Journal of Nonlinear and Convex Analysis
Volume 15 or 16, Number **Proof**, 2014 or
2015, 9999–

ESSENTIAL INTERSECTION AND APPROXIMATION RESULTS FOR ROBUST OPTIMIZATION

CHRISTIAN HESS*, RAFFAELLO SERI, AND CHRISTINE CHOIRAT

Abstract. We examine the concept of essential intersection of a random set in the framework of robust optimization programs and ergodic theory. Using a recent extension of Birkhoff's Ergodic Theorem developed by the present authors, it is shown that essential intersection can be represented as the countable intersection of random sets involving an asymptotically mean stationary transformation. This is applied to the approximation of a robust optimization program by a sequence of simpler programs with only a finite number of constraints. We also discuss some formulations of robust optimization programs that have appeared in the literature and we make them more precise, especially from the probabilistic point of view. We show that the essential intersection appears naturally in the correct formulation.

1. Introduction

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a metric space E and two extended-real-valued functions $f : \Omega \times E \to \overline{\mathbb{R}}$ and $h : E \to \overline{\mathbb{R}}$, consider the optimization problem:

$$\min h(x) \quad \text{subject to} \quad f(\omega, x) \leq 0 \quad \text{for all} \quad \omega \in \Omega. \tag{1.1}$$

If E is a finite dimensional Euclidean space and if Ω is infinite, the problem involves a finite number of unknowns under an infinite number of constraints, which justifies the name of *semi-infinite programming* that is often given to this kind of problems (see, e.g., [18, 23]). A natural situation where such a program finds application is for designing optimal solutions that are robust against uncertain events, subsumed under the parameter ω. As an example, it is possible to show that semi-infinite programs encompass *minimax problems*, in which the decision maker has to select the best strategy in response to the worst possible situation (see, e.g., [30] for the relation between minimax and semi-infinite programs). Another situation strictly linked to the present one arises in *robust feasibility problems*, whose objective is to find $x \in E$ such that $f(\omega, x) \leq 0$ for any $\omega \in \Omega$. The field is also known under the alternative name of *robust optimization* (see [4, 5]). In the sequel we shall use this second name.

Despite having been in use for a long time when ω is a random element of a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, the formulations customarily discussed in the literature, among which (1.1), are not satisfactory from a probabilistic point of view and need to be made more precise. In fact, due to the random nature of ω in (1.1), the

2010 *Mathematics Subject Classification.* 26E25, 37A30, 37A40, 49N99, 52A22, 52A41, 60F15.
Key words and phrases. Robust optimization, set-valued map, random set, essential intersection, Birkhoff's Ergodic Theorem, asymptotically mean stationary processes, strong law of large numbers.

*Corresponding author.

constraints $f(\omega, x) \leq 0$ need not be satisfied for all $\omega \in \Omega$, but only on a subset of probability one. One objective of the present paper is to explain this point and to provide alternative correct formulations (see Remark 5.2). This requires some concepts from the theory of random sets and, in particular, that of essential intersection. We provide a quick presentation of this object, the needed properties and a few examples. Then, we establish a representation formula for the essential intersection, which is the main result of this paper. In turn, the above formula is used to establish several results on the scenario approximation of robust programs.

As robust stochastic optimization problems require optimizing a function on the intersection of an infinite collection of sets and are notoriously difficult to deal with, it is customary to make recourse to approximate solutions. These are often obtained by replacing the original set of constraints with the intersection of a finite number of sets extracted from the previous infinite collection (see among others [24–26, 29]). These methods are sometimes called *outer approximation* or *discretization* methods. We examine this kind of problem and we show that, under suitable conditions, the original problem can be approximated through the optimization of the objective function on the intersection of a finite collection of sets, sampled from an asymptotically mean stationary (in the following, ams) stochastic process. This is useful because the collection of ams processes (see, e.g., [12]) is the largest class of stochastic processes for which it is possible to prove a Birkhoff-type Ergodic Theorem (see [17]). This makes possible, in practical computations, to replace the original infinite set of constraints with only a finite number of them associated with real data, that often display dependence and local nonstationarity. It is well-known that the Strong Law of Large Numbers can be deduced from the Birkhoff-Ergodic Theorem. However, our approximation results can be adapted, not only to the case of independent identically distributed (i.i.d.) observations, but also to the case of pairwise independent identically distributed observations (see Remark 5.13). On the other hand, our results are valid in infinite dimensional Banach spaces, which can be useful for dealing with optimization problems on functional spaces (e.g., Calculus of Variations, Optimal Control, ...). There are versions in the strong topology and in the weak topology.

The paper is organized as follows. In Section 2 we set the notation and introduce the needed preliminaries. In Section 3 we provide several results on essential intersection. Most of them are required later. There is an exception with Theorem 3.16 that involves the notion of image measure. We think that this result is important, because it serves in robust optimization when image measures are involved and one has to switch between two probability spaces. The main result of Section 4 is Theorem 4.1 that gives a representation formula for essential intersection in the framework of ams dynamical systems. A short discussion follows as well as applications and examples. In the first part of Section 5, we explain why the notion of essential intersection is needed to give an appropriate formulation of robust optimization programs. In the second part, using the representation formula of Section 4, we prove a result on the stochastic approximations of robust optimization problems (Theorem 5.3). This result admits several extensions or variants. In particular, we present an extension involving the weak topology in an infinite dimensional Banach space and another for sequences of measurable selections. Finally, we briefly examine the stochastic approximation of optimization problems when the constraints are given by an i.i.d. sequence of random lower semi-continuous functions. This situation is a special case of stationary sequences and is often encountered in applications. In a final remark, we explain why our results are also adaptable to the more general case of pairwise i.i.d. observations.

2. Definitions and preliminaries

In this section we set the notation and terminology and we compile some basic facts that mainly concern Ergodic Theory and the theory of Random Sets.

In the sequel, E generally denotes a Polish space.[1] When a linear structure is needed, E is assumed to be a separable Banach space. The *topological closure* (resp. *interior*) of a subset C of E is denoted by $\mathrm{cl}(C)$ (resp. $\mathrm{int}(C)$). The distance function of C is denoted by $d(\cdot, C)$ and defined by

$$d(x,C) = \inf_{y\in C} d(x,y) \qquad x \in C. \tag{2.1}$$

The *open ball* of radius r centered at x is denoted by $B(x,r)$.

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and an $\mathcal{A}-$measurable transformation $T : \Omega \to \Omega$, $(\Omega, \mathcal{A}, \mathbb{P}, T)$ is often referred to as a *dynamical system.* The transformation T is said to be *null-preserving* if the probability $\mathbb{P}T^{-1}$ is absolutely continuous with respect to $\mathbb{P}$, which is denoted by $\mathbb{P}T^{-1} \ll \mathbb{P}$. The transformation T is said to be *measure-preserving* if $\mathbb{P}\left(T^{-1}A\right) = \mathbb{P}(A)$ for all $A \in \mathcal{A}$. Equivalently, $\mathbb{P}$ is said to be *stationary* with respect to T. We also say that T preserves the $\mathbb{P}-$measure. The sets $A \in \mathcal{A}$ that satisfy $T^{-1}A = A$ are called *$T-$invariant sets* (or simply *invariant*) and constitute a sub$-\sigma-$field $\mathcal{I}$ of $\mathcal{A}$. The notion of *$\mathbb{P}-$almost surely invariant set* is also useful. The class of these sets constitutes a $\sigma-$field which is equal to the $\mathbb{P}-$completion of $\mathcal{I}$. A measurable and measure-preserving transformation T is said to be *ergodic* if $\mathbb{P}(A) = 0$ or 1 for all invariant sets A. Equivalently, the sub$-\sigma-$field $\mathcal{I}$ reduces to the trivial $\sigma-$field $\{\Omega, \emptyset\}$ (up to the $\mathbb{P}-$null sets).

The probability $\mathbb{P}$ is said to be *asymptotically mean stationary* (ams) with respect to T if the sequence $\frac{1}{n}\sum_{j=0}^{n-1}\mathbb{P}\left(T^{-j}A\right)$ is convergent for all $A \in \mathcal{A}$. From the Vitali-Hahn-Saks Theorem, it is known that $\lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1}\mathbb{P}T^{-j}$ is a probability measure denoted by $\mathbb{P}^\star$ and referred to as the *asymptotic mean of* $\mathbb{P}$. The probability $\mathbb{P}^\star$ is stationary with respect to T and coincides with $\mathbb{P}$ on invariant sets. Further, it is not difficult to prove that $\mathbb{P}$ is ams if and only if for each bounded real-valued random variable X the limit

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} X\left(T^i\omega\right)$$

exists for $\mathbb{P}-$almost all $\omega \in \Omega$ (see, e.g., [21, Theorem 4.10]).

Given a dynamical system $(\Omega, \mathcal{A}, \mathbb{P}, T)$ and $A \in \mathcal{A}$, a point $\omega \in \Omega$ is said to be *recurrent* with respect to A if there exists a positive integer n such that $T^n\omega \in A$. An event A having a positive probability is said to be *recurrent* if almost every point of

[1] A Polish space E is a separable topological space whose topology can be given by a metric for which E is complete (in particular a Euclidean space is Polish).

A is recurrent with respect to A. A dynamical system is said to be *recurrent* if every event is a recurrent event. It is also said that the transformation T is recurrent. The notion of infinitely recurrent event and infinitely recurrent dynamical system is also useful. Given $A \in \mathcal{A}$ the set of all ω that return in A infinitely often (denoted by i.o.) is denoted by $A_{i.o.}$ and defined by

$$A_{i.o.} = \bigcap_{m\geq 0} \bigcup_{k\geq m} T^{-k}A.$$

The transformation T is said to be *infinitely recurrent* if for all events A, one has $\mathbb{P}(A \setminus A_{i.o.}) = 0$. By the Poincaré Recurrence Theorem, every stationary dynamical system is recurrent. Further, it is known that a dynamical system $(\Omega, \mathcal{A}, \mathbb{P}, T)$ is recurrent if and only if it is infinitely recurrent (see, e.g., Theorem 6.4.2 in [13]).

A real-valued random variable X is said to be *quasi-integrable* if either $\mathbb{E}X^+$ or $\mathbb{E}X^-$ is finite, where $X^+ = \max\{X, 0\}$ (resp. $X^- = \max\{-X, 0\}$) stands for the positive (resp. the negative) part of X. For any $A \in \mathcal{A}$, the *(probabilistic) indicator function* of A is denoted by 1_A and defined by $1_A(\omega) = 1$ if $\omega \in A$, 0 otherwise. Another kind of indicator function for subsets of E will be introduced below.

The following result is contained in [17] (Theorem 3). It extends the Birkhoff Ergodic Theorem in two directions. Firstly, it is valid for quasi-integrable extended-real-valued random variables under stationary, but not necessarily ergodic, transformations. The random variables may even take infinite values on a set of positive measure. Secondly, it is also valid in the case of ams transformations. This result will be used in Section 4 for proving our result on the representation of the essential intersection (Theorem 4.1). We shall only need the ergodic case.

Theorem 2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P}, T)$ be an* ams *ergodic dynamical system with asymptotic mean $\mathbb{P}^\star$ and X be an extended-real-valued random variable defined on $(\Omega, \mathcal{A})$. Also assume that X is $\mathbb{P}^\star-$quasi-integrable. Then, for $\mathbb{P}$ and $\mathbb{P}^\star-$almost every $\omega \in \Omega$, one has*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} X\left(T^i\omega\right) = \mathbb{E}^\star\left(X\right),$$

where each side can be equal to $+\infty$ or $-\infty$, and where $\mathbb{E}^\star(X)$ denotes the expectation taken on $(\Omega, \mathcal{A}, \mathbb{P}^\star)$.

Remark 2.2. In the classical version of Birkhoff's Ergodic Theorem (see, e.g., [7] or [21]), $\mathbb{P}$ is assumed to be stationary with respect to T and X is assumed to be $\mathbb{P}-$integrable. When $\mathbb{P}$ is stationary one has $\mathbb{P} = \mathbb{P}^\star$ and thus $\mathbb{E}^\star(X) = \mathbb{E}(X)$.

Now we recall some basic facts on random sets. Given a Polish space E, the set of all subsets of E is denoted by 2^E. Basically, a random set is a set-valued map $\Gamma : \Omega \to 2^E$ having some sort of measurability property. Here, we shall use graph measurability. The *graph* of Γ is denoted by $\mathrm{Gr}(\Gamma)$ and defined by

$$\mathrm{Gr}(\Gamma) = \{(\omega, x) \in \Omega \times E : x \in \Gamma(\omega)\}.$$

In this framework, Γ is said to be a *random set* if $\mathrm{Gr}(\Gamma)$ is a member of the product σ-field $\mathcal{A} \otimes \mathcal{B}(E)$. Equivalently, Γ is said to be *graph-measurable*. Other synonyms for random sets are encountered such as 'measurable set-valued map', 'measurable multifunction' or 'measurable correspondence'. From the definition, it follows that the countable intersection and the countable union of random sets is still a random set. In the sequel, we mainly consider closed-valued random sets, also called *random closed sets.* We need the following characterization (see, e.g., [1], [16] or [22]).

Proposition 2.3. *Let E be a Polish space and $\Gamma : \Omega \to 2^E$ be a set-valued map. Consider the following two statements.*

(a) $\mathrm{Gr}(\Gamma) \in \mathcal{A} \otimes \mathcal{B}(E)$*, i.e.,* Γ *is a random set.*

(b) *For every open subset U of E, the set $\Gamma^- U$ defined by*

$$\Gamma^- U = \{\omega \in \Omega : \Gamma(\omega) \cap U \neq \emptyset\}$$

is a member of $\mathcal{A}$.

If Γ is closed-valued then implication (b) $\Longrightarrow$ (a) holds. Conversely, provided $\mathcal{A}$ is replaced by $\mathcal{A}_{\mathbb{P}}$ (the $\mathbb{P}-$completion of $\mathcal{A}$), implication (a) $\Longrightarrow$ (b) also holds. Consequently, if Γ is closed-valued and $\mathcal{A}$ is complete, statements (a) and (b) are equivalent.

Let us present two examples of random sets that shall be used several times in the sequel. Given an extended-real-valued function $\varphi : E \to \overline{\mathbb{R}}$, the *epigraph* (or *upper graph*) of φ is denoted by $\mathrm{epi}(\varphi)$ and defined by

$$\mathrm{epi}(\varphi) = \{(x, \alpha) \in E \times \mathbb{R} : \varphi(x) \leq \alpha\}. \tag{2.2}$$

For any real β, the *level-set* of φ at height β is denoted by $L(\varphi, \beta)$ and defined by

$$L(\varphi, \beta) = \{x \in E : \varphi(x) \leq \beta\}. \tag{2.3}$$

When φ is lower semi-continuous (in short lsc), $\mathrm{epi}(\varphi)$ and $L(\varphi, \beta)$ are closed. These sets are convex when E is a Banach space and φ is a convex function. Given an $\mathcal{A}\otimes\mathcal{B}(E)-$measurable function $f : \Omega \times E \to \overline{\mathbb{R}}$ and a real β, we define the set-valued maps Δ and Γ_β by

$$\Delta : \omega \mapsto \mathrm{epi}(f(\omega, \cdot)) = \{(x, \alpha) \in E \times \mathbb{R} : f(\omega, x) \leq \alpha\}, \tag{2.4}$$

$$\Gamma_\beta(\omega) = L(f(\omega, \cdot), \beta) = \{x \in E : f(\omega, x) \leq \beta\}. \tag{2.5}$$

These maps are readily seen to be random sets (where Δ takes on its values in the Polish space $E \times \mathbb{R}$). Δ is often called the *epigraphical multifunction* associated with f and Γ_β is called the *level-set multifunction* (at height β) associated with f. When $f(\omega, \cdot)$ is lsc (resp. convex) for $\mathbb{P}-$almost all $\omega \in \Omega$, f is often referred to as a *random* lsc *function* (resp. a *random convex function*).

A short review on the Painlevé-Kuratowski convergence (in short $PK-$convergence) is also in order (see, e.g., [1] or [3]). Given a sequence $(C_n)_{n\geq 0}$ of subsets of E, its lower limit and its upper limit are respectively denoted by $\mathrm{Li}\, C_n$ and $\mathrm{Ls}\, C_n$ and defined by

$$\mathrm{Li}\, C_n = \{x \in E : x = \lim_{n\to+\infty} x_n \,,\, x_n \in C_n \quad \forall n\}$$
$$\mathrm{Ls}\, C_n = \{x \in E : x = \lim_{k\to+\infty} x_k \,,\, x_k \in C_{n_k} \quad \forall k\}$$

where $(C_{n_k})_{k\geq 0}$ denotes an infinite subsequence of (C_n). These two sets are closed and the inclusion $\mathrm{Li}\, C_n \subseteq \mathrm{Ls}\, C_n$ easily follows from the definition. If $\mathrm{Li}\, C_n = \mathrm{Ls}\, C_n$ and if C denotes the common value, the sequence (C_n) is said to *$PK-$converge* to C. This is denoted by $C = PK - \lim_n C_n$. In particular if the sequence (C_n) is nonincreasing and if we set $C = \bigcap_n \mathrm{cl}(C_n)$, one has $C = PK - \lim_n C_n$. When the C_n's are compact, the convergence also holds in the sense of Hausdorff distance (see, e.g., [3]).

We need another notion of indicator function, different from the one that we have introduced above in the framework of a probability space. It characterizes subsets of E, and is often used in Optimization Theory and Convex Analysis (see, e.g., [27]). The context will allow for avoiding any ambiguity. For any subset C of E, the *(convex analysis) indicator function* of C is denoted by χ_C and defined for any $x \in E$ by $\chi_C(x) = 0$ if $x \in C$, $\chi_C(x) = +\infty$ otherwise. For any pair (C_1, C_2) of subsets, the following equalities hold true

$$\chi_{C_1 \cap C_2} = \sup(\chi_{C_1}, \chi_{C_2}) = \chi_{C_1} + \chi_{C_2} \quad \text{and} \quad \chi_{C_1 \cup C_2} = \inf(\chi_{C_1}, \chi_{C_2})\,. \tag{2.6}$$

The extension to any finite sequence of subsets is straightforward. This second kind of indicator function is convenient to express properties concerning random sets. For example, a set-valued map $\Gamma : \Omega \to 2^E$ is graph-measurable (i.e. is a random set) if and only if the function $(\omega, x) \mapsto \chi_{\Gamma(\omega)}(x)$ is $\mathcal{A} \otimes \mathcal{B}(E)-$measurable. It is also convenient to express constraints appearing in optimization problems (see, e.g., [27]). Formula (4.1) in Section 4 will show another useful example.

3. Essential intersection

In this section we introduce the notion of essential intersection, which is defined for a random set whose values are subsets of a Polish space E. The term 'essential' is used by analogy with the essential infimum or essential supremum of a random variable in Probability Theory (see, e.g., Example 3.12). This concept seems to have been introduced by Hiriart-Urruty [20] in view of applications to stochastic optimization. Essential intersection is useful to characterize properties of random sets that are satisfied almost surely. We present its most relevant properties in connection with our goals, as well as a few examples.

3.1. **Definition and elementary results.** Let Γ be a random set defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{N}_{\mathbb{P}}$ be the set of $\mathbb{P}-$null sets. The *essential intersection*[2] of Γ with respect to $(\Omega, \mathcal{A}, \mathbb{P})$ is the subset of E, denoted by $\wedge(\Gamma)$ or $\wedge_{\mathbb{P}}(\Gamma)$, and defined by

$$\wedge(\Gamma) = \{x \in E : x \in \Gamma(\omega),\ \mathbb{P} - a.s.\}. \tag{3.1}$$

Thus, x is a member of $\wedge(\Gamma)$ if and only if $x \in \Gamma(\omega)$ for $\mathbb{P}-$almost all $\omega \in \Omega$. More precisely, we can write

$$\wedge(\Gamma) = \bigcup_{N \in \mathcal{N}_{\mathbb{P}}} I(\Gamma, N) \tag{3.2}$$

where, for every $N \in \mathcal{N}_{\mathbb{P}}$, the set $I(\Gamma, N)$ is defined by

$$I(\Gamma, N) = \bigcap_{\omega \in \Omega \setminus N} \Gamma(\omega). \tag{3.3}$$

[2]It is sometimes called the 'continuous' intersection, in reference to the case where the probability space is nonatomic.

A simple and useful property is given by the following implication, valid for all $N_1, N_2 \in \mathcal{N}_{\mathbb{P}}$,

$$N_1 \subseteq N_2 \Rightarrow I(\Gamma, N_1) \subseteq I(\Gamma, N_2).$$

Example 3.1. Observe that the sets $I(\Gamma, N)$ may be much smaller than $\wedge(\Gamma)$. For example, consider the case where $\Omega = E = [0,1]$, the unit interval endowed with the Lebesgue measure $\mathbb{P}$. If we define the random set Γ by $\Gamma(\omega) = E \setminus \{\omega\}$, it is readily checked that $\wedge(\Gamma) = E$, whereas $I(\Gamma, N) = N$ for every $\mathbb{P}-$null set N.

Remark 3.2. (Essential intersection and null sets). (i) If the random set Γ is modified on a $\mathbb{P}-$null set, the essential intersection is unchanged. Moreover, the essential intersection does not depend explicitly upon the probability $\mathbb{P}$, but rather on the set $\mathcal{N}_{\mathbb{P}}$ of all $\mathbb{P}-$null sets. Consequently, the essential intersection is not modified if one replaces $\mathbb{P}$ by an equivalent probability measure.

(ii) Given two probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$ on $(\Omega, \mathcal{A})$, if $\mathbb{P}_1$ is absolutely continuous with respect to $\mathbb{P}_2$, it follows that $\mathcal{N}_{\mathbb{P}_2} \subseteq \mathcal{N}_{\mathbb{P}_1}$, which in turn implies $\wedge_{\mathbb{P}_2}(\Gamma) \subseteq \wedge_{\mathbb{P}_1}(\Gamma)$.

Remark 3.3. Essential intersection is stable by countable intersection of random sets: it follows from the definition that if $(\Gamma_k)_{k\geq 0}$ is a sequence of random sets and if Γ is defined by $\Gamma = \bigcap_{k\geq 0} \Gamma_k$, then $\wedge(\Gamma) = \bigcap_{k\geq 0} \wedge(\Gamma_k)$.

Remark 3.4 (Sufficient subfamilies of null sets)**.** In (3.2) it is often enough to take the union over a strict subfamily of $\mathcal{N}_{\mathbb{P}}$. The following notion will be useful. A subfamily $\mathcal{N}_0$ of $\mathcal{N}_{\mathbb{P}}$ is said to be sufficient in $\mathcal{N}_{\mathbb{P}}$ (for computing the essential intersection of Γ) if for each $N \in \mathcal{N}_{\mathbb{P}}$ there exists $N_0 \in \mathcal{N}_0$ such that $I(\Gamma, N) \subseteq I(\Gamma, N_0)$. In this case, the equivalent formula is valid

$$\wedge(\Gamma) = \bigcup_{N\in\mathcal{N}_0} I(\Gamma, N).$$

The notion of sufficient subfamily of null sets is convenient and will be used several times in the sequel.

It is interesting to know when properties of Γ are transmitted to $\wedge(\Gamma)$. The following simple result provides two examples.

Proposition 3.5. (a) *Let E be a Polish space. If Γ is closed-valued, then the essential intersection $\wedge(\Gamma)$ is closed.*

(b) *If E is a separable Banach space and if Γ is convex-valued, then $\wedge(\Gamma)$ is convex.*

Proof of (a). Consider a sequence $(x_k)_{k\geq 1}$ in $\wedge(\Gamma)$ and assume that it converges to $x \in E$. For each $k \geq 1$, one can find a null set $N_k \in \mathcal{N}_{\mathbb{P}}$ such that $x_k \in I(\Gamma, N_k)$. If we define the null set N by $N = \bigcup_{k\geq 1} N_k$, then we have $x_k \in I(\Gamma, N)$ for all $k \geq 1$. Since $I(\Gamma, N)$ is the intersection of closed sets, it is closed, which entails $x \in I(\Gamma, N) \subseteq \wedge(\Gamma)$ and proves the closedness of $\wedge(\Gamma)$. The proof of (b) is similar. $\square$

Using the same kind of arguments, it is possible to show that if $\Gamma(\omega)$ is compact for $\mathbb{P}-$almost all ω, then $\wedge(\Gamma)$ is compact. Moreover, a stronger result holds true as the following proposition shows.

Proposition 3.6. *If* Γ *is a closed-valued random set satisfying Condition* (**K**) *hereafter:*

(**K**) *there exists* $A_0 \in \mathcal{A}$ *of positive measure such that* $\Gamma(\omega)$ *is compact for all* $\omega \in A_0$,

then $\wedge(\Gamma)$ *is compact.*

Proof. Consider a sequence $(x_k)_{k\geq 1}$ in $\wedge(\Gamma)$ and define $B_0 \in \mathcal{A}$ by

$$B_0 = \{\omega \in \Omega : x_k \in \Gamma(\omega) \qquad \forall k \geq 1\}.$$

It is readily checked that $\mathbb{P}(B_0) = 1$, which implies $\mathbb{P}(A_0 \cap B_0) = \mathbb{P}(A_0) > 0$ and shows that $A_0 \cap B_0$ is nonempty. By Condition (**K**), for any $\omega \in A_0 \cap B_0$, it is possible to find a subsequence (x_{k_i}) of (x_k) and $x \in \Gamma(\omega)$ such that $x = \lim_{i\to+\infty} x_{k_i}$. By Proposition 3.5 (a), $\wedge(\Gamma)$ is closed, which implies $x \in \wedge(\Gamma)$. □

In the framework of ams ergodic dynamical systems, condition (**K**) implies the compactness of $\Gamma(T^i\omega)$ for infinitely many indices.

Proposition 3.7. *If* $(\Omega, \mathcal{A}, \mathbb{P}, T)$ *is an ams ergodic dynamical system with asymptotic mean* $\mathbb{P}^\star$ *and* Γ *is a random set satisfying condition* (**K**), *then* $\Gamma(T^i\omega)$ *is compact for infinitely many indices* i.

Proof. Since $\mathbb{P}^\star$ is stationary with respect to T, Poincaré's Recurrence Theorem implies that $\mathbb{P}^\star$−almost every point of A_0 is recurrent, whence infinitely recurrent. Thus, the following inclusion holds up to a $\mathbb{P}^\star$−null set:

$$A_0 \subseteq B_0 \stackrel{def}{=} \bigcap_{m\geq 0} \bigcup_{i\geq m} T^{-i}(A_0).$$

It is readily checked that B_0 is invariant, which implies $\mathbb{P}(B_0) = \mathbb{P}^\star(B_0)$. Further, $\mathbb{P}(B_0)$ is positive by condition (**K**). This entails $\mathbb{P}(B_0) = 1$, because T is ergodic with respect to $\mathbb{P}$. □

If the values of Γ are open it cannot be deduced that $\wedge(\Gamma)$ is open. This is true only when $\mathcal{A}$ is finite. However, it immediately follows from the definition that if there exists an open ball $B(x,r)$ such that $B(x,r) \subseteq \Gamma(\omega)$ $\mathbb{P}$−a.s., then $B(x,r) \subseteq \wedge(\Gamma)$. Conversely, when $\wedge(\Gamma)$ has a nonempty interior and Γ is closed-valued, the following proposition shows that $\Gamma(\omega)$ has a nonempty interior for almost all $\omega \in \Omega$. This result will serve in the proof of Theorem 4.1.

Proposition 3.8. *If the values of* Γ *are closed and if the open ball* $B(x,r)$ *is contained in* $\wedge(\Gamma)$, *then this ball is contained in* $\Gamma(\omega)$ *for* $\mathbb{P}$−*almost all* $\omega \in \Omega$.

Proof. It suffices to show that it is possible to find a null set N such that $B(x,r) \subseteq I(\Gamma, N)$. Let D be a dense countable subset of E. For all $y \in D \cap B(x,r)$ there exists a null set N_y such that $y \in I(\Gamma, N_y)$. Defining the null set N by

$$N = \bigcup_{y\in D\cap B(x,r)} N_y,$$

it is easily checked that

$$B(x,r) \subseteq \mathrm{cl}\big(D \cap B(x,r)\big) \subseteq I(\Gamma, N) \subseteq \wedge(\Gamma).$$

□

Remark 3.9 (A counterexample)**.** In Proposition 3.8, the closedness of the values of Γ is necessary as the following counterexample shows. Consider $E = \mathbb{R}$, $\Omega = (0,1)$ (the open unit interval) endowed with the Lebesgue measure $\mathbb{P}$, and the multifunction Γ defined by

$$\Gamma(\omega) = (0,1) \setminus \{\omega\} \qquad \omega \in \Omega.$$

Observe that the interval $(0,1)$ is equal to the open ball of radius $1/2$ centered at $1/2$ in E and, on the other hand, that the values of Γ are not closed. For every $N \in \mathcal{N}_{\mathbb{P}}$, one has $I(\Gamma, N) = N$, which yields $\wedge(\Gamma) = (0,1)$. Clearly, the inclusion $(0,1) \subseteq \Gamma(\omega)$ is false for all $\omega \in \Omega$.

Remark 3.10 (A more general definition)**.** (i) In the definition of the essential intersection, the set Ω can be replaced with any subset $A \in \mathcal{A}$ of positive measure. In this case, we use the notation

$$\wedge(\Gamma, A) = \bigcup_{N \in \mathcal{N}_{\mathbb{P}}(A)} \bigcap_{\omega \in A \setminus N} \Gamma(\omega). \tag{3.4}$$

where $\mathcal{N}_{\mathbb{P}}(A) = \{N \in \mathcal{N}_{\mathbb{P}} : N \subseteq A\}$ is the set of $\mathbb{P}-$null sets that are contained in A. Clearly, this defines the essential intersection of Γ with respect to the probability space $(A, \mathcal{A}_A, \mathbb{P}_A)$, where $\mathcal{A}_A = \{B \in \mathcal{A} : B \subseteq A\}$ and $\mathbb{P}_A$ is the restriction of $\mathbb{P}$ to $\mathcal{A}_A$, namely $\mathbb{P}_A(B) = \mathbb{P}(B)/\mathbb{P}(A)$ for all $B \in \mathcal{A}_A$. We say that $\wedge(\Gamma, A)$ is the essential intersection of Γ on A. A useful example is given by Proposition 3.11

(ii) Since $A \setminus N = A \setminus (A \cap N)$ for all $N \in \mathcal{N}_{\mathbb{P}}$, (3.4) admits the equivalent expression

$$\wedge(\Gamma, A) = \bigcup_{N \in \mathcal{N}_{\mathbb{P}}} \bigcap_{\omega \in A \setminus N} \Gamma(\omega). \tag{3.5}$$

(iii) The following implication immediately follows from (3.5)

$$A_1 \subseteq A_2 \qquad \Longrightarrow \qquad \wedge(\Gamma, A_2) \subseteq \wedge(\Gamma, A_1).$$

(iv) For example, if the random set Γ is single-valued, i.e. reduces to a random variable $X : \Omega \to E$ (by identifying $X(\omega)$ with the singleton $\{X(\omega)\}$), then given $A \in \mathcal{A}$, one has $\wedge(X, A) \neq \emptyset$ if and only if X is almost surely constant on A. If the constant is denoted by x_0 then $\wedge(X, A) = \{x_0\}$.

Assume that Ω is a metric space endowed with its Borel $\sigma-$field $\mathcal{A} = \mathcal{B}(\Omega)$ and that $A = S$, where $S = \operatorname{supp}(\mathbb{P})$ denotes the *support* of $\mathbb{P}$, namely the smallest closed set of Ω with full $\mathbb{P}-$measure. In this case, a more precise expression of the essential intersection of Γ can be given.

Proposition 3.11. *Under the above hypotheses, one has*

$$\wedge(\Gamma) = \wedge(\Gamma, S). \tag{3.6}$$

Proof. If N is a $\mathbb{P}-$null set, so is $N \cup S^c$, which implies that

$$\mathcal{N}_1 = \{N \cup S^c : N \in \mathcal{N}_{\mathbb{P}}\}$$

is a sufficient subfamily of $\mathcal{N}_{\mathbb{P}}$ (see Remark 3.4). Thus,

$$\wedge(\Gamma) \;=\; \bigcup_{M \in \mathcal{N}_1} \bigcap_{\omega \in \Omega \setminus M} \Gamma(\omega) = \bigcup_{N \in \mathcal{N}_{\mathbb{P}}} \bigcap_{\omega \in \Omega \setminus (N \cup S^c)} \Gamma(\omega)$$

$$= \bigcup_{N \in \mathcal{N}_{\mathbb{P}}} \bigcap_{\omega \in S \setminus N} \Gamma(\omega) = \wedge(\Gamma, S).$$

□

Example 3.12. Let E still denote a Polish space. Consider an $\mathcal{A} \otimes \mathcal{B}(E)-$function $f : \Omega \times E \to \overline{\mathbb{R}}$ and the set-valued map $\Delta : \omega \mapsto \text{epi}(f(\omega, \cdot))$ as in (2.4). Also consider the function $g : E \to \overline{\mathbb{R}}$ defined for every $x \in E$ as the essential supremum of the random variable $f(\cdot, x)$. It is known that $g(x)$ is the infimum of the set of real numbers α satisfying $\mathbb{P}(\{\omega \in \Omega : f(\omega, x) > \alpha\}) = 0$. This is denoted by $g(x) = \text{ess.sup f}(\cdot, \text{x})$. It is not difficult to check that the epigraph of g is the essential intersection of the random set Δ, namely

$$\wedge(\Delta) = \text{epi}(g). \tag{3.7}$$

Thus, the epigraph of the essential supremum is the essential intersection of the random set Δ. Also observe that $\wedge(\Delta) \neq \emptyset$ if and only if g is not identically $+\infty$. Further, if $\varphi(\omega, \cdot)$ is lsc for $\mathbb{P}-$almost all $\omega \in \Omega$, then the random set Δ is a.s. closed-valued. Thus, Proposition 3.5 (a) shows that $\wedge(\Delta)$ is closed, which implies that g is lsc. Similarly, by Proposition 3.5 (b), if E is a separable Banach space and if $\varphi(\omega, \cdot)$ is convex for $\mathbb{P}-$almost all $\omega \in \Omega$, then g is convex.

The following result concerns the inf-compactness of g. Recall that a function $\varphi : E \to \overline{\mathbb{R}}$ is said to be *inf-compact* if for every real β, the subset $\{x \in E : \varphi(x) \leq \beta\}$ is compact.

Proposition 3.13. *Let f be a random lsc function. If there exists $A_0 \in \mathcal{A}$ of positive measure such that $f(\omega, \cdot)$ is inf-compact for $\mathbb{P}-$almost all $\omega \in A_0$, then g is inf-compact. Further, if $(\Omega, \mathcal{A}, \mathbb{P}, T)$ is an ams ergodic dynamical system, then $f(T^i\omega, \cdot)$ is inf-compact for infinitely many indices i.*

Proof. For each real β consider, as at (2.5), the random set Γ_β given by

$$\Gamma_\beta(\omega) = \{x \in E : f(\omega, x) \leq \beta\}. \tag{3.8}$$

One has $g(x) \leq \beta$ if and only if there exists a $\mathbb{P}-$null set N such that

$$f(\omega, x) \leq \beta \qquad \forall \omega \in \Omega \setminus N$$

which yields

$$\wedge(\Gamma_\beta) = \{x \in E : g(x) \leq \beta\}. \tag{3.9}$$

In view of the hypothesis, Γ_β satisfies Condition (**K**) of Proposition 3.6. Thus, $\wedge(\Gamma_\beta)$ is compact, which implies the inf-compactness of g. The second statement follows from Proposition 3.7. □

Example 3.14. Let E be a separable Banach space. It is known (see, e.g., [1, Section 7.1]) and easy to check that a convex cone C such that $C \cap (-C) = \{0\}$ allows for defining an order relation $\leq_C$ on E, namely for any pair $(x, y) \in E^2$, one has

$$x \leq_C y \iff y - x \in C.$$

Consider the following two examples.

(i) Given an $E-$valued random variable X, define the random set Γ by

$$\Gamma(\omega) = X(\omega) + C \qquad \omega \in \Omega.$$

Then, it is not hard to see that the essential intersection of Γ is the set of upper bounds of X with respect to $\leq_C$ and that the following equivalence holds

$$\wedge(\Gamma) \neq \emptyset \quad \Longleftrightarrow \quad X \text{ is a.s. bounded from above with respect to } \leq_C .$$

Equivalently, one can find $x_0 \in E$ such that $X(\omega) \leq_C x_0$ for $\mathbb{P}-$almost all $\omega \in \Omega$. In particular, when $E = \mathbb{R}$ and $C = [0, +\infty)$, the order relation $\leq_C$ reduces to $\leq$, the usual order relation on the set of reals. Then, $\wedge(\Gamma)$ is nonempty if and only if X is a.s. bounded from above by some real x_0.

(ii) Given a subset $F \subseteq E$, a member x of F is said to be *maximal* with respect to $\leq_C$ if there exists no y in F such that $x \leq_C y$ and $y \neq x$ or, equivalently, if $F \cap C = \{x\}$. Further, let $\Theta : \Omega \to 2^E$ be a random set. It is not hard to check that x is maximal in $\wedge(\Theta)$ if and only if x is maximal in $\Theta(\omega)$ for $\mathbb{P}-$almost all $\omega \in \Omega$. This concept is useful when one studies equilibria in Mathematical Economics and Game Theory (see, e.g., [19], [2]).

3.2. **Essential intersection and image measure.** In many situations, a first random set, say Γ, is defined on some probability space and a second random set Δ is defined by composing Γ with a random variable Y. In such a situation, it is much useful to elucidate the connection between the essential intersection of Γ and that of Δ. More precisely, let $Y : \Omega \to E$ be a random variable defined on $(\Omega, \mathcal{A}, \mathbb{P})$ and $\Gamma : E \to 2^F$ be a random set defined on the measurable space $(E, \mathcal{B}(E))$, where F stands for another Polish space endowed with its Borel $\sigma-$field $\mathcal{B}(F)$. Also consider the composed random set Δ defined by

$$\Delta(\omega) = \Gamma(Y(\omega)) \qquad \omega \in \Omega. \tag{3.10}$$

Further, denote by $\mathbb{P}_Y$ the *image (or pushforward) measure* of $\mathbb{P}$ by Y. It is defined by

$$\mathbb{P}_Y(B) = \mathbb{P}(Y^{-1}(B)) \qquad B \in \mathcal{B}(E).$$

Recall that $Y^{-1}(B)$ denotes the *preimage* of B by Y, namely

$$Y^{-1}(B) = \{\omega \in \Omega : Y(\omega) \in B\}.$$

At this point, we need a technical hypothesis (**TH**) that reads as follows

(**TH**) For any $\mathbb{P}-$null set $M \subseteq \Omega$ one has

$$\mathbb{P}[Y^{-1}(Y(M))] = \mathbb{P}(M) = 0. \tag{3.11}$$

Remark 3.15. (i) Hypothesis (**TH**) can be also formulated by

$$M \in \mathcal{N}_{\mathbb{P}} \Longleftrightarrow Y(M) \in \mathcal{N}_{\mathbb{P}_Y}. \tag{3.12}$$

(ii) As to (3.11), observe that the inequality

$$\mathbb{P}[Y^{-1}(Y(M))] \geq \mathbb{P}(M). \tag{3.13}$$

is always true. It follows from the inclusion $M \subseteq Y^{-1}(Y(M))$. Similarly, in equivalence (3.12), the implication $\Longleftarrow$ is always true.

(iii) Hypothesis (**TH**) will hold if the random variable $Y : \Omega \to E$ is one-to-one (injective). This is not very restrictive, because if Y is not one-to-one it is possible to replace $(\Omega, \mathcal{A}, \mathbb{P})$ by another probability space $(\Omega', \mathcal{A}', \mathbb{P}')$ and $Y : \Omega \to E$ by another random variable $Y' : \Omega' \to E$ so that Y and Y' have the same distribution on $(E, \mathcal{B}(E))$, and Y' is one-to-one. Indeed, it is enough to choose Ω' as the quotient set of Ω with respect to the equivalence relation $\mathcal{R}$ defined by $\omega_1 \, \mathcal{R} \, \omega_2$ if and only if $Y(\omega_1) = Y(\omega_2)$. If we define class$(\omega)$ by

$$\text{class}(\omega) = \{\omega' \in \Omega : \omega \, \mathcal{R} \, \omega'\}$$

then the random variable Y' can be defined by

$$Y'(\text{class}(\omega)) = Y(\omega).$$

It follows from the definition that this defines Y' without ambiguity and that Y' is one-to-one.

The following result shows the connection between the essential intersection of Γ with respect to $(\Omega, \mathcal{A}, \mathbb{P})$ and the essential intersection of $\Delta = \Gamma \circ Y$ (defined at (3.10)) with respect to $(E, \mathcal{B}(E), \mathbb{P}_Y)$.

Theorem 3.16. *Under Hypothesis* (**TH**), *one has*

$$\wedge_{\mathbb{P}}(\Delta) = \wedge_{\mathbb{P}_Y}(\Gamma).$$

Proof. It is convenient to use the notation $\mathbb{Q} = \mathbb{P}_Y$. For each $M \in \mathcal{N}_{\mathbb{P}}$, one can find $N \in \mathcal{N}_{\mathbb{Q}}$ such that $M \subseteq Y^{-1}(N)$. Indeed, it is enough to choose $N = Y(M)$, which is a member of $\mathcal{N}_{\mathbb{Q}}$ by (**TH**) and Remark 3.15 (i). Thus, the family $Y(\mathcal{N}_{\mathbb{P}}) = \{Y(M) : M \in \mathcal{N}_{\mathbb{P}}\}$ is sufficient in $\mathcal{N}_{\mathbb{Q}}$ (see Remark 3.4), whence

$$\wedge_{\mathbb{P}}(\Delta) = \bigcup_{M \in \mathcal{N}_{\mathbb{P}}} \bigcap_{\omega \in \Omega \setminus M} \Delta(\omega) = \bigcup_{N \in \mathcal{N}_{\mathbb{Q}}} \bigcap_{\omega \in \Omega \setminus Y^{-1}(N)} \Gamma(Y(\omega)).$$

Letting $y = Y(\omega)$ we get

$$\wedge_{\mathbb{P}}(\Delta) = \bigcup_{N \in \mathcal{N}_{\mathbb{Q}}} \bigcap_{y \in E \setminus N} \Gamma(y) = \wedge_{\mathbb{Q}}(\Gamma).$$

□

4. A representation formula

In this section we present a representation formula for the essential intersection as well as some applications.

4.1. **Statement of the representation formula.** Given a random set $\Gamma : \Omega \to 2^E$, it follows from the definition that the (convex analysis) indicator function of $\wedge(\Gamma)$ evaluated at $x \in E$ is the expectation of $\chi_\Gamma(x)$ with respect to probability $\mathbb{P}$, namely

$$\chi_{\wedge(\Gamma)}(x) = \mathbb{E}\chi_\Gamma(\cdot)(x) = \int_\Omega \chi_{\Gamma(\omega)}(x)\mathbb{P}(\mathrm{d}\omega) \qquad x \in E. \tag{4.1}$$

The following result provides a representation formula for the essential intersection in connection with ams transformations. It shows that the essential intersection of Γ is almost surely equal to the random set Π_0 defined by

$$\Pi_0(\omega) = \bigcap_{i \geq 0} \Gamma(T^i\omega) \qquad \omega \in \Omega. \tag{4.2}$$

Theorem 4.1. *Let E be a separable Banach space, $(\Omega, \mathcal{A}, \mathbb{P}, T)$ be a dynamical system, $\mathcal{I}$ be the sub$-\sigma-$field of $T-$invariant subsets and $\Gamma : \Omega \to 2^E$ be a random set. Also assume the following conditions.*

(a) *On $(\Omega, \mathcal{A})$, the probability $\mathbb{P}$ is* ams *with respect to T, with stationary mean $\mathbb{P}^\star$,*
(b) *$\mathcal{I} = \{\Omega, \emptyset\}$ (i.e. T is ergodic),*
(c) *T is null-preserving,*
(d) *the values of Γ are nonempty, closed and convex,*
(e) *$\wedge_{\mathbb{P}}(\Gamma)$ (in short $\wedge(\Gamma)$) is nonempty.*

Under the above hypotheses the following equality holds

$$\wedge^\star(\Gamma) = \Pi_0(\omega) \qquad \mathbb{P} - a.s.\ (and\ \mathbb{P}^\star) - a.s. \tag{4.3}$$

where $\wedge^\star(\Gamma) = \wedge_{\mathbb{P}^\star}(\Gamma)$ denotes the essential intersection of Γ taken on $(\Omega, \mathcal{A}, \mathbb{P}^\star)$ and Π_0 is defined by (4.2).

Remark 4.2. Theorem 4.1 shows that the essential intersection of a random set Γ can be expressed as a countable intersection depending on ω and involving the values of Γ at $T^i\omega$ $(i \geq 0)$, where $T : \Omega \to \Omega$ is a given null-preserving measurable transformation. However, the values of the random set Π_0 may be difficult to evaluate, because this random set involves an infinite sequence of subsets. It is thus useful to look for an approximation of Π_0 by a finite intersection of the form $\Pi_m(\omega) = \bigcap_{i \geq m} \Gamma(T^i\omega)$. According to the results recalled in Section 2, the approximation holds in the sense of Painlevé-Kuratowski. More precisely, one has $\Pi_0(\omega) = PK - \lim_{m \to +\infty} \Pi_m(\omega)$ for $\mathbb{P}-$almost $\omega \in \Omega$. As recalled in Section 2 the convergence also holds in the sense of Hausdorff distance when Γ is compact-valued.

Remark 4.3. Following Remark 3.2 (i), when $\mathbb{P}$ and $\mathbb{P}^\star$ are equivalent, i.e. when $\mathcal{N}_{\mathbb{P}} = \mathcal{N}_{\mathbb{P}^\star}$, one has

$$\wedge^\star(\Gamma) = \wedge(\Gamma), \tag{4.4}$$

so that the representation formula is also valid for $\wedge(\Gamma)$. Equality (4.4) holds in particular when $\mathbb{P}$ is stationary with respect to T, because the probabilities $\mathbb{P}$ and $\mathbb{P}^\star$ coincide. In view of Condition (c), another situation where (4.4) holds is when T is invertible ([13, Corollary 6.3.2]). Otherwise, Equality (4.4) may be false. However, the inclusion $\mathcal{N}_{\mathbb{P}} \subseteq \mathcal{N}_{\mathbb{P}^\star}$ remains true, because T is assumed to be null-preserving. Thus, the inclusion $\wedge_{\mathbb{P}}(\Gamma) \subseteq \wedge_{\mathbb{P}^\star}(\Gamma)$ holds, which entails

$$\wedge_{\mathbb{P}}(\Gamma) \subseteq \Pi_0(\omega) \qquad \mathbb{P} - a.s.\ (and\ \mathbb{P}^\star) - a.s. \tag{4.5}$$

4.2. **Examples.** Let us explain how Theorem 4.1 can be applied in the situations of Examples 3.12 and 3.14.

Example 4.4. Continuation of Example 3.12. Assume that it is possible to apply Theorem 4.1 to the random set $\Delta : \omega \mapsto \text{epi}\big(f(\omega, \cdot)\big)$. In particular, assume that g is not identically $+\infty$ and that $f(\omega, \cdot)$ is lsc and convex for $\mathbb{P}-$almost all $\omega \in \Omega$. This yields, for $\mathbb{P}\,(\text{and}\,\mathbb{P}^\star)-$almost all $\omega \in \Omega$,

$$\text{epi}(g) = \bigcap_{m \geq 0} \Delta(T^m\omega) = \bigcap_{m \geq 0} \text{epi}\, f(T^m\omega, \cdot)$$

or, equivalently,

$$g(x) = \sup_{m\geq 0} f(T^m\omega, x) \qquad x \in E. \tag{4.6}$$

Remark 4.2 applied to the random set Δ shows that the function $x \mapsto g(x) = \sup_{m\geq 0} f(T^m\omega, x)$ can be approximated by a finite supremum of the following form

$$f_m(\omega, x) = \sup_{i\leq m} f(T^i\omega, x) \qquad x \in E \qquad m \geq 1.$$

The quality of the approximation will be good provided m is large enough.

Example 4.5. Continuation of Example 3.14 (i). Applying Theorem 4.1 to the random set Γ produces the following equalities, valid for $\mathbb{P}-$almost all $\omega \in \Omega$,

$$\wedge^\star(\Gamma) = \bigcap_{m\geq 0} \Gamma(T^m\omega) = \bigcap_{m\geq 0} \Big(X(T^m\omega) + C\Big).$$

The right-hand side involves the set of upper bounds of $X(T^m\omega)$ for $m \geq 0$ with respect to the order relation $\leq_C$. As in Example 4.4, $\Pi_m(\omega) = \bigcap_{i\leq m} \big(X(T^i\omega) + C\big)$ can provide a good approximation of $\wedge^\star(\Gamma)$.

4.3. **Proof of Theorem 4.1.** In the proof, D (resp. D') denotes a dense countable subset of E (resp. $\wedge^\star(\Gamma)$). In particular, we have $\wedge^\star(\Gamma) = \mathrm{cl}(D')$. We proceed in three steps. In the first two steps we assume that $\wedge(\Gamma)$ has a non empty interior. In Step 3, it is shown that this condition can be removed.

Step 1. Let us show that $\Pi_0(\omega)$ has a nonempty interior for $\mathbb{P}-$almost all $\omega \in \Omega$. By hypothesis, one can find an open ball $B(x, r)$ such that $B(x, r) \subseteq \wedge(\Gamma)$. Consider the set

$$A_0 = \{\omega \in \Omega : B(x, r) \subseteq \Gamma(\omega)\}.$$

It can be written as the countable intersection

$$A_0 = \bigcap_{y\in D\cap B(x,r)} \{\omega \in \Omega : y \in \Gamma(\omega)\} \tag{4.7}$$

because one has $B(x, r) \subseteq \mathrm{cl}(D \cap B(x, r)) \subseteq \Gamma(\omega)$ for all $\omega \in A_0$. Further, recalling that $\mathrm{Gr}(\Gamma) \in \mathcal{A} \otimes \mathcal{B}(E)$ and appealing to a standard result on product measurable spaces, it is readily seen that each set of the countable intersection in the righthand side of (4.7) belongs to $\mathcal{A}$, which proves $A_0 \in \mathcal{A}$. Also note that by Proposition 3.8, $\mathbb{P}(A_0) = 1$. Now, consider the set

$$B_0 = \{\omega \in \Omega : B(x, r) \subseteq \Pi_0(\omega)\}. \tag{4.8}$$

It is readily seen that

$$B_0 = \bigcap_{i\geq 0} T^{-i}(A_0)$$

whence

$$\mathbb{P}(B_0^c) = \mathbb{P}\Big(\bigcup_{i\geq 0} T^{-i}(A_0^c)\Big) \leq \sum_{i\geq 0} \mathbb{P}(T^{-i}(A_0^c)) = 0.$$

The rightmost equality holds because T is assumed to be null-preserving. Thus, $\mathbb{P}(B_0) = 1$ and one can find a $\mathbb{P}-$null set N_0 such that for all $\omega \in \Omega \setminus N_0$ the set $\Pi_0(\omega)$ has a nonempty interior.

Step 2. Consider $x \in E$. The random variable defined by $\omega \to \chi_{\Gamma(\omega)}(x)$ is nonnegative, whence quasi-integrable with respect to any probability measure. Appealing to hypotheses (a) and (b), it is possible to apply Theorem 2.1, which entails the existence of a $\mathbb{P}$ and $\mathbb{P}^\star-$null set N_x such that

$$\chi_{\wedge^\star(\Gamma)}(x) = \mathbb{E}^\star\big(\chi_\Gamma(\cdot)(x)\big) = \lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \chi_{\Gamma(T^i\omega)}(x) \qquad \omega \in \Omega \setminus N_x\,. \tag{4.9}$$

Let N_1 be the $\mathbb{P}-$null set defined by

$$N_1 = \bigcup_{x\in D} N_x\,.$$

Since the only possible values of the indicator functions are 0 and $+\infty$, the following equalities hold for all $\omega \in \Omega \setminus N_1$ and $x \in D$,

$$\chi_{\wedge^\star(\Gamma)}(x) = \sum_{i\geq 0} \chi_{\Gamma(T^i\omega)}(x) = \sup_{i\geq 0} \chi_{\Gamma(T^i\omega)}(x) = \chi_{\Pi_0(\omega)}(x). \tag{4.10}$$

Now, let us prove the inclusion $\Pi_0(\omega) \subseteq \wedge^\star(\Gamma)$ for all $\omega \in \Omega \setminus (N_0 \cup N_1)$. Consider $x \in D \cap \Pi_0(\omega)$. Since $\chi_{\Pi_0(\omega)}(x) = 0$, Equation (4.10) entails $x \in \wedge^\star(\Gamma)$. We thus have

$$D \cap \Pi_0(\omega) \subseteq \wedge^\star(\Gamma). \tag{4.11}$$

The left-hand side of (4.11) is nonempty, because $\Pi_0(\omega)$ has a nonempty interior by Step 1. By Proposition 3.5 (a) we know that $\wedge^\star(\Gamma)$ is closed. Further, since $\Pi_0(\omega)$ is a convex set with nonempty interior, taking the closure in both sides of (4.11) yields

$$\Pi_0(\omega) = \mathrm{cl}(D \cap \Pi_0(\omega)) \subseteq \wedge^\star(\Gamma).$$

Now, let us prove the converse inclusion, namely $\wedge^\star(\Gamma) \subseteq \Pi_0(\omega)$. Defining the null set N_2 by

$$N_2 = \bigcup_{x\in D'} N_x$$

and using again (4.10), it is readily seen that for all $x \in D'$ and $\omega \in \Omega \setminus N_2$ one has $x \in \Pi_0(\omega)$, which yields the inclusion

$$D' \subseteq \Pi_0(\omega). \tag{4.12}$$

Taking the closure in both sides of (4.12) yields the desired inclusion. Thus, we have shown that (4.3) holds for all $\omega \in \Omega \setminus (N_0 \cup N_1 \cup N_2)$.

Step 3. Now, let us prove that Equality (4.3) still holds when $\wedge(\Gamma)$ is no longer assumed to have a nonempty interior. For each integer $k \geq 1$ consider the random set Γ_k defined by

$$\Gamma_k(\omega) = \{x \in E : d(x, \Gamma(\omega)) \leq 1/k\}.$$

For all $\omega \in \Omega$, we have

$$\Gamma(\omega) = \bigcap_{k\geq 1} \Gamma_k(\omega). \tag{4.13}$$

By condition (e), $\wedge(\Gamma)$ is nonempty. For any member $x_0 \in \wedge(\Gamma)$, the inclusion $B(x_0, 1/k) \subseteq \Gamma_k(\omega)$ holds for $\mathbb{P}-$almost all $\omega \in \Omega$ and $k \geq 1$, which implies that the interior of $\wedge(\Gamma_k)$ is nonempty. By the result of Step 2, it is possible to construct a $\mathbb{P}-$null set N such that

$$\wedge^\star(\Gamma_k) = \bigcap_{i\geq 0} \Gamma_k(T^i\omega) \tag{4.14}$$

for all $\omega \in \Omega \setminus N$ and $k \geq 1$. In view of Remark 3.3 and Equation (4.13), taking the intersection over k on both sides of (4.14) yields

$$\wedge^\star(\Gamma) = \bigcap_{k\geq 1} \wedge^\star(\Gamma_k) = \bigcap_{k\geq 1}\bigcap_{i\geq 0} \Gamma_k(T^i\omega) = \bigcap_{i\geq 0}\bigcap_{k\geq 1} \Gamma_k(T^i\omega) = \bigcap_{i\geq 0} \Gamma(T^i\omega),$$

which ends the proof. □

Remark 4.6. Another version of Theorem 4.1 was already stated in [10, Theorem 2.7] as a possible application of the main result of that paper.[3] However, the hypotheses were different from those of Theorem 4.1 and we have realized that we were not able to prove this former version. Clearly the integrability hypothesis on Γ should be replaced with condition (e) of Theorem 4.1. Further, it seems difficult to remove the convexity hypothesis on Γ and the null-preserving hypothesis on T.

5. Robust stochastic optimization

This section is devoted to robust stochastic optimization problems. First, we discuss some formulations often encountered in the literature and give a significant correction to one of them. Then, using the representation formula of Theorem 4.1, we provide several approximation results for robust stochastic optimization problem.

5.1. **Mathematical formulation.** Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a $\mathcal{B}(E)-$measurable function $h : E \to \mathbb{R}$ and a random lsc function, namely an $\mathcal{A} \otimes \mathcal{B}(E)-$measurable function $f : \Omega \times E \to \mathbb{R}$ such that $f(\omega, \cdot)$ is lsc for $\mathbb{P}-$almost all $\omega \in \Omega$. In the following we examine formulations of what is usually known in the literature as a *robust program* (see, e.g., [6]). Consider the following stochastic optimization program:

(O) Minimize $h(x)$ subject to

$$f(\omega, x) \leq 0 \qquad\qquad \mathbb{P} - a.s.$$

If we define the random set Γ by

$$\Gamma(\omega) = \{x \in E : f(\omega, x) \leq 0\} \qquad\qquad \omega \in \Omega, \tag{5.1}$$

which is a special case of (3.8) (with $\beta = 0$), then the above program admits the following equivalent formulation:

(O′) Minimize $h(x)$ subject to $x \in \wedge(\Gamma)$, where $\wedge(\Gamma)$ was given by (3.9).

Example 5.1. Continuation of Example 3.1. Assume that $\Omega = E = [0, 1]$, and consider the function f defined by $f(\omega, x) = 0$ if $\omega \neq x$ and $f(\omega, x) = 1$ if $\omega = x$. The random set Γ defined by (5.1) satifies $\Gamma(\omega) = E \setminus \{\omega\}$, for all $\omega \in \Omega$. This is a special case of Example 3.1. We thus have $I(\Gamma, N) = N$ for all $N \in \mathcal{N}_\mathbb{P}$ and

[3] By the same authors as this one.

$\wedge(\Gamma) = E$, which shows that the constraints set given by $I(\Gamma, N)$ may be much smaller than that given by $\wedge(\Gamma)$.

Remark 5.2. Sometimes, Problem (O) or (O′) is formulated in the literature as follows:

(O″) Minimize $h(x)$ subject to

$$f(\omega, x) \leq 0 \qquad\qquad \forall\, \omega \in \Omega.$$

It is worth observing that, in the above example, the set of constraints of Problem (O″) is given by $I(\Gamma, \emptyset) = \emptyset$. Hence, the solution of (O″) is trivial and, of course, of little interest. Consequently, in a probabilistic framework, the notion of essential intersection appears to be quite useful for providing relevant and precise formulations of robust optimization programs, as in (O′).

5.2. **Approximation results.** The robust program given by (O) or (O′) is often difficult to solve in practice, because an infinite number of constraints is involved. This is why it is customary to approximate it with a simplified version in which the set of constraints is replaced with a simpler one obtained through random sampling or deterministic approximation. Let $\{\omega_i\}_{i=0}^{n-1}$ be a finite sequence of points of Ω drawn according to the probability measure $\mathbb{P}$. These points can be members of either a random sample or a deterministic point-set. They can also be obtained through the observation of real situations (e.g., in hydrology or climatology), so that it is reasonable to suppose that each value exhibits some sort of dependence upon the past. It is therefore relevant to consider the situation where the sequence $\{\omega_i\}_{i=0}^{n-1}$ is obtained through the iteration of a transformation T (i.e. $\omega_i = T^i\omega_0$). The resulting optimization problem is sometimes called a *scenario program* (see [8]), since every sequence $\{\omega_i\}_{i=0}^{n-1}$ corresponds to a particular scenario:

(O_n) Minimize $h\,(x)$ subject to

$$f\,(\omega_i, x) \leq 0 \qquad i = 0, \ldots, n-1.$$

An alternative equivalent formulation is:

(O'_n) Minimize $h\,(x)$ subject to

$$x \in \bigcap_{i=0}^{n-1} \Gamma\,(\omega_i)$$

where $\Gamma\,(\omega_i) = \{x \in E : f\,(\omega_i, x) \leq 0\}$.

As already seen, Theorem 4.1 implies that for $\mathbb{P}$ and $\mathbb{P}^\star-$almost all $\omega_0 \in \Omega$

$$\wedge^\star\,(\Gamma) = \bigcap_{i \geq 0} \Gamma\,(\omega_i)$$

where $\omega_i = T^i\omega_0$. The above formula suggests that the finite intersection

$$\Pi_m(\omega_0) = \bigcap_{i \leq m} \Gamma(\omega_i) \tag{5.2}$$

can give a consistent approximation of $\wedge^\star(\Gamma)$ provided m is large enough. Indeed, this approximation holds in the sense of $PK-$ convergence or even in that of Hausdorff distance when the constraints sets are compact (see Remark 4.2).

Theorem 5.3 hereafter is the main result of this section. It shows a typical situation where the solution of problem (O) or (O′) (where $\mathbb{P}$ is replaced with $\mathbb{P}^\star$) can be approximated by solutions of problems like (O_n) or ($\mathrm{O'}_n$). To make it more precise, it is convenient to introduce the following optimization problems for all $\omega \in \Omega$ and all nonnegative integers m, namely

$$\big(O_m(\omega)\big) : \text{ minimize } h(x) \text{ under } x \in \Pi_m(\omega)$$

where $\Pi_m(\omega)$ is defined at (5.2).

Theorem 5.3. *Let $(\Omega, \mathcal{A}, \mathbb{P}, T)$ be a dynamical system, E be a separable Banach space, $h : E \to \overline{\mathbb{R}}$ and $f : \Omega \times E \to \overline{\mathbb{R}}$ be extended-real-valued functions. Also assume the following conditions:*

(a) *On $(\Omega, \mathcal{A})$, the probability $\mathbb{P}$ is* ams *with respect to T, with stationary mean $\mathbb{P}^\star$,*
(b) *$\mathcal{I} = \{\Omega, \emptyset\}$, that is, T is ergodic,*
(c) *T is null-preserving,*
(d) *f is $\mathcal{A} \otimes \mathcal{B}(E)-$measurable and $f(\omega, \cdot)$ is* lsc *and convex for $\mathbb{P}-$almost all $\omega \in \Omega$,*
(e) *the function $x \to g(x) = \operatorname{ess.sup} \mathrm{f}(\cdot, \mathrm{x})$ satisfies $L(g, 0) \neq \emptyset$ (see (2.3)),*
(f) *there exists $A_0 \in \mathcal{A}$ of positive measure such that $f(\omega, \cdot)$ is inf-compact for $\mathbb{P}-$almost all $\omega \in A_0$,*
(g) *h is* lsc*,*
(h) *Problem (O′) admits a unique solution $\overline{x}$, namely*

$$\overline{x} \in \arg\min h \qquad \textit{subject to} \qquad x \in \wedge^\star(\Gamma) \, .$$

Under the above conditions, for $\mathbb{P}$ and $\mathbb{P}^\star-$almost all $\omega \in \Omega$ and each sequence $(\overline{x}_m)_{m\geq 1}$ such that $\overline{x}_m$ is a solution of $\big(O_m(\omega)\big)$ for each m, one has $\overline{x}_m \to \overline{x}$ and $h(\overline{x}_m) \to h(\overline{x})$.

The following lemma, of a purely deterministic nature, serves for proving Theorem 5.3. It concerns the convergence of infima and minimizers of a given function h when the minimization of h is carried out on a subset C_n, where the sequence $(C_n)_{n\geq 0}$ is non increasing.

Lemma 5.4. *Let E be a metric space, $h : E \to \overline{\mathbb{R}}$ be a* lsc *extended-real-valued function, $(C_n)_{n\geq 1}$ be a nonincreasing sequence of nonempty closed subsets of E and let C_∞ denote*

$$C_\infty = \bigcap_{n\geq 1} C_n.$$

Consider the following optimization problems

$$(O_\infty) \qquad \textit{minimize } h(x) \textit{ subject to } x \in C_\infty$$

and for each $n \geq 1$

$$(O_n) \qquad \textit{minimize } h(x) \textit{ subject to } x \in C_n \, .$$

Also consider a sequence $(\overline{x}_n)_{n\geq 1}$ such that, for each $n \geq 1$, $\overline{x}_n$ is a solution of problem (O_n), namely $h(\overline{x}_n) = \inf_{x\in C_n} h(x)$.

(i) If $\overline{x}_\infty$ *is a cluster point of* $(\overline{x}_n)_{n\geq 1}$*, then* $\overline{x}_\infty$ *is a solution of* (O_∞) *and*

$$h(\overline{x}_\infty) = \lim_{n\to+\infty} h(\overline{x}_n). \tag{5.3}$$

(ii) In addition, if problem (O_∞) *admits only one solution, then the whole sequence* $(\overline{x}_n)$ *converges to* $\overline{x}_\infty$*.*

Remark 5.5. There will exist a cluster point for the sequence $(\overline{x}_n)$ if at least one of the sets C_n is compact. Since the sequence is non increasing, this allows for a finite number of C_n's to be non compact.

Proof of Theorem 5.3. Consider the random set Γ defined by (5.1). Condition (d) shows that the values of Γ are closed and convex $\mathbb{P}-$a.s. From (3.9) we know that

$$\wedge(\Gamma) = \{x \in E : g(x) \leq 0\}.$$

Condition (e) implies that $\wedge(\Gamma)$ is nonempty. Consequently, according to Theorem 4.1, it is possible to find a $\mathbb{P}$ and $\mathbb{P}^\star-$null set N_1 such that the following equality holds for all $\omega \in \Omega \setminus N_1$

$$\wedge^\star(\Gamma) = \bigcap_{m\geq 0} \Gamma(T^m\omega).$$

or equivalently

$$\wedge^\star(\Gamma) = \Big\{x \in E : \sup_{m\geq 0} f(T^m\omega, x) \leq 0\Big\}.$$

Further, by Condition (f) there exists $A_0 \in \mathcal{A}$ of positive measure such that $\Gamma(\omega)$ is compact for $\mathbb{P}-$almost all $\omega \in A_0$. This makes possible to apply Proposition 3.7 to Γ, which shows that $\Gamma(T^i\omega)$ is compact for infinitely many indices $i \geq 0$ and $\mathbb{P}-$almost all $\omega \in \Omega$. Therefore, one can find a $\mathbb{P}-$null set N_2 with the following property: for all $\omega \in \Omega \setminus N_2$ there exists an integer $m_0(\omega)$ such that $\Pi_m(\omega) = \bigcap_{i\leq m} \Gamma(T^i\omega)$ is compact for all $m \geq m_0(\omega)$ and for all $\omega \in \Omega \setminus N_2$. Setting $N = N_1 \cup N_2$ it only remains to apply Lemma 5.4 (ii) and Remark 5.5 for each $\omega \in \Omega \setminus N$ to the sequence (C_n) defined by $C_n = \Pi_n(\omega)$ $(n \geq 1)$ and to $C_\infty = \wedge^\star(\Gamma)$. □

Proof of Lemma 5.4 One can find a subsequence $(\overline{x}_{n_k})_{k\geq 1}$ of $(\overline{x}_n)$ such that $\overline{x}_\infty = \lim_{k\to+\infty} \overline{x}_{n_k}$. Observe that $\overline{x}_\infty$ is a member of C_∞, because the limit of a sequence is not changed by removing a finite number of terms and because the sequence (C_n) is nonincreasing. Consequently, one has

$$h(\overline{x}_\infty) \leq \liminf_{k\to+\infty} h(\overline{x}_{n_k}) = \lim_{n\to+\infty} h(\overline{x}_n) = \lim_{n\to+\infty} \inf_{x\in C_n} h(x) \leq \inf_{x\in C_\infty} h(x) \leq h(\overline{x}_\infty).$$

The first inequality holds because h is lsc and the second one holds because the sequence $(h(\overline{x}_n))$ is nondecreasing. The above relationships imply that $\overline{x}_\infty$ is a solution of (O_∞) and

$$h(\overline{x}_\infty) = \lim_{n\to+\infty} h(\overline{x}_n)$$

which proves part (i). Part (ii) follows from a well-known result on compact metric spaces. □

Remark 5.6. (i) If Condition (f) of Theorem 5.3 is strengthened by assuming that $\mathbb{P}(A_0) = 1$, namely that $f(\omega, \cdot)$ is inf-compact for $\mathbb{P}-$almost all $\omega \in \Omega$, then, $f(T^n\omega, \cdot)$ enjoys the same property for all $n \geq 0$. It follows that the integer $m_0(\omega)$ appearing in the proof of Theorem 5.3 can be taken equal to 1 $\mathbb{P}-$a.s.

(ii) If h is inf-compact, then Condition (f) of Theorem 5.3 can be removed. The constraint sets only need to be closed and convex.

(iii) If Problem (O') is no longer assumed to have only one solution, but if the sequence $(\overline{x}_n)$ has a cluster point $\overline{x}$, then we can still assert that $\overline{x}$ is a solution of (O') and that $h(\overline{x}_m) \to h(\overline{x})$.

The following corollary deals with the case of sequences of measurable selections that are minimizers for $\mathbb{P}-$almost all $\omega \in \Omega$. Recall that the existence of such measurable selections can be proved by standard arguments (see, e.g., [14] and the references therein).

Corollary 5.7. *Under the same hypotheses as in Theorem 5.3, for all sequences* $\overline{f}_n : \Omega \to E$ *of measurable functions such that for* $\mathbb{P}-$*almost all* $\omega \in \Omega$ *and for all* $n \geq 1$, $\overline{f}_n(\omega)$ *is a minimizer of* h *subject to*

$$\overline{f}_n(\omega) \in \bigcap_{i=0}^{n-1} \Gamma\left(T^i\omega\right)$$

one has $\overline{f}_n(\omega) \to \overline{x}$ *and* $h\left(\overline{f}_n(\omega)\right) \to h(\overline{x})$.

Proof. Appeal again to Lemma 5.4 (ii) and Remark 5.5. □

The next result shows that if the objective function h is assumed to be convex, then it is possible to weaken the inf-compactness condition, which is only assumed to hold in the weak topology of E.

Theorem 5.8. *Let* $(\Omega, \mathcal{A}, \mathbb{P}, T)$ *be a dynamical system,* E *be a separable Banach space,* $h : E \to \overline{\mathbb{R}}$ *and* $f : \Omega \times E \to \overline{\mathbb{R}}$ *be extended-real-valued functions. Also assume the same conditions (a) to (h) of Theorem 5.3 apart from Conditions (f) and (g) that are replaced with:*

(f') *there exists* $A_0 \in \mathcal{A}$ *of positive measure such that* $f(\omega, \cdot)$ *is weakly inf-compact for* $\mathbb{P}-$*almost all* $\omega \in A_0$,

(g') h *is* lsc *and convex.*

Then, for $\mathbb{P}-$*almost all* $\omega \in \Omega$ *and for all sequences* $(\overline{x}_n)$ *such that*

$$\overline{x}_n \in \arg\min h \qquad \text{subject to} \qquad x \in \bigcap_{i=0}^{n-1} \Gamma\left(T^i\omega\right). \tag{5.4}$$

one has $\overline{x}_n \to \overline{x}$ *in the weak topology and* $h(\overline{x}_n) \to h(\overline{x})$.

Proof. We only sketch the proof. It is based on a variant of Lemma 5.4 when E is a Banach space and h is assumed to be lsc and convex, whence lsc in the weak topology. Further, using the weak compactness assumption of Condition (f) and appealing to Eberlein's Theorem, it is possible to extract a weakly converging subsequence. The rest of the proof is like that of Theorem 5.3. □

A version of Corollary 5.7 for the weak topology could be also stated in the same lines and is left to the reader.

5.3. **The case of i.i.d. sequences.** Apart from the formulation we have used so far, another one exists in Ergodic Theory: a sequence of random variables $X_0, X_1, \dots$ is said to be *stationary* if the random vectors $(X_0, \dots, X_n)$ and $(X_k, \dots, X_{n+k})$ have the same distribution for all integers $n, k \geq 1$. Those formulations are in some sense equivalent. More precisely, given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a measure-preserving transformation T and a random variable X, the sequence $(X_n)_{n\geq 0}$ defined by $X_n(\omega) = X(T^n\omega)$ is stationary. Conversely, given a stationary sequence $(X_n)_{n\geq 0}$ it is possible to construct another sequence $(\hat{X}_n)_{n\geq 0}$, another probability space $(\Omega', \mathbb{A}', \mathbb{P}')$ and a measure preserving transformation $T' : \Omega' \to \Omega'$ such that (X_n) and $(\hat{X}_n)$ have the same distribution (see, e.g., [7, Proposition 6.11]). Consequently, Theorems 4.1 and 5.3 could be rewritten in the setting of stationary sequences. The purpose of this subsection is to examine briefly the special case of i.i.d. sequences, which is often encountered in applications. The extension to the case of pairwise i.i.d. sequences is briefly addressed in Remark 5.13.

The distribution of a random set can be defined like that of any random variable. If we regard a random set as a measurable map $\Gamma : \Omega \to 2^E$, it is necessary to introduce a $\sigma-$field on 2^E. A convenient and popular one is the so-called *Effros–σ–field* $\mathcal{E}$ (see, e.g., [3], [15] or [22]), which is generated by the following subsets of 2^E

$$\mathcal{F}^- U = \{F \in 2^E : F \cap U \neq \emptyset\}$$

where U ranges over the set $\mathcal{U}$ of all subsets of E. This corresponds to the definition given in Section 2 (statement (b) of Proposition 2.3). Indeed, a random set Γ satisfies $\Gamma^- U = \Gamma^{-1}(\mathcal{F}^- U) \in \mathcal{A}$ for all $U \in \mathcal{U}$. The *distribution of* Γ is the probability $\mathbb{P}_\Gamma$ defined on $(2^E, \mathcal{E})$ by

$$\mathbb{P}_\Gamma(\mathcal{F}^- U) = \mathbb{P}(\Gamma^{-1}(\mathcal{F}^- U)) = \mathbb{P}(\Gamma^- U)$$

for all $U \in \mathcal{U}$.

For a short introduction to the distribution of random sets and their independence, we refer the reader to [15] or [22]. We only recall a useful criterion for random closed sets.

Proposition 5.9. *Let E be a Polish space. Given two random sets Γ_1 and Γ_2 with closed values in E, the following three statements are equivalent:*

(i) *Γ_1 and Γ_2 have the same distribution.*

(ii) *For all open subsets U of E, $\mathbb{P}(\Gamma_1^- U) = \mathbb{P}(\Gamma_2^- U)$.*

(iii) *For any finite subset Y of E (or of some dense countable subset) the $\mathbb{R}^k-$valued random vectors $(d(y, \Gamma_1))_{y\in Y}$ and $(d(y, \Gamma_2))_{y\in Y}$ have the same distribution (where the distance function was defined at (2.1)).*

Similarly, two random sets Γ_1 and Γ_2 are said to be *independent* if the equality

$$\mathbb{P}_{(\Gamma_1, \Gamma_2)} = \mathbb{P}_{\Gamma_1} \otimes \mathbb{P}_{\Gamma_2}$$

holds on the product measurable space $(2^E \times 2^E, \mathcal{E} \otimes \mathcal{E})$, where (Γ_1, Γ_2) denotes the map $\omega \to (\Gamma_1(\omega), \Gamma_2(\omega))$.

We present two results that are a transposition to the i.i.d. case of Theorems 4.1 and 5.3. The first one provides a version of the representation formula (4.3).

Theorem 5.10. *Let E be a separable Banach space and $\Gamma : \Omega \to 2^E$ a random set satisfying the following conditions.*

(i) *The values of Γ are nonempty, closed and convex.*

(ii) *$\wedge(\Gamma)$ is nonempty.*

Then, given an i.i.d. sequence $(\Gamma_n)_{n\geq 0}$ of random sets having the same distribution as Γ, the following equality holds for $\mathbb{P}-$almost all $\omega \in \Omega$

$$\wedge(\Gamma) = \Pi_0(\omega)$$

where Π_0 is the random set defined by $\Pi_0(\omega) = \bigcap_{n\geq 0} \Gamma_n(\omega)$.

Sketch of the Proof. It is enough to observe as above that an i.i.d. sequence is a special case of a stationary sequence. □

Remark 5.11. It is interesting to mention that a direct proof of Theorem 5.10, very similar to that of Theorem 4.1, can be given as well. Now, the analogue of Equality (4.9) can be obtained by applying the strong law of large numbers to the sequence of indicator functions $\chi_{\Gamma_i}(x)$ $(i \geq 0)$, where $x \in E$ is fixed. This gives

$$\chi_{\wedge(\Gamma)}(x) = \mathbb{E}\big(\chi_{\Gamma}(\cdot)(x)\big) = \lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_{\Gamma_i(\omega)}(x) \qquad \omega \in \Omega \setminus N_x \tag{5.5}$$

where N_x denotes a suitable $\mathbb{P}-$null set (generally depending on x).

In order to state the i.i.d. version of Theorem 5.3 on the convergence of minimizers, we consider once more a random convex lsc function $f : \Omega \times E :\to \overline{\mathbb{R}}$ and the random closed set Γ defined by $\Gamma(\omega) = \{x \in E : f(\omega, x) \leq 0\}$. The distribution of a random lsc function can be defined via the random set $\omega \mapsto \mathrm{epi} f(\omega, \cdot)$ and a criterion like Proposition 5.9 can be derived. Also consider an i.i.d. sequence $(f_n)_{n\geq 0}$ of random convex lsc functions having the same distribution as f and the random closed sets Γ_n defined by

$$\Gamma_n(\omega) = L(f_n(\omega, \cdot), 0) = \{x \in E : f_n(\omega, x) \leq 0\}.$$

As already seen, the following optimization problems $\big(O_m(\omega)\big)$ are defined for all $\omega \in \Omega$ and all nonnegative integers $m \geq 0$

$$\big(O_m(\omega)\big) : \text{ minimize } h(x) \text{ under } x \in \Pi_m(\omega)$$

where $\Pi_m(\omega)$ is defined this time by $\Pi_m(\omega) = \bigcap_{n\leq m} \Gamma_n(\omega)$.

Theorem 5.12. *Let E be a separable Banach space, $h : E \to \overline{\mathbb{R}}$ and $f : \Omega \times E \to \overline{\mathbb{R}}$ be extended-real-valued functions. Also assume the following conditions:.*

(i) *f is $\mathcal{A} \otimes \mathcal{B}(E)-$measurable and $f(\omega, \cdot)$ is lsc and convex for $\mathbb{P}-$almost all $\omega \in \Omega$,*

(ii) *the function $x \to g(x) = \mathrm{ess.\,sup\,f(\cdot, x)}$ satisfies $L(g, 0) \neq \emptyset$,*

(iii) *there exists $A_0 \in \mathcal{A}$ of positive measure such that $f(\omega, \cdot)$ is inf-compact for $\mathbb{P}-$almost all $\omega \in A_0$,*

(iv) *h is lower semi-continuous,*

(v) *Problem (O') admits a unique solution $\overline{x}$, namely*

$$\overline{x} \in \arg\min h \qquad \textit{subject to} \qquad x \in \wedge(\Gamma)\,.$$

Also consider an i.i.d. sequence $(f_n)_{n\geq 0}$ *of convex* lsc *functions having the same distribution as* f. *Under the above conditions, for* $\mathbb{P}-$*almost all* $\omega \in \Omega$ *and each sequence* $(\overline{x}_m)_{m\geq 1}$ *such that* $\overline{x}_m$ *is a solution of* $\big(O_m(\omega)\big)$ *for each* m, *one has* $\overline{x}_m \to \overline{x}$ *and* $h\left(\overline{x}_m\right) \to h\left(\overline{x}\right)$.

Remark 5.13. Other versions of Theorems 5.10 and 5.12 can be proved using the Etemadi strong law of large numbers for pairwise independent and identically distributed random variables (see [11]). In this case, where mutual independence is replaced with pairwise independence, it is no longer possible to appeal to Birkhoff's Ergodic Theorem. However, a direct proof can be given by mimicking that of Theorem 4.1 and by observing that, by Theorem 1 of [11], Equality (5.5) also holds for pairwise i.i.d. sequences of indicator functions.

Remark 5.14. Approximation results similar to those of this section are given in [28] (Theorem 4.1 and Corollary 4.4). The formulation is different, but it can be observed that the compactness assumptions are stronger than those of the present paper. Further, in view of Remark 4.6 it seems that in Corollary 4.4 of [28] the integrability condition (Hypothesis (iii)) is unnecessary.

References

[1] C. Aliprantis and K. Border, *Infinite-Dimensional Analysis*, Springer-Verlag, Berlin, 1999.

[2] R. J. Aumann and L. S. Shapley, *Values of Non-Atomic Games*, Princeton University Press, Princeton, 1974.

[3] G. Beer, *Topologies on Closed and Closed Convex Sets, Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, 1993.

[4] A. Ben-Tal, L. El Ghaoui and A. Nemirovski, *Robust Optimization*, Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2009.

[5] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Math. Oper. Res. **23** (1998), 769–805.

[6] D. Bai, T. Carpenter and J. Mulvey, *Making a case for robust models*, Management Science **43** (1997), 895–907.

[7] L. Breiman, *Probability*, vol. 7 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992, corrected reprint of the 1968 original.

[8] G. Calafiore and M.C. Campi, *Uncertain convex programs: randomized solutions and confidence levels*, Math. Program., Ser. A **102** (2005), 25–46.

[9] C. Choirat, C. Hess and R. Seri, *A functional version of the Birkhoff ergodic theorem for a normal integrand: a variational approach*, Tech. Rep. No. 0101, Cahiers du Centre de Recherche Viabilité-Jeux-Contrôle, Université Paris Dauphine (2001).

[10] C. Choirat, C. Hess and R. Seri, *A functional version of the Birkhoff ergodic theorem for a normal integrand: a variational approach*, Ann. Probab. **31** (2003), 63–92.

[11] N. Etemadi, *An elementary proof of the strong law of large numbers*, Z. Wahrschein. **55** (1981), 119–122.

[12] R. M. Gray and J. Kieffer, *Asymptotically mean stationary measures*, Ann. Probab. **8** (1980), 962–973.

[13] R. M. Gray, *Probability, Random Processes and Ergodic Properties*, Springer-Verlag, New York.1987.

[14] C. Hess, *Epi-convergence of sequences of normal integrands and strong consistency of the maximum likelihood estimator*, Ann. Statist. **24** (1996), 1298–1315.

[15] C. Hess, *From the distribution of unbounded random sets to the multivalued strong law of large numbers in nonreflexive Banach spaces*, J. Convex Anal. **6** (1999), 163–182.

[16] C. Hess, *Set-valued integration and set-valued probability theory: an overview*, in Handbook of Measure Theory, Endre Pap (edi.), Elsevier Science B.V. North Holland, 2002, pp. 617–673. edited by .
[17] C. Hess, R. Seri and C. Choirat, *Ergodic theorems for extended real-valued random variables*, Stochastic Processes Appl. **120** (2010), 1908–1919.
[18] R. Hettich and K.O. Kortanek, *Semi-infinite programming: theory, methods, and applications*, SIAM Rev. **35** (1993), 380–429.
[19] W. Hildenbrand, *Core and Equilibria of a Large Economy*, Princeton University Press, Princeton, 1974.
[20] J.-B. Hiriart-Urruty, *Contribution à la programmation mathématique : cas déterministe et stochastique*, Ph.D. thesis, Université de Clermont-Ferrand II, 1977.
[21] U. Krengel, *Ergodic theorems*, vol. 6 of de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1985.
[22] I. Molchanov, *Theory of Random Sets*, Springer Verlag, London, 2005.
[23] E. Polak, *On the mathematical foundations of nondifferentiable optimization in engineering design*, SIAM Rev. **29** (1987), 21–89.
[24] R. Reemtsen, *Discretization methods for the solution of semi-infinite programming problems*, J. Optim. Theory Appl. **71** (1991), 85–103.
[25] R. Reemtsen, *Some outer approximation methods for semi-infinite optimization problems*, J. Comput. Appl. Math. **53** (1994), 87–108.
[26] R. Reemtsen, *Semi-infinite Programming: Discretization Methods*, Encyclopedia of Optimization, 2009, pp. 3417–3424.
[27] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
[28] R. Seri and C. Choirat, *Scenario approximation of robust and chance-constrained programs*, J. Optim. Theory Appl. **158** (2013), 590–614.
[29] G. Still, *Discretization in semi-infinite programming: the rate of convergence*, Math. Program. Ser. A **91** (2001), 53–69.
[30] S. Žakovic and B. Rustem, *Semi-infinite programming and applications to minimax problems. Applied mathematical programming and modelling*, Ann. Oper. Res. **124** (2003), 81–110.

Manuscript received ,
revised ,

Christian Hess
Département MIDO, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 PARIS CEDEX 16, France
E-mail address: Christian.Hess@dauphine.fr

Raffaello Seri
Dipartimento di Economia, Università degli Studi dell'Insubria, via Monte Generoso 71, 21100 Varese, Italy
E-mail address: raffaello.seri@uninsubria.it

Christine Choirat
IQSS Harvard University, 1737 Cambridge Street, CGIS Knafel Building, Cambridge, MA 02138, USA
E-mail address: cchoirat@iq.harvard.edu