MOREAU ENVELOPE OF SUPREMUM FUNCTIONS WITH APPLICATIONS TO INFINITE AND STOCHASTIC PROGRAMMING*

PEDRO PÉREZ-AROS† AND EMILIO VILCHES‡

Abstract. In this paper, we investigate the Moreau envelope of the supremum of a family of convex, proper, and lower semicontinuous functions. Under mild assumptions, we prove that the Moreau envelope of a supremum is the supremum of Moreau envelopes, which allows us to approximate possibly nonsmooth supremum functions by smooth functions that are also the suprema of functions. Consequently, we propose and study approximated optimization problems from infinite and stochastic programming for which we obtain zero-duality results and optimality conditions without the verification of constraint qualification conditions.

Key words. Moreau envelope, subdifferential calculus, supremum function, infinite programming, semi-infinite programming, stochastic programming.

AMS subject classifications. 49J53; 90C15; 90C34; 90C25

1. Introduction.

The Moreau envelope (also called Moreau regularization) is ubiquitous in optimization, convex analysis, and variational analysis. It appears as a natural way to regularize a convex function through an associated optimization problem. Since its introduction by J.J. Moreau in [26], it has allowed both theoretical and practical development of the areas mentioned above (see [1, 3, 21, 22, 29] and the references therein). Furthermore, its study is strongly linked to the development of proximal algorithms to solve and approximate optimization problems.

Supremum functions play a fundamental role in diverse areas such as optimization, convex analysis, and variational analysis. For instance, in optimization, they enable to model robust objective or constraints functions for real-world applications. Hence, several authors have developed subdifferential calculus and optimality conditions for general optimization problems using such a class of mappings. We refer to [9,17,18,27] and the references therein, where it is shown that the supremum functions can be used as a universal approach to study the whole theory of convex subdifferential calculus. It is also important to mention that supremum functions are also used in nonconvex analysis (see, e.g., [24,28,30]). Despite the significant theoretical advances in the study of this type of functions, from a practical point of view, it is not easy to deal with the nonsmoothness of supremum functions. Thus, considering appropriate regularizations could be a good alternative for applications.

In this paper, we study the Moreau envelope of supremum functions. We prove that, under concave-convex assumptions on the family of functions, the Moreau envelope of a supremum is the supremum of Moreau envelopes, which allows us to approximate possibly nonsmooth supremum functions by smooth functions that are also the suprema of functions. Consequently, we obtain optimality conditions and zero-

* **Funding:** E. Vilches was partially funded by ANID Chile under grants Fondecyt de Iniciación N° 11180098 and Fondecyt regular N° 1200283. P. Pérez-Aros was partially supported by ANID Chile under grants Fondecyt regular N° 1190110, Fondecyt regular N° 1200283 and Programa Regional Mathamsud 20-Math-08 CODE: MATH190003.

†Instituto de Ciencias de la Ingeniería, Universidad de O'Higgins, Rancagua, Chile (pedro.perez@uoh.cl)

‡Instituto de Ciencias de la Ingeniería, Universidad de O'Higgins, Rancagua, Chile (emilio.vilches@uoh.cl)

duality results for infinite and stochastic programming of approximated optimization problems, where we replace the original functions by its Moreau envelope. The obtained results do not require constraint qualification conditions, which, in general, are difficult to verify.

The paper is organized as follows. After some preliminaries, in Section 3, by using an inf-sup theorem due to J.J. Moreau [25], we show that the Moreau envelope of a supremum is the supremum of Moreau envelopes, and we provide a formula for its gradient. In Section 4, we obtain optimality conditions and zero-duality gap results for infinite programming problems. In Section 5, we furnish optimality conditions for some stochastic optimization problems. The paper ends with some final comments.

2. Preliminaries.

2.1. Basic tools from convex analysis. Throughout this paper, $\mathcal{H}$ is a Hilbert space endowed with an inner product $\langle\cdot,\cdot\rangle$. The weak topology in $\mathcal{H}$ is denoted by w and its respective convergence by $\rightharpoonup$. For a set $A\subset\mathcal{H}$, we denote by $\operatorname{co}(A)$ the *convex hull* generated by A. The *indicator* and the *support function* of A are defined, respectively, by

$$\delta_A(x)=\begin{cases}0 & \text{if } x\in A,\\ +\infty & \text{if } x\notin A,\end{cases}\qquad \sigma_A(x)=\sup_{y\in A}\langle x,y\rangle.$$

The *distance function* from x to A is given by $\operatorname{dist}(x,A)=\inf\{\|x-y\|:y\in A\}$. Moreover, if A is nonempty, closed and convex, the *projection* of $x\in\mathcal{H}$ onto A is denoted by $\operatorname{proj}_A(x)$.

Given a function $g:\mathcal{H}\to\mathbb{R}_\infty:=\mathbb{R}\cup\{+\infty\}$, the *domain* and the *epigraph* of g are

$$\operatorname{dom} g:=\{x\in\mathcal{H}:\ g(x)<+\infty\}\text{ and }\operatorname{epi} g:=\{(x,\alpha)\in\mathcal{H}\times\mathbb{R}:\ g(x)\le\alpha\},$$

respectively. We say that g is *proper* if $\operatorname{dom} g\neq\emptyset$. The class of all proper, convex, and lower semicontinuous functions is denoted by $\Gamma_0(\mathcal{H})$.

The Fenchel conjugate function of g is given by

$$g^*(x^*):=\sup\{\langle x^*,z\rangle-g(z):z\in\mathcal{H}\}\text{ for all } x^*\in\mathcal{H},$$

and the biconjugate is given by $g^{**}=(g^*)^*$.

The (*Moreau-Rockafellar*) *convex subdifferential* of g at a point $x\in\mathcal{H}$, where g is finite, is the set

$$\partial g(x):=\{x^*\in\mathcal{H}:\ \langle x^*,x\rangle=g^*(x^*)+g(x)\},$$

otherwise, we set $\partial g(x)=\emptyset$.

The next lemma will be used through the paper.

LEMMA 2.1. *Let $(g_c)_{c\in\mathcal{C}}\subset\Gamma_0(\mathcal{H})$ be a family of functions indexed by a convex set $\mathcal{C}$. If the map $c\to g_c(x)$ is concave for all $x\in\mathcal{H}$, then*

a) The map $\mathcal{C}\times\mathcal{H}\ni(c,x^*)\to g_c^*(x^*)$ is convex.

b) In addition, if $(\mathcal{C},\tau)$ is topological space and the map $c\to g_c(x)$ is τ-upper semicontinuous for all $x\in\mathcal{H}$, then the function $(c,x^*)\to g_c^*(x^*)$ is $\tau\times w$-lower semicontinous.

Proof. a) For $(c_1,x_1^*),(c_2,x_2^*)\in\mathcal{C}\times\mathcal{H}$ and $\alpha\in[0,1]$ fixed, let us consider the convex combination

$$(c_\alpha,x_\alpha^*):=\alpha(c_1,x_1^*)+(1-\alpha)(c_2,x_2^*).$$

Then, by the concavity of $c \mapsto g_c(w)$ for $w \in \mathcal{H}$ arbitrary,

$$g_{c_\alpha}(w) \geq \alpha g_{c_1}(w) + (1-\alpha) g_{c_2}(w).$$

Hence, by the latter inequality and the definition of the convex conjugate,

$$\begin{aligned}\langle w, x_\alpha^* \rangle - g_{c_\alpha}(w) &\leq \alpha \langle w, x_1^* \rangle + (1-\alpha)\langle w, x_2^* \rangle - \alpha g_{c_1}(w) - (1-\alpha) g_{c_2}(w) \\ &\leq \alpha \left(\langle w, x_1^* \rangle - g_{c_1}(w) \right) + (1-\alpha) \left(\langle w, x_2^* \rangle - (1-\alpha) g_{c_2}(w) \right) \\ &\leq \alpha g_{c_1}^*(x_1^*) + (1-\alpha) g_{c_2}^*(x_2^*).\end{aligned}$$

Since $w \in \mathcal{H}$ is arbitrary, we get that

$$g_{c_\alpha}^*(x_\alpha^*) \leq \alpha g_{c_1}^*(x_1^*) + (1-\alpha) g_{c_2}^*(x_2^*),$$

which proves the convexity of the map $(c, x^*) \to g_c^*(x^*)$.
b) By virtue of the upper semicontinuity of $c \mapsto g_c(w)$ for all $w \in \mathcal{H}$, the function

$$(c, x^*) \to \langle w, x^* \rangle - g_c(w)$$

is $\tau \times w$-lower semicontinuous for all $w \in \mathcal{H}$. Hence, since the conjugate is the supremum of $\tau \times w$-lower semicontinuous functions, the above function is also $\tau \times w$-lower semicontinuous, which proves b). ◻

Let $g\colon \mathcal{H} \to \mathbb{R}_\infty$ be a function in $\Gamma_0(\mathcal{H})$. For any $\lambda > 0$, the Moreau envelope of g of index λ is the function $\mathrm{e}_\lambda g\colon \mathcal{H} \to \mathbb{R}$, defined by

$$\mathrm{e}_\lambda g(x) := \inf_{y \in \mathcal{H}} \left(g(y) + \frac{1}{2\lambda} \|x - y\|^2 \right) \text{ for all } x \in \mathcal{H}.$$

The latter infimum is attained at a unique point $\mathrm{P}_\lambda g(x) \in \mathcal{H}$, which satisfies

$$\mathrm{e}_\lambda g(x) = g(\mathrm{P}_\lambda g(x)) + \frac{1}{2\lambda} \|x - \mathrm{P}_\lambda g(x)\|^2 \text{ and } \mathrm{P}_\lambda g(x) + \lambda \partial g(\mathrm{P}_\lambda g(x)) \ni x.$$

Moreover, the operator $\mathrm{P}_\lambda g(x) = (I + \lambda \partial g)^{-1}\colon \mathcal{H} \to \mathcal{H}$ is everywhere defined and nonexpansive. It is called the resolvent of index λ of $A = \partial g$. The Moreau envelope $\mathrm{e}_\lambda g$ is convex and continuously differentiable. For each $x \in \mathcal{H}$, its gradient is given by

$$\nabla \mathrm{e}_\lambda g(x) = \frac{1}{\lambda}(x - \mathrm{P}_\lambda g(x)). \tag{2.1}$$

The operator $A_\lambda = \frac{1}{\lambda}(I - \mathrm{P}_\lambda g)$ is called the Yosida approximation of index λ of the maximal monotone operator $A = \partial g$. This operator is Lipschitz continuous of constant $1/\lambda$ and satisfies

$$A_\lambda x = \nabla \mathrm{e}_\lambda g(x) \text{ and } A_\lambda(x) \in \partial g(\mathrm{P}_\lambda g(x)) \text{ for all } x \in \mathcal{H}.$$

The following proposition summarizes the main properties of the Moreau envelope in Hilbert spaces. We refer to [1,2,4] for more details.

Proposition 2.2. *Let $g\colon \mathcal{H} \to \mathbb{R}_\infty$ be a function in $\Gamma_0(\mathcal{H})$. Then the following hold.*

1. *Monotone convergence:* $\mathrm{e}_\lambda g(x) \nearrow g(x)$ *as* $\lambda \searrow 0$ *for all* $x \in \mathcal{H}$.
2. *Convergence of resolvents:* $\mathrm{P}_\lambda g(x) \to x$ *as* $\lambda \to 0$ *for all* $x \in \overline{\operatorname{dom} g}$ *and* $\mathrm{P}_\lambda g \to \operatorname{proj}_{\overline{\operatorname{dom} g}}(x)$ *as* $\lambda \to 0$ *for all* $x \in \mathcal{H}$.
3. *Conjugate:* $(\mathrm{e}_\lambda g)^*(x^*) = g^*(x^*) + \frac{\lambda}{2}\|x^*\|^2$ *for all* $x^* \in \mathcal{H}$.
4. *Lower epi-convergence: If* $x_k \rightharpoonup x$ *and* $\lambda_k \searrow 0$*, then* $g(x) \leq \liminf_{k \to \infty} \mathrm{e}_{\lambda_k} g(x_k)$.

2.2. Inf-sup Theorem. One of the main tools used in this article is the following inf-sup theorem. Its proof is based on the inf-sup theorem proved by J.J. Moreau in [25]. We refer to [34] for historical perspectives and other variants of inf-sup theorems.

PROPOSITION 2.3. *Let $\mathcal{C}$ be a convex set and let $g : \mathcal{C} \times \mathcal{H} \to \mathbb{R}_\infty$ be a function such that*

- *The function $x \mapsto g(c, x)$ belongs to $\Gamma_0(\mathcal{H})$ for all $c \in \mathcal{C}$.*
- *The function $c \mapsto g(c, x)$ is concave for all $x \in \mathcal{H}$.*
- *There exists $c_0 \in \mathcal{C}$ and $k_0 > \inf_{x\in\mathcal{H}} g(c_0, x)$ such that $\{x \in \mathcal{H} : g(c_0, x) \leq k_0\}$ is bounded.*

Then, $\inf\limits_{x\in\mathcal{H}} \sup\limits_{c\in\mathcal{C}} g(c, x) = \sup\limits_{c\in\mathcal{C}} \inf\limits_{x\in\mathcal{H}} g(c, x)$.

Proof. Let $\mathcal{F}$ be the vector space spanned by $\mathcal{C}$. We extend the function g to the whole space $\mathcal{F} \times \mathcal{H}$ by

$$\hat{g}(c, x) := \begin{cases} g(c, x) & \text{if } (c, x) \in \mathcal{C} \times \mathcal{H} \\ -\infty & \text{if } (c, x) \notin \mathcal{C} \times \mathcal{H}. \end{cases}$$

Since $\hat{g}$ satisfies the assumptions of [25, Proposition 2], we obtain that

$$\inf_{x\in\mathcal{H}} \sup_{c\in\mathcal{F}} \hat{g}(c, x) = \sup_{c\in\mathcal{F}} \inf_{x\in\mathcal{H}} \hat{g}(c, x).$$

Therefore,

$$\inf_{x\in\mathcal{H}} \sup_{c\in\mathcal{C}} g(c, x) = \inf_{x\in\mathcal{H}} \sup_{c\in\mathcal{F}} \hat{g}(c, x) = \sup_{c\in\mathcal{F}} \inf_{x\in\mathcal{H}} \hat{g}(c, x) = \sup_{c\in\mathcal{C}} \inf_{x\in\mathcal{H}} g(c, x),$$

which proves the result. ∎

2.3. Perturbation function and duality. Let $\mathcal{C}$ be a family of indexes. Following [14, 15], let $\mathcal{X}_c$ be a family of Banach spaces indexed by $c \in \mathcal{C}$. We denote by 0_c the zero vector in $\mathcal{X}_c$. Consider for each $c \in \mathcal{C}$ a perturbation function $\Phi_c : \mathcal{H} \times \mathcal{X}_c \to \mathbb{R}_\infty$. For each $x^* \in \mathcal{H}$, the primal problem is given by

$$(\mathrm{P}_{x^*}) \qquad \inf_{x\in\mathcal{H}} \sup_{c\in\mathcal{C}} \left(\Phi_c(x, 0_c) - \langle x^*, x\rangle\right),$$

and its corresponding dual is given by

$$(\mathrm{D}_{x^*}) \qquad \sup_{c\in\mathcal{C}, x_c^*\in\mathcal{X}_c^*} -\Phi_c^*(x^*, x_c^*).$$

The following result establishes a zero duality gap for convex perturbation functions (see [14, Proposition 3.2] and [15, 23]).

PROPOSITION 2.4. *Let $\Phi_c \in \Gamma_0(\mathcal{H} \times \mathcal{X}_c)$ for $c \in \mathcal{C}$ such that $\sup_{c\in\mathcal{C}} \Phi_c(\cdot, 0)$ is proper. Then, the following statements are equivalent:*

a) $\inf(\mathrm{P}_{x^*}) = \max(\mathrm{D}_{x^*})$.

b) The set $\mathcal{W} = \bigcup\limits_{c\in\mathcal{C}} \{(x^*, \alpha) \in \mathcal{H} \times \mathbb{R} : \exists x_c^* \in \mathcal{X}_c^*, \text{ s.t. } \Phi_c^*(x^*, x_c^*) \leq \alpha\}$ *is closed and convex.*

3. Moreau envelope of the supremum function. In this section, we study the Moreau envelope of the supremum function

$$f(x) := \sup_{c\in\mathcal{C}} f_c(x) \text{ for all } x \in \mathcal{H}, \tag{3.1}$$

where $(f_c)_{c\in\mathcal{C}}$ is a family of functions in $\Gamma_0(\mathcal{H})$ indexed by a convex set $\mathcal{C}\neq\emptyset$.

The following structural result, whose proof is based on Proposition 2.3, establishes that the Moreau envelope of a supremum is the supremum of Moreau envelopes under the concavity of the maps $c\mapsto f_c(x)$ for all $x\in\mathcal{H}$.

THEOREM 3.1. *Let $\mathcal{C}\neq\emptyset$ be a convex set and $(f_c)_{c\in\mathcal{C}}$ be a family of functions in $\Gamma_0(\mathcal{H})$ such the supremum function $f:=\sup_{c\in\mathcal{C}} f_c$ is proper and the function $c\mapsto f_c(x)$ is concave for all $x\in\mathcal{H}$. Then, for all $\lambda>0$*

$$\mathrm{e}_\lambda f(x)=\sup_{c\in\mathcal{C}}\mathrm{e}_\lambda f_c(x)\quad \text{for all } x\in\mathcal{H}. \tag{3.2}$$

Moreover, if $\mathcal{C}$ is compact and the function $c\mapsto f_c(x)$ is upper semicontinuous for all $x\in\mathcal{H}$, then the supremum in (3.2) is attained.

Proof. Fix $\lambda>0$ and $x\in\mathcal{H}$. Let us consider the function $g:\mathcal{C}\times\mathcal{H}\to\mathbb{R}_\infty$ defined by

$$g(c,y):=f_c(y)+\frac{1}{2\lambda}\|y-x\|^2.$$

It is clear that $\sup_{c\in\mathcal{C}} g(c,y)=f(y)+\frac{1}{2\lambda}\|y-x\|^2$. Also, for any $c_0\in\mathcal{C}$ fixed, the function f_{c_0} belongs to $\Gamma_0(\mathcal{H})$, thus, it has an affine minorant. Then, the set $\{y\in\mathcal{H}: g(c_0,y)\le\alpha\}$ is bounded for all $\alpha\in\mathbb{R}$. Hence, by virtue of Proposition 2.3, the following equality holds:

$$\sup_{c\in\mathcal{C}}\inf_{y\in\mathcal{H}} g(c,y)=\inf_{y\in\mathcal{H}}\sup_{c\in\mathcal{C}} g(c,y).$$

Therefore, for all $x\in\mathcal{H}$

$$\begin{aligned}\sup_{c\in\mathcal{C}}\mathrm{e}_\lambda f_c(x)&=\sup_{c\in\mathcal{C}}\inf_{y\in H}\left(f_c(y)+\frac{1}{2\lambda}\|y-x\|^2\right)\\&=\sup_{c\in\mathcal{C}}\inf_{y\in\mathcal{H}} g(c,y)\\&=\inf_{y\in\mathcal{H}}\sup_{c\in\mathcal{C}} g(c,y)\\&=\inf_{y\in\mathcal{H}}\left(f(y)+\frac{1}{2\lambda}\|y-x\|^2\right)\\&=\mathrm{e}_\lambda f(x),\end{aligned}$$

which ends the proof. ∎

The next result provides a framework where the concave assumption on Theorem 3.1 automatically holds. For a given set $T\neq\emptyset$, we define the T-dimensional simplex $\Delta(T)$ as the set of all functions $(\alpha_t)\in\mathbb{R}_+^T$ such that $\alpha_t\neq 0$ for finitely many $t\in T$, and $\sum\alpha_t=1$.

COROLLARY 3.2. *Let T be a nonempty set and $(g_t)_{t\in T}$ be a family of functions on $\Gamma_0(\mathcal{H})$ such that $g:=\sup_{t\in T} g_t$ is proper. Then, for all $\lambda>0$*

$$\mathrm{e}_\lambda g(x)=\sup_{(\alpha_t)\in\Delta(T)}\mathrm{e}_\lambda\left(\sum_{t\in T}\alpha_t g_t\right)(x)\quad \text{for all } x\in\mathcal{H}.$$

Proof. Consider $\mathcal{C} := \Delta(T)$ and the family $f_c(x) = \sum\limits_{t\in T} \alpha_t g_t(x)$ for $c = (\alpha_t) \in \mathcal{C}$. It is clear that $(f_c)_{c\in\mathcal{C}}$ satisfies the assumptions of Theorem 3.1 and $g = \sup_{c\in\mathcal{C}} f_c$. Therefore,

$$\mathrm{e}_\lambda\, g(x) = \sup_{c\in\mathcal{C}} \mathrm{e}_\lambda\, f_c(x) = \sup_{(\alpha_t)\in\Delta(T)} \mathrm{e}_\lambda \left(\sum_{t\in T} \alpha_t g_t\right)(x) \quad \text{for all } x \in \mathcal{H},$$

which ends the proof. ∎

An prominent case of the above result corresponds to the case where T is a finite set. We write this important particular case as a corollary.

Corollary 3.3. *Let $g_i : \mathcal{H} \to \mathbb{R}_\infty$ with $i = 1, \ldots, n$ be a finite family of functions on $\Gamma_0(\mathcal{H})$ such that $\bigcap_{i=1}^n \operatorname{dom} g_i \neq \emptyset$. Then, for all $\lambda > 0$*

$$\mathrm{e}_\lambda\, g(x) = \sup_{(\alpha_i)_{i=1}^n\in\Delta_n} \mathrm{e}_\lambda \left(\sum_{i=1}^n \alpha_i g_i\right)(x) \quad \textit{for all } x \in \mathcal{H},$$

where $\Delta_n := \{(\alpha_i)_{i=1}^n \in \mathbb{R}^n_+ :\ \sum_{i=1}^n \alpha_i = 1\}$.

The rest of this section is dedicated to calculating the gradient of the Moreau envelope of the supremum function defined in (3.1). The following formula (see [10, Corollary 3.13] and also [9, 18, 27, 28]) will be useful for this purpose.

Proposition 3.4. *Let $(g_t)_{t\in T} \subset \Gamma_0(\mathcal{H})$, and suppose that $g := \sup_{t\in T} g_t$ is finite and continuous at $\bar{x}$. Assume additionally that T is a compact topological space and for every $z \in \operatorname{dom} g$, the function $t \mapsto g_t(z)$ is upper semicontinuous. Then*

$$\partial g(\bar{x}) = \operatorname{co}\{x^* \in \partial g_t(\bar{x}) : g_t(\bar{x}) = g(\bar{x})\}.$$

To calculate the gradient of the Moreau envelope of the supremum function f defined in (3.1), we introduce the following notation: for $\lambda > 0$ and $x \in \mathcal{H}$ given, we define the active set of $\mathrm{e}_\lambda f$ at x by

$$\mathcal{C}_\lambda(x) := \{c \in \mathcal{C} : \mathrm{e}_\lambda\, f(x) = \mathrm{e}_\lambda\, f_c(x)\}.$$

Theorem 3.5. *Let $\mathcal{C}$ be a nonempty compact and convex set and $(f_c)_{c\in\mathcal{C}} \subset \Gamma_0(\mathcal{H})$ be a family of functions such that $f = \sup_{c\in\mathcal{C}} f_c$ is proper and the map $\mathcal{C} \ni c \mapsto f_c(x)$ is concave and upper semicontinous for all $x \in \mathcal{H}$. Then, for all $\lambda > 0$ and $x \in \mathcal{H}$, the set $\mathcal{C}_\lambda(x)$ is nonempty,*

$$\mathrm{P}_\lambda\, f(x) = \mathrm{P}_\lambda\, f_c(x) \ \textit{and}\ \nabla\, \mathrm{e}_\lambda\, f(x) = \frac{x - \mathrm{P}_\lambda\, f_c(x)}{\lambda} \ \textit{for all } c \in \mathcal{C}_\lambda(x). \tag{3.3}$$

Proof. By virtue of Theorem 3.1, we have that $\mathrm{e}_\lambda\, f(x) = \sup_{c\in\mathcal{C}} \mathrm{e}_\lambda\, f_c(x)$ for all $x \in \mathcal{H}$. Since for each $y \in \mathcal{H}$ the function $c \to f_c(y)$ is upper semicontinuous, the function $c \to f_c(y) + \frac{1}{2\lambda}\|x - y\|^2$ is also upper semicontinuous for each $y \in \mathcal{H}$. Consequently, since the infimum of upper semicontinuous functions is upper semicontinuous, the function $c \to \mathrm{e}_\lambda\, f_c(x)$ is also upper semicontinuous for all $x \in \mathcal{H}$. Moreover, since the Moreau envelope is continuously differentiable, Proposition 3.4 implies that

$$\{\nabla\, \mathrm{e}_\lambda\, f(x)\} = \partial\, \mathrm{e}_\lambda\, f(x) = \operatorname{co}\{\nabla\, \mathrm{e}_\lambda\, f_c(x) : c \in \mathcal{C}_\lambda(x)\}.$$

Finally, $\dfrac{x - \mathrm{P}_\lambda\, f_c(x)}{\lambda} \in \partial\, \mathrm{e}_\lambda\, f(x)$ for any $c \in \mathcal{C}_\lambda(x)$, which implies (3.3). ∎

4. Optimality and duality in infinite programming. In this section, we investigate the regularization of general problems arising in infinite and semi-infinite programming. To cover several classes of problems, we start our analysis by considering the class of Robust (objective) optimization problems:

$$(\mathcal{P}) \qquad \min_{x\in\mathcal{Q}} f(x) := \sup_{c\in\mathcal{C}} f_c(x),$$

where $(f_c)_{c\in\mathcal{C}}$ is a family of functions on $\Gamma_0(\mathcal{H})$ and $\mathcal{Q}$ is a closed convex set. For $\lambda > 0$, we consider the following regularization version of problem ($\mathcal{P}$):

$$(\mathcal{P}_\lambda) \qquad \min_{x\in\mathcal{Q}} \mathrm{e}_\lambda f(x).$$

The formal relation between these optimization problems is given in the following proposition.

PROPOSITION 4.1. *Let x_k be an ϵ_k-minimizer of Problem ($\mathcal{P}_{\lambda_k}$) for $k \in \mathbb{N}$, where $\varepsilon_k \to 0$ and $\lambda_k \searrow 0$. If $x_k \rightharpoonup \bar{x}$, then $\bar{x}$ is a minimizer of ($\mathcal{P}$).*

Proof. Since x_k is an ϵ_k-minimizer of Problem ($\mathcal{P}_{\lambda_k}$) and by virtue of Proposition 2.2, it follows that for all $k \in \mathbb{N}$

$$\mathrm{e}_{\lambda_k} f(x_k) \le \mathrm{e}_{\lambda_k} f(x) + \epsilon_k \le f(x) + \epsilon_k \quad \text{for all } x \in \mathcal{Q}. \qquad \square$$

Hence, according to Proposition 2.2, $f(\bar{x}) \le \liminf \mathrm{e}_{\lambda_k} f(x_k) \le f(x)$ for all $x \in \mathcal{Q}$, which proves the result.

The following result gives necessary and sufficient optimality conditions for the optimization problem ($\mathcal{P}_\lambda$).

THEOREM 4.2. *Let $\mathcal{C} \neq \emptyset$ be a nonempty compact convex set and $(f_c)_{c\in\mathcal{C}} \subset \Gamma_0(\mathcal{H})$ be a family of functions such that $f := \sup_{c\in\mathcal{C}} f_c$ is proper and the map $c \mapsto f_c(x)$ is concave and upper semicontinuous for all $x \in \mathcal{H}$. Then, $\bar{x} \in \mathcal{Q}$ is a solution of Problem ($\mathcal{P}_\lambda$) if and only if there exist $c \in \mathcal{C}_\lambda(\bar{x})$, which is nonempty, such that*

$$\mathrm{P}_\lambda f_c(\bar{x}) \in \bar{x} + N_{\mathcal{Q}}(\bar{x}). \tag{4.1}$$

Proof. Let us consider the function $\varphi := \mathrm{e}_\lambda f + \delta_{\mathcal{Q}}$. Then, $\bar{x} \in \mathcal{Q}$ is a minimum of ($\mathcal{P}_\lambda$) if and only if $0_{\mathcal{H}} \in \partial\left(\mathrm{e}_\lambda f + \delta_{\mathcal{Q}}\right)(\bar{x})$. Hence, by the smoothness of the Moreau envelope and the sum rule for the the convex subdifferential (see, e.g., [4, Corollary 16.38]), we obtain that

$$\partial\left(\mathrm{e}_\lambda f + \delta_{\mathcal{Q}}\right)(\bar{x}) = \nabla \mathrm{e}_\lambda f(\bar{x}) + N_{\mathcal{Q}}(\bar{x}).$$

Therefore, by virtue of Theorem 3.5, $\bar{x} \in \mathcal{Q}$ is a solution of ($\mathcal{P}_\lambda$) if and only if there exists $c \in \mathcal{C}_\lambda(\bar{x})$ such that

$$0_{\mathcal{H}} \in \frac{\bar{x} - \mathrm{P}_\lambda f_c(\bar{x})}{\lambda} + N_{\mathcal{Q}}(\bar{x}),$$

which implies the result. $\square$

Remark 4.3. It is important to emphasize that Theorem 4.2 does not require any qualification conditions. Moreover, the concave assumption can be obtained for arbitrary family of functions by considering the generalized simplex as in the Corollary 3.2.

Now, we proceed to apply Theorem 4.2 to obtain necessary and sufficient optimality conditions for concrete classes of optimization problems.

Infinite Programming Problems. Let us consider the following infinite programming problem:

$$\text{(I)} \qquad \begin{aligned} &\min g(x) \\ \text{s.t.} \quad &x \in \mathcal{Q} \text{ and } f_c(x) \leq 0 \text{ for all } c \in \mathcal{C}, \end{aligned}$$

where $g \in \Gamma_0(\mathcal{H})$, $\mathcal{Q} \subset \mathcal{H}$ is a closed and convex set, $(f_c)_{c\in\mathcal{C}} \subset \Gamma_0(\mathcal{H})$ and $\mathcal{C}$ is a compact and convex set. Given $\lambda > 0$ and $\eta \in [0, +\infty)$, we consider the following regularization form of the Problem (I)

$$(\mathrm{I}_{\lambda,\eta}) \qquad \begin{aligned} &\min \mathrm{e}_\lambda\, g(x) \\ \text{s.t.} \quad &x \in \mathcal{Q} \text{ and } \mathrm{e}_\lambda\, f_c(x) \leq \eta \text{ for all } c \in \mathcal{C}. \end{aligned}$$

The following result relates (I) and $(\mathrm{I}_{\lambda,\eta})$. The proof follows similar arguments to the given in the proof of Proposition 4.1, so we omit it.

PROPOSITION 4.4. *Let (x_k) be a sequence of ϵ_k-minimizers of $(\mathrm{I}_{\lambda_k,\eta_k})$ with $\eta_k \to 0$, $\varepsilon_k \to 0$ and $\lambda_k \searrow 0$. If $x_k \rightharpoonup \bar{x}$, then $\bar{x}$ is a minimizer of (I).*

The next theorem gives necessary and sufficient optimality conditions for the problem $(\mathrm{I}_{\lambda,\eta})$.

THEOREM 4.5. *Let $g \in \Gamma_0(\mathcal{H})$, $\mathcal{C}$ be a nonempty, compact and convex set and $(f_c)_{c\in\mathcal{C}} \subset \Gamma_0(\mathcal{H})$ be a family of functions such that the supremum function $f := \sup_{c\in\mathcal{C}} f_c$ is proper and the map $c \mapsto f_c(x)$ is concave for all $x \in \mathcal{H}$. Let $\lambda > 0$ and $\eta \in [0, +\infty)$, and assume that the optimal value of Problem $(\mathrm{I}_{\lambda,\eta})$ is finite. Then, if $\bar{x}$ is a solution of Problem $(\mathrm{I}_{\lambda,\eta})$ there are $\alpha_1, \alpha_2 \geq 0$ and $c \in \mathcal{C}$ with $\alpha_1 + \alpha_2 = 1$ such that*

$$\alpha_2(\mathrm{e}_\lambda\, f(\bar{x}) - \eta) = 0 \text{ and } 0_{\mathcal{H}} \in \bar{x} - \alpha_1 \mathrm{P}_\lambda\, g(\bar{x}) - \alpha_2 \mathrm{P}_\lambda\, f_c(\bar{x}) + N_{\mathcal{Q}}(\bar{x}). \tag{4.2}$$

Conversely, if $\bar{x}$ is feasible for $(\mathrm{I}_{\lambda,\eta})$, satisfies (4.2) and $\mathrm{proj}_{\mathcal{Q}}(\mathrm{P}_\lambda\, f_c(\bar{x})) \neq \bar{x}$, then $\bar{x}$ is a mimimun of $(\mathrm{I}_{\lambda,\eta})$.

Proof. Let us denote by β the optimal value of Problem $(\mathrm{I}_{\lambda,\eta})$ and consider the optimization problem

$$\min_{x\in\mathcal{Q}} h(x), \tag{4.3}$$

where $h(x) := \max\{\mathrm{e}_\lambda\, g(x) - \beta + \eta, \mathrm{e}_\lambda\, f(x)\}$. Then, $\bar{x}$ solves $(\mathrm{I}_{\lambda,\eta})$ if and only if $\bar{x}$ solves (4.3). Therefore, $\bar{x}$ solves $(\mathrm{I}_{\lambda,\eta})$ if and only if

$$0_{\mathcal{H}} \in \partial\left(h + \delta_{\mathcal{Q}}\right)(\bar{x}).$$

By the continuity of h and the sum rule for the convex subdifferential (see, e.g., [4, Corollary 16.38]), we obtain that

$$0_{\mathcal{H}} \in \partial h(\bar{x}) + N_{\mathcal{Q}}(\bar{x}).$$

Moreover, due to Proposition 3.4, the subdifferential of the above max function is given by the convex hull of (sub)-gradients which attains the maximum. In our particular case (recall that $\mathrm{e}_\lambda\, g(\bar{x}) = \beta$)

$$\partial h(\bar{x}) = \begin{cases} \mathrm{co}\{\nabla \mathrm{e}_\lambda\, g(\bar{x}), \nabla \mathrm{e}_\lambda\, f(\bar{x})\} & \text{if } \mathrm{e}_\lambda\, f(\bar{x}) = \eta, \\ \nabla \mathrm{e}_\lambda\, g(\bar{x}) & \text{if } \mathrm{e}_\lambda\, f(\bar{x}) < \eta. \end{cases}$$

Finally, by virtue of Theorem 3.5, the gradient $\nabla \mathrm{e}_\lambda f(\bar{x}) = \frac{\bar{x} - \mathrm{P}_\lambda f_c(\bar{x})}{\lambda}$ for any $c \in \mathcal{C}_\lambda(\bar{x})$. Therefore, (4.2) holds.
Conversely, notice that if the multiplier α_1 in (4.2) is equal to zero, then

$$\mathrm{P}_\lambda f_c(\bar{x}) \in \bar{x} + N_{\mathcal{Q}}(\bar{x}),$$

which means that $\mathrm{proj}_{\mathcal{Q}}(\mathrm{P}_\lambda f_c(\bar{x})) = \bar{x}$. Thus, (4.2) holds with $\alpha_1 \neq 0$. Hence, for every feasible point y of Problem ($\mathrm{I}_{\lambda,\eta}$), we have that

$$\begin{aligned}
0 &\leq \alpha_1(\mathrm{e}_\lambda g(y) - \mathrm{e}_\lambda g(\bar{x})) + \alpha_2(\mathrm{e}_\lambda f_c(y) - \mathrm{e}_\lambda f_c(\bar{x}))\\
&\leq \alpha_1(\mathrm{e}_\lambda g(y) - \mathrm{e}_\lambda g(\bar{x})) + \alpha_2(\eta - \mathrm{e}_\lambda f_c(\bar{x}))\\
&= \alpha_1(\mathrm{e}_\lambda g(y) - \mathrm{e}_\lambda g(\bar{x}))
\end{aligned}$$

where in the last inequality equality we have used that $\alpha_2(\eta - \mathrm{e}_\lambda f_c(\bar{x})) = 0$. Therefore, dividing by α_1, we get that $\mathrm{e}_\lambda g(y) \geq \mathrm{e}_\lambda g(\bar{x})$ for every feasible point y of Problem ($\mathrm{I}_{\lambda,\eta}$), which ends the proof. ∎

Conic programming. Let us consider the following conic programming problem

$$\text{(4.4)} \qquad \begin{aligned} &\min g(x)\\ &\text{s.t. } x \in \mathcal{Q} \text{ and } F(x) \in -\mathcal{K}, \end{aligned}$$

where g is a function in $\Gamma_0(\mathcal{H})$, $\mathcal{Q} \subset \mathcal{H}$ is a closed and convex set, $\mathcal{K} \subset \mathcal{Z}$ is a closed convex cone of a Banach space $\mathcal{Z}$ and $F : \mathrm{dom}\, F \to \mathcal{Z}$. We recall that F is $\mathcal{K}$-convex (see, e.g., [7,20,35]), if $\langle x^*, F\rangle$ is convex for all $x^* \in \mathcal{K}^+$, where

$$\langle x^*, F\rangle(x) = \begin{cases} \langle x^*, F(x)\rangle & \text{if } x \in \mathrm{dom}\, F,\\ +\infty & \text{if } x \notin \mathrm{dom}\, F. \end{cases}$$

and $\mathcal{K}^+ := \{x^* \in \mathcal{Z} : \langle x^*, w\rangle \geq 0 \text{ for all } w \in \mathcal{K}\}$ is the positive polar cone of $\mathcal{K}$. The Problem (4.4) can be equivalently written as

$$(\mathcal{R}) \qquad \begin{aligned} &\min g(x)\\ &\text{s.t. } x \in \mathcal{Q} \text{ and } \sup_{x^* \in \mathcal{K}^+ \cap \mathbb{B}} \langle x^*, F\rangle(x) \leq 0, \end{aligned}$$

where $\mathbb{B}$ is the unit ball in $\mathcal{Z}$. Thus, for $\lambda > 0$ and $\eta \in [0, +\infty)$, we propose the following regularized problem:

$$(\mathcal{R}_{\lambda,\eta}) \qquad \begin{aligned} &\min \mathrm{e}_\lambda g(x)\\ &\text{s.t. } x \in \mathcal{Q} \text{ and } \mathrm{e}_\lambda \left[\sup_{x^* \in \mathcal{K}^+ \cap \mathbb{B}} \langle x^*, F\rangle\right](x) \leq \eta. \end{aligned}$$

As a consequence of Theorem 4.5, we get optimality conditions for the problem ($\mathcal{R}_{\lambda,\eta}$).

COROLLARY 4.6. *Fix $\lambda > 0$ and $\eta \in [0, +\infty)$. Assume that g belongs to $\Gamma_0(\mathcal{H})$, $\mathcal{Q}$ is a closed and convex set, $\mathcal{K}$ is a closed convex cone and F be a $\mathcal{K}$-convex function such that the optimal value of ($\mathcal{R}_{\lambda,\eta}$) is finite. Then, if $\bar{x}$ is a optimal solution of Problem ($\mathcal{R}_{\lambda,\eta}$), then there exist $\alpha_1, \alpha_2 \geq 0$ and $x^* \in \mathcal{K}^+ \cap \mathbb{B}$ with $\alpha_1 + \alpha_2 = 1$ such that*

$$\text{(4.5)} \qquad \alpha_2(\langle x^*, F(\bar{x})\rangle - \eta) = 0 \text{ and } \alpha_1 \mathrm{P}_\lambda g(\bar{x}) + \alpha_2 \mathrm{P}_\lambda \langle x^*, F\rangle(\bar{x}) \in \bar{x} + N_{\mathcal{Q}}(\bar{x}).$$

Conversely, if $\bar{x}$ is a feasible for ($\mathcal{R}_{\lambda,\eta}$), satisfies (4.5) and $\mathrm{proj}_{\mathcal{Q}}(\mathrm{P}_\lambda \langle x^*, F\rangle(\bar{x})) \neq \bar{x}$, then $\bar{x}$ is a optimal solution of ($\mathcal{R}_{\lambda,\eta}$).

4.1. Duality in infinite programming. In this section, associated to a point $x^* \in \mathcal{H}$, we consider the Primal optimization problem

$$
(\mathrm{P}(x^*)) \qquad \begin{aligned} &\inf \left(\mathrm{e}_\lambda f(x) - \langle x^*, x\rangle\right) \\ &\text{s.t.} \quad x \in \mathcal{Q}, \end{aligned}
$$

The for each $c \in \mathcal{C}$ let us consider $\mathcal{X}_c := \mathcal{H} \times \mathcal{H}$ and the perturbation function $\Phi_c : \mathcal{H} \times \mathcal{X}_c \to \mathbb{R}_\infty$ given by

$$
\Phi_c(x, y, z) = \mathrm{e}_\lambda f_c(x + y) + \delta_{\mathcal{Q}}(x + z)
$$

Proposition 4.7. *Let $\mathcal{C} \neq \emptyset$ be a compact convex set and $(f_c)_{c\in\mathcal{C}} \subset \Gamma_0(\mathcal{H})$ be a family of functions with f proper and such that for all $x \in \mathcal{H}$, the map $c \mapsto f_c(x)$ is concave and upper semicontinuous. Moreover, assume that $\mathcal{Q}$ is a nonempty closed convex set.Then, we have that*

$$
\Phi_c^*(x^*, y^*, z^*) = \begin{cases} f_c^*(y^*) + \frac{\lambda}{2}\|y^*\|^2 + \sigma_{\mathcal{Q}}(z^*) & \text{if } x^* = y^* + z^* \\ +\infty & \text{if } x^* \neq y^* + z^* \end{cases}
$$

Proof. Let us compute the conjugate of Φ_c. Fix $(x^*, y^*, z^*) \in \mathcal{H} \times \mathcal{X}_c$, then

$$
\begin{aligned}
\Phi_c^*(x^*, y^*, z^*) &= \sup_{(x,y,z)\in\mathcal{H}\times\mathcal{X}_c} \{\langle x^*, x\rangle + \langle y^*, y\rangle + \langle z^*, z\rangle - \Phi_c(x, y, z)\} \\
&= \sup_{(x,u,w)\in\mathcal{H}\times H\times\mathcal{Q}} \{\langle x^*, x\rangle + \langle y^*, u - x\rangle + \langle z^*, w - x\rangle - \mathrm{e}_\lambda f_c(u)\} \\
&= \sup_{(x,u,w)\in\mathcal{H}\times H\times\mathcal{Q}} \{\langle x^* - y^* - z^*, x\rangle + \langle y^*, u\rangle + \langle z^*, w\rangle - \mathrm{e}_\lambda f_c(u)\} \\
&= \begin{cases} f_c^*(y^*) + \frac{\lambda}{2}\|y^*\|^2 + \sigma_{\mathcal{Q}}(z^*) & \text{if } x^* = y^* + z^* \\ +\infty & \text{if } x^* \neq y^* + z^* \end{cases}
\end{aligned}
$$

∎

Given $\lambda > 0$ and $x^* \in \mathcal{H}$, we consider the dual problem of ($\mathrm{P}(x^*)$) as (see Section 2.3).

$$
(\mathrm{D}(x^*)) \qquad \begin{aligned} &-\inf\left(f_c^*(y^*) + \frac{\lambda}{2}\|y^*\|^2 + \sigma_{\mathcal{Q}}(z^*)\right) \\ &\text{s.t. } c \in \mathcal{C} \text{ and } (y^*, z^*) \in \mathcal{X}_c \text{ such that } x^* = y^* + z^*. \end{aligned}
$$

Theorem 4.8. *Let $\mathcal{C} \neq \emptyset$ be a compact convex set and $(f_c)_{c\in\mathcal{C}} \subset \Gamma_0(\mathcal{H})$ be a family of functions with f being proper and such that for all $x \in \mathcal{H}$, the map $c \mapsto f_c(x)$ is concave and upper semicontinuous. Moreover, assume that $\mathcal{Q}$ is a nonempty, closed and convex set. Then*

$$
\inf \mathrm{P}(x^*) = \max \mathrm{D}(x^*), \qquad \textit{for all } x^* \in \mathcal{H}. \tag{4.6}
$$

Proof. Let us consider the set

$$
\mathcal{W} := \bigcup_{c\in\mathcal{C}} \{(x^*, \alpha) \in \mathcal{H} \times \mathbb{R} : \exists (y^*, z^*) \in \mathcal{X}_c \text{ s.t. } \Phi_c^*(x^*, y^*, z^*) \leq \alpha\}.
$$

We are going to prove that $\mathcal{W}$ is closed and convex, which by Proposition 2.4 and Proposition 4.7, implies that (4.6) holds.

Convexity: Let us consider $\beta \in [0,1]$ and $(x_1^*, \alpha_1), (x_2^*, \alpha_2) \in \mathcal{W}$, then by definition of the set $\mathcal{W}$ there are $c_1, c_2 \in \mathcal{C}$ and $y_1^*, z_1^*, y_2^*, z_2^* \in \mathcal{H}$ such that $\Phi_{c_1}^*(x_i^*, y_i^*, z_i^*) \leq \alpha_i$ for $i = 1, 2$. Then, by Proposition 4.7 we have that $x_i^* = y_i^* + z_i^*$ and $f(y_i^*) + \frac{\lambda}{2}\|y_i^*\|^2 + \sigma_{\mathcal{Q}}(z_i^*) \leq \alpha_i$. Then, define

$$(c_\beta, x_\beta^*, y_\beta^*, z_\beta^*, \alpha_\beta) := \beta(c_1, x_1^*, y_1^*, z_1^*, \alpha_1) + (1-\beta)(c_2, x_2^*, y_2^*, z_2^*, \alpha_2),$$

we get that $x_\beta^* = y_\beta^* + z_\beta^*$ and by convexity of the involved functions (see Lemma 2.1)

$$\begin{aligned} f_{c_\beta}^*(y_\beta^*) + \frac{\lambda}{2}\|y_\beta^*\|^2 + \sigma_{\mathcal{Q}}(z_\beta^*) \leq & \beta\left(f_{c_1}^*(y_1^*) + \frac{\lambda}{2}\|y_1^*\|^2 + \sigma_{\mathcal{Q}}(z_1^*)\right) \\ & + (1-\beta)\left(f_{c_2}(y_2^*) + \frac{\lambda}{2}\|y_2^*\|^2 + \sigma_{\mathcal{Q}}(z_2^*)\right) \\ & = \beta\alpha_1 + (1-\beta)\alpha_2 = \alpha_\beta, \end{aligned}$$

which shows the convexity of $\mathcal{W}$.

Closedness: Let us consider a sequence $(x_n^*, \alpha_n) \in \mathcal{W}$ converging to (x^*, α), then by definition of the set $\mathcal{W}$ and Proposition 4.7 we have that there are $c_n \in \mathcal{C}$ and $y_n^*, z_n^* \in \mathcal{H}$ such that $x_n^* = y_n^* + z_n^*$ and

$$f_{c_n}^*(y_n^*) + \frac{\lambda}{2}\|y_n^*\|^2 + \sigma_{\mathcal{Q}}(z_n^*) \leq \alpha_n, \text{ for all } n \in \mathbb{N}. \tag{4.7}$$

By compactness of $\mathcal{C}$ we can assume, by passing to a subsequence (subnet if $\mathcal{C}$ is not first countable) that $c_n \to c \in \mathcal{C}$. Now, let us notice that the above inequality implies that

$$h^{**}(x_n^*) + \frac{\lambda}{4}\|y_n^*\|^2 \leq f(y_n^*) + \frac{\lambda}{2}\|y_n^*\|^2 + \sigma_{\mathcal{Q}}(z_n^*) \leq \alpha_n, \text{ for all } n \in \mathbb{N}, \tag{4.8}$$

where $h(x^*) := \inf\{f^*(w_1^*) + \frac{\lambda}{4}\|w_1^*\|^2 + \sigma_{\mathcal{Q}}(w_2^*) : x^* = w_1^* + w_2^*\}$, that is h, is the inf-convolution between $(\mathrm{e}_{\lambda/2} f)^*$ and $\sigma_{\mathcal{Q}}$. Let us check that h^{**} is proper. Indeed, by [4, Proposition 13.21] we have that $h^*(x) = \mathrm{e}_{\lambda/2} f(x) + \delta_{\mathcal{Q}}(x)$ is proper convex and lsc. Therefore, its conjugate function h^{**} is also proper. Now, since h^{**} is proper and lsc we have that $\inf_{n\in\mathbb{N}} h^{**}(x_n^*) > -\infty$, so using (4.8) we have that (y_n^*) is bounded, and so is (z_n^*). Therefore, by passing to a subsequence we can assume that (y_n^*) and (z_n^*) converges weakly to some y^* and z^*, repectively. Now, by Lemma 2.1 we have that

$$\liminf f_{c_n}^*(y_n^*) \geq f_c^*(y^*). \tag{4.9}$$

Finally, using (4.8) and (4.9) we get that

$$f_c^*(y^*) + \frac{\lambda}{2}\|y^*\|^2 + \sigma_{\mathcal{Q}}(z^*) \leq \alpha \text{ with } x^* = y^* + z^*,$$

Therefore, $(x^*, \alpha) \in \mathcal{W}$ and that completes the proof. ∎

5. Optimality and duality in stochastic programming. In this section we use our formulae and results to study optimality conditions and zero duality gap to regularizations of general stochastic problems using the Moreau regularization.

5.1. Normal integrands and expectation functional. We start this section by recalling some definitions and notations. For a topological space $\mathcal{T}$, we denote its Borel σ-algebra by $\mathcal{B}(\mathcal{T})$. Let $(\Omega, \mathcal{A})$ be a measurable space. A set-valued map $M : \Omega \rightrightarrows \mathcal{H}$ is measurable if for every open set $U \subset \mathcal{H}$ the set $M^{-1}(U) := \{\omega \in \Omega : M(\omega) \cap U \neq \emptyset\}$ is measurable. The set-valued map M is said to be graph measurable when $\operatorname{graph} M \in \mathcal{A} \otimes \mathcal{B}(\mathcal{H})$. Here, it is important to recall that when $(\Omega, \mathcal{A})$ is a complete measure space (for some measure) a set-valued map with closed valued is measurable if and only if it is graph measurable. A function $g : \Omega \times \mathcal{H} \to \mathbb{R}_\infty$ is called a normal integrand if the epigraph multifunction, $\omega \rightrightarrows \operatorname{epi} g(\omega, \cdot)$ is measurable with closed values. We say that g is proper if the map $x \mapsto g(\omega, x)$ is proper for almost all $\omega \in \Omega$. In addition, when $g_\omega := g(\omega, \cdot)$ is convex for all $\omega \in \Omega$, the function g is called a convex normal integrand. Again, if the measure space is complete, g is a normal integrand if and only if g is measurable and the map g_ω is lsc, for all $\omega \in \Omega$. We refer to [8, 31, 32] for more details.

Now, let us consider a finite measure μ over $\mathcal{A}$. Let $g : \Omega \times \mathcal{H} \to \mathbb{R}_\infty$ be a measurable function. We define the expectation functional associated to g (also called continuous or integral sum) as the function $\mathbb{E}_g^\mu : \mathcal{H} \to [-\infty, +\infty]$ given by

$$\mathbb{E}_g^\mu(x) := \int_\Omega g^+(\omega, x) d\mu(\omega) + \int_\Omega g^-(\omega, x) d\mu(\omega),$$

where $g^+(\omega, x) := \max\{g(\omega, x), 0\}$ and $g^-(\omega, x) := \min\{g(\omega, x), 0\}$. Here, we use the convention $+\infty + -\infty = +\infty$.

The following is a well-known result for expectation functionals (see, e.g., [6, 11, 12, 33]).

Proposition 5.1. *Let $\mathcal{H}$ be a separable Hilbert space and let g be a convex normal integrand such that there exist $x^* \in L^1(\Omega, \mathcal{H})$ and $\alpha \in L^1(\Omega, \mathbb{R})$ such that*

$$g(\omega, x) \geq \langle x^*(\omega), x \rangle + \alpha(\omega), \text{ for almost all } \omega \in \Omega.$$

Then, $\mathbb{E}_g^\mu$ is convex and lsc. Moreover, if x is a point where $\mathbb{E}_g^\mu$ is continuous, then we have

$$\partial \mathbb{E}_g^\mu(x) = \int_\Omega \partial_x g(\omega, x) d\mu(\omega), \tag{5.1}$$

where $\int_\Omega \partial_x g(\omega, x) d\mu(\omega)$ denotes the Aumann integral *of the convex subdifferential $\partial_x g(\omega, x)$, that is,*

$$\int_\Omega \partial_x g(\omega, x) d\mu(\omega) := \left\{ \int_\Omega x^*(\omega) d\mu(\omega) : x^*(\omega) \in \partial_x g(\omega, x) \text{ a.a. } \omega \in \Omega \right\}.$$

We recall that a net of measures μ_α converges weakly to μ if and only if for every bounded continuous real function $\phi : \mathcal{S} \to \mathbb{R}$ one has

$$\lim_\alpha \int_\mathcal{S} \phi(s) d\mu_\alpha = \int_\mathcal{S} \phi(s) d\mu.$$

We refer to [5, Chapter 8] for more details about the weak convergence of measures.

5.2. Distributionally Robust Optimization. Let $\mathcal{S}$ be a complete and separable metric space. We denote by $\mathcal{A}$ the borel σ-algebra on $\mathcal{S}$. Given a collection Θ

of probability measures on $(\mathcal{S},\mathcal{A})$, we consider the following *Distributionally Robust Optimization* problem

$$(DR) \qquad \min_{x\in\mathcal{Q}} \sup_{\nu\in\Theta} \mathbb{E}^{\nu}_{\varphi}(x),$$

where $\varphi\colon \mathcal{S}\times\mathcal{H}\to\mathbb{R}_\infty$ is a convex normal integrand and $\mathcal{Q}\subset\mathcal{H}$ is a closed and convex set. Distributionally robust optimization is a paradigm for decision-making under uncertainty where the uncertain problem data is governed by a probability distribution that is itself subject to uncertainty [37]. We refer to [13, 16, 36, 37] for more details and applications of this model. For $\lambda>0$, we propose the following approximation of (DR):

$$(DR_\lambda) \qquad \min_{x\in\mathcal{Q}} \mathrm{e}_\lambda \left[\sup_{\nu\in\Theta}\mathbb{E}^{\nu}_{\varphi}\right](x).$$

The following corollary, which is a consequence of Theorem 4.2, provides optimality conditions for the approximate problem (DR_λ).

COROLLARY 5.2. *Let $\mathcal{H}$ be a separable Hilbert space and Θ be a collection of probability measures on $(\mathcal{S},\mathcal{A})$, which is assumed to be convex and compact for the weak topology. Let $\varphi\colon\mathcal{S}\times\mathcal{H}\to\mathbb{R}_\infty$ be a convex normal integrand such that*

(i) For all $x\in\mathcal{H}$, the map $s\mapsto\varphi(s,x)$ is upper semicontinuous and bounded from above.

(ii) There are integrable functions $x^*\colon\mathcal{S}\to\mathcal{H}$ and $\alpha\colon\mathcal{S}\to\mathbb{R}$ such that

$$\varphi(s,x)\geq\langle x^*(s),x\rangle+\alpha(s), \text{ for all } x\in\mathcal{H} \text{ and } s\in\mathcal{S}. \tag{5.2}$$

Then, for every $\lambda>0$, $\bar{x}$ is a solution of (DR_λ) if and only if there exist $\mu\in\Theta$ with $\mathrm{e}_\lambda\mathbb{E}^{\mu}_{\varphi}(\bar{x})=\sup_{\nu\in\Theta}\mathrm{e}_\lambda\mathbb{E}^{\nu}_{\varphi}(\bar{x})$ and $\bar{u}\in\bar{x}+N_{\mathcal{Q}}(\bar{x})$ such that

$$\bar{x}-\bar{u}\in\lambda\int_{\mathcal{S}}\partial_x\varphi(s,\bar{u})d\mu(s).$$

Proof. We observe that the map $(\nu,x)\to\mathbb{E}^{\nu}_{\varphi}(x)$ is concave in ν and convex in x. On the one hand, by virtue of (ii) and Proposition 5.1, $\mathbb{E}^{\nu}_{\varphi}\in\Gamma_0(\mathcal{H})$ for all $\nu\in\Theta$. On the other hand, due to [5, Corollary 8.2.5], the map $\nu\to\mathbb{E}^{\nu}_{\varphi}(x)$ is upper semicontinuous for all $x\in\mathcal{H}$. Besides, by assumption (i), the map $x\mapsto\sup_{\nu\in\Theta}\mathbb{E}^{\nu}_{\varphi}(x)$ is continuous over $\mathcal{H}$, in particular, proper. Therefore, by Theorem 4.2, $\bar{x}$ solves (DR_λ) if and only if there exist $\mu\in\Theta$ with $\mathrm{e}_\lambda\mathbb{E}^{\mu}_{\varphi}(\bar{x})=\sup_{\nu\in\Theta}\mathbb{E}^{\nu}_{\varphi}(\bar{x})$ such that

$$\mathrm{P}_\lambda\mathbb{E}^{\mu}_{\varphi}(\bar{x})\in\bar{x}+N_{\mathcal{Q}}(\bar{x}).$$

Moreover, $u:=\mathrm{P}_\lambda\mathbb{E}^{\mu}_{\varphi}(\bar{x})$ is characterized as the unique point satisfying $\bar{x}\in\bar{u}+\lambda\partial\mathbb{E}^{\mu}_{\varphi}(\bar{u})$ (see, e.g., [4, Proposition 16.34]). Finally, by using Proposition 5.1, we obtain that $\bar{u}$ satisfies

$$\bar{x}-\bar{u}\in\lambda\int_{\mathcal{S}}\partial_x\varphi(s,\bar{u})d\mu(s),$$

which concludes the proof. ∎

5.3. Stochastic robust programming. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space. For each $\omega \in \Omega$, we consider the problem

$$\min_{x \in \mathcal{Q}} f(\omega, x) := \sup_{x \in \mathcal{C}} f_c(\omega, x), \tag{5.3}$$

where $(f_c)_{c\in\mathcal{C}}$ is a family of functions defined from $\Omega \times \mathcal{H}$ into $\mathbb{R}_\infty$. We assume that $\mathcal{C}$ is a bounded, closed and convex set of a separable Hilbert space $\mathcal{Z}$. To deal with (5.3), we introduce the concept of concave-convex normal integrands.

Definition 5.3. *Let $(f_c)_{c\in\mathcal{C}}$ be a family of functions defined from $\Omega \times \mathcal{H}$ into $\mathbb{R}_\infty$. We say $(f_c)_{c\in\mathcal{C}}$ is a family of concave-convex normal integrands if:*

(a) The set-valued map $\Omega \times \mathcal{C} \rightrightarrows \operatorname{epi} f_c(\omega, \cdot) := \{(x, \alpha) \in \mathcal{H} \times \mathbb{R} : f_c(\omega, x) \le \alpha\}$ is $\mathcal{A} \otimes \mathcal{B}(\mathcal{C})$-measurable with nonempty, closed and convex values.

(b) The set-valued map $\Omega \times \mathcal{H} \rightrightarrows \operatorname{hypo} f_{(\cdot)}(\omega, x) := \{(c, \beta) \in \mathcal{C} \times \mathbb{R} : f_c(\omega, x) \ge \beta\}$ is $\mathcal{A} \otimes \mathcal{B}(\mathcal{H})$-measurable with nonempty, closed and convex values.

In particular, the above definition implies that the following maps are convex normal integrand:

(i) $(\omega, x) \mapsto f_c(\omega, x)$, for all $c \in \mathcal{C}$;

(ii) $(\omega, c) \mapsto -f_c(\omega, x)$, for all $x \in \mathcal{H}$.

The next proposition shows that the supremum function defined in (5.3) is also a normal integrand.

Proposition 5.4. *Let $\mathcal{H}$ be a separable Hilbert space and $(f_c)_{c\in\mathcal{C}}$ be a family of concave-convex normal integrands. Then, the supremum function $f := \sup_{c\in\mathcal{C}} f_c$ is also a convex normal integrand.*

Proof. Let us consider the set

$$B := \{(\omega, x) \colon \Omega \times \mathcal{H} \colon \operatorname{hypo} f_{(\cdot)}(\omega, x) \neq \emptyset\}.$$

Due to condition (b) from Definition 5.3, the set B is measurable. Moreover, it is easy to see that $f(\omega, x) = -\infty$ for $(\omega, x) \notin B$. Furthermore, by virtue of [8, Theorem III. 9], we are able to find sequences of measurable functions $c_n \colon B \to \mathcal{C}$ and $\beta_n \colon B \to \mathbb{R}$ such that for all $(\omega, x) \in B$

i) $(c_n(\omega, x), \beta_n(\omega, x)) \in \operatorname{hypo} f_{(\cdot)}(\omega, x)$ for all $n \in \mathbb{N}$;

ii) $\operatorname{cl} \{(c_n(\omega, x), \beta_n(\omega, x)) : n \in \mathbb{N}\} = \operatorname{hypo} f_{(\cdot)}(\omega, x)$.

We claim that

$$f(\omega, x) = \sup_{n\in\mathbb{N}} \beta_n(\omega, x) \text{ for all } (\omega, x) \in B. \tag{5.4}$$

Indeed, fix $(\omega, x) \in B$. On the one hand, due to i), it follows that for all $n \in \mathbb{N}$

$$f(\omega, x) \ge f_{c_n(\omega,x)}(\omega, x) \ge \beta_n(\omega, x),$$

which implies that $f(\omega, x) \ge \sup_{n\in\mathbb{N}} \beta_n(\omega, x)$. On the other hand, due to ii), for each $(c_*, \beta_*) \in \operatorname{hypo} f_{(\cdot)}(\omega, x)$, there exists a sequence $(c_{n_k}(\omega, x), \beta_{n_k}(\omega, x)) \to (c_*, \beta_*)$. Then,

$$\beta_* = \lim_k \beta_{n_k}(\omega, x) \le \sup_{n\in\mathbb{N}} \beta_n(\omega, x).$$

Since (c_*, β_*) is arbitrary, we conclude that $f(\omega, x) \le \sup_{n\in\mathbb{N}} \beta_n(\omega, x)$. Thus, (5.4) holds. From (5.4), we deduce that f is $\mathcal{A} \otimes \mathcal{B}(\mathcal{H})$-measurable. Finally, for each fixed $\omega \in \Omega$ the function $f(\omega, \cdot)$ is convex and lsc, hence by virtue of [19, Proposition 9.3], the function f is a convex normal integrand. ∎

LEMMA 5.5. *Let $\mathcal{H}$ be a separable Hilbert space and let $f\colon \Omega\times\mathcal{H}\to\mathbb{R}_\infty$ be a convex normal integrand. Assume that there exist $x^*\in L^2(\Omega,\mathcal{H})$ and $\alpha\in L^1(\Omega,\mathbb{R})$ such that for all $x\in\mathcal{H}$*

$$f(\omega,x)\geq\langle x^*(\omega),x\rangle+\alpha(\omega),\ \text{almost all } \omega\in\Omega. \tag{5.5}$$

Then, for every $\lambda>0$, the function $\mathrm{e}_\lambda f$ is a convex normal integrand and

$$\mathrm{e}_\lambda f_\omega(x)\geq\langle x^*(\omega),x\rangle-\frac{\lambda}{2}\|x^*(\omega)\|^2+\alpha(\omega)\ \text{for all } x\in\mathcal{H} \text{ and a.a. } \omega\in\Omega. \tag{5.6}$$

Here $\mathrm{e}_\lambda f_\omega(x)$ denotes the Moreau envelope of the function $f_\omega\colon x\mapsto f(\omega,x)$ for a fixed $\omega\in\Omega$.

Proof. The first assertion follows from [32, Exercise 14.38]. To prove (5.6), from (5.5) there exists a null set N such that

$$f(\omega,\mathrm{P}_\lambda f_\omega(x))\geq\langle x^*(\omega),\mathrm{P}_\lambda f_\omega(x)\rangle+\alpha(\omega),\ \text{for all } \omega\in\Omega\setminus N. \tag{5.7}$$

Thus, for all $\omega\in\Omega\setminus N$

$$\begin{aligned}\mathrm{e}_\lambda f_\omega(x)&=f_\omega(\mathrm{P}_\lambda f_\omega(x))+\frac{1}{2\lambda}\|x-\mathrm{P}_\lambda f_\omega(x)\|^2\\&\geq\langle x^*(\omega),\mathrm{P}_\lambda f_\omega(x)\rangle+\frac{1}{2\lambda}\|x-\mathrm{P}_\lambda f_\omega(x)\|^2+\alpha(\omega)\\&\geq\inf_{z\in\mathcal{H}}\left(\langle x^*(\omega),z\rangle+\frac{1}{2\lambda}\|x-z\|^2\right)+\alpha(\omega)\\&=\langle x^*(\omega),x\rangle-\frac{\lambda}{2}\|x^*(\omega)\|^2+\alpha(\omega),\end{aligned}$$

which ends the proof. ∎

Now, we consider the expectation version of the optimization problem (5.3):

$$\min_{x\in\mathcal{Q}}\mathbb{E}^{\mathbb{P}}_f(x), \tag{E}$$

where $f\colon\Omega\times\mathcal{H}\to\mathbb{R}_\infty$ is a function and $\mathcal{Q}\subset\mathcal{H}$ is a closed and convex set. Besides, we consider regularized version of (E).

$$\min_{x\in\mathcal{Q}}\mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}(x), \tag{E$_\lambda$}$$

where $e_\lambda f(\omega,x):=\mathrm{e}_\lambda f_\omega(x)$ denotes the Moreau envelope of the function $f_\omega\colon x\mapsto f(\omega,x)$ for a fixed $\omega\in\Omega$. According to Proposition 5.4 and Lemma 5.5, problems (E) and (E$_\lambda$) are well-defined. Moreover, we have the following relationship between these problems.

PROPOSITION 5.6. *Under the assumptions of Lemma 5.5, consider sequences of $\lambda_k\searrow 0$ and $\epsilon_k\to 0^+$, and let (x_k) be a sequence of ϵ_k-minimizer of problem (E$_\lambda$) for each $k\in\mathbb{N}$, respectively. If $x_k\rightharpoonup\bar{x}$, then $\bar{x}$ is a minimum of Problem (E).*

Proof. On the one hand, for all $k\in\mathbb{N}$ we have

$$\mathbb{E}^{\mathbb{P}}_{\mathrm{e}_{\lambda_k} f}(x_k)\leq\mathbb{E}^{\mathbb{P}}_{\mathrm{e}_{\lambda_k} f}(x)+\epsilon_k\leq\mathbb{E}^{\mathbb{P}}_f(x)+\epsilon_k,\ \text{for all } x\in\mathcal{Q}. \tag{5.8}$$

On the other hand, by Proposition 2.2 we have

$$f_\omega(\bar{x}) \le \liminf_{k\to\infty} \mathrm{e}_{\lambda_k} f_\omega(x_k) \text{ for all } \omega \in \Omega. \tag{5.9}$$

Hence, by Lemma 5.5, we can apply Fatou's Lemma to conclude that

$$\mathbb{E}^{\mathbb{P}}_f(\bar{x}) \le \int_\Omega \liminf_{k\to\infty} \mathrm{e}_{\lambda_k} f_\omega(x_k) d\mathbb{P} \le \liminf_{k\to\infty} \int_\Omega \mathrm{e}_{\lambda_k} f_\omega(x_k) d\mathbb{P} = \liminf_{k\to\infty} \mathbb{E}^{\mathbb{P}}_{\mathrm{e}_{\lambda_k} f}(x_k), \tag{5.10}$$

where in the first inequality we used (5.9). Finally, taking into account (5.8) and (5.10) we conclude the proof. ◻

Now, we provide optimality for the regularized problem.

THEOREM 5.7. *Suppose that $\mathcal{H}$ is separable and let f be a convex normal integrand satisfying (5.5) with $\mathbb{E}^{\mathbb{P}}_f$ is proper. Then, for any $\lambda > 0$, $\bar{x}$ is a minimum of Problem (E_λ) if and only if*

$$\mathbb{E}^{\mathbb{P}}\left(\mathrm{P}_\lambda f_\omega(\bar{x})\right) \in \bar{x} + N_{\mathcal{Q}}(\bar{x}). \tag{5.11}$$

Here f_ω denotes the map $x \mapsto f(\omega, x)$ for a fixed $\omega \in \Omega$.

Proof. Since $\mathbb{E}^{\mathbb{P}}_f$ is proper, there exists $x_0 \in \mathcal{H}$ such that $\mathbb{E}^{\mathbb{P}}_f(x_0) \in \mathbb{R}$. Then, by using Lemma 5.5, for all $x \in \mathcal{H}$ and almost all $\omega \in \Omega$

$$f_\omega(x_0) + \frac{1}{2\lambda}\|x - x_0\|^2 \ge \mathrm{e}_\lambda f_\omega(x) \ge \langle x^*(\omega), x\rangle - \frac{\lambda}{2}\|x^*(\omega)\|^2 + \alpha(\omega).$$

Thus, according to Proposition 5.1, $\mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}$ is convex, lower semicontinuous and finite in the whole space. Consequently, $\mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}$ is also continuous on $\mathcal{H}$. Hence, by Fermat's rule, $\bar{x}$ is a minimum of Problem (E_λ) if and only if $0_{\mathcal{H}} \in \partial \mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}(\bar{x}) + N_{\mathcal{Q}}(\bar{x})$. Now, by using (2.1) and Proposition 5.1,

$$\partial \mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}(\bar{x}) = \mathbb{E}^{\mathbb{P}}\left(\frac{\bar{x} - \mathrm{P}_\lambda f_\omega(\bar{x})}{\lambda}\right).$$

Therefore,

$$\begin{aligned} & 0_{\mathcal{H}} \in \partial \mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}(\bar{x}) + N_{\mathcal{Q}}(\bar{x}) \\ \Leftrightarrow\ & 0_{\mathcal{H}} \in \mathbb{E}^{\mathbb{P}}\left(\frac{\bar{x} - \mathrm{P}_\lambda f_\omega(\bar{x})}{\lambda}\right) + N_{\mathcal{Q}}(\bar{x}) \\ \Leftrightarrow\ & 0_{\mathcal{H}} \in \bar{x} - \mathbb{E}^{\mathbb{P}}\left(\mathrm{P}_\lambda f_\omega(\bar{x})\right) + N_{\mathcal{Q}}(\bar{x}), \end{aligned}$$

which ends the proof. ◻

5.4. Duality in stochastic robust programming. In this section, we consider a duality scheme for the problem (E_λ). For $\lambda > 0$ and $x^* \in \mathcal{H}$ fixed, let us consider the following primal optimization problem:

$$\inf_{x\in\mathcal{Q}} \left(\mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}(x) - \langle x^*, x\rangle\right), \tag{$\mathrm{P}_s(x^*)$}$$

where f is the supremum function defined in (5.3). Also, we consider the following spaces

$$\begin{aligned} L^2(\Omega, \mathcal{C}) &:= \{c \in L^2(\Omega, \mathcal{Z})\colon c(\omega) \in \mathcal{C} \text{ a.a. } \omega \in \mathcal{C}\}, \\ \mathcal{X}_c &:= L^2(\Omega, \mathcal{H}) \times \mathcal{H}, \text{ for } c \in L^2(\Omega, \mathcal{C}), \end{aligned}$$

and the perturbation function $\Phi_c : \mathcal{H} \times \mathcal{X}_c \to \mathbb{R}_\infty$ given by

$$\Phi_c(x, y, z) = \int_\Omega \mathrm{e}_\lambda\, f_{c(\omega)}(\omega, x + y(\omega))d\mathbb{P} + \delta_{\mathcal{Q}}(x + z),$$

where $\mathcal{Q} \subset \mathcal{H}$ is a closed and convex set and $\mathrm{e}_\lambda\, f_{c(\omega)}(\omega, x+y(\omega)) := \mathrm{e}_\lambda \left(f_{c(\omega)}(\omega, \cdot)\right)(x+y(\omega))$ for all $\omega \in \Omega$. The following result shows that the perturbation function is well-defined and the objective function of (E_λ) can be recovered through its maximization.

Proposition 5.8. *Let $(f_c)_{c\in\mathcal{C}}$ be a family of concave.convex normal integrands such that the supremum function $f := \sup_{c\in\mathcal{C}} f_c$ is proper. Then, for all $c \in L^2(\Omega, \mathcal{C})$, the map $(\omega, x) \mapsto f_{c(\omega)}(\omega, x)$ is a convex normal integrand. Moreover, if there exist $x^* \in L^2(\Omega, \mathcal{H})$ and $\alpha \in L^1(\Omega, \mathbb{R})$ such that*

$$f_c(\omega, x) \geq \langle x^*(\omega), x\rangle + \alpha(\omega) \text{ for all } (\omega, c, x) \in \Omega \times \mathcal{C} \times \mathcal{H}, \tag{5.12}$$

then the map $(\omega, x) \mapsto e_\lambda f_{c(\omega)}(\omega, x)$ is a convex normal integrand, Φ_c is well-defined for all $c \in L^2(\Omega, \mathcal{C})$ and

$$\sup_{c\in L^2(\Omega,\mathcal{C})} \Phi_c(x, 0_{\mathcal{X}_c}) = \mathbb{E}^{\mathbb{P}}_{\mathrm{e}_\lambda f}(x) \text{ for all } x \in \mathcal{Q}. \tag{5.13}$$

Proof. First, according to [19, Proposition 9.4 and Remark 9.4] (see also [32, Corollary 14.34]), the function $(\omega, c, x) \to f_c(\omega, x)$ is $\mathcal{A}\otimes\mathcal{B}(\mathcal{C})\otimes\mathcal{H}$-measurable. Hence, for each $c \in L^2(\Omega, \mathcal{C})$, the functions $(\omega, x) \to f_{c(\omega)}(\omega, x)$ is $\mathcal{A} \otimes \mathcal{B}(\mathcal{H})$-measurable and convex and lower semicontinuous with respect to x. Hence, it defines a convex normal integrand.
Now, if (5.13) holds, then the map $(\omega, x) \mapsto f_{c(\omega)}(\omega, x)$ satisfies (5.5). Thus, by applying Lemma 5.5 to this map, we get that $\mathrm{e}_\lambda\, f_{c(\omega)}(\omega, x)$ is a convex normal integrand and satisfies (5.6). In particular, for each $y \in L^2(\Omega, \mathcal{H})$ and $x \in \mathcal{H}$, the function $\omega \mapsto e_\lambda f_{c(\omega)}(\omega, x + y(\omega))$ is measurable. Therefore, Φ_c is well-defined for all $c \in L^2(\Omega, \mathcal{C})$. Finally, to prove (5.13), fix $x \in \mathcal{Q}$ and consider the function

$$h_x(\omega, c, z) := f_c(\omega, z) + \frac{\lambda}{2}\|x - z\|^2 \quad (\omega, c, z) \in \Omega \times \mathcal{C} \times \mathcal{H}.$$

This function, being the sum of convex normal integrands, is also a convex normal integrand (see, e.g., [32, Proposition 14.44]). Hence, by virtue of [32, Proposition 14.40], the function $(\omega, c) \mapsto \inf_{z\in\mathcal{H}} h_x(\omega, c, z) = e_\lambda f_c(\omega, x)$ is measurable (in both variables), and upper semicontinuous and concave with respect to c, that is, the map $(\omega, c) \mapsto -e_\lambda f_c(\omega, x)$ is a convex normal integrand. Now, due to (5.12), we have that $-\int_\Omega \mathrm{e}_\lambda\, f_{c(\omega)}(\omega, x)d\mathbb{P} < +\infty$, for each $c \in L^2(\Omega, \mathcal{C})$. Then, interchanging the infimum and the integral (see, e.g., [8, 11, 19, 31]), we get that

$$\inf_{c\in L^2(\Omega,\mathcal{C})} \int_\Omega -\mathrm{e}_\lambda\, f_{c(\omega)}(\omega, x)d\mu = \int_\Omega \inf_{c\in\mathcal{C}} -\mathrm{e}_\lambda\, f_c(\omega, x)d\mu.$$

Therefore,

$$\sup_{c\in L^2(\Omega,\mathcal{C})} \Phi_c(x,0_{\mathcal{X}_c}) = -\inf_{c\in L^2(\Omega,\mathcal{C})} \int_\Omega -\mathrm{e}_\lambda\, f_{c(\omega)}(\omega,x)d\mu$$
$$= -\int_\Omega \inf_{c\in\mathcal{C}} -\mathrm{e}_\lambda\, f_c(\omega,x)d\mu$$
$$= \int_\Omega \sup_{c\in C} \mathrm{e}_\lambda\, f_c(\omega,x)d\mu$$
$$= \int_\Omega \mathrm{e}_\lambda\, f(\omega,x)d\mu,$$

where we have used Theorem 3.1 in the last equality, which proves (5.12). ∎

LEMMA 5.9. *Suppose, in addition the assumptions of Proposition 5.8, that* $\mathbb{E}_f^{\mathbb{P}}$ *is proper. Then, the function*

$$(\omega,c,x^*) \mapsto f_c^*(\omega,x^*) = \sup_{x\in\mathcal{H}} \langle x^*,x\rangle - f_c(\omega,x), \tag{5.14}$$

is a convex normal integrand. Moreover, there exist $x \in L^2(\Omega,\mathcal{H})$ *and* $\alpha \in L^1(\Omega,\mathbb{R})$ *such*

$$f_c^*(\omega,x^*) \geq \langle x^*,x(\omega)\rangle + \alpha(\omega) \text{ for all } (\omega,c,x) \in \Omega\times\mathcal{C}\times\mathcal{H}. \tag{5.15}$$

Proof. Let us recall that $(\Omega,\mathcal{A},\mathbb{P})$ is a complete probability space. On the one hand, since the set-valued map $(\omega,c) \to \operatorname{epi} f_c(\omega,\cdot)$ is measurable with respect to $\mathcal{A}\otimes\mathcal{B}(\mathcal{C})$, the conjugate function defined in (5.14) is measurable with respect to $\mathcal{A}\otimes\mathcal{B}(\mathcal{C})\otimes\mathcal{B}(\mathcal{H}) = \mathcal{A}\otimes\mathcal{B}(\mathcal{C}\times\mathcal{H})$. Moreover, by Lemma 2.1, the function $(c,x^*) \mapsto f_c^*(\omega,x^*)$ is convex and lower semicontinuous for all $\omega\in\Omega$. Hence, the function defined in (5.15) is a convex normal integrand.
On the other hand, since $\mathbb{E}_f^{\mathbb{P}}$ is proper, there exists $x_0\in\mathcal{H}$ such that the map $\omega\mapsto f(\omega,x_0)$ is integrable. For $\omega\in\Omega$ define $x(\omega) := x_0$ and $\alpha(\omega) := -f(\omega,x_0)$. Then, by the definition of convex conjugate, for all $(\omega,c,x)\in\Omega\times\mathcal{C}\times\mathcal{H}$

$$f_c^*(\omega,x^*) \geq \langle x^*,x_0\rangle - f_c(\omega,x_0) \geq \langle x^*,x_0\rangle - f(\omega,x_0) = \langle x^*,x(\omega)\rangle + \alpha(\omega),$$

which ends the proof. ∎

Remark 5.10. It is worth noting that to prove (5.15) it suffices that the integral function $L^2(\Omega,\mathcal{H}) \ni x(\cdot) \to \int_\Omega f(\omega,x(\omega))d\mathbb{P}$ is proper.

PROPOSITION 5.11. *Suppose, in addition to the assumptions of Proposition 5.8, that* $\mathbb{E}_f^{\mathbb{P}}$ *is proper. Then, we have that*

$$\Phi_c^*(x^*,y^*,z^*) = \begin{cases} \int_\Omega \left(f_{c(\omega)}^*(\omega,y^*(\omega)) + \frac{\lambda}{2}\|y^*(\omega)\|^2 \right) d\mathbb{P} & \text{if } x^* = \int_\Omega y^*(\omega)d\mathbb{P} + z^*, \\ +\sigma_{\mathcal{Q}}(z^*) & \\ +\infty & \text{if } x^* \neq \int_\Omega y^*(\omega)d\mathbb{P} + z^*. \end{cases}$$

Proof. Indeed,

$$\begin{aligned}\Phi_c^*(x^*,y^*,z^*) &= \sup_{(x,y,z)\mathcal{H}\times\mathcal{X}_c}\Bigg(\langle x^*,x\rangle + \int_\Omega \langle y^*(\omega),y(\omega)\rangle d\mathbb{P} + \langle z^*,z\rangle \\ &\qquad - \int_\Omega \mathrm{e}_\lambda\, f_{c(\omega)}(\omega, x+y(\omega))d\mathbb{P} - \delta_{\mathcal{Q}}(x+z)\Bigg) \\ &= \sup_{x\in\mathcal{H}}\langle x^* - \int_\Omega y^*(\omega)d\mathbb{P} - z^*, x\rangle + \sup_{v\in\mathcal{Q}}\langle z^*,v\rangle \\ &\quad + \sup_{u\in L^2(\Omega,\mathcal{H})}\int_\Omega \left(\langle y^*(\omega),u(\omega)\rangle - \mathrm{e}_\lambda\, f_{c(\omega)}(\omega,x+y(\omega))\right)d\mathbb{P}\end{aligned}$$

where we have used the change of variables $u = y + x$ and $v = z + x$.
Moreover, $\sup_{v\in\mathcal{Q}}\langle z^*,v\rangle = \sigma_{\mathcal{Q}}(z^*)$ and

$$\sup_{x\in\mathcal{H}}\langle x^* - \int_\Omega y^*(\omega)d\mathbb{P} - z^*, x\rangle = \begin{cases} 0 & \text{if } x^* = \int_\Omega y^*(\omega)d\mathbb{P} + z^*, \\ +\infty & \text{if } x^* \neq \int_\Omega y^*(\omega)d\mathbb{P} + z^*.\end{cases}$$

Since $\mathbb{E}_f^{\mathbb{P}}$ is proper we have that there exist at least one $u_0 \in L^2(\Omega,\mathcal{H})$ such that

$$\int_\Omega \mathrm{e}_\lambda\, f_{c(\omega)}(\omega, u_0(\omega))d\mathbb{P} < +\infty$$

Then by Rockafellar's formula for the conjugate of a integral function (see, e.g., [8, 11, 19, 31]), we obtain that

$$\begin{aligned}\sup_{u\in L^2(\Omega,\mathcal{H})}\int_\Omega &\left(\langle y^*(\omega),u(\omega)\rangle - \mathrm{e}_\lambda\, f_{c(\omega)}(\omega,x+y(\omega))\right)d\mathbb{P} \\ &= \int_\Omega \left(\mathrm{e}_\lambda\, f_{c(\omega)}(\omega,\cdot)\right)^*(y^*(\omega))d\mathbb{P} \\ &= \int_\Omega \left(f^*_{c(\omega)}(\omega,y^*(\omega)) + \frac{\lambda}{2}\|y^*(\omega)\|^2\right)d\mathbb{P},\end{aligned}$$

which ends the proof. ∎

Given $\lambda > 0$ and x^* fixed, we consider the dual problem for ($\mathrm{P}_s(x^*)$).

$$(\mathrm{D}_s(x^*)) \quad \begin{aligned}&-\inf\left\{\int_\Omega \left(f^*_{c(\omega)}(\omega,y^*(\omega)) + \frac{\lambda}{2}\|y^*(\omega)\|^2\right)d\mathbb{P} + \sigma_{\mathcal{Q}}(z^*)\right\} \\ &\text{s.t. } c \in L^2(\Omega,\mathcal{C}) \text{ and } (y^*,z^*)\in\mathcal{X}_c \text{ such that } x^* = \int_\Omega y^*(\omega)d\mathbb{P} + z^*.\end{aligned}$$

Now, we establish the main result of this section.

THEOREM 5.12. *Let $\mathcal{C} \neq \emptyset$ be a convex and bounded set of a Hilbert space $\mathcal{Z}$ and $(f_c)_{c\in\mathcal{C}}$ be a family of concave-convex normal integrands such that* (5.12) *holds and* $\operatorname{dom}\mathbb{E}_f^{\mathbb{P}} \cap \mathcal{Q} \neq \emptyset$. *Then,* $\inf \mathrm{P}_s(x^*) = \max \mathrm{D}_s(x^*)$ *for all* $x^* \in \mathcal{H}$.

Proof. By Proposition 2.4, we have to prove that the following set is closed and convex.

$$\mathcal{W} := \bigcup_{c\in L^2(\Omega,\mathcal{C})}\left\{(x^*,\alpha)\in\mathcal{H}\times\mathbb{R} : \exists (y^*,z^*)\in\mathcal{X}_c^* \text{ s.t. } \Phi_c^*(x^*,y^*,z^*) \le \alpha\right\}.$$

Convexity: Let us consider $\beta \in [0,1]$ and $(x_1^*, \alpha_1), (x_2^*, \alpha_2) \in \mathcal{W}$, then by definition of the set $\mathcal{W}$ and Proposition 5.11, there are $c_1, c_2 \in L^2(\Omega, \mathcal{C})$, $y_1^*, y_2^* \in L^2(\Omega, \mathcal{H})$ and $z_1^*, z_2^* \in \mathcal{H}$ such that

$$\int_\Omega \left(f^*_{c_i(\omega)}(\omega, y_i^*(\omega)) \right) d\mathbb{P} + \frac{\lambda}{2} \int_\Omega \|y_i^*(\omega)\|^2 d\mathbb{P} + \sigma_{\mathcal{Q}}(z_i^*) \leq \alpha_i,$$

and $x_i^* = \int_\Omega y_i^*(\omega) d\mathbb{P} + z_i^*$ for $i = 1, 2$. Define

$$(c_\beta, x_\beta^*, y_\beta^*, z_\beta^*, \alpha_\beta) := \beta(c_1, x_1^*, y_1^*, z_1^*, \alpha_1) + (1-\beta)(c_2, x_2^*, y_2^*, z_2^*, \alpha_2).$$

Then, $x_\beta^* = \int_\Omega y_\beta^*(\omega) d\mathbb{P} + z_\beta^*$ and, by virtue of Lemma 2.1, we get

$$\begin{aligned} \Phi^*_{c_\beta}(x_\beta^*, y_\beta^*, z_\beta^*) &\leq \beta \Phi^*_{c_1}(x_1^*, y_1^*, z_1^*) + (1-\beta)\Phi^*_{c_2}(x_2^*, y_2^*, z_2^*) \\ &= \beta\alpha_1 + (1-\beta)\alpha_2 = \alpha_\beta, \end{aligned}$$

which prove the convexity of $\mathcal{W}$.

Closedness: Let us consider a sequence $(x_n^*, \alpha_n) \in \mathcal{W}$ converging to (x^*, α). Then, by definition of $\mathcal{W}$ and Proposition 5.11, there are $c_n \in L^2(\Omega, \mathcal{C})$, $y_n^* \in L^2(\Omega, \mathcal{H})$ and $z_n^* \in \mathcal{H}$ such that $x_n^* = \int_\Omega y_n^*(\omega) d\mathbb{P} + z_n^*$ and

$$\int_\Omega f^*_{c_n(\omega)}(\omega, y_n^*(\omega)) d\mathbb{P} + \frac{\lambda}{2} \int_\Omega \|y_n^*(\omega)\|^2 d\mathbb{P} + \sigma_{\mathcal{Q}}(z_n^*) \leq \alpha_n \text{ for all } n \in \mathbb{N}. \tag{5.16}$$

Since $\mathcal{C} \subset \mathcal{Z}$ is a closed, convex and bounded, then $L^2(\Omega, \mathcal{C})$ is also closed, convex and bounded in $L^2(\Omega, \mathcal{Z})$. So, without loss of generality, we can assume that $c_n \rightharpoonup c \in L^2(\Omega, \mathcal{C})$. Now, on the one hand, by inequality (5.15), there are $x \in L^2(\Omega, \mathcal{H})$ and $\alpha \in L^1(\Omega, \mathbb{R})$ such that for almost all $\omega \in \Omega$ and $n \in \mathbb{N}$

$$\begin{aligned} f^*_{c_n(\omega)}(\omega, y_n^*(\omega)) + \frac{\lambda}{6}\|y_n^*(\omega)\|^2 &\geq \langle y_n^*(\omega), x(\omega)\rangle + \alpha(\omega) + \frac{\lambda}{6}\|y_n^*(\omega)\|^2 \\ &\geq \inf_{y^* \in \mathcal{H}} \left(\langle y^*, x(\omega)\rangle + \frac{\lambda}{6}\|y^*\|^2 \right) + \alpha(\omega) \\ &\geq -\frac{3}{2\lambda}\|x(\omega)\|^2 + \alpha(\omega). \end{aligned} \tag{5.17}$$

On the other hand,

$$\begin{aligned} \frac{\lambda}{6}\int_\Omega \|y_n^*(\omega)\|^2 d\mathbb{P} + \sigma_{\mathcal{Q}}(z_n^*) &\geq \frac{\lambda}{6}\left\| \int_\Omega y_n^*(\omega) d\mathbb{P} \right\|^2 + \sigma_{\mathcal{Q}}(z_n^*) \\ &\geq h(x_n^*) := \inf\left\{ \frac{\lambda}{6}\|v^*\|^2 + \sigma_{\mathcal{Q}}(w^*) : v^* + w^* = x_n^* \right\}, \end{aligned} \tag{5.18}$$

where $h \in \Gamma_0(\mathcal{H})$ (see, e.g., [4, Proposition 12.14]). Therefore, by using (5.16), (5.17) and (5.18), we conclude that (y_n^*) is bounded in $L^2(\Omega, \mathcal{H})$. Thus, $y_n^* \rightharpoonup y^* \in L^2(\Omega, \mathcal{H})$ up to a subsequence and, by the equality $x_n^* = \int_\Omega y_n(\omega) d\mathbb{P} + z_n^*$, the sequence $z_n^* \rightharpoonup z^* \in \mathcal{H}$.

Now, by Lemma 5.9, the integral functional defined from $L^2(\Omega, \mathcal{Z}) \times L^2(\Omega, \mathcal{H})$ into $\mathbb{R}_\infty$ by

$$(c, x^*) \to \mathcal{I}_{f^*}(c, x^*) := \begin{cases} \int_\Omega f^*_{c(\omega)}(\omega, x(\omega)) d\mathbb{P} & \text{if } c \in L^2(\Omega, \mathcal{C}), \\ +\infty & \text{otherwise}, \end{cases}$$

is well-defined and weak lower semicontinuous in $L^2(\Omega, \mathcal{Z}) \times L^2(\Omega, \mathcal{H})$. In particular, passing to the limits in (5.16), we obtain that

$$\int_\Omega f^*_{c(\omega)}(\omega, y^*(\omega))d\mathbb{P} + \frac{\lambda}{2}\int_\Omega \|y^*(\omega)\|^2 d\mathbb{P} + \sigma_{\mathcal{Q}}(z^*) \leq \alpha,$$

which proves $(x^*, \alpha) \in \mathcal{W}$, and, consequently, $\mathcal{W}$ is closed. □

6. Final comments. In this work, under concave-convex assumptions, we have proved that the Moreau envelope commutes with the supremum function of a family of convex functions. This unexpected result enables us to compute the gradient of envelopes of supremum function and propose regularization of optimization problems that contain supremum functions in the objective or the constraint. Moreover, we provide optimality conditions and zero-duality results for optimization problems in infinite and stochastic programming. We emphasize that the obtained results do not require the verification of classical qualification conditions, which, in general, are challenging to check even in the finite-dimensional setting.

We expect that our results can be used to design algorithms to approximate optimization problems to infinite and semi-infinite programming with robust constraint or objective function.

REFERENCES

[1] H. Attouch and D. Azé, *Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry-Lions method*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 10 (1993), pp. 289–312, https://doi.org/10.1016/S0294-1449(16)30214-1.

[2] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces*, vol. 17 of MOS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, second ed., 2014, https://doi.org/10.1137/1.9781611973488. Applications to PDEs and optimization.

[3] H. Attouch and J. Peypouquet, *Convergence rate of proximal inertial algorithms associated with Moreau envelopes of convex functions*, in Splitting algorithms, modern operator theory, and applications, Springer, Cham, [2019] ©2019, pp. 1–44.

[4] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, CMS Books Math./Ouvrages Math. SMC, Springer, Cham, second ed., 2017, https://doi.org/10.1007/978-3-319-48311-5. With a foreword by Hédy Attouch.

[5] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007, https://doi.org/10.1007/978-3-540-34514-5.

[6] J. F. Bonnans, *Convex and stochastic optimization*, Universitext, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-14977-2.

[7] R. S. Burachik and V. Jeyakumar, *A dual condition for the convex subdifferential sum formula with applications*, J. Convex Anal., 12 (2005), pp. 279–290.

[8] C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, vol. 580 of Lecture Notes in Math., Springer-Verlag, Berlin-New York, 1977.

[9] 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, https://doi.org/10.1137/15M1045375.

[10] 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.

[11] R. Correa, A. Hantoute, and P. Pérez-Aros, *Characterizations of the subdifferential of convex integral functions under qualification conditions*, J. Funct. Anal., 277 (2019), pp. 227–254, https://doi.org/10.1016/j.jfa.2019.02.007.

[12] R. Correa, A. Hantoute, and P. Pérez-Aros, *Qualification conditions-free characterizations of the ε-subdifferential of convex integral functions*, Appl. Math. Optim., (2019), https://doi.org/10.1007/s00245-019-09604-y.

[13] E. Delage and Y. Ye, *Distributionally robust optimization under moment uncertainty with application to data-driven problems*, Oper. Res., 58 (2010), pp. 595–612, https://doi.org/10.1287/opre.1090.0741.

[14] N. Dinh, M. A. Goberna, M. A. López, and M. Volle, *Convexity and closedness in stable robust duality*, Optim. Lett., 13 (2019), pp. 325–339, https://doi.org/10.1007/s11590-018-1311-5.
[15] N. Dinh, T. H. Mo, G. Vallet, and M. Volle, *A unified approach to robust Farkas-type results with applications to robust optimization problems*, SIAM J. Optim., 27 (2017), pp. 1075–1101, https://doi.org/10.1137/16M1067925.
[16] J. Goh and M. Sim, *Distributionally robust optimization and its tractable approximations*, Oper. Res., 58 (2010), pp. 902–917, https://doi.org/10.1287/opre.1090.0795.
[17] A. Hantoute and M. A. López, *A complete characterization of the subdifferential set of the of an arbitrary family of convex functions*, J. Convex Anal., 15 (2008), pp. 831–858.
[18] 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, https://doi.org/10.1137/070700413.
[19] S. Hu and N. S. Papageorgiou, *Handbook of multivalued analysis. Vol. I*, vol. 419 of Math. Appl., Kluwer Academic Publishers, Dordrecht, 1997, https://doi.org/10.1007/978-1-4615-6359-4.
[20] V. Jeyakumar, Z. Y. Wu, G. M. Lee, and N. Dinh, *Liberating the subgradient optimality conditions from constraint qualifications*, J. Global Optim., 36 (2006), pp. 127–137, https://doi.org/10.1007/s10898-006-9003-6.
[21] A. Jourani, L. Thibault, and D. Zagrodny, *Differential properties of the Moreau envelope*, J. Funct. Anal., 266 (2014), pp. 1185–1237, https://doi.org/10.1016/j.jfa.2013.11.008.
[22] A. Jourani and E. Vilches, *Moreau-Yosida regularization of state-dependent sweeping processes with nonregular sets*, J. Optim. Theory Appl., 173 (2017), pp. 91–116, https://doi.org/10.1007/s10957-017-1083-6.
[23] G. Y. Li, V. Jeyakumar, and G. M. Lee, *Robust conjugate duality for convex optimization under uncertainty with application to data classification*, Nonlinear Anal., 74 (2011), pp. 2327–2341, https://doi.org/10.1016/j.na.2010.11.036.
[24] B. Mordukhovich and T. 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, https://doi.org/10.1137/110857738.
[25] J. J. Moreau, *Théorèmes "*inf-sup*"*, C. R. Acad. Sci. Paris, 258 (1964), pp. 2720–2722.
[26] J. J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France, 93 (1965), pp. 273–299.
[27] P. Pérez-Aros, *Formulae for the conjugate and the subdifferential of the supremum function*, J. Optim. Theory Appl., 180 (2019), pp. 397–427, https://doi.org/10.1007/s10957-018-1350-1.
[28] P. Pérez-Aros, *Subdifferential formulae for the supremum of an arbitrary family of functions*, SIAM J. Optim., 29 (2019), pp. 1714–1743, https://doi.org/10.1137/17M1163141.
[29] P. Pérez-Aros, D. Salas, and E. Vilches, *Determination of convex functions via subgradients of minimal norm*, Math. Program., (2020).
[30] P. Pérez-Aros, D. Salas, and E. Vilches, *On formulae for the ioffe geometric subdifferential of supremum function*, J. Convex Anal., 27 (2020), pp. 487–508.
[31] R. T. Rockafellar, *Integral functionals, normal integrands and measurable selections*, in Nonlinear Operators and the Calculus of Variations, J. P. Gossez, E. J. Lami Dozo, J. Mawhin, and L. Waelbroeck, eds., vol. 543 of Lect. Notes Math., Springer Berlin Heidelberg, 1976, pp. 157–207.
[32] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317 of Grundlehren Math. Wiss., Springer-Verlag, Berlin, 1998.
[33] A. Ruszczyński and A. Shapiro, eds., *Stochastic programming*, vol. 10 of Handbooks Oper. Res. Management Sci., Elsevier Science B.V., Amsterdam, 2003.
[34] S. Simons, *Minimax theorems and their proofs*, in Minimax and applications, vol. 4 of Nonconvex Optim. Appl., Kluwer Acad. Publ., Dordrecht, 1995, pp. 1–23, https://doi.org/10.1007/978-1-4613-3557-3_1.
[35] T. Strömberg, *The operation of infimal convolution*, Dissertationes Math. (Rozprawy Mat.), 352 (1996), p. 58.
[36] W. van Ackooij, D. D. Escobar, M. Glanzer, and G. C. Pflug, *Distributionally robust optimization with multiple time scales: valuation of a thermal power plant*, Computational Management Science, (2019), https://doi.org/10.1007/s10287-019-00358-0.
[37] W. Wiesemann, D. Kuhn, and M. Sim, *Distributionally robust convex optimization*, Oper. Res., 62 (2014), pp. 1358–1376, https://doi.org/10.1287/opre.2014.1314.