$

Provided that f^* is proper, by Moreau's theorem we have

$$f^{**} = \overline{\operatorname{co}} f, \tag{4}$$

where $f^{**} := (f^*)^*$. For example, if $\{f_i, \; i \in I\}$ is a nonempty family of proper lsc convex functions, then

$$(\sup\nolimits_{i \in I} f_i)^* = \overline{\operatorname{co}}(\inf\nolimits_{i \in I} f_i^*), \tag{5}$$

provided that the supremum function $\sup_{i \in I} f_i$ is proper. Thus, given a nonempty family of closed convex sets $A_i \subset X$, $i \in I$, such that $\cap_{i \in I} A_i \neq \emptyset$, we have $\mathrm{I}_{\bigcap_{i\in I} A_i}(x) = \sup_{i\in I} \mathrm{I}_{A_i}(x)$ and, so, by taking the conjugate in the equalities $\mathrm{I}_{\bigcap_{i\in I} A_i}(x) = \sup_{i\in I} \mathrm{I}_{A_i}(x) = \sup_{i\in I} \sigma^*_{A_i}(x)$, we obtain

$$\sigma_{\bigcap_{i\in I} A_i} = (\mathrm{I}_{\bigcap_{i\in I} A_i})^* = (\sup_{i\in I} \mathrm{I}_{A_i})^* = \overline{\mathrm{co}}(\inf_{i\in I} \sigma_{A_i}).$$

3 Preliminary results in the continuous framework

In Sect. 4 we develop a compactification process addressed to give new characterizations of the subdifferential of the pointwise supremum, with the aim of unifying both the compact and non-compact settings. In this section we gather some preliminary results in the continuous setting.

We give a family of convex functions $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, and the associated supremum function $f := \sup_{t\in T} f_t$. When $\mathrm{ri}(\mathrm{dom}\, f) \neq \emptyset$ and $f_{|\mathrm{aff}(\mathrm{dom}\, f)}$ is continuous on $\mathrm{ri}(\mathrm{dom}\, f)$, by [9, Corollary 3.9] we know that

$$\partial f(x) = \overline{\mathrm{co}} \left\{ \bigcup\nolimits_{t\in T(x)} \partial (f_t + \mathrm{I}_{\mathrm{dom}\, f})(x) \right\}.$$

Consequently, if f is continuous somewhere in its domain, then we obtain ([9, Theorem 3.12])

$$\partial f(x) = \overline{\mathrm{co}} \left\{ \bigcup\nolimits_{t\in T(x)} \partial f_t(x) \right\} + \mathrm{N}_{\mathrm{dom}\, f}(x),$$

and the closure is removed in finite dimensions (see, also, [9, Theorem 3.12])). In particular, when f is continuous at the reference point x, the normal cone above collapses to θ and we recover Valadier's formula [30].

More generally, in the absence of continuity assumptions on the supremum function f we proved in [10, Proposition 1] the following result in which we denote, for any given $x \in \mathrm{dom}\, f$,

$$\mathcal{F}(x) := \{L \subset X : \ L \text{ is a finite-dimensional linear subspace containing } x\}, \quad (6)$$

$$T_\varepsilon(x) := \{t \in T : f_t(x) \geq f(x) - \varepsilon\}, \ \varepsilon \geq 0,$$

with $T_0(x) := T(x)$.

Proposition 1 ([10, Proposition 1]) *Fix $x \in X$ and suppose that, for some $\varepsilon > 0$, the set $T_\varepsilon(x)$ is compact Hausdorff and, for each net $(t_i)_i \subset T_\varepsilon(x)$ converging to t,*

$$\limsup\nolimits_i f_{t_i}(z) \leq f_t(z) \text{ for all } z \in \mathrm{dom}\, f; \quad (7)$$

that is, the functions $f_{(\cdot)}(z)$ are upper semi-continuos (usc, in brief) relatively to $T_\varepsilon(x)$. Then we have

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co} \left\{ \bigcup\nolimits_{t\in T(x)} \partial (f_t + \mathrm{I}_{L\cap \mathrm{dom}\, f})(x) \right\}. \quad (8)$$

For the sake of completeness, we provide next a sketch of the proof.

Sketch of the proof First, we recall that (see relation (8) in [8, p. 1109])

$$\partial f(x) = \bigcap\nolimits_{L \in \mathcal{F}(x)} \partial (f + \mathrm{I}_L)(x). \tag{9}$$

Next, we fix $L \in \mathcal{F}(x)$ and denote by h_t, $t \in T$, the restriction of the function $f_t + \mathrm{I}_L$ to the subspace L. If $h := \sup_{t \in T} h_t$, then $\operatorname{dom} h = (\operatorname{dom} f) \cap L$,

$$\{t \in T : h_t(x) \geq h(x) - \varepsilon\} = T_\varepsilon(x), \text{ for all } \varepsilon \geq 0,$$

and assumption (7) reads, for every net $(t_i)_i \subset T_\varepsilon(x)$ converging to t,

$$\limsup\nolimits_i h_{t_i}(z) \leq h_t(z) \text{ for all } z \in \operatorname{dom} h.$$

Consequently, [8, Theorem 3] applies and yields the following equation in the dual of L,

$$\partial h(x) = \operatorname{co}\left\{\bigcup\nolimits_{t \in T(x)} \partial (h_t + \mathrm{I}_{\operatorname{dom} h})(x)\right\}.$$

Therefore, by an extension argument from the subspace L to the whole space X, the last equation gives rise to

$$\partial (f + \mathrm{I}_L)(x) = \operatorname{co}\left\{\bigcup\nolimits_{t \in T(x)} \partial (f_t + \mathrm{I}_{L \cap \operatorname{dom} f})(x)\right\},$$

and the desired relation follows then by inserting this last relation in (9).

In the general setting, when either T is not compact and/or some of the mappings $t \to f_t(z)$, $z \in \operatorname{dom} f$, fail to be usc, the active index set $T(x)$ as well as the subdifferential sets $\partial f_t(x)$ may be empty. To overcome this situation, the following result given in [15, Theorem 4] (see, also, [14] for finite dimensions) appeals to the ε-subdifferentials of the data functions and the ε-active set $T_\varepsilon(x)$.

Proposition 2 *If*

$$\operatorname{cl} f = \sup_{t \in T} (\operatorname{cl} f_t), \tag{10}$$

then for every $x \in X$

$$\partial f(x) = \bigcap\nolimits_{\varepsilon > 0, L \in \mathcal{F}(x)} \overline{\operatorname{co}}\left\{\bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L \cap \operatorname{dom} f}(x)\right\}. \tag{11}$$

Also here, the intersection over the L's is dropped out if $\operatorname{ri}(\operatorname{dom} f) \neq \emptyset$ ([15, Corollary 8]). Moreover, if f is continuous somewhere, so that (10) holds automatically ([15, Corollary 9]), then the last formula reduces to

$$\partial f(x) = \mathrm{N}_{\operatorname{dom} f}(x) + \bigcap\nolimits_{\varepsilon > 0} \overline{\operatorname{co}}\left\{\bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right\}.$$

Hence, provided that f is continuous at x, we obtain the formula in [31] (where the underlying space X is additionally assumed to be normed).

Formula (11) is proved in [23, Corollary 4.11] under the weaker assumption

$$f^{**} = \sup_{t \in T} f_t^{**}.$$

Under this assumption on the biconjugate, in [23, Theorem 4.1] an alternative formula for the subdifferential of f is provided by replacing the subspaces $L \in \mathcal{F}(x)$ by the family of segments $\{[z, x],\ z \in \operatorname{dom} f\}$ (see also [18] for the use of other families of convex sets instead of $\mathcal{F}(x)$).

Condition (10) guarantees the possibility of characterizing $\partial f(x)$ by means of the f_t's, and not via the augmented functions $f_t + \mathrm{I}_{L \cap \operatorname{dom} f}$ as in Proposition 1. Thus, to complete the analysis, we give next a consequence of (11), which avoids to appeal to condition (10).

Proposition 3 *For every $x \in X$,*

$$\partial f(x) = \bigcap\nolimits_{\varepsilon > 0, L \in \mathcal{F}(x)} \overline{\mathrm{co}} \left\{ \bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon (f_t + \mathrm{I}_{L \cap \operatorname{dom} f})(x) \right\}. \tag{12}$$

Proof Fix $x \in \operatorname{dom} f$ and $L \in \mathcal{F}(x)$, and denote

$$g_t := f_t + \mathrm{I}_{L \cap \operatorname{dom} f},\ t \in T; \quad g := \sup\nolimits_{t \in T} g_t.$$

We have $\operatorname{dom} g_t = L \cap \operatorname{dom} f$ and

$$\operatorname{dom} g \cap (\cap_{t \in T} \operatorname{ri}(\operatorname{dom} g_t)) = (L \cap \operatorname{dom} f) \cap \operatorname{ri}(\operatorname{dom} f \cap L) = \operatorname{ri}(\operatorname{dom} f \cap L) \neq \emptyset,$$

so that, by [15, Corollary 9(iv)], the family $\{g_t,\ t \in T\}$ satisfies condition (10). At the same time we have, for all $\varepsilon \geq 0$,

$$\{t \in T : g_t(x) \geq g(x) - \varepsilon\} = T_\varepsilon(x).$$

Then, since that $\partial f(x) \subset \partial(f + \mathrm{I}_{L \cap \operatorname{dom} f})(x) = \partial g(x)$, by Proposition 2 we obtain that

$$\begin{aligned} \partial f(x) &\subset \bigcap\nolimits_{\varepsilon > 0} \overline{\mathrm{co}} \left\{ \bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon g_t(x) + \mathrm{N}_{L \cap \operatorname{dom} g}(x) \right\} \\ &\subset \bigcap\nolimits_{\varepsilon > 0} \overline{\mathrm{co}} \left\{ \bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon (f_t + \mathrm{I}_{L \cap \operatorname{dom} f})(x) \right\}, \end{aligned}$$

and the inclusion "$\subset$" in (12) follows as L was arbitrarily chosen. The opposite inclusion is straightforward, and we are done. □

4 Compactification approach to the subdifferential

Our main objective in this section is to give a new characterization for $\partial f(x)$, which covers both formula (8) in the compact-continuous setting, using the active set and the exact subdifferential, and formula (12) in the non-compact non-continuous framework, given in terms of ε-active indices and ε-subdifferentials. To this aim, we develop a compactification approach which works by extending the original index set T to a compact set $\widehat{T}$, and building new appropriate functions f_γ, $\gamma \in \widehat{T}$, that satisfy property (7) of Proposition 1. To make the paper self-contained, we resume here the main features of the compactification process, which can be also found in [10].

We start by assuming that T is endowed with some topology τ, for instance the discrete topology. If

$$\mathcal{C}(T, [0, 1]) := \{\varphi : T \to [0, 1] : \varphi \text{ is } \tau\text{-continuous}\}, \tag{13}$$

we consider the product space $[0, 1]^{\mathcal{C}(T,[0,1])}$, which is compact for the product topology (by Tychonoff theorem). We regard the index set T as a subset of $[0, 1]^{\mathcal{C}(T,[0,1])}$. For this purpose we consider the continuous embedding $\mathfrak{w} : T \to [0, 1]^{\mathcal{C}(T,[0,1])}$ which assigns to each $t \in T$ the evaluation function $\mathfrak{w}(t) = \gamma_t$, defined as

$$\gamma_t(\varphi) := \varphi(t),\ \ \varphi \in \mathcal{C}(T, [0, 1]). \tag{14}$$

The closure of $\mathfrak{w}(T)$ in $[0, 1]^{\mathcal{C}(T,[0,1])}$ for the product topology is the compact set

$$\widehat{T} := \text{cl}(\mathfrak{w}(T)), \tag{15}$$

which is the so-called *Stone–Čech compactification* of T, also denoted by βT. The convergence in $\widehat{T}$ is the pointwise convergence; i.e., for $\gamma \in \widehat{T}$ and a net $(\gamma_i)_i \subset \widehat{T}$ we have $\gamma_i \to \gamma$ if and only if

$$\gamma_i(\varphi) \to \gamma(\varphi) \text{ for all } \varphi \in \mathcal{C}(T, [0, 1]). \tag{16}$$

Hence, provided that T is completely regular (when endowed with the discrete topology, for instance), the mapping $\mathfrak{w}$ is an homeomorphism between T and $\mathfrak{w}(T)$, and if $\gamma_i = \gamma_{t_i}$ and $\gamma = \gamma_t$ for some $t, t_i \in T$, then $\gamma_i \to \gamma$ if and only if $t_i \to t$ in T.

Next, we enlarge the original family $\{f_t,\ t \in T\}$ by introducing the functions $f_\gamma : X \to \overline{\mathbb{R}}$, $\gamma \in \widehat{T}$, defined by

$$f_\gamma(z) := \limsup_{\gamma_t \to \gamma,\ t \in T} f_t(z). \tag{17}$$

It can be easily verified that the functions f_γ, $\gamma \in \widehat{T}$, are all convex and satisfy $\sup_{\gamma \in \widehat{T}} f_\gamma \le f$. Moreover, if $(t_n)_n \subset T$ verifies $f(z) = \lim_n f_{t_n}(z)$, with $z \in X$, then there exist a subnet $(t_i)_i$ of $(t_n)_n$ and $\gamma \in \widehat{T}$ such that $\gamma_{t_i} \to \gamma$. Hence,

$$f_\gamma(z) \ge \limsup_i f_{t_i}(z) = \lim_i f_{t_i}(z) = \lim_n f_{t_n}(z) = f(z),$$

and so $\sup_{\gamma \in \widehat{T}} f_\gamma \geq f$. In other words, the functions f_γ provide the same supremum f as the original f_t's; i.e.,

$$\sup_{\gamma \in \widehat{T}} f_\gamma = \sup_{t \in T} f_t = f.$$

If $f(x) \in \mathbb{R}$ and $\varepsilon \geq 0$, then the extended ε-active index set of f at x is

$$\widehat{T}_\varepsilon(x) := \left\{\gamma \in \widehat{T} : f_\gamma(x) \geq f(x) - \varepsilon\right\}, \tag{18}$$

with $\widehat{T}(x) := \widehat{T}_0(x)$; when $f(x) \notin \mathbb{R}$ we set $\widehat{T}_\varepsilon(x) := \emptyset$ for all $\varepsilon \geq 0$. By the compactness of $\widehat{T}$ and the simple fact that, for each $t \in T$,

$$\begin{aligned} f_{\gamma_t}(x) = \limsup_{\gamma_s \to \gamma_t} f_s(x) &= \sup\left\{\lim_i f_{t_i}(x),\ \gamma_{t_i} \to \gamma_t\right\} \\ &\geq \sup\left\{\lim_i f_{t_i}(x),\ t_i \to t\right\} \geq f_t(x), \end{aligned}$$

we verify that $\widehat{T}_\varepsilon(x) \neq \emptyset$. Also, the closedness of $\widehat{T}_\varepsilon(x)$ is established by using a diagonal process.

The way that the functions f_γ, $\gamma \in \widehat{T}$, are constructed ensures the fulfillment of the upper semi-continuity property required in Proposition 1. More precisely, assuming that $f(x) \in \mathbb{R}$ and $\varepsilon \geq 0$, for every net $(\gamma_i)_i \subset \widehat{T}_\varepsilon(x)$ with an accumulation point $\gamma \in \widehat{T}_\varepsilon(x)$, and every $z \in \operatorname{dom} f$, we verify that

$$\limsup_i f_{\gamma_i}(z) \leq f_\gamma(z). \tag{19}$$

Indeed, we may assume without loss of generality that $\gamma_i \to \gamma$ and $\limsup_i f_{\gamma_i}(z) = \lim_i f_{\gamma_i}(z) = \alpha \in \mathbb{R}$. Next, for each i there exists a net $(t_{ij})_j \subset T$ such that

$$\gamma_{t_{ij}} \to_j \gamma_i,\ \ f_{\gamma_i}(z) = \lim_j f_{t_{ij}}(z);$$

that is, $(\gamma_{t_{ij}}, f_{t_{ij}}(z)) \to_j (\gamma_i, f_{\gamma_i}(z))$ and $(\gamma_i, f_{\gamma_i}(z)) \to_i (\gamma, \alpha)$. Then we can find a diagonal net $(t_{ij_i})_i \subset T$ such that $(\gamma_{t_{ij_i}}, f_{t_{ij_i}}(z)) \to_i (\gamma, \alpha)$, and we obtain

$$f_\gamma(z) \geq \limsup_i f_{t_{ij_i}}(z) = \alpha = \limsup_i f_{\gamma_i}(z).$$

The compactification process above covers in a natural way the compact framework. Namely, if T is compact Hausdorff (hence, complete regular), then the family $\left\{f_\gamma, \gamma \in \widehat{T}\right\}$ above turns out to be the family of the usc regularization of the functions $f_{(\cdot)}(z)$, given by

$$\bar{f}_t(z) := \limsup_{s \to t} f_s(z).$$

In this case, the indexed set T does not change; i.e., $\widehat{T}=T$. Consequently, if additionally the functions $f_{(\cdot)}(z)$, $z\in\operatorname{dom} f$, are already usc, then we recover the classical compact and continuous setting, originally proposed in [30].

The following theorem characterizes $\partial f(x)$ in terms of the functions f_γ introduced in (17) and the compact set $\widehat{T}(x)$, when τ is any topology on T. This result is crucial in the subsequent sections.

Theorem 4 *Let $f_t : X\to\overline{\mathbb{R}}$, $t\in T$, be convex functions and $f=\sup_{t\in T} f_t$. Then, for every $x\in X$,*

$$\partial f(x)=\bigcap\nolimits_{L\in\mathcal{F}(x)}\operatorname{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)}\partial(f_\gamma+\mathrm{I}_{L\cap\operatorname{dom} f})(x)\right\}. \tag{20}$$

Proof First, we consider that the topology τ in T is the discrete topology τ_d, so that $C(T,[0,1]):=[0,1]^T$ and $\widehat{T}$ is compact. Moreover, since (T,τ_d) is completely regular, $\widehat{T}$ is Hausdorff (see, i.e., [26, §38]). Since $f=\sup_{\gamma\in\widehat{T}} f_\gamma$ and (19) holds, Proposition 1 applies and yields

$$\partial f(x)=\bigcap\nolimits_{L\in\mathcal{F}(x)}\operatorname{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}^d(x)}\partial(f_\gamma^d+\mathrm{I}_{L\cap\operatorname{dom} f})(x)\right\}, \tag{21}$$

where f_γ^d and $\widehat{T}^d(x)$ are defined as in (17) and (18), respectively, but with respect to the topology τ_d.

Now, let τ be any topology, so that $\tau\subset\tau_d$ and, for any $(\gamma_{t_i})_i\subset\widehat{T}$,

$$\begin{aligned}\gamma_{t_i}\to_{\tau_d}\gamma &\iff \varphi(t_i)\to\gamma(\varphi)\text{ for all }\varphi\in[0,1]^T\\ &\implies \varphi(t_i)\to\gamma(\varphi)\text{ for all }\varphi\in C(T,[0,1])\\ &\iff \gamma_{t_i}\to_\tau\gamma;\end{aligned}$$

hence, for every $z\in X$,

$$f_\gamma^d(z)=\limsup_{\gamma_t\to_{\tau_d}\gamma,\ t\in T} f_t(z)\le\limsup_{\gamma_t\to_\tau\gamma,\ t\in T} f_t(z)=f_\gamma(z).$$

Moreover, since for all $\gamma\in\widehat{T}^d(x)$ we have

$$f(x)=f_\gamma^d(x)\le f_\gamma(x)\le f(x),$$

we deduce that

$$\widehat{T}^d(x)\subset\widehat{T}(x)\text{ and }\partial(f_\gamma^d+\mathrm{I}_{L\cap\operatorname{dom} f})(x)\subset\partial(f_\gamma+\mathrm{I}_{L\cap\operatorname{dom} f})(x). \tag{22}$$

Thus, by (21),

$$\partial f(x)\subset\bigcap\nolimits_{L\in\mathcal{F}(x)}\operatorname{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)}\partial(f_\gamma+\mathrm{I}_{L\cap\operatorname{dom} f})(x)\right\},$$

and (20) follows as the opposite inclusion is straightforward. □

It is worth observing, from the inclusions in (22), that the discrete topology provides the simplest characterization of $\partial f(x)$, since it possibly involves less and smaller sets. Also observe that the intersection over finite-dimensional L in (20) is superfluous in finite dimensions.

Theorem 4 covers the classical Valadier's setting as we show in the following corollary, where the main result is a simpler version of Proposition 1. The formula in assertions (ii) is a global version of [10, Corollary 3]. Nevertheless, we give here an alternative proof based on Theorem 4.

Corollary 5 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be convex functions and assume that T is compact Hausdorff and the mappings $f_{(\cdot)}(z)$, $z \in \operatorname{dom} f$, are usc. If $f = \sup_{t\in T} f_t$, then*

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x)} \operatorname{co}\left\{\bigcup\nolimits_{t\in T(x)} \partial(f_t + \mathrm{I}_{L\cap \operatorname{dom} f})(x)\right\}. \tag{23}$$

In addition, the following assertions hold:

(i) *If $X = \mathbb{R}^n$ and f is continuous at x, then*

$$\partial f(x) = \operatorname{co}\left\{\bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\}.$$

(ii) *If f is continuous at x, then*

$$\partial f(x) = \overline{\operatorname{co}}\left\{\bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\}.$$

Proof Relation (23) is immediate from (20), as $\widehat{T} = T$ and $\widehat{T}(x) = T(x)$ in the current setting.

Assume now that $X = \mathbb{R}^n$, so that (23) reads

$$\partial f(x) = \operatorname{co}\left\{\bigcup\nolimits_{t\in T(x)} \partial(f_t + \mathrm{I}_{\operatorname{dom} f})(x)\right\},$$

and the classical Moreau-Rockafellar subdifferential sum rule entails

$$\partial f(x) = \operatorname{co}\left\{\bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x).$$

Thus (i) holds true since that $\mathrm{N}_{\operatorname{dom} f}(x) = \{\theta\}$ by the continuity of f at x.

Finally, to establish (ii) we observe that the continuity of f at x also ensures the continuity of all the f_t's at x, and so (23) gives rise, again thanks to the classical Moreau-Rockafellar subdifferential sum rule, to

$$\begin{aligned}\partial f(x) &= \bigcap\nolimits_{L\in\mathcal{F}(x)} \operatorname{co}\left\{\bigcup\nolimits_{t\in T(x)} (\partial f_t(x) + L^{\perp} + \mathrm{N}_{\operatorname{dom} f}(x))\right\}\\ &= \bigcap\nolimits_{L\in\mathcal{F}(x)} \operatorname{co}\left\{L^{\perp} + \bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\}\\ &\subset \overline{\operatorname{co}}\left\{\bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\},\end{aligned}$$

□

where the last inclusion uses (2). Thus we are done since the converse of the last inclusion can be easily checked.

Let us also observe that when T admits a one-point compactification $T_\Omega := T \cup \{\Omega\}$ ($\Omega \notin T$), which occurs if and only if T is locally compact Hausdorff (hence, complete regular), instead of $\left\{f_\gamma,\ \gamma \in \widehat{T}\right\}$ we can use the family $\left\{f_{\gamma_t},\ t \in T;\ f_\Omega\right\}$, where

$$f_\Omega(z) := \limsup_{t \to \Omega} f_t(z),\ z \in X. \tag{24}$$

Indeed, in this case the Stone–Čech compactification of T is

$$\widehat{T} := \{\gamma_t,\ t \in T\} \cup \left\{\lim_i \gamma_{t_i} : (t_i)_i \subset T,\ t_i \to \Omega\right\},$$

where the limits $\lim_i \gamma_{t_i}$ and $t_i \to \Omega$ are in $[0,1]^{C(T,[0,1])}$ and T_Ω, respectively. In this way we obtain, for all $t \in T$,

$$f_{\gamma_t} = \limsup_{\gamma_s \to \gamma_t,\ s \in T} f_s = \limsup_{s \to t,\ s \in T} f_s,\ \text{for } t \in T, \tag{25}$$

due to the topological identification of T with $\mathfrak{w}(T)$, and

$$f_\gamma = \limsup_{\gamma_t \to \gamma,\ t \in T} f_t = \limsup_{\gamma_t \to \gamma,\ t \to \Omega,\ t \in T} f_t, \text{for } \gamma \in \widehat{T} \setminus T.$$

Now, we observe that

$$\sup_{\gamma \in \widehat{T} \setminus T} f_\gamma = \sup_{\gamma \in \widehat{T} \setminus T} \limsup_{\gamma_t \to \gamma,\ t \to \Omega,\ t \in T} f_t = \limsup_{t \to \Omega} f_t = f_\Omega.$$

It is clear that the family $\left\{f_{\gamma_t},\ t \in T;\ f_\Omega\right\}$ and the (one-point compactification) index set $T \cup \{\Omega\}$ satisfy the assumptions of Proposition 1, together with $f = \sup\left\{f_{\gamma_t},\ t \in T;\ f_\Omega\right\}$. Thus, it suffices to consider Theorem 4 with this new family $\left\{f_{\gamma_t},\ t \in T;\ f_\Omega\right\}$ instead of the one of the original f_γ's.

In the particular case when $T = \mathbb{N}$, endowed with the discrete topology, for each $n \in \mathbb{N}$ we obtain

$$f_{\gamma_n} = \limsup_{\gamma_k \to \gamma_n,\ k \in \mathbb{N}} f_k = \limsup_{k \to n,\ k \in \mathbb{N}} f_k = f_n,$$

so that the family to consider in Theorem 4 is

$$\{f_n,\ n \in \mathbb{N};\ f_\infty\},$$

where

$$f_\infty = \limsup_{n \to \infty} f_n.$$

Corollary 6 *Assume that T is locally compact Hausdorff. Then, for every $x \in X$, formula* (20) *holds with*

$$\widehat{T}(x) = \begin{cases} \left\{\gamma_t,\ t \in T,\ f_{\gamma_t}(x) = f(x)\right\}, & \text{if } f_\Omega(x) < f(x), \\ \left\{\gamma_t,\ t \in T,\ f_{\gamma_t}(x) = f(x),\ \Omega\right\}, & \text{if } f_\Omega(x) = f(x), \end{cases}$$

and, when $T = \mathbb{N}$,

$$\widehat{T}(x) = \begin{cases} \{n \in \mathbb{N},\ f_n(x) = f(x)\}, & \text{if } f_\infty(x) < f(x), \\ \{n \in \mathbb{N},\ f_n(x) = f(x),\ \infty\}, & \text{if } f_\infty(x) = f(x). \end{cases}$$

5 From non-continuous to continuous. Enhanced formulas

We give in this section some new characterizations of $\partial f(x)$, which provide additional insight to Theorem 4 and that are applied later on in Sect. 6.

According to Theorem 4, $\partial f(x)$ only involves the active functions f_γ, i.e., when $\gamma \in \widehat{T}(x)$. The idea behind the following result is to replace these f_γ's by the new functions $\tilde{f}_\gamma : X \to \mathbb{R}_\infty$, $\gamma \in \widehat{T}$, defined as

$$\tilde{f}_\gamma(z) := \limsup_{\gamma_t \to \gamma,\, f_t(x) \to f(x),\, t \in T} f_t(z), \tag{26}$$

considering only those nets $(t_i)_i \subset T$ associated with functions f_{t_i} approaching the supremum function f at the nominal point x. Observe that if $\gamma \in \widehat{T} \setminus \widehat{T}(x)$, then $\tilde{f}_\gamma \equiv -\infty$ by the convention $\sup \emptyset = -\infty$, and this function is ignored when taking the supremum.

Remember that T is endowed with any topology.

Theorem 7 *For every $x \in X$ we have*

$$\partial f(x) = \bigcap_{L \in \mathcal{F}(x)} \operatorname{co} \left\{ \bigcup_{\gamma \in \widehat{T}(x)} \partial(\tilde{f}_\gamma + \mathrm{I}_{L \cap \operatorname{dom} f})(x) \right\}, \tag{27}$$

where $\tilde{f}_\gamma$ and $\widehat{T}(x)$ are defined in (26) *and* (18)*, respectively.*

Proof We only need to check the inclusion "$\subset$ " when τ is the discrete topology τ_d, and $\partial f(x) \neq \emptyset$; hence, f is lsc at x and proper, and we may suppose, without loss of generality, that $x = \theta$ and $f(\theta) = 0$. Let us fix a closed convex neighborhood U of θ such that $f(z) \geq -1$, for all $z \in U$, and denote by $g_t : X \to \mathbb{R}_\infty$, $t \in T$, the functions given by

$$g_t(z) := \max\{f_t(z), -1\}. \tag{28}$$

Thus, for all $z \in U$,

$$f(z) = \max\{f(z), -1\} = \sup_{t \in T} \max\{f_t(z), -1\} = \sup_{t \in T} g_t(z),$$

and so, applying (20), with the discrete topology τ_d on T, to the family $\{g_t,\ t \in T\}$,

$$\partial f(\theta) = \partial(\sup_{t\in T} g_t)(\theta) = \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcup\nolimits_{\gamma\in\widetilde{T}(\theta)} \partial(g_\gamma + \mathrm{I}_{L\cap \mathrm{dom}\, f})(\theta)\right\}, \quad (29)$$

where $g_\gamma := \limsup\limits_{\gamma_t\to\gamma,\, t\in T} g_t$ and $\widetilde{T}(\theta) := \left\{\gamma \in \widehat{T} : g_\gamma(\theta) = 0\right\}$.

Let us first verify that

$$\widetilde{T}(\theta) = \widehat{T}(\theta). \quad (30)$$

Indeed, if $\gamma \in \widetilde{T}(\theta)$ so that

$$0 = g_\gamma(\theta) = \limsup_{\gamma_t\to\gamma,\, t\in T} g_t(\theta) \leq \max\left\{f_\gamma(\theta), -1\right\} \leq \max\{f(\theta), -1\} = 0,$$

then $f_\gamma(\theta) = 0$ and, so, $\gamma \in \widehat{T}(\theta)$. Conversely, if $\gamma \in \widehat{T}(\theta)$, then

$$0 = f_\gamma(\theta) \leq g_\gamma(\theta) \leq \sup_{\gamma\in\widehat{T}} g_\gamma(\theta) = \sup_{t\in T} g_t(\theta) = f(\theta) = 0,$$

and so $\gamma \in \widetilde{T}(\theta)$.

Next, we fix $\gamma \in \widetilde{T}(\theta)$ and, by the definition of this set, let $(\bar{t}_i)_i \subset T$ be a net such that $\gamma_{\bar{t}_i} \to \gamma$ and $\lim_i g_{\bar{t}_i}(\theta) = 0$; hence,

$$\lim_i f_{\bar{t}_i}(\theta) = \lim_i g_{\bar{t}_i}(\theta) = 0. \quad (31)$$

We also introduce the functions φ_z, $z \in \mathrm{dom}\, f$, defined on T as follows

$$\varphi_z(t) := (\max\{f(z)+1, 1\})^{-1}(g_t(z)+1),$$

which are τ_d-continuous functions such that $\varphi_z(t) \in [0, 1]$ for all $t \in T$, because

$$-1 \leq g_t(z) \leq \max\{f(z), -1\} < +\infty \text{ for all } t \in T \text{ and } z \in \mathrm{dom}\, f.$$

Hence, for every $\gamma_{t_i} \to \gamma$ we have $\varphi_z(t_i) \to_i \gamma(\varphi_z)$, and this entails

$$g_{t_i}(z) \to_i -1 + (\max\{f(z)+1, 1\})\gamma(\varphi_z) \in \mathbb{R}. \quad (32)$$

Consequently, by taking into account that $\gamma_{\bar{t}_i} \to \gamma$ and $\lim_i f_{\bar{t}_i}(\theta) = 0$ (see (31)) we obtain

$$g_\gamma = \limsup_{\gamma_t\to\gamma,\, t\in T} g_t = \lim_{\gamma_t\to\gamma} g_t = \lim_{\gamma_t\to\gamma,\ f_t(\theta)\to 0} g_t, \quad (33)$$

which leads us to

$$g_\gamma + \mathrm{I}_{L\cap \mathrm{dom}\, f} = \lim_{\gamma_t \to \gamma,\, f_t(\theta)\to 0} (g_t + \mathrm{I}_{L\cap \mathrm{dom}\, f}) \tag{34}$$

$$\leq \max\left\{ \limsup_{\gamma_t \to \gamma,\, f_t(\theta)\to 0} (f_t + \mathrm{I}_{L\cap \mathrm{dom}\, f}), -1 \right\}. \tag{35}$$

But the two functions on the left and the right have the same value 0 at θ, and so

$$\partial(g_\gamma + \mathrm{I}_{L\cap \mathrm{dom}\, f})(\theta) \subset \partial\left(\max\left\{ \limsup_{\gamma_t \to \gamma,\, f_t(\theta)\to 0} (f_t + \mathrm{I}_{L\cap \mathrm{dom}\, f}), -1 \right\}\right)(\theta)$$

$$= \partial\left(\limsup_{\gamma_t \to \gamma,\, f_t(\theta)\to 0} f_t + \mathrm{I}_{L\cap \mathrm{dom}\, f}\right)(\theta) = \partial\left(\tilde{f}_\gamma + \mathrm{I}_{L\cap \mathrm{dom}\, f}\right)(\theta),$$

where the first equality comes from Proposition 1, applied to the finite family $\left\{\tilde{f}_\gamma, -1\right\}$. Finally, the desired inclusion follows thanks to (29) and (30). □

Let us introduce a function which asigns to each given $\gamma \in \widehat{T}(x)$ a net $(t_i^\gamma)_i \subset T$ such that

$$\gamma_{t_i^\gamma} \to \gamma,\ f_{t_i^\gamma}(x) \to f(x). \tag{36}$$

Then, according to (35),

$$\lim_{\gamma_t \to \gamma} (g_t + \mathrm{I}_{\mathrm{dom}\, f}) = \lim_i (g_{t_i^\gamma} + \mathrm{I}_{L\cap \mathrm{dom}\, f}) \leq \max\left\{ \limsup_i (f_{t_i^\gamma} + \mathrm{I}_{L\cap \mathrm{dom}\, f}), -1\right\},$$

and we obtain, reasoning as above,

$$\partial f(x) = \bigcap_{L\in\mathcal{F}(x)} \mathrm{co}\left\{ \bigcup_{\gamma\in\widehat{T}(x)} \partial(\limsup_i f_{t_i^\gamma} + \mathrm{I}_{L\cap \mathrm{dom}\, f})(x)\right\}. \tag{37}$$

The use of the functions g_t allows us to formulate $\partial f(x)$ involving only limits instead of upper limits. In fact, from (29), (30) and (33) we get

$$\partial f(x) = \bigcap_{L\in\mathcal{F}(x)} \mathrm{co}\left\{ \bigcup_{\gamma\in\widehat{T}(x)} \partial\left(\lim_{\gamma_t \to \gamma,\, f_t(x)\to f(x)} (g_t + \mathrm{I}_{L\cap \mathrm{dom}\, f})\right)(x)\right\}. \tag{38}$$

Corollary 8 *Suppose that the function f is finite and continuous somewhere. Then, for every $x \in X$,*

$$\partial f(x) = \overline{\mathrm{co}}\left\{ \bigcup_{\gamma\in\widehat{T}(x)} \partial(\limsup_i f_{t_i^\gamma})(x)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x) \tag{39}$$

$$= \mathrm{co}\left\{ \bigcup_{\gamma\in\widehat{T}(x)} \partial(\limsup_i f_{t_i^\gamma})(x)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x)\ (\textit{if } X = \mathbb{R}^n), \tag{40}$$

where (t_i^γ) *is defined in* (36).

Proof Suppose, without loss of generality, that $x = \theta$ and $f(\theta) = 0$. According to (37), and using (3),

$$\begin{aligned}\partial f(\theta) &= \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(\theta)} \partial(\limsup\nolimits_i f_{t_i^\gamma} + \mathrm{I}_{L\cap\mathrm{dom}\, f})(\theta)\right\}\\ &= \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \left(\mathrm{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(\theta)} \partial(\limsup\nolimits_i f_{t_i^\gamma})(\theta)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(\theta) + L^\perp\right),\end{aligned}$$

and (40) follows. To prove (39) we first obtain, due to the last relation and (2),

$$\partial f(\theta) \subset \mathrm{cl}\,(A+B) = \partial\sigma_{A+B}(\theta) = \partial(\sigma_A + \sigma_B)(\theta), \tag{41}$$

where $A := \mathrm{co}\left\{\bigcup_{\gamma\in\widehat{T}(\theta)} \partial(\limsup_i f_{t_i^\gamma})(\theta)\right\}$ and $B := \mathrm{N}_{\mathrm{dom}\, f}(\theta)$.

Since $\limsup_i f_{t_i^\gamma} \le f$ and both functions coincide at θ, we have $A \subset \partial f(\theta)$. There also exist $m \ge 0$, $x_0 \in \mathrm{dom}\, f$ and θ-neighborhood $U \subset X$ such that $f(x_0 + y) \le m$, for all $y \in U$. Then

$$\sigma_A(x_0 + y) \le \sigma_{\partial f(\theta)}(x_0 + y) \le f(x_0 + y) \le m \text{ for all } y \in U; \tag{42}$$

that is, σ_A is continuous at x_0. Consequently, since $\sigma_B(x_0) \le 0$, (41) and (3) entail

$$\partial f(\theta) \subset \partial\sigma_A(\theta) + \partial\sigma_B(\theta) = \mathrm{cl}(A) + B,$$

and the inclusion "$\subset$" in (39) follows. The opposite inclusion is straightforward. □

The following corollary provides a characterization of $\partial f(x)$ in terms only of the active original functions f_t's.

Corollary 9 *Fix* $x \in X$. *If for each net* $(t_i)_i \subset T$ *satisfying* $f_{t_i}(x) \to f(x)$, *there exist a subnet* $(t_{i_j})_j \subset T$ *of* $(t_i)_i$ *and an index* $t \in T$ *such that*

$$\limsup\nolimits_j f_{t_{i_j}}(z) \le f_t(z) \textit{ for all } z \in \mathrm{dom}\, f, \tag{43}$$

then we have

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co}\left\{\bigcup\nolimits_{t\in T(x)} \partial(f_t + \mathrm{I}_{L\cap\mathrm{dom}\, f})(x)\right\}.$$

Proof Given any $\gamma \in \widehat{T}(x)$ such that $\gamma_{t_i} \to \gamma$ and $f_{t_i}(x) \to f(x)$, for some net $(t_i)_i \subset T$, we choose a subnet $(t_{i_j}^\gamma)_j$ in (36) satisfying (43) for a certain $t^\gamma \in T$. Then $t^\gamma \in T(x)$, taking into account (43) with $z = x$, and by (37)

$$\begin{aligned}\partial f(x) &= \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial(\limsup\nolimits_j f_{t_{i_j}^\gamma} + \mathrm{I}_{L\cap\mathrm{dom}\, f})(x)\right\}\\ &\subset \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial(f_{t^\gamma} + \mathrm{I}_{L\cap\mathrm{dom}\, f})(x)\right\},\end{aligned}$$

where the last inclusion holds as $\limsup_j f_{t^{\gamma}_{i_j}} + \mathrm{I}_{L\cap \operatorname{dom} f} \leq f_{t^\gamma} + \mathrm{I}_{L\cap \operatorname{dom} f}$, by (43), and these two functions take the same value at x. The inclusion "$\subset$" follows as we have shown that $t^\gamma \in T(x)$. The opposite inclusion is immediate. □

6 From continuous to non-continuous

In this section, we consider again a family $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, of convex functions defined on X, and the supremum function $f := \sup_{t\in T} f_t$. Based on the results of the previous section we provide characterizations of $\partial f(x)$ involving only the f_t's and not the regularized ones, i.e, the f_γ's. We shall need the following technical lemmas. In what follows, cl^s stands for the strong topology on X^* (usually denoted by $\beta(X^*, X)$).

Lemma 10 *Assume that the convex functions f_t, $t \in T$, are proper, lsc, and such that $f_{|\operatorname{aff}(\operatorname{dom} f)}$ is continuous on $\operatorname{ri}(\operatorname{dom} f)$, assumed nonempty. Let $x \in \operatorname{dom} f$ and the net $(z_i^*)_{i\in I} \subset X^*$ such that*

$$\lim_i (\langle z_i^*, x\rangle - \inf_{t\in T} f_t^*(z_i^*)) = f(x), \tag{44}$$

and for all $z \in \operatorname{dom} f$

$$\limsup_i \left(\langle z_i^*, z\rangle - \inf_{t\in T} f_t^*(z_i^*)\right) > -\infty. \tag{45}$$

Then there exist a subnet $(z_{i_j}^*)_j$ of $(z_i^*)_i$ and $z^* \in X^*$ such that

$$z^* \in \mathrm{cl}\left(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + (\operatorname{aff}(\operatorname{dom} f))^\perp\right),\ \textit{for all } \varepsilon > 0, \tag{46}$$

and

$$\left\langle z_{i_j}^* - z^*, z\right\rangle \to_j 0,\ \textit{for all } z \in \operatorname{aff}(\operatorname{dom} f). \tag{47}$$

In particular, if $\operatorname{dom} f$ is finite-dimensional, then (46) also holds with cl^s instead of cl.

Proof We may assume that $x = \theta$ and $f(\theta) = 0$, and denote $E := \operatorname{aff}(\operatorname{dom} f)$ which is a closed subspace with dual E^*. We also denote $h := \inf_{t\in T} f_t^*$, so that (see (4))

$$h^* = (\inf_{t\in T} f_t^*)^* = \sup_{t\in T} f_t^{**} = \sup_{t\in T} f_t = f, \tag{48}$$

and

$$h^*(\theta) + h(z_i^*) = f(\theta) + h(z_i^*) = h(z_i^*) \to 0. \tag{49}$$

Hence, for every fixed $\varepsilon > 0$, there is some $i_0 \in I$ such that for all $i \succeq i_0$

$$h^*(\theta) + h(z_i^*) = \sup_{t \in T} f_t(\theta) + \inf_{t \in T} f_t^*(z_i^*) = h(z_i^*) < \varepsilon, \tag{50}$$

and so

$$(z_i^*)_{i \succeq i_0} \subset \partial_\varepsilon h^*(\theta) = \partial_\varepsilon f(\theta). \tag{51}$$

Now, using the continuity assumption, we choose $x_0 \in \operatorname{dom} f$, a θ-neighborhood $U \subset X$ and $r \geq 0$ such that

$$f(x_0 + y) \leq r \text{ for all } y \in U \cap E, \tag{52}$$

and, by (45) with $z = x_0$ and (49),

$$\limsup_i \langle z_i^*, x_0 \rangle > -\infty.$$

Therefore we may assume, up to some subnet, that $\inf_i \langle z_i^*, x_0 \rangle > -\infty$ and, so, by (51) and (52), there is some $m > 0$ such that

$$\langle z_i^*, y \rangle \leq f(x_0 + y) + \varepsilon - \inf_i \langle z_i^*, x_0 \rangle \leq m, \text{ for all } y \in U \cap E \text{ and for all } i; \tag{53}$$

that is $(z_i^*)_i \subset (U \cap E)^\circ$. Since the last set is weak*-compact in E^*, by the Alaoglu-Banach-Bourbaki theorem, there exist a subnet $(z_{i_j|E}^*)_j$ and $\tilde{z}^* \in E^*$ such that

$$\left\langle z_{i_j|E}^* - \tilde{z}^*, u \right\rangle \to_j 0 \text{ for all } u \in E, \tag{54}$$

where the subscript "$|E$" denotes the restriction to E.

Moreover, by the Hahn-Banach theorem, $\tilde{z}^* \in E^*$ is extended to some $z^* \in X^*$, which satisfies

$$\left\langle z_{i_j}^* - z^*, u \right\rangle = \left\langle z_{i_j|E}^* - \tilde{z}^*, u \right\rangle \to_j 0 \text{ for all } u \in E, \tag{55}$$

Now, using (50), we see that for each i there exists $t_i \in T$ such that

$$f_{t_i}(\theta) + f_{t_i}^*(z_i^*) \leq f_{t_i}^*(z_i^*) < \varepsilon,$$

entailing that $z_i^* \in \partial_\varepsilon f_{t_i}(\theta)$ and

$$-f_{t_i}(\theta) = \langle z_i^*, \theta \rangle - f_{t_i}(\theta) \leq f_{t_i}^*(z_i^*) < \varepsilon;$$

that is, $t_i \in T_\varepsilon(\theta)$ and so,

$$z_i^* \in \bigcup\nolimits_{t \in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta).$$

We fix a weak* (strong, when dom f is finite-dimensional) θ-neighborhood $V \subset X^*$. Since E^* is isomorphic to the quotient space $X^*{}_{\diagup E^\perp}$, then [12]

$$V_{|E} := \left\{ u^*_{|E} : u^* \in V \right\} \in \mathcal{N}_{E^*},$$

where $u^*_{|E}$ denotes the restriction of u^* to E^*. Consequently, writing

$$z^*_{i_j|E} \in A := \left\{ u^*_{|E} \in E^* : u^* \in \bigcup\nolimits_{t \in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) \right\},$$

and passing to the limit on j, (55) leads us to

$$z^*_{|E} \in A + V_{|E}. \tag{56}$$

In other words, there are $u^* \in \bigcup_{t \in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta)$ and $v^* \in V$ such that $z^*_{|E} = u^*_{|E} + v^*_{|E}$; that is,

$$\langle z^*, u \rangle = \langle u^* + v^*, u \rangle \text{ for all } u \in E,$$

implying that

$$z^* \in u^* + v^* + E^\perp \subset \bigcup\nolimits_{t \in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) + E^\perp + V.$$

The conclusion follows then by intersecting over V and, after, over $\varepsilon > 0$. □

In the current framework, $\widehat{X^*}$ is the Stone–Čech compactification of X^*, with respect to the discrete topology, and the mappings $\gamma_{z^*} : [0,1]^{X^*} \to [0,1]$, $z^* \in X^*$, are defined as in (14), so that the convergence $\gamma_{z_i^*} \to \gamma$ for a net $(z_i^*)_i \subset X^*$ and $\gamma \in \widehat{X^*}$ means

$$\varphi(z_i^*) \to \gamma(\varphi) \text{ for all } \varphi \in [0,1]^{X^*}.$$

Lemma 11 *Assume in Lemma 10 that the net $(\gamma_{z_i^*})_i$ converges in $\widehat{X^*}$. Then for the function*

$$\psi(z) := \limsup_i \left(\langle z_i^*, z \rangle - \inf_{t \in T} f_t^*(z_i^*) + \mathrm{I}_{\operatorname{dom} f}(z) \right), \; z \in X,$$

we have

$$\begin{aligned} \partial\psi(x) &\subset \mathrm{N}_{\operatorname{dom} f}(x) + \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + (\operatorname{aff}(\operatorname{dom} f))^\perp \right) \\ &\subset \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) + \mathrm{N}_{\operatorname{dom} f}(\theta) \right), \end{aligned}$$

with cl^s *instead of* cl *when* dom f *is finite-dimensional.*

Proof We may suppose that $x = \theta$ and $f(\theta) = 0$. By Lemma 10 there exist a subnet $(z^*_{i_j})_j$ of $(z^*_i)_i$ and

$$z^* \in \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + (\operatorname{aff}(\operatorname{dom} f))^\perp\right)$$

such that $(z^*_{i_j})_j$ weak*-converges to z^* in E^* (where $E = \operatorname{aff}(\operatorname{dom} f)$).

We introduce the functions $g_{u^*} : X \to \mathbb{R}_\infty$, $u^* \in X^*$, defined as

$$g_{u^*} := \max\left\{u^* - h(u^*), -1\right\},$$

where $h = \inf_{t\in T} f^*_t$ (already used in the proof of Lemma 10). Observe that (recall (48))

$$-1 \le g_{u^*} \le \max\left\{h^*, -1\right\} = \max\left\{f, -1\right\},$$

and

$$\varphi_z(u^*) := \frac{g_{u^*}(z) + 1}{\max\left\{f(z) + 1, 1\right\}} \in [0, 1]\,,\ \text{ for all } z \in \operatorname{dom} f.$$

Hence, since φ_z is obviously continuous on X^*, endowed with the discrete topology, the convergence assumption of $(\gamma_{z^*_i})_i$ ensures that, for each $z \in \operatorname{dom} f$, the net

$$\gamma_{z^*_i}(\varphi_z) = \frac{g_{z^*_i}(z) + 1}{\max\left\{f(z) + 1, 1\right\}}$$

also converges, as well as the net $(g_{z^*_i}(z))_i$. Then, taking into account (44) and (47), we obtain

$$\begin{aligned}\lim_i g_{z^*_i}(z) &= \lim_i \max\left\{\langle z^*_i, z\rangle - h(z^*_i), -1\right\}\\ &= \lim_j \max\left\{\left\langle z^*_{i_j}, z\right\rangle, -1\right\} = \max\left\{\langle z^*, z\rangle, -1\right\},\end{aligned}$$

which gives

$$\limsup_i \langle z^*_i, z\rangle \le \limsup_i (\max\{\langle z^*_i, z\rangle, -1\}) = \max\{\langle z^*, z\rangle, -1\}.$$

But both functions $\limsup_i z^*_i + \mathrm{I}_{\operatorname{dom} f}$ and $\max\{z^*, -1\} + \mathrm{I}_{\operatorname{dom} f}$ coincide at θ, and so

$$\partial\left(\limsup_i (z^*_i + \mathrm{I}_{\operatorname{dom} f})\right)(\theta) \subset \partial(\max\{z^* + \mathrm{I}_{\operatorname{dom} f}, -1\})(\theta),$$

and (20) applied to the (finite) family $\{z^* + \mathrm{I}_{\mathrm{dom}\, f}, -1\}$ yields (recall (44))

$$\begin{aligned}\partial\psi(\theta) &= \partial\left(\lim_i \sup(z_i^* + \mathrm{I}_{\mathrm{dom}\, f})\right)(\theta) \\ &\subset z^* + \mathrm{N}_{\mathrm{dom}\, f}(\theta). \\ &\subset \mathrm{N}_{\mathrm{dom}\, f}(\theta) + \bigcap\nolimits_{\varepsilon>0} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) + (\mathrm{aff}(\mathrm{dom}\, f))^\perp\right) \\ &\subset \bigcap\nolimits_{\varepsilon>0} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) + \mathrm{N}_{\mathrm{dom}\, f}(\theta)\right).\end{aligned}$$

□

Theorem 12 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be convex functions and $f = \sup_{t\in T} f_t$. Then, for every $x \in X$,*

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^s\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon(f_t + \mathrm{I}_{L\cap\mathrm{dom}\, f})(x)\right)\right\}. \qquad (57)$$

If, in addition,

$$\mathrm{cl}\, f = \sup_{t\in T}(\mathrm{cl}\, f_t), \qquad (58)$$

then

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L\cap\mathrm{dom}\, f}(x)\right)\right\}. \qquad (59)$$

Remark 1 **(before the proof)** *Formula* (57) *leads straightforwardly to the following characterization of $\partial f(x)$, using the strong closure*

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x),\varepsilon>0} \overline{\mathrm{co}}^s\left\{\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon(f_t + \mathrm{I}_{L\cap\mathrm{dom}\, f})(x)\right\},$$

improving the one of Proposition 3, which is given in terms of the weak-closure. However, on despite that both formulas involve similar elements, the order in taking the intersection over ε leads to different interpretations of $\partial f(x)$. For instance, if T is finite, $T = T(x)$ and f is continuous, then* (57) *reads*

$$\partial f(x) = \mathrm{co}\left\{\bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\},$$

giving Valadier's formula (see, e.g., [30]), *while Proposition 3 yields*

$$\partial f(x) = \bigcap\nolimits_{\varepsilon>0} \overline{\mathrm{co}}\left\{\bigcup\nolimits_{t\in T(x)} \partial_\varepsilon f_t(x)\right\},$$

which turns out to be the Brøndsted formula ([1]; *see, also,* [15, Corollary 12]).

Proof The inclusions "$\supset$" in both formulas are straightforward. We may suppose, without loss of generality, that $x = \theta,\ f(\theta) = 0$ and $\partial f(\theta) \neq \emptyset$; hence,

$$\partial(\mathrm{cl}\, f)(\theta) = \partial f(\theta) \text{ and } f(\theta) = (\mathrm{cl}\, f)(\theta) = 0. \tag{60}$$

We proceed in three steps:
Step 1. We assume that all the f_t's are proper and lsc; hence, (58) obviously holds. We fix $L \in \mathcal{F}(\theta)$, and define the functions

$$\tilde{f}_t := f_t + \mathrm{I}_L,\ t \in T,\ \text{ and } h := \inf\nolimits_{t\in T} \tilde{f}_t^*. \tag{61}$$

The $\tilde{f}_t$'s are proper and lsc, and we have (see (4))

$$(f + \mathrm{I}_L)(z) = \sup\nolimits_{t\in T} \tilde{f}_t(z) = \sup\nolimits_{t\in T} \tilde{f}_t^{**}(z) = (\inf\nolimits_{t\in T} \tilde{f}_t^*)^*(z) = h^*(z); \tag{62}$$

that is,

$$(f + \mathrm{I}_L)(z) = \sup\left\{\langle z, z^*\rangle - h(z^*),\ z^* \in X^*\right\},$$

and (37) applied with $T = X^*$ (endowed with the discrete topology) yields

$$\partial(f + \mathrm{I}_L)(\theta) \subset \mathrm{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{X^*}(\theta)} \partial\left(\limsup_i (z_i^{*\gamma} - h(z_i^{*\gamma}) + \mathrm{I}_{L\cap \mathrm{dom}\, f})\right)(\theta)\right\}, \tag{63}$$

where $\widehat{X^*}(\theta)$ represents the set $\widehat{T}(\theta)$ given in (18); that is,

$$\widehat{X^*}(\theta) = \left\{\gamma \in \widehat{X^*} : \limsup_{\gamma_{z^*}\to\gamma}(-h(z^*)) = 0\right\},$$

and $(z_i^{*\gamma})_i \subset X^*$ is a fixed net such that $\gamma_{z_i^{*\gamma}} \to \gamma$ and $h(z_i^{*\gamma}) \to 0$ (by (36)). Consequently, for every $\gamma \in \widehat{X^*}(\theta)$, Lemma 11 applies and yields

$$\begin{aligned}&\partial\left(\limsup_i (z_i^{*\gamma} - h(z_i^{*\gamma}) + \mathrm{I}_{L\cap \mathrm{dom}\, f})\right)(\theta)\\ &\quad\subset \bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^s\left(\bigcup\nolimits_{t\in T_\varepsilon^1(\theta)} \partial_\varepsilon \tilde{f}_t(\theta) + \mathrm{N}_{L\cap \mathrm{dom}\, f}(\theta)\right),\end{aligned} \tag{64}$$

where

$$T_\varepsilon^1(\theta) := \left\{t \in T : \tilde{f}_t(\theta) \geq -\varepsilon\right\} = T_\varepsilon(\theta). \tag{65}$$

Indeed, condition (45) is satisfied when the left-hand side in (64) is nonempty, and thus the function $\limsup_i (z_i^{*\gamma} - h(z_i^{*\gamma}) + \mathrm{I}_{L\cap \mathrm{dom}\, f})$ is proper. Consequently, combining

(63), (64) and (65),

$$\partial(f+\mathrm{I}_L)(\theta) \subset \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^s\left(\bigcup\nolimits_{t\in T_\varepsilon(\theta)} \partial_\varepsilon \tilde{f}_t(\theta) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right)\right\}, \tag{66}$$

and the inclusion "$\subset$" in (57) follows since $\partial f(\theta) \subset \partial(f+\mathrm{I}_L)(\theta)$ and

$$\partial_\varepsilon \tilde{f}_t(\theta) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta) \subset \partial_\varepsilon(f_t + \mathrm{I}_{L\cap \operatorname{dom} f})(\theta).$$

Moreover, due to the fact that $\partial_\varepsilon \tilde{f}_t(\theta) \subset \mathrm{cl}(\partial_\varepsilon f_t(\theta) + L^\perp)$ (see, e.g., [17]), (66) implies that

$$\begin{aligned}\partial f(\theta) &\subset \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^s\left(\bigcup\nolimits_{t\in T_\varepsilon(\theta)} \mathrm{cl}(\partial_\varepsilon f_t(\theta) + L^\perp) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right)\right\}\\ &\subset \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^s\left(\mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right)\right)\right\}\\ &= \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(\theta)} \partial_\varepsilon f_t(\theta) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right)\right\},\end{aligned} \tag{67}$$

which yields the inclusion "$\subset$" in (59).

Step 2. We suppose that (58) holds and we fix $L \in \mathcal{F}(\theta)$. By (60) we choose a θ-neighborhood $U \subset X$ such that

$$f(z) \geq (\mathrm{cl}\, f)(z) \geq -1, \text{ for all } z \in U, \tag{68}$$

and denote $S := \{t \in T : \ \mathrm{cl}\, f_t \text{ is proper}\}$. We define the functions

$$g_t := \mathrm{cl}\, f_t, \text{ if } t \in S, \text{ and } g_t := \max\{\mathrm{cl}\, f_t, -1\}, \text{ otherwise.}$$

Then (see the proof of [15, Theorem 4], page 871) g_t is proper, lsc and convex,

$$g(z) := \sup_{t\in T} g_t(z) = (\mathrm{cl}\, f)(z), \text{ for all } z \in U;$$

hence, $g(\theta) = 0$,

$$\{t \in T : g_t(\theta) \geq -\varepsilon\} \subset T_\varepsilon(\theta) \cap S, \ \forall \varepsilon \in]0,1[,$$
$$\partial_\varepsilon g_t(\theta) \subset \partial_{2\varepsilon} f_t(\theta), \ \partial_\varepsilon(g_t + \mathrm{I}_{L\cap \operatorname{dom} f})(\theta) \subset \partial_{2\varepsilon}(f_t + \mathrm{I}_{L\cap \operatorname{dom} f})(\theta), \ \forall \varepsilon \in]0,1[,$$

and

$$\partial f(\theta) = \partial(\mathrm{cl}\, f)(\theta) = \partial g(\theta). \tag{69}$$

Consequently, by Step 1,

$$\begin{aligned}\partial f(\theta) &= \partial g(\theta)\\ &= \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^{s}\left(\bigcup\nolimits_{t\in T,\ g_t(\theta)\geq-\varepsilon} \partial_{\varepsilon}(g_t+\mathrm{I}_{L\cap\mathrm{dom}\, g})(\theta)\right)\right\}\\ &\subset \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^{s}\left(\bigcup\nolimits_{t\in T_{\varepsilon}(\theta)} \partial_{2\varepsilon}(f_t+\mathrm{I}_{L\cap\mathrm{dom}\, f})(\theta)\right)\right\},\end{aligned}$$

entailing the desired inclusion "$\subset$ " in (57). Similarly, (67) yields

$$\begin{aligned}\partial f(\theta) &= \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcap\nolimits_{0<\varepsilon<1} \mathrm{cl}\left(\bigcup\nolimits_{t\in T,\ g_t(\theta)\geq-\varepsilon} \partial_{\varepsilon} g_t(\theta)+\mathrm{N}_{L\cap\mathrm{dom}\, g}(\theta)\right)\right\}\\ &\subset \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcap\nolimits_{0<\varepsilon<1} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_{\varepsilon}(\theta)} \partial_{2\varepsilon} f_t(\theta)+\mathrm{N}_{L\cap\mathrm{dom}\, f}(\theta)\right)\right\}, \qquad (70)\end{aligned}$$

which easily leads to the inclusion "$\subset$ " in (59).

Step 3. We prove (57) in the general case, without assuming (58). We fix $L \in \mathcal{F}(\theta)$ and define

$$\hat{f}_t := f_t + \mathrm{I}_{L\cap\mathrm{dom}\, f},$$

so that

$$f_L := \sup_{t\in T} \hat{f}_t = f + \mathrm{I}_{L\cap\mathrm{dom}\, f} = f + \mathrm{I}_L,$$
$$\hat{f}_t(\theta) = f_t(\theta),\ \ f_L(\theta) = 0,\ \text{ and }\ \mathrm{dom}\, f_L = L \cap \mathrm{dom}\, f.$$

Moreover, by arguing as in the proof of Proposition 3, the family $\left\{\hat{f}_t,\ t \in T\right\}$ satisfies condition (58) and we have (see (9))

$$\partial f(\theta) = \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \partial(f+\mathrm{I}_L)(\theta) = \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \partial f_L(\theta).$$

Applying Step 2 to the family $\left\{\hat{f}_t,\ t \in T\right\}$ we get

$$\begin{aligned}\partial f(\theta) &= \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \partial f_L(\theta)\\ &\subset \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^{s}\left(\bigcup\nolimits_{t\in T,\ \hat{f}_t(\theta)\geq-\varepsilon} \partial_{\varepsilon}(\hat{f}_t+\mathrm{I}_{L\cap\mathrm{dom}\, f_L})(\theta)\right)\right\}\\ &= \bigcap\nolimits_{L\in\mathcal{F}(\theta)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^{s}\left(\bigcup\nolimits_{T_{\varepsilon}(\theta)} \partial_{\varepsilon}(f_t+\mathrm{I}_{L\cap\mathrm{dom}\, f})(\theta)\right)\right\},\end{aligned}$$

and the inclusion "$\subset$" in (57) follows. $\square$

The following corollary closing this section considers a frequent hypothesis in the literature.

Corollary 13 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be convex functions. If $f = \sup_{t\in T} f_t$ is finite and continuous at some point, then for every $x \in X$*

$$\begin{aligned}\partial f(x) &= \mathrm{N}_{\operatorname{dom} f}(x) + \overline{\operatorname{co}}\left\{\bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right)\right\}\\ &= \mathrm{N}_{\operatorname{dom} f}(x) + \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right)\right\} \ (\textit{if } X = \mathbb{R}^n).\end{aligned}$$

Proof The proof is similar to the one of Theorem 12, but with the use of the formulas in Corollary 8 instead of formula (37). □

We close this section with an extension of Theorem 12 to nonconvex functions. We also refer to [24], and references therein, for other studies on the subdifferential of the supremum of nonconvex functions.

Corollary 14 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be a family of non-necessarily convex functions and $f := \sup_{t\in T} f_t$. Assume that*

$$f^{**} = \sup_{t\in T} f_t^{**}.$$

Then (59) holds.

Proof It suffices to prove the inclusion "$\subset$" in (57) for x such that $\partial f(x) \neq \emptyset$; hence, f^* is proper, $f(x) = f^{**}(x)$ and $\partial f(x) = \partial(\overline{\operatorname{co}} f)(x) = \partial f^{**}(x)$. Thus, applying the second statement in Theorem 11 to the family $\left\{f_t^{**},\ t \in T\right\}$,

$$\begin{aligned}\partial f(x) &= \partial f^{**}(x)\\ &= \bigcap\nolimits_{L\in\mathcal{F}(x)} \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon^1(x)} \partial_\varepsilon f_t^{**}(x) + \mathrm{N}_{L\cap \operatorname{dom} f^{**}}(x)\right)\right\},\end{aligned}$$

where $T_\varepsilon^1(x) := \left\{t \in T : f_t^{**}(x) \geq f(x) - \varepsilon\right\}$. Observe that every $t \in T_\varepsilon^1(x)$ satisfies

$$f_t(x) \geq f_t^{**}(x) \geq f(x) - \varepsilon \geq f_t(x) - \varepsilon;$$

hence, $t \in T_\varepsilon(x)$ and $\partial_\varepsilon f_t^{**}(x) \subset \partial_{2\varepsilon} f_t(x)$. Additionally, the inequality $f^{**} \leq f$ implies that $\mathrm{N}_{L\cap \operatorname{dom} f^{**}}(x) \subset \mathrm{N}_{L\cap \operatorname{dom} f}(x)$, and the desired inclusion follows. □

7 Two applications in optimization

First, in this section, we apply the previous results to extend the classical Fenchel duality to the nonconvex framework. This will lead us to recover some of the results in [3–5] (see, also, [23]), relating the solution set of a nonconvex optimization problem and its convexified relaxation. Second, we establish Fritz-John and KKT optimality conditions for convex semi-infinite optimization problems, improving similar results in [8].

Given a function $g : X \to \mathbb{R}_\infty$, we let $f : X^* \to \overline{\mathbb{R}}$ be the Fenchel conjugate of g. When g is proper, lsc and convex, the classical Fenchel duality, together with (4), yields

$$\partial f = (\partial g)^{-1}. \tag{71}$$

We extend this relation to non-necessarily convex functions. We denote below the closure with respect to the weak topology in X by cl^w.

Proposition 15 *Assume that the function f is proper. Then, for every $x^* \in X^*$,*

$$\partial f(x^*) = \bigcap\nolimits_{L\in\mathcal{F}(x^*)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^w \left((\partial_\varepsilon g)^{-1}(x^*) + \mathrm{N}_{L\cap \mathrm{dom}\, f}(x^*)\right)\right\}.$$

If, in addition, f is finite and (weak-) continuous somewhere, then*

$$\begin{aligned}\partial f(x^*) &= \overline{\mathrm{co}}\left\{\left((\partial(\mathrm{cl}^w g))^{-1}(x^*)\right)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x^*)\\ &= \mathrm{co}\left\{\left((\partial(\mathrm{cl}\, g))^{-1}(x^*)\right)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x^*) \text{ (if } X = \mathbb{R}^n),\end{aligned}$$

where $\mathrm{cl}^w g$ is the weak-lsc hull of g.

Proof We define the convex functions $f_x : X^* \to \overline{\mathbb{R}}$, $x \in X$, as

$$f_x(x^*) := \langle x, x^*\rangle - g(x), \ \ x \in \mathrm{dom}\, g,$$

so that f_x is weak*-continuous and $f = \sup_{x\in\mathrm{dom}\, g} f_x$. Then, according to formula (59), for every $x^* \in X^*$ we have

$$\partial f(x^*) = \bigcap\nolimits_{L\in\mathcal{F}(x^*)} \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^w \left(\bigcup\nolimits_{x\in T_\varepsilon(x^*)} \partial_\varepsilon f_x(x^*) + \mathrm{N}_{L\cap \mathrm{dom}\, f}(x^*)\right)\right\},$$

where

$$T_\varepsilon(x^*) := \left\{x \in \mathrm{dom}\, g : f_x(x^*) \geq f(x^*) - \varepsilon\right\} = (\partial_\varepsilon g)^{-1}(x^*).$$

Consequently, the first formula comes from the fact that $\partial_\varepsilon f_x(x^*) = \{x\}$.

Assume now that f is finite and weak*-continuous somewhere. Then, arguing in a similar way, but using Corollary 13 instead of (59),

$$\begin{aligned}\partial f(x^*) &= \overline{\mathrm{co}}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^w \left((\partial_\varepsilon g)^{-1}(x^*)\right)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x^*)\\ &= \mathrm{co}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl} \left((\partial_\varepsilon g)^{-1}(x^*)\right)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x^*) \text{ (if } X = \mathbb{R}^n).\end{aligned}$$

The desired formulas follow as

$$\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}^w \left((\partial_\varepsilon g)^{-1}(x^*)\right) = (\partial(\mathrm{cl}^w g))^{-1}(x^*), \tag{72}$$

according to [6, Lemma 2.3]. □

Observing that $\operatorname{Argmin}(\overline{\operatorname{co}}g) = \partial f(\theta)$, the previous proposition gives:

Corollary 16 *Assume that the function f is proper. Then we have*

$$\operatorname{Argmin}(\overline{\operatorname{co}}g) = \bigcap_{L\in\mathcal{F}(\theta)} \operatorname{co}\left\{\bigcap_{\varepsilon>0} \operatorname{cl}^{w}\left(\varepsilon\text{- Argmin}\, g + \mathrm{N}_{L\cap\operatorname{dom} f}(\theta)\right)\right\}.$$

If, in addition, f is finite and continuous at some point, then

$$\begin{aligned}\operatorname{Argmin}(\overline{\operatorname{co}}g) &= \overline{\operatorname{co}}(\operatorname{Argmin}(\operatorname{cl}^{w} g)) + \mathrm{N}_{\operatorname{dom} f}(\theta)\\ &= \operatorname{co}(\operatorname{Argmin}(\operatorname{cl} g)) + \mathrm{N}_{\operatorname{dom} f}(\theta)\ (\text{if } X = \mathbb{R}^n).\end{aligned}$$

When X is a normed space, the set $\partial f(x^*)$ is also seen as a subset of the bidual space, whereas Proposition 15 characterizes only the part of $\partial f(x^*)$ in the subspace X of X^{**}. A light adaptation of Proposition 15 allows us to have a complete picture of $\partial f(x^*)$, as a proper set of the bidual space X^{**}. In such a setting, we denote the weak*-topology $\sigma(X^{**}, X^*)$ in X^{**} by w^{**}, and introduce the function $\overline{g}^{w^{**}} : X^{**} \to \overline{\mathbb{R}}$ defined by

$$\overline{g}^{w^{**}}(y) = \liminf_{x\to^{w^{**}} y} g(x),\ \ y \in X^{**}.$$

We refer, e.g., to [2, Chapter 1] for these concepts.

Proposition 17 *Assume that X is a normed space and X^* is endowed with the dual norm topology. If the function f is proper, then for every $x^* \in X^*$*

$$\partial f(x^*) = \bigcap_{L\in\mathcal{F}(x^*)} \operatorname{co}\left\{\bigcap_{\varepsilon>0} \operatorname{cl}^{w^{**}}\left((\partial_\varepsilon g)^{-1}(x^*) + \mathrm{N}_{L\cap\operatorname{dom} f}(x^*)\right)\right\}.$$

If, in addition, f is finite and (norm-) continuous somewhere, then

$$\partial f(x^*) = \overline{\operatorname{co}}\left\{(\partial\overline{g}^{w^{**}})^{-1}(x^*)\right\} + \mathrm{N}_{\operatorname{dom} f}(x^*).$$

Proof Following similar arguments as those used in [4], we apply Proposition 15 in the duality pair $((X^{**}, w^{**}), (X^*, \|\|_*))$, replacing the function g there by the function $\hat{g}$ defined on X^{**} as

$$\hat{g}(y) = g(y),\ \ \text{if } y \in X^{**};\ +\infty,\ \text{otherwise}.$$

Observe that the w^{**}-lsc hull of $\hat{g}$ is precisely the function $\overline{g}^{w^{**}}$. □

Now, as in [8,10], we consider the following convex semi-infinite optimization problem

$$(\mathcal{P}) : \operatorname{Inf} f_0(x),\ \ \text{subject to } f_t(x) \le 0,\ t \in T,$$

where T is a given set, and $f_0,\ f_t : \mathbb{R}^n \to \mathbb{R}_\infty,\ t \in T$, are proper and convex. We assume, without loss of generality, that $0 \notin T$, and denote

$$f := \sup_{t \in T} f_t.$$

The following result establishes new Fritz-John and KKT optimality conditions for problem $(\mathcal{P})$, improving similar results in [8,10]. Here we adopt the convention $\mathbb{R}_+ \emptyset = \{0_n\}$.

Proposition 18 *Let $\bar{x}$ be an optimal solution of $(\mathcal{P})$ such that $f(\bar{x}) = 0$. Then we have*

$$0_n \in \operatorname{co}\left\{\partial(f_0 + \mathrm{I}_{\operatorname{dom} f})(\bar{x}) \cup \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(\bar{x})} \partial_\varepsilon (f_t + \mathrm{I}_{\operatorname{dom} f \cap \operatorname{dom} f_0})(\bar{x})\right)\right\}. \tag{73}$$

Moreover, if the Slater condition holds; that is, $f(x_0) < 0$ for some $x_0 \in \operatorname{dom} f_0$, then

$$0_n \in \partial(f_0 + \mathrm{I}_{\operatorname{dom} f})(\bar{x}) + \operatorname{cone} \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(\bar{x})} \partial_\varepsilon (f_t + \mathrm{I}_{\operatorname{dom} f \cap \operatorname{dom} f_0})(\bar{x})\right) \tag{74}$$

and, provided in addition that f is continuous at some point in $\operatorname{dom} f_0 \cap \operatorname{dom} f$,

$$0_n \in \partial f_0(\bar{x}) + \operatorname{cone} \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(\bar{x})} \partial_\varepsilon f_t(\bar{x})\right) + \mathrm{N}_{\operatorname{dom} f}(\bar{x}). \tag{75}$$

Proof We consider the function $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, defined as

$$g(x) := \sup\{f_0(x) - f_0(\bar{x}),\ f_t(x), t \in T\} = \max\{f_0(x) - f_0(\bar{x}),\ f(x)\},$$

so that $\operatorname{dom} g = \operatorname{dom} f_0 \cap \operatorname{dom} f$. Then $\bar{x}$ is a global minimum of g; that is, $0_n \in \partial g(\bar{x})$. To proceed, we first apply Proposition 1 to the (finite) family $\{f_0 - f_0(\bar{x}),\ f\}$ and obtain

$$0_n \in \operatorname{co}\left\{\partial(f_0 + \mathrm{I}_{\operatorname{dom} f})(\bar{x}) \cup \partial(f + \mathrm{I}_{\operatorname{dom} f_0})(\bar{x})\right\}. \tag{76}$$

But Theorem 12, applied to the family $\left\{f_t + \mathrm{I}_{\operatorname{dom} f_0},\ t \in T\right\}$, yields

$$\partial(f + \mathrm{I}_{\operatorname{dom} f_0})(\bar{x}) = \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(\bar{x})} \partial_\varepsilon (f_t + \mathrm{I}_{\operatorname{dom} f \cap \operatorname{dom} f_0})(\bar{x})\right)\right\}, \tag{77}$$

and (73) follows from (76).

Finally, it can be easily seen from (76) that the Slater condition precludes that $0_n \in \partial(f + \mathrm{I}_{\operatorname{dom} f_0})(\bar{x})$. So, (74) follows from (73). Under the supplementary continuity condition, Corollary 13 ensures that

$$\begin{aligned}\partial(f + \mathrm{I}_{\operatorname{dom} f_0})(\bar{x}) &= \mathrm{N}_{\operatorname{dom} f_0}(\bar{x}) + \partial f(\bar{x}) \\ &= \mathrm{N}_{\operatorname{dom} f_0}(\bar{x}) + \mathrm{N}_{\operatorname{dom} f}(x) + \operatorname{co}\left\{\bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\bigcup\nolimits_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right)\right\},\end{aligned}$$

and (75) follows, taking into account (3) and

$$
\begin{aligned}
0_n &\in \partial(f_0 + \mathrm{I}_{\mathrm{dom}\, f})(\bar{x}) + \mathbb{R}_+ \partial(f + \mathrm{I}_{\mathrm{dom}\, f_0})(\bar{x}) \\
&= \partial f_0(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f}(\bar{x}) + \mathbb{R}_+ \partial(f + \mathrm{I}_{\mathrm{dom}\, f_0})(\bar{x}) \\
&= \partial f_0(\bar{x}) + \mathrm{cone}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right)\right\} + \mathrm{N}_{\mathrm{dom}\, f_0}(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f}(x) \\
&\subset \partial f_0(\bar{x}) + \mathrm{cone}\left\{\bigcap\nolimits_{\varepsilon>0} \mathrm{cl}\left(\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x).
\end{aligned}
$$

□

8 Conclusions

The main conclusion of this work is that the compactification method proposed in the paper allows us to move from the non-continuous setting to the continuous one and the other way around, as well as to develop a unifying theory which inspires new results and applications. The main results in relation to the subdifferential of the supremum are stated in Theorems 4, 7, and 12, which are established in the most general framework, free of assumptions on the index set and the data functions. Our results cover most of the existing formulas such as those obtained in [7–11,14–16,18–20,22,23,27,29–31]. The Fritz-John and KKT conditions for convex semi-infinite optimization are expressed in the most general scenario and, consequently, extend some previous results which can be found in [11,13,16,20].

References

1. Brøndsted, A.: On the subdifferential of the supremum of two convex functions. Math. Scand. **31**, 225–230 (1972)
2. Bonnans, J., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer Series in Operations Research. Springer, New York (2000)
3. Correa, R., Hantoute, A.: New formulas for the Fenchel subdifferential of the conjugate function. Set-Valued Var. Anal. **18**, 405–422 (2010)
4. Correa, R., Hantoute, A.: Subdifferential of the conjugate function in general Banach spaces. TOP **20**, 328–346 (2012)
5. Correa, R., Hantoute, A.: Lower semicontinuous convex relaxation in optimization. SIAM J. Optim. **23**, 54–73 (2013)
6. Correa, R., Hantoute, A., García, Y.: Nonconvex integration using ε-subdifferentials. Optimization **67**, 2205–2227 (2018)
7. Correa, R., Hantoute, A., López, M.A.: Weaker conditions for subdifferential calculus of convex functions. J. Funct. Anal. **271**, 1177–1212 (2016)
8. Correa, R., Hantoute, A., López, M.A.: Moreau–Rockafellar type formulas for the subdifferential of the supremum function. SIAM J. Optim. **29**, 1106–1130 (2019)
9. Correa, R., Hantoute, A., López, M.A.: Valadier-like formulas for the supremum function II: the compactly indexed case. J. Convex Anal. **26**, 299–324 (2019)
10. Correa, R., Hantoute, A., López, M.A.: Subdifferential of the supremum via compactification of the index set. Vietnam J. Math. **48**, 569–588 (2020)
11. Dinh, N., Goberna, M.A., López, M.A.: From linear to convex systems: consistency, Farkas' lemma and applications. J. Convex Anal. **13**, 113–133 (2006)
12. Fabian, M., Habala, P., Háyek, P., Montesinos, V., Pelant, J., Zizler, V.: Functional Analysis and Infinite-Dimensional Geometry. Springer, New York (2001)
13. Goberna, M.A., López, M.A.: Linear Semi-infinite Optimization. Wiley, Chichester (1998)

14. Hantoute, A., López, M.A.: A complete characterization of the subdifferential set of the supremum of an arbitrary family of convex functions. J. Convex Anal. **15**, 831–858 (2008)
15. Hantoute, A., López, M.A., Zălinescu, C.: Subdifferential calculus rules in convex analysis: a unifying approach via pointwise supremum functions. SIAM J. Optim. **19**, 863–882 (2008)
16. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms. I. Fundamentals. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 305. Springer, Berlin (1993)
17. Hiriart-Urruty, J.-B., Phelps, R.R.: Subdifferential calculus using ε-subdifferentials. J. Funct. Anal. **118**, 154–166 (1993)
18. Ioffe, A.D.: A note on subdifferentials of pointwise suprema. Top **20**, 456–466 (2012)
19. Ioffe, A.D., Levin, U.L.: Subdifferentials of convex functions. Trudy Moskov Mat. Obshch **26**, 3–73 (1972). (Russian)
20. Ioffe, A.D., Tikhomirov, V.H.: Theory of Extremal Problems. Studies in Mathematics and Its Applications, vol. 6. North-Holland, Amsterdam (1979)
21. Li, C., Ng, K.F.: Subdifferential calculus rules for supremum functions in convex analysis. SIAM J. Optim. **21**, 782–797 (2011)
22. Lopez, O., Thibault, L.: Sequential formula for subdifferential of upper envelope of convex functions. J. Nonlinear Convex Anal. **14**, 377–388 (2013)
23. López, M.A., Volle, M.: A formula for the set of optimal solutions of a relaxed minimization problem. Applications to subdifferential calculus. J. Convex Anal. **17**, 1057–1075 (2010)
24. Mordukhovich, B.S., Nghia, T.T.A.: Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data. SIAM J. Optim. **23**, 406–431 (2013)
25. Moreau, J.-J.: Fonctionnelles Convexes. In: Leray, J. (ed.) Séminaire sur les équations aux dérivées partielles, vol. 2, pp. 1–108. (1965). (French)
26. Munkres, J.: Topology, 2nd edn. Prentice Hall, Upper Saddle River (2000)
27. Pschenichnyi, B.N.: Convex programming in a normalized space. Kibernetika **5**, 46–54 (1965). ((Russian); translated as Cybernetics **1**, 46–57 (1966))
28. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
29. Solov'ev, V.N.: The subdifferential and the directional derivatives of the maximum of a family of convex functions. Izvestiya RAN: Ser. Mat. **65**, 107–132 (2001)
30. Valadier, M.: Sous-différentiel d'une borne supérieure d'une somme continue de fonctions convexes. C. R. Acad. Sci. Paris Sér. A-B **268**, A39–A42 (1969)
31. Volle, M.: Sous-différentiel d'une enveloppe supérieure de fonctions convexes. C. R. Acad. Sci. Paris Sér. I Math. **317**, 845–849 (1993)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

R. Correa[1,2] · A. Hantoute[3,4] · M. A. López[4,5]

✉ A. Hantoute
hantoute@ua.es

R. Correa
rafael.correa@uoh.cl

M. A. López
marco.antonio@ua.es

1 Universidad de O'Higgins, Rancagua, Chile

2 DIM-CMM of Universidad de Chile, Santiago, Chile

[3] Center for Mathematical Modeling (CMM), Universidad de Chile, Santiago, Chile

[4] Universidad de Alicante, Alicante, Spain

[5] CIAO, Federation University, Ballarat, Australia