Approximations of upper and lower probabilities by measurable selections[*]

Enrique Miranda[†], Inés Couso and Pedro Gil

Department of Statistics and Operations Research
University of Oviedo

Abstract

A random set can be regarded as the result of the imprecise observation of a random variable. Following this interpretation, we study to which extent the upper and lower probabilities induced by the random set keep all the information about the values of the probability distribution of the random variable. We link this problem to the existence of selectors of a multi-valued mapping and with the inner approximations of the upper probability, and prove that under fairly general conditions (although not in all cases), the upper and lower probabilities are an adequate tool for modelling the available information. In doing this, we generalise a number of results from the literature. Finally, we study the particular case of consonant random sets and we also derive a relationship between Aumann and Choquet integrals.

Keywords: Random sets, Dempster-Shafer upper and lower probabilities, imprecise probabilities, core, Aumann integral, Choquet integral.

1 Introduction

Random sets, or multi-valued mappings, have been used by several authors in the context of imprecise or incomplete information [45]. They have also been applied in many different contexts, such as image analysis [21, 34], medicine [13] or statistics with coarse data [24]. In this paper, we follow the interpretation given to random sets by Kruse and Meyer [33] and we regard them as the imprecise observation of a measurable mapping.

---

[*]This is an updated version, with proofs and additional comments, of a paper [42] presented at ECSQARU'09, the Tenth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty.

[†]Corresponding author. C-Calvo Sotelo, s/n, 33007 Oviedo, Spain. Tel: (+34) 985102955, Fax: (+34) 985103354. e-mail addresses: (mirandaenrique, couso, pedro) @uniovi.es

Dempster [15] summarised the probabilistic information of the random set by means of the upper and lower probabilities, which constitute a generalisation to a context of imprecise information of the concept of probability induced by a random variable. The upper and lower probabilities of a random set are plausibility and belief functions in the context of evidence theory [47], and capacities of infinite order under Choquet's terminology [8]. This type of set functions has been thoroughly studied in the literature [9, 44], not only in the context of multi-valued mappings [16, 22, 47], as a powerful alternative to probability measures that is able to deal with uncertain, or vague, knowledge. Nevertheless, under Kruse and Meyer's interpretation, the most precise piece of information that the random set gives about the measurable mapping imprecisely observed is the set of the probability distributions of the measurable selections. The relationship between this set and Dempster-Shafer upper and lower probabilities has already been studied by some authors (see for instance [7, 11, 23, 26]). In this paper we investigate some additional aspects of this relationship. Let us introduce first some concepts and notation.

Consider a probability space $(\Omega, \mathcal{A}, P)$, a measurable space $(X, \mathcal{A}')$, and a measurable mapping $U_0 : \Omega \to X$. We will refer to U_0 as the **original random variable**. There may be some imprecision in the observation of the values of U_0. Following Kruse and Meyer ([33]), a possible model for this situation is to consider a multi-valued mapping $\Gamma : \Omega \to \mathcal{P}(X)$, in the sense that for all $\omega \in \Omega$ we are certain that $U_0(\omega)$ belongs to the set $\Gamma(\omega)$. As a consequence, we shall assume throughout that $\Gamma(\omega)$ is non-empty for all ω. We will call the multi-valued mapping **closed** (resp., compact, open, complete) when $\Gamma(\omega)$ is a closed (resp., compact, open, complete) subset of X for all $\omega \in \Omega$.

Let us define

$$S(\Gamma) := \{U : \Omega \to X \mid U \text{ measurable}, U(\omega) \in \Gamma(\omega) \ \forall \omega\}. \tag{1}$$

This is the class of the **measurable selections** of Γ. Taking into account our interpretation of Γ, all we know about U_0 is that it is one of the elements of $S(\Gamma)$. Concerning the probability distribution of U_0, it will belong to

$$\mathcal{P}(\Gamma) = \{P_U \mid U \in S(\Gamma)\}, \tag{2}$$

the class of the probability distributions induced on $\mathcal{A}'$ by the measurable selections. In particular, the probability that the value of U_0 belongs to an event $A \in \mathcal{A}'$, i.e. $P_{U_0}(A)$, is an element of $\mathcal{P}(\Gamma)(A) := \{P_U(A) \mid U \in S(\Gamma)\}$.

Hence, the interpretation of a multi-valued mapping as a model for the imprecise observation of a random variable provides us with a Bayesian sensitivity

2

analysis model for the probability distribution of this variable: the set of probability distributions $\mathcal{P}(\Gamma)$. There is, however, another set of probabilities that shall also be interesting for our purposes. It is based on the notions of upper and lower probabilities induced by multi-valued mapping.

Definition 1. *[15, 44] Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a multi-valued mapping. Given $A \in \mathcal{A}'$, its **upper inverse** by Γ is $\Gamma^*(A) := \{\omega \mid \Gamma(\omega) \cap A \neq \emptyset\}$, and its **lower inverse**, $\Gamma_*(A) := \{\omega \mid \Gamma(\omega) \subseteq A\}$.*

Following Nguyen [44], a multi-valued mapping is said to be **strongly measurable** when $\Gamma^*(A) \in \mathcal{A}$ for all $A \in \mathcal{A}'$. In that case, we will refer to Γ as **random set**. Taking into account the relationship $\Gamma_*(A) = (\Gamma^*(A^c))^c$, valid for all $A \in \mathcal{A}'$, a random set satisfies $\Gamma_*(A) \in \mathcal{A}$ for all $A \in \mathcal{A}'$. We will use the shorter notation $A^* = \Gamma^*(A)$ and $A_* = \Gamma_*(A)$ when no confusion arises. Although there are other measurability conditions for multi-valued mappings (see for instance [27]), we shall only consider in this paper the strong measurability; this condition is necessary if we want to define the upper and lower probabilities of the random set on $\mathcal{A}'$, as we see next.

Definition 2. *[15] Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a non-empty random set. The **upper probability** induced by Γ on $\mathcal{A}'$ is defined as $P^*(A) = P(A^*) \ \forall A \in \mathcal{A}'$, and the **lower probability** is given by $P_*(A) = P(A_*) \ \forall A \in \mathcal{A}'$.*

The upper probability of a random set is ∞-alternating and lower continuous, and the lower probability is ∞-monotone and upper continuous [44]. They are moreover conjugate functions, meaning that $P^*(A) = 1 - P_*(A^c) \ \forall A \in \mathcal{A}'$. If the final space is finite, they are a plausibility and a belief function, respectively. We shall sometimes use the notation $P^*_\Gamma := P^*$ and $P_{*\Gamma} := P_*$, if there is ambiguity about the random set inducing the upper and lower probabilities. It is easy to see that $A_* \subseteq U^{-1}(A) \subseteq A^*$ for every $A \in \mathcal{A}'$ and every $U \in S(\Gamma)$. This implies that the class $\mathcal{P}(\Gamma)$ defined in Eq. (2) is included in

$$M(P^*) = \{Q : \mathcal{A}' \to [0, 1] \text{ probability s.t. } Q(A) \leq P^*(A) \ \forall A\}, \qquad (3)$$

which is called the **core** of P^*.

The upper probability of a random set generalises the concept of probability distribution of a random variable, and is sometimes used as a model of the probabilistic information of the random set [17, 19, 35, 43]. In this paper, we shall investigate if it is appropriate to do so when Γ has the interpretation

3

considered by Kruse and Meyer. Specifically, we are going to study under which conditions we can use the upper probability to represent the information about the probability that our original random variable takes values in some arbitrary set $A \in \mathcal{A}'$. In other words, we shall investigate under which conditions the equality $\mathcal{P}(\Gamma)(A) = [P_*(A), P^*(A)]$ holds. This is important because, as we shall show, when these two sets are not equal the use of the upper and the lower probability could carry some serious loss of information.

The study of the equality $\mathcal{P}(\Gamma)(A) = [P_*(A), P^*(A)]$ can be split into two different subproblems: on the one hand, we need to study the convexity of the set $\mathcal{P}(\Gamma)(A)$; and on the other, we need to determine whether the supremum and infimum values of this set coincide with the upper and lower probabilities of A, respectively. Because of the duality existing between P^* and P_*, it suffices to study one of the two equalities.

Although this problem has already been investigated by some authors ([4, 23, 26]), this has always been done as a support for other mathematical considerations, and hence the sufficient conditions established for the equalities $P^*(A) = \sup \mathcal{P}(\Gamma)(A)$ and $P_*(A) = \inf \mathcal{P}(\Gamma)(A)$ assume some hypotheses on the random set that are not really necessary for the equalities to hold. We shall see nevertheless that the problem is not trivial, and we shall extend some of the established results.

In our treatment, we shall try to be as general as possible. As a consequence, we shall not assume the completeness of the initial probability space or the closedness of its images. This is not unusual in a context of imprecise information [1, 2, 30].

As far as we know, the most important result on this problem is the following:

Theorem 1. *[10, Prop. 3] Let $(\Omega, \mathcal{A}, P)$ be a probability space, (X, τ) a Polish space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a compact random set. Then, $P^*(A) = \sup \mathcal{P}(\Gamma)(A)$ for every A in β_X, the Borel σ-field associated to τ.*

Recall here that a *Polish* space is a separable and completely metrizable topological space. We shall also use later *Souslin* spaces, which are the images of Polish spaces by continuous mappings.

We shall study the equality $\mathcal{P}(\Gamma)(A) = [P_*(A), P^*(A)]$ in detail in our next section. First, we will investigate under which conditions $P^*(A)$ and $P_*(A)$ are, respectively, the supremum and infimum values of $\mathcal{P}(\Gamma)(A)$. For this, we shall use some results on the existence of measurable selections and on the inner approximations of the upper probability. Secondly, we will study the convexity of $\mathcal{P}(\Gamma)(A)$. Finally, in Section 3 we shall show some of the consequences of our

results.

Remark 1. *It is important to remark here that, due to our interpretation of random sets as a model for the imprecise observation of a random variable, in our developments we shall always assume that: (i) the initial probability space is fixed, and (ii) that the measurable selections have to take images in Γ for all the elements on the initial space. This makes our setting more restrictive than other works in the literature, where also* almost-surely *selections are considered (i.e., selections where the image is included in the random set with probability one), and where the initial space is not considered important, using then the so-called* allocations *of probability [45]. It is not difficult to show that the set $\mathcal{P}(\Gamma)$ in Eq. (2) is included in the set of allocations of the upper probability P^*.*

On the other hand, this also implies that the sufficient conditions we will establish for the approximations of upper and lower probabilities by distributions of measurable selections are also valid under these more general contexts. ♦

2 $[P_*(A), P^*(A)]$ as a model of $P_{U_0}(A)$

2.1 Study of the equality $P^*(A) = \sup \mathcal{P}(\Gamma)(A)$

Let us study first if the upper and lower probabilities of an event A are the most precise bounds of $P_{U_0}(A)$ that we can give, taking into account the information given by Γ. As we shall show in the following example, this is not always the case: in fact, it may happen that $[P_*(A), P^*(A)] = [0, 1]$ while $\mathcal{P}(\Gamma)(A) = \{0\}$. In such an extreme case, the set of the distributions of the measurable selections would provide precise information, while the upper and lower probabilities would give no information at all. The example we give is based on [28, Example 5]; the differences are that Himmelberg et al. consider a weaker notion of measurability and give an example of a multi-valued mapping which is measurable in that sense and for which $S(\Gamma) = \emptyset$.

Example 1. *Let P be an absolutely continuous probability measure on $(\mathbb{R}, \beta_{\mathbb{R}})$, and let $P^{\mathbb{N}}$ be the probability measure on $\beta_{\mathbb{R}^n}$ determined by the product of countably many copies of P. Consider $\Omega := \{F \subseteq \mathbb{R} \text{ countable}\}$ and $\mathcal{A} := \sigma(\{\mathcal{F}_A \mid A \in \beta_{\mathbb{R}}\})$, with $\mathcal{F}_A := \{F \in \Omega \mid F \cap A \neq \emptyset\}$. Let us define the mapping $g : \mathbb{R}^{\mathbb{N}} \to \Omega$ by $g(z) = \{x \in \mathbb{R} \mid x = z_n \text{ for some } n\}$. Then, $g^{-1}(\mathcal{F}_A) = \bigcup_n \left(\prod_{i=1}^{n-1} \mathbb{R} \times A \times \prod_{i>n} \mathbb{R} \right) \in \beta_{\mathbb{R}^{\mathbb{N}}}$, so g is a measurable mapping. Let us denote by Q the probability measure it induces on $\mathcal{A}$. Consider now the*

multi-valued mapping

$$\Gamma : \Omega \;\rightarrow\; \mathcal{P}(\mathbb{R})$$
$$F \;\hookrightarrow\; F \cup \{0\}$$

For every $A \in \beta_{\mathbb{R}}$, it follows that $\Gamma^*(A) = \Omega$ if $0 \in A$, and $\Gamma^*(A) = \mathcal{F}_A$ otherwise. Hence, Γ is strongly measurable. Consider $B = \mathbb{R} \setminus \{0\}$. Then $\Gamma^*(B) = \mathcal{F}_B$, whence $P^*(B) = Q(\mathcal{F}_B) = P^{\mathbb{N}}(g^{-1}(\mathcal{F}_B)) = P^{\mathbb{N}}(\cup_n(\prod_{i=1}^{n-1}\mathbb{R} \times B \times \prod_{i>n}\mathbb{R})) = 1 - P^{\mathbb{N}}(\{0,0,0,\dots\}) = 1$, taking into account that $P^{\mathbb{N}}$ is the product of an infinite number of copies of a continuous probability measure.

Now, if $U \in S(\Gamma)$ satisfies $P_U(B) > 0$, U is also a measurable selection of the multi-valued mapping $\Gamma_1 : \Omega \to \mathcal{P}(\mathbb{R})$ given by $\Gamma_1(F) = F$ if $F \in U^{-1}(B)$, $\Gamma_1(F) = F \cup \{0\}$ otherwise. However, reasoning as in [28, Example 5], it can be checked that Γ_1 does not have measurable selections. As a consequence, $\mathcal{P}(\Gamma) = \{\delta_0\}$. Hence, $\mathcal{P}(\Gamma)(B) = \{0\}$ and $[P_*(B), P^*(B)] = [0, 1]$. ♦

This example shows that the use of the upper and lower probabilities may carry some serious loss of information. Assume for instance that the continuous probability P considered in the example is the Lebesgue measure on $[0, 1]$, extended in the trivial way to $\beta_{\mathbb{R}}$: $P(A) = P(A \cap [0, 1])$ for all $A \in \beta_{\mathbb{R}}$. If we use P^*, P_* to model the expectation of the original random variable, by means of the corresponding Choquet integrals, we obtain the interval $[0, 1]$ as the set of possible expectations; however, we know, because $\mathcal{P}(\Gamma) = \{\delta_0\}$, that the expectation of the original random variable is 0.

Hence, it is necessary to consider some additional hypotheses in the random set if we want to guarantee that the upper probability of a set A, $P^*(A)$, is the supremum of the set $\mathcal{P}(\Gamma)(A)$ of its probability values by the measurable selections. In our next result, we prove that the supremum of $\mathcal{P}(\Gamma)(A)$ is indeed a maximum:

Proposition 2. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ be a measurable space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. Then, $\mathcal{P}(\Gamma)(A)$ has a maximum and a minimum value for every $A \in \mathcal{A}'$.*

Proof: Let us show that $\mathcal{P}(\Gamma)(A)$ has a maximum value; the proof of the second part follows using the duality between the upper and lower probabilities. Consider $\alpha = \sup \mathcal{P}(\Gamma)(A)$. For all $n \in \mathbb{N}$, there is some $V_n \in S(\Gamma)$ s.t. $\alpha - P_{V_n}(A) < \frac{1}{n}$. Take $D_1 = V_1^{-1}(A), D_n = V_n^{-1}(A) \setminus [\cup_{i=1}^{n-1} D_i] \; \forall n \geq 2$, and define $U_1 = \sum_{n=1}^{\infty} V_n I_{D_n} + V_1 I_{(\cup_{n=1}^{\infty} D_n)^c}$. The map U_1 is a measurable selection of Γ and satisfies $P_{U_1}(A) \geq \alpha$, whence this supremum is indeed a maximum. ∎

Let us now define

$$\mathcal{H}_\Gamma := \{A \in \mathcal{A}' \mid P^*(A) = \max \mathcal{P}(\Gamma)(A)\}. \tag{4}$$

We can then rephrase our goal in this section by stating that we are interested in providing conditions for the equality between $\mathcal{H}_\Gamma$ and $\mathcal{A}'$. To see that they do not coincide in general, check that $\mathcal{H}_\Gamma := \{B \in \beta_\mathbb{R} : P(B) = 0 \text{ or } 0 \in B\}$ in Example 1. It is also easy to modify the example in order to obtain a random set without measurable selections. In that case the class $\mathcal{H}_\Gamma$ would be empty.

In the following proposition, we state that a set $A \in \mathcal{A}'$ belongs to the class $\mathcal{H}_\Gamma$ given by Eq.(4) if and only if some random set that we can derive from Γ has measurable selections. The proof of the result is straightforward, and is therefore omitted: it suffices to take into account that a measurable set A belongs to $\mathcal{H}_\Gamma$ if and only if there is a measurable selection of Γ taking values in A whenever possible (up to a set of zero probability).

Proposition 3. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. A set $A \in \mathcal{A}'$ belongs to $\mathcal{H}_\Gamma$ if and only if there is some $H \in \mathcal{A}$ with $P(H) = 0$ such that $S(\Gamma_{A,H}) \neq \emptyset$, where*
$$\Gamma_{A,H}(\omega) = \begin{cases} \Gamma(\omega) \cap A & \text{if } \omega \in A^* \setminus H \\ \Gamma(\omega) & \text{otherwise.} \end{cases}$$

In the sequel, we shall use the notation Γ_A for $\Gamma_{A,\emptyset}$, i.e, Γ_A is the random set given by $\Gamma_A(\omega) = \begin{cases} \Gamma(\omega) \cap A & \text{if } \omega \in A^* \\ \Gamma(\omega) & \text{otherwise.} \end{cases}$

Proposition 3 shall be useful later on when studying which sets belong to $\mathcal{H}_\Gamma$, because the existence of measurable selections of a random set is one of the most important problems in this framework, and there are therefore many results that may be applicable together with this proposition; see the survey on the existence of measurable selections by Wagner [48]. Together with some of the results in [48], we shall prove and use another sufficient condition for the existence of measurable selections. For this, we must introduce the notion of *reducible σ-field*.

Definition 3. *Consider a measurable space $(X, \mathcal{A}')$. Given $x \in X$, we define the* **minimal measurable set** *generated by x as $[x] := \bigcap\{A \in \mathcal{A}' \mid x \in A\}$. The σ-field $\mathcal{A}'$ is called* **reducible** *when $[x] \in \mathcal{A}'$ for all $x \in X$, and $(X, \mathcal{A}')$ is called then a* **reducible measurable space**.

We can easily see that most σ-fields in our context are reducible: for instance, the Borel σ-field generated by a T_1 topology (and hence also by a metric) is

reducible, because in that case we have $[x] = \{x\}$ $\forall x \in X$. To see that the notion is not trivial, we give next an example of a non-reducible σ-field:

Example 2. *Let $\leq$ be a well-order on the set of real numbers (existing by Zermelo's theorem), and let $\mathcal{P}_x^< := \{y \in \mathbb{R} \mid y < x\}$ and $\mathcal{P}_x^\leq := \{y \in \mathbb{R} \mid y \leq x\}$ denote the sets of strict predecessors and predecessors of x under $\leq$, respectively. Let us also define the notation $\mathcal{P}_x^\geq := (\mathcal{P}_x^<)^c$.*

There is some $x_0 \in \mathbb{R}$ such that $\mathcal{P}_{x_0}^<$ uncountable and such that $\mathcal{P}_x^<$ is countable for any $x < x_0$: it suffices to take the set of points with an uncountable number of predecessors, and select its first element, existing because $\leq$ is a well-order. Consider $X := \mathcal{P}_{x_0}^\leq$, and let us define $\mathcal{B} := \{\emptyset, \ A \subseteq \mathcal{P}_{x_0}^< \ countable\} \cup \{\mathcal{P}_x^\geq \cup A \mid A \subseteq \mathcal{P}_x^<, x \in \mathcal{P}_{x_0}^<\}$.

- *Given a countable set $A \subseteq \mathcal{P}_{x_0}^<$, $\sup A$ always exists, because $\leq$ is a well-order, and it belongs to $\mathcal{P}_{x_0}^<$. Taking this into account, we can deduce that $\mathcal{B}$ is closed under complementation. Since it is immediate that it is closed under countable unions and that $\emptyset, X$ belong to $\mathcal{B}$, we deduce that $\mathcal{B}$ is a σ-field.*

- *Note now that x_0 does not have a previous element under the order $\leq$: otherwise, we contradict the uncountability of $\mathcal{P}_{x_0}^<$. Hence, the minimal measurable set generated by x_0 is $[x_0] = \cap_{x<x_0}\mathcal{P}_x^\geq = \{x_0\}$. This set does not belong to $\mathcal{B}$ and as a consequence this σ-field is not reducible.* ♦

We have already mentioned that a random set may not possess measurable selections, and that we need to make some requirements in order to guarantee that the set $S(\Gamma)$ given by Eq. (1) is non-empty. The existing results usually make some assumptions on the images of the random set and on the structure of the final σ-field. In our next result, we give a sufficient condition for the existence of measurable selections where the only thing we require in $\mathcal{A}'$ is its reducibility, which, as we have already said, is a very weak assumption:

Lemma 4. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a reducible measurable space, and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. If there is some countable $\{x_n\}_n \subseteq X$ s.t. $\cup_n \Gamma^*([x_n]) = \Omega$, then $S(\Gamma) \neq \emptyset$.*

Proof: Let us denote $A_1 = \Gamma^*([x_1]), A_n = \Gamma^*([x_n]) \setminus \cup_{i=1}^{n-1}\Gamma^*([x_i])$ $\forall n \geq 2$, and let us define $\Gamma' := \sum_n (\Gamma \cap [x_n])I_{A_n}$. Then, $\{A_n\}_n$ constitutes a measurable partition of Ω, and satisfies moreover $A_n \subseteq \Gamma^*([x_n])$ $\forall n$. Hence, $\Gamma'(\omega) \neq \emptyset$ $\forall \omega$. Apply the axiom of choice and take $U : \Omega \to X$ satisfying $U(\omega) \in \Gamma'(\omega) \subseteq \Gamma(\omega)$ $\forall \omega$. Taking into account that the minimal measurable sets are pairwise

disjoint, we deduce that given $\omega_1 \in A_n, \omega_2 \in A_m, n \neq m$, $\Gamma'(\omega_1) \cap \Gamma'(\omega_2) = \emptyset$. This implies that $U^{-1}(B)$ is a countable union of elements of $\{A_n\}_n \subseteq \mathcal{A}$ for every $B \in \mathcal{A}'$. Hence, U is measurable, and consequently $S(\Gamma) \neq \emptyset$. $\blacksquare$

Proposition 5. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a reducible measurable space, and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set s.t. $S(\Gamma) \neq \emptyset$. For any countable subset $\{x_n\}_n$ of X, $\cup_n [x_n] \in \mathcal{H}_\Gamma$.*

Proof: Let U be a measurable selection of Γ, and define $A_1 := [x_1]^*$, $A_n := [x_n]^* \setminus (\cup_{i=1}^{n-1} [x_i]^*)$ for any $n \geq 2$. Consider the multi-valued mapping

$$\Gamma' := \sum_n (\Gamma \cap [x_n]) I_{A_n} + U I_{(\cup_n A_n)^c}.$$

Let U_1 be a mapping satisfying $U_1(\omega) \in \Gamma'(\omega)$ for all ω. Such a mapping exists from the axiom of choice. It follows that U_1 is a measurable selection of Γ and that $U_1^{-1}(\cup_n [x_n]) = \cup_n A_n = \cup_n [x_n]^*$, whence $P_U(\cup_n [x_n]) = P^*(\cup_n [x_n])$. $\blacksquare$

We turn now to another property of random sets that shall be useful in our quest for sufficient conditions for the equality between $\mathcal{P}(\Gamma)(A)$ and the interval $[P_*(A), P^*(A)]$: the existence of inner approximations of P^*. We shall investigate under which hypotheses there is some subclass $\mathcal{A}'_1$ of $\mathcal{A}'$ such that P^* is the inner set function of its restriction to $\mathcal{A}'_1$. The interest of this problem for our purposes lies in the following proposition:

Proposition 6. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. If $B \in \mathcal{A}'$ satisfies $P^*(B) = \sup_n P^*(A_n)$ for some increasing sequence $\{A_n\}_n \subseteq \mathcal{H}_\Gamma$ of subsets of B, then $B \in \mathcal{H}_\Gamma$.*

Proof: We deduce from the hypotheses and from the monotonicity of a probability measure that

$$P^*(B) = \sup_n P^*(A_n) = \sup_n \sup_{U \in S(\Gamma)} P_U(A_n)$$
$$= \sup_{U \in S(\Gamma)} P_U(\cup_n A_n) \leq \sup_{U \in S(\Gamma)} P_U(B) \leq P^*(B).$$

Hence, $P^*(B) = \sup \mathcal{P}(\Gamma)(B)$, and applying Proposition 2 we deduce that B belongs to $\mathcal{H}_\Gamma$. $\blacksquare$

We deduce that if P^* satisfies

$$P^*(A) = \sup_{B \subseteq A, B \in \mathcal{H}_\Gamma} P^*(B) \ \forall A \in \mathcal{A}', \tag{5}$$

9

it also satisfies $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for every $A \in \mathcal{A}'$ (i.e., $\mathcal{H}_\Gamma$ is actually equal to $\mathcal{A}'$). This will be helpful for our purposes because in some cases it will be easier to prove the equality $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for some specific types of sets, such as closed or compact sets, and to show then that the upper probability can be approximated from below from these sets. In particular, Proposition 6 and the lower continuity of P^* implies that $\mathcal{H}_\Gamma$ is closed under countable unions.

In the language of measure theory, Eq. (5) means that P^* is the *inner set function* of its restriction to $\mathcal{H}_\Gamma$, or that it is *inner regular* with respect to $\mathcal{H}_\Gamma$. There are some results about the inner regularity of upper probabilities in the literature (see [7, 35]). In this respect, we have proven the following:

Lemma 7. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, (X, τ) a Polish space and consider a closed random set $\Gamma : \Omega \to \mathcal{P}(X)$. For every $A \in \beta_X$, $P^*(A) = \sup_{K \subseteq A \; compact} P^*(K)$.*

Proof: From [31, Theorem 30.13], the result follows if we show that for any compact subset K of X, $P^*(K) = \inf_{K \subseteq G \; open} P^*(G)$.

- Let us consider first of all the case of $X = \mathbb{R}^\mathbb{N}$. Let $\Pi_n : \mathbb{R}^\mathbb{N} \to \mathbb{R}^n$, $\pi_n^m : \mathbb{R}^m \to \mathbb{R}^n$, $m \geq n$ denote the projection operators, and let K be a compact subset of $\mathbb{R}^\mathbb{N}$. Π_n is continuous for every n, whence $K_n := \Pi_n(K) \subseteq \mathbb{R}^n$ is compact. As a consequence, $K_n' := \overline{B(K_n; \frac{1}{n})} = \{x \in \mathbb{R}^n \mid d(x, K_n) \leq \frac{1}{n}\}$ is also a compact subset of $\mathbb{R}^n$, because it is closed and bounded. Consider $K_n'' := K_n' \times \mathbb{R} \times \dots$, and let us prove that $P^*(K) = \inf_n P^*(K_n'')$. It suffices to show that $K^* = \cap_n (K_n'')^*$.

  Given $\omega \notin K^*$, $\Gamma(\omega) \cap K = \emptyset \Rightarrow K \subseteq \Gamma(\omega)^c$ open. For any $x \in K$, there is $G_x = G_x' \times \mathbb{R} \times \dots$, with $G_x' \subseteq \mathbb{R}^n$ open for some $n \in \mathbb{N}$, s.t. $x \in G_x \subseteq \Gamma(\omega)^c$. Hence, $K \subseteq \cup_{x \in K} G_x \subseteq \Gamma(\omega)^c$, whence there are $x_1, \dots, x_m \in K$ such that $K \subseteq G_{x_1} \cup \dots \cup G_{x_m} = (G_{x_1}' \times \mathbb{R} \times \dots) \cup \dots \cup (G_{x_m}' \times \mathbb{R} \times \dots) \subseteq \Gamma(\omega)^c$, and as a consequence there are $j \in \mathbb{N}, G \subseteq \mathbb{R}^j$ open such that $K \subseteq G \times \mathbb{R} \times \dots \subseteq \Gamma(\omega)^c$.

  For this j, there must be some $l \in \mathbb{N}$ s.t. $\overline{B(K_j; \frac{1}{l})} \subseteq G$: otherwise, $\{\overline{B(K_j; \frac{1}{l})} \cap G^c\}_l$ would be a decreasing sequence of non-empty compact sets on a Hausdorff space; then, their global intersection would be non-empty, and there would exist $x \in \cap_l (\overline{B(K_j; \frac{1}{l})} \cap G^c) = K_j \cap G^c$. But $K \subseteq G \times \mathbb{R} \times \dots$ implies that $K_j = \Pi_j(K) \subseteq \Pi_j(G \times \mathbb{R} \times \dots) = G$, a contradiction.

  Therefore, there is $l \in \mathbb{N}$, (which we may assume $l \geq j$) s.t. $\overline{B(K_j; \frac{1}{l})} \subseteq G$. It follows easily that $\overline{B(K_l; \frac{1}{l})} \times \mathbb{R} \times \dots \subseteq G \times \mathbb{R} \times \dots \subseteq \Gamma(\omega)^c$, and as

a consequence $K_l'' \cap \Gamma(\omega) = \emptyset$ and $K^* = \cap_n (K_n'')^*$. This means that $P^*(K) = \inf_n P^*(K_n'')$. Now, for every $n \in \mathbb{N}$, $K \subseteq B(K_n; \frac{1}{n}) \times \mathbb{R} \times \cdots \subseteq K_n''$, whence $P^*(K) = \inf_n P^*(B(K_n; \frac{1}{n}) \times \mathbb{R} \times \dots)$. Hence, $P^*(K) = \inf_{K \subseteq G \text{ open}} P^*(G)$. Applying now [31, Theorem 30.13], we deduce that $P^*(A) = \sup_{K \subseteq A \text{ compact}} P^*(K)$.

- Secondly, if X is a closed subset of $\mathbb{R}^{\mathbb{N}}$, the closed (resp., compact) subsets of X are also closed (resp., compact) in $\mathbb{R}^{\mathbb{N}}$. In particular, Γ is also closed-valued in $\mathbb{R}^{\mathbb{N}}$. Hence, given $A \in \beta_X \subseteq \beta_{\mathbb{R}^{\mathbb{N}}}$, $P^*(A) = \sup\{P^*(K) \mid K \subseteq A, K \in \beta_{\mathbb{R}^{\mathbb{N}}} \text{ compact}\} = \sup\{P^*(K) \mid K \subseteq A, K \in \beta_X \text{ compact}\}$.

- Consider finally an arbitrary Polish space X. From [31, Theorem 4.17], it is homeomorphic to a closed subset of $\mathbb{R}^{\mathbb{N}}$. Let us denote this set C, and let $f : X \to C$ be the homeomorphism. Consider $\Gamma' := f \circ \Gamma : \Omega \to \mathcal{P}(C)$.

  - Given $A \in \beta_C$, $\Gamma'^*(A) = \{\omega \mid f(\Gamma(\omega)) \cap A \neq \emptyset\} = \{\omega \mid \Gamma(\omega) \cap f^{-1}(A) \neq \emptyset\} = \Gamma^*(f^{-1}(A)) \in \mathcal{A}$. Hence, Γ' is strongly measurable.

  - Taking into account that Γ is a closed random set and f is a closed mapping, we deduce that Γ' is a closed random set on a closed subset of $\mathbb{R}^{\mathbb{N}}$. Applying the previous part, $P_{\Gamma'}^*(A) = \sup_{K \subseteq A \text{ compact}} P_{\Gamma'}^*(K)$ for every $A \in \beta_C$.

  - Given $A \in \beta_X$, $f(A)$ belongs to β_C. From the previous point, it follows that $P_{\Gamma'}^*(f(A)) = \sup_{K \subseteq f(A) \text{ compact}} P_{\Gamma'}^*(K)$, whence given $\epsilon > 0$ there is some compact set $K \subseteq f(A)$ such that $P_{\Gamma'}^*(f(A)) - P_{\Gamma'}^*(K) < \epsilon$. This implies that $P_{\Gamma}^*(A) - P_{\Gamma}^*(f^{-1}(K)) < \epsilon$, and $f^{-1}(K) \subseteq A$ is compact because f^{-1} is continuous. Hence, $P^*(A) = \sup_{K \subseteq A \text{ compact}} P^*(K)$. This completes the proof. $\blacksquare$

This lemma generalises a result in [35, Section 2.1]. Let us establish now sufficient conditions for the equality between $\mathcal{H}_\Gamma$ and $\mathcal{A}'$.

Theorem 8. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$, a random set. Under any of the following conditions:*

1. *Ω is complete, X is Souslin and $Gr(\Gamma) \in \mathcal{A} \otimes \beta_X$*

2. *X is a separable metric space and Γ is compact*

3. *X is a Polish space and Γ is closed*

4. *X is a σ-compact metric space and Γ is closed*

11

5. X is a separable metric space and Γ is open

6. $\mathcal{C}_\Gamma := \{\Gamma^*(B) : B \in \mathcal{A}'\}$ is countable and $\mathcal{A}'$ is reducible

7. $\mathcal{A}'$ is reducible and Γ has a countable range,

$P^*(A) = \max \mathcal{P}(\Gamma)(A)$ and $P_*(A) = \min \mathcal{P}(\Gamma)(A) \ \forall A \in \mathcal{A}'$.

Proof: Consider $A \in \mathcal{A}'$, and let us show that under any of these conditions, $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ (the equality $P_*(A) = \min \mathcal{P}(\Gamma)(A)$ will follow using conjugacy).

1. The graph of the multi-valued mapping Γ_A belongs to $\mathcal{A} \otimes \beta_X$ when $Gr(\Gamma)$ does, and, from [27, Theorem 5.2], Γ_A has a measurable selection under these conditions. Applying Proposition 3, we deduce that $A \in \mathcal{H}_\Gamma$.

2,3,4. The result in these three cases follows from Proposition 6 if we show that i) given a compact set K, the random set Γ_K possesses measurable selections, whence the compact sets belong to $\mathcal{H}_\Gamma$; and ii) given $A \in \beta_X$, $P^*(A) = \sup_{K \subseteq A \text{ compact}} P^*(A)$.

    - For point 2, these two facts follow from [6, Theorem III.8] and [7, Lemma A3].

    - For point 3, they are a consequence of [6, Theorem III.8] and Lemma 7.

    - For point 4, it suffices to apply [27, Theorem 5.6] and [39, Proposition 2.6].

5. Let $\{x_n\}_n, \{y_n\}_n$ be two countable families dense on (X,d) and (A, d_A), respectively. Taking into account that Γ is open-valued, we deduce that $[\cup_n \Gamma_A^*(\{y_n\})] \cup [\cup_n \Gamma_A^*(\{x_n\})] = \Omega$. From Lemma 4, $S(\Gamma_A)$ is non-empty, and applying Proposition 3, $P^*(A) = \max \mathcal{P}(\Gamma)(A)$.

6,7. First, we are going to show that under any of these two conditions, Γ has measurable selections.

    - It suffices to show that there exists $\{x_n\}_n \subseteq X$ s.t. $\Gamma^*(\cup_n[x_n]) = \Omega$ and to apply Lemma 4. Take $x_1 \in X$. If $\Gamma^*([x_1]) = \Omega$, the result holds. Otherwise, we select $\omega_1 \notin \Gamma^*([x_1])$ and $x_2 \in \Gamma(\omega_1)$ (note that as a consequence $\Gamma^*([x_1]) \subsetneq \Gamma^*([x_1] \cup [x_2])$), and repeat the process. The countability of $\mathcal{C}_\Gamma$ guarantees that this process is countable.

    - Let us denote $\Gamma = \sum_n A_n I_{B_n}$, and consider, for any $n \in \mathbb{N}$, $x_n \in A_n$. Then, $\Gamma^*(\cup_n[x_n]) = \cup_n B_n = \Omega$. Applying Lemma 4, $S(\Gamma) \neq \emptyset$.

Now, let us see that Γ_A also has measurable selections:

- Given $B \in \mathcal{A}'$, $\Gamma_A^*(B) = \Gamma^*(B \cap A) \cup (\Gamma^*(B) \setminus \Gamma^*(A))$. Hence, $\mathcal{C}_{\Gamma_A}$ is countable when $\mathcal{C}_\Gamma$ is.

- If $\Gamma = \sum_n A_n I_{B_n}$ for some countable partition $\{B_n\}_n$ of Γ, $\Gamma_A = \sum_n (A_n \cap A) I_{B_n \cap A^*} + \sum_n A_n I_{B_n \cap (A^*)^c}$. Therefore, the range of Γ_A is also countable.

Applying the previous points we deduce that $S(\Gamma_A) \neq \emptyset$ and, using now Proposition 3, $A \in \mathcal{H}_\Gamma$. ∎

The second and third points of Theorem 8 generalise Theorem 1. Moreover, this theorem also generalises the results mentioned in the proofs of [4, Proposition 2.7] and [23, Theorem 1].

We conclude therefore that the upper and lower probabilities of a random set provide, under fairly general conditions, the tightest available bounds for the probabilities induced by the original random variable. They are hence an adequate tool under the interpretation of Kruse and Meyer.

In our last proposition in this section, we provide a sufficient condition for the equality $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ to hold for every set A in a field that is included in the σ-field $\mathcal{A}'$. This property shall be useful in Section 3.3, when we relate the probability distributions in $\mathcal{P}(\Gamma)$ and $M(P^*)$. Recall that a complete random set is one whose images are complete subsets of the final space, i.e., subsets for which any Cauchy sequence has a limit within the set. In particular, this means that the images of the random set are closed.

Theorem 9. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, (X, d) a separable metric space, let $\Gamma : \Omega \to \mathcal{P}(X)$ be a complete random set. For every A in $\mathcal{Q}(\tau(d))$, the field generated by the open balls, $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ and $P_*(A) = \min \mathcal{P}(\Gamma)(A)$.*

Proof: Given C closed, Γ_C is also a complete random set on a separable metric space. From [6, Theorem III.6], we deduce that $S(\Gamma_C)$ is non-empty, whence all closed sets belong to $\mathcal{H}_\Gamma$. Any element from the field $\mathcal{Q}(\tau(d))$ is a countable union of closed sets (see for instance [37, Lemma 2.3.2]). Applying now the lower continuity of P^* and Proposition 6, we conclude that $\mathcal{Q}(\tau(d))$ is included in $\mathcal{H}_\Gamma$. ∎

It is an open problem at this stage whether, for this type of random sets, $\mathcal{H}_\Gamma$ coincides with $\mathcal{A}'$. An affirmative answer to this question would generalise the second and third points from Theorem 8. One possible approach would be

to study whether $\mathcal{H}_\Gamma$ is closed under countable intersections: in that case $\mathcal{H}_\Gamma$ would include the monotone class generated by the field $\mathcal{Q}(\tau(d))$, which is the Borel σ-field β_X. Another approach would be to use convergence properties of random sets, such as those discussed in [9] and [45, Chapter 7].

2.2 Convexity of $\mathcal{P}(\Gamma)(A)$

As we mentioned in the introduction, the study of the equality between $\mathcal{P}(\Gamma)(A)$ and $[P_*(A), P^*(A)]$ can be split into two different subproblems: the equality between $P^*(A), P_*(A)$ and the maximum and minimum values of $\mathcal{P}(\Gamma)(A)$ and the convexity of this last set. We focus our attention now on this second problem. We introduce first the following definition:

Definition 4. *[5] Let $(\Omega, \mathcal{A}, P)$ be a probability space. A set $B \in \mathcal{A}$ is not an* **atom** *when for every $\epsilon \in (0,1)$ there is some measurable $B_\epsilon \subsetneq B$ such that $P(B_\epsilon) = \epsilon P(B)$.*

Proposition 10. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ be a measurable space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. Let $U_1, U_2 \in S(\Gamma)$ satisfy $P_{U_1}(A) = \max \mathcal{P}(\Gamma)(A), P_{U_2}(A) = \min \mathcal{P}(\Gamma)(A)$. Then $\mathcal{P}(\Gamma)(A)$ is convex $\Leftrightarrow$ $U_1^{-1}(A) \setminus U_2^{-1}(A)$ is not an atom.*

    Proof: Note first that $P(U^{-1}(A) \setminus U_1^{-1}(A)) = P(U_2^{-1}(A) \setminus U^{-1}(A)) = 0$ for every $U \in S(\Gamma)$; otherwise, we obtain a contradiction with either $P_{U_1}(A) = \max \mathcal{P}(\Gamma)(A)$ or $P_{U_2}(A) = \min \mathcal{P}(\Gamma)(A)$.

($\Rightarrow$) Take $\alpha \in (0,1)$. If $\mathcal{P}(\Gamma)(A)$ is convex, there is some $V \in S(\Gamma)$ s.t. $P_V(A) = P_{U_2}(A) + \alpha(P_{U_1}(A) - P_{U_2}(A))$. The measurable set $B = V^{-1}(A) \cap (U_1^{-1}(A) \setminus U_2^{-1}(A)) \subseteq U_1^{-1}(A) \setminus U_2^{-1}(A)$ satisfies $P(B) = \alpha P(U_1^{-1}(A) \setminus U_2^{-1}(A))$.

($\Leftarrow$) Conversely, take $\alpha \in (0,1)$. Then, there is some measurable $B \subseteq U_1^{-1}(A) \setminus U_2^{-1}(A)$ such that $P(B) = \alpha P(U_1^{-1}(A) \setminus U_2^{-1}(A))$. Define $V := U_1 I_B + U_2 I_{B^c}$. It follows that $V \in S(\Gamma)$ and moreover $P_V(A) = P(B) + P_{U_2}(A) = \alpha P_{U_1}(A) + (1 - \alpha) P_{U_2}(A)$. Hence, $\mathcal{P}(\Gamma)(A)$ is convex. $\blacksquare$

We deduce that whenever the equalities $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ and $P_*(A) = \min \mathcal{P}(\Gamma)(A)$ hold, $\mathcal{P}(\Gamma)(A) = [P_*(A), P^*(A)]$ if and only if $A^* \setminus A_*$ is not an atom of the initial probability space. This immediately implies the following:

Corollary 11. *Under any of the conditions listed in Theorem 8,*

$$[P_*(A), P^*(A)] = \mathcal{P}(\Gamma)(A) \; \forall A \in \mathcal{A}' \Leftrightarrow \forall A \in \mathcal{A}', A^* \setminus A_* \text{ is not an atom of } \mathcal{A}.$$

The right-hand side of this equivalence holds trivially whenever the initial probability space is non-atomic; however, as we show in [37, Example 3.3.2], there are instances of random sets defined on a purely atomic probability space where $[P_*(A), P^*(A)] = \mathcal{P}(\Gamma)(A) \; \forall A \in \mathcal{A}'$.

3   Some implications of the previous results

3.1   Consonant random sets

One particular type of random sets which is of interest in practice are the **consonant** random sets, which are those whose images are nested. They have been studied in connection with possibility and maxitive measures in a number of works ([12, 18, 20, 39]). Since a possibility measure is usually defined on all subsets of its possibility space, we are going to assume in this section that the final σ-field is $\mathcal{P}(X)$, which is in particular reducible.

In this paper we are going to consider the following notion of consonant random sets. Other possibilities can be found in [39].

Definition 5. *A random set $\Gamma : \Omega \to \mathcal{P}(X)$ is called* **consonant** *when the following two conditions hold:*

- *For every $\omega_1, \omega_2 \in \Omega$, either $\Gamma(\omega_1) \subseteq \Gamma(\omega_2)$ or $\Gamma(\omega_2) \subseteq \Gamma(\omega_1)$.*

- *Every $A \subseteq \Omega$ has a countable subset B for which $\cap_{\omega \in A} \Gamma(\omega) = \cap_{\omega \in B} \Gamma(\omega)$.*

This definition is a generalisation of the so-called **antitone**[12] random sets, where the initial probability space is $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ and where $x \leq y \Rightarrow \Gamma(x) \supseteq \Gamma(y)$.

Proposition 12. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{P}(X))$ a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a consonant random set. Then $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for all $A \subseteq X$.*

Proof: Let us show first of all that a consonant random set has measurable selections. Let us consider a countable set $\{\omega_n\}_n$ such that $\cap_n \Gamma(\omega_n) = \cap_{\omega \in \Omega} \Gamma(\omega)$. We may assume without loss of generality that $\Gamma(\omega_n) \subsetneq \Gamma(\omega_{n-1})$ for all $n \geq 2$. If $\cap_n \Gamma(\omega_n) \neq \emptyset$, then a constant mapping on an element of $\cap_n \Gamma(\omega_n)$ is a measurable selection of Γ. Assume then that $\cap_n \Gamma(\omega_n) \neq \emptyset$. For each natural number n, let us consider $x_n \in \Gamma(\omega_n) \setminus \Gamma(\omega_{n+1})$. Then $\cup_n \{x_n\}^* = \Omega$: otherwise, there is some $\omega \in \Omega$ such that $\Gamma(\omega) \subsetneq \Gamma(\omega_n)$ for all n (because Γ is consonant), and then $\Gamma(\omega) \subseteq \cap_n \Gamma(\omega_n) = \emptyset$, a contradiction. This shows that $\cup_n \{x_n\}^* = \Omega$ and Lemma 4 implies that $S(\Gamma) \neq \emptyset$.

We prove next that if Γ is consonant the upper probability P^* is the inner approximation of its restriction to countable sets. Let A be a subset of X, and let $B = A^*$. If there is some $\omega \in B$ such that $\Gamma(\omega) \subseteq \Gamma(\omega')$ for every other $\omega' \in B$, then given $x \in \Gamma(\omega) \cap A$, $A^* = \{x\}^*$. If there is no such ω, then we consider a countable subset $\{\omega_n\}_n$ of B such that $\cap_n \Gamma(\omega_n) = \cap_{\omega \in B} \Gamma(\omega)$. If there is some $x \in A \bigcap \cap_{\omega \in B} \Gamma(\omega)$, it follows that $A^* = \{x\}^*$ and the result holds. If there is no such x, we consider, for each natural number n, $x_n \in \Gamma(\omega_n) \cap A$. Now, for every $\omega \in B$, there must be some n such that $\Gamma(\omega_n) \subseteq \Gamma(\omega)$: otherwise, $\Gamma(\omega) \subseteq \cap_n \Gamma(\omega_n) = \cap_{\omega' \in B} \Gamma(\omega')$, a contradiction with our previous assumption. We deduce that $\omega \in \{x_n\}^*$ for some n and as a consequence $A^* = \cup_n \{x_n\}^*$.

Since countable sets belong to $\mathcal{H}_\Gamma$ from Proposition 5, we deduce from Proposition 6 that $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for any $A \subseteq X$. $\blacksquare$

We can deduce from this proof and from [39, Propositions 2.4 and 5.2] that the upper probability P^* of a consonant random set is a possibility measure.

An open problem at this point is whether Proposition 12 can be generalised to weaker notions of consonancy for random sets, such as those considered in [39].

3.2 Relationship between the Aumann and the Choquet integral

Our results allow us also to relate the Choquet [8, 16] integral of a bounded function with respect to the upper and lower probabilities of a random set and the set of its integrals with respect to the measurable selections. This set is related to the Aumann integral of the random set, whose definition we recall:

Definition 6. *[3] Let $(\Omega, \mathcal{A}, P)$ be a probability space, and let $\Gamma : \Omega \to \mathcal{P}(\mathbb{R}^n)$ be a random set. Its* **Aumann integral** *is given by*

$$(A) \int \Gamma dP := \left\{ \int f dP : f \in L^1(P), f(\omega) \in \Gamma(\omega) \ a.s \right\}.$$

This is the definition of expectation of a random set which is more interesting in the context of this paper; there are, however, other definitions of expectation of a random set which could be used, and which in many cases coincide with the Aumann integral, such as the Debreu [14] or the Herer integral [25]. We refer to [43, Chapter 2] for a deeper study of this subject.

Note that is this definition we consider the set of the integrals with respect to the *almost-surely integrable* selections, which are those integrable mappings whose images are included in the random set with probability one. Given a

random set $\Gamma : \Omega \to \mathcal{P}(X)$ and a measurable mapping $f : X \to \mathbb{R}$, it is not difficult to see that $f \circ \Gamma : \Omega \to \mathcal{P}(\mathbb{R})$ is also a random set.

We first establish a result showing that our approximation result can be extended to finite chains. The intuition of this result is clear once we realise that one of the properties of the Choquet integral is its *comonotone additivity* [16], which implies in particular that it is additive over finite sums of chained sets.

Lemma 13. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ be a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. If $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for all $A \in \mathcal{A}'$, then for any finite chain $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n$ there is some $U \in S(\Gamma)$ such that $P_U(A_i) = P^*(A_i)$ for every $i = 1, \ldots, n$.*

Proof: By assumption for every $i = 1, \ldots, n$ there is some $U_i \in S(\Gamma)$ such that $P_{U_i}(A_i) = P^*(A_i)$. Let us define $B_0 = \emptyset$, $B_i = U_i^{-1}(A_i), C_i = B_i \setminus (\cup_{j=1}^{i-1} B_j)$ for $i = 1, \ldots, n$. Then $\{C_1, \ldots, C_n, (\cup_{j=1}^n C_j)^c\}$ constitutes a measurable partition of Ω.

Consider the mapping $U := \sum_{i=1}^n U_i I_{C_i} + U_n I_{(\cup_{j=1}^n C_j)^c}$. This is a measurable selection of Γ because it is a measurable combination of elements of $S(\Gamma)$. Moreover, for $i = 1, \ldots, n$,

$$U^{-1}(A_i) \supseteq \cup_{j=1}^i C_i = \cup_{j=1}^i B_i \supseteq B_i = U_i^{-1}(A_i),$$

whence $P_U(A_i) \geq P_{U_i}(A_i) = P^*(A_i)$ and consequently $P_U(A_i) = P^*(A_i)$ for $i = 1, \ldots, n$. ∎

Theorem 14. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ be a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. If $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for all $A \in \mathcal{A}'$, then for any bounded random variable $f : X \to \mathbb{R}$,*

$$(C) \int f \, dP^* = \sup_{U \in S(\Gamma)} \int f \, dP_U, \ (C) \int f \, dP_* = \inf_{U \in S(\Gamma)} \int f \, dP_U.$$

As a consequence,

$$(C) \int f \, dP^* = \sup(A) \int (f \circ \Gamma) dP, \ (C) \int f \, dP_* = \inf(A) \int (f \circ \Gamma) dP.$$

Proof: We are going to establish the result for P^*; the equalities for P_* follow by conjugacy.

Assume first of all that the random variable f is simple. Then, there is some natural number n, a partition $\{A_1, \ldots, A_n\}$ of X made up with measurable sets

17

and $x_1 > x_2 > \cdots > x_n$ such that $f = \sum_{i=1}^{n} x_i I_{A_i}$. Define $S_i = \cup_{j=1}^{i} A_j$, $i = 1, \ldots, n$, $S_0 = \emptyset$. Then

$$(C) \int f dP^* = \sum_{i=1}^{n} x_i (P^*(S_i) - P^*(S_{i-1})).$$

Since $S_0 \subseteq S_1 \subseteq S_2 \subseteq \cdots \subseteq S_n = X$, it follows from Lemma 13 that there is some $U \in S(\Gamma)$ such that $P_U(S_i) = P^*(S_i)$ for $i = 0, \ldots, n$. As a consequence,

$$\sum_{i=1}^{n} x_i (P^*(S_i) - P^*(S_{i-1})) = \sum_{i=1}^{n} x_i (P_U(S_i) - P_U(S_{i-1})) = \int f dP_U, \quad (6)$$

and since $\int f dP_V \le (C) \int f dP^*$ for every $V \in S(\Gamma)$, we deduce that $\int f dP_V = (C) \int f dP^*$.

Now, any bounded random variable f is the uniform limit of some increasing sequence $(f_n)_n$ of simple random variables. Applying [16, Theorem 8.1], we deduce that

$$(C) \int f dP^* = \sup_n (C) \int f_n dP^* = \sup_n \sup_{U \in S(\Gamma)} \int f_n dP_U = \sup_{U \in S(\Gamma)} \int f dP_U,$$

where the second equality follows from Eq. (6) and the third, from the dominated convergence theorem.

We turn now to the second equality. Let V be an almost-everywhere measurable selection of $f \circ \Gamma$. Then for every measurable A, $P_V(A) = P(V^{-1}(A)) \le P((f \circ \Gamma)^*(A)) = P(\Gamma^*(f^{-1}(A)))$. As a consequence, $\int V dP \le \int f dP^*$, whence $\sup(A) \int (f \circ \Gamma) dP \le (C) \int f dP^*$. Let us prove the converse inequality. For any $\epsilon > 0$, there is some $U_\epsilon \in S(\Gamma)$ such that $(C) \int f dP^* - \int f dP_U < \epsilon$. The mapping $f \circ U$ is a measurable selection of the random set $f \circ \Gamma$, whence $\int f dP_U = \int f \circ U dP \le \sup(A) \int (f \circ \Gamma) dP$, and as a consequence

$$(C) \int f dP^* - \sup(A) \int (f \circ \Gamma) dP < \epsilon.$$

We conclude that $(C) \int f dP^* = \sup(A) \int (f \circ \Gamma) dP$. $\blacksquare$

Using this result together with Theorem 8, we can generalise [7, Theorem 3.2]. On the other hand, in [43] it is showed the Aumann integral is a closed convex set as soon as the images of $f \circ \Gamma$ are closed convex subsets of $\mathbb{R}$ and the initial probability space is non-atomic (it coincides then with the so-called *selection expectation*). Hence, under those conditions Theorem 14 implies that

$$\left[(C) \int f dP_*, (C) \int f dP^* \right] = (A) \int (f \circ \Gamma) dP.$$

3.3 Measurable selections and the core of P^*

As we said in the introduction, the set $\mathcal{P}(\Gamma)$ of distributions of the selections is included in the core $M(P^*)$ of the upper probability, which is given by Eq. (3). This set can be more imprecise than $\mathcal{P}(\Gamma)$; on the other hand, it has the advantage of being convex and it is uniquely determined by the function P^*. This makes $M(P^*)$ easier to handle for practical purposes than $\mathcal{P}(\Gamma)$.

We can use our results on the equality between $\mathcal{P}(\Gamma)(A)$ and $[P_*(A), P^*(A)]$ to derive conclusions on the relationship between $\mathcal{P}(\Gamma)$ and $M(P^*)$. For instance, the results we have established allow us to relate the extreme points of the convex set of probability measures $M(P^*)$ to distributions of measurable selections. These extreme points are studied in some detail in [38]. Using our results in that paper, we prove in [40] the following result, which relates the closures of these two sets under the weak topology:

Theorem 15. *[40, Theorem 4.4] Let $(\Omega, \mathcal{A}, P)$ be a probability space, (X, d) a separable metric space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Let $\{x_n\}_n$ be a countable dense subset of X and let $\mathcal{J} := \{B(x_i, q) : q \in \mathbb{Q}, i \in \mathbb{N}\}$. If $P^*(A) = \max \mathcal{P}(\Gamma)(A)$ for all A in the field $\mathcal{Q}(\mathcal{J})$ generated by the countable class $\mathcal{J}$:*

1. $\overline{M(P^*)} = \overline{Conv(\mathcal{P}(\Gamma))}$.

2. $\overline{M(P^*)} = \overline{\mathcal{P}(\Gamma)} \Leftrightarrow \overline{\mathcal{P}(\Gamma)}$ is convex.

Note that not only we can apply this result together with Theorem 8 and Proposition 12, but also with Theorem 9, because for any separable metric space the field generated by the open balls includes in particular the field generated by $\mathcal{J}$. Hence, under very general situations, we can relate the core of the upper probability with the distributions of the measurable selections. Moreover, $\overline{\mathcal{P}(\Gamma)}$ is a convex set as soon as the initial probability space is non-atomic [40, Theorem 4.7]; this allows us to derive conditions for applying the second point of Theorem 15. On the other hand, the equality $\overline{\mathcal{P}(\Gamma)} = \overline{M(P^*)}$ does not imply in general that $\mathcal{P}(\Gamma)$ coincides with $M(P^*)$; an example and sufficient conditions for this equality can be found in [41].

As a side result, we also deduce that under any of the conditions listed in Theorem 8 and Proposition 12, P^* is the upper envelope of its core $M(P^*)$. This relates our work to the problem studied by Krätschmer in [32], and also to some results from [29, 46].

4 Conclusions

The results we have established show that the upper and lower probabilities of the random set are informative enough in most (but not in all) cases about the values taken by the probability distribution of the original random variable. Indeed, the features of Example 1 and the sufficient conditions listed in Theorem 8 make us conclude that we can use the upper and lower probabilities in all cases of practical interest. Moreover, the problem we have studied allows us to derive relationships between the core of the upper probability and the set of distributions of the measurable selections, and between the Aumann and Choquet integrals.

We would like to discuss here in more detail the case of *finite* random sets, i.e., those where X is a finite space and $\mathcal{A}' = \mathcal{P}(X)$. They have been studied for instance in [45, Chapter 3] and [36]. It follows from point 5 in Theorem 8 that for them the upper and lower probabilities can be approximated by the distributions of the measurable selections.

Hence, Theorem 14 implies that for any random variable $f : X \to \mathbb{R}$, its Choquet integral with respect to the upper probability is the maximum of its Aumann integral (note that in the proof of that theorem we establish that the supremum is a maximum for simple random variables, and this is always the case when X is a finite set). Moreover, the property of consonancy is here equivalent to the images of the random set being pair-wise nested, and that this is equivalent to the upper probability of the random set being a possibility measure [39, 47]. Finally, the sets $\mathcal{P}(\Gamma)$ and $M(P^*)$ are now closed under the weak topology (which is here equivalent to the Euclidean topology), and therefore Theorem 15 implies that $M(P^*)$ is the convex hull of $\mathcal{P}(\Gamma)$.

We have already pointed out in a few places some of the open problems derived from our results. More generally, it would be interesting to investigate the suitability of the upper and the lower probabilities when we have some additional information on the distribution of the original random variable (for instance that it belongs to some parametric family). Another interesting possibility would be to consider the case where we model the imprecise observation of U_0 by means of a fuzzy random variable.

References

[1] Alvarez, D. (2006), On the calculation of the bounds of probability of events using infinite random sets. *International Journal of Approximate Reasoning*, 43:241–267.

[2] Aubin, J.P. and Frankowska, H. (1990), *Set-valued analysis*. Birkhäuser, Boston.

[3] Aumann, J. (1965), Integral of set-valued functions. *Journal of Mathematical Analysis and Applications*, 12:1–12.

[4] Arstein, Z. and Hart, S. (1981), Law of large numbers for random sets and allocation processes. *Mathematics of Operations Research*, 6(4):485–492.

[5] Billingsley, P. (1986), *Probability and measure*. Wiley, New York.

[6] Castaing, C. and Valadier, M. (1977), *Convex analysis and measurable multifunctions*. Springer, Berlin.

[7] Castaldo, A., Maccheroni, F., and Marinacci, M. (2004), Random correspondences as bundles of random variables. *Sankhya*, 66(3): 409–427.

[8] Choquet, G. (1953), Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295.

[9] Couso, I. (1999), *Teoría de la Probabilidad con datos imprecisos. Algunos aspectos*. PhD Thesis, University of Oviedo. In Spanish.

[10] Couso, I., Montes, S. and Gil, P. (2002), Second order possibility measure induced by a fuzzy random variable, in: C. Bertoluzza, M. A. Gil, and D. A. Ralescu, eds., *Statistical modeling, analysis and management of fuzzy data*. Springer, Heidelberg, pp.127-144.

[11] Couso, I., Sánchez, L. and Gil, P. (2004), Imprecise distribution function associated to a random set. *Information Sciences*, 159(1–2):109–123.

[12] de Cooman, G. and Aeyels, D. (2000), A random set description of a possibility measure and its natural extension. *IEEE Transactions on Systems, Man and Cybernetics*, 30:124–130.

[13] Cressie, N. and Laslett, G. (1987), *Random set theory and problems of modeling*. SIAM review, 29:557-574.

[14] Debreu, G. (1967), Integration of correspondences, in *Proceedings of the Fifth Berkeley Sympos. on Math. Stat. and Probability*, 351-372.

[15] Dempster, A. P. (1967), Upper and lower probabilities induced by a multi-valued mapping. *Annals of Mathematical Statistics*, 38:325–339.

[16] Denneberg, D. (1994), *Non-additive measure and integral*. (Kluwer, Dordretch).

[17] Dhompongsa, S., Kaewkhao, A. and Saejung, S. (2007), On topological properties of Choquet weak convergence of capacity functionals of random sets. *Information Sciences*, 177:1852–1859.

[18] Dubois, D. and Prade, H. (1987), The mean value of a fuzzy number. *Fuzzy Sets and Systems*, 24: 279–300.

[19] Feng, D. and Nguyen, H.T. (2007), Choquet weak convergence of capacity functionals of random sets. *Information Sciences*, 177(16): 3239–3250.

[20] Goodman, I. R., (1982), Fuzzy sets as equivalence classes of possibility random sets, in: R. R. Yager, ed., *Fuzzy Sets and Possibility Theory: Recent Developments*. Pergamon, Oxford, pp. 327–343.

[21] Goutsias, J. (1997), Morphological analysis of random sets: an introduction, in: J. Goutsias, R. Mahler and H.T. Nguyen, eds, *Random sets: theory and applications*. Springer, New York, pp.3–26.

[22] Grabisch, M., Nguyen, H.T. and Walker, E.A. (1995), *Fundamentals of uncertainty calculi with applications to fuzzy inference*. Kluwer, Dordretch.

[23] Hart, S. and Köhlberg, E. (1974), Equally distributed correspondences. *Journal of Mathematical Economics*, 1(2):167–674.

[24] Heitjan, D. and Rubin, D. (1991), Ignorability and coarse data. *Annals of Statistics*, 19:2244–2253.

[25] Herer, W. (1990), Mathematical expectation and martingales of random subsets of a metric space. *Probability and Mathematical Statistics*, 11(2):291-304.

[26] Hess, C. (1999), The distribution of unbounded random sets and the multi-valued strong law of large numbers in nonreflexive Banach spaces. *Journal of Convex Analysis*, 6(1):163–182.

[27] Himmelberg, C. J. (1975), Measurable relations. *Fundamenta Mathematicae*, 87:53–72.

[28] Himmelberg, C. J., Parthasarathy, T. and Van Vleck, F. S. (1981), On measurable relations. *Fundamenta Mathematicae*, 111(2):161–167.

[29] Huber, P. J. and Strassen, V. (1973), Minimax tests and the Neyman-Pearson lemma for capacities. *Annals of Statistics*, 1(2):251–263.

[30] C. Joslyn. Distributional representations of random interval measurements. In B. Ayyub and M. Gupta, editors, *Uncertainty Analysis in Engineering and the Sciences*, pages 37–52. Kluwer Academic Publishers, Boston, 1997.

[31] Kechris, A. S. (1995), *Classical Descriptive Set Theory.* Springer Verlag, New York.

[32] Krätschmer, V. (2003) When fuzzy measures are upper envelopes of probability measures. *Fuzzy Sets and Systems*, 138(3):455–468.

[33] Kruse, R. and Meyer, K. D. (1987), *Statistics with vague data.* D. Reidel Publishing Company, Dordrecht.

[34] Li, S., Ogura, Y. and Kreinovich, V., (2002), *Limit theorems and applications of set-valued and fuzzy set-valued random variables.* Kluwer, Dordretch.

[35] Mathéron, G. (1975), *Random sets and integral geometry.* (Wiley, New York).

[36] Miranda, E., Couso, I. and Gil, P. (2002), Upper probabilities and selectors of random sets, in: P. Grzegorzewski, O. Hryniewicz, M. A. Gil, eds., *Soft methods in probability, statistics and data analysis.* Physica-Verlag, Heidelberg.

[37] Miranda, E. (2003), *Análisis de la información probabilística de los conjuntos aleatorios.* PhD Thesis, University of Oviedo. In Spanish.

[38] Miranda, E., Couso, I. and Gil, P. (2003), Extreme points of credal sets generated by 2-alternating capacities. *International Journal of Approximate Reasoning*, 33(2):95–115.

[39] Miranda, E., Couso, I. and Gil, P. (2004), A random set characterisation of possibility measures. *Information Sciences*, 168(1–4):51–75.

[40] Miranda, E., Couso, I. and Gil, P. (2005). Random sets as imprecise random variables. *Journal of Mathematical Analysis and Applications*, 307(1):32–47.

[41] Miranda, E., Couso, I. and Gil, P. (2005). Random intervals as a model for imprecise information. *Fuzzy Sets and Systems*, 154(3): 386–412.

[42] Miranda, E., Couso, I. and Gil, P. (2009). Upper probabilities attainable by distributions of measurable selections. Proceedings of the *Tenth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, pp. 335–346. Verona (Italy).

[43] Molchanov, I. (2005), *Theory of random sets*. Springer, London.

[44] Nguyen, H. T. (1978), On random sets and belief functions. *Journal of Mathematical Analysis and Applications*, 65(3):531–542.

[45] Nguyen, H.T. (2006), *An introduction to random sets*. Chapman and Hall. Boca Raton, Florida.

[46] Philippe, F., Debs G. and Jaffray, J.-Y., (1999), Decision making with monotone lower probabilities of infinite order. *Mathematics of Operations Research*, 24(3):767-784.

[47] Shafer, G. (1976), *A mathematical theory of evidence*. Princeton University Press, New Jersey.

[48] Wagner, D. H. (1977), Survey of measurable selection theorems. *SIAM Journal Control and Optimization*, 15(5):859–903.