Available online at www.sciencedirect.com
SCIENCE DIRECT
J. Math. Anal. Appl. 307 (2005) 32–47
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

# Random sets as imprecise random variables ☆

Enrique Miranda [a,*], Inés Couso [b], Pedro Gil [b]

[a] *Department of Informatics, Statistics and Telematics, Rey Juan Carlos University, C-Tulipán, s/n, 28933 Móstoles, Madrid, Spain*

[b] *Department of Statistics and Operations Research, University of Oviedo, Spain*

Received 30 March 2004

Available online 5 February 2005

Submitted by U. Stadtmueller

**Abstract**

Given a random set coming from the imprecise observation of a random variable, we study how to model the information about the probability distribution of this random variable. Specifically, we investigate whether the information given by the upper and lower probabilities induced by the random set is equivalent to the one given by the class of the probabilities induced by the measurable selections; together with sufficient conditions for this, we also give examples showing that they are not equivalent in all cases.


*Keywords:* Random sets; Upper and lower probabilities; Measurable selections; Choquet integral; Weak convergence

## 1. Introduction

Random sets, or measurable multi-valued mappings, constitute a useful generalisation of random variables, and have been successfully applied in such different fields as econ-

☆ This is an updated and expanded version of a paper of E. Miranda, I. Couso and P. Gil, Study of the probabilistic information of a random set, in: Proceedings of the 3rd ISIPTA Conference, Lugano, Switzerland, 2003, pp. 383–395 [22] presented at ISIPTA'03, the 3rd International Symposium on Imprecise Probabilities and Their Applications.

* Corresponding author.

*E-mail addresses:* emiranda@escet.urjc.es (E. Miranda), couso@pinon.ccu.uniovi.es (I. Couso), pedro@uniovi.es (P. Gil).

0022-247X/$ – see front matter 
doi:10.1016/j.jmaa.2004.10.022

omy [12] or stochastic geometry [18]. They have also been given different interpretations, like the behavioral [25] or the evidential one [8]. In this paper, we follow the interpretation given by Kruse and Meyer in [16], and regard them as a model for the imprecise observation of a random variable $U_0$. We assume that for some elements of the initial space we cannot tell their image by $U_0$ (due to some inaccuracy during the observation process, or simply to the existence of missing data), and we consider then a subset of the final space which is sure to include these images. This reasoning leads naturally to the definition of a random set, for which there are a number of possible ways of summarizing the information about the probability induced by the imprecisely observed random variable. The most important ones are the class of probability distributions of the measurable selections (that we shall denote $P(\Gamma)$) and those bounded between the upper and lower probabilities the random set induces (denoted $M(P^*)$ in this paper). Although working with the upper and lower probabilities leads to a number of mathematical simplifications [26,28], the information they provide is in general more imprecise than the one given by the set of distributions of the measurable selections [20,23]; our aim in this paper is to study the relationships between these models in order to understand the information conveyed by each of them.

In Section 2, we introduce some concepts and notations that we will use in the rest of the paper, and recall some previous works on the subject. In Section 3, we investigate the information that the upper and lower probabilities give about the values of the probability distribution induced by the original random variable. This is a first step towards the comparison of the models of this probability distribution, which is carried out in Section 4. Starting with a study of the extreme points of $M(P^*)$ and their relationship with $P(\Gamma)$, we prove several relationships between the upper and lower probabilities and the class of probabilities of the measurable selections that hold under fairly general conditions, and generalise some results from the literature. The paper concludes in Section 5 with some additional comments and remarks.

## 2. Preliminary concepts

Let us introduce some notation that we will use throughout the paper. We will denote a probability space by $(\Omega, \mathcal{A}, P)$, a measurable space by $(X, \mathcal{A}')$ and a multi-valued mapping, $\Gamma : \Omega \to \mathcal{P}(X)$. On the other hand, $(X, d)$ will denote a metric space, and $(X, \tau)$ will denote a topological space. Given a subset $A$ of a topological space, $\partial(A)$ will denote its boundary. Given a class of sets $\mathcal{H}$, $\mathcal{F}(\mathcal{H})$ and $\sigma(\mathcal{H})$ will denote, respectively, the field and the $\sigma$-field generated by $\mathcal{H}$. In the particular case where we consider the Borel $\sigma$-field generated by a topology $\tau$ on $X$, we will also denote $\beta_X = \sigma(\tau)$. The topology associated to a metric $d$ over $X$, i.e., the one generated by the open balls, will be denoted by $\tau(d)$. A topological space is said to be *Polish* when it is separable and complete for some compatible metric $d$, and it is called *Souslin* if it is the continuous image of a Polish space. A multi-valued mapping will be called open (respectively complete, closed, compact) if $\Gamma(\omega)$ is an open (respectively complete, closed, compact) subset of $X$ for every $\omega \in \Omega$. Given a random variable $U : \Omega \to \mathbb{R}$, $P_U$ and $F_U$ will denote, respectively, its induced probability and its distribution function. Finally, $\lambda_A$ will denote the Lebesgue measure on a set $A \in \beta_{\mathbb{R}}$ and $\mathcal{P}_{\mathcal{A}'}$ will denote the set of probabilities that can be defined on a $\sigma$-field $\mathcal{A}'$.

Formally, a random set is a multi-valued mapping that satisfies some measurability condition. Although in the literature we can find different conditions, such as the weak-, the strong-, or the graph-measurability [13,14], we will only work in this paper with the strong measurability: this condition is necessary if we want to be able to define the upper and lower probabilities on the final $\sigma$-field (and consequently, if we want the discussion carried in this paper to be possible).

**Definition 2.1.** 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 $\Gamma$ is $\Gamma^*(A) = \{\omega \in \Omega \mid \Gamma(\omega) \cap A \neq \emptyset\}$, and its *lower inverse* is $\Gamma_*(A) = \{\omega \in \Omega \mid \emptyset \neq \Gamma(\omega) \subseteq A\}$. The multi-valued mapping $\Gamma$ is said to be *strongly measurable* when $\Gamma^*(A)$ and $\Gamma_*(A)$ belong to $\mathcal{A}$ for all $A \in \mathcal{A}'$.

When there is no possible confusion about the multi-valued mapping we are working with, we will use the notation $A^* := \Gamma^*(A)$ and $A_* := \Gamma_*(A)$. By a *random set* we will mean throughout a strongly measurable multi-valued mapping.

**Definition 2.2** [8]. Given a random set $\Gamma : \Omega \to \mathcal{P}(X)$, the *upper probability* of $A \in \mathcal{A}'$ is $P^*_\Gamma(A) = \frac{P(A^*)}{P(X^*)}$, and the *lower probability*, $P_{*\Gamma}(A) = \frac{P(A_*)}{P(X^*)}$.

Note that the upper and lower probabilities of a random set can be defined on $\mathcal{A}'$ because we are assuming that $\Gamma$ is strongly measurable. This is not the case with other (weaker) measurability conditions. Let us also remark that, because $A^* = [(A^c)_*]^c$ for any $A \subseteq X$, it is $P^*_\Gamma(A) = 1 - P_{*\Gamma}(A^c)$ for all $A \in \mathcal{A}'$, i.e., these two functions are *conjugate*. $P^*_\Gamma$ and $P_{*\Gamma}$ are $\infty$-alternating and $\infty$-monotone capacities, respectively [26] and in particular satisfy Walley's axioms of coherence [28]. When there is no ambiguity about which random set is inducing the upper and lower probabilities, we will denote $P^* := P^*_\Gamma$ and $P_* := P_{*\Gamma}$.

As we pointed out in the introduction, we are regarding random sets as a model of the imprecise observation of random variables. Hence, we consider a random variable $U_0 : \Omega \to X$ (which we call *original* random variable) and assume that for every $\omega$ in the initial space all we know about $U_0(\omega)$ is that it belongs to the set $\Gamma(\omega)$. This idea has two immediate consequences: first, we may assume that $\Gamma(\omega)$ is non-empty for every $\omega$ in the initial space, whence $P^*(A) = P(A^*)$ and $P_*(A) = P(A_*)$ for all $A \in \mathcal{A}'$; and more importantly, our knowledge about $U_0$ is given by the class of *measurable selections* (or *selectors*) of $\Gamma$,

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

In particular, the probability distribution of $U_0$ belongs to

$$P(\Gamma) := \{P_U \mid U \in S(\Gamma)\}, \tag{1}$$

and our information about $P_{U_0}(A)$ is given by the set of values

$$P(\Gamma)(A) := \{P_U(A) \mid U \in S(\Gamma)\}. \tag{2}$$

Equations (1) and (2) are the most precise pieces of information that $\Gamma$ gives about the probability distribution of $U_0$, and about the values $\{P_{U_0}(A) \mid A \in \mathcal{A}'\}$, respectively. In general, these two pieces of information are not equivalent: we can derive the sets in Eq. (2) from Eq. (1), but there may be different sets of probabilities whose sets of values on $\mathcal{A}'$ coincide. We may then consider

$$\Delta(\Gamma) := \{Q \text{ probability} \mid Q(A) \in P(\Gamma)(A)\ \forall A \in \mathcal{A}'\}, \tag{3}$$

which is the biggest set of probabilities compatible with the sets in Eq. (2). It was first introduced by Couso in [6]. It is clear that $P(\Gamma) \subseteq \Delta(\Gamma)$. When they coincide, the information about $P_{U_0}$ is equivalent to the information about the values this probability takes. On the other hand, we can also consider the class

$$M(P^*) := \{Q \text{ probability} \mid Q(A) \leqslant P^*(A)\ \forall A \in \mathcal{A}'\} \tag{4}$$

of probabilities dominated by $P^*$, or (following the notation of Levi [17]) *credal set* generated by $P^*$, which has been more thoroughly studied in the literature ([4,6,8], among others). Given a set $A \in \mathcal{A}'$, its lower inverse $A_*$ is the greatest subset of $\Omega$ which is certain to be included in $U_0^{-1}(A)$, and its upper inverse $A^*$ is the smallest subset of $\Omega$ which is sure to include $U_0^{-1}(A)$ as a subset. Taking into account that all we know about $U_0$ is that it is a measurable selection of $\Gamma$, we deduce that $P_*(A) \leqslant P_U(A) \leqslant P^*(A)$ for all $U \in S(\Gamma), A \in \mathcal{A}'$. This implies that $\Delta(\Gamma) \subseteq M(P^*)$, and as a consequence $P(\Gamma) \subseteq \Delta(\Gamma) \subseteq M(P^*)$. However, both these inclusions can be strict, as the following example shows.

**Example 2.1** [6]. Let us consider the probability space $(\Omega, \mathcal{A}, P)$, where $\Omega = \{\omega_1, \omega_2\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ and $P(\{\omega_1\}) = \frac{1}{3}$ and the random set $\Gamma : \Omega \to \mathcal{P}(\{1, 2, 3\})$ given by $\Gamma(\omega_1) = \{1, 2, 3\}$, $\Gamma(\omega_2) = \{1, 2\}$. Then, it is easy to verify that

$$P(\Gamma) = \left\{(1, 0, 0), \left(\frac{2}{3}, \frac{1}{3}, 0\right), \left(\frac{2}{3}, 0, \frac{1}{3}\right), \left(\frac{1}{3}, \frac{2}{3}, 0\right), (0, 1, 0), \left(0, \frac{2}{3}, \frac{1}{3}\right)\right\},$$

where a vector $(p_1, p_2, p_3)$ denotes $(p(\{1\}), p(\{2\}), p(\{3\}))$. The probability measure given by $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ belongs to $\Delta(\Gamma) \setminus P(\Gamma)$. On the other hand, $(0.5, 0.3, 0.2) \in M(P^*) \setminus \Delta(\Gamma)$. Hence, in this case we have $P(\Gamma) \subsetneq \Delta(\Gamma) \subsetneq M(P^*)$.

The set $M(P^*)$ is convex and is uniquely determined by the upper probability. Hence, it is easier to handle than $P(\Gamma)$; we see from the example that it may also be more imprecise. The goal of this paper is to study the relationships between $P(\Gamma)$ and $M(P^*)$, determining under which conditions the upper probability of the random set keeps all the information about the probability distribution of the original random variable (as we will show, if we are only interested in the values taken by $P_{U_0}$ in the final $\sigma$-field, the set $\Delta(\Gamma)$ allows us to express the problem in terms of sets of probabilities instead of subsets of $[0, 1]$). This problem was studied in [20] for the case where $X$ is finite, in [24] for random intervals, and in [4] for compact random sets on Polish spaces. We will generalise some of the results from these references in this paper. On the other hand, other aspects of the sets of probabilities induced by the measurable selections or the upper and lower probabilities were investigated in [1,6,10,11]. Their relevance to this problem will be detailed later.

## 3. $P^*(A)$, $P_*(A)$ as a model for $P_{U_0}(A)$

We begin our study by comparing the information that the sets of probabilities defined by Eqs. (1), (3), and (4) give about the probability that the image of $U_0$ belongs to a certain set $A$ of the final $\sigma$-field. This information is given by the sets

$$\{p \in [0,1] \mid \exists Q \in P(\Gamma)\colon Q(A) = p\}, \tag{5}$$

$$\{p \in [0,1] \mid \exists Q \in \Delta(\Gamma)\colon Q(A) = p\} \quad \text{and} \tag{6}$$

$$\{p \in [0,1] \mid \exists Q \in M(P^*)\colon Q(A) = p\}, \tag{7}$$

respectively. As we see from Example 2.1, $P(\Gamma)$ can be strictly included in $\Delta(\Gamma)$. Nevertheless, it can be checked that the sets given by Eqs. (5) and (6) coincide with the class $P(\Gamma)(A)$ defined in Eq. (2). On the other hand, it is easy to see that the set defined in Eq. (7) is actually the interval $[P_*(A), P^*(A)]$. Let us study then under which conditions $P(\Gamma)(A)$ and $[P_*(A), P^*(A)]$ coincide. For this, we must determine under which conditions the maximum and minimum values of $P(\Gamma)(A)$ coincide, respectively, with the upper and lower probabilities of $A$, and also when $P(\Gamma)(A)$ is convex. We studied these two problems in [23]. Concerning the first, we showed that $P_*(A)$ and $P^*(A)$ are not equal in general to the minimum and maximum values of $P(\Gamma)(A)$. We also provided in that paper sufficient conditions for these equalities, which we summarize in the following theorem.

**Theorem 3.1** [23]. *Consider $(\Omega, \mathcal{A}, P)$ a probability space, $(X, \tau)$ a topological space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Under any of the following conditions*:

(1) *$\Omega$ is complete, $X$ is Souslin and* $\mathrm{Gr}(\Gamma) \in \mathcal{A} \otimes \beta_X$;
(2) *$X$ is a separable metric space and $\Gamma$ is compact*;
(3) *$X$ is a Polish space and $\Gamma$ is closed*;
(4) *$X$ is a $\sigma$-compact metric space and $\Gamma$ is closed*;
(5) *$X$ is a separable metric space and $\Gamma$ is open,*

$$P^*(A) = \max P(\Gamma)(A) \quad \text{and} \quad P_*(A) = \min P(\Gamma)(A) \quad \forall A \in \beta_X.$$

*Moreover, if*

(6) *$X$ is a separable metric space and $\Gamma$ is complete, then*

$$P^*(A) = \max P(\Gamma)(A), \qquad P_*(A) = \min P(\Gamma)(A) \quad \forall A \in \mathcal{F}\big(\tau(d)\big).$$

As we show in [23, Example 1], the equalities $P^*(A) = \max P(\Gamma)(A)$ and $P_*(A) = \min P(\Gamma)(A)$ do not hold in general, and therefore we must look for sufficient conditions such as those listed in this theorem; in fact, it may even happen that $\Gamma$ does not possess any measurable selection, and in that case both $P(\Gamma)$ and $\Delta(\Gamma)$ would be empty.

Let us make now a small digression concerning this theorem. When the equality $P^*(A) = \max P(\Gamma)(A)$ holds for every set $A$ in the final $\sigma$-field, the upper probability is the upper envelope of the set $P(\Gamma)$. We already know from the coherence of $P^*$ that it is the upper envelope of the class of *finitely* additive probabilities it dominates [28]; our theorem gives sufficient conditions for $P^*$ to be the upper envelope of the class of *countably* additive probabilities belonging to $P(\Gamma)$, which is in general a subclass of $M(P^*)$. This is related to the problem studied by Krätschmer in [15]. A similar comment can be made for $P_*$.

Our result is also related to some properties proven by Couso. In [6], she showed that the equality $P^*(A) = \sup P(\Gamma)(A)\ \forall A \in \mathcal{A}'$ implies the equality, for any bounded random variable, of its Choquet integral ([9]) respect to the upper probability of $\Gamma$ and the supremum of its integrals respect to the distributions of the measurable selections. This fact, together with Theorem 3.1 produces the following result, which generalizes [4, Theorem 1].

**Theorem 3.2.** *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 satisfying any of the conditions* (1) *to* (5) *from Theorem* 3.1. *Then, for any bounded random variable $V : X \to \mathbb{R}$, it is $(C)\int V\,dP^* = \sup\{\int V\,dQ \mid Q \in P(\Gamma)\}$ and $(C)\int V\,dP_* = \inf\{\int V\,dQ \mid Q \in P(\Gamma)\}$.*

Let us remark that in particular, under the hypotheses of this theorem, given a finite chain $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_m$ of elements of $\mathcal{A}'$, there exists an element $U$ of $S(\Gamma)$ s.t. $P_U(A_i) = P^*(A_i)$ for all $i = 1, \ldots, m$.

The second necessary condition for the equality $P(\Gamma)(A) = [P_*(A), P^*(A)]$ is the convexity of $P(\Gamma)(A)$. This property does not hold in general either. It can be characterized in terms of a property of the initial probability space. We need to remark that, as it is proven in [23], the set $P(\Gamma)(A)$ always has a maximum and a minimum value.

**Proposition 3.3** [23]. *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. Consider $A \in \mathcal{A}'$ and let $U_1, U_2 \in S(\Gamma)$ satisfy $P_{U_1}(A) = \max P(\Gamma)(A)$, $P_{U_2}(A) = \min P(\Gamma)(A)$. Then,*

$$P(\Gamma)(A) \text{ is convex} \quad \Leftrightarrow \quad U_1^{-1}(A) \setminus U_2^{-1}(A) \text{ is not an atom.}$$ [1]

The right-hand side holds trivially, and $P(\Gamma)(A)$ is consequently convex for all $A \in \mathcal{A}'$, when the initial probability space is non-atomic. This is for instance the case when we have some additional information stating that the probability distribution $P_{U_0}$ is continuous. Nevertheless, the non-atomicity of $(\Omega, \mathcal{A}, P)$ is not necessary for $P(\Gamma)(A)$ to be convex, as we showed in [20, Remark 1]. We can also see in the Example 1 from this reference that $U_1^{-1}(A) \setminus U_2^{-1}(A)$ is not necessarily an atom (and, consequently, that $P(\Gamma)(A)$ is not always a convex set). If we join now Theorem 3.1 and Proposition 3.3, we derive the following corollary:

**Corollary 3.4.** *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 satisfying any of the conditions* (1) *to* (5) *from Theorem* 3.1. *Then, for any $A \in \mathcal{A}'$,*

$$P(\Gamma)(A) = \big[P_*(A), P^*(A)\big] \quad \Leftrightarrow \quad A^* \setminus A_* \text{ is not an atom.}$$

[1] By this we mean that for every $\alpha \in (0,1)$ there is some measurable $B \subseteq U_1^{-1}(A) \setminus U_2^{-1}(A)$ with $P(B) = \alpha P(U_1^{-1}(A) \setminus U_2^{-1}(A))$.

This means that, under some conditions, the sets $P(\Gamma)$, $\Delta(\Gamma)$ and $M(P^*)$ provide the same information about the values of the probability distribution of the original random variable. Moreover, these conditions are not very restrictive: on the one hand, most random sets used for practical purposes satisfy one of the conditions (1) to (5) from Theorem 3.1; on the other hand, when $P_{U_0}$ is continuous $(\Omega, \mathcal{A}, P)$ is necessarily non-atomic, and then $A^* \setminus A_*$ is not an atom for any $A \in \mathcal{A}'$.

## 4. $P^*$, $P_*$ as a model for $P_{U_0}$

Let us study next the relationships between the sets $P(\Gamma)$, $\Delta(\Gamma)$ and $M(P^*)$, which model the information about the probability distribution $P_{U_0}$. It can easily be checked that $\Delta(\Gamma)$ coincides with $M(P^*)$ if and only if the sets $P(\Gamma)(A)$ and $[P_*(A), P^*(A)]$ coincide for all $A \in \mathcal{A}'$. Hence, our Corollary 3.4 gives sufficient conditions for the equality $\Delta(\Gamma) = M(P^*)$; as we argued before, these conditions are not very restrictive. Nevertheless, we showed in [20] that this equality does not imply the one between $P(\Gamma)$ and $M(P^*)$, not even when the final space is finite. This is another way to see that the information provided by $\Gamma$ about the probability distribution of $U_0$ is not equivalent, in general, to the one about the values of this distribution.

Although a possible approach to the study of the relationships between $P(\Gamma)$ and $M(P^*)$ would be to study the relationship between $P(\Gamma)$ and $\Delta(\Gamma)$ and combine the results with the ones mentioned in the previous paragraph, it will be more fruitful for this paper to study directly the relation between $P(\Gamma)$ and $M(P^*)$. Our course of reasoning will be based on the form of the extreme points of $M(P^*)$, and will use the following supporting result:[2]

**Theorem 4.1** [5]. *Let $X = \{x_1, \dots, x_n\}$ be a finite set and consider a 2-alternating capacity $\mu : \mathcal{P}(X) \to [0,1]$. For any permutation $\pi \in S^n$, let $Q_\pi$ be the probability on $\mathcal{P}(X)$ determined by the equations*

$$Q_\pi\big(\{x_{\pi(1)}, \dots, x_{\pi(j)}\}\big) = \mu\big(\{x_{\pi(1)}, \dots, x_{\pi(j)}\}\big) \quad \forall j = 1, \dots, n.$$

*Then*, $\mathrm{Ext}(M(\mu)) = \{Q_\pi \mid \pi \in S^n\}$ *and* $M(\mu) = \mathrm{Conv}(\{Q_\pi \mid \pi \in S^n\})$.

Using this result, we proved in [20] that given a random set $\Gamma : \Omega \to \mathcal{P}(X)$ taking values on a finite space, all the extreme points of $M(P^*)$ belong to $P(\Gamma)$; hence, $P(\Gamma)$ and $M(P^*)$ coincide if and only if $P(\Gamma)$ is convex. It would be interesting to see whether such an equivalence holds for more general final spaces, not necessarily finite. Although the direct implication holds in general, the converse does not hold necessarily when the cardinal of $X$ is infinite, as the following example shows.

[2] This theorem is an extension, for 2-alternating capacities, of a result established by Dempster [8] for random sets on finite spaces. Other proofs of this result in different contexts can be found in [3,27].

**Example 4.1.** Consider the probability space $((0,1), \beta_{(0,1)}, \lambda_{(0,1)})$, the measurable space $((0,1), \beta_{(0,1)})$, and the multi-valued mapping

$$\Gamma : (0,1) \to \mathcal{P}\big((0,1)\big), \quad \omega \hookrightarrow (0,\omega).$$

It is strongly measurable: given $A \in \beta_{(0,1)}$ non-empty, it is

$$\Gamma^*(A) = \big(\inf\big\{A \cap (0,1)\big\}, 1\big),$$

and trivially $\Gamma^*(\emptyset) = \emptyset$. We are going to prove that $P(\Gamma)$ coincides with the set of probabilities $\mathcal{C} := \{Q \in \mathcal{P}_{\beta_{(0,1)}} \mid \exists N_Q \in \beta_{(0,1)}, \lambda_{(0,1)}(N_Q) = 0,$ s.t. $Q((0,x]) > x\ \forall x \in (0,1) \setminus N_Q\}$. Note that any element of $\mathcal{C}$ satisfies $Q((0,x]) \geqslant x\ \forall x \in (0,1)$: it suffices to use the right-continuity of the distribution function associated to a probability.

($\subseteq$) Let $U$ be a measurable selection of $\Gamma$. Then,

- Given $x \in (0,1)$, $P_U((0,x]) \geqslant P_*((0,x]) = \lambda_{(0,1)}((0,x]) = x$.
- Let us denote $N_U := \{x \in (0,1) \mid P_U((0,x]) = x\}$. This is a subset of $(0,1)$, which is totally bounded; moreover, the right-continuity of the distribution function of $U$, $F_U$, implies that the limit of a decreasing sequence of elements of $N_U$ also belongs to $N_U$. Hence, for any $n \in \mathbb{N}$, there exist $x_1^n, \ldots, x_{m_n}^n \in N_U$ s.t. $N_U = N_U \cap (\bigcup_{i=1}^{m_n}[x_i^n, x_i^n + \frac{1}{n}])$.

  Let us define $A_n = \{\omega \in (0,1) \mid \omega - U(\omega) \geqslant \frac{1}{n}\}$. It is clear that the sequence $\{A_n\}_n$ is increasing and that $(0,1) = \bigcup_n A_n$, because $U$ is a selection of $\Gamma$. Given $x \in N_U$ and $n \in \mathbb{N}$,

$$\begin{aligned} x = P_U\big((0,x]\big) &= \lambda_{(0,1)}\big(\{\omega \mid U(\omega) \leqslant x\}\big) \geqslant \lambda_{(0,1)}\big((0,x)\big) \\ &\quad + \lambda_{(0,1)}\left(\left[x, x+\frac{1}{n}\right] \cap A_n\right) = x + \lambda_{(0,1)}\left(\left[x, x+\frac{1}{n}\right] \cap A_n\right), \end{aligned}$$

  whence $\lambda_{(0,1)}([x, x+\frac{1}{n}] \cap A_n) = 0$, and this implies that

$$\begin{aligned} &\lambda_{(0,1)}(N_U \cap A_n) \leqslant \sum_{i=1}^{m_n} \lambda_{(0,1)}\left(\left[x_i^n, x_i^n + \frac{1}{n}\right] \cap A_n\right) = 0 \\ &\quad\Rightarrow\quad \lambda_{(0,1)}(N_U \cap A_n) = 0 \quad \forall n. \end{aligned}$$

  Therefore, $\lambda_{(0,1)}(N_U) = \lim_n \lambda_{(0,1)}(N_U \cap A_n) = 0$, whence $P_U \in \mathcal{C}$.

($\supseteq$) Conversely, consider $Q \in \mathcal{C}$, and let $N'_Q := \{x \in (0,1) \mid Q((0,x)) = x\}$. Taking into account that $Q((0,x)) = Q((0,x])$ except for the countable number of discontinuity points of $F_Q$, we deduce that $Q((0,x)) > x$ for all $x \in (0,1)$ except for a null set (whence $\lambda_{(0,1)}(N'_Q) = 0$), and, from the right-continuity of $F_Q$, it is $Q((0,x)) \geqslant x\ \forall x$. Let $U : (0,1) \to \mathbb{R}$ be the quantile function of $Q$, $U(\omega) = \inf\{x \mid \omega \leqslant Q((0,x])\}$, and let $\{B_n\}_n$ be the measurable partition of $(0,1)$ given by $B_1 := (\frac{1}{2}, 1)$, $B_n := (\frac{1}{2^n}, \frac{1}{2^{n-1}}]\ \forall n \geqslant 2$. Let us define

$$U_1 : (0,1) \to (0,1), \quad \omega \hookrightarrow \begin{cases} U(\omega), & \text{if } \omega \notin N'_Q, \\ \frac{1}{2^n}, & \text{if } \omega \in N'_Q \cap B_n. \end{cases}$$

– Consider $\omega \in (0,1)$. If $\omega \in N'_Q$, there is some $n$ s.t. $\omega \in N'_Q \cap B_n$, and then $U_1(\omega) = \frac{1}{2^n} < \omega$. If $\omega \notin N'_Q$, then $Q((0,\omega)) > \omega$, whence there is some $\omega' < \omega$ with $Q((0,\omega']) > \omega$. Hence, $U_1(\omega) = U(\omega) = \inf\{x \mid \omega \leqslant Q((0,x])\} \leqslant \omega' < \omega$. On the other hand, if it were $U(\omega) = 0$, then it would be $Q(\emptyset) \geqslant \omega > 0$, a contradiction. Hence, $U_1(\omega) \in (0,\omega)\ \forall \omega$. This shows that $U_1$ is a selection of $\Gamma$.
– Given $A \in \beta_{(0,1)}$,

$$U_1^{-1}(A) = \left(U_1^{-1}(A) \cap N'_Q\right) \cup \left(U_1^{-1}(A) \cap \left(N'_Q\right)^c\right)$$
$$= \left(\bigcup_{\{n \mid \frac{1}{2^n} \in A\}} N'_Q \cap B_n\right) \cup \left(U^{-1}(A) \cap \left(N'_Q\right)^c\right) \in \beta_{(0,1)},$$

taking into account that $U$ is measurable and $N'_Q, \{B_n\}_n$ belong to $\beta_{(0,1)}$. Hence, $U_1$ is a measurable mapping.
– The quantile function $U$ of $Q$ satisfies $P_U = Q$. Taking into account that $U_1(\omega) = U(\omega)$ for all $\omega \notin \mathbb{N}'_Q$ and that $\lambda_{(0,1)}(N'_Q) = 0$, we conclude that $P_{U_1} = Q$.

Therefore, $P(\Gamma)$ coincides with $\mathcal{C}$, and it is immediate to verify that this set of probabilities is convex. Consider now the Lebesgue measure $\lambda_{(0,1)}$ on $\beta_{(0,1)}$. It satisfies $\lambda_{(0,1)}(A) \leqslant \lambda_{(0,1)}((\inf\{A \cap (0,1)\}, 1)) = P^*(A)\ \forall A \in \beta_{(0,1)}$ non-empty (and trivially $\lambda_{(0,1)}(\emptyset) = P^*(\emptyset) = 0$), whence $\lambda_{(0,1)} \in M(P^*)$. However, $\lambda_{(0,1)}((0,x]) = x\ \forall x$, whence $\lambda_{(0,1)} \notin \mathcal{C} = P(\Gamma)$. Hence, the convexity of $P(\Gamma)$ does not imply its equality with $M(P^*)$.

In [21], we investigated the form of the extreme points of $M(\mu)$ when $\mu$ is a 2-alternating and upper continuous capacity defined on a separable metric space $X$. The idea in that paper was to approximate a probability $Q : \beta_X \to [0,1]$ by a sequence of probabilities that coincide with $Q$ on a sequence of finite fields. In this section, we are going to use a similar construction, this time applied to the upper probability induced by a random set (which is not, in general, upper continuous). We will work with the topology of the weak convergence, whose main properties can be found in [2]. Together with the well-known Portmanteau's theorem, we will also use the following result:

**Proposition 4.2** [2]. *Let $(X,d)$ be a separable metric space, and consider a class $\mathcal{U} \subseteq \beta_X$ closed under finite intersections and such that every open set is a finite or countable union of elements from $\mathcal{U}$. Let $\{P_n\}_n \cup P$ be a family of probability measures on $\beta_X$ such that* $\lim_n P_n(A) = P(A)\ \forall A \in \mathcal{U}$. *Then, $\{P_n\}_n$ converges weakly to $P$.*

The following lemma will be used later:

**Lemma 4.3.** *Let $(X,d)$ be a metric space, and consider $\mathcal{Q}$ a convex set of probabilities defined on $\beta_X$. Then, its closure $\bar{\mathcal{Q}}$ in the topology of the weak convergence is also convex.*

**Proof.** Consider $P_1, P_2 \in \bar{\mathcal{Q}}, \alpha \in (0,1)$, and let us show that $\alpha P_1 + (1-\alpha)P_2 \in \bar{\mathcal{Q}}$. There are two sequences $\{P_n^1\}_n, \{P_n^2\}_n \subseteq \mathcal{Q}$ such that $\lim_n P_n^1 = P_1, \lim_n P_n^2 = P_2$. Let us consider, for every natural number $n$, the probability $\alpha P_n^1 + (1-\alpha)P_n^2$. It belongs to $\mathcal{Q}$

because this set is convex by hypothesis. Let $A$ be a $\alpha P_1 + (1-\alpha)P_2$-continuity set. Then, $0 = (\alpha P_1 + (1-\alpha)P_2)(\partial A) = \alpha P_1(\partial A) + (1-\alpha)P_2(\partial A)$, whence $A$ is also a $P_1$ and $P_2$-continuity set. Applying Portmanteau's theorem, $\lim_n P_n^1(A) = P_1(A)$ and $\lim_n P_n^2(A) = P_2(A)$, and as a consequence $\lim_n(\alpha P_n^1 + (1-\alpha)P_n^2)(A) = \alpha P_1(A) + (1-\alpha)P_2(A)$. Using again Portmanteau's theorem, we deduce that the sequence $\{\alpha P_n^1 + (1-\alpha)P_n^2\}_n$ converges weakly to $\alpha P_1 + (1-\alpha)P_2$. Therefore, this probability belongs to $\bar{\mathcal{Q}}$ and this set is convex. □

Consider a separable metric space $(X, d)$, and let $\{x_n\}_n$ be a countable dense subset of $X$. Let us define the class $\{B(x_i; q_j) \mid i \in \mathbb{N},\ q_j \in \mathbb{Q}\}$. This is a countable basis of the topology $\tau(d)$, and we denote it $\{B_n\}_n$. For every natural $n$, let $\mathcal{F}_n$ denote the field generated by $\{B_1, \ldots, B_n\}$, and let $\mathcal{F}(\{B_n\}_n)$ be the field generated by $\{B_n\}_n$. It can easily be checked that $\mathcal{F}(\{B_n\}_n) = \bigcup_n \mathcal{F}_n$; moreover, the class $\mathcal{F}(\{B_n\}_n)$ satisfies the hypotheses of Proposition 4.2: it is clear that it is closed under finite intersections, because it is a field; on the other hand, any open set is a countable union of elements from $\{B_n\}_n$, and in particular from $\mathcal{F}(\{B_n\}_n)$. Any element of $\mathcal{F}_n$ is a (finite and disjoint) union of elements from $\mathcal{D}_n := \{C_1 \cap C_2 \cap \cdots \cap C_n \mid C_i \in \{B_i, B_i^c\}\ \forall i\colon 1, \ldots, n\}$. Let us denote this class $\mathcal{D}_n := \{E_1^n, \ldots, E_{k_n}^n\}$.

Next, we prove the main theorem of the paper. It establishes a relationship between $P(\Gamma)$ and $M(P^*)$ which holds, taking into account our results from the previous section, under fairly general conditions.

**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 such that $P^*(A) = \max P(\Gamma)(A)\ \forall A \in \mathcal{F}(\{B_n\}_n)$. Then,*

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

**Proof.** (1) It is clear that $\overline{\mathrm{Conv}(P(\Gamma))} \subseteq \overline{M(P^*)}$. Conversely, consider $Q_1 \in M(P^*)$, and fix $n \in \mathbb{N}$. Consider the finite measurable space $(\mathcal{D}_n, \mathcal{P}(\mathcal{D}_n))$, and let us define the multi-valued mapping

$$\Gamma_n : \Omega \to \mathcal{P}(\mathcal{D}_n), \quad \omega \hookrightarrow \{E_i^n \mid \Gamma(\omega) \cap E_i^n \neq \emptyset\}.$$

This mapping is strongly measurable: given $I \subseteq \{1, \ldots, k_n\}$,

$$\Gamma_n^*(\{E_i^n\}_{i \in I}) = \{\omega \mid \exists i \in I,\ E_i^n \in \Gamma_n(\omega)\} = \{\omega \mid \exists i \in I,\ \Gamma(\omega) \cap E_i^n \neq \emptyset\}$$
$$= \Gamma^*\left(\bigcup_{i \in I} E_i^n\right) \in \mathcal{A}.$$

Let $Q$ be the probability measure on $\mathcal{P}(\mathcal{D}_n)$ determined by the equalities

$$Q(\{E_i^n\}) = Q_1(E_i^n) \quad \forall i = 1, \ldots, k_n. \tag{8}$$

It belongs to $M(P^*_{\Gamma_n})$:

$$Q\left(\{E_i^n\}_{i\in I}\right) = Q_1\left(\bigcup_{i\in I} E_i^n\right) \leqslant P^*_\Gamma\left(\bigcup_{i\in I} E_i^n\right) = P^*_{\Gamma_n}\left(\{E_i^n\}_{i\in I}\right)$$

for any $I \subseteq \{1, \dots, k_n\}$.

From Theorem 4.1, $M(P^*_{\Gamma_n}) = \mathrm{Conv}(\{Q_\pi \mid \pi \in S^{k_n}\})$, where the probability measure $Q_\pi : \mathcal{P}(\mathcal{D}_n) \to [0,1]$ is determined by

$$Q_\pi\left(\{E^n_{\pi(1)}, \dots, E^n_{\pi(j)}\}\right) = P^*_{\Gamma_n}\left(\{E^n_{\pi(1)}, \dots, E^n_{\pi(j)}\}\right) = P^*_\Gamma\left(\bigcup_{i=1}^{j} E^n_{\pi(i)}\right) \quad \forall j = 1, \dots, k_n.$$

For any of these extreme points, there is some $P_\pi \in P(\Gamma)$ with $P_\pi(E^n_j) = Q_\pi(\{E^n_j\})\ \forall j = 1, \dots, k_n$: it suffices to take into account that, from Theorem 3.2, we can approximate $P^*_\Gamma$ on a finite chain, and then make a correspondence similar to that of Eq. (8) between the restriction to $\mathcal{F}_n$ of a probability defined on $\beta_X$ and a probability on $\mathcal{P}(\mathcal{D}_n)$. As a consequence, given the probability $Q \in \mathrm{Conv}(\{Q_\pi \mid \pi \in S^{k_n}\})$ defined through Eq. (8), there exists $P_n \in \mathrm{Conv}(P(\Gamma))$ such that $P_n(E^n_j) = Q(\{E^n_j\}) = Q_1(E^n_j)\ \forall j = 1, \dots, k_n$, whence $P_n(A) = Q_1(A)\ \forall A \in \mathcal{F}_n$. The sequence $\{P_n\}_n \subseteq \mathrm{Conv}(P(\Gamma))$ satisfies $\lim_n P_n(A) = Q_1(A)$ for all $A \in \mathcal{F}(\{B_n\}_n)$. Applying Proposition 4.2, we conclude that $\{P_n\}_n$ converges weakly to $Q_1$, whence $M(P^*) \subseteq \overline{\mathrm{Conv}(P(\Gamma))}$. This implies the desired equality.

(2) The direct implication follows if we take into account that, from Lemma 4.3, the closure of $M(P^*)$ is convex. For the converse implication, assume that $\overline{P(\Gamma)}$ is convex. Then, applying the first point of this theorem, it is

$$\overline{M(P^*)} = \overline{\mathrm{Conv}(P(\Gamma))} \subseteq \overline{\mathrm{Conv}(\overline{P(\Gamma)})} = \overline{\overline{P(\Gamma)}} = \overline{P(\Gamma)},$$

and this implies that $\overline{P(\Gamma)} = \overline{M(P^*)}$. □

Let us remark that the conclusions of this theorem do not necessarily hold when the random set does not satisfy $P^*(A) = \max P(\Gamma)(A)$ for all $A \in \mathcal{F}(\{B_n\}_n)$; as we said before, in that case it may happen (see [23, Example 1]) that $P(\Gamma)$ is empty, or even if it is not, that it contains only one distribution, while the class $M(P^*)$ is much larger. In such a situation our theorem would not hold.

The first point of this theorem generalises a result established in [24] for random intervals, and also a result established by Castaldo and Marinacci for compact random sets on Polish spaces in [4]. On the other hand, the second point of the theorem extends the result mentioned before for the finite case: as we check in [19], when the final space is finite the classes $P(\Gamma)$ and $M(P^*)$ are closed, whence the equivalence $\overline{P(\Gamma)} = \overline{M(P^*)} \Leftrightarrow \overline{P(\Gamma)}$ convex becomes $P(\Gamma) = M(P^*) \Leftrightarrow P(\Gamma)$ convex.

As we mentioned in the introduction, the main advantage of $M(P^*)$ over $P(\Gamma)$ as a model of the information concerning $P_{U_0}$ is that it can be represented in terms of the upper probability, and its disadvantage is that it usually produces a loss of precision. Even when $\overline{M(P^*)} = \overline{\mathrm{Conv}(P(\Gamma))}$, the upper probability can lose some important information respect to $P(\Gamma)$, as we showed in [24]; this loss is smaller when the closures of $P(\Gamma)$ and $M(P^*)$

coincide. Taking into account the second point of this theorem, it becomes interesting to establish sufficient conditions for the convexity of $\overline{P(\Gamma)}$. The first one is given in the following proposition, which is an immediate consequence of Lemma 4.3.

**Proposition 4.5.** *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, d)$ be a metric space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. If $P(\Gamma)$ is convex, then $\overline{P(\Gamma)}$ is convex.*

Next, we are going to prove that if the initial probability space is non-atomic, the closure $\overline{P(\Gamma)}$ is convex and therefore it coincides (under the hypotheses of Theorem 4.4) with $\overline{M(P^*)}$. We need the following supporting result, where we use the notations established before Theorem 4.4.

**Lemma 4.6.** *Let $(\Omega, \mathcal{A}, P)$ be a non-atomic probability space, $(X, d)$ a separable metric space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Then, the class of probabilities*

$$\mathcal{H}_n := \Big\{ Q : \mathcal{P}(\mathcal{D}_n) \to [0,1] \text{ probability} \;\Big|\; \exists Q' \in P(\Gamma) \text{ such that } Q\big(\big\{E_i^n\big\}\big) = Q'\big(E_i^n\big)\ \forall i = 1, \dots, k_n \Big\}$$

*is convex for every $n$.*

**Proof.** Fix $n \in \mathbb{N}$, and consider $P_1, P_2 \in \mathcal{H}_n$, $\alpha \in (0,1)$. Then, there exist $U_1, U_2 \in S(\Gamma)$ with $P_{U_1}(E_i^n) = P_1(\{E_i^n\})$, $P_{U_2}(E_i^n) = P_2(\{E_i^n\})$ $\forall i = 1, \dots, k_n$. Let us consider the measurable partition of $\Omega$ given by $\{C_{ij} \mid i, j = 1, \dots, k_n\}$ with $C_{ij} = U_1^{-1}(E_i^n) \cap U_2^{-1}(E_j^n)$; from the non-atomicity of $(\Omega, \mathcal{A}, P)$, there is, for every $i, j$, some measurable $D_{ij} \subseteq C_{ij}$ such that $P(D_{ij}) = \alpha P(C_{ij})$. Define $D = \bigcup_{i,j} D_{ij}$ and the mapping $U : \Omega \to \mathbb{R}$ by

$$U := U_1 I_D + U_2 I_{D^c}.$$

Taking into account that $U_1$ and $U_2$ are selectors of $\Gamma$, we deduce that $U(\omega) \in \Gamma(\omega)$ $\forall \omega$. Besides, $U$ is measurable, because $U_1$ and $U_2$ are measurable and $D \in \mathcal{A}$. Hence, $U \in S(\Gamma)$. Moreover,

$$\begin{aligned} P_U\big(E_l^n\big) &= P\big(U_1^{-1}\big(E_l^n\big) \cap D\big) + P\big(U_2^{-1}\big(E_l^n\big) \cap D^c\big) \\ &= \sum_{i=1}^{k_n} P(D_{li}) + \sum_{j=1}^{k_n} \big(P(C_{jl}) - P(D_{jl})\big) \\ &= \sum_{i=1}^{k_n} \alpha P(C_{li}) + \sum_{j=1}^{k_n} (1-\alpha) P(C_{jl}) \\ &= \alpha P_{U_1}\big(E_l^n\big) + (1-\alpha) P_{U_2}\big(E_l^n\big) \quad \forall l = 1, \dots, k_n, \end{aligned}$$

and we deduce from this that $\alpha P_1 + (1-\alpha) P_2$ belongs to $\mathcal{H}_n$. $\square$

**Theorem 4.7.** *Let $(\Omega, \mathcal{A}, P)$ be a non-atomic probability space, $(X, d)$ a separable metric space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Then, $\overline{P(\Gamma)}$ is convex.*

**Proof.** Let us show first that $\mathrm{Conv}(P(\Gamma))$ is a subset of $\overline{P(\Gamma)}$. Consider $P_1, P_2 \in P(\Gamma)$, $\alpha \in (0,1)$. Define, for every natural $n$, probabilities $P_n^1$ and $P_n^2$ on $\mathcal{P}(\mathcal{D}_n)$ by $P_n^i(\{E_j^n\}) = P_i(E_j^n)$, $\forall j = 1, \dots, k_n$, $i = 1, 2$. Then, $P_n^1$, $P_n^2$ belong to $\mathcal{H}_n$, and applying the previous lemma, there exists $P_n' \in P(\Gamma)$ s.t.

$$P_n'(A) = \alpha P_n^1\bigl(\{A\}\bigr) + (1-\alpha)P_n^2\bigl(\{A\}\bigr) = \alpha P_1(A) + (1-\alpha)P_2(A), \quad \forall A \in \mathcal{F}_n.$$

Now, applying Proposition 4.2, we deduce that the sequence $\{P_n'\}_n$ converges weakly to $\alpha P_1 + (1-\alpha)P_2$ and as a consequence this probability belongs to $\overline{P(\Gamma)}$. Therefore,

$$\overline{P(\Gamma)} \subseteq \overline{\mathrm{Conv}\bigl(P(\Gamma)\bigr)} \subseteq \overline{P(\Gamma)},$$

whence $\overline{P(\Gamma)} = \overline{\mathrm{Conv}(P(\Gamma))}$. Applying Lemma 4.3, this set of probabilities is convex. This completes the proof. □

The conclusion of this theorem does not hold in general when the initial probability space is atomic: even in the finite case, where $P(\Gamma)$ is closed and therefore it is $P(\Gamma) = \overline{P(\Gamma)}$ there are examples where it is not convex (see [20, Example 1]).

Now, using Theorems 3.1, 4.4, and 4.7, we can establish conditions on the images of the random set that guarantee the equality between the closures of $P(\Gamma)$ and $M(P^*)$.

**Corollary 4.8.** *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, d)$ a separable metric space and let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. Under any of the following conditions*:

(1) *$\Omega$ is complete, $X$ is Souslin*[3] *and* $\mathrm{Gr}(\Gamma) \in \mathcal{A} \otimes \beta_X$,
(2) *$\Gamma$ is open,*
(3) *$\Gamma$ is complete,*
(4) *$X$ is $\sigma$-compact and $\Gamma$ is closed,*

$\overline{M(P^*)} = \overline{\mathrm{Conv}(P(\Gamma))}$. *If in addition $(\Omega, \mathcal{A}, P)$ is non-atomic, then $\overline{M(P^*)} = \overline{P(\Gamma)}$.*

**Proof.** The first part follows from Theorem 3.1 and the first point of Theorem 4.4. For the second part, it suffices to apply the second point of Theorems 4.4 and 4.7. □

This corollary extends some results from [11]: it is proven there that given two closed random sets $\Gamma_1, \Gamma_2$ taking values on a separable Banach space, the equality between $P^*_{\Gamma_1}$ and $P^*_{\Gamma_2}$ implies that

$$\overline{\mathrm{Conv}\bigl(P(\Gamma_1)\bigr)} = \overline{\mathrm{Conv}\bigl(P(\Gamma_2)\bigr)}.$$

Similar results can be found in [1,10], in those cases with other hypotheses on the random set: in [1], Arstein and Hart show that $P^*_{\Gamma_1} = P^*_{\Gamma_2} \Rightarrow \overline{P'(\Gamma_1)} = \overline{P'(\Gamma_2)}$, where $\Gamma_i$ is a closed random set on $\mathbb{R}^n$ and $P'(\Gamma_i)$ is the set of distributions of its *integrable* selections, for $i = 1, 2$. On the other hand, Hart and Köhlberg prove in [10] that $P^*_{\Gamma_1} = P^*_{\Gamma_2}$ implies

[3] Although a Souslin space is not in general metrizable, this extra hypothesis is necessary for the result.

$$
\begin{tikzcd}
 & P(\Gamma)\text{ closed} \arrow[dl] \arrow[ddr] & \\
P(\Gamma)=M(P^*) \arrow[d] & & \\
P(\Gamma)\text{ convex} & & M(P^*)\text{ closed}
\end{tikzcd}
$$

Fig. 1. Some relationships between $P(\Gamma)$ and $M(P^*)$ when their closures coincide.

that $\overline{P(\Gamma_1)} = \overline{P(\Gamma_2)}$, where $\Gamma_i$ is an integrably bounded random set defined between a non-atomic complete probability space and $\mathbb{R}^n$. We have proven that the equality between $P^*_{\Gamma_1}$ and $P^*_{\Gamma_2}$ implies the equality between $\overline{\mathrm{Conv}(P(\Gamma_1))}$ and $\overline{\mathrm{Conv}(P(\Gamma_2))}$ only requiring $\Gamma_1, \Gamma_2$ to be complete on a separable metric space; moreover, we have showed that these two sets of probabilities coincide with $\overline{M(P^*_{\Gamma_1})}$ and $\overline{M(P^*_{\Gamma_2})}$, respectively.

The corollary provides sufficient conditions for the equality between the closures, in the topology of the weak convergence, of $P(\Gamma)$ and $M(P^*)$. Under those conditions, if $P(\Gamma)$ is closed, it coincides with $M(P^*)$, and then the upper probability provides an accurate representation of the information concerning the probability distribution of $U_0$. We think it is interesting at this point to clarify the relationship between a number of topological conditions. This will avoid confusions and will help to understand the meaning of the relationships we have established. Whenever it is $\overline{P(\Gamma)} = \overline{M(P^*)}$, it is easy to see that the implications in Fig. 1 hold.

None of the converses of these implications is true in general.

**Example 4.2.**

(1) Let us start showing that the equality between $P(\Gamma)$ and $M(P^*)$ does not imply that $P(\Gamma)$ is closed. Consider the probability space $((0,1), \beta_{(0,1)}, \lambda_{(0,1)})$, which is non-atomic, and let $\Gamma : (0,1) \to \mathcal{P}(\mathbb{R})$ be given by $\Gamma(\omega) = (0,1)$ for all $\omega \in (0,1)$. Then, $M(P^*) = \{Q : \beta_{\mathbb{R}} \to [0,1] \text{ probability} \mid Q((0,1)) = 1\}$. Consider $Q \in M(P^*)$, and let $U : (0,1) \to \mathbb{R}$ denote its quantile function. Then, it is easy to see that $U$ is a selector of $\Gamma$ and satisfies $P_U = Q$, whence $P(\Gamma) = M(P^*)$. However, the sequence of degenerate probability measures on $\frac{1}{n}$, $\{\delta_{\frac{1}{n}}\}_n \subseteq M(P^*)$, converges weakly to $\delta_0 \notin M(P^*)$. Hence, neither $M(P^*)$ or $P(\Gamma)$ is closed.

(2) Let us see now that $P(\Gamma)$ is not necessarily closed when $M(P^*)$ is closed. Consider the probability space $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ and the random closed interval $\Gamma_1 : [0,1] \to \mathcal{P}(\mathbb{R})$ given by $\Gamma_1(\omega) = [-\omega, \omega]$ for all $\omega \in [0,1]$. Then, $M(P^*)$ is closed (this is indeed the case for all compact random sets on Polish spaces). However, we check in [24] that $P(\Gamma_1)$ is not convex, and, taking into account Fig. 1, it is not closed either.

(3) Let us show finally that the convexity of $P(\Gamma)$ does not imply the equality $P(\Gamma) = M(P^*)$. Let $\Gamma$ be the random set from Example 4.1. Then, the set of probabilities $P(\Gamma)$ is convex, but it does not coincide with $M(P^*)$.

## 5. Conclusions and open problems

The results established in this paper allow us to shed some light into the problem of the representation of the information provided by a random set. On the one hand, given an (unknown) selector $U_0$, the information about its induced probability distribution is not equivalent to the information about the values of this probability: the former is given by the class $P(\Gamma)$ of the distributions of the selectors, while the latter is given by the class of sets $\{P(\Gamma)(A) \mid A \in \mathcal{A}'\}$, which is in a one-to-one correspondence with the set of probabilities $\Delta(\Gamma)$. We can deduce from our results that these two sets of probabilities do not coincide except in very particular cases. On the other hand, the sets $\Delta(\Gamma)$ and $M(P^*)$ will be equivalent under fairly general conditions, and so will the closures of $P(\Gamma)$ and $M(P^*)$. In those cases, the upper probability keeps most of the information given by the random set about $P_{U_0}$, but it may produce nonetheless a loss of precision. We want to stress that the conditions for the equality $\overline{P(\Gamma)} = \overline{M(P^*)}$ are sufficiently general, because in practice it is common to consider closed (or open) random sets taking values on $\mathbb{R}^n$, and also the non-atomicity of the initial probability space is fairly usual (as we have said, it holds whenever there is a continuous random variable starting on this space).

We want to point out three open problems from this paper: first, it would be interesting to establish sufficient conditions for the equality between $P(\Gamma)$ and $M(P^*)$. Under those conditions, the upper probability would suffice to summarize the information about the probability distribution of the original random variable. We obtained some conditions of that type in [20] for random sets on finite spaces, and in [24] for random intervals. We would like to know if the equality $P(\Gamma) = M(P^*)$ holds under more general situations. Secondly, it would be important to compare the information provided by the sets of probabilities $P(\Gamma)$ and $\overline{P(\Gamma)}$ about some parameters of the probability induced by $U_0$; this would allow us, taking into account the sufficient conditions we have proven for $P(\Gamma) \subseteq M(P^*) \subseteq \overline{M(P^*)} = \overline{P(\Gamma)}$, to determine when to use $P(\Gamma)$ or $M(P^*)$ to represent the information given by the random set $\Gamma$. Taking into account Theorem 3.2, we think that it is probable that both sets of probabilities keep the same information for the expectation operator (see also [7]). And finally, it may be interesting to make an analogous study under additional hypotheses over $U_0$, such as the continuity of its probability distribution. A possible approach to this problem would be to approximate in some way the distribution of any measurable selection by a sequence of distributions of continuous measurable selections. As suggested by one of the referees, and also related to this problem, it would also be interesting the study of random sets or multi-valued mappings as a model of the imprecise observation of a mapping (not necessarily measurable). The study of the probabilistic information would be, however, more involved; one possible approach would be to extend probability measure on the initial space to $\mathcal{P}(\Omega)$ (probably no longer as a probability measure, but as an inner/outer measure). This would allow us to give a measure to the upper and lower inverses of the subsets of the final space (and also on their inverses by the selections); another alternative, as pointed out by the referee, would be to approximate the lower and upper inverses by sets where the probability is defined. We think that it may be more interesting in this non-measurable context to use the approach considered in [25] and to compare, using the results from that paper, the lower previsions induced by the multi-valued mapping and the selectors.

## Acknowledgments

The research in this paper has been partially supported by FEDER-MCYT, grant number BFM2001-3515. We also acknowledge the referees for their suggestions and comments.

## References

[1] Z. Arstein, S. Hart, Law of large numbers for random sets and allocation processes, Math. Oper. Res. 6 (1981) 485–492.
[2] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[3] L.M. de Campos, M.J. Bolaños, Representation of fuzzy measures through probabilities, Fuzzy Sets and Systems 31 (1989) 23–36.
[4] A. Castaldo, M. Marinacci, Random correspondences as bundles of random variables, in: Proceedings of the 2nd ISIPTA Conference, Ithaca, New York, 2001, pp. 77–82.
[5] A. Chateauneuf, J.-Y. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, Math. Soc. Sci. 17 (1989) 263–283.
[6] I. Couso, Teoría de la probabilidad para datos imprecisos. Algunos aspectos, PhD thesis, University of Oviedo, 1999 (in Spanish).
[7] I. Couso, E. Miranda, G. de Cooman, A possibilistic interpretation of the expectation of a fuzzy random variable, in: M. López Díaz, M.A. Gil, P. Grzegorzewski, O. Hryniewicz, J. Lawry (Eds.), Soft Methodology and Random Information Systems, Springer-Verlag, Heidelberg, 2004, pp. 133–140.
[8] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. of Math. Stat. 38 (1967) 325–339.
[9] D. Denneberg, Non-additive Measure and Integral, Kluwer Academic, Dordrecht, 1994.
[10] S. Hart, E. Köhlberg, Equally distributed correspondences, J. Math. Econ. 1 (1974) 167–674.
[11] C. Hess, The distribution of unbounded random sets and the multivalued strong law of large numbers in nonreflexive Banach spaces, J. Convex Anal. 6 (1999) 163–182.
[12] W. Hildenbrand, Core and Equilibria of a Large Economy, Princeton Univ. Press, Princeton, NJ, 1974.
[13] C.J. Himmelberg, Measurable relations, Fund. Math. 87 (1975) 53–72.
[14] C.J. Himmelberg, T. Parthasarathy, F.S. Van Vleck, On measurable relations, Fund. Math. 111 (1981) 161–167.
[15] V. Krätschmer, When fuzzy measures are upper envelopes of probability measures, Fuzzy Sets and Systems 138 (2003) 455–468.
[16] R. Kruse, K.D. Meyer, Statistics with Vague Data, Reidel, Dordrecht, 1987.
[17] I. Levi, The Enterprise of Knowledge, MIT Press, Cambridge, UK, 1980.
[18] G. Mathéron, Random Sets and Integral Geometry, Wiley, New York, 1975.
[19] E. Miranda, Análisis de la información probabilística de los conjuntos aleatorios, PhD thesis, University of Oviedo, 2003 (in Spanish).
[20] E. Miranda, I. Couso, P. Gil, 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, 2002, pp. 126–133.
[21] E. Miranda, I. Couso, P. Gil, Extreme points of credal sets generated by 2-alternating capacities, Internat. J. Approx. Reason. 33 (2003) 95–115.
[22] E. Miranda, I. Couso, P. Gil, Study of the probabilistic information of a random set, in: Proceedings of the 3rd ISIPTA Conference, Lugano, Switzerland, 2003, pp. 383–395.
[23] E. Miranda, I. Couso, P. Gil, Upper probabilities attainable by distributions of measurable selections, Technical report, 2003.
[24] E. Miranda, I. Couso, P. Gil, Random intervals as a model for imprecise information, Technical report, 2004.
[25] E. Miranda, G. de Cooman, I. Couso, Lower previsions induced by multi-valued mappings, J. Statist. Planning and Inference, 2004, in press.
[26] H.T. Nguyen, On random sets and belief functions, J. Math. Anal. Appl. 65 (1978) 531–542.
[27] L.S. Shapley, Cores of convex games, Internat. J. Game Theory 1 (1971) 11–26.
[28] P. Walley, Statistical reasoning with imprecise probabilities, Chapman & Hall, London, 1991.