SIAM J. OPTIM.
Vol. 29, No. 2, pp. 1106–1130

MOREAU–ROCKAFELLAR-TYPE FORMULAS FOR THE SUBDIFFERENTIAL OF THE SUPREMUM FUNCTION*

RAFAEL CORREA†, ABDERRAHIM HANTOUTE‡, AND MARCO A. LÓPEZ-CERDÁ§

Abstract. We characterize the subdifferential of the supremum function of finitely and infinitely indexed families of convex functions. The main contribution of this paper consists of providing formulas for such a subdifferential under weak continuity assumptions. The resulting formulas are given in terms of the exact subdifferential of the data functions at the reference point, and not at nearby points as in [Valadier, *C. R. Math. Acad. Sci. Paris*, 268 (1969), pp. 39–42]. We also derive new Fritz John- and KKT-type optimality conditions for semi-infinite convex optimization, omitting the continuity/closedness assumptions in [Dinh et al., *ESAIM Control Optim. Calc. Var.*, 13 (2007), pp. 580–597]. When the family of functions is finite, we use continuity conditions concerning only the active functions, and not all the data functions as in [Rockafellar, *Proc. Lond. Math. Soc.* (3), 39 (1979), pp. 331–355; Volle, *Acta Math. Vietnam.*, 19 (1994), pp. 137–148].

Key words. supremum function, convex functions, subdifferential calculus rules, qualification conditions

AMS subject classifications. 26B05, 26J25, 49H05

DOI. 10.1137/18M1169370

1. Introduction.

Our aim is to characterize the subdifferential $\partial f(x)$ of the supremum function

$$f := \sup_{t\in T} f_t, \tag{1}$$

where $f_t : X \to \mathbb{R}\cup\{+\infty\}$, $t \in T$, are proper convex functions defined in a locally convex topological vector space X. We establish formulas involving only the exact subdifferentials $\partial f_t(x)$, for active indices at x, up to the normal cone to the effective domain of f, $\mathrm{N}_{\mathrm{dom}\, f}(x)$. Two cases are studied: either T is finite, or the set of ε-active indices $T_\varepsilon(x)$ is compact in a Hausdorff topological space T and the data functions $f_t(z)$, $z \in \mathrm{dom}\, f$, are upper semicontinuous as functions of t on the set $T_\varepsilon(x)$.

Both the finite and the infinite-dimensional settings are considered. In the finite-dimensional framework, we prove in Theorem 3 the following formula for the so-called compact case:

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

*Received by the editors February 6, 2018; accepted for publication (in revised form) January 3, 2019; published electronically April 16, 2019.
http://www.siam.org/journals/siopt/29-2/M116937.html
Funding: Research of the first and second authors is supported by CONICYT grants, Fondecyt 1150909 and 1151003, and Proyecto grant PIA AFB-170001. Research of the second and third authors is supported by MINECO of Spain and FEDER of EU, grant MTM2014-59179-C2-1-P. Research of the third author is also supported by the Australian Research Council, Project DP160100854.
†Universidad de O'Higgins, Rancagua, Chile and DIM-CMM of Universidad de Chile, Beauchef 851, Santiago, Chile (rcorrea@dim.uchile.cl).
‡Center for Mathematical Modeling (CMM), Universidad de Chile, Beauchef 851, Santiago, Chile (ahantoute@dim.uchile.cl).
§Department of Mathematics, Alicante University, Campus de San Vicent, 03080, Alicante, Spain and CIAO, Federation University, Mt Helen Campus, Ballarat Central, VIC 3350, Australia (Marco.Antonio@ua.es).

where $\mathrm{I}_{\operatorname{dom} f}$ is the indicator function of $\operatorname{dom} f$ and $T(x) := \{t \in T \mid f_t(x) = f(x)\}$ is the set of active indices at x. In addition, when the relative interior of the effective domains of the active data functions overlap, that is,

$$\operatorname{ri}(\operatorname{dom} f_t) \cap \operatorname{dom} f \neq \emptyset \quad \text{for all } t \in T(x),$$

we prove that

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

In the infinite-dimensional framework, when T is finite, we show in Theorem 9 that if all the active functions, except perhaps one of them, namely f_{k_0}, are continuous at a common point in $\operatorname{dom} f$, then

$$\partial f(x) = \operatorname{co}\left\{\bigcup_{k \in T(x) \setminus \{k_0\}} \partial f_k(x) \bigcup \partial (f_{k_0} + \mathrm{I}_{\operatorname{dom} f})(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x). \tag{2}$$

The last formula extends well-known results in [25] and [30]. More precisely, the following result of [30] requires that all the functions $f_k,\ k \in T$, except perhaps one of them (not only the active ones as in our formula (2)) are continuous at some point in $\operatorname{dom} f$:

$$\partial f(x) = \operatorname{co}\left\{\bigcup_{k \in T(x)} \partial f_k(x)\right\} + \sum_{k \in T} \mathrm{N}_{\operatorname{dom} f_k}(x).$$

This formula reduces to the Rockafellar characterization [25, Theorem 4]

$$\partial f(x) = \operatorname{co}\left\{\bigcup_{k \in T(x)} \partial f_k(x)\right\},$$

valid when $T(x) = T$ and all the subdifferentials $\partial f_k(x)$, $k \in T$, are nonempty. Observe that the equality $\mathrm{N}_{\operatorname{dom} f}(x) = \sum_{k \in T} \mathrm{N}_{\operatorname{dom} f_k}(x)$ is a consequence of the current continuity condition, due to the the sum subdifferential rule [23].

It is worth observing that formula (2) is also related to the following characterization given in [1], which uses approximate subdifferentials instead of the exact ones, and requires the lower semicontinuity of all the functions f_k, $k \in T$, as well as the condition $T(x) = T$:

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\operatorname{co}}\left\{\bigcup_{k \in T(x)} \partial_\varepsilon f_k(x)\right\}. \tag{3}$$

If, instead of being lower semicontinuous (lsc), the data functions f_k, $k \in T$, are required to satisfy the weaker closure condition

$$\operatorname{cl} f = \sup_{k \in T} \operatorname{cl} f_k, \tag{4}$$

then formula (3) also holds (see [12, Corollary 12]). This condition was introduced in [12] as a common lower semicontinuity-like condition guaranteeing the fulfilment of several subdifferential calculus rules in the recent literature. Moreover, a variant of (4) was shown in [16, Theorem 3.1] to be necessary for the validity of formula (2) in the Banach setting. In contrast to some results in [6], a feature of the present paper is that we succeed in removing this condition, increasing in this way the validity of Theorems 1 and 4 in [6].

We apply these results to derive new Fritz John- and KKT-type optimality conditions for semi-infinite convex optimization, omitting the standard continuity assumptions. More precisely, we deal with the problem

$$(\mathcal{P}) \qquad \inf_{\substack{f_t(x)\leq 0,\ t\in T\\ x\in C}} f_0(x),$$

where $C \subset \mathbb{R}^n$ is convex, T is a Hausdorff topological space, and $f_t : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, for $t \in T \cup \{0\}$ ($0 \notin T$), are proper and convex. Then, we prove in Corollary 6 that if a point $\bar{x} \in C \subset \mathbb{R}^n$ is optimal for problem $(\mathcal{P})$, then there exist a (possibly empty) finite set $\widehat{T}(\bar{x}) \subset A(\bar{x}) := \{t \in T \mid f_t(\bar{x}) = 0\}$ such that $\partial f_t(\bar{x}) \neq \emptyset$ for all $t \in \widehat{T}(\bar{x})$, and scalars $\lambda_t > 0$ for all $t \in \widehat{T}(\bar{x})$ satisfying

$$0_n \in \partial f_0(\bar{x}) + \sum_{t\in\widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T} \mathrm{N}_{\operatorname{dom} f_t}(\bar{x}), \tag{5}$$

provided that the Slater constraint qualification and some natural assumptions, including the interiority condition (37), hold. Here 0_n is the zero vector in $\mathbb{R}^n$ and $\sum_{\emptyset} = \{0_n\}$. It also turns out that (5) is equivalent to

$$0_n \in \partial f_0(\bar{x}) + \sum_{t\in\widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T\setminus\widehat{T}(\bar{x})} \mathrm{N}_{\operatorname{dom} f_t}(\bar{x}),$$

since for $t \in \widehat{T}(\bar{x})$ one has $\partial f_t(\bar{x}) + \mathrm{N}_{\operatorname{dom} f_t}(\bar{x}) = \partial f_t(\bar{x})$.

It is worth mentioning that alternative KKT conditions exist in the literature, obtained via many different approaches: approximate subdifferentials of the data functions [3, 13], exact subdifferentials at close points [28], asymptotic KKT conditions [19] for linear semi-infinite programming, Farkas–Minkowski-type closedness criteria [8] in convex semi-infinite optimization, and strong CHIP-like qualifications (where CHIP stands for conical hull intersection property) for convex optimization with not necessarily convex C^1-constraints [2] (see also [9] for locally Lipschitz constraints), among others. We also refer the reader to [32] and references therein for KKT conditions in the framework of subsmooth semi-infinite optimization, and to [10] for analysis of the relationships among KKT rules and Lagrangian dualities.

We refer the reader to [12, Theorem 4] for a complete characterization of the subdifferential of the supremum, involving the approximate subdifferential of the ε-active functions, which does not require any continuity assumption. The compactly indexed case is treated in [5] using the same finite-dimensional reduction approach as in [12]. In the framework of the last section, we succeeded in avoiding this reduction tool when obtaining the desired extension of Rockafellar's result (Theorem 9).

For variants of [12, Theorem 4], see [16, Theorem 3.1] and [14]. In Banach spaces, [20] gives a formula using the exact subdifferentials of the data functions but at points close to the reference point. The locally convex version of this result is investigated in [5]. We also cite here [27], which deals with the directional derivative of the supremum function under certain conditions on the index set. The paper [22] approaches the subdifferential of the supremum of (nonconvex) uniformly Lipschitz continuous functions. Applications of [12, Theorem 4] gave rise in [3, 4] to new calculus rules for the subdifferential of the sum.

The paper is organized as follows. After section 2, which provides notation, we establish in section 3 some general results on the subdifferential of the supremum function. Section 4 focuses on the finite-dimensional case, where Theorem 3 is the main result. In the same section, we derive Fritz John- and KKT-type conditions for semi-infinite convex optimization in Theorem 5, and Corollaries 6 and 7, respectively. In section 5 we deal with the case of finitely many convex functions in locally convex spaces. Theorem 9 is the most relevant result in this final section.

2. Notation. In this paper X stands for a (real) separated locally convex space (lcs), whose topological dual, denoted by X^*, is endowed with the weak*-topology. Hence, X and X^* form a dual pair by means of the canonical bilinear form $\langle x, x^*\rangle = \langle x^*, x\rangle := x^*(x)$, $(x, x^*) \in X \times X^*$. The zero vectors are denoted by θ, and the convex, closed, and balanced neighborhoods of θ are called θ-neighborhoods. The families of such θ-neighborhoods in X and in X^* are denoted by $\mathcal{N}_X$ and $\mathcal{N}_{X^*}$, respectively. Recall that 0_n is the zero vector in $\mathbb{R}^n$.

Given a nonempty set A in X (or in X^*), by $\operatorname{co} A$, $\operatorname{aff} A$, and $\operatorname{span} A$, we denote the *convex hull*, the *affine hull*, and the *linear hull* of A, respectively. Moreover, $\operatorname{cl} A$ and $\overline{A}$ are indistinctly used for denoting the *closure* of A (the *weak*-closure* if $A \subset X^*$). Thus, $\overline{\operatorname{co}} A := \operatorname{cl}(\operatorname{co} A)$, $\overline{\operatorname{aff}} A := \operatorname{cl}(\operatorname{aff} A)$, etc. We use $\operatorname{ri} A$ to denote the (topological) *relative interior* of A (i.e., the interior of A in the topology relative to $\operatorname{aff} A$ when this set is closed, and the empty set otherwise). The *polar* of A is the set

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

The following standard conventions are adopted within the paper:

$$\emptyset + A = \emptyset \text{ and } \operatorname{co}\emptyset = \emptyset. \tag{6}$$

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

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

$$\sigma_A(x^*) := \sup\{\langle x^*, a\rangle \mid a \in A\}, \ x^* \in X^*, \tag{7}$$

with the convention $\sigma_\emptyset \equiv -\infty$. We say that a convex function $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ is proper if its *(effective) domain,* $\operatorname{dom}\varphi := \{x \in X \mid \varphi(x) < +\infty\}$, is nonempty. The *epigraph* of φ is the set $\operatorname{epi}\varphi := \{(x, \lambda) \in X \times \mathbb{R} \mid \varphi(x) \leq \lambda\}$. The *lsc hull* of φ is the function $\operatorname{cl}\varphi$ such that $\operatorname{epi}(\operatorname{cl}\varphi) = \operatorname{cl}(\operatorname{epi}\varphi)$.

The *subdifferential* of φ at a point x where $\varphi(x)$ is finite is the weak*-closed convex set

$$\partial\varphi(x) := \{x^* \in X^* \mid \varphi(y) - \varphi(x) \geq \langle x^*, y - x\rangle \text{ for all } y \in X\}.$$

If $\varphi(x) \notin \mathbb{R}$, then we set $\partial\varphi(x) := \emptyset$. If $\varphi(x) = (\operatorname{cl}\varphi)(x)$, then

$$\partial\varphi(x) = \partial(\operatorname{cl}\varphi)(x). \tag{8}$$

In particular, this holds when $\partial\varphi(x) \neq \emptyset$. One can easily verify that, for every $x \in \operatorname{dom}\varphi$,

$$\partial\varphi(x) = \bigcap_{L \in \mathcal{F}(x)} \partial(\varphi + \mathrm{I}_{L \cap \operatorname{dom}\varphi})(x), \tag{9}$$

where

$$\mathcal{F}(x) := \{\text{finite-dimensional linear subspaces } L \subset X \text{ containing } x\}.$$

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

$$\mathrm{N}_A(x) := \{x^* \in X^* \mid \langle x^*, y - x\rangle \le 0 \text{ for all } y \in A\} \text{ if } x \in A,$$

and $\mathrm{N}_A(x) := \emptyset$ if $x \in X \setminus A$.

A family of convex sets $\{A_i,\ i \in I\}$ such that $\cap_{i\in I} A_i \neq \emptyset$ has the *strong conical hull intersection property (the strong CHIP) at* $x \in \cap_{i\in I} A_i$ if

$$\begin{aligned}\mathrm{N}_{\cap_{i\in I} A_i}(x) &= \sum_{i\in I} \mathrm{N}_{A_i}(x) \\ &:= \left\{\sum_{i\in J} a_i,\ a_i \in \mathrm{N}_{A_i}(x),\ J \text{ being a finite subset of } I\right\}.\end{aligned}$$

This notion was introduced in [7] and extended to infinite families of convex sets in [17] and [18].

3. First results on the subdifferential of the supremum function. We devote this section to providing some general results on the subdifferential of the supremum function, which is used later on in the present work. We consider a family of proper convex functions $f_t : X \to \mathbb{R} \cup \{+\infty\}$, $t \in T$, defined in a locally convex topological vector space X, together with the supremum function

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

The set of ε-*active indices* at $x \in X$ is

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

when $f(x) \in \mathbb{R}$, and $T_\varepsilon(x) := \emptyset$ otherwise. We write $T(x)$ instead of $T_0(x)$. In section 4 we apply the following result, which extends the validity of Theorem 1 in [6] since the closedness condition (4) is omitted. Observe that if X is the Euclidean space $\mathbb{R}^n$ and f is proper, then $\mathrm{ri}(\mathrm{dom}\, f) \neq \emptyset$ and $f_{|\mathrm{aff}(\mathrm{dom}\, f)}$ is continuous on this set (see [26, Theorem 10.1]).

PROPOSITION 1. *Suppose that the function* $f_{|\mathrm{aff}(\mathrm{dom}\, f)}$ *is continuous on* $\mathrm{ri}(\mathrm{dom}\, f)$, *which is assumed to be nonempty. Let* $x \in \mathrm{dom}\, f$ *be such that for some* $\varepsilon_0 > 0$,

(i) *the set* $T_{\varepsilon_0}(x)$ *is compact in the Hausdorff topological space* T;

(ii) *for each* $z \in \mathrm{dom}\, f$, *the function* $t \mapsto f_t(z)$ *is upper semicontinuous (usc, for short) on* $T_{\varepsilon_0}(x)$.

Then

$$\partial f(x) = \overline{\mathrm{co}}\left\{\bigcup_{t\in T(x)} \partial(f_t + \mathrm{I}_{\mathrm{dom}\, f})(x)\right\}. \tag{10}$$

Proof. We consider the proper convex functions

$$g_t := f_t + \mathrm{I}_{\mathrm{dom}\, f},\ t \in T, \text{ and } g := \sup_{t\in T} g_t,$$

so that $g = \sup_{t\in T}(f_t + \mathrm{I}_{\mathrm{dom}\, f}) = f + \mathrm{I}_{\mathrm{dom}\, f} = f$, and

$$\mathrm{dom}\, g_t = \mathrm{dom}\, f_t \cap \mathrm{dom}\, f = \mathrm{dom}\, f \text{ for all } t \in T. \tag{11}$$

Hence, for each $t \in T$, since $g_t \le g = f$ and $\mathrm{dom}\, g_t = \mathrm{dom}\, f$, so that $\mathrm{aff}(\mathrm{dom}\, f) = \mathrm{aff}(\mathrm{dom}\, g_t)$, the current continuity assumptions on f imply that $g_{t|\mathrm{aff}(\mathrm{dom}\, f)}$ is locally uniformly upper bounded at each point in $\mathrm{ri}(\mathrm{dom}\, f)$. So, $g_{t|\mathrm{aff}(\mathrm{dom}\, f)}$ is continuous on $\mathrm{ri}(\mathrm{dom}\, f)$ [24], and we obtain that

$$\mathrm{cl}\, g_t(y) = g_t(y) \quad \text{for all } y \in \mathrm{ri}(\mathrm{dom}\, f). \tag{12}$$

Observe that the g_t's satisfy conditions (i) and (ii), since

$$\{t \in T \mid g_t(x) \ge g(x) - \varepsilon_0\} = T_{\varepsilon_0}(x), \tag{13}$$

and the functions

$$t \mapsto g_t(z) = f_t(z),\ z \in \mathrm{dom}\, f, \tag{14}$$

are usc on $T_{\varepsilon_0}(x)$.

Now, let us proceed by showing that

$$\mathrm{cl}\, g = \sup_{t\in T}(\mathrm{cl}\, g_t).$$

On the one hand, since

$$\sup_{t\in T}(\mathrm{cl}\, g_t)(y) \le \sup_{t\in T} g_t(y) = g(y) \quad \text{for all } y \in X,$$

we deduce that

$$\sup_{t\in T}(\mathrm{cl}\, g_t) \le \mathrm{cl}\, g \tag{15}$$

as a consequence of the lower semicontinuity of the function on the left-hand side. On the other hand, in order to prove the converse inequality, we fix $y \in X$ such that $\sup_{t\in T}(\mathrm{cl}\, g_t)(y) < +\infty$. Now, due to the inequality $\mathrm{cl}(f_t)+\mathrm{I}_{\mathrm{cl}(\mathrm{dom}\, f)} \le f_t+\mathrm{I}_{\mathrm{dom}\, f}$ and the lower semicontinuity of the function $\mathrm{cl}(f_t)+\mathrm{I}_{\mathrm{cl}(\mathrm{dom}\, f)}$, we have $\mathrm{cl}(f_t)+\mathrm{I}_{\mathrm{cl}(\mathrm{dom}\, f)} \le \mathrm{cl}(f_t + \mathrm{I}_{\mathrm{dom}\, f})$, yielding

$$\begin{aligned} \sup_{t\in T}(\mathrm{cl}(f_t) + \mathrm{I}_{\mathrm{cl}(\mathrm{dom}\, f)})(y) &\le \sup_{t\in T}(\mathrm{cl}(f_t + \mathrm{I}_{\mathrm{dom}\, f}))(y) \\ &= \sup_{t\in T}(\mathrm{cl}\, g_t)(y) < +\infty, \end{aligned}$$

and this implies that $y \in \mathrm{cl}(\mathrm{dom}\, f)$.

Let us pick a point $x_0 \in \mathrm{ri}(\mathrm{dom}\, f) = \mathrm{ri}(\mathrm{dom}\, g_t)$ (by (11)) and consider $x_\lambda := \lambda y + (1-\lambda)x_0$ for $\lambda \in]0,1[$. By the accessibility lemma (see, e.g., [26]), for each $t \in T$ we have that $x_\lambda \in \mathrm{ri}(\mathrm{dom}\, f) = \mathrm{ri}(\mathrm{dom}\, g_t)$, and so (12) leads us to

$$\mathrm{cl}\, g_t(x_\lambda) = g_t(x_\lambda) \text{ for all } \lambda \in]0,1[\,.$$

Consequently,

$$\begin{aligned} g(x_\lambda) &= \sup_{t\in T} g_t(x_\lambda) \\ &= \sup_{t\in T}(\mathrm{cl}\, g_t)(x_\lambda) \\ &\le \lambda \sup_{t\in T}(\mathrm{cl}\, g_t)(y) + (1-\lambda)\sup_{t\in T}(\mathrm{cl}\, g_t)(x_0) \\ &\le \lambda \sup_{t\in T}(\mathrm{cl}\, g_t)(y) + (1-\lambda) f(x_0), \end{aligned}$$

and, taking the lower limit as $\lambda \uparrow 1$, we get

$$(\mathrm{cl}\, g)(y) \le \liminf_{\lambda \uparrow 1} g(x_\lambda) \le \sup_{t \in T} (\mathrm{cl}\, g_t)(y),$$

yielding the converse inequality of (15).

Finally, the g_t's satisfy the assumption of [6, Theorem 1], which gives us

$$\begin{aligned} \partial f(x) = \partial g(x) &= \overline{\mathrm{co}} \left\{ \bigcup_{\{t \in T \mid g_t(x) = g(x)\}} \partial (g_t + \mathrm{I}_{\mathrm{dom}\, g})(x) \right\} \\ &= \overline{\mathrm{co}} \left\{ \bigcup_{t \in T(x)} \partial (f_t + \mathrm{I}_{\mathrm{dom}\, f} + \mathrm{I}_{\mathrm{dom}\, g})(x) \right\} \\ &= \overline{\mathrm{co}} \left\{ \bigcup_{t \in T(x)} \partial (f_t + \mathrm{I}_{\mathrm{dom}\, f})(x) \right\}. \end{aligned}$$

∎

The following proposition improves Theorem 4 in [6], and it also gets rid of condition (4).

PROPOSITION 2. *Let $x \in \mathrm{dom}\, f$ be such that, for some $\varepsilon_0 > 0$,*
(i) *the set $T_{\varepsilon_0}(x)$ is compact,*
(ii) *for each $z \in \mathrm{dom}\, f$ the function $t \mapsto f_t(z)$ is usc on $T_{\varepsilon_0}(x)$.*
Then

$$\partial f(x) = \bigcap_{L \in \mathcal{F}(x)} \overline{\mathrm{co}} \left\{ \bigcup_{t \in T(x)} \partial (f_t + \mathrm{I}_{L \cap \mathrm{dom}\, f})(x) \right\}. \tag{16}$$

Proof. To prove (16) we recall that (see (9))

$$\partial f(x) = \bigcap_{L \in \mathcal{F}(x)} \partial (f + \mathrm{I}_{L \cap \mathrm{dom}\, f})(x). \tag{17}$$

Fix $L \in \mathcal{F}(x)$ and proceed by checking that the functions

$$h_t := f_t + \mathrm{I}_{L \cap \mathrm{dom}\, f},\ t \in T, \text{ and } h := \sup_{t \in T} h_t$$

satisfy the assumptions of Proposition 1. Firstly, since

$$h = \sup_{t \in T} (f_t + \mathrm{I}_{L \cap \mathrm{dom}\, f}) = f + \mathrm{I}_{L \cap \mathrm{dom}\, f} = f + \mathrm{I}_L,$$

we have that $\mathrm{dom}\, h = \mathrm{dom}\, f \cap L$ $(\subset L)$, and so $\mathrm{ri}(\mathrm{dom}\, h) \neq \emptyset$ and $h_{|\mathrm{aff}(\mathrm{dom}\, h)}$ is continuous on $\mathrm{ri}(\mathrm{dom}\, h)$ (since L is finite-dimensional and $\mathrm{dom}\, f \cap L$ is nonempty as it contains x). Secondly, due to the definition of the functions h_t and h, we have that $h_t(x) = f_t(x) + \mathrm{I}_{L \cap \mathrm{dom}\, f}(x) = f_t(x)$ and $h(x) = (f + \mathrm{I}_L)(x) = f(x)$, and so

$$\{t \in T \mid h_t(x) \ge h(x) - \varepsilon_0\} = \{t \in T \mid f_t(x) \ge f(x) - \varepsilon_0\} = T_{\varepsilon_0}(x)$$

and

$$h_t(z) = f_t(z) \text{ for all } z \in \mathrm{dom}\, h = \mathrm{dom}\, f \cap L \subset \mathrm{dom}\, f,$$

and, therefore, the functions h_t, $t \in T$, and h also satisfy conditions (i) and (ii) in Proposition 1. Consequently, by applying this proposition we obtain that

$$\begin{aligned}\partial(f + \mathrm{I}_{L\cap \operatorname{dom} f})(x) &= \partial h(x)\\ &= \overline{\mathrm{co}}\left\{\bigcup_{\{t\in T \mid h_t(x)=h(x)\}} \partial(h_t + \mathrm{I}_{\operatorname{dom} h})(x)\right\}\\ &= \overline{\mathrm{co}}\left\{\bigcup_{T(x)} \partial(f_t + \mathrm{I}_{L\cap \operatorname{dom} f} + \mathrm{I}_{\operatorname{dom} h})(x)\right\}\\ &= \overline{\mathrm{co}}\left\{\bigcup_{T(x)} \partial(f_t + \mathrm{I}_{L\cap \operatorname{dom} f} + \mathrm{I}_{\operatorname{dom} f})(x)\right\}\\ &= \overline{\mathrm{co}}\left\{\bigcup_{T(x)} \partial(f_t + \mathrm{I}_{L\cap \operatorname{dom} f})(x)\right\}.\end{aligned}$$

Then the conclusion follows by (17), intersecting over $L \in \mathcal{F}(x)$. □

4. Qualification conditions in finite dimensions. The first theorem in this section yields a simple characterization in the finite-dimensional setting of the subdifferential of the supremum function $f = \sup_{t\in T} f_t$, where the $f_t : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $t \in T$, are proper and convex.

We use the following qualification condition:

$$\operatorname{ri}(\operatorname{dom} f_t) \cap \operatorname{dom} f \neq \emptyset \text{ for all } t \in T(x). \tag{18}$$

Due to the accessibility lemma, we can show that (18) is equivalent to

$$\operatorname{ri}(\operatorname{dom} f_t) \cap \operatorname{ri}(\operatorname{dom} f) \neq \emptyset \text{ for all } t \in T(x). \tag{19}$$

THEOREM 3. *Let $x \in \mathbb{R}^n$ be such that, for some $\varepsilon_0 > 0$,*
(i) *the set $T_{\varepsilon_0}(x)$ is compact,*
(ii) *for each $z \in \operatorname{dom} f$ the function $t \mapsto f_t(z)$ is usc on $T_{\varepsilon_0}(x)$.*
Then

$$\partial f(x) = \mathrm{co}\left\{\bigcup_{t\in T(x)} \partial(f_t + \mathrm{I}_{\operatorname{dom} f})(x)\right\} \tag{20}$$

and, under condition (18),

$$\partial f(x) = \mathrm{co}\left\{\bigcup_{t\in T(x)} \partial f_t(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x). \tag{21}$$

Proof. To start, observe that the following inclusions always hold:

$$\mathrm{co}\left\{\bigcup_{t\in T(x)} \partial f_t(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x) \subset \mathrm{co}\left\{\bigcup_{t\in T(x)} \partial(f_t + \mathrm{I}_{\operatorname{dom} f})(x)\right\} \subset \partial f(x). \tag{22}$$

If $\operatorname{dom} f = \emptyset$, then $\partial f(x) = \emptyset$ and (22) leads to (20) and (21). Even when $\operatorname{dom} f \neq \emptyset$ but $\partial f(x) = \emptyset$, these formulas also hold. Consequently, in the rest of the proof we shall suppose that $\partial f(x) \neq \emptyset$, which leads to $f(x) = (\operatorname{cl} f)(x) \in \mathbb{R}$ (recall (8)). In particular, the function f is proper, so $\operatorname{dom} f \neq \emptyset$ and, therefore, $\operatorname{ri}(\operatorname{dom} f) \neq \emptyset$.

Then, according to Proposition 1, conditions (i) and (ii) imply that

$$\partial f(x) = \operatorname{cl} E, \text{ where } E := \operatorname{co}\left\{ \bigcup_{t \in T(x)} \partial(f_t + \mathrm{I}_{\operatorname{dom} f})(x) \right\}. \tag{23}$$

We are going to prove that the set E is closed. To this end, we take a sequence $(z_i)_{i \geq 1} \subset E$ that converges to some $z \in \mathbb{R}^n$; hence, as $E \subset \partial f(x)$, we have $z \in \partial f(x)$. So, taking Charathéodory's theorem into account, for each $i \geq 1$ there are scalars $\lambda_{i,1}, \ldots, \lambda_{i,n+1} \geq 0$ and elements

$$z_{i,1} \in \partial(f_{t_{i,1}} + \mathrm{I}_{\operatorname{dom} f})(x), \ \ \ldots, \ \ z_{i,n+1} \in \partial(f_{t_{i,n+1}} + \mathrm{I}_{\operatorname{dom} f})(x),$$

for indices $t_{i,1}, \ldots, t_{i,n+1} \in T(x)$, such that $\lambda_{i,1} + \cdots + \lambda_{i,n+1} = 1$ and

$$z_i = \lambda_{i,1} z_{i,1} + \cdots + \lambda_{i,n+1} z_{i,n+1}. \tag{24}$$

We may assume, without loss of generality, that

$$\lambda_{i,k} \to \lambda_k \geq 0, \ \ k = 1, \ldots, n+1, \ \text{ and } \lambda_1 + \cdots + \lambda_{n+1} = 1.$$

Also, due to conditions (i) and (ii), we can find a common directed set $\mathbb{D}$ such that the nets $(t_{i,k})_{i \in \mathbb{D}}$ converge, say

$$t_{i,k} \to_{\mathbb{D}} t_k \in T(x), \ \ k = 1, \ldots, n+1. \tag{25}$$

At this step, we show that the nets $(\lambda_{i,k} z_{i,k})_{i \in \mathbb{D}}$, $k = 1, \ldots, n+1$, converge. Indeed, since $\operatorname{ri}(\operatorname{dom} f) \neq \emptyset$ and $x \in \operatorname{dom} f$, thanks to the accessibility lemma and the continuity of f on each one of the segments $[x, v]$, $v \in \operatorname{ri}(\operatorname{dom} f)$ (recall that $f(x) = (\operatorname{cl} f)(x)$), we may choose $x_0 \in \operatorname{ri}(\operatorname{dom} f)$ close enough to x to guarantee that $f(x_0) - f(x) + 1 \geq 0$, and some $r > 0$ such that $x_0 + (r\mathbb{B}) \cap F \subset \operatorname{dom} f$, where $\mathbb{B}$ is the unit closed ball in $\mathbb{R}^n$, $F := \operatorname{span}(\operatorname{dom} f - x_0)$, and

$$f(x_0 + y) \leq f(x_0) + 1 \ \text{ for all } y \in (r\mathbb{B}) \cap F.$$

Hence, for all $y \in (r\mathbb{B}) \cap F$, $i \in \mathbb{D}$, and $k = 1, \ldots, n+1$, $x_0 + y \in x_0 + (r\mathbb{B}) \cap F \subset \operatorname{dom} f$ and

$$\begin{aligned} \langle z_{i,k}, x_0 + y - x \rangle &\leq (f_{t_{i,k}} + \mathrm{I}_{\operatorname{dom} f})(x_0 + y) - (f_{t_{i,k}} + \mathrm{I}_{\operatorname{dom} f})(x) \\ &= f_{t_{i,k}}(x_0 + y) - f_{t_{i,k}}(x) \\ &\leq f(x_0 + y) - f(x) \leq f(x_0) - f(x) + 1, \end{aligned}$$

and this yields, multiplying by $\lambda_{i,k}$,

$$\langle \lambda_{i,k} z_{i,k}, x_0 + y - x \rangle \leq \lambda_{i,k}(f(x_0) - f(x) + 1) \leq f(x_0) - f(x) + 1. \tag{26}$$

In particular, for $y = 0_n$ we get

$$\langle \lambda_{i,k} z_{i,k}, x_0 - x \rangle \leq f(x_0) - f(x) + 1, \quad k = 1, \ldots, n+1.$$

Then, since $\langle z_i, x_0 - x\rangle \to_{\mathbb{D}} \langle z, x_0 - x\rangle$, and due to (24), the last relations entail the existence of some $m \geq 0$ such that

$$\langle \lambda_{i,k} z_{i,k}, x_0 - x\rangle \geq -m \text{ for all } i \in \mathbb{D} \text{ and } k = 1, \dots, n+1.$$

Thus, (26) gives rise to, for all $y \in (r\mathbb{B}) \cap F$, $i \in \mathbb{D}$, and $k = 1, \dots, n+1$,

$$\langle \lambda_{i,k} z_{i,k}, y\rangle \leq \langle \lambda_{i,k} z_{i,k}, x - x_0\rangle + f(x_0) - f(x) + 1 \leq m + f(x_0) - f(x) + 1,$$

that is, if

$$\rho := m + f(x_0) - f(x) + 1,$$

we have

$$(\lambda_{i,k} z_{i,k})_i \subset \rho r^{-1}(\mathbb{B} \cap F)^\circ = \rho r^{-1}\mathbb{B} + F^\perp,$$

and so there exists $(v_{i,k})_i \subset F^\perp$ such that $(\lambda_{i,k} z_{i,k} + v_{i,k})_i \subset \rho r^{-1}\mathbb{B}$; hence, without loss of generality, there must exist $w_1, \dots, w_{n+1} \in \mathbb{R}^n$ such that

$$\lambda_{i,k} z_{i,k} + v_{i,k} \to w_k,\ k = 1, \dots, n+1. \tag{27}$$

Moreover, writing (recall (24))

$$z_i = (\lambda_{i,1} z_{i,1} + v_{i,1}) + \cdots + (\lambda_{i,n+1} z_{i,n+1} + v_{i,n+1}) - \sum_{k=1,\dots,n+1} v_{i,k},$$

and since $z_i \to z$, we conclude that (without loss of generality)

$$\sum_{k=1,\dots,n+1} v_{i,k} \to u = z - \sum_{k=1,\dots,n+1} w_k. \tag{28}$$

In particular, observing that $F = \operatorname{span}(\operatorname{dom} f - x_0) = \operatorname{span}(\operatorname{dom} f - x)$, we have for all $y \in \operatorname{dom} f$,

$$\langle \lambda_{i,k} z_{i,k} - w_k, y - x\rangle = \langle \lambda_{i,k} z_{i,k} + v_{i,k} - w_k, y - x\rangle \to 0.$$

Now, since $(v_{i,k})_i \subset F^\perp$, (28) leads us to

$$u \in F^\perp = (\operatorname{dom} f - x)^\perp. \tag{29}$$

Let us analyze the following two possibilities: if $\lambda_k > 0$, then $\lambda_{i,k} > 0$ eventually, and so (27) implies that $z_{i,k} + \lambda_{i,k}^{-1} v_{i,k} \to \lambda_k^{-1} w_k$. Moreover, for all $y \in \operatorname{dom} f$ we have, eventually,

$$\begin{aligned}\left\langle z_{i,k} + \lambda_{i,k}^{-1} v_{i,k}, y - x\right\rangle = \langle z_{i,k}, y - x\rangle &\leq (f_{t_{i,k}} + \mathrm{I}_{\operatorname{dom} f})(y) - (f_{t_{i,k}} + \mathrm{I}_{\operatorname{dom} f})(x)\\ &= f_{t_{i,k}}(y) - f_{t_{i,k}}(x),\end{aligned}$$

which at the limit gives us, by condition (ii), (25), and the fact that $t_k \in T(x)$,

$$\left\langle \lambda_k^{-1} w_k, y - x\right\rangle \leq \limsup_{i \in \mathbb{D}} f_{t_{i,k}}(y) - f_{t_k}(x) \leq f_{t_k}(y) - f_{t_k}(x),$$

that is,

$$\lambda_k^{-1} w_k \in \partial(f_{t_k} + \mathrm{I}_{\operatorname{dom} f})(x). \tag{30}$$

Otherwise, if $\lambda_k = 0$, then for all $y \in \operatorname{dom} f$ we have (eventually)

$$\langle \lambda_{i,k} z_{i,k} + v_{i,k}, y - x\rangle = \langle \lambda_{i,k} z_{i,k}, y - x\rangle \leq \lambda_{i,k}(f_{t_{i,k}}(y) - f_{t_{i,k}}(x)) \leq \lambda_{i,k}(f(y) - f(x)),$$

which at the limit gives us

$$\langle w_k, y - x\rangle \leq \limsup_{i\in\mathbb{D}} \lambda_{i,k}(f(y) - f(x)) = 0;$$

that is, $w_k \in \mathrm{N}_{\operatorname{dom} f}(x)$, and so

$$\sum_{k \text{ s.t. } \lambda_k = 0} w_k \in \mathrm{N}_{\operatorname{dom} f}(x). \tag{31}$$

To summarize, using (24) together with (28), (30), (31), and (29),

$$\begin{aligned}
z = \lim_{i\in\mathbb{D}} z_i &= \lim_{i\in\mathbb{D}} \left(\sum_{\lambda_k > 0} \lambda_{i,k}(z_{i,k} + \lambda_{i,k}^{-1} v_{i,k}) + \sum_{\lambda_k = 0} (\lambda_{i,k} z_{i,k} + v_{i,k}) - \sum_{k=1,\dots,n+1} v_{i,k} \right) \\
&= \sum_{\lambda_k>0} w_k + \sum_{\lambda_k=0} w_k - u \\
&\in \sum_{\lambda_k>0} \lambda_k \partial(f_{t_k} + \mathrm{I}_{\operatorname{dom} f})(x) + \mathrm{N}_{\operatorname{dom} f}(x) + (\operatorname{dom} f - x)^{\perp} \\
&= \sum_{\lambda_k>0} \lambda_k (\partial(f_{t_k} + \mathrm{I}_{\operatorname{dom} f})(x) + \mathrm{N}_{\operatorname{dom} f}(x)) \\
&\subset \sum_{\lambda_k>0} \lambda_k \partial(f_{t_k} + \mathrm{I}_{\operatorname{dom} f} + \mathrm{I}_{\operatorname{dom} f})(x) \\
&= \sum_{\lambda_k>0} \lambda_k \partial(f_{t_k} + \mathrm{I}_{\operatorname{dom} f})(x) \subset E,
\end{aligned}$$

showing that E is closed, and (23) reads

$$\partial f(x) = \operatorname{cl} E = E = \operatorname{co}\left\{ \bigcup_{t\in T(x)} \partial(f_t + \mathrm{I}_{\operatorname{dom} f})(x) \right\}. \tag{32}$$

We finish the proof by using condition (19), which guarantees the exact subdifferential sum rule [26]. This allows us to simplify (20) and write

$$\begin{aligned}
\partial f(x) &= \operatorname{co}\left\{ \bigcup_{t\in T(x)} \partial(f_t + \mathrm{I}_{\operatorname{dom} f})(x) \right\} \\
&= \operatorname{co}\left\{ \bigcup_{t\in T(x)} \partial f_t(x) + \mathrm{N}_{\operatorname{dom} f}(x) \right\} \\
&= \operatorname{co}\left\{ \bigcup_{t\in T(x)} \partial f_t(x) \right\} + \mathrm{N}_{\operatorname{dom} f}(x);
\end{aligned}$$

thus, (21) is proved. ∎

In the following proposition we compare condition (18) with the usual Rockafellar condition

$$\bigcap_{t\in T} \mathrm{ri}(\mathrm{dom}\, f_t) \neq \emptyset, \tag{33}$$

guaranteeing the following standard sum rule [26] (when T is finite):

$$\partial\left(\sum_{t\in T} f_t\right) = \sum_{t\in T} \partial f_t. \tag{34}$$

PROPOSITION 4. *Assume that T is finite. Then condition* (33) *and*

$$\mathrm{ri}(\mathrm{dom}\, f_t) \cap \mathrm{dom}\, f \neq \emptyset, \quad \textit{for all } t\in T, \tag{35}$$

are equivalent and under either of them we have, for all $x \in X$,

$$\partial f(x) = \mathrm{co}\left\{\bigcup_{t\in T(x)} \partial f_t(x)\right\} + \sum_{t\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(x).$$

Proof. On the one hand, condition (33) and Theorem 6.5 in [26] imply that, for all $t \in T$,

$$\begin{aligned}\mathrm{ri}(\mathrm{dom}\, f_t) \cap \mathrm{ri}(\mathrm{dom}\, f) &= \mathrm{ri}(\mathrm{dom}\, f_t) \cap \mathrm{ri}\left(\bigcap_{i\in T} \mathrm{dom}\, f_i\right)\\ &= \mathrm{ri}(\mathrm{dom}\, f_t) \cap \bigcap_{i\in T} \mathrm{ri}(\mathrm{dom}\, f_i)\\ &= \bigcap_{i\in T} \mathrm{ri}(\mathrm{dom}\, f_i) \neq \emptyset,\end{aligned}$$

and (35) follows. On the other hand, if condition (35) holds, then we choose $x_i \in \mathrm{ri}(\mathrm{dom}\, f_i) \cap \mathrm{dom}\, f$ and define $\bar{x} := \sum_{i\in T} \frac{1}{|T|} x_i$, where $|T|$ (> 1) is the cardinal of T. Then for each $i_0 \in T$ we have

$$\sum_{i\in T\setminus\{i_0\}} \frac{1}{(|T|-1)} x_i \in \mathrm{dom}\, f \subset \mathrm{dom}\, f_{i_0},$$

and so, by the accessibility lemma,

$$\bar{x} = \frac{1}{|T|} x_{i_0} + \left(\frac{|T|-1}{|T|}\right) \sum_{i\in T\setminus\{i_0\}} \frac{1}{(|T|-1)} x_i \in \mathrm{ri}(\mathrm{dom}\, f_{i_0}).$$

In other words, $\bar{x} \in \bigcap_{i\in T} \mathrm{ri}(\mathrm{dom}\, f_i)$ and (33) holds.

Now, we fix $x \in X$. From the paragraph above, (18) holds, and the last statement of the proposition comes straightforwardly from Theorem 3 due to the relation $\mathrm{N}_{\mathrm{dom}\, f}(x) = \sum_{t\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(x)$. The last equality is a consequence of (34) when applied to the indicator functions of $\mathrm{dom}\, f_t$, $t \in T$. ∎

Now we consider the semi-infinite convex optimization problem

$$\inf_{\substack{f_t(x)\le 0,\ t\in T\\ x\in C}} f_0(x), \tag{$\mathcal{P}$}$$

where $C \subset \mathbb{R}^n$ is convex, T is a Hausdorff topological space, and the $f_t : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, for $t \in T \cup \{0\}$ (we assume, without loss of generality, that $0 \notin T$), are proper and convex. Let us define $g := \sup_{t\in T} f_t$,

$$D := \operatorname{dom} f_0 \cap \operatorname{dom} g,$$

and, for x being a feasible point of $(\mathcal{P})$,

$$A(x) := \{t \in T \mid f_t(x) = 0\}.$$

The following theorem provides different Fritz John-type necessary optimality conditions for problem $(\mathcal{P})$.

THEOREM 5. *Let $\bar{x} \in C$ be a feasible point of $(\mathcal{P})$ such that $A(\bar{x}) \neq \emptyset$, and assume that for some $\varepsilon_0 > 0$,*

(i) *the set $A_{\varepsilon_0}(\bar{x}) := \{t \in T \mid f_t(\bar{x}) \geq -\varepsilon_0\}$ is compact,*

(ii) *for each $z \in D \cap C$ the function $t \mapsto f_t(z)$ is usc on $A_{\varepsilon_0}(\bar{x})$.*

Then $\bar{x}$ is optimal for $(\mathcal{P})$ if and only if one of the following conditions holds:

(a)

$$0_n \in \operatorname{co}\left\{\partial(f_0 + \mathrm{I}_{D\cap C})(\bar{x}) \cup \bigcup_{t\in A(\bar{x})} \partial(f_t + \mathrm{I}_{D\cap C})(\bar{x})\right\}.$$

(b)

$$0_n \in \operatorname{co}\left\{\partial f_0(\bar{x}) \cup \bigcup_{t\in A(\bar{x})} \partial f_t(\bar{x})\right\} + \mathrm{N}_{D\cap C}(\bar{x}),$$

provided that

$$\operatorname{ri}(\operatorname{dom} f_t) \cap \operatorname{ri}(D \cap C) \neq \emptyset \quad \textit{for all } t \in A(\bar{x}) \cup \{0\}. \tag{36}$$

(c)

$$0_n \in \operatorname{co}\left\{\partial f_0(\bar{x}) \cup \bigcup_{t\in A(\bar{x})} \partial f_t(\bar{x})\right\} + \mathrm{N}_C(\bar{x}) + \sum_{t\in T\cup\{0\}} \mathrm{N}_{\operatorname{dom} f_t}(\bar{x}),$$

provided that T is compact, for each $z \in \bigcap_{t\in T\cup\{0\}} \operatorname{dom} f_t \cap C$ the function $t \mapsto f_t(z)$ is usc on T, and the family $\{C,\ \operatorname{dom} f_t,\ t \in T \cup \{0\}\}$ is strong CHIP at $\bar{x}$. In particular, this happens when T is finite and

$$\bigcap_{t\in T\cup\{0\}} \operatorname{ri}(\operatorname{dom} f_t) \cap \operatorname{ri}(C) \neq \emptyset. \tag{37}$$

Proof. (a) It is well known that $\bar{x}$ is optimal of $(\mathcal{P})$ if and only if $\bar{x}$ is an unconstrained minimum of the function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, defined as

$$f(x) := \sup\{f_0(x) - f_0(\bar{x}),\ f_1(x),\ f_t(x),\ t \in T\},$$

where

$$f_1(x) = \mathrm{I}_C(x) - 2\varepsilon_0$$

(assuming that $1 \notin T$), and this happens if and only if $0_n \in \partial f(\bar{x})$. On the one hand, since $f(\bar{x}) = 0$ and $f_1(\bar{x}) = -2\varepsilon_0 < -\varepsilon_0 < 0$, the set of ε-active indices at $\bar{x}$ for the supremum function f is

$$\{0\} \cup A_{\varepsilon_0}(\bar{x}),$$

which is compact due to assumption (i) (we are considering here the Hausdorff topological space $\widetilde{T} := T \cup \{0, 1\}$, with 0 and 1 being isolated points). On the other hand, using assumption (ii), for all $z \in \operatorname{dom} f = D \cap C$ the mapping $t \mapsto f_t(z)$ is usc on the set $\{0\} \cup A_{\varepsilon_0}(\bar{x})$. Consequently, Theorem 3 applies and yields

$$\partial f(\bar{x}) = \operatorname{co}\left\{ \bigcup_{t \in A(\bar{x}) \cup \{0\}} \partial (f_t + \mathrm{I}_{D \cap C})(\bar{x}) \right\}.$$

Thus, the equivalence with condition (a) follows.

(b) This follows as in (a) but applying (21) instead of (20) in Theorem 3.

(c) The compactness of T and the upper semicontinuity of the functions $t \mapsto f_t(z)$, $z \in \bigcap_{t \in T \cup \{0\}} \operatorname{dom} f_t \cap C$, imply that

$$D \cap C = \bigcap\nolimits_{t \in T \cup \{0\}} \operatorname{dom} f_t \cap C.$$

Then (c) comes from (b) and the definition of the strong CHIP property. The second statement is also straightforward, taking into account that (37) implies the strong CHIP property when T is finite. ∎

We derive next the KKT condition for problem $(\mathcal{P})$ under the following Slater qualification condition:

$$\sup_{t \in T} f_t(x_0) < 0 \text{ for some } x_0 \in C \cap \operatorname{dom} f_0. \tag{38}$$

Corollary 6. *If in* (c) *of Theorem* 5 *we assume additionally that condition* (38) *holds, then there exist a (possibly empty) finite set $\widehat{T}(\bar{x}) \subset A(\bar{x})$ such that $\partial f_t(\bar{x}) \neq \emptyset$ for $t \in \widehat{T}(\bar{x})$ and scalars $\lambda_t > 0$ for $t \in \widehat{T}(\bar{x})$ satisfying*

$$0_n \in \partial f_0(\bar{x}) + \sum_{t \in \widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t \in T} \mathrm{N}_{\operatorname{dom} f_t}(\bar{x}), \tag{39}$$

with the convention that $\sum_\emptyset = \{0_n\}$.

Proof. According to Theorem 5(c), we have

$$0_n \in \operatorname{co}\left\{ \bigcup_{t \in A(\bar{x}) \cup \{0\}} \partial f_t(\bar{x}) \right\} + \mathrm{N}_C(\bar{x}) + \sum_{t \in T \cup \{0\}} \mathrm{N}_{\operatorname{dom} f_t}(\bar{x}), \tag{40}$$

and so, by (6),

$$\bigcup_{t \in A(\bar{x}) \cup \{0\}} \partial f_t(\bar{x}) \neq \emptyset.$$

If $\partial f_0(\bar{x})$ does not intervene in (40), i.e., defining $g := \max_{t\in T} f_t$,

$$0_n \in \mathrm{co}\left\{\bigcup_{t\in A(\bar{x})} \partial f_t(\bar{x})\right\} + \mathrm{N}_{\mathrm{dom}\, f_0}(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x})$$

$$\subset \mathrm{co}\left\{\bigcup_{t\in A(\bar{x})} \partial f_t(\bar{x})\right\} + \mathrm{N}_{\mathrm{dom}\, f_0}(\bar{x}) + \mathrm{N}_C(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, g}(\bar{x}), \tag{41}$$

then, since

$$\emptyset \neq \mathrm{co}\left\{\bigcup_{t\in A(\bar{x})} \partial f_t(\bar{x})\right\} \subset \partial g(\bar{x}), \tag{42}$$

relation (41) gives rise to

$$\begin{aligned} 0_n &\in \partial g(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f_0}(\bar{x}) + \mathrm{N}_C(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, g}(\bar{x}) \\ &\subset \partial(g + \mathrm{I}_{\mathrm{dom}\, g} + \mathrm{I}_{\mathrm{dom}\, f_0} + \mathrm{I}_C)(\bar{x}) = \partial(g + \mathrm{I}_{C\cap \mathrm{dom}\, f_0})(\bar{x}), \end{aligned}$$

which contradicts the Slater condition, as

$$0 = g(\bar{x}) = (g + \mathrm{I}_{C\cap \mathrm{dom}\, f_0})(\bar{x}) \leq (g + \mathrm{I}_{C\cap \mathrm{dom}\, f_0})(x_0) = g(x_0) < 0.$$

Otherwise, if $\partial f_0(\bar{x})$ intervenes in (40) (hence, $\partial f_0(\bar{x}) \neq \emptyset$), then there would exist scalars $\alpha > 0$ and $\alpha_t \geq 0$, $t \in A(\bar{x})$, with only finitely many of them being positive, such that $\alpha + \sum_{t\in T(\bar{x})} \alpha_t = 1$,

$$0_n \in \alpha\partial f_0(\bar{x}) + \sum_{t\in A(\bar{x})} \alpha_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T\cup\{0\}} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x})$$

$$= \alpha\partial f_0(\bar{x}) + \sum_{t\in A(\bar{x})} \alpha_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x}) \text{ if } \alpha < 1, \tag{43}$$

and

$$0_n \in \partial f_0(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{i\in T\cup\{0\}} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x})$$

$$= \partial f_0(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{i\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x}) \text{ if } \alpha = 1, \tag{44}$$

since $\alpha\partial f_0(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f_0}(\bar{x}) = \alpha\partial f_0(\bar{x})$. Then (43) leads us to

$$0_n \in \partial f_0(\bar{x}) + \sum_{t\in A(\bar{x})} \alpha^{-1}\alpha_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x}),$$

which combined with (44) yield the existence of a (possibly empty) finite set $\widehat{T}(\bar{x}) \subset A(\bar{x})$ such that $\partial f_t(\bar{x}) \neq \emptyset$ for $t \in \widehat{T}(\bar{x})$, and scalars $\lambda_t > 0$ for $t \in \widehat{T}(\bar{x})$ satisfying

$$0_n \in \partial f_0(\bar{x}) + \sum_{t\in \widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \sum_{t\in T} \mathrm{N}_{\mathrm{dom}\, f_t}(\bar{x}). \qquad \square$$

Example 1. In $(\mathcal{P})$ take $n = 1$, $C = \mathbb{R}$, $T = \{1\}$, $f_0(x) = x$, and $f_1(x) = -\sqrt{x}$ if $x \geq 0$ and $+\infty$ if not. Then $\bar{x} = 0$ is the unique optimal point of $(\mathcal{P})$, $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathrm{dom}\, f_1) = \mathbb{R}_{++} \neq \emptyset$, $\mathrm{N}_C(0) = \{0\}$, and the Slater condition holds. Since $\partial f_0(0) = \{1\}$, $\partial f_1(0) = \emptyset$, and $\mathrm{N}_{\mathrm{dom}\, f_1}(0) = \mathrm{N}_{\mathbb{R}_+}(0) = \mathbb{R}_-$, we see that Corollary 6 is verified with $\widehat{T}(0) = \emptyset$:

$$0 \in \partial f_0(0) + \mathrm{N}_{\mathbb{R}_+}(0) = 1 + \mathbb{R}_-.$$

It turns out that we cannot get rid of the term $\mathrm{N}_{\mathrm{dom}\, f_1}(0)$.

Other KKT optimality conditions were established in [8] for problem $(\mathcal{P})$ when C is a closed convex set in an infinite-dimensional space, and the convex functions f_0, f_t, $t \in T$, are proper and lsc. In [8] the authors appealed to some conditions related to the so-called locally Farkas–Minkowski property and the basic constraint qualification. Previously, in [11, Chapter 7] KKT conditions for convex semi-infinite optimization were derived for finite-valued functions using a closedness condition that is implied by some version of Slater's qualification, first considered in [21].

The following KKT conditions for an ordinary (T finite) convex optimization problem are obtained from Theorem 5, where the strong CHIP property used in Theorem 5(c) is replaced by appropriate continuity conditions on the constraint functions. For simplicity, we shall assume that

$$C \subset \mathrm{dom}\, f_0.$$

(Otherwise, we shall consider the abstract constraint $x \in C \cap \mathrm{dom}\, f_0$ instead of $x \in C$.)

COROLLARY 7. *Let T be finite and let $\bar{x} \in C$ be optimal for $(\mathcal{P})$ such that $A(\bar{x}) \neq \emptyset$. If condition (38) holds and the functions f_t, $t \in T$, are continuous at some common interior point in C ($\subset \mathrm{dom}\, f_0$), then there exist a (possibly empty) set $\widehat{T}(\bar{x}) \subset A(\bar{x})$ such that $\partial f_t(\bar{x}) \neq \emptyset$ for $t \in \widehat{T}(\bar{x})$ and scalars $\lambda_t > 0$ for $t \in \widehat{T}(\bar{x})$ satisfying*

$$0_n \in \partial f_0(\bar{x}) + \sum_{t \in \widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \mathrm{N}_{\cap_{t \in T} \mathrm{dom}\, f_t}(\bar{x}), \tag{45}$$

with the convention that $\sum_{\emptyset} = \{0_n\}$. Consequently, if $\bar{x}$ is the mentioned common continuity point, then

$$0_n \in \partial f_0(\bar{x}) + \sum_{t \in \widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}).$$

Proof. By Theorem 5(a) we have that (as T is finite)

$$0_n \in \mathrm{co} \left\{ \partial(f_0 + \mathrm{I}_{D \cap C})(\bar{x}) \cup \bigcup_{t \in A(\bar{x})} \partial(f_t + \mathrm{I}_{D \cap C})(\bar{x}) \right\},$$

that is, there exist scalars $\alpha \geq 0$ and $\alpha_t \geq 0$, $t \in A(\bar{x})$, such that $\alpha + \sum_{t \in A(\bar{x})} \alpha_t = 1$ and

$$0_n \in \alpha \partial(f_0 + \mathrm{I}_{D \cap C})(\bar{x}) + \sum_{t \in A(\bar{x})} \alpha_t \partial(f_t + \mathrm{I}_{D \cap C})(\bar{x}).$$

As in the proof of Corollary 6, due to the Slater condition the last relation entails the existence of a (possibly empty) set $\widehat{T}(\bar{x}) \subset A(\bar{x})$ such that $\partial(f_t + \mathrm{I}_{D\cap C})(\bar{x}) \neq \emptyset$ for $t \in \widehat{T}(\bar{x})$ and scalars $\lambda_t > 0$ for $t \in \widehat{T}(\bar{x})$ satisfying

$$0_n \in \partial(f_0 + \mathrm{I}_{D\cap C})(\bar{x}) + \sum_{t\in\widehat{T}(\bar{x})} \lambda_t \partial(f_t + \mathrm{I}_{D\cap C})(\bar{x}). \tag{46}$$

Now, due to the Moreau–Rockafellar sum rule, the continuity assumption on the functions f_t, $t \in T$, ensures that (recall that $g = \max_{t\in T} f_t$)

$$\partial(f_0 + \mathrm{I}_{D\cap C})(\bar{x}) = \partial(f_0 + \mathrm{I}_C + \mathrm{I}_{\operatorname{dom} g})(\bar{x}) = \partial f_0(\bar{x}) + \mathrm{N}_{\operatorname{dom} g}(\bar{x}),$$

and, for all $t \in \widehat{T}(\bar{x})$,

$$\begin{aligned}\partial(f_t + \mathrm{I}_{D\cap C})(\bar{x}) &= \partial(f_t + \mathrm{I}_C + \mathrm{I}_{\operatorname{dom} g})(\bar{x}) \\ &= \partial(f_t + \mathrm{I}_C)(\bar{x}) + \mathrm{N}_{\operatorname{dom} g}(\bar{x}) \\ &= \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \mathrm{N}_{\operatorname{dom} g}(\bar{x});\end{aligned}$$

hence, $\partial f_t(\bar{x}) \neq \emptyset$. In other words, using (46) and the fact that $\operatorname{dom} g = \cap_{t\in T} \operatorname{dom} f_t$,

$$0_n \in \partial f_0(\bar{x}) + \sum_{t\in\widehat{T}(\bar{x})} \lambda_t \partial f_t(\bar{x}) + \mathrm{N}_C(\bar{x}) + \mathrm{N}_{\cap_{t\in T} \operatorname{dom} f_t}(\bar{x}). \qquad \square$$

Remark 1. Many results in convex analysis and optimization do not require the lower semicontinuity of the involved functions (or the closedness of the constraint sets). This is the case of the Moreau–Rockafellar sum rule for the subdifferential of the sum of convex not-necessarily lsc functions (see, also, the different results gathered in [31, Theorem 2.8.7]).

Let us illustrate the issue raised by the lack of closedness conditions. Consider the optimization problem $(\mathcal{P})$, $\inf_{(x,y)\in C} f(x,y)$, where

$$f(x,y) := \begin{cases} x^2 & \text{if } x < 0, \\ 1 & \text{if } x = 0, \\ +\infty & \text{if } x > 0, \end{cases} \qquad C := \left\{(x,y) \in \mathbb{R}^2 : x > 0\right\} \cup \{(0,0)\}.$$

Observe that $(0,0)$ is an optimal solution of $(\mathcal{P})$, and it is also optimal for the regularized problem $(\mathcal{P}_r)$, $\inf_{(x,y)\in \operatorname{cl} C} (\operatorname{cl} f)(x,y)$, although the optimal set of (P_r) is much larger. The KKT optimality conditions for $(\mathcal{P}_r)$ are

$$(0,0) \in \partial(\operatorname{cl} f)(0,0) + \mathrm{N}_{\operatorname{cl} C}(0,0) = \mathbb{R} \times \{0\},$$

but $\partial f(0,0) = \emptyset$, and this precludes the existence of KKT optimality conditions for problem $(\mathcal{P})$ involving only the original data, f and C.

We close this section by making a short discussion to relate problem $(\mathcal{P})$ to its regularization

$$(\mathcal{P}_r) \qquad \inf_{\substack{(\operatorname{cl} f_t)(x) \le 0,\ t\in T \\ x\in \operatorname{cl} C}} (\operatorname{cl} f_0)(x).$$

This discussion aims to clarify the role played by the closedness assumption, and to compare the optimality conditions for $(\mathcal{P})$ and $(\mathcal{P}_r)$.

In the above example of Remark 1, $(\mathcal{P}_r)$ admits KKT necessary optimality conditions, whereas $(\mathcal{P})$ does not. This is due to the failure of assumption (36) as

$$\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathrm{dom}\, f_0 \cap C) = \emptyset.$$

Nevertheless, under such a condition (36), the lower semicontinuity of f_0 and the closedness of $\mathrm{dom}\, f_0 \cap C$ are implicitly evoked. To see this, consider for simplicity that only the abstract constraint $x \in C$ is present, and $(\mathcal{P})$ and $(\mathcal{P}_r)$ are equivalently written as follows:

$$\begin{aligned} &(\mathcal{P}) \qquad \inf_{x \in \mathrm{dom}\, f_0 \cap C} f_0(x), \\ &(\mathcal{P}_r) \qquad \inf_{x \in \mathrm{cl}(\mathrm{dom}\, f_0 \cap C)} (\mathrm{cl}\, f_0)(x). \end{aligned} \tag{47}$$

The relation between $(\mathcal{P})$ and $(\mathcal{P}_r)$ under (36) is analyzed in the following corollary, where the optimality conditions for both problems turn out to be equivalent.

COROLLARY 8. *Let $\bar{x}$ be optimal for $(\mathcal{P})$ in* (47). *If condition* (36) *holds, i.e.,*

$$\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathrm{dom}\, f_0 \cap C) \neq \emptyset, \tag{48}$$

then

(i) *$\bar{x}$ is also optimal for $(\mathcal{P}_r)$ in* (47)*;*

(ii) *the optimality conditions for $(\mathcal{P}_r)$ hold at $\bar{x}$, i.e.,*

$$0_n \in \partial(\mathrm{cl}\, f_0)(\bar{x}) + \mathrm{N}_{\overline{\mathrm{dom}\, f_0 \cap C}}(\bar{x});$$

(iii) *$(\mathcal{P})$ and $(\mathcal{P}_r)$ have the same optimal value, i.e., $v(\mathcal{P}) = v(\mathcal{P}_r)$;*

(iv) *$(\mathcal{P})$ and $(\mathcal{P}_r)$ satisfy the same associated optimality conditions, i.e.,*

$$0_n \in \partial f_0(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f_0 \cap C}(\bar{x}) \iff 0_n \in \partial(\mathrm{cl}\, f_0)(\bar{x}) + \mathrm{N}_{\overline{\mathrm{dom}\, f_0 \cap C}}(\bar{x}). \tag{49}$$

Proof. First, it is known that a function and its closure have the same infimum; hence,

$$v(\mathcal{P}) = \inf_{x \in X} (f_0 + \mathrm{I}_{\mathrm{dom}\, f_0 \cap C})(x) = \inf_{x \in X} \mathrm{cl}(f_0 + \mathrm{I}_{\mathrm{dom}\, f_0 \cap C})(x),$$

and (48) yields (using [26, Theorem 9.3])

$$v(\mathcal{P}) = \inf_{x \in X} \Big((\mathrm{cl}\, f_0)(x) + \mathrm{I}_{\overline{\mathrm{dom}\, f_0 \cap C}}(x) \Big) = v(\mathcal{P}_r),$$

that is, (iii) follows.

(i) This holds because $\bar{x} \in \mathrm{dom}\, f_0 \cap C \subset \overline{\mathrm{dom}\, f_0 \cap C}$ and, for all $x \in \overline{\mathrm{dom}\, f_0 \cap C}$,

$$\begin{aligned} (\mathrm{cl}\, f_0)(\bar{x}) \leq f_0(\bar{x}) = v(\mathcal{P}) = v(\mathcal{P}_r) \\ = \inf_{x \in X} \Big((\mathrm{cl}\, f_0)(x) + \mathrm{I}_{\overline{\mathrm{dom}\, f_0 \cap C}}(x) \Big) \leq (\mathrm{cl}\, f_0)(x). \end{aligned} \tag{50}$$

(ii) Since

$$\mathrm{ri}(\mathrm{dom}(\mathrm{cl}\, f_0)) \cap \mathrm{ri}(\overline{(\mathrm{dom}\, f_0) \cap C}) = \mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}((\mathrm{dom}\, f_0) \cap C) \neq \emptyset,$$

condition (48) holds for $(\mathcal{P}_r)$, so (ii) follows by Theorem 5(b).

(iv) This assertion follows because $(\mathrm{cl}\, f_0)(\bar{x}) = f_0(\bar{x})$, which comes from (50), since in this case $\partial(\mathrm{cl}\, f_0)(\bar{x}) = \partial f_0(\bar{x})$ (recall (8)) and $\mathrm{N}_{\mathrm{dom}\, f_0 \cap C}(\bar{x}) = \mathrm{N}_{\overline{\mathrm{dom}\, f_0 \cap C}}(\bar{x})$. ∎

5. Infinite-dimensional qualification conditions for the max function. This section deals with the maximum function

$$f := \max_{k \in T := \{1,\dots,p\}} f_k,$$

where $p \geq 2$, of a finite family of proper convex functions, $f_k : X \to \mathbb{R} \cup \{+\infty\}$, $k \in T$, defined on an lcs X. This model constitutes a relevant particular case of the compactly indexed setting studied in [6]. We give here the main characterization of ∂f, which holds under much weaker conditions than the continuity of the supremum function f, which is frequently used.

For a better understanding of the similarity of our conditions to those used in the literature for the sum rule of the subdifferential, observe that for any pair of convex functions f_1 and f_2, $f := \max\{f_1, f_2\}$, we have that

$$\begin{aligned} x^* \in \partial f(x) \Leftrightarrow (x^*, -1) &\in \mathrm{N}_{\mathrm{epi}\, f}(x, f(x)) \\ &= \mathrm{N}_{\mathrm{epi}\, f_1 \cap \mathrm{epi}\, f_2}(x, f(x)) = \partial(\mathrm{I}_{\mathrm{epi}\, f_1 \cap \mathrm{epi}\, f_2})(x, f(x)) \\ &= \partial(\mathrm{I}_{\mathrm{epi}\, f_1} + \mathrm{I}_{\mathrm{epi}\, f_2})(x, f(x)). \end{aligned}$$

Thus, qualification conditions ensuring the possibility of decomposing the subdifferential of the sum $\partial(\mathrm{I}_{\mathrm{epi}\, f_1} + \mathrm{I}_{\mathrm{epi}\, f_2})$ would lead to a characterization of ∂f in terms of ∂f_1 and ∂f_2. This idea can obviously be applied to finitely many functions $f_1, \dots, f_p$ via a continuity condition affecting all of them (except perhaps one); see [25, 30, 31]. In contrast, we introduce here a new approach allowing us to relax the continuity assumption in the mentioned references, confining it to the active functions f_k, $k \in T(x)$ (except perhaps one of them).

THEOREM 9. *Given a fixed $x \in X$, we assume that each one of the functions f_k, $k \in T(x)$, except perhaps one of them, say f_{k_0}, is continuous at some point in* $\mathrm{dom}\, f$. *Then*

$$\partial f(x) = \mathrm{co}\left\{ \bigcup_{k \in T(x) \setminus \{k_0\}} \partial f_k(x) \bigcup \partial(f_{k_0} + \mathrm{I}_{\mathrm{dom}\, f})(x) \right\} + \mathrm{N}_{\mathrm{dom}\, f}(x). \tag{51}$$

Proof. The inclusion "$\supset$" is straightforward. To prove the inclusion "$\subset$" we may assume that $\partial f(x) \neq \emptyset$; hence, $f(x) = (\mathrm{cl}\, f)(x) \in \mathbb{R}$. Thus, we may suppose that $x = \theta$ and $f(\theta) = (\mathrm{cl}\, f)(\theta) = 0$. For the sake of simplicity, we write

$$T(\theta) = \{1, 2, \dots, m, m+1\},$$

with $p \geq m+1$, and $k_0 = m+1$. By the current assumption, for each $k \in \{1, \dots, m\}$ we take $x_k \in \mathrm{dom}\, f$ such that f_k is continuous at x_k, all of which we may suppose are equal, say $x_k = \hat{x}$ for $k = 1, \dots, m$; indeed, due to the accessibility lemma the point $\hat{x} := \frac{1}{m} \sum_{k=1,\dots,m} x_k \in \mathrm{dom}\, f$ also satisfies the continuity condition of the theorem.

Now, we take $M \geq 0$ and a θ-neighborhood $W \subset X$ such that, for all $w \in W$,

$$f_k(\hat{x} + w) \leq M, \;\; k = 1, 2, \dots, m. \tag{52}$$

We introduce the family of finite-dimensional linear subspaces

$$\widetilde{\mathcal{F}}(\theta) := \{\mathrm{span}\{L, \hat{x}\} \mid L \in \mathcal{F}(\theta)\},$$

together with

$$\widetilde{\mathcal{N}}_{X^*} := \{V \in \mathcal{N}_{X^*} \mid \sigma_V(\hat{x}) \leq 1\},$$

and endow the Cartesian product $\widetilde{\mathcal{F}}(\theta) \times \widetilde{\mathcal{N}}_{X^*}$ with the partial order "$\succeq$" defined as follows: $\alpha_2 \succeq \alpha_1$, with $\alpha_1 := (L_1, V_1) \in \widetilde{\mathcal{F}}(\theta) \times \widetilde{\mathcal{N}}_{X^*}$, $\alpha_2 := (L_2, V_2) \in \widetilde{\mathcal{F}}(\theta) \times \widetilde{\mathcal{N}}_{X^*}$, if and only if

$$L_1 \subset L_2,\ V_2 \subset V_1;$$

in this way, $(\widetilde{\mathcal{F}}(\theta) \times \widetilde{\mathcal{N}}_{X^*}, \succeq)$ becomes a directed set.

Take $x^* \in \partial f(\theta)$. By the Moreau–Rockafellar theorem, and according to Proposition 2 (applied with the discrete topology on T), we have that

$$\begin{aligned}
x^* &\in \bigcap_{L \in \mathcal{F}(\theta)} \overline{\mathrm{co}} \left\{ \bigcup_{k \in T(\theta)} \partial(f_k + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta) \right\} \\
&\subset \bigcap_{L \in \widetilde{\mathcal{F}}(\theta)} \overline{\mathrm{co}} \left\{ \bigcup_{k \in T(\theta)} \partial(f_k + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta) \right\} \\
&= \bigcap_{L \in \widetilde{\mathcal{F}}(\theta)} \overline{\mathrm{co}} \left\{ \bigcup_{k \in \{1,\ldots,m\}} (\partial f_k(\theta) + \mathrm{N}_{L \cap \operatorname{dom} f}(\theta)) \bigcup \partial(f_{m+1} + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta) \right\} \\
&= \bigcap_{(L,V) \in \widetilde{\mathcal{F}}(\theta) \times \widetilde{\mathcal{N}}_{X^*}} \Bigg(\mathrm{co} \Big\{ \bigcup\nolimits_{k \in \{1,\ldots,m\}} (\partial f_k(\theta) + \mathrm{N}_{L \cap \operatorname{dom} f}(\theta)) \\
&\qquad\qquad \bigcup \partial(f_{m+1} + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta) + V \Big\} \Bigg).
\end{aligned}$$

Hence, for each $\alpha := (L, V) \in \widetilde{\mathcal{F}}(\theta) \times \widetilde{\mathcal{N}}_{X^*}$, there exist $(\lambda_{1,\alpha}, \ldots, \lambda_{m+1,\alpha}) \in \Delta_{m+1}$ (the canonical simplex), $y^*_{k,\alpha} \in \partial f_k(\theta)$ and $u^*_{k,\alpha} \in \mathrm{N}_{L \cap \operatorname{dom} f}(\theta)$, $k = 1, \ldots, m$, $y^*_{m+1,\alpha} \in \partial(f_{m+1} + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta)$, and $z^*_\alpha \in V$ such that

$$x^* = \lambda_{1,\alpha}(y^*_{1,\alpha} + u^*_{1,\alpha}) + \cdots + \lambda_{m,\alpha}(y^*_{m,\alpha} + u^*_{m,\alpha}) + \lambda_{m+1,\alpha} y^*_{m+1,\alpha} + z^*_\alpha, \tag{53}$$

or, equivalently,

$$x^* = \lim_\alpha (\lambda_{1,\alpha}(y^*_{1,\alpha} + u^*_{1,\alpha}) + \cdots + \lambda_{m,\alpha}(y^*_{m,\alpha} + u^*_{m,\alpha}) + \lambda_{m+1,\alpha} y^*_{m+1,\alpha}). \tag{54}$$

We may suppose, without loss of generality, that

$$\lim_\alpha (\lambda_{1,\alpha}, \ldots, \lambda_{m+1,\alpha}) = (\lambda_1, \ldots, \lambda_{m+1}) \in \Delta_{m+1}.$$

Let us first verify that the nets $(\lambda_{k,\alpha} y^*_{k,\alpha})_\alpha$, $k = 1, \ldots, m$, weak*-converge in X^*. Indeed, given $k \in \{1, \ldots, m\}$, relation (52) yields, for all $w \in W$ and α,

$$\langle y^*_{k,\alpha}, \hat{x} + w \rangle \leq f_k(\hat{x} + w) - f_k(\theta) = f_k(\hat{x} + w) \leq M; \tag{55}$$

in particular, for $w = \theta$ it holds that

$$\langle y^*_{k,\alpha}, \hat{x} \rangle \leq f_k(\hat{x}) - f_k(\theta) \leq f(\hat{x}) \leq \max\{0, f(\hat{x})\},$$

while, as $y^*_{m+1,\alpha} \in \partial(f_{m+1} + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta)$,

$$\langle y^*_{m+1,\alpha}, \hat{x} \rangle \leq f_{m+1}(\hat{x}) - f_{m+1}(\theta) \leq f(\hat{x}) \leq \max\{0, f(\hat{x})\},$$

that is,

$$\left\langle \lambda_{k,\alpha} y_{k,\alpha}^*, \hat{x} \right\rangle \leq \max\{0, f(\hat{x})\} \text{ for all } \alpha \text{ and } k \in \{1, \ldots, m+1\}.$$

Hence, since we have that $\hat{x} \in L \cap \operatorname{dom} f$ for $L \in \widetilde{\mathcal{F}}(\theta)$, and $u_{k,\alpha}^* \in \mathrm{N}_{L \cap \operatorname{dom} f}(\theta)$, the inequalities

$$\left\langle u_{k,\alpha}^*, \hat{x} \right\rangle \leq 0 \text{ for all } \alpha \text{ and } k \in \{1, \ldots, m\}$$

imply that the nets

$$\left(\left\langle \lambda_{k,\alpha}(y_{k,\alpha}^* + u_{1,\alpha}^*), \hat{x} \right\rangle \right)_\alpha, \ k \in \{1, \ldots, m\}, \ \left(\left\langle \lambda_{m+1,\alpha} y_{m+1,\alpha}^*, \hat{x} \right\rangle \right)_\alpha$$

are bounded from above. In addition, due to the definition of $\widetilde{\mathcal{N}}_{X^*}$, we have that $\langle z_\alpha^*, \hat{x} \rangle \leq 1$ for all α, and so the net $(\langle z_\alpha^*, \hat{x} \rangle)_\alpha$ is also bounded from above. Consequently, thanks to the following inequalities derived from (53),

$$\begin{aligned}
\langle x^*, \hat{x} \rangle &= \left\langle \lambda_{1,\alpha}(y_{1,\alpha}^* + u_{1,\alpha}^*), \hat{x} \right\rangle + \cdots + \left\langle \lambda_{m,\alpha}(y_{m,\alpha}^* + u_{m,\alpha}^*), \hat{x} \right\rangle \\
&\quad + \left\langle \lambda_{m+1,\alpha} y_{m+1,\alpha}^*, \hat{x} \right\rangle + \langle z_\alpha^*, \hat{x} \rangle \\
&\leq \left\langle \lambda_{1,\alpha} y_{1,\alpha}^*, \hat{x} \right\rangle + \cdots + \left\langle \lambda_{m,\alpha} y_{m,\alpha}^*, \hat{x} \right\rangle \\
&\quad + \left\langle \lambda_{m+1,\alpha} y_{m+1,\alpha}^*, \hat{x} \right\rangle + \langle z_\alpha^*, \hat{x} \rangle,
\end{aligned}$$

we infer that the nets $(\langle \lambda_{k,\alpha} y_{k,\alpha}^*, \hat{x} \rangle)_\alpha$, $k \in \{1, \ldots, m+1\}$, are bounded.

Now, for each $k \in \{1, \ldots, m\}$, by (55),

$$\left\langle \lambda_{k,\alpha} y_{k,\alpha}^*, \hat{x} + w \right\rangle \leq \lambda_{k,\alpha} M \leq M \text{ for all } w \in W \text{ and } \alpha,$$

and taking the boundedness of the net $(\langle \lambda_{k,\alpha} y_{k,\alpha}^*, \hat{x} \rangle)_\alpha$ into account, we deduce that $(\lambda_{k,\alpha} y_{k,\alpha}^*)_\alpha \subset rW^\circ$ for some $r \geq 0$. Consequently, by the Alaoglu–Bourbaki theorem we may suppose, without loss of generality, that $(\lambda_{k,\alpha} y_{k,\alpha}^*)_\alpha$ weak*-converges to some $\ell_k^* \in X^*$. Due to (54), we deduce that the net $(v_\alpha^*)_\alpha$, defined as

$$v_\alpha^* := \lambda_{1,\alpha} u_{1,\alpha}^* + \cdots + \lambda_{m,\alpha} u_{m,\alpha}^* + \lambda_{m+1,\alpha} y_{m+1,\alpha}^*, \tag{56}$$

also weak*-converges to some $\ell_{m+1}^* \in X^*$. More specifically, if $k \in \{1, \ldots, m\}$ is such that $\lambda_k > 0$, then

$$\ell_k^* \in \lambda_k \partial f_k(\theta), \tag{57}$$

while for the other case, when $\lambda_k = 0$, by taking the limit on α in the inequality

$$\left\langle \lambda_{k,\alpha} y_{k,\alpha}^*, z \right\rangle \leq \lambda_{k,\alpha} f_k(z) \leq \lambda_{k,\alpha} f(z), \ z \in \operatorname{dom} f,$$

we observe that

$$\ell_k^* \in \mathrm{N}_{\operatorname{dom} f}(\theta). \tag{58}$$

Let us analyze the behavior of the net $(v_\alpha^*)_\alpha$ defined in (56), which has already been proved to converge to ℓ_{m+1}^*. Take $z \in \operatorname{dom} f$, $L_0 := \operatorname{span}\{\hat{x}, z\}$ $(\in \widetilde{\mathcal{F}}(\theta))$, and $\alpha_0 := (L_0, X^*)$, so that $z \in L \cap \operatorname{dom} f$ for all $\alpha = (L, V) \succeq \alpha_0$. Fix $\alpha \succeq \alpha_0$. Since, by definition, $u_{k,\alpha}^* \in \mathrm{N}_{L \cap \operatorname{dom} f}(\theta)$, $k = 1, \ldots, m$, and $y_{m+1,\alpha}^* \in \partial(f_{m+1} + \mathrm{I}_{L \cap \operatorname{dom} f})(\theta)$, we obtain that

$$\begin{aligned}
\langle v_\alpha^*, z \rangle &= \left\langle \lambda_{1,\alpha} u_{1,\alpha}^* + \cdots + \lambda_{m,\alpha} u_{m,\alpha}^* + \lambda_{m+1,\alpha} y_{m+1,\alpha}^*, z \right\rangle \\
&\leq \lambda_{m+1,\alpha} \left\langle y_{m+1,\alpha}^*, z \right\rangle \\
&\leq \lambda_{m+1,\alpha}(f_{m+1}(z) - f_{m+1}(\theta)) \\
&= \lambda_{m+1,\alpha} f_{m+1}(z),
\end{aligned}$$

which, by taking limits, gives

$$\left\langle \ell_{m+1}^{*}, z\right\rangle \leq \lambda_{m+1} f_{m+1}(z);$$

that is, as z is an arbitrary point in $\operatorname{dom} f$ and $f_{m+1}(z) \leq f(z) < +\infty$,

$$\ell_{m+1}^{*} \in \begin{cases} \mathrm{N}_{\operatorname{dom} f}(\theta) & \text{when } \lambda_{m+1}=0, \\ \lambda_{m+1}\partial(f_{m+1}+\mathrm{I}_{\operatorname{dom} f})(\theta) & \text{when } \lambda_{m+1}>0. \end{cases} \tag{59}$$

We proceed by defining $T_{+}(\theta) := \{k=1,\ldots,m \mid \lambda_k > 0\}$. Then in virtue of (54) we get

$$\begin{aligned} x^{*} &= \lim_{\alpha}\left(\sum_{k=1,\ldots,m} \lambda_{k,\alpha} y_{k,\alpha}^{*} + v_{\alpha}^{*}\right) \\ &= \lim_{\alpha} \sum_{k\in T_{+}(\theta)} \lambda_{k,\alpha} y_{k,\alpha}^{*} + \lim_{\alpha} \sum_{k\in\{1,\ldots,m\}\setminus T_{+}(\theta)} \lambda_{k,\alpha} y_{k,\alpha}^{*} + \lim_{\alpha} v_{\alpha}^{*} \\ &\subset \sum_{k\in T_{+}(\theta)} \lambda_k \partial f_k(\theta) + \mathrm{N}_{\operatorname{dom} f}(\theta) + \ell_{m+1}^{*}. \end{aligned} \tag{60}$$

At this step, and in order to specify the nature of ℓ_{m+1}^{*}, we distinguish two cases.

(a) If $\lambda_{m+1} > 0$, then by (59) we have $\ell_{m+1}^{*} \in \lambda_{m+1}\partial(f_{m+1}+\mathrm{I}_{\operatorname{dom} f})(\theta)$, so that (60) gives us

$$\begin{aligned} x^{*} &\in \sum_{k\in T_{+}(\theta)} \lambda_k \partial f_k(\theta) + \mathrm{N}_{\operatorname{dom} f}(\theta) + \lambda_{m+1}\partial(f_{m+1}+\mathrm{I}_{\operatorname{dom} f})(\theta) \\ &\subset \operatorname{co}\left\{\bigcup_{k=1,\ldots,m} \partial f_k(\theta) \bigcup \partial(f_{m+1}+\mathrm{I}_{\operatorname{dom} f})(\theta)\right\} + \mathrm{N}_{\operatorname{dom} f}(\theta). \end{aligned}$$

(b) Otherwise, if $\lambda_{m+1} = 0$, then by (59) we have $\ell_{m+1}^{*} \in \mathrm{N}_{\operatorname{dom} f}(\theta)$, so that (60) yields

$$\begin{aligned} x^{*} &\in \sum_{k\in T_{+}(\theta)} \lambda_k \partial f_k(\theta) + \mathrm{N}_{\operatorname{dom} f}(\theta) + \ell_{m+1}^{*} \\ &\subset \operatorname{co}\left\{\bigcup_{k=1,\ldots,m} \partial f_k(\theta)\right\} + \mathrm{N}_{\operatorname{dom} f}(\theta) \\ &\subset \operatorname{co}\left\{\bigcup_{k=1,\ldots,m} \partial f_k(\theta) \bigcup \partial(f_{m+1}+\mathrm{I}_{\operatorname{dom} f})(\theta)\right\} + \mathrm{N}_{\operatorname{dom} f}(\theta). \end{aligned}$$

The proof is finished. □

Remark 2. If in Theorem 9 each one of the functions f_k, $k \in T(x)$, is continuous at some point of $\operatorname{dom} f$, then

$$\partial(f_{k_0}+\mathrm{I}_{\operatorname{dom} f})(x) = \partial f_{k_0}(x) + \mathrm{N}_{\operatorname{dom} f}(x),$$

and so, due to the Moreau–Rockafellar sum rule for the subdifferentials,

$$\mathrm{N}_{\operatorname{dom} f}(x) = \sum_{k\in T(x)} \mathrm{N}_{\operatorname{dom} f_k}(x) + \mathrm{N}_{\cap_{k\in T\setminus T(x)} \operatorname{dom} f_k}(x).$$

Thus, Theorem 9 gives

$$\partial f(x) = \operatorname{co}\left\{ \bigcup_{k\in T(x)} \partial f_k(x) \right\} + \sum_{k\in T(x)} \mathrm{N}_{\operatorname{dom} f_k}(x) + \mathrm{N}_{\cap_{k\in T\setminus T(x)} \operatorname{dom} f_k}(x).$$

As a consequence of the previous theorem we obtain the following result given in [30].

Corollary 10. *Assume that all the functions f_k, $k \in \{1,\dots,p\}$, except perhaps one of them, f_{k_0}, are continuous at some point in* $\operatorname{dom} f$. *Then for all $x \in X$,*

$$\partial f(x) = \operatorname{co}\left\{ \bigcup_{k\in T(x)} \partial f_k(x) \right\} + \sum_{k\in T} \mathrm{N}_{\operatorname{dom} f_k}(x). \tag{61}$$

Proof. Fix $x \in X$. First, we observe that the proper functions $\mathrm{I}_{\operatorname{dom} f_k}$, $k \in T \setminus \{k_0\}$, are continuous at $x_0 \in \operatorname{dom} f \subset \operatorname{dom} f_{k_0} = \operatorname{dom}(\mathrm{I}_{\operatorname{dom} f_{k_0}})$. So, by the Moreau–Rockafellar subdifferential sum rule we have that

$$\mathrm{N}_{\operatorname{dom} f}(x) = \sum_{k\in T} \mathrm{N}_{\operatorname{dom} f_k}(x). \tag{62}$$

First, if $k_0 \notin T(x)$, (51) and (62) yield (61). If $k_0 \in T(x)$, we write

$$\begin{aligned} \partial(f_{k_0} + \mathrm{I}_{\operatorname{dom} f})(x) &= \partial\Bigg(f_{k_0} + \sum_{k\in T} \mathrm{I}_{\operatorname{dom} f_k} \Bigg)(x) \\ &= \partial\Bigg(f_{k_0} + \sum_{k\in T\setminus k_0} \mathrm{I}_{\operatorname{dom} f_k} \Bigg)(x) \\ &= \partial f_{k_0}(x) + \sum_{k\in T\setminus k_0} \mathrm{N}_{\operatorname{dom} f_k}(x). \end{aligned} \tag{63}$$

Therefore, $\partial(f_{k_0} + \mathrm{I}_{\operatorname{dom} f})(x) = \emptyset$ if $\partial f_{k_0}(x) = \emptyset$, and again (51) and (62) provide (61).

Finally, we analyze the case in which $\partial f_{k_0}(x) \neq \emptyset$. The fact that $\partial f_{k_0}(x) = \partial f_{k_0}(x) + \mathrm{N}_{\operatorname{dom} f_{k_0}}(x)$, together with (63), gives rise to

$$\begin{aligned} \partial(f_{k_0} + \mathrm{I}_{\operatorname{dom} f})(x) &= \partial f_{k_0}(x) + \sum_{k\in T\setminus k_0} \mathrm{N}_{\operatorname{dom} f_k}(x) \\ &= \partial f_{k_0}(x) + \sum_{k\in T} \mathrm{N}_{\operatorname{dom} f_k}(x) \\ &= \partial f_{k_0}(x) + \mathrm{N}_{\operatorname{dom} f}(x). \end{aligned}$$

Next, by Theorem 9 we obtain that

$$\begin{aligned}\partial f(x) &= \operatorname{co}\left\{\bigcup_{k\in T(x)\setminus\{k_0\}} \partial f_k(x) \bigcup \partial(f_{k_0}+\mathrm{I}_{\operatorname{dom} f})(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x)\\
&= \operatorname{co}\left\{\bigcup_{k\in T(x)\setminus\{k_0\}} \partial f_k(x) \bigcup (\partial f_{k_0}(x)+\mathrm{N}_{\operatorname{dom} f}(x))\right\} + \mathrm{N}_{\operatorname{dom} f}(x)\\
&\subset \operatorname{co}\left\{\bigcup_{k\in T(x)\setminus\{k_0\}} (\partial f_k(x)+\mathrm{N}_{\operatorname{dom} f}(x)) \bigcup (\partial f_{k_0}(x)+\mathrm{N}_{\operatorname{dom} f}(x))\right\} + \mathrm{N}_{\operatorname{dom} f}(x)\\
&= \operatorname{co}\left\{\bigcup_{k\in T(x)} (\partial f_k(x)+\mathrm{N}_{\operatorname{dom} f}(x))\right\} + \mathrm{N}_{\operatorname{dom} f}(x)\\
&= \operatorname{co}\left\{\bigcup_{k\in T(x)} \partial f_k(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x).\end{aligned}$$

Since the reverse of the last inclusion always holds, we deduce that

$$\partial f(x) = \operatorname{co}\left\{\bigcup_{k\in T(x)} \partial f_k(x)\right\} + \mathrm{N}_{\operatorname{dom} f}(x),$$

and, finally, the conclusion of the corollary follows due to (62). ☐

The previous corollary leads to the following formula given in [25, Theorem 4], when $T(x)=T$ and $\partial f_k(x)\neq\emptyset$ for all $k\in T$:

$$\partial f(x) = \operatorname{co}\left\{\bigcup_{k\in T(x)} \partial f_k(x)\right\}.$$

Remark 3. The continuity condition of Corollary 10 implies (4), as established in [12, Corollary 9(iii)]. Thus, removing (4) within the subdifferential calculus of section 3 allowed us to obtain Theorem 8 without requiring any lower semicontinuity-like assumption on the functions.

Acknowledgment. We are thankful to both referees for their insightful suggestions and comments.

REFERENCES

[1] A. Brandsted, *On the subdifferential of the supremum of two convex functions*, Math. Scand., 31 (1972), pp. 225–230.
[2] N. H. Chieu, V. Jeyakumar, G. Li, and H. Mohebi, *Constraint qualifications for convex optimization without convexity of constraints: New connections and applications to best approximation*, European J. Oper. Res., 265 (2018), pp. 19–25.
[3] R. Correa, A. Hantoute, and M. A. López, *Weaker conditions for subdifferential calculus of convex functions*, J. Funct. Anal., 271 (2016), pp. 1177–1212.
[4] R. Correa, A. Hantoute, and M. A. López, *Towards supremum-sum subdifferential calculus free of qualification conditions*, SIAM J. Optim., 26 (2016), pp. 2219–2234.
[5] R. Correa, A. Hantoute, and M. A. López, *Valadier-like formulas for the supremum function* I, J. Convex Anal., 25 (2018), pp. 1253–1278.

[6] R. Correa, A. Hantoute, and M. A. López *Valadier-like formulas for the supremum function II: The compactly indexed case*, J. Convex Anal., 26 (2019), pp. 299–324.
[7] F. Deutsch, W. Li, and J. Ward, *A dual approach to constrained interpolation from a convex subset of Hilbert space*, J. Approx. Theory, 90 (1997), pp. 385–414.
[8] N. Dinh, M. A. Goberna, M. A. López, and T. Q. Son, *New Farkas-type constraint qualifications in convex infinite programming*, ESAIM Control Optim. Calc. Var., 13 (2007), pp. 580–597.
[9] J. Dutta and C. S. Lalitha, *Optimality conditions in convex optimization revisited*, Optim. Lett., 7 (2013), pp. 221–229.
[10] D. H. Fang, C. Li, and K. F. Ng, *Constraint qualifications for optimality conditions and total Lagrange dualities in convex infinite programming*, Nonlinear Anal., 73 (2010), pp. 1143–1159.
[11] M. A. Goberna and M.A. López, *Linear Semi-Infinite Optimization*, John Wiley, Chichester, 1998.
[12] A. Hantoute, M. A. López, and C. Zălinescu, *Subdifferential calculus rules in convex analysis: A unifying approach via pointwise supremum functions*, SIAM J. Optim., 19 (2008), pp. 863–882.
[13] J.-B. Hiriart-Urruty and R. R. Phelps, *Subdifferential calculus using ε-subdifferentials*, J. Funct. Anal., 118 (1993), pp. 154–166.
[14] A. D. Ioffe, *A note on subdifferentials of pointwise suprema*, TOP, 20 (2012), pp. 456–466.
[15] P.-J. Laurent, *Approximation et Optimisation*, Hermann, Paris, 1972.
[16] C. Li and K. F. Ng, *Subdifferential calculus rules for supremum functions in convex analysis*, SIAM J. Optim., 21 (2011), pp. 782–797.
[17] C. Li, K. F. Ng, and T. K. Pong, *The SECQ, linear regularity, and the strong CHIP for an infinite system of closed convex sets in normed linear spaces*, SIAM J. Optim., 18 (2007), pp. 643–665.
[18] C. Li, K. F. Ng, and T. K. Pong, *Constraint qualifications for convex inequality systems with applications in constrained optimization*, SIAM J. Optim., 19 (2008), pp. 163–187.
[19] Y. Liu and M. A. Goberna, *Asymptotic optimality conditions for linear semi-infinite programming*, Optimization, 65 (2016), pp. 387–414.
[20] O. López and L. Thibault, *Sequential formula for subdifferential of upper envelope of convex functions*, J. Nonlinear Convex Anal., 14 (2013), pp. 377–388.
[21] M. A. López and E. Vercher, *Optimality conditions for nondifferentiable convex semi-infinite programming*, Math. Program., 27 (1983), pp. 307–319.
[22] B. Mordukhovich and T. T. A. Nghia, *Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data*, SIAM J. Optim., 23 (2013), pp. 406–431.
[23] J. J. Moreau, *Fonctionnelles convexes*, Séminaire Jean Leray (1966-1967), Collège de France, Paris, 1966.
[24] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd ed., Springer, Berlin, 1993.
[25] R. T. Rockafellar, *Directionally Lipschitzian functions and subdifferential calculus*, Proc. Lond. Math. Soc. (3), 39 (1979), pp. 331–355.
[26] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[27] V. N. Solov'ev, *The subdifferential and the directional derivatives of the maximum of a family of convex functions*, Izv. Ross. Akad. Nauk Ser. Mat., 65 (2001), pp. 107–132.
[28] L. Thibault, *Sequential convex subdifferential calculus and sequential Lagrange multipliers*, SIAM J. Control Optim., 35 (1997), pp. 1434–1444.
[29] M. Valadier, *Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes*, C. R. Math. Acad. Sci. Paris, 268 (1969), pp. 39–42.
[30] M. Volle, *On the subdifferential of an upper envelope of convex functions*, Acta Math. Vietnam., 19 (1994), pp. 137–148.
[31] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, River Edge, NJ, 2002.
[32] X. Y. Zheng and K. F. Ng, *Subsmooth semi-infinite and infinite optimization problems*, Math. Program., 134 (2012), pp. 365–393.