# New Tour on the Subdifferential of Supremum via Finite Sums and Suprema

**A. Hantoute**[1,2] · **M. A. López-Cerdá**[1,3]

## Abstract

This paper provides new characterizations for the subdifferential of the pointwise supremum of an arbitrary family of convex functions. The main feature of our approach is that the normal cone to the effective domain of the supremum (or to finite-dimensional sections of it) does not appear in our formulas. Another aspect of our analysis is that it emphasizes the relationship with the subdifferential of the supremum of finite subfamilies, or equivalently, finite weighted sums. Some specific results are given in the setting of reflexive Banach spaces, showing that the subdifferential of the supremum can be reduced to the supremum of a countable family.

Dedicated to Prof. Franco Giannessi on his 85th birthday.

Communicated by Boris S. Mordukhovich.

✉ M. A. López-Cerdá
  marco.antonio@ua.es

  A. Hantoute
  hantoute@ua.es

[1] Universidad de Alicante, Alicante, Spain

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

[3] CIAO, Federation University, Ballarat, Australia

 Springer

## 1 Introduction

The characterization of the subdifferential of the pointwise supremum of a family of functions has attracted the attention of many researchers. Their interest comes from the fact that a huge number of important functions in convex analysis and optimization (like the Fenchel conjugate, the sum, the composition with affine mappings, etc.) can be expressed as suprema of this type. Accordingly, many publications in the last decades dealt with supremum functions and their subdifferentials and, among the most remarkable, we quote here the following ones: Brøndsted [1], Ioffe [9], Ioffe and Levin [10], Ioffe and Tikhomirov [11], Levin [12], Pschenichnyi [16], Rockafellar [17], Valadier [19], etc. See [18] to trace out the historical evolution of the topic.

More precisely, given the pointwise supremum $f := \sup_{t \in T} f_t$ of a family of convex functions $f_t : X \to \mathbb{R} \cup \{+\infty\}$, $t \in T$, $T$ being a non-empty and arbitrary set, defined on a separated locally convex space $X$, many authors addressed the problem of characterizing the subdifferential of the supremum, $\partial f(x)$, at any point $x \in \mathrm{dom}\, f$, the effective domain of $f$. These characterizations are usually given in terms of (approximate-) subdifferentials of the data functions, $\partial_\varepsilon f_t(x)$, $t \in T$, $\varepsilon \geq 0$, and, in the most general cases, in terms also of the normal cone to (finite-dimensional sections of) the effective domain of $f$, $\mathrm{N}_{L \cap \mathrm{dom}\, f}(x)$. For instance, if $f_t \in \Gamma_0(X)$, $t \in T$, where $\Gamma_0(X)$ is the family of proper convex and lower semicontinuous (lsc, in brief) functions, then the following key formula is proved in [7, Theorem 4] (see [14, Theorem 4] and [13] for related formulas):

$$\partial f(x) = \bigcap_{L \in \mathcal{F}(x), \varepsilon > 0} \overline{\mathrm{co}} \left( \bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L \cap \mathrm{dom}\, f}(x) \right), \tag{1}$$

where $\overline{\mathrm{co}}$ stands for the $w^*$-closed convex hull,

$$T_\varepsilon(x) := \{t \in T : f_t(x) \geq f(x) - \varepsilon\}, \tag{2}$$

and

$$\mathcal{F}(x) := \{L \subset X : L \text{ is a finite-dimensional linear subspace such that } x \in L\}.$$

In the so-called compact setting, which stands for assuming that $T$ is compact and the mappings $t \mapsto f_t(z)$, $z \in X$, are upper semicontinuous (usc, in brief), the following result, involving only the active functions at the reference point $x$, is established in [4, Theorem 3.8]:

$$\partial f(x) = \bigcap_{L \in \mathcal{F}(x), \varepsilon > 0} \overline{\mathrm{co}} \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L \cap \mathrm{dom}\, f}(x) \right), \tag{3}$$

where $T(x) := T_0(x)$ (see (2)).

In order to get simpler formulas, without these normal cones, one possibility is to impose additional assumptions as the continuity of $f$ at $x$, in which case (1) gives rise

to ([7, Corollary 10]; see, also, [20], for normed spaces):

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}}(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)).$$

The operation of taking the pointwise supremum is exclusive to convex analysis and has no equivalence in differential calculus. Since the sum operation is fundamental in classical calculus, many authors have been naturally led to establish a relationship between these two operations. In other words, they aimed to transform the supremum into a sum, in order to use the classical tools dealing with differentiable functions like Fermat's rule and many others.

In the case of finitely many functions $f_1, \cdots, f_n$, with $f = \max_{1\le k\le n} f_k$, it is well-known that for every $x \in X$ and $\varepsilon \ge 0$ ([21, Corollary 2.8.11], see also Lemma 11 in Appendix for an alternative proof based on the minimax theorem)

$$\partial_\varepsilon f(x) = \bigcup_{\eta\in[0,\varepsilon],\lambda\in S(x,\varepsilon-\eta)} \partial_\eta \left(\sum_{1\le k\le n} \lambda_k f_k\right)(x), \qquad (4)$$

with

$$S(x,\varepsilon-\eta) := \{\lambda \in \Delta_n : \sum_{1\le k\le n} \lambda_k f_k(x) \ge f(x) + \eta - \varepsilon\},$$

and $\Delta_n$ being the canonical simplex in $\mathbb{R}^n$.

The purpose of this paper is to establish new characterizations of $\partial f(x)$, in which only the data functions $f_t$'s appear and without involving the extra term $N_{\mathrm{dom}\, f}(x)$; namely, we provide the following more general formulas

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left( \left(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right) + \left(\bigcup_{J\in \mathcal{T}_\varepsilon(x)} \{0,\varepsilon\}\partial_\varepsilon f_J(x)\right) \right), \qquad (5)$$

where

$$\mathcal{T}_\varepsilon(x) := \{J \subset T : J \text{ finite and } \max_{t\in J} f_t(x) \ge f(x) - \varepsilon\},$$

or equivalently, using (4),

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left( \left(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right) + \left(\bigcup_{\substack{J\in \mathcal{T}_\varepsilon(x) \\ \lambda\in S_J(x,\varepsilon)}} \{0,\varepsilon\}\partial_\varepsilon \left(\sum_{t\in J} \lambda_t f_t\right)(x)\right) \right),$$

$$(6)$$

where

$$S_J(x,\varepsilon) := \{\lambda \in \Delta_{|J|} : \sum_{t\in J} \lambda_t f_t(x) \ge f(x) - \varepsilon\}.$$

Preliminary results in this direction have been obtained in [5] for the compact setting.

Both formulas (5) and (6) highlight the role played by the almost active functions at the reference point, whereas the normal cone which is present in (1) is now replaced by weighted finite maxima and sums. Formula (6) naturally covers some other formulas from the literature, as those established in [8, Theorem 2] for the case of exact sub-differentials (see, also, [15, Theorem 1]). At the same time, we prove that the choice of the involved convex combinations in (5) and (6) can be made more precise in the so-called compact setting; in fact, we establish that for any fixed $t_0 \in T(x)$ we have that

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\mathrm{co}} \left( \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) \right) \cup \left( \bigcup_{t \in T \setminus T(x)} \partial_\varepsilon (\rho_{t,\varepsilon} f_t + (1 - \rho_{t,\varepsilon}) f_{t_0})(x) \right) \right), \quad (7)$$

with

$$\rho_{t,\varepsilon} := \frac{\varepsilon}{2f(x) - 2f_t(x) + \varepsilon}, \, t \in T \setminus T(x), \varepsilon > 0.$$

The paper is structured as follows. After the section devoted to present the notation and preliminary results used in the paper, Sect. 3 provides, in Theorem 2, a representation of $\partial f(x)$, by means of specific convex combinations of the $f_t$'s which involve at most two functions. Proposition 3, first result in Sect. 4, dealing with the non-compact setting, provides the reduction of the index set $T$ to countable subsets. In this section, Theorems 4 and 5 give non-compact counterparts of the characterizations of $N_{\mathrm{dom}\,f}(x)$ and $\partial f(x)$ established in [5]. Some technical results and/or proofs are transferred to appendix, with the purpose of simplifying the presentation of the more relevant results in the paper.

## 2 Notation and Preliminary Results

Let $X$ be a (real) separated locally convex space (lcs, for short), whose topological dual space, $X^*$, is endowed with the $w^*$-topology; hence, $X^{**} := (X^*)^* \equiv X$. The spaces $X$ and $X^*$ are paired in duality by the bilinear form $(x^*, x) \in X^* \times X \mapsto \langle x^*, x \rangle := x^*(x)$. The zero vectors in $X$ and $X^*$ are denoted by $\theta$. We adopt the conventions $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$ and $0(+\infty) = +\infty$.

Given $n \geq 1$, the $n$-canonical simplex in $\mathbb{R}^n$ is

$$\Delta_n := \{(\lambda_1, \cdots, \lambda_n) \geq 0 : \lambda_1 + \cdots + \lambda_n = 1\}.$$

Given two sets $A$ and $B$ in $X$ (or in $X^*$), and $\Lambda \subset \mathbb{R}$, we define

$$A + B := \{a + b : a \in A, b \in B\} \text{ and } \Lambda A := \{\lambda a : \lambda \in \Lambda, a \in A\}. \quad (8)$$

By $\mathrm{co}(A)$ and $\mathrm{cone}(A)$, we denote the *convex* and the *conical convex hulls* of the non-empty set $A$, respectively. In the topological side, $\mathrm{cl}(A)$ and $\overline{A}$ are indistinctly

used for denoting the *closure* of $A$. When $A \subset X^*$, the closure is taken with respect to the $w^*$-topology, unless something else is explicitly stated.

Associated with a non-empty set $A \subset X$, we define the *negative dual cone* and the *orthogonal subspace* of $A$ as follows

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

respectively. Observe that $A^- = (\overline{\text{cone}}(A))^-$. These concepts are defined similarly for sets in $X^*$. The so-called *bipolar theorem* establishes that

$$A^{--} := (A^-)^- = \overline{\text{cone}}(A). \tag{9}$$

If $A \subset X$, we define the *normal cone* to $A$ at $x$ by

$$\mathrm{N}_A(x) := \begin{cases} (A - x)^-, & \text{if } x \in A, \\ \emptyset, & \text{if } x \in X \setminus A. \end{cases}$$

If $A \neq \emptyset$ is convex and closed, $A_\infty$ represents its *recession cone* defined by

$$A_\infty := \{y \in X : x + \lambda y \in A \text{ for some } x \in A \text{ and all } \lambda \geq 0\}.$$

Given a function $f : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$, its *(effective) domain* is $\mathrm{dom}\, f := \{x \in X : f(x) < +\infty\}$, and $f$ is *proper* when $\mathrm{dom}\, f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$. The *closed convex hull* of $f$, denoted by $\overline{\text{co}}\, f$, is the largest lsc convex function dominated by $f$. If $f$ is convex, then $\overline{\text{co}}\, f = \mathrm{cl}\, f$, the *closed hull* of $f$. For $x \in X$ and $\varepsilon \geq 0$, the $\varepsilon$-*subdifferential* (or the *approximate subdifferential*) of $f$ at $x$ is

$$\partial_\varepsilon f(x) = \{x^* \in X^* : f(y) \geq f(x) + \langle x^*, y - x \rangle - \varepsilon \ \text{ for all } y \in X\}, \tag{10}$$

when $f(x) \in \mathbb{R}$, and $\partial_\varepsilon f(x) := \emptyset$ when $f(x) \notin \mathbb{R}$. The *subdifferential* of $f$ at $x$ is $\partial f(x) := \partial_0 f(x)$. The $\varepsilon$-*directional derivative* of $f$ at $x \in f^{-1}(\mathbb{R})$ in the direction $u \in X$ is defined by

$$f'_\varepsilon(x; u) := \inf_{s > 0} \frac{f(x + su) - f(x) + \varepsilon}{s},$$

so that

$$\mathrm{dom}\, f'_\varepsilon(x, \cdot) = \mathbb{R}_+ (\mathrm{dom}\, f - x). \tag{11}$$

If $f \in \Gamma_0(X)$, $x \in \mathrm{dom}\, f$, and $\varepsilon > 0$, then $\partial_\varepsilon f(x) \neq \emptyset$ and we have

$$\mathrm{N}_{\mathrm{dom}\, f}(x) = (\partial_\varepsilon f(x))_\infty, \tag{12}$$

and

$$\sigma_{\partial_\varepsilon f(x)}(\cdot) = f'_\varepsilon(x, \cdot). \tag{13}$$

Formula (12) is also valid for $\varepsilon = 0$ provided that $\partial f(x) \neq \emptyset$.

The *Fenchel conjugate* of $f$ is the function $f^* : X^* \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ given by

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

and it is well-known that, for all $x \in f^{-1}(\mathbb{R})$ and $\varepsilon \geq 0$,

$$\partial_\varepsilon f(x) = \{x^* \in X^* : f(x) + f^*(x^*) \leq \langle x^*, x \rangle + \varepsilon\},$$

and $\partial f(x) = \cap_{\varepsilon > 0} \partial_\varepsilon f(x)$.

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

$$\sigma_A(x^*) := \sup\{\langle x^*, x \rangle : x \in A\}, x^* \in X^*,$$

with $\sigma_\emptyset \equiv -\infty$, and

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

It is known that, if $A$ is a closed convex set,

$$A_\infty = (\mathrm{dom}\, \sigma_A)^-, \tag{14}$$

or equivalently, by using (9),

$$(A_\infty)^- = \mathrm{cl}(\mathrm{dom}\, \sigma_A). \tag{15}$$

Next, given a finite family $\{f_k, 1 \leq k \leq n\} \subset \Gamma_0(X)$, we consider the maximum function $f = \max_{1 \leq k \leq n} f_k$. We suppose that $f$ is proper and denote

$$\varphi(\lambda, x) := \sum_{1 \leq k \leq n} \lambda_k f_k(x) - \mathrm{I}_{\mathbb{R}^n_+}(\lambda), \lambda \in \mathbb{R}^n, x \in X. \tag{16}$$

The adopted convention $0(+\infty) = +\infty$ entails $0 f_k = \mathrm{I}_{\mathrm{dom}\, f_k}$. Then $\varphi(\Delta_n, \mathrm{dom}\, f) \subset \mathbb{R}$, $\varphi(\cdot, x)$ is concave and usc for every $x \in \mathrm{dom}\, f$, and $\varphi(\lambda, \cdot)$ is convex and lsc for every $\lambda \in \Delta_n$. Thus, since $\Delta_n$ is compact in $\mathbb{R}^n$ and $\mathrm{dom}\, f$ is non-empty and convex, the minimax theorem ensures that (see, e.g., [21, Theorem 2.10.2])

$$\max_{\lambda \in \Delta_n} \inf_{x \in \mathrm{dom}\, f} \varphi(\lambda, x) = \inf_{x \in \mathrm{dom}\, f} \max_{\lambda \in \Delta_n} \varphi(\lambda, x). \tag{17}$$

Moreover, since

$$f(x) = \max_{\lambda \in \Delta_n} \varphi(\lambda, x), \text{ for all } x \in \mathrm{dom}\, f,$$

relation (17) leads us to

$$\max_{\lambda \in \Delta_n} \inf_{x \in X} \varphi(\lambda, x) = \inf_{x \in X} f(x). \tag{18}$$

As a consequence of this, for every $x \in \text{dom } f$ and $\varepsilon \geq 0$ we obtain that (see Lemma 11 in Appendix)

$$\partial_\varepsilon f(x) = \bigcup_{\lambda \in \Delta_n} \partial_{\varepsilon + \varphi(\lambda, x) - f(x)} \varphi(\lambda, \cdot)(x) \tag{19}$$

$$= \bigcup_{\eta \in [0, \varepsilon], \lambda \in S(x, \varepsilon - \eta)} \partial_\eta \varphi(\lambda, \cdot)(x), \tag{20}$$

where $S(x, \varepsilon - \eta) := \{\lambda \in \Delta_n, \eta \in [0, \varepsilon], \varphi(\lambda, x) \geq f(x) + \eta - \varepsilon\}$. Notice that formula (19) constitutes a slight improvement of [21, Corollary 2.8.11] as it involves only one precise value of the parameter $\eta$.

The arguments used in Lemma 11 to prove (19) and (20) are specific to finite families of functions, and so they cannot be extended to families with infinitely many functions, where the following simplices in $\mathbb{R}^T$,

$$\left\{\lambda \in \mathbb{R}_+^T : \lambda(t) \equiv \lambda_t = 0 \text{ except for finitely many } t\text{'s and } \sum_{\lambda_t > 0} \lambda_t = 1\right\}, \tag{21}$$

may be not compact.

## 3 The Compact Setting

We give in this section some additional results to those established in [5]. We consider a non-empty family $\{f_t, \ t \in T\} \subset \Gamma_0(X)$ such that

$$\begin{array}{l} T \text{ is Hausdorff compact,} \\ \text{and, for each } z \in X, \text{ the mapping } t \longmapsto f_t(z) \text{ is upper semicontinuous.} \end{array} \tag{22}$$

The associated supremum function is

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

and assumptions (22) ensure that (see [5, Lemma 5])

$$\text{dom } f = \cap_{t \in T} \text{dom } f_t \tag{23}$$

and, for every $x \in \text{dom } f$,

$$\mathbb{R}_+(\text{dom } f - x) = \cap_{t \in T} \mathbb{R}_+(\text{dom } f_t - x). \tag{24}$$

Moreover, (22) also yields

$$T(x) := \{t \in T : f_t(x) = f(x)\} \neq \emptyset.$$

Assuming $\inf_{t \in T} f_t(x) > -\infty$, we have proved in [5, Theorem 12] that

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\mathrm{co}} \left( \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) \right) + \left( \bigcup_{t \in T \setminus T(x)} \{0, \varepsilon\} \, \partial_\varepsilon f_t(x) \right) \right). \quad (25)$$

This formula involves the active functions $f_t$, $t \in T(x)$, as the same time as the non-active ones $f_t$, $t \in T \setminus T(x)$, but with these last ones being affected by the weighting parameter $\varepsilon > 0$. The main ingredient we used to establish (25) is the following relation ([5, Theorem 6])

$$N_{\mathrm{dom}\, f}(x) = \left[ \overline{\mathrm{co}} \left( \bigcup_{t \in T} \partial_\varepsilon f_t(x) \right) \right]_\infty, \text{ for every } \varepsilon > 0.$$

We give next an equivalent description of the elements in $N_{\mathrm{dom}\, f}(x)$, which highlights the role played by the active and non-active functions.

**Lemma 1** *Assume that* (22) *holds. Consider* $x \in \mathrm{dom}\, f$ *and fix* $t_0 \in T(x)$. *Then we have that*

$$N_{\mathrm{dom}\, f}(x) \subset \left[ \overline{\mathrm{co}} \left( \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) \right) \cup \left( \bigcup_{t \in T \setminus T(x)} \partial_\varepsilon (\mu_t f_t + (1 - \mu_t) f_{t_0})(x) \right) \right) \right]_\infty,$$
$$(26)$$

*for every* $\varepsilon > 0$ *and* $0 < \mu_t < 1$.

**Proof** We fix $\varepsilon > 0$ and $0 < \mu_t < 1$, for all $t \in T \setminus T(x)$, and denote

$$E_\varepsilon := \bigcup_{t \in T} \partial_\varepsilon \tilde{f}_t(x),$$

where

$$\tilde{f}_t := \begin{cases} f_t, & \text{if } t \in T(x), \\ \mu_t f_t + (1 - \mu_t) f_{t_0}, & \text{if } t \in T \setminus T(x). \end{cases} \quad (27)$$

The sets $T(x)$ and $E_\varepsilon$ are non-empty thanks to (22) and the lower semicontinuity of the $f_t$'s. Since

$$\mathrm{cl}(\mathrm{dom}\, \sigma_{E_\varepsilon}) = ([\overline{\mathrm{co}}\, (E_\varepsilon)]_\infty)^-,$$

by (15), and $(N_{\mathrm{dom}\, f}(x))^- = (\mathrm{dom}\, f - x)^{--} = \mathrm{cl}(\mathbb{R}_+(\mathrm{dom}\, f - x))$, by (9), desired relation (26) is equivalent to

$$\mathrm{cl}(\mathrm{dom}\, \sigma_{E_\varepsilon}) \subset \mathrm{cl}(\mathbb{R}_+(\mathrm{dom}\, f - x)).$$

To prove this inclusion we take, using (13),

$$z \in \operatorname{dom} \sigma_{E_\varepsilon} = \operatorname{dom} \left( \sigma_{\bigcup_{t \in T} \partial_\varepsilon \tilde{f}_t(x)} \right) = \operatorname{dom} \left( \sup_{t \in T} \sigma_{\partial_\varepsilon \tilde{f}_t(x)} \right)$$

$$= \operatorname{dom} \left( \sup_{t \in T} (\tilde{f}_t)'_\varepsilon(x; \cdot) \right) \subset \cap_{t \in T} \operatorname{dom} (\tilde{f}_t)'_\varepsilon(x; \cdot). \qquad (28)$$

Hence, since for every $t \in T \setminus T(x)$, by (11),

$$\begin{aligned}
\operatorname{dom}(\tilde{f}_t)'_\varepsilon(x; \cdot) &= \mathbb{R}_+ (\operatorname{dom} \tilde{f}_t - x) \\
&= \mathbb{R}_+ ((\operatorname{dom}(\mu_t f_t) \cap \operatorname{dom}((1 - \mu_t) f_{t_0})) - x) \\
&= \mathbb{R}_+ ((\operatorname{dom} f_t \cap \operatorname{dom} f_{t_0}) - x) \\
&= (\mathbb{R}_+ \operatorname{dom}(f_t - x)) \cap (\mathbb{R}_+ \operatorname{dom}(f_{t_0} - x)),
\end{aligned}$$

relation (28) entails

$$\begin{aligned}
z &\in \left( \cap_{t \in T(x)} \operatorname{dom}(f_t)'_\varepsilon(x; \cdot) \right) \cap \left( \cap_{t \in T \setminus T(x)} \operatorname{dom}(\tilde{f}_t)'_\varepsilon(x; \cdot) \right) \\
&= \left( \cap_{t \in T(x)} \mathbb{R}_+ (\operatorname{dom} f_t - x) \right) \cap \left( \cap_{t \in T \setminus T(x)} (\mathbb{R}_+ \operatorname{dom}(f_t - x)) \cap (\mathbb{R}_+ \operatorname{dom}(f_{t_0} - x)) \right) \\
&= \cap_{t \in T} \mathbb{R}_+ \operatorname{dom}(f_t - x),
\end{aligned}$$

and, so (24) gives rise to

$$z \in \cap_{t \in T} \mathbb{R}_+ \operatorname{dom}(f_t - x) = \mathbb{R}_+ (\operatorname{dom} f - x) \subset \operatorname{cl}(\mathbb{R}_+ (\operatorname{dom} f - x)).$$

Hence, $\operatorname{dom} \sigma_{E_\varepsilon} \subset \operatorname{cl}(\mathbb{R}_+ (\operatorname{dom} f - x))$ and the desired inclusion follows. $\qquad \square$

The main purpose of this section is to obtain another representation of $\partial f(x)$, which involves appropriate convex combinations of the non-active $f_t$'s. In the non-compact setting, instead of considering two-elements convex combinations as in the compact framework, we shall appeal to all finite-elements convex combinations of the $f_t$'s (see Theorem 5 below).

**Theorem 2** *Assume that hypothesis* (22) *fulfills. Consider* $x \in \operatorname{dom} f$ *and choose* $t_0 \in T(x)$. *Then we have that*

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\operatorname{co}} \left( \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) \right) \cup \left( \bigcup_{t \in T \setminus T(x)} \partial_\varepsilon (\rho_{t,\varepsilon} f_t + (1 - \rho_{t,\varepsilon}) f_{t_0})(x) \right) \right),$$
(29)

*where*

$$\rho_{t,\varepsilon} := \frac{\varepsilon}{2 f(x) - 2 f_t(x) + \varepsilon}, \, t \in T \setminus T(x), \varepsilon > 0.$$

**Proof** Let us suppose, for simplicity, that $f(x) = 0$, and observe that for each given $\varepsilon > 0$ we have, for every $t \in T \setminus T(x)$,

$$0 < \rho_{t,\varepsilon} < 1 \text{ and } \rho_{t,\varepsilon} f_t(x) > -\frac{\varepsilon}{2}. \tag{30}$$

Let us also denote

$$\tilde{f}_{t,\varepsilon} := \begin{cases} f_t, & \text{if } t \in T(x), \\ \rho_{t,\varepsilon} f_t + (1 - \rho_{t,\varepsilon}) f_{t_0}, & \text{if } t \in T \setminus T(x). \end{cases}$$

Then, for all $t \in T \setminus T(x)$,

$$\tilde{f}_{t,\varepsilon}(x) = \rho_{t,\varepsilon} f_t(x) > -\frac{\varepsilon}{2},$$

and so, observing that $\tilde{f}_{t,\varepsilon} \leq \max\{f_t, f_{t_0}\} \leq f$,

$$\bigcup_{t \in T \setminus T(x)} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) \subset \partial_{\frac{3\varepsilon}{2}} f(x) \subset \partial_{2\varepsilon} f(x).$$

Thus, since we also have

$$\bigcup_{t \in T(x)} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) \subset \partial_\varepsilon f(x) \subset \partial_{2\varepsilon} f(x),$$

we conclude that

$$\bigcup_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) \subset \partial_{2\varepsilon} f(x),$$

and the inclusion "$\supset$" follows by taking the closed convex hull and intersecting over $\varepsilon > 0$.

To establish the inclusion "$\subset$", we fix $\varepsilon > 0$ and $L \in \mathcal{F}(x)$. Next, by applying Lemma 1 to the family $\{\tilde{f}_{t,\varepsilon}, t \in T; I_L\}$ we obtain that

$$N_{L \cap \text{dom} f}(x) \subset \left[ \overline{\text{co}} \left( \bigcup_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) \cup L^\perp \right) \right]_\infty = \left[ \overline{\text{co}} \left( \bigcup_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) + L^\perp \right) \right]_\infty,$$

where the last equality comes from (47). Therefore, by (3),

$$\partial f(x) \subset \overline{\text{co}} \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) + N_{L \cap \text{dom} f}(x) \right)$$

$$= \overline{\text{co}} \left( \bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) + \left[ \overline{\text{co}} \left( \bigcup_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) + L^\perp \right) \right]_\infty \right)$$

$$\subset \overline{\text{co}} \left( \bigcup_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) + L^\perp \right) = \text{cl} \left( \text{co} \left( \bigcup_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x) \right) + L^\perp \right).$$

Intersecting over the $L$'s in $\mathcal{F}(x)$ we get

$$\partial f(x) \subset \bigcap\nolimits_{L \in \mathcal{F}(x)} \text{cl}\left(\text{co}\left(\bigcup\nolimits_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x)\right) + L^\perp\right) = \overline{\text{co}}\left(\bigcup\nolimits_{t \in T} \partial_\varepsilon \tilde{f}_{t,\varepsilon}(x)\right),$$

where the last equality is due to the fact that, for every $A \subset X^*$ (see ([3, Lemma 3])),

$$\bigcap\nolimits_{L \in \mathcal{F}(x)} \text{cl}\left(A + L^\perp\right) = \text{cl}\,(A). \tag{31}$$

$\square$

In the particular case when all the $f_t$'s are active at $x$, that is, $T(x) = T$, formula (29) reduces to

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\text{co}}\left(\bigcup_{t \in T} \partial_\varepsilon f_t(x)\right),$$

which extends the well-known Brøndsted formula [1] to infinite index sets. Another illustration of Theorem 2 is the alternative proof of formula (51) in Appendix.

## 4 Non-Compact Framework

This section is devoted to give new characterizations of $N_{\text{dom}\,f}(x)$ and $\partial f(x)$, without any additional assumptions on the family $\{f_t, t \in T\} \subset \Gamma_0(X)$.

The first result, whose proof is postponed to Appendix, provides the reduction of the index set $T$ to countable subsets within the normal cone of dom $f$.

**Proposition 3** *Consider a family $\{f_t, t \in T\} \subset \Gamma_0(X)$ and $f = \sup_{t \in T} f_t$. Given $x \in \text{dom}\, f$ and $u^* \in N_{\text{dom}\,f}(x)$, for each $L \in \mathcal{F}(x)$ there is a sequence $(t_n)_n \subset T$ such that*

$$u^* \in N_{\text{dom}(\sup_{n \geq 1} f_{t_n}) \cap L}(x).$$

The following result provides the non-compact counterpart of the characterizations of $N_{\text{dom}\,f}(x)$ established in [5].

**Theorem 4** *Consider the family $\{f_t, t \in T\} \subset \Gamma_0(X)$ and $f := \sup_{t \in T} f_t$. Given $x \in \text{dom}\, f$, for every $\varepsilon > 0$ we have that*

$$N_{\text{dom}\,f}(x) \subset \left(\overline{\text{co}}\left(\bigcup\nolimits_{J \in \mathcal{T}} \partial_\varepsilon f_J(x)\right)\right)_\infty, \tag{32}$$

*where $\mathcal{T} := \{J \subset T, |J| < +\infty\}$ and*

$$f_J := \max\{f_t, t \in J\}, J \in \mathcal{T}. \tag{33}$$

*In addition,* (32) *becomes an equality when*

$$\inf_{J \in \mathcal{T}} f_J(x) > -\infty. \tag{34}$$

**Remark 1** (Before the proof ) *Condition* (34) *is not very restrictive, indeed, it suffices to choose* $t_0 \in T$ *and consider the family* $\{\max\{f_t, f_{t_0}\}, \ t \in T\}$. *This new family obviously satisfies condition* (34),

$$\inf_{J \in \widetilde{\mathcal{T}}} f_J(x) \geq f_{t_0}(x),$$

*where* $\widetilde{\mathcal{T}} := \{J \in \mathcal{T} : t_0 \in J\}$, *and, consequently, Theorem* 4 *yields*

$$\mathrm{N}_{\mathrm{dom}\, f}(x) = \left( \overline{\mathrm{co}} \left( \bigcup\nolimits_{J \in \widetilde{\mathcal{T}}} \partial_\varepsilon f_J(x) \right) \right)_\infty.$$

**Proof** Take $u^* \in \mathrm{N}_{\mathrm{dom}\, f}(x)$ and $\varepsilon > 0$. Then, by Proposition 3, for every fixed $L \in \mathcal{F}(x)$ there exists a sequence $(t_n)_n \subset T$ such that

$$u^* \in \mathrm{N}_{\mathrm{dom}(\sup_{n \geq 1} f_{t_n}) \cap L}(x).$$

We denote $J_n := \{t_1, \cdots, t_n\}$, $n \geq 1$, and introduce the functions

$$\hat{f}_n := f_{J_n} + \mathrm{I}_L, n \geq 1,$$

where $f_{J_n} = \max\{f_t, \ t \in J_n\}$ (see (33)). So, $(\hat{f}_n)_n$ is non-decreasing and

$$\sup\nolimits_{n \geq 1} \left( f_{t_n} + \mathrm{I}_L \right) = \sup\nolimits_{n \geq 1} \hat{f}_n \text{ and } \mathrm{dom} \left( \sup\nolimits_{n \geq 1} f_{t_n} \right) \cap L = \mathrm{dom} \left( \sup\nolimits_{n \geq 1} \hat{f}_n \right).$$

In addition, according to Lemma 12 and (46), we have that

$$\begin{aligned}
\partial_{\frac{\varepsilon}{2}} (\sup\nolimits_{n \geq 1} \hat{f}_n)(x) &= \bigcap_{\delta > 0} \mathrm{cl} \left( \bigcup\nolimits_{k \geq 1} \bigcap\nolimits_{n \geq k} \partial_{\frac{\varepsilon}{2} + \delta} \hat{f}_n(x) \right) \\
&\subset \mathrm{cl} \left( \bigcup\nolimits_{k \geq 1} \partial_\varepsilon (f_{J_k} + \mathrm{I}_L)(x) \right) \\
&\subset \mathrm{cl} \left( \bigcup\nolimits_{k \geq 1} \mathrm{cl}(\partial_\varepsilon f_{J_k}(x) + L^\perp) \right) \\
&\subset \overline{\mathrm{co}} \left( \bigcup\nolimits_{k \geq 1} \partial_\varepsilon f_{J_k}(x) + L^\perp \right).
\end{aligned}$$

Therefore, using (12),

$$\begin{aligned}
u^* \in \mathrm{N}_{\mathrm{dom}(\sup_{n \geq 1} f_{t_n}) \cap L}(x) &= (\partial_{\frac{\varepsilon}{2}} (\sup\nolimits_{n \geq 1} \hat{f}_n)(x))_\infty \\
&\subset \left( \overline{\mathrm{co}} \left( \bigcup\nolimits_{k \geq 1} \partial_\varepsilon f_{J_k}(x) + L^\perp \right) \right)_\infty,
\end{aligned}$$

that is, for all $L \in \mathcal{F}(x)$,

$$u^* \in \left( \overline{\mathrm{co}} \left( \bigcup_{k \geq 1} \partial_\varepsilon f_{J_k}(x) + L^\perp \right) \right)_\infty \subset \left( \overline{\mathrm{co}} \left( \bigcup_{J \in \mathcal{T}} \partial_\varepsilon f_J(x) + L^\perp \right) \right)_\infty,$$

and so

$$
\begin{aligned}
u^* &\in \bigcap_{L \in \mathcal{F}(x)} \left( \overline{\mathrm{co}} \left( \bigcup_{J \in \mathcal{T}} \partial_\varepsilon f_J(x) + L^\perp \right) \right)_\infty \\
&= \left( \bigcap_{L \in \mathcal{F}(x)} \overline{\mathrm{co}} \left( \bigcup_{J \in \mathcal{T}} \partial_\varepsilon f_J(x) + L^\perp \right) \right)_\infty \\
&= \left( \overline{\mathrm{co}} \left( \bigcup_{J \in \mathcal{T}} \partial_\varepsilon f_J(x) \right) \right)_\infty,
\end{aligned}
$$

where the last equality is a consequence of (31).

For the converse inclusion, observe that (34) implies the existence of a constant $M$ such that

$$\inf_{J \in \mathcal{T}} f_J(x) \geq M (> -\infty).$$

Then, for every $J \in \mathcal{T}$ and $x^* \in \partial_\varepsilon f_J(x)$,

$$
\begin{aligned}
\langle x^*, y - x \rangle &\leq f_J(y) - f_J(x) + \varepsilon \\
&\leq f(y) - f(x) + (f(x) - M + \varepsilon), \text{ for all } y \in X;
\end{aligned}
$$

in other words, $\partial_\varepsilon f_J(x) \subset \partial_{\varepsilon + f(x) - M} f(x)$ and so

$$\left( \overline{\mathrm{co}} \left( \bigcup_{J \in \mathcal{T}} \partial_\varepsilon f_J(x) \right) \right)_\infty \subset \left( \partial_{\varepsilon + f(x) - M} f(x) \right)_\infty = \mathrm{N}_{\mathrm{dom}\, f}(x),$$

where the last equality comes from (12). $\qquad \square$

Next, we give the main result in this section, which constitutes a non-compact counterpart of Theorem 2.

**Theorem 5** *Consider the family $\{ f_t, \, t \in T \} \subset \Gamma_0(X)$ and $f := \sup_{t \in T} f_t$. Then for every $x \in \mathrm{dom}\, f$ we have that*

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\mathrm{co}} \left( \left( \bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) \right) + \{0, \varepsilon\} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \right), \quad (35)$$

*where*

$$\mathcal{T}_\varepsilon(x) := \{ J \in \mathcal{T} : f_J(x) \geq f(x) - \varepsilon \}. \quad (36)$$

**Proof** Fix $x \in \mathrm{dom}\, f$ and $\varepsilon > 0$ so that, by formula (1), and whichever $L \in \mathcal{F}(x)$ we take, one has

$$\partial f(x) \subset \overline{\mathrm{co}} \left( \bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L \cap \mathrm{dom}\, f}(x) \right). \quad (37)$$

Now we pick $t_0 \in T_\varepsilon(x)$, and denote $\widehat{\mathcal{T}} := \{J \in \mathcal{T} : t_0 \in J\}$,

$$f_J := \max\{f_t, t \in J\} + I_L, \quad J \in \widehat{\mathcal{T}},$$

so that

$$f_J(x) \geq f_{t_0}(x) \geq f(x) - \varepsilon, \text{ for every } J \in \widehat{\mathcal{T}}, \tag{38}$$

and

$$\widehat{\mathcal{T}} \subset T_\varepsilon(x). \tag{39}$$

Then, by Remark 1, and taking into account (47), (48), and (39),

$$N_{L \cap \mathrm{dom}\, f}(x) = \left(\overline{\mathrm{co}} \left(\bigcup_{J \in \widehat{\mathcal{T}}} \partial_\varepsilon f_J(x)\right)\right)_\infty$$
$$\subset \left(\overline{\mathrm{co}} \left(\left(\bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right) \bigcup \left(\bigcup_{J \in \widehat{\mathcal{T}}} \partial_\varepsilon f_J(x)\right)\right)\right)_\infty \subset (\overline{\mathrm{co}}\,(E_\varepsilon))_\infty,$$

where we have denoted

$$E_\varepsilon := \left(\bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right) + \{0, \varepsilon\} \left(\bigcup_{J \in T_\varepsilon(x)} \partial_\varepsilon f_J(x)\right). \tag{40}$$

So, (37) gives rise to

$$\partial f(x) \subset \overline{\mathrm{co}} \left(\bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + (\overline{\mathrm{co}}\,(E_\varepsilon))_\infty\right) \subset \overline{\mathrm{co}}\,(E_\varepsilon),$$

that is, the desired inclusion "$\subset$" follows once we intersect over $\varepsilon > 0$.

To verify the opposite inclusion, by (36) we easily observe that

$$\left(\bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x)\right) \bigcup \left(\bigcup_{J \in T_\varepsilon(x)} \partial_\varepsilon f_J(x)\right) \subset \partial_{2\varepsilon} f(x),$$

and so,

$$\bigcap_{\varepsilon > 0} \overline{\mathrm{co}}\,(E_\varepsilon) \subset \bigcap_{\varepsilon > 0} \overline{\mathrm{co}}\,(\partial_{2\varepsilon} f(x) + \{0, \varepsilon\} \partial_{2\varepsilon} f(x))$$
$$\subset \bigcap_{\varepsilon > 0} [1, 1 + \varepsilon]\, \partial_{2\varepsilon} f(x) = \partial f(x).$$

$\square$

For $x \in \mathrm{dom}\, f$, $\delta \geq 0$ and $J \in \mathcal{T}$, we denote

$$S_J(x, \delta) := \left\{\lambda \in \Delta_{|J|} : \sum_{t \in J} \lambda_t f_t(x) \geq f(x) - \delta\right\}.$$

Observe that

$$S_J(x, 0) = \left\{\lambda \in \Delta_{|J|} : \sum_{t \in J} \lambda_t f_t(x) = f(x)\right\}.$$

Theorem 5 leads us to the characterization below, involving the finite suprema $f_J$ or sums $\sum_{t \in J} \lambda_t f_t$.

**Corollary 6** *Consider the family* $\{f_t, \ t \in T\} \subset \Gamma_0(X)$ *and* $f := \sup_{t \in T} f_t$. *Then for every* $x \in \operatorname{dom} f$ *we have that*

$$\partial f(x) = \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \tag{41}$$

*and, consequently,*

$$\partial f(x) = \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \in \mathcal{T}, \lambda \in S_J(x,\varepsilon)} \partial_\varepsilon \left( \sum_{t \in J} \lambda_t f_t \right)(x) \right). \tag{42}$$

*Proof* Fix $x \in \operatorname{dom} f$ and $\varepsilon > 0$. Since $\{\{t\} : t \in T_\varepsilon(x)\} \subset \mathcal{T}_\varepsilon(x)$, we have that (recall the definition of $E_\varepsilon$ in (40))

$$E_\varepsilon = \left( \bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) \right) + \{0, \varepsilon\} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right)$$
$$\subset [1, 1 + \varepsilon] \overline{\operatorname{co}} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right),$$

and so, by Theorem 5 and Lemma 13 (for the second inclusion),

$$\partial f(x) \subset \bigcap_{\varepsilon > 0} [1, 1 + \varepsilon] \overline{\operatorname{co}} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right)$$
$$= \bigcap_{\varepsilon > 0} \overline{\operatorname{co}} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right)$$
$$\subset \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_{2\varepsilon} f_J(x) \right)$$
$$\subset \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \in \mathcal{T}_{2\varepsilon}(x)} \partial_{2\varepsilon} f_J(x) \right).$$

Hence, the inclusion "$\subset$" in (41) follows.

To verify the opposite inclusion, take $x^* \in \partial_\varepsilon f_J(x)$, $J \in \mathcal{T}_\varepsilon(x)$, and $\varepsilon > 0$. Then, for every $y \in X$,

$$\langle x^*, y - x \rangle \le f_J(y) - f_J(x) + \varepsilon \le f(y) - (f(x) - \varepsilon) + \varepsilon = f(y) - f(x) - 2\varepsilon,$$

and so $\partial_\varepsilon f_J(x) \subset \partial_{2\varepsilon} f(x)$. Thus,

$$\bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \subset \bigcap_{\varepsilon > 0} \partial_{2\varepsilon} f(x) = \partial f(x),$$

and we are done with the first statement.

Finally, using (20), formula (41) implies that

$$\partial f(x) = \bigcap_{\varepsilon>0} \mathrm{cl}\left(\bigcup_{J\in\mathcal{T}_\varepsilon(x),\eta\in[0,\varepsilon],\lambda\in S_J(x,\varepsilon-\eta)} \partial_\eta\left(\sum_{t\in J}\lambda_t f_t\right)(x)\right)$$
$$\subset \bigcap_{\varepsilon>0} \mathrm{cl}\left(\bigcup_{J\in\mathcal{T}_\varepsilon(x),\lambda\in S_J(x,\varepsilon)} \partial_\varepsilon\left(\sum_{t\in J}\lambda_t f_t\right)(x)\right)$$
$$= \bigcap_{\varepsilon>0} \mathrm{cl}\left(\bigcup_{J\in\mathcal{T},\lambda\in S_J(x,\varepsilon)} \partial_\varepsilon\left(\sum_{t\in J}\lambda_t f_t\right)(x)\right),$$

where we use the inclusions $S_J(x,\varepsilon-\eta) \subset S_J(x,\varepsilon)$ and $\partial_\eta g(x) \subset \partial_\varepsilon g(x)$, for any convex function $g$. Moreover, the converse inclusion follows by observing that

$$\partial_\varepsilon\left(\sum_{t\in J}\lambda_t f_t\right)(x) \subset \partial_{2\varepsilon} f(x),$$

for all $J \in \mathcal{T}$ and $\lambda \in S_J(x,\varepsilon)$. $\qquad\square$

**Remark 2** Let us emphasize at this point that the main feature of our approach is to provide characterizations of $\partial f(x)$, which are independent of the effective domains of the involved functions and the associated normal cones. For comparative purposes, we quote here the following formula, given in [15, Theorem 1],

$$\partial f(x) = \bigcap_{\varepsilon>0} \mathrm{cl}\left(\bigcup_{J\in\mathcal{T},\lambda\in S_J(x,\varepsilon)} \partial_\varepsilon\left(\sum_{t\in J}\lambda_t f_t + \mathrm{I}_D\right)(x)\right), \qquad (43)$$

with $D$ being any subset of $X$ satisfying

$$\mathrm{dom}\, f \subset D \subset \bigcap_{t\in T} \mathrm{dom}\, f_t.$$

Observe that formula (43) requires the use of the augmented functions $f_t + \mathrm{I}_D$ and not the exact ones $f_t$'s as in (42). The following example illustrates the difference between (42) and (43).

**Example 1** Consider the support function of a non-empty set $T \subset X^*$,

$$\sigma_T(x) := \sup_{t\in T} \langle t, x\rangle.$$

Here, $f = \sigma_T = \sup_{t\in T} f_t$ with $f_t(x) := \langle t, x\rangle$, $t \in T$, in $\Gamma_0(X)$. On the one hand, for every $x \in X$, formula (42) yields

$$\partial f(x) = \bigcap_{\varepsilon>0} \mathrm{cl}\left(\bigcup_{J\in\mathcal{T},\lambda\in S_J(x,\varepsilon)} \partial_\varepsilon\left(\sum_{t\in J}\lambda_t f_t\right)(x)\right)$$
$$= \bigcap_{\varepsilon>0} \mathrm{cl}\left(\bigcup_{J\in\mathcal{T},\lambda\in S_J(x,\varepsilon)}\left(\sum_{t\in J}\lambda_t t\right)\right)$$
$$= \bigcap_{\varepsilon>0} \mathrm{cl}\left(\{x^* \in \mathrm{co}\, T : \langle x^*, x\rangle \geq f(x) - \varepsilon\}\right)$$
$$= \bigcap_{\varepsilon>0}\{x^* \in \overline{\mathrm{co}}T : \langle x^*, x\rangle \geq f(x) - \varepsilon\},$$

which is well-known (see, for instance, [6, (5) in page 834]); actually, it is a consequence of the Fenchel equality as

$$\sigma_T(x) + I_{\overline{co}T}(x^*) \leq \langle x^*, x \rangle + \varepsilon \iff x^* \in \partial_\varepsilon \sigma_T(x) \text{ for all } \varepsilon \geq 0.$$

On the other hand, if we apply formula (43) choosing $D = \operatorname{dom} \sigma_T$, then we obtain that

$$\partial f(x) = \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \in \mathcal{T}, \lambda \in S_J(x,\varepsilon)} \left( \sum_{t \in J} \lambda_t t \right) + N_D^\varepsilon(x) \right).$$

Hence, using Lemma 14, we derive the following alternative representation of $\partial f(x)$,

$$\partial f(x) = \bigcap_{\varepsilon > 0} \operatorname{cl} \left\{ \begin{array}{l} x^* + y^* : x^* \in \operatorname{co} T, \langle x^*, x \rangle \geq f(x) - \varepsilon, \\ y^* \in (\overline{co}T)_\infty, -\varepsilon \leq \langle y^*, x \rangle \leq 0 \end{array} \right\},$$

which appeals to the extra term $\{y^* \in (\overline{co}T)_\infty : -\varepsilon \leq \langle y^*, x \rangle \leq 0\}$.

We apply Corollary 6 to provide a new proof for the characterization of the normal cone to sublevel sets given in [8, Corollary 7] (see, also, [2] and references therein).

**Corollary 7** *Consider a function $g \in \Gamma_0(X)$ and let $x \in X$ such that $g(x) = 0$. Then we have that*

$$N_{[g \leq 0]}(x) = \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{\mu > 0} \partial_\varepsilon (\mu g)(x) \right). \tag{44}$$

*Proof* We define the functions

$$f_t := tg, t > 0, \text{ and } f := \sup_{t > 0} f_t.$$

Obviously, $\{f_t, t \in T\} \subset \Gamma_0(X)$ and $f_t(x) = f(x) = 0$ for all $t > 0$. Therefore, since that $f = I_{[g \leq 0]}$, by formula (42) we obtain that

$$\begin{aligned} N_{[g \leq 0]}(x) &= \partial f(x) \\ &= \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \subset ]0, +\infty[, |J| < \infty, \lambda \in S_J(x,\varepsilon)} \partial_\varepsilon \left( \sum_{t \in J} \lambda_t tg \right)(x) \right), \end{aligned}$$

where

$$S_J(x, \varepsilon) = \left\{ \lambda \in \Delta_{|J|} : \sum_{t \in J} \lambda_t tg(x) \geq -\varepsilon \right\} = \Delta_{|J|}.$$

Hence,

$$\begin{aligned} N_{[g \leq 0]}(x) &= \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{J \subset ]0, +\infty[, |J| < \infty, \lambda \in \Delta_{|J|}} \partial_\varepsilon \left( \sum_{t \in J} \lambda_t tg \right)(x) \right) \\ &= \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \left\{ \bigcup \partial_\varepsilon (\mu g)(x) : \mu = \sum_{t \in J} \lambda_t t, J \subset ]0, +\infty[, |J| < \infty, \lambda \in \Delta_{|J|} \right\} \right) \\ &= \bigcap_{\varepsilon > 0} \operatorname{cl} \left( \bigcup_{\mu > 0} \partial_\varepsilon (\mu g)(x) \right), \end{aligned}$$

and we are done. □

The following corollary gives more insight to the conclusion of Corollary 6 in reflexive Banach spaces.

**Corollary 8** *If X is a reflexive Banach space, then* (41) *and* (42) *also hold when the closure is taken with respect to the strong (norm) topology.*

**Proof** It suffices to prove formula (41). Given $x \in \text{dom } f$, by Corollary 6 we have that

$$
\begin{aligned}
\partial f(x) &= \bigcap_{\varepsilon>0} \text{cl} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \\
&\subset \bigcap_{\varepsilon>0} \overline{\text{co}} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \\
&= \bigcap_{\varepsilon>0} \text{cl}^{\|\cdot\|_*} \left( \text{co} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \right),
\end{aligned}
$$

due to Mazur's theorem. Next, taking into account (54), we obtain that

$$
\begin{aligned}
\partial f(x) &\subset \bigcap_{\varepsilon>0} \text{cl}^{\|\cdot\|_*} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_{2\varepsilon} f_J(x) \right) \\
&\subset \bigcap_{\varepsilon>0} \text{cl}^{\|\cdot\|_*} \left( \bigcup_{J \in \mathcal{T}_{2\varepsilon}(x)} \partial_{2\varepsilon} f_J(x) \right).
\end{aligned}
$$

Hence, using again Corollary 6, and taking into account that $\text{cl}^{\|\cdot\|_*}(A) \subset \text{cl}(A)$, for any $A \subset X^*$,

$$
\begin{aligned}
\partial f(x) &\subset \bigcap_{\varepsilon>0} \text{cl}^{\|\cdot\|_*} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) \\
&\subset \bigcap_{\varepsilon>0} \text{cl} \left( \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right) = \partial f(x).
\end{aligned}
$$

□

The following result shows that the subdifferential of the supremum can be reduced to the supremum of a countable family.

**Proposition 9** *Assume that X is a reflexive Banach space. Given an arbitrary family* $\{f_t, t \in T\} \subset \Gamma_0(X)$ *and* $f = \sup_{t \in T} f_t$, *for every* $x \in \text{dom } f$ *we have that*

$$
\partial f(x) = \bigcup_{J \in \mathcal{T}_c(x)} \partial f_J(x),
$$

*where*

$$
\mathcal{T}_c(x) := \{J \subset T : J \text{ countable}, \ f_J(x) = f(x)\}.
$$

**Proof** Take $x^* \in \partial f(x)$. Then, by Corollary 8, for each $n \geq 1$ there exists $J_n \in \mathcal{T}_{\frac{1}{n}}(x)$ such that

$$
x^* \in \partial_{\frac{1}{n}} f_{J_n}(x) + \frac{1}{n} B_{X^*}, \tag{45}
$$

where $B_{X^*}$ is the closed unit ball in $X^*$. Moreover, denoting $J := \cup_{n \geq 1} J_n$, for every $z^* \in \partial_{\frac{1}{n}} f_{J_n}(x)$ we have that

$$
\begin{aligned}
\langle z^*, y - x \rangle &\leq f_{J_n}(y) - f_{J_n}(x) + \frac{1}{n} \\
&\leq f_J(y) - f(x) + \frac{1}{n} + \frac{1}{n} \\
&\leq f_J(y) - f_J(x) + \frac{2}{n},
\end{aligned}
$$

showing that $z^* \in \partial_{\frac{2}{n}} f_J(x)$, that is, $\partial_{\frac{1}{n}} f_{J_n}(x) \subset \partial_{\frac{2}{n}} f_J(x)$. Hence, (45) gives rise to

$$
x^* \in \partial_{\frac{2}{n}} f_J(x) + \frac{1}{n} B_{X^*},
$$

that is, $x^* = u_n^* + v_n^*$, for $u_n^* \in \partial_{\frac{2}{n}} f_J(x)$ and $v_n^* \in \frac{1}{n} B_{X^*}$, $n \geq 1$. Hence, $v_n^* \to \theta$ and we obtain, for every $y \in X$,

$$
\begin{aligned}
\langle x^*, y - x \rangle &= \langle u_n^* + v_n^*, y - x \rangle \\
&= \lim_{n \to \infty} \langle u_n^* + v_n^*, y - x \rangle \\
&= \lim_{n \to \infty} \langle u_n^*, y - x \rangle \\
&\leq \limsup_{n \to \infty} (f_J(y) - f_J(x) + \frac{2}{n}) \\
&= f_J(y) - f_J(x),
\end{aligned}
$$

which shows that $x^* \in \partial f_J(x)$. Moreover, since that

$$
f_J(x) \geq f_{J_n}(x) \geq f(x) - \frac{1}{n}, \quad \text{for all } n \geq 1,
$$

we deduce that $f_J(x) = f(x)$, that is, $J \in \mathcal{T}_c(x)$. We are done since the opposite inclusion holds straightforwardly. $\qquad \square$

## 5 Concluding Remarks

This paper is intended to establish new characterizations of the subdifferential of the pointwise supremum of an arbitrary family of convex functions which are free of the normal cone to the effective domain of the supremum (or to finite-dimensional sections of it). These characterizations involve both (almost) active and non-(almost) active functions, the last ones being affected by a weighting parameter. Main formulas (5) and (6) highlight the role played by the almost active functions at the reference point. Formula (6) covers some other formulas in the literature; e.g., [8, Theorem 2] in the case of exact subdifferentials (see, also, [15, Theorem 1]).

The first part of the paper deals with the so-called compact scenario in which we assumed that the index set is compact and that the functions are upper semicontinuous with respect to the index. In this part, we first provide an explicit representation of the subdifferential of the supremum in Theorem 2, in terms of the active functions in one side, plus specific two-elements convex combinations in the other side.

In the second part of the paper, these compactness/upper semicontinuity assumptions are removed, and main Theorem 5 constitutes a non-compact counterpart of Theorem 2. We also aimed in the paper to emphasize the relationship of the subdifferential of the supremum function with the subdifferential of finite weighted sums. This is the purpose of (42) in Corollary 6.

Some consequences of the main results in the setting of reflexive Banach spaces are also analyzed. In particular, it turns out that formulas (41) and (42) are valid when the closure is taken with respect to the strong (norm) topology. The last proposition in the paper shows that, in this setting, the subdifferential of the supremum of the whole family can be reduced to the supremum of a countable subfamilies.

## Appendix

We start this Appendix by recalling the Hiriart-Urruty & Phelps formula (e.g., [21, Corollary 2.6.7]): for every $f, g \in \Gamma_0(X)$ and $x \in (\text{dom } f) \cap (\text{dom } g)$, we have that

$$\partial_\varepsilon (f + g)(x) = \text{cl}\left( \bigcup_{\substack{\varepsilon_1, \varepsilon_2 \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} \left( \partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2} g(x) \right) \right), \text{ for all } \varepsilon > 0. \quad (46)$$

The following lemma can be found in [5, Lemmas 2 and 3].

**Lemma 10** *Given non-empty sets $A_1, \cdots, A_k$ in $X$, and $m > 0$, we have that*

$$\left[ \overline{\text{co}}(\cup_{i=1,\cdots,k} A_i) \right]_\infty = (\overline{\text{co}}(A_1 + \cdots + A_k))_\infty \quad (47)$$

$$= (\overline{\text{co}}(A_1 + \cdots + A_{k-1} + mA_k))_\infty. \quad (48)$$

Formulas (50) and (51) in the following lemma can be found, for instance, in [21, Corollary 2.8.11]. Here, we give an alternative proof based on the minimax theorem (see (18)). However, it is not clear how this approach can be extended to infinite families of functions.

**Lemma 11** *Consider a finite family* $\{f_k, \ 1 \ \leq \ k \ \leq \ n\} \ \subset \ \Gamma_0(X)$, *and* $f \ = \ \max_{1 \leq k \leq n} f_k$. *If the function* $\varphi$ *is defined as in* (16), *i.e.,*

$$\varphi(\lambda, x) := \sum\nolimits_{1 \leq k \leq n} \lambda_k f_k(x) - I_{\mathbb{R}^n_+}(\lambda), \lambda \in \mathbb{R}^n, x \in X,$$

*then for every* $x \in \operatorname{dom} f$ *and* $\varepsilon \geq 0$ *we have that*

$$\partial_\varepsilon f(x) = \bigcup\nolimits_{\lambda \in \Delta_n} \partial_{\varepsilon + \varphi(\lambda, x) - f(x)} \varphi(\lambda, \cdot)(x) \tag{49}$$

$$= \bigcup\nolimits_{\eta \in [0, \varepsilon], \lambda \in S(x, \varepsilon - \eta)} \partial_\eta \varphi(\lambda, \cdot)(x), \tag{50}$$

*where* $S(x, \varepsilon - \eta) := \{\lambda \in \Delta_n : \varphi(\lambda, x) \geq f(x) + \eta - \varepsilon\}$. *In particular, we have that*

$$\partial f(x) = \bigcup\nolimits_{\lambda \in S(x)} \partial \varphi(\lambda, \cdot)(x), \tag{51}$$

*with* $S(x) := \{\lambda \in \Delta_n : \varphi(\lambda, x) = f(x)\}$.

**Proof** Given $\bar{x} \in X$ and $\varepsilon \geq 0$, we first assume that $\theta \in \partial_\varepsilon f(\bar{x})$. So, $\bar{x}$ is an $\varepsilon$-minimizer of the function $f$, and (18) yields some $\bar{\lambda} \in \Delta_n$ such that

$$f(\bar{x}) \leq \inf_{x \in X} f(x) + \varepsilon = \max_{\lambda \in \Delta_n} \inf_{x \in X} \varphi(\lambda, x) + \varepsilon = \inf_{x \in X} \varphi(\bar{\lambda}, x) + \varepsilon,$$

that is,

$$\varphi(\bar{\lambda}, \bar{x}) \leq \inf_{x \in X} \varphi(\bar{\lambda}, x) + \varepsilon - (f(\bar{x}) - \varphi(\bar{\lambda}, \bar{x})).$$

Therefore $\eta := \varepsilon + \varphi(\bar{\lambda}, \bar{x}) - f(\bar{x}) \in [0, \varepsilon]$ and $\bar{x}$ is an $\eta$-minimizer of the function $\varphi(\bar{\lambda}, \cdot)$, that is, $\theta \in \partial_\eta \varphi(\bar{\lambda}, \cdot)(\bar{x})$.

More generally, if $x^* \in \partial_\varepsilon f(\bar{x})$, then $\theta \in \partial_\varepsilon(f - x^*)(\bar{x})$ and the argument above yields some $\bar{\lambda} \in \Delta_n$ and $\eta \in [0, \varepsilon]$ such that $\theta \in \partial_\eta(\varphi(\bar{\lambda}, \cdot) - x^*)(\bar{x})$. Hence, $x^* \in \partial_\eta \varphi(\bar{\lambda}, \cdot)(\bar{x})$ and we conclude that

$$\partial_\varepsilon f(x) \subset \bigcup\nolimits_{\lambda \in \Delta_n} \partial_{\varepsilon + \varphi(\lambda, x) - f(x)} \varphi(\lambda, \cdot)(x) \tag{52}$$

$$\subset \bigcup\nolimits_{\lambda \in \Delta_n, \eta \in [0, \varepsilon], \varphi(\lambda, x) \geq f(x) + \eta - \varepsilon} \partial_\eta \varphi(\lambda, \cdot)(x). \tag{53}$$

Conversely, if $x^* \in \partial_\eta \varphi(\lambda, \cdot)(\bar{x})$ for some $\lambda \in \Delta_n$ and $\eta \in [0, \varepsilon]$ such that $\varphi(\lambda, \bar{x}) \geq f(\bar{x}) + \eta - \varepsilon$, then

$$\langle x^*, y - \bar{x} \rangle \leq \varphi(\lambda, y) - \varphi(\lambda, \bar{x}) + \eta \leq f(y) - f(\bar{x}) + \varepsilon,$$

and $x^* \in \partial_\varepsilon f(\bar{x})$. In other words, (49) and (50) hold true. $\qquad \square$

The following result is also used in the current work, but it has its own interest.

**Lemma 12** *Let $(f_n)_n \subset \Gamma_0(X)$ be a non-decreasing sequence, and denote $f :=$ $\sup_{n \geq 1} f_n$. Then, for all $x \in \operatorname{dom} f$ and $\varepsilon \geq 0$,*

$$\partial_\varepsilon f(x) = \bigcap_{\delta > 0} \operatorname{cl}\left(\bigcup_{k \geq 1} \bigcap_{n \geq k} \partial_{\varepsilon + \delta} f_n(x)\right).$$

*Proof* Take $x^* \in \partial_\varepsilon f(x)$ and fix $\delta > 0$. Then

$$f(x) + f^*(x^*) \leq \langle x^*, x \rangle + \varepsilon < \langle x^*, x \rangle + \varepsilon + \delta.$$

We know that

$$f^* = (\sup_{n \geq 1} f_n)^* = \overline{\operatorname{co}}(\inf_{n \geq 1} f_n^*),$$

and so, since $(f_n^*)_n$ is non-increasing, the function $\inf_{n \geq 1} f_n^*$ is convex and $f^*$ is the closed hull of $\inf_{n \geq 1} f_n^*$, i.e.,

$$f^* = \operatorname{cl}(\inf_{n \geq 1} f_n^*).$$

Then

$$f(x) + \operatorname{cl}(\inf_{n \geq 1} f_n^*)(x^*) = f(x) + f^*(x^*) < \langle x^*, x \rangle + \varepsilon + \delta,$$

and, so, there exists a net $(x_i^*)_i \subset X^*$ $w^*$-converging to $x^*$ such that

$$f(x) + \left(\inf_{n \geq 1} f_n^*\right)(x_i^*) < \langle x_i^*, x \rangle + \varepsilon + \delta, \text{ for all } i.$$

In other words, for each $i$, there exists $k \geq 1$ such that for all $n \geq k$

$$f_n(x) + f_n^*(x_i^*) \leq f(x) + f_k^*(x_i^*) < \langle x_i^*, x \rangle + \varepsilon + \delta,$$

that is,

$$x_i^* \in \bigcap_{n \geq k} \partial_{\varepsilon + \delta} f_n(x) \subset \bigcup_{k \geq 1} \bigcap_{n \geq k} \partial_{\varepsilon + \delta} f_n(x),$$

and, by taking the limit on $i$,

$$x^* \in \operatorname{cl}\left(\bigcup_{k \geq 1} \bigcap_{n \geq k} \partial_{\varepsilon + \delta} f_n(x)\right).$$

The direct inclusion follows by intersecting over $\delta > 0$. The opposite inclusion is straightforward. □

The following technical lemma is needed in the proof of Corollary 6.

**Lemma 13** *For every $\varepsilon > 0$ we have that*

$$\text{co}\left(\bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x)\right) \subset \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_{2\varepsilon} f_J(x). \tag{54}$$

**Proof** If we take $x_i^* \in \partial_\varepsilon f_{J_i}(x)$ with $J_i \in \mathcal{T}_\varepsilon(x)$, $i = 1, 2$, then, for every $\alpha \in \, ]0, 1[ \,$, it is straightforward that

$$x^* := \alpha x_1^* + (1 - \alpha) x_2^* \in \partial_\varepsilon (\alpha f_{J_1} + (1 - \alpha) f_{J_2})(x).$$

Moreover, for every $x^* \in \partial_\varepsilon (\alpha f_{J_1} + (1 - \alpha) f_{J_2})(x)$ we have

$$\begin{aligned}
\langle x^*, y - x \rangle &\leq (\alpha f_{J_1} + (1 - \alpha) f_{J_2})(y) - (\alpha f_{J_1} + (1 - \alpha) f_{J_2})(x) + \varepsilon \\
&\leq (\max\{f_{J_1}, f_{J_2}\})(y) - \alpha(f(x) - \varepsilon) - (1 - \alpha)(f(x) - \varepsilon) + \varepsilon \\
&= (\max\{f_{J_1}, f_{J_2}\})(y) - f(x) + 2\varepsilon \\
&\leq (\max\{f_{J_1}, f_{J_2}\})(y) - (\max\{f_{J_1}, f_{J_2}\})(x) + 2\varepsilon,
\end{aligned}$$

which shows that

$$\partial_\varepsilon (\alpha f_{J_1} + (1 - \alpha) f_{J_2})(x) \subset \partial_{2\varepsilon}(\max\{f_{J_1}, f_{J_2}\})(x).$$

Hence, (54) follows by taking $J := J_1 \cup J_2 \ (\in \mathcal{T}_\varepsilon(x))$. $\qquad\square$

**Proof of Proposition 3** Fix positive integers $m, n$ with $m > f(x)$, and take $\delta > 0$. Since

$$nu^* \in \mathrm{N}_{\mathrm{dom}\,f}(x) \subset \mathrm{N}_{L \cap \mathrm{dom}\,f}(x),$$

we have that, for every $y \in L$,

$$f(y) \leq m \implies \langle nu^*, y - x \rangle \leq 0 < \delta,$$

that is,

$$\langle nu^*, y - x \rangle \geq \delta, \, y \in L \implies f(y) > m,$$

and so

$$\langle nu^*, y - x \rangle \geq \delta, \, y \in L \implies \exists t \in T \text{ such that } f_t(y) > m.$$

In other words,

$$\langle nu^*, y - x \rangle \geq \delta, \, y \in L \implies y \in \bigcup_{t \in T} [f_t > m],$$

and this shows that

$$\left\{ y \in B_\delta(x) : \langle nu^*, y - x \rangle \geq \delta \right\} \subset \left\{ y \in L : \langle nu^*, y - x \rangle \geq \delta \right\}$$
$$\subset \bigcup_{t \in T} [f_t > m], \qquad (55)$$

where $B_\delta(x)$ denotes the ball in $L$ centered at $x$ with radius $\delta$ ($L$ endowed with the relative topology of $X$ is isomorphic to an Euclidean space and, consequently, $B_\delta(x)$ is compact). Therefore, since the sets $[f_t > m]$ are open, by the lower semicontinuity of the $f_t$'s, due to (55) we find a finite set $\left\{ t_1^{(n,m)}, \cdots, t_{k_{(n,m)}}^{(n,m)} \right\} \subset T$, $k_{(n,m)} \geq 1$, such that

$$\left\{ y \in B_\delta(x) : \langle nu^*, y - x \rangle \geq \delta \right\} \subset \bigcup_{i=1,\cdots,k_{(n,m)}} \left[ f_{t_i^{(n,m)}} > m \right].$$

Equivalently, if we define

$$g_{(n,m)} := \max_{i=1,\cdots,k_{(n,m)}} f_{t_i^{(n,m)}},$$

then we have

$$\left[ g_{(n,m)} \leq m \right] = \bigcap_{i=1,\cdots,k_{(n,m)}} \left[ f_{t_i^{(n,m)}} \leq m \right]$$
$$\subset (X \setminus B_\delta(x)) \cup \left\{ y \in X : \langle nu^*, y - x \rangle < \delta \right\}.$$

Also, by denoting

$$g := \sup_{n,m \geq 1} g_{(n,m)},$$

we have that

$$[g \leq m] \subset \bigcap_{n \geq 1} \left[ g_{(n,m)} \leq m \right],$$

and we obtain

$$[g \leq m] \cap B_\delta(x) \subset \left( \bigcap_{n \geq 1} \left[ g_{(n,m)} \leq m \right] \right) \bigcap B_\delta(x)$$
$$\subset \left\{ y \in X : \langle nu^*, y - x \rangle < \delta \right\}. \qquad (56)$$

Hence, since $x \in [g \leq m] \cap B_\delta(x)$ (remember that $m > f(x)$),

$$nu^* \in \mathrm{N}_{[g \leq m] \cap B_\delta(x)}^\delta(x), \text{ for all } n \geq 1,$$

and by taking $n \uparrow +\infty$ we obtain that

$$u^* \in \mathrm{N}_{[g \leq m] \cap B_\delta(x)}(x) = \mathrm{N}_{[g \leq m] \cap L}(x), \text{ for all } m > f(x).$$

Therefore

$$u^* \in \bigcap_{m > f(x)} \mathrm{N}_{[g \leq m] \cap L}(x)$$
$$\subset \mathrm{N}_{\bigcup_{m > f(x)} [g \leq m] \cap L}(x) = \mathrm{N}_{(\mathrm{dom}\, g) \cap L}(x),$$

and we conclude the proof since

$$g = \sup_{n,m \geq 1} g_{(n,m)} = \sup_{n,m \geq 1, i = 1, \cdots, k_{(n,m)}} f_{t_i^{(n,m)}}$$

is the supremum of a countable family.

In the following lemma, the result corresponding to $\varepsilon = 0$ can be found in [6, (8) in page 835].

**Lemma 14** *Given a non-empty set $A \subset X$, for every $x \in \mathrm{dom}\, \sigma_A$ and $\varepsilon \geq 0$ we have that*

$$\mathrm{N}^\varepsilon_{\mathrm{dom}\, \sigma_A}(x) = \mathrm{N}^\varepsilon_{((\overline{\mathrm{co}}A)_\infty)^-}(x) = (\overline{\mathrm{co}}A)_\infty \cap \{x^* \in X^* : -\varepsilon \leq \langle x^*, x \rangle \leq 0\}. \quad (57)$$

**Proof** Fix $x \in \mathrm{dom}\, \sigma_A$ and $\varepsilon \geq 0$. The first equality comes from (15). Take $x^* \in \mathrm{N}^\varepsilon_{((\overline{\mathrm{co}}A)_\infty)^-}(x)$, so that

$$\langle x^*, \alpha y - x \rangle \leq \varepsilon \text{ for all } y \in ((\overline{\mathrm{co}}A)_\infty)^- \text{ and } \alpha > 0. \quad (58)$$

Then, by dividing on $\alpha$ and next making $\alpha \uparrow +\infty$ we obtain that $x^* \in ((\overline{\mathrm{co}}A)_\infty)^{--} = (\overline{\mathrm{co}}A)_\infty$, using again (9). Moreover, by taking $y = x$ in (58), as $x \in \mathrm{dom}\, \sigma_A \subset ((\overline{\mathrm{co}}A)_\infty)^-$ (by (15)), we obtain that

$$\langle x^*, (\alpha - 1)x \rangle \leq \varepsilon, \text{ for all } \alpha \geq 0.$$

If $\alpha = 0$, then $-\varepsilon \leq \langle x^*, x \rangle$, and if $\alpha$ goes to $+\infty$, then $\langle x^*, x \rangle \leq 0$. By summarizing, we have proved that $x^* \in (\overline{\mathrm{co}}A)_\infty \cap \{x^* \in X^* : -\varepsilon \leq \langle x^*, x \rangle \leq 0\}$.

Conversely, if $x^* \in (\overline{\mathrm{co}}A)_\infty \cap \{x^* : -\varepsilon \leq \langle x^*, x \rangle \leq 0\}$, then

$$\langle x^*, y - x \rangle \leq \langle x^*, -x \rangle \leq \varepsilon \text{ for all } y \in ((\overline{\mathrm{co}}A)_\infty)^-,$$

and so $x^* \in \mathrm{N}^\varepsilon_{((\overline{\mathrm{co}}A)_\infty)^-}(x)$.

# References

1. Brøndsted, A.: On the subdifferential of the supremum of two convex functions. Math. Scand. **31**, 225–230 (1972)
2. Cabot, A., Thibault, L.: Sequential formulae for the normal cone to sublevel sets. Trans. Am. Math. Soc. **366**, 6591–6628 (2014)
3. Correa, R., Hantoute, A., López, M.A.: Valadier-like formulas for the supremum function I. J. Convex Anal. **25**, 1253–1278 (2018)
4. 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)
5. Correa, R., Hantoute, A., López, M.A.: Alternative representations of the normal cone to the domain of supremum functions and subdifferential calculus. Set-Valued Var. Anal (2021). https://doi.org/10.1007/s11228-021-00583-3
6. 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)
7. 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)
8. Hantoute, A., Svensson, A.: A General representation of $\delta$-normal sets to sublevels of convex functions. Set-Valued Var. Anal. **25**, 651–678 (2017)
9. Ioffe, A.D.: A note on subdifferentials of pointwise suprema. TOP **20**, 456–466 (2012)
10. Ioffe, A.D., Levin, V.L.: Subdifferentials of convex functions. Trudy Moskov Mat. Obshch. **26**, 3–73 (1972). **(Russian)**
11. Ioffe, A.D., Tikhomirov, V.H.: *Theory of Extremal Problems*. Studies in Mathematics and its Applications, Vol. 6, North-Holland, Amsterdam (1979)
12. Levin, V.L., An application of Hellys theorem in convex programming, problems of best approximation and related questions. Mat. Sb., Nov. Ser. 79, : 250–263. English transl.: Math. USSR, Sb. **8**, 235–247 (1969)
13. Li, C., Ng, K.F.: Subdifferential calculus rules for supremum functions in convex analysis. SIAM J. Optim. **21**, 782–797 (2011)
14. López, M.A., Volle, M.: A formula for the set of optimal solutions of relaxed minimization problems. Applications to subdifferential calculus. J. Convex Anal. **17**, 1057–1075 (2010)
15. López, M.A., Volle, M.: On the subdifferential of the supremum of an arbitrary family of extended real-valued functions. RACSAM **105**, 3–21 (2011)
16. Pschenichnyi, B.N.: Convex programming in a normalized space. Kibernetika **5**, 46–54 (1965) (Russian); translated as Cybernetics **1**, 46–57 (1965)
17. Rockafellar, R.T.: Directionally Lipschitzian functions and subdifferential calculus. Proc. Lond. Math. Soc. **39**, 331–355 (1979)
18. Tikhomirov, V.M.: *Analysis II: Convex Analysis and Approximation Theory*. RX Gamkrelidze (Ed.), Encyclopedia of Mathematics Vol **14** (1990)
19. Valadier, M.: Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes. C. R. Acad. Sci. Paris Sér. A-B **268**, A39–A42 (1969)
20. Volle, M.: Sous-différentiel d'une enveloppe supérieure de fonctions convexes. C. R. Acad. Sci. Paris Sér. I Math. **317**, 845–849 (1993)
21. Zălinescu, C.: Convex Analysis in General Vector Spaces. World Scientific Publishing Company, River Edge (2002)