COHERENT UPDATING OF NON-ADDITIVE MEASURES

ENRIQUE MIRANDA, IGNACIO MONTES

ABSTRACT. The conditions under which a 2-monotone lower prevision can be uniquely updated (in the sense of focusing) to a conditional lower prevision are determined. Then a number of particular cases are investigated: completely monotone lower previsions, for which equivalent conditions in terms of the focal elements of the associated belief function are established; random sets, for which some conditions in terms of the measurable selections can be given; and minitive lower previsions, which are shown to correspond to the particular case of vacuous lower previsions.

Keywords: Coherent lower previsions, n-monotonicity, belief functions, minitive measures, natural extension, regular extension.

1. INTRODUCTION

The theory of imprecise probabilities contains a wide variety of mathematical models which are of interest in situations where it is unfeasible to determine, with certain guarantees, the probability model associated with an experiment. It includes for instance 2-monotone capacities [4], belief functions [43], possibility and necessity measures [24, 56], random sets [17, 41] or coherent lower previsions [48]. Under any of them, one important problem is that of updating the model under the light of new information. Unfortunately, this matter is far from settled, and quite a number of different rules has been proposed (see for instance [52] for an overview in the case of possibility measures). Among the most popular are Dempster's rule of conditioning [17, 43], regular extension [9, 29] and natural extension [48].

In order to be able to choose one rule above the others for a particular problem, it is essential to have a clear interpretation of the mathematical model we are using, and of what we mean by *updating* in our context. In this paper we deal with *epistemic* probabilities, where we model degrees of (partial) knowledge from a subject, and more specifically we focus on the behavioural approach championed by Peter Walley [48], that has its roots in the works on subjective probability by Bruno de Finetti [16]. This approach regards the lower and upper probabilities of an event as its supremum and infimum acceptable betting rates, and focuses on a consistency notion between these betting rates called *coherence*. Although this may seem restrictive at first, we argue that this is not the case, for a number of reasons:

- Imprecise probability models satisfying the notion of coherence (from now on *coherent lower previsions*) are always the envelopes of a convex set of probability measures. As a consequence, the behavioural approach is also compatible with a Bayesian sensitivity analysis interpretation.
- Almost all models of imprecise probabilities considered in the literature can be seen as particular instances of coherent lower previsions [51], and as a consequence our results shall be applicable also to them.

With respect to the meaning of updating, although in the case of a precise probability model this is relatively straightforward, and it amounts to applying Bayes' rule, when we move to the imprecise case the situation becomes more complicated. There are basically two main scenarios:

- *Belief revision* [1, 30], where we modify either the generic knowledge or the factual evidence about the problem under the light of new knowledge/evidence. The modification is usually done under the principle of minimal change. In the context of imprecise probabilities it gives rise to rules such as Dempster's rule of conditioning [17, 27].
- *Focusing* [27, Section 6], where we condition our generic knowledge on factual evidence. This produces rules such as the regular extension [29].

Within Walley's behavioural approach to imprecise probabilities, the interpretation of the lower prevision of a gamble conditional on an event B corresponds to the *current* supremum acceptable betting rate we would establish for the gamble, assuming that later we come to know that the outcome of the experiment belongs to B. As such, it tells us which are the predictions associated with our current model, and therefore the process of updating corresponds to a problem of focusing.

Using this interpretation, Walley proposes in [48, Chapter 6] a notion of coherence that tells us if the conditional betting rates are compatible with the unconditional ones. However, this notion does not suffice to uniquely determine the conditional models from the unconditional ones. This was shown for instance in [37], where it was established that in general we may have an infinite number of conditional models compatible with the unconditional one, and that the smallest and greatest such models are respectively determined by the procedures called *natural* and *regular* extension, whose underlying differences we shall discuss later in the paper. Here, we investigate under which conditions the natural and the regular extensions coincide, and as a consequence there is only one conditional model that is coherent with the unconditional one. This would mean that in those cases it is not necessary to choose between the natural and the regular extensions (or any of other coherent rules that lie between them).

The rest of the paper is organised as follows: we shall recall the basics from the theory of coherent lower previsions in Section 2. Then we shall focus on a particular case of coherent lower previsions: those satisfying the property of 2-monotonicity [4, 14]; these have the advantage that, unlike general coherent lower previsions, they are uniquely determined by their restrictions to events (a 2-monotone lower probability) by means of the Choquet integral.

After establishing a necessary and sufficient condition for the uniqueness of the coherent extensions in Section 3, we focus on two particular cases of 2-monotone lower previsions. First, in Section 4 we consider completely monotone lower previsions, which correspond to the Choquet integral with respect to a belief function [14]. We show that the necessary and sufficient condition mentioned above can be simplified by means of the focal elements of the belief function. Moreover, completely monotone lower previsions are associated with random sets, and from this we characterize the equality between the natural and the regular extensions in terms of the images of the random set; we also give an equivalent expression of the regular extension in terms of the measurable selections.

In Section 5 we focus on a second instance of 2-monotone lower previsions: the Choquet integral functionals with respect to Boolean necessity measures. Taking into account some recent results [12, 13], these are related to the so-called *vacuous* lower previsions, which model a situation of complete ignorance about the outcome of an experiment. Interestingly, we show that both the natural and regular extensions also produce vacuous models, although they do not coincide in general; moreover, there exist also non-vacuous models coherent with our unconditional lower prevision.

One interesting fact stems from our results in this last section: that the problem of checking the coherence between the unconditional and the conditional models is not equivalent for lower probabilities and for lower previsions; and this even when the lower previsions, both in the unconditional and the conditional case, are uniquely determined by their associated lower probabilities. Indeed, in [52] it is proved that the smallest and greatest conditional possibility measures that are coherent with an unconditional possibility measure are the ones determined by Dempster's rule and by natural extension, respectively. This leads the authors to propose the harmonic mean between these two possibility measures as an updating rule. As we shall show, if we consider the upper previsions determined from these unconditional and conditional possibility measures by means of the Choquet integral, we obtain models that are not necessarily coherent.

We conclude the paper with some additional remarks in Section 6.

2. Preliminary concepts

Let us introduce the main concepts of the theory of coherent lower previsions we shall use in this paper. We refer to [48] for a more detailed exposition of the theory, and in particular of the behavioural interpretation of the concepts we shall introduce below. A survey of the theory can be found in [36].

2.1. **Coherent lower previsions.**

Consider a possibility space Ω, that we shall assume in this paper to be *finite*. A *gamble* is a real-valued functional defined on Ω. We shall denote by $\mathcal{L}(\Omega)$ the set of all gambles on Ω. One instance of gambles are the indicators of events. Given a subset A of Ω, the indicator function of A is the gamble that takes the value 1 on the elements of A and 0 elsewhere. We shall denote this gamble by I_A, or by A when no confusion is possible.

A *lower prevision* is a functional $\underline{P}$ defined on a set of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$. Given a gamble f, $\underline{P}(f)$ is understood to represent a subject's supremum acceptable buying price for f, in the sense that for any $\epsilon > 0$ the transaction $f - \underline{P}(f) + \epsilon$ is acceptable to him.

Using this interpretation, we can derive a notion of coherence:

Definition 1. A lower prevision $\underline{P} : \mathcal{L}(\Omega) \to \mathbb{R}$ is called *coherent* if and only if it satisfies the following properties for every $f, g \in \mathcal{L}(\Omega)$ and every $\lambda > 0$:

(C1) $\underline{P}(f) \geq \min f$.
(C2) $\underline{P}(\lambda f) = \lambda \underline{P}(f)$.
(C3) $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$.

The interpretation of this notion is that the acceptable buying prices encompassed by $\{\underline{P}(f) : f \in \mathcal{L}(\Omega)\}$ are consistent with each other, in the sense defined in [48, Section 2.5]. In particular, when $\underline{P}$ satisfies (C3) with equality for every $f, g \in \mathcal{L}(\Omega)$, it is called a *linear* prevision. Any coherent lower prevision is the *lower envelope* of the set of linear previsions that dominate it, i.e.,

$$\underline{P}(f) = \min\{P(f) : P \text{ linear prevision}, P \geq \underline{P}\}.$$

The conjugate functional $\overline{P}$ of a coherent lower prevision $\underline{P}$ is called a coherent *upper* prevision, and it is given by $\overline{P}(f) = -\underline{P}(-f)$ for every $f \in \mathcal{L}(\Omega)$. It corresponds to the upper envelope of the set of linear previsions that dominate $\underline{P}$.

A coherent lower prevision defined on indicators of events only is called a *coherent lower probability*. Its conjugate function $\overline{P}$, given by $\overline{P}(A) = 1 - \underline{P}(A^c)$ for every subset A of Ω, is a coherent *upper* probability. In particular, the restriction of a linear prevision to indicators of events is a probability measure. In fact, coherent lower previsions can be equivalently defined as lower envelopes of closed and convex sets of probability measures, and as such they can also be given a Bayesian sensitivity analysis interpretation.

One particular case of coherent lower previsions are the *vacuous* ones. They correspond to the case where the only information we have is that the outcome of the experiment belongs to a subset A of Ω. In that case, our coherent lower prevision is given by

$$\underline{P}(f) = \min_{\omega \in A} f(\omega) \text{ for every } f \in \mathcal{L}(\Omega). \tag{1}$$

Although a linear prevision is uniquely determined by the probability measure that is its restriction to events, this is not the case for lower previsions: a coherent lower probability will have in general more than one coherent extension to the set of all gambles. This is the reason why the theory is established in terms of gambles instead of events. Interestingly, there are some cases where the restriction to events uniquely determines the coherent lower prevision. One instance that shall be important in this paper is that where the restriction to events is 0–1-valued:

Lemma 1. [48, Note 4, Section 3.2.6] *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\Omega)$ whose restriction to events is 0–1-valued*[1]*. Then $\underline{P}$ is the unique coherent extension of its restriction to events, and it is given by*

$$\underline{P}(f) = \max_{D:\underline{P}(D)=1} \min_{\omega \in D} f(\omega) \quad \textit{for every } f \in \mathcal{L}(\Omega);$$

moreover, the class $\{D \subseteq \Omega : \underline{P}(D) = 1\}$ is a filter on Ω.

This applies in particular to the vacuous lower previsions in Eq. (1).

2.2. **Conditional lower previsions.** The theory of coherent lower previsions can also be extended to the conditional case in the following way: given a partition[2] $\mathcal{B}$ of the set of outcomes, a *conditional lower prevision* on $\mathcal{L}(\Omega)$ is a functional $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{L}(\Omega)$ that to any gamble f and any $B \in \mathcal{B}$ assigns the value $\underline{P}(f|B)$; this value represents a subject's supremum acceptable buying price for f, if he comes to know later that the outcome of the experiment belongs to the subset B of Ω. By putting all these values together, we end up with the gamble

$$\underline{P}(f|\mathcal{B}) := \sum_{B \in \mathcal{B}} I_B \, \underline{P}(f|B).$$

[1]Note that this implies that the restriction to events of $\underline{P}$ is a (Boolean) necessity measure.

[2]We are using partitions here because we are sticking to Walley's formulation of the theory in [48, Chapter 6]. Nevertheless, this assumption is not important in the context of this paper, where the possibility space Ω is finite, as Walley's theory is then equivalent to the one formulated by Peter Williams in [53], which does not make use of partitions. Note also that in the mathematical developments we shall make from Section 3 onwards we shall deal with each of the conditioning events separately.

Similarly to conditions (C1)–(C3), we can establish a notion of coherence for conditional lower previsions.

Definition 2. A conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{L}(\Omega)$ is *separately coherent* when the following properties hold for every $f, g \in \mathcal{L}(\Omega), \lambda > 0$ and $B \in \mathcal{B}$:

(SC1) $\underline{P}(f|B) \geq \min_{\omega \in B} f(\omega)$.
(SC2) $\underline{P}(\lambda f|B) = \lambda \underline{P}(f|B)$.
(SC3) $\underline{P}(f + g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$.

The behavioural interpretation of this notion is that, for every set B in the partition $\mathcal{B}$, the acceptable conditional buying prices encompassed by $\underline{P}(\cdot|B)$ are consistent with each other. A separately coherent conditional lower prevision satisfies $\underline{P}(B|B) = 1$ for every $B \in \mathcal{B}$.

If we start with a coherent lower prevision $\underline{P}$ and consider a partition $\mathcal{B}$ of the space Ω, there is in general not a unique way of updating $\underline{P}$ into a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$. This is related to the problem of conditioning on sets of probability zero, which has attracted a lot of attention in the literature [20, 35]. In particular, de Finetti's behavioural approach [16] is at the root of a number of interesting studies on the verification of coherence of conditional assessments, both in the precise [3, 6], and the imprecise case [2, 5]. In this paper we shall consider the approach considered by Walley in [48, Chapter 6].

In the next section we detail how the conditional lower prevision may be derived and we formulate the problem under study. Note that under Walley's interpretation, a conditional lower prevision models the consequences of the assessments encompassed by the unconditional lower prevision $\underline{P}$. In other words, the elicitation of the conditional lower prevision is a model of focusing, in the sense discussed in the introduction. For a study of coherence in the dynamic case, we refer to [45, 57].

2.3. **Formulation of the problem.** Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$, let $\mathcal{B}$ be a partition of Ω. Assume we want to update the coherent lower prevision $\underline{P}$ into a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{L}(\Omega)$.

One strategy to derive $\underline{P}(\cdot|\mathcal{B})$ from $\underline{P}$ is to verify that the assessments present in these two lower previsions are compatible with each other. This gives rise to the concept of *joint coherence*, which is studied in much detail in [48, Chapters 6 and 7]. In the context of this paper, where the possibility space Ω is finite, joint coherence is characterized in the following way:

Proposition 1. [48, Theorem 6.5.4] *Consider a coherent lower prevision $\underline{P}$ and a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{L}(\Omega)$, where Ω is a finite space. They are jointly coherent when*

$$\underline{P}(B(f - \underline{P}(f|B))) = 0 \quad \textit{for every } f \in \mathcal{L}(\Omega),\ B \in \mathcal{B}. \tag{2}$$

In the particular case where $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are precise previsions, that we can then denote $P, P(\cdot|\mathcal{B})$, Eq. (2) becomes

$$P(B(f - P(f|B))) = 0 \quad \text{for every } f \in \mathcal{L}(\Omega), B \in \mathcal{B}.$$

This means that whenever $P(B) > 0$ it must be $P(f|B) = \frac{P(Bf)}{P(B)}$, i.e., the conditional prevision must be defined by means of Bayes' rule. Because of this fact, Eq. (2) is called the *Generalised Bayes Rule.* The rationale behind it is that the assessments modelled by $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ should be consistent with each other, in the sense that a combination of them should not allow us to raise the already specified supremum acceptable buying price for a gamble. More specifically, it turns out to be equivalent to the following two conditions:

$$\sup\left[f-\underline{P}(f)+g-\underline{P}(g|\mathcal{B})-(h-\underline{P}(h))\right]\geq 0\ \forall f,g,h\in\mathcal{L}(\Omega)$$
$$\sup\left[f-\underline{P}(f)+g-\underline{P}(g|\mathcal{B})-B(h-\underline{P}(h|B))\right]\geq 0\ \forall f,g,h\in\mathcal{L}(\Omega),B\in\mathcal{B}.$$

See [48, Section 6.3] for more details on this behavioural interpretation.

Unfortunately, this rule does not permit to determine $\underline{P}(\cdot|\mathcal{B})$ uniquely, in the sense that for a given coherent lower prevision there may be more than one conditional lower prevision satisfying Eq. (2) with it. If we fix the conditioning event B, we can distinguish a number of cases.

The first one is when $\underline{P}(B)>0$; then for every gamble f there is a unique real number λ such that $\underline{P}(B(f-\lambda))=0$, and as a consequence there is only one conditional lower prevision $\underline{P}(\cdot|B)$ that is compatible with $\underline{P}$.

The second case is when $\overline{P}(B)=0$. In that case, any real number λ satisfies $\underline{P}(B(f-\lambda))=0$, and this means that any conditional lower prevision $\underline{P}(\cdot|B)$ is compatible with $\underline{P}$. Thus, Walley's definition of joint coherence does not provide any guidance as to how to update our unconditional model. Although we shall not deal with this case in the paper, there are a number of possibilities that may be considered, such as the *full conditional measures* of Dubins [20] or the theory of zero layers by Coletti and Scozzafava [6, 32].

The last and most interesting case is that where the conditioning event has zero lower probability and positive upper probability, i.e., that of $\underline{P}(B)=0<\overline{P}(B)$. In that case, there may be an infinite number of conditional lower previsions that are compatible with $\underline{P}$; there were characterised in [37], where it was proven that they are bounded by the so-called natural and regular extensions.

Definition 3. Given $B\in\mathcal{B}$, the *natural extension* $\underline{E}(\cdot|B)$ induced by $\underline{P}$ is given by:

$$\underline{E}(f|B):=\begin{cases}\inf\{P(f|B):P\geq\underline{P}\} & \text{if } \underline{P}(B)>0\\ \min_{\omega\in B}f(\omega) & \text{otherwise}\end{cases}$$

for any gamble $f\in\mathcal{L}(\Omega)$.

The natural extension is vacuous when the conditioning event has zero lower probability, and is uniquely determined by Eq. (2) otherwise. Although it produces a conditional lower prevision that is coherent with $\underline{P}$, it is arguably too uninformative. A more informative alternative is called the regular extension:

Definition 4. Given $B\in\mathcal{B}$, the *regular extension* $\underline{R}(\cdot|B)$ induced by $\underline{P}$ is given by:

$$\underline{R}(f|B):=\begin{cases}\inf\{P(f|B):P(B)>0,P\geq\underline{P}\} & \text{if } \overline{P}(B)>0\\ \min_{\omega\in B}f(\omega) & \text{otherwise}\end{cases}$$

for any gamble $f\in\mathcal{L}(\Omega)$.

Hence, regular extension corresponds to applying Bayes' rule whenever possible on the set of precise models compatible with our conditional lower prevision, and to take then the lower envelope of the resulting set of conditional previsions. It has been proposed as an updating rule in a number of works in the literature [9, 29, 31, 34, 49]. In particular, it has been advocated also from the decision-theoretic point of view in [28, 42], under the name *full Bayesian updating.* From the behavioural point of view, it is based on an axiom of desirability that is stronger than the ones assumed in the definition of the natural extension: that a gamble f such that $\underline{P}(f) = 0 < \overline{P}(f)$ should be desirable for our subject. See [48, Appendix J] for more details.

Remark 1. The difference between the natural and the regular extension can also be understood if we consider a sensitivity analysis interpretation. Consider thus that our model is given by a set of previsions $\mathcal{M}$ (or $\mathcal{M}(\underline{P})$), and that we want to update this model assuming that a set $B \subseteq \Omega$ has been observed. Then we can express $\mathcal{M}$ as the union of two disjoint sets by $\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2$, where:

- $\mathcal{M}_1 := \{P \in \mathcal{M} : P(B) > 0\}$. Each of these P determines a conditional model $P(\cdot|B)$ by means of Bayes' rule.
- $\mathcal{M}_2 := \{P \in \mathcal{M} : P(B) = 0\}$. For these P we cannot apply Bayes' rule, and any conditional model $P(\cdot|B)$ is compatible with them.

Then if we maintain our sensitivity analysis interpretation, we may consider the set of conditional models that are compatible with at least one of the models in $\mathcal{M}$. There are a number of possibilities:

- $\mathcal{M}_2 = \emptyset$, or in other words, $\underline{P}(B) > 0$. Then we consider the set of conditional models derived from the elements in $\mathcal{M}_1$. The lower envelope of this set produces both the natural and the regular extension, which in this case coincide.
- $\mathcal{M}_1 = \emptyset$, i.e., $\overline{P}(B) = 0$. Then we consider the set of conditional models derived from the elements in $\mathcal{M}_2$, which is the set of all $P(\cdot|B)$. Its lower envelope is the vacuous lower prevision. Again here the natural and the regular extensions coincide.
- $\mathcal{M}_1 \neq \emptyset \neq \mathcal{M}_2$, i.e., $\underline{P}(B) = 0 < \overline{P}(B)$. Then the regular extension considers only the conditional models derived from the elements of $\mathcal{M}_1$, while the natural extension takes also into account the ones derived from $\mathcal{M}_2$ (and becomes thus vacuous). ♦

We see that the natural extension is dominated by the regular extension, and they coincide when the conditioning event B has either positive lower probability or zero upper probability. However, the natural extension may be strictly smaller than the regular extension when the conditioning event has zero lower probability and positive upper probability (see for instance Example 3 later on).

Our interest in these two updating rules for coherent lower previsions lies in the following theorem:

Proposition 2. [37, Theorem 9] *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\Omega)$ and $\mathcal{B}$ a partition of Ω such that $\overline{P}(B) > 0$ for any $B \in \mathcal{B}$. Then a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ is coherent with $\underline{P}$ if and only if $\underline{P}(f|B) \in [\underline{E}(f|B), \underline{R}(f|B)]$ for every $f \in \mathcal{L}(\Omega)$ and every $B \in \mathcal{B}$.*

Since Eq. (2) holds trivially when $\overline{P}(B) = 0$, the problem of coherently updating $\underline{P}$ into a conditional lower prevision $\underline{P}(f|\mathcal{B})$ only needs to be studied for those conditioning events B with positive upper probability, and then Proposition 2 tells us that the possible models lie between the natural and the regular extension[3]. In

[3]The key issue in the difference between the natural and the regular extensions seems to be that of conditioning on a set of probability zero (see again the comments in Remark 1): either we avoid this possibility whenever possible (like with the regular extension) or for those P with $P(B) = 0$ we specify a set of full conditional measures. The most extreme case is that where we

this paper, we are going to characterise the equality between these two conditional lower previsions. We shall focus on one interesting particular case of coherent lower previsions: the 2-monotone ones.

3. Updating 2-monotone lower previsions

One important instance of coherent lower previsions are the n-monotone ones, which were first introduced by Choquet in [4]:

Definition 5. A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ is called *n-monotone* if and only if

$$\underline{P}\left(\bigvee_{i=1}^{p} f_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\dots,p\}} (-1)^{|I|+1} \underline{P}\left(\bigwedge_{i\in I} f_i\right) \tag{3}$$

for all $2 \leq p \leq n$, and all $f_1, \dots, f_p$ in $\mathcal{L}(\Omega)$, where $\vee$ denotes the point-wise maximum and $\wedge$ the point-wise minimum.

In particular, a coherent lower probability $\underline{P} : \mathcal{P}(\Omega) \to [0,1]$ is n-monotone when

$$\underline{P}\left(\bigcup_{i=1}^{p} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\dots,p\}} (-1)^{|I|+1} \underline{P}\left(\bigcap_{i\in I} A_i\right) \tag{4}$$

for all $2 \leq p \leq n$, and all subsets $A_1, \dots, A_p$ of Ω.

Remark 2. We mentioned in Section 2 that a coherent lower prevision is not determined uniquely by its restriction to events, in the sense that a coherent lower probability may be the restriction of many different coherent lower previsions. Interestingly, we can uniquely determine this extension when we require in addition the property of n-monotonicity, in the following sense: given a n-monotone lower probability, its *only* n-monotone extension to $\mathcal{L}(\Omega)$ is the Choquet integral [18] with respect to this non-additive measure [14, 47], so we have that

$$\underline{P}(f) := (C)\int f d\underline{P} = \inf f + \int_{\inf f}^{\sup f} \underline{P}(f \geq t)dt \quad \text{for every } f \in \mathcal{L}(\Omega).$$

Note, however, that there is a subtlety here: if $\underline{P}$ is a n-monotone lower prevision on $\mathcal{L}(\Omega)$, its restriction to events is a n-monotone lower probability; however, a n-monotone lower probability on $\mathcal{P}(\Omega)$ can be in general extended by many different coherent lower previsions. Out of these, only one is n-monotone on all gambles: the one determined by the Choquet integral.♦

A coherent lower prevision on $\mathcal{L}(\Omega)$ that is n-monotone for all $n \in \mathbb{N}$ is called *completely monotone*, and its restriction to events is a *belief function*. One example of completely monotone coherent lower previsions are the vacuous ones in Eq. (1); the linear previsions are another one, and they moreover satisfy Eq. (3) with equality for every n.

In particular, a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ is 2-monotone if and only if it satisfies Eq. (3) for $n = 2$, that is, if and only if

$$\underline{P}(f \vee g) + \underline{P}(f \wedge g) \geq \underline{P}(f) + \underline{P}(g)$$

---

consider the set of *all* full conditional measures, and this produces the natural extension. Indeed, the separately coherent conditional lower previsions $\underline{P}(\cdot|B)$ between the natural and the regular extension can be seen to correspond to a subset of full conditional measures for those $P \geq \underline{P}$ s.t. $P(B) = 0$.

for every $f, g \in \mathcal{L}(\Omega)$. On the other hand, we deduce from Eq. (4) that a coherent lower probability on $\mathcal{P}(\Omega)$ is 2-monotone whenever

$$\underline{P}(A \cup B) + \underline{P}(A \cap B) \geq \underline{P}(A) + \underline{P}(B) \text{ for every } A, B \subseteq \Omega.$$

This property is also called *supermodularity* or *convexity*.

In this section, we are going to determine under which conditions a 2-monotone lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ can be uniquely updated to a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ that is coherent with $\underline{P}$, in the sense of Eq. (2). In order to do this, we shall use an equivalent expression for the conditional lower probability determined by regular extension:

Proposition 3. [47, Theorem 7.2] *Let $\underline{P}$ be a 2-monotone lower prevision on $\mathcal{L}(\Omega)$, and consider $B \subseteq \Omega$ such that $\overline{P}(B) > 0$. Then for any event A,*

$$\underline{R}(A|B) = \begin{cases} \dfrac{\underline{P}(A \cap B)}{\underline{P}(A \cap B) + \overline{P}(A^c \cap B)} & \text{if } \overline{P}(A^c \cap B) > 0, \\ 1 & \text{otherwise.} \end{cases} \tag{5}$$

Moreover, $\underline{R}(\cdot|B)$ is 2-monotone on events.

The proof makes use of the fact that, for any 2-monotone lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ and any $A \subseteq B \subseteq \Omega$, there is some $P \geq \underline{P}$ such that $P(A) = \underline{P}(A)$ and $P(B) = \underline{P}(B)$ [47, Corollary 6.5]. This allows to deduce that

$$\underline{R}(A|B) \leq \frac{\underline{P}(A \cap B)}{\underline{P}(A \cap B) + \overline{P}(A^c \cap B)} \quad \text{if } \overline{P}(A^c \cap B) > 0,$$

while the converse inequality holds for any coherent lower prevision.

A similar result is also presented in [28]. We shall show in Example 4 later on that in general $\underline{R}(\cdot|B)$ need not be 2-monotone on gambles, because it will not correspond to the Choquet integral with respect to its restriction to events (see again Remark 2). Indeed, we shall see that 2-monotonicity on gambles is only guaranteed when the conditioning event has zero lower probability and positive upper probability.

To see that Eq. (5) does not hold without the assumption of 2-monotonicity, consider the following example:

Example 1. Consider $\Omega = \{a, b, c, d\}$ and let P_1, P_2 be the linear previsions determined by the mass functions p_1, p_2 given by

	a	b	c	d
p_1	0.5	0.5	0	0
p_2	0.25	0.25	0.25	0.25

It has been shown in [47, Section 6] that the lower envelope $\underline{P}$ of $\{P_1, P_2\}$ is a coherent lower prevision that is not 2-monotone. Consider $B = \{a, b\}$ and $A = \{a\}$. Then $\overline{P}(A^c \cap B) = \overline{P}(\{b\}) = 0.5 > 0$, and

$$\frac{\underline{P}(A \cap B)}{\underline{P}(A \cap B) + \overline{P}(A^c \cap B)} = \frac{\underline{P}(\{a\})}{\underline{P}(\{a\}) + \overline{P}(\{b\})} = \frac{0.25}{0.25 + 0.5} = \frac{1}{3};$$

on the other hand any $P \geq \underline{P}$ is given by $\alpha P_1 + (1 - \alpha) P_2$, where $\alpha \in [0, 1]$; since $P_1(\{a\}) = P_1(\{b\})$ and $P_2(\{a\}) = P_2(\{b\})$, it follows that any $P \geq \underline{P}$ must satisfy

$P(\{a\}) = P(\{b\})$ too, whence

$$\underline{R}(A|B) = \inf\{P(\{a\}|\{a,b\}) : P \geq \underline{P}, P(\{a,b\}) > 0\} = 0.5.$$

Hence, Eq. (5) does not hold. ♦

The key in this example is that there is no linear prevision $P \geq \underline{P}$ such that $P(A \cap B) = \underline{P}(A \cap B) = P_2(\{a\})$ and $P(A^c \cap B) = \overline{P}(A^c \cap B) = P_1(\{b\})$. Note also that in the example $\underline{P}(B) > 0$ and as a consequence the natural and the regular extensions coincide.

Example 2. If we want instead an example when the natural and regular extensions do not coincide and the regular extension is not given by (5), it suffices to take the same coherent lower prevision as in Example 1, and consider $B = \{c, d\}$ and $A = \{c\}$. In that case the lower probability of B is zero, so $\underline{E}(A|B) = 0$. On the other hand,

$$\begin{aligned} \underline{R}(A|B) &= \inf\{P(A|B) : P(B) > 0, P \geq \underline{P}\} \\ &= \inf\left\{\tfrac{\alpha p_1(\{c\})+(1-\alpha)p_2(\{c\})}{\alpha p_1(\{c,d\})+(1-\alpha)p_2(\{c,d\})} : \alpha \in [0,1)\right\} = \tfrac{1}{2}. \end{aligned}$$

If we apply Eq. (5) we obtain that

$$\frac{\underline{P}(A \cap B)}{\underline{P}(A \cap B) + \overline{P}(A^c \cap B)} = \frac{\underline{P}(\{c\})}{\underline{P}(\{c\}) + \overline{P}(\{d\})} = \frac{0}{0 + 0.5} = 0.$$

Hence, natural and regular extensions do not coincide and the latter cannot be computed by means of Eq. (5). ♦

From Proposition 3 we deduce the following:

Proposition 4. *Let $\underline{P}$ be a 2-monotone lower prevision on $\mathcal{L}(\Omega)$ and consider $B \subseteq \Omega$ such that $\underline{P}(B) = 0 < \overline{P}(B)$. Then for any gamble f it holds that*

$$\underline{R}(f|B) = \min_{\omega \in C} f(\omega),$$

where C is the smallest subset of B satisfying $\underline{R}(C|B) = 1$.

Proof. First of all, since $\underline{P}(B) = 0$ implies that $\underline{P}(A \cap B) = 0$, we deduce from Proposition 3 that $\underline{R}(A|B) = 0$ if $\overline{P}(A^c \cap B) > 0$. The same formula tells us also that $\underline{R}(A|B) = 1$ if $\overline{P}(A^c \cap B) = 0$. Hence, for such a B the restriction to events of $\underline{R}(\cdot|B)$ is 0–1-valued. Applying Lemma 1, $\underline{R}(\cdot|B)$ is the unique coherent extension to $\mathcal{L}(\Omega)$ of its restriction to events, and it is given by

$$\underline{R}(f|B) = \max_{D : \underline{P}(D|B)=1} \min_{\omega \in D} f(\omega).$$

Since we are dealing here with finite spaces, it follows that the filter of sets $\{D \subseteq \Omega : \underline{R}(D|B) = 1\}$ is fixed and has thus a smallest set C. Hence, $\underline{R}(f|C) = \min_{\omega \in C} f(\omega)$ and the proof is complete. □

Interestingly, this shows that, if the lower prevision $\underline{P}$ satisfies 2-monotonicity, when the conditioning event B has zero lower probability and positive upper probability, the regular extension $\underline{R}(\cdot|B)$ is a completely monotone lower prevision, even if the lower prevision $\underline{P}$ we start from is not completely monotone.

Using these results, we can determine in which cases the natural and regular extensions coincide:

Proposition 5. *Let $\underline{P}$ be a 2-monotone lower prevision on $\mathcal{L}(\Omega)$, and consider $B \subseteq \Omega$ with $\overline{P}(B) > 0 = \underline{P}(B)$. The following are equivalent:*

(1) $\underline{E}(f|B) = \underline{R}(f|B)$ *for every* $f \in \mathcal{L}(\Omega)$.
(2) $\underline{E}(A|B) = \underline{R}(A|B)$ *for every* $A \subseteq \Omega$.
(3) $\overline{P}(\{\omega\}) > 0$ *for every* $\omega \in B$.

Proof. ($1 \Leftrightarrow 2$) The direct implication is trivial. For the converse, note that the restrictions to events of both $\underline{E}(\cdot|B)$ and $\underline{R}(\cdot|B)$ are 0–1-valued. Applying Lemma 1, we deduce that $\underline{E}(\cdot|B) = \underline{R}(\cdot|B)$ on gambles if and only if they agree on events.

($2 \Rightarrow 3$) To see that the second statement implies the third, note that if there is some ω in B for which $\overline{P}(\{\omega\}) = 0$, then given $A = B \setminus \{\omega\}$, applying Eq. (5) it holds that $\underline{R}(A|B) = 1$ while $\underline{E}(A|B) = 0$, because $\underline{P}(B) = 0$.

($3 \Rightarrow 2$) Conversely, since $\underline{P}(B) = 0$, it follows that $\underline{E}(A|B) = 0$ for any $A \subseteq \Omega$ such that $B \cap A^c \neq \emptyset$. Now, if $\overline{P}(\{\omega\}) > 0$ for every $\omega \in B$, we deduce that $\overline{P}(A^c \cap B) > 0$ for every $A \subseteq \Omega$ with $B \cap A^c \neq \emptyset$. Applying Proposition 3, we deduce that

$$\underline{R}(A|B) = \frac{\underline{P}(A \cap B)}{\underline{P}(A \cap B) + \overline{P}(A^c \cap B)} = \frac{0}{\overline{P}(A^c \cap B)} = 0,$$

where the one-but-last equality holds because $\underline{P}(A \cap B) = \underline{P}(B) = 0$.

On the other hand, for any $A \subseteq \Omega$ such that $A^c \cap B = \emptyset$, it holds that $B \subseteq A$, and thus $\underline{E}(A \mid B) = \min_{\omega \in B} I_A(\omega) = 1 = \underline{R}(A|B)$. We conclude that the regular and natural extensions coincide. □

The equivalence between the first and the third statements can also be derived easily from Proposition 4: $\underline{R}(\cdot|B) = \underline{E}(\cdot|B)$ if and only if there is no proper subset C of B such that $\underline{R}(C|B) = 1$. From Proposition 3, this holds if and only if $\overline{P}(B \cap C^c) > 0$ for every proper subset C of B, and this in turn is equivalent to the third statement.

Now, since the natural and the regular extensions coincide whenever the conditioning event has positive lower or zero upper probability, we immediately deduce the following:

Theorem 1. *Let $\underline{P}$ be a 2-monotone lower prevision on $\mathcal{L}(\Omega)$, and let $\mathcal{B}$ be a partition of Ω. Then:*

$$\underline{E}(\cdot|\mathcal{B}) = \underline{R}(\cdot|\mathcal{B}) \Leftrightarrow \overline{P}(\{\omega\}) > 0 \text{ for every } \omega \in B \in \mathcal{B} \text{ s.t. } \underline{P}(B) = 0 < \overline{P}(B).$$

To see that this result cannot be extended to arbitrary coherent lower previsions, it suffices to consider the coherent lower prevision $\underline{P}$ in Example 2, for which $\overline{P}(\{\omega\}) > 0$ for every $\omega \in \Omega$ but where the natural and the regular extensions do not necessarily coincide: we saw there that given $B = \{c, d\}$ and $A = \{c\}$, we have $\underline{R}(A|B) = 0.5 > 0 = \underline{E}(A|B)$.

4. Coherent updating of completely monotone lower previsions

We consider next the case where the unconditional lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ is completely monotone. Then its restriction to events is a belief function, i.e., it satisfies Eq. (4) for every natural number n; its conjugate upper probability is a plausibility function.

One of the most important updating rules for plausibility functions is Dempster's rule of conditioning [17, 43]. Given a plausibility function $\overline{P}$ on Ω and a conditioning event B with $\overline{P}(B) > 0$, the conditional plausibility is defined by

$$\overline{P}(A|B) := \frac{\overline{P}(A \cap B)}{\overline{P}(B)}.$$

However, this conditional upper probability is not coherent with the unconditional upper probability $\overline{P}$ [54][4]; see also [48, Section 5.13] and [50]. This is not surprising because, as discussed by Dubois et al. in [7, 22, 26, 27], Dempster's rule of conditioning is meant for a problem of belief revision, where we update our model by considering new evidence. On the contrary, in this paper we are dealing instead with a problem of *focusing*, for which it makes more sense to consider the rules of natural and regular extension.

Taking into account that a completely monotone lower prevision is in particular 2-monotone, given a conditioning event B with $\overline{P}(B) > 0$, we can compute the regular extension by means of Eq. (5). This formula has also been derived in a few papers ([29, Theorem 3.4]; [31, Proposition 4]; and see also [9, 17]). Moreover, it has been established in [29, 31, 46] that the restriction of $\underline{R}(\cdot|B)$ to events is a belief function whenever the conditioning event B satisfies $\underline{P}(B) > 0$.

The equality between the natural and the regular extensions of $\underline{P}$ is characterised by Theorem 1. In this section, we give equivalent conditions in terms of the focal elements of $\underline{P}$.

Definition 6. [43] Given a belief function $\underline{P}$ on $\mathcal{P}(\Omega)$, its *Möbius inverse* is the map $m : \mathcal{P}(\Omega) \to [0, 1]$ given by

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \underline{P}(B) \ \text{ for every } A \subseteq \Omega.$$

It holds that

$$\underline{P}(A) = \sum_{B \subseteq A} m(B), \tag{6}$$

and m is called a *basic probability assignment* within the evidential theory of Shafer. For the plausibility function $\overline{P}$ that is conjugate to $\underline{P}$, it holds that

$$\overline{P}(A) = \sum_{B \cap A \neq \emptyset} m(B) \text{ for every } A \subseteq \Omega. \tag{7}$$

For the results in this section, it shall be interesting to work with the focal elements of the belief function:

Definition 7. [43] Given a belief function $\underline{P}$ with Möbius inverse m, a subset $B \subseteq \Omega$ is called a *focal element* when $m(B) > 0$. The union F of all the focal elements of $\underline{P}$ is called the *support*[5] of the belief function $\underline{P}$.

[4]Interestingly, Dempster's rule of conditioning is coherent in the particular case of possibility measures, as we shall discuss in Section 5.

[5]This is called *core* by Shafer. We have followed a reviewer's suggestion and called it *support* instead. We think this terminology is better because it extends the idea of support from probability measures. Note also that the term *core* is sometimes used to refer to the credal set associated with the belief function.

We shall be particularly interested in those belief functions whose focal elements are a covering of the possibility space Ω. Following Daniel [8], we call these belief functions non-exclusive:

Definition 8. [8] A belief function $\underline{P}$ with support F is called *non-exclusive* when $F = \Omega$, and it is called *exclusive* otherwise.

Since $\overline{P}(F^c) = \sum_{B \cap F^c \neq \emptyset} m(B) = 0$, given an exclusive belief function, any set included in F^c will have zero upper probability. Equivalently, if $\underline{P}$ is a non-exclusive belief function, any subset B of Ω has a positive plausibility.

We have mentioned repeatedly that $\underline{E}(\cdot|B) = \underline{R}(\cdot|B)$ when $\underline{P}(B) > 0$ or $\overline{P}(B) = 0$. Hence, the natural and regular extensions will agree as soon as there is no conditioning event with zero lower probability and positive upper probability. This situation is characterised by the following definition:

Definition 9. A belief function is called *scattered* if for every focal element B, it holds that $m(\{\omega\}) > 0$ for every $\omega \in B$.

Scattered and non-exclusive belief functions can be characterised in the following way:

Proposition 6. *Let $\underline{P}$ be a belief function on $\mathcal{P}(\Omega)$.*

(1) $\underline{P}$ is scattered if and only if for any $B \subseteq \Omega$ either $\overline{P}(B) = 0$ or $\underline{P}(B) > 0$.
(2) $\underline{P}$ is non-exclusive if and only if for any $B \subseteq \Omega$, $\overline{P}(B) > 0$.
(3) $\underline{P}$ is non-exclusive and scattered if and only if $\underline{P}(B) > 0$ for every $B \subseteq \Omega$.

Proof. (1) Assume first of all that $\underline{P}$ is scattered, and let F be the support of $\underline{P}$. Given a subset B of Ω, there are two possibilities: either (a) $B \subseteq F^c$, and then $\overline{P}(B) = 0$; or (b) $B \cap F \neq \emptyset$. In this second case the definition of the support and of a scattered belief function implies that $\underline{P}(B) > 0$.

Conversely, if $\underline{P}$ is not scattered there is some focal element B such that $m(\{\omega\}) = 0$ for some $\omega \in B$. Then Eqs. (6) and (7) imply that $\underline{P}(\{\omega\}) = 0 < \overline{P}(\{\omega\})$, a contradiction.

(2) The second statement follows trivially from Definition 8.
(3) The third statement is a consequence of the first and the second. □

When the conditioning event B has zero lower probability and positive upper probability the equality between the natural and the regular extensions is characterised by Proposition 5: we need that $\overline{P}(\{\omega\}) > 0$ for every $\omega \in B$; in the case of belief functions, this is equivalent to $B \subseteq F$, the support of the belief function. From this we deduce the following result:

Proposition 7. *Let $\underline{P}$ be a completely monotone lower prevision on $\mathcal{L}(\Omega)$, and let μ denote the belief function that is the restriction of $\underline{P}$ to events. Then,*

$$\underline{E}(\cdot|B) = \underline{R}(\cdot|B)\ \forall B \subseteq \Omega \Leftrightarrow \mu \text{ is either non-exclusive or scattered.}$$

Proof. First of all, if μ is scattered it follows from Proposition 6 that for any $B \subseteq \Omega$ either $\overline{P}(B) = 0$ or $\underline{P}(B) > 0$, and in any of these two cases $\underline{E}(\cdot|B) = \underline{R}(\cdot|B)$.

On the other hand, if μ is non-exclusive, it follows from Proposition 6 that $\overline{P}(B) > 0$ for every $B \subseteq \Omega$. Applying Proposition 5, we deduce that in this case $\underline{R}(\cdot|B) = \underline{E}(\cdot|B)$.

Conversely, assume that the belief function μ is both exclusive and not scattered. Then there is some $B \subseteq \Omega$ such that $\underline{P}(B) = 0 < \overline{P}(B)$. Applying (6) and (7), we deduce that there is some focal element C such that $C \cap B \neq \emptyset \neq C \cap B^c$. On the other hand, if μ is exclusive there is some $\omega \in F^c$, where F denotes the support of μ. Hence, given $D = B \cup \{\omega\}$,

$$\underline{P}(D) = \sum_{C \text{ focal}:C \subseteq D} m(C) = \sum_{C \text{ focal}:C \subseteq B} m(C) = \underline{P}(B) = 0,$$

Moreover, from the monotonicity of $\overline{P}$ it follows that $\overline{P}(D) \geq \overline{P}(B) > 0$. Now, since $\omega \in D$ satisfies $\overline{P}(\{\omega\}) = 0$, we deduce that $\underline{R}(\{\omega\}^c|D) = 1 > 0 = \underline{E}(\{\omega\}^c|D)$. □

This result allows to easily provide an example of a completely monotone lower prevision whose natural and regular extensions do not coincide[6]:

Example 3. Consider $\Omega = \{a, b, c, d\}$, and let $\underline{P}$ be the completely monotone lower prevision given by

$$\underline{P}(f) = \min\{f(b), f(c)\} \; \forall f \in \mathcal{L}(\Omega).$$

The restriction to events of $\underline{P}$ is the belief function associated with the basic probability assignment m where

$$m(\{b, c\}) = 1 \text{ and } m(C) = 0 \text{ for every } C \neq \{b, c\}.$$

Obviously, this belief function is exclusive. If we take $B = \{a, b\}$ and $A = \{b\}$, then any probability $P \geq \underline{P}$ satisfying $P(B) > 0$ must satisfy $P(\{b\}) > 0$, because $P(\{a\}) \leq \overline{P}(\{a\}) = 0$. But then P will satisfy $P(A|B) = 1$, and from this we deduce that

$$\underline{R}(A|B) = 1 > 0 = \underline{E}(A|B),$$

where $\underline{E}(A|B)$ is equal to 0 because $\underline{P}(B) = 0$. Hence, the natural and regular extensions do not coincide. ♦

Moreover, for completely monotone lower previsions we can give an alternative expression of the regular extension to that in Proposition 4.

Proposition 8. *Let $\underline{P}$ be a completely monotone lower prevision, and let F be the support of its associated belief function. Then for any $B \subseteq \Omega$ such that $\underline{P}(B) = 0 < \overline{P}(B)$,*

$$\underline{R}(f|B) = \min_{\omega \in B \cap F} f(\omega) \quad \textit{for every } f \in \mathcal{L}(\Omega).$$

Proof. From Proposition 4, it suffices to show that $B \cap F$ is the smallest subset C of B such that $\underline{R}(C|B) = 1$.

On the one hand, since $\overline{P}(F^c) = 0$, and as a consequence also $\overline{P}(F^c \cap B) = 0$, we deduce from Proposition 3 that $\underline{R}(B \cap F|B) = 1$.

Conversely, consider a subset C of B such that $\underline{R}(C|B) = 1$. Then if it were $\overline{P}(C^c \cap B) > 0$, then Proposition 3 would imply that

$$\underline{R}(C|B) = \frac{\underline{P}(C \cap B)}{\underline{P}(C \cap B) + \overline{P}(C^c \cap B)} = \frac{0}{0 + \overline{P}(C^c \cap B)} = 0,$$

taking into account that $\underline{P}(C \cap B) \leq \underline{P}(B) = 0$. This is a contradiction.

We deduce that $\overline{P}(C^c \cap B) = 0$, and as a consequence it must be $C^c \cap B \subseteq F^c$. Hence, $F \subseteq C \cup B^c$ and therefore $F \cap B \subseteq C \cap B = C$. □

We can summarise the above results in the following theorem. Its proof is immediate and therefore omitted.

[6]Recall that the coherent lower prevision in Example 2 was not 2-monotone.

Theorem 2. *Let $\underline{P}$ be a completely monotone lower prevision on $\mathcal{L}(\Omega)$ and let $\mathcal{B}$ be a partition of Ω. If the restriction to events μ of $\underline{P}$ is either non-exclusive or scattered, then $\underline{E}(\cdot|\mathcal{B}) = \underline{R}(\cdot|\mathcal{B})$.*

Note that the sufficient condition in this theorem is not necessary: for instance, it may be that μ is neither non-exclusive nor scattered and $\mu(B) > 0$ for every B in the partition $\mathcal{B}$, and then $\underline{E}(\cdot|\mathcal{B}) = \underline{R}(\cdot|\mathcal{B})$. The result in the theorem proves the equality between the natural and the regular extension for all the conditioning events, that is, irrespective of the partition $\mathcal{B}$.

4.1. **Random Sets.** One context where completely monotone lower previsions arise naturally is that of measurable multi-valued mappings, or random sets [17, 41].

Definition 10. Let $(X, \mathcal{A}, P)$ be a probability space, $(\Omega, \mathcal{P}(\Omega))$ a measurable space, where Ω is finite, and $\Gamma : X \to \mathcal{P}(\Omega)$ a non-empty multi-valued mapping. It is called a *random set* when it satisfies the following measurability condition:

$$\Gamma_*(A) := \{x \in X : \Gamma(x) \subseteq A\} \in \mathcal{A} \quad \text{ for every } A \subseteq \Omega.$$

Its associated *lower probability* $P_{*\Gamma} : \mathcal{P}(\Omega) \to [0, 1]$ is the functional given by

$$P_{*\Gamma}(A) = P(\Gamma_*(A)) \text{ for every } A \subseteq \Omega, \tag{8}$$

and it is a belief function.

The focal elements of $P_{*\Gamma}$ are given by

$$\{A \subseteq \Omega : P(\Gamma^{-1}(A)) > 0\},$$

and its Möbius inverse is given by $m = P \circ \Gamma^{-1}$, where Γ^{-1} is defined by:

$$\Gamma^{-1}(A) = \{x : \Gamma(x) = A\} \quad \text{ for every } A \subseteq \Omega.$$

The conjugate plausibility measure is denoted by P^*_Γ and called the *upper probability* of the random set Γ. It satisfies

$$P^*_\Gamma(A) = 1 - P_{*\Gamma}(A^c) = P(\{x : \Gamma(x) \cap A \neq \emptyset\}),$$

where the set $\{x : \Gamma(x) \cap A \neq \emptyset\}$ is the *upper inverse* of A by Γ, and is usually denoted by $\Gamma^*(A)$. The Choquet integral with respect to $P_{*\Gamma}$ is a completely monotone lower prevision on $\mathcal{L}(\Omega)$. If we want to update this completely monotone lower prevision, we can use the natural or the regular extensions, that, by Proposition 7, coincide on all gambles if and only if $P_{*\Gamma}$ is either non-exclusive or scattered. These two properties can be easily characterised in terms of the images of Γ:

Proposition 9. *Let $(X, \mathcal{A}, P)$ be a probability space, Ω a finite set and $\Gamma : X \to \mathcal{P}(\Omega)$ a random set with associated lower probability $P_{*\Gamma}$. Let F denote the support of $P_{*\Gamma}$.*

(1) $P_{*\Gamma}$ is non-exclusive $\Leftrightarrow F = \Omega \Leftrightarrow P^*_\Gamma(B) > 0\ \forall B \subseteq X \Leftrightarrow P^*_\Gamma(\{\omega\}) > 0$ for all $\omega \in \Omega \Leftrightarrow P(\{x : \omega \in \Gamma(x)\}) > 0$ for all $\omega \in \Omega$.

(2) $P_{*\Gamma}$ is scattered $\Leftrightarrow \forall \omega \in F,\ P(\Gamma^{-1}(\omega)) > 0$.

Moreover, $\underline{E}(\cdot|B) = \underline{R}(\cdot|B)$ for all $B \subseteq \Omega$ if and only if $P_{*\Gamma}$ is either non-exclusive or scattered.

Proof. The result follows immediately from Propositions 6 and 7. □

One interesting interpretation of random sets is the *epistemic* one, where they are seen as models for the imprecise knowledge of a random variable [33]. In that case, our information about this random variable is provided by the *measurable selections* of Γ: those measurable mappings $U: X \to \Omega$ such that $U(x) \in \Gamma(x)\ \forall x \in X$. We shall denote by $S(\Gamma)$ the set of measurable selections of Γ and by $P(\Gamma)$ the set of the probability measures they induce. This set is included in the class $\mathcal{M}(P_{*\Gamma})$ of probabilities that dominate $P_{*\Gamma}$. Although both sets do not coincide in general, when Ω is finite it can be checked that:

Proposition 10. [39, Theorem 1] *Let $\Gamma : X \to \mathcal{P}(\Omega)$ be a random set, where Ω is finite. Then $Ext(M(P_{*\Gamma})) \subseteq P(\Gamma)$ and $M(P_{*\Gamma}) = Conv(Ext(M(P_{*\Gamma})))$.*

In other words, the lower probability of the random set can also be determined by the set of measurable selections. The epistemic interpretation is thus related to that of Bayesian sensitivity analysis we have mentioned before; in both cases we assume the existence of a precise model that is imprecisely observed. The difference lies in that the epistemic interpretation focuses on the random variable and the Bayesian sensitivity analysis on the probability distribution it induces; however, both give rise to the same coherent lower prevision (or probability).

Moreover, from [17] $M(P_{*\Gamma})$ has a finite number of extreme points, that are related to the permutations of the final space.

For any subset B such that $P^*_\Gamma(B) > 0$, the regular extension of $P_{*\Gamma}$ can be expressed by:

$$\underline{R}(f \mid B) = \min\{P(f|B) : P \in M(P_{*\Gamma}), P(B) > 0\} \quad \text{for every } f \in \mathcal{L}(\Omega).$$

The epistemic interpretation can be carried on towards the regular extension, in the sense that the regular extension can be computed as the envelope of the conditional expectations associated with the measurable selections:

Proposition 11. *Let $(X, \mathcal{A}, P)$ be a probability space, Ω a finite set and $\Gamma : X \to \mathcal{P}(\Omega)$ a random set with associated lower probability $P_{*\Gamma}$. Consider $B \subseteq \Omega$ with $P^*_\Gamma(B) > 0$. Then,*

$$\underline{R}(f \mid B) = \min\{P(f \mid B) : P \in P(\Gamma), P(B) > 0\} \quad \textit{for every } f \in \mathcal{L}(\Omega).$$

Proof. Since $P(\Gamma) \subseteq \mathcal{M}(P_{*\Gamma})$, one of the inequalities is obvious. Let us prove the other one. Let $\{P_1, \ldots, P_m\}$ be the class of the extreme points of $\mathcal{M}(P_{*\Gamma})$, and consider $P \geq P_{*\Gamma}$ such that $P(B) > 0$. From Proposition 10, there are $\alpha_1, \ldots, \alpha_m \geq 0$ such that $\alpha_1 + \cdots + \alpha_m = 1$ and $P = \alpha_1 P_1 + \ldots + \alpha_m P_m$.

Assume ex-absurdo that

$$P(f|B) < \min\{P_i(f|B) : P_i(B) > 0\}.$$

Then it follows that $P(f|B)P_i(B) < P_i(fI_B)$ if $P_i(B) > 0$, and $P(f|B)P_i(B) = P_i(fI_B)$ if $P_i(B) = 0$. Since $P(B) > 0$, there must be some i in $\{1, \ldots, m\}$ such that $P_i(B) > 0$ and $\alpha_i > 0$. As a consequence,

$$\begin{aligned} P(f \mid B) = \frac{P(fI_B)}{P(B)} &= \frac{\alpha_1 P_1(fI_B) + \cdots + \alpha_m P_m(fI_B)}{\alpha_1 P_1(B) + \cdots + \alpha_m P_m(B)} \\ &> \frac{\alpha_1 P(f \mid B) P_1(B) + \ldots + \alpha_m P(f \mid B) P_m(B)}{\alpha_1 P_1(B) + \ldots + \alpha_m P_m(B)} = P(f \mid B), \end{aligned}$$

a contradiction. We deduce that

$$\underline{R}(f|B) = \min\{P(f|B) : P \in P(\Gamma), P(B) > 0\} = \min\{P_i(f|B) : P_i(B) > 0\}. \quad \square$$

To conclude this section, we use random sets to establish that the conditional lower prevision we obtain when we update a completely monotone lower prevision by means of Generalised Bayes Rule is not necessarily completely monotone, even if its restriction to events is always a belief function, as shown in [29, 31]. The key here is that the conditional lower prevision determined by Generalised Bayes Rule does not correspond to the Choquet integral with respect to its restriction to events. To us, this seems to indicate that, even if complete monotonicity is a useful mathematical property, its behavioural interpretation remains unclear; in particular, it is a property that is not necessarily preserved by the updating procedures from the behavioural theory of imprecise probabilities.

Example 4. Consider the probability space $(X, \mathcal{P}(X), P)$, where $X = \{a, b, c, d, e\}$, and P is the probability measure determined by the equalities $P(a) = P(b) = 1/8$, and $P(c) = P(d) = P(e) = 1/4$. Let Γ be the multi-valued mapping $\Gamma : X \to \mathcal{P}(\{1, 2, 3, 4\})$ given by

$$\Gamma(a) = \{1\},\ \Gamma(b) = \{2\},\ \Gamma(c) = \{1, 4\},\ \Gamma(d) = \{2, 4\},\ \Gamma(e) = \{3, 4\}.$$

Let $P_{*\Gamma}$ denote the lower probability induced by this random set. This is a belief function, and the lower prevision $\underline{P}$ on $\mathcal{L}(\{1, 2, 3, 4\})$ given by $\underline{P}(f) = (C) \int f dP_{*\Gamma}$ is a completely monotone lower prevision.

It follows from Eq. (8) that

$$P_{*\Gamma}(\{1, 2, 3\}) = P(\{x : \Gamma(x) \subseteq \{1, 2, 3\}\}) = P(\{a, b\}) = \frac{1}{4} > 0.$$

As a consequence, the natural and regular extensions coincide, and we deduce from Proposition 11 that

$$\underline{R}(f|\{1, 2, 3\}) = \min\{P(f|\{1, 2, 3\}) : P \in P(\Gamma)\}. \tag{9}$$

Let us consider the gamble f on $\{1, 2, 3, 4\}$ given by $f(\omega) = 4 - \omega$ for all $\omega \in \{1, 2, 3, 4\}$. This gamble can be expressed by $f = 1\ \mathbb{I}_{1,2,3} + 1\ \mathbb{I}_{1,2} + 1\ \mathbb{I}_1$, i.e., it is the sum of comonotone functions. Since the Choquet integral with respect to a monotone set function is comonotone additive ([18, Proposition 5.1]), we deduce that the Choquet integral of f with respect to $\underline{R}(\cdot \mid \{1, 2, 3\})$ is given by:

$$(C) \int f\ d\underline{R}(\cdot|\{1, 2, 3\}) = 1 + \underline{R}(\{1, 2\}|\{1, 2, 3\}) + \underline{R}(\{1\}|\{1, 2, 3\}).$$

We deduce from Eq. (9) that

$$\underline{R}(\{1\}|\{1, 2, 3\}) = \frac{1}{6} \ \text{ and } \underline{R}(\{1, 2\}|\{1, 2, 3\}) = \frac{1}{2};$$

to see this, it suffices to determine the measurable selections that determine the smallest conditional probabilities. In the first case the measurable selection we get is given by the vector $(U_1(a), U_1(b), U_1(c), U_1(d), U_1(e)) = (1, 2, 4, 2, 3)$, and in the second case it is given by $(U_2(a), U_2(b), U_2(c), U_2(d), U_2(e)) = (1, 2, 4, 4, 3)$. From this, we deduce that

$$(C) \int f\ d\underline{R}(\cdot|\{1, 2, 3\}) = 5/3.$$

On the other hand, the smallest value of $\{P(f|\{1, 2, 3\}) : P \in P(\Gamma)\}$ is given by $7/4 > 5/3$, considering again the measurable selection U_2. This means that $\underline{R}(f|\{1, 2, 3\}) > (C) \int f d\underline{R}(\cdot|\{1, 2, 3\})$.

But it has been established in [14, 47] that if we have a 2-monotone lower probability on all events (as is the case for $\underline{R}(\cdot|\{1,2,3\})$, the *only* 2-monotone extension to all gambles is the Choquet integral. This means that the conditional lower prevision $\underline{R}(\cdot|\{1,2,3\})$ is not 2-monotone on $\mathcal{L}(\{1,2,3\})$. ♦

Note that in this example the natural and regular extensions coincide, and as a consequence there is only one conditional lower prevision $\underline{P}(\cdot|\{1,2,3\})$ satisfying Eq. (2) with $\underline{P}$. This means that in this case complete monotonicity cannot be preserved by Walley's updating procedure.

5. Coherent updating of minimum-preserving previsions

We narrow our focus a bit further by considering now the particular case of completely monotone lower previsions that are minimum-preserving, i.e., lower previsions $\underline{P}$ such that

$$\underline{P}(f \wedge g) = \min\{\underline{P}(f), \underline{P}(g)\}$$

for every pair of gambles f, g on Ω. They correspond to the Choquet integral with respect to their restriction to events, which is a necessity measure that we shall denote N. The conjugate upper prevision $\overline{P}$ is the Choquet integral with respect to the possibility measure Π that is determined by N using duality.

From Proposition 7, we deduce the following:

Corollary 1. *Let $\underline{P}$ be a minimum-preserving coherent lower prevision. Then $\underline{E}(\cdot|B) = \underline{P}(\cdot|B)$ for all $B \subseteq \Omega$ if and only if either of the following conditions holds:*

(i) $\overline{P}(\{\omega\}) > 0$ for all $\omega \in \Omega$.
(ii) $\underline{P}(\{\omega\}) = 1$ for some $\omega \in \Omega$.

Proof. Let N be the necessity measure that results from restricting $\underline{P}$ to events. From [43], all the focal elements of N are nested. It follows that $\underline{P}$ is non-exclusive if and only if Ω is a focal element; if this is the case then by Eq. (7) $\overline{P}(\{\omega\}) = \sum_{\omega \in B} m(B) \geq m(\Omega) > 0$ for every $\omega \in \Omega$. Conversely, if Ω is not a focal element, then the support F of N is a proper subset of Ω. Thus, if we take $\omega \notin F$, it holds that $\overline{P}(\{\omega\}) = 0$.

On the other hand, we deduce from Definition 9 that μ is scattered if and only if it corresponds to a degenerate probability measure: since the focal elements are nested, at most one singleton $\{\omega\}$ can be a focal element. Hence, $\underline{P}$ is scattered if and only if the only focal element is $\{\omega\}$, i.e., if and only if $\underline{P}$ corresponds to the assessment $\underline{P}(\{\omega\}) = 1$.

The result follows then from Proposition 7. □

The result in Corollary 1 can be simplified further taking into account an interesting result that de Cooman and Aeyels proved in [12] (see also [13]): a coherent *upper* prevision $\overline{P}$ on $\mathcal{L}(\Omega)$ is maximum-preserving if and only if its restriction to events is a 0–1-valued possibility measure. This is interesting because for every natural number n there is a one-to-one correspondence between n-monotone lower previsions and n-monotone lower probabilities, by means of the Choquet integral: the restriction to events of any n-monotone lower prevision is a n-monotone lower probability, and the Choquet integral w.r.t. the latter is a n-monotone lower prevision. However, the correspondence only holds in one direction for the particular case of minimum-preserving lower previsions: although the restriction to events of a minimum-preserving lower prevision is again minimum-preserving (that is, a necessity measure), the Choquet integral with respect to a necessity measure will not be in general minimum-preserving (although by the result mentioned above it shall be completely monotone).

If we assume then that the restriction to events of $\underline{P}$ is 0–1-valued and define $F := \{\omega : \overline{P}(\{\omega\}) = 1\}$, it turns out that F is the support of the possibility measure, since $\overline{P}(F^c) = 0$, and moreover it is the only focal element of $\overline{P}$: if there was another focal element $F' \subsetneq F$, then we would have

$$0 < m(F) \leq \overline{P}(F \setminus F') = \sum_{B \cap (F \setminus F') \neq \emptyset} m(B) \leq 1 - m(F') < 1,$$

a contradiction. Thus, it must be $m(F) = 1$, and as a consequence $\underline{P}$ is the vacuous lower prevision on F:

$$\underline{P}(f) = \min_{\omega \in F} f(\omega) \quad \forall f \in \mathcal{L}(\Omega).$$

In the context of game theory this means that the restriction of $\underline{P}$ to events is the *unanimity game* ([19, 44]) associated with F. These are important because they correspond to the extreme points in the class of belief functions.

Now, given a conditioning event $B \subseteq F$, there are a number of possibilities:

- $F \subseteq B^c$. Then $\overline{P}(B) = 0$ and both the natural and regular extensions are vacuous on B.
- $B \cap F \neq \emptyset \neq B^c \cap F$. Then $\underline{P}(B) = 0 < 1 = \overline{P}(B)$, whence $\underline{E}(\cdot|B)$ is vacuous on B and $\underline{R}(\cdot|B)$ is vacuous on $B \cap F$. Hence, in that case the natural and regular extensions do not coincide.
- $F \subseteq B$. Then $\underline{P}(B) = 1$, whence $\underline{E}(\cdot|B)$ and $\underline{R}(\cdot|B)$ coincide. It follows from their definition that they correspond to the vacuous lower prevision on F.

Note that in this case the restriction of $\underline{P}$ to events is only scattered when F is a singleton (i.e., when $\underline{P}$ corresponds to the expectation operator with respect to a degenerate probability measure), and $\underline{P}$ is non-exclusive if and only if $F = \Omega$, meaning that $\underline{P}$ corresponds to the vacuous model. Hence, we only have the equality between the natural and the regular extensions for all $B \subseteq \Omega$ in these two extreme cases. They correspond to particular cases of *quasi-Bayesian belief functions* in the sense of [8].

We summarise the coherent updating of a minimum-preserving lower prevision in the following theorem. Its proof is immediate, and therefore omitted.

Theorem 3. *Let $\underline{P}$ be a minimum-preserving lower prevision on $\mathcal{L}(\Omega)$, and consider a partition $\mathcal{B}$ of Ω. Consider $F \subseteq \Omega$ such that $\underline{P}(f) = \min_{\omega \in F} f(\omega)\ \forall f \in \mathcal{L}(\Omega)$. Given $B \in \mathcal{B}$ and $f \in \mathcal{L}(\Omega)$,*

(1) $\underline{E}(f|B) = \begin{cases} \min_{\omega \in B} f(\omega) & \text{if } F \not\subseteq B \\ \min_{\omega \in F} f(\omega) & \text{if } F \subseteq B. \end{cases}$

(2) $\underline{R}(f|B) = \begin{cases} \min_{\omega \in B \cap F} f(\omega) & \text{if } F \not\subseteq B^c \\ \min_{\omega \in B} f(\omega) & \text{if } F \subseteq B^c. \end{cases}$

(3) $\underline{E}(f|B) = \underline{R}(f|B)$ *if and only if either* $F \subseteq B$ *or* $F \subseteq B^c$.

(4) *Moreover, a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ is coherent with $\underline{P}$ if and only if*

$$\min_{\omega\in B} f(\omega) \leq \underline{P}(f|B) \leq \min_{\omega\in B\cap F} f(\omega) \quad \forall f\in\mathcal{L}(\Omega), \forall B\in\mathcal{B} \ s.t\ B\cap F\neq\emptyset\neq B^c\cap F.$$

Remark 3. This theorem helps to illustrate the difference between the process of focusing we are considering in this paper and one of belief revision: in the latter, if we start with the vacuous lower prevision on F and observe some set B such that $B\cap F^c\neq\emptyset$, we should remove the elements from $\mathcal{M}(\underline{P})$ that are not compatible with the observation of B (i.e., those with $P(B)=0$), and this means that it makes no sense to consider a vacuous updated prevision on a set that is not included in $B\cap F$. This would rule out the use of the natural extension.

However, in our context if we update the models $P\in\mathcal{M}(\underline{P})$ for which $P(B)=0$ (i.e., those with $P(B^c\cap F)=1$), then we can consider any conditional linear prevision because Bayes' rule holds trivially, and then their lower envelope, which is the natural extension, produces the vacuous lower prevision on F. See again Remark 1. ♦

Remark 4. This result can be applied in particular in the context of multi-valued mappings. Assume that $\Gamma: X\to\mathcal{P}(\Omega)$ is a random set, and let $P_{*\Gamma}$ be its associated lower probability. Then it follows from [38, Proposition 7] that $P_{*\Gamma}$ is a necessity measure if and only if there is some $A\in\mathcal{A}$ with $P(A)=1$ such that for every $\omega_1,\omega_2\in A$, either $\Gamma(\omega_1)\subseteq\Gamma(\omega_2)$ or $\Gamma(\omega_2)\subseteq\Gamma(\omega_1)$. Taking into account Proposition 9 and Corollary 1, we deduce that $\underline{E}(\cdot\mid B)=\underline{R}(\cdot\mid B)$ for all $B\subseteq\Omega$ if and only if one of the following conditions holds:

(1) $\Gamma=\Omega$ a.s.
(2) There is some $\omega\in\Omega$ such that $\Gamma=\{\omega\}$ a.s.

The first of these conditions is equivalent to $P_{*\Gamma}$ being non-exclusive, and the second one is equivalent to $P_{*\Gamma}$ being scattered. ♦

From Theorem 3, the bounds determined by natural and regular extension are both minimum-preserving, and as a consequence they correspond to the Choquet integral of their respective restrictions to events. However, not every conditional lower prevision between them is also minimum-preserving, as we can see in the following example:

Example 5. Consider the finite set $\Omega=\{a,b,c,d\}$ and let $\underline{P}$ be the coherent lower prevision from Example 3, given by

$$\underline{P}(f)=\min\{f(b),f(c)\} \text{ for every } f\in\mathcal{L}(\Omega).$$

Define $\underline{P}(\cdot|B)$ as $\underline{P}(f|B)=0.5\cdot\underline{E}(f|B)+0.5\cdot\underline{R}(f|B)$ for every $B\subseteq\Omega$ and every gamble $f\in\mathcal{L}(\Omega)$. Then it follows from Proposition 2 that for every partition $\mathcal{B}$ of Ω, the separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ thus defined is coherent with $\underline{P}$.

To see that $\underline{P}(\cdot|B)$ is not minimum-preserving for all subsets B of Ω, consider $B=\{a,b\}$. Then it follows from Theorem 3 that

$$\underline{E}(f|B)=\min\{f(a),f(b)\} \text{ and } \underline{R}(f|B)=f(b) \text{ for every } f\in\mathcal{L}(\Omega).$$

Take now the following gambles:

	a	b	c	d
f	0	1	0	0
g	1	0.5	0	0

Then $\underline{E}(f|B) = 0$ and $\underline{R}(f|B) = 1$, whence $\underline{P}(f|B) = 0.5$; and $\underline{E}(g|B) = 0.5 = \underline{R}(g|B)$, whence $\underline{P}(g|B) = 0.5$. However, the minimum of f and g is given by

	a	b	c	d
$f \wedge g$	0	0.5	0	0

whence $\underline{E}(f \wedge g|B) = 0$, $\underline{R}(f \wedge g|B) = 0.5$ and therefore $\underline{P}(f \wedge g|B) = 0.25 < \min\{\underline{P}(f|B), \underline{P}(g|B)\}$. ♦

5.1. **Comparison with the updating of possibility measures.** Next, we compare our work with the problem of updating a possibility measure. This problem has received quite some attention in the literature [10, 11, 23, 25, 50]. In particular, in [52] it was studied which of the rules for defining a conditional possibility satisfy the property of coherence with the unconditional model.

The conjugate of a possibility measure is a 2-monotone lower probability, and as such is has only one 2-monotone extension to the space of all gambles: its Choquet integral. However, the coherence of this 2-monotone lower probability with a conditional 2-monotone lower probability is not equivalent to the coherence of the lower previsions they induce by means of the Choquet integral. This is the reason behind the apparent contradiction with the results in [52]: it is shown there that Dempster's rule is a coherent updating rule for updating a possibility measure, even if it can be more informative than the conditional possibility we obtain by regular extension.

To make this clearer, let us study the results in [52] in more detail. The authors consider two finite sets $\mathcal{X}$ and $\mathcal{Y}$. They take a possibility measure $\Pi(\cdot,\cdot)$ on $\mathcal{P}(\mathcal{X}\times\mathcal{Y})$ and look for the smallest and greatest conditional possibility measures $\Pi(\cdot|Y)$ that satisfy coherence with Π. Note that, since we are dealing with upper previsions now, it follows from conjugacy and Proposition 2 that a conditional upper prevision $\overline{P}(\cdot|\mathcal{B})$ is coherent with $\overline{P}$ if and only if $\overline{P}(f|B) \in [\overline{R}(f|B), \overline{E}(f|B)]$ for every gamble f and every $B \subseteq \mathcal{X}\times\mathcal{Y}$ s.t. $\overline{P}(B) > 0$, where $\overline{R}(\cdot|B)$ and $\overline{E}(\cdot|B)$ are the conjugate upper previsions of the regular and natural extensions, respectively.

In [52], the focus is on conditional upper *probabilities* instead of previsions, and in particular on those conditional possibility measures $\Pi(\cdot|Y)$ that satisfy coherence with the unconditional possibility measure Π[7]. They prove in [52, Theorem 4] that the greatest such conditional possibility measure is given by natural extension, whose possibility distribution is

$$\pi_{NE}(x|y) = \begin{cases} \dfrac{\pi(x,y)}{\pi(x,y) + 1 - \max\{\pi(x,y), \Pi(\{y\}^c)\}} & \text{if } \Pi(\{y\}^c) < 1 \\ 1 & \text{if } \Pi(\{y\}^c) = 1, \end{cases}$$

while the smallest such conditional possibility measure is determined by Dempster's rule[8], which produces the possibility measure associated with the following

[7]Here the notation $\Pi(\cdot|Y)$ refers to the conditional upper prevision associated with the partition $\{\mathcal{X}\times\{y\} : y \in \mathcal{Y}\}$ of $\mathcal{X}\times\mathcal{Y}$.

[8]One reason why Dempster's rule is coherent in the case of possibility measures and not for other, more general models, such as plausibility functions, can be found in the characterisation of

possibility distribution:

$$\pi_{DE}(x|y) = \begin{cases} \dfrac{\pi(x,y)}{\pi(y)} & \text{if } \pi(y) > 0. \\ 1 & \text{if } \pi(y) = 0. \end{cases} \tag{10}$$

Then in [52], it is advocated to use the harmonic mean between Dempster's rule and natural extension as an informative updating rule for updating a possibility measure Π. Their reason for this is that the conditional possibility it produces is more informative than the one determined by natural extension, and also that it avoids some of the problems associated with Dempster's rule[9]. The harmonic mean determines the possibility measure defined by the possibility distribution

$$\pi_{HM}(x|y) = \begin{cases} \dfrac{2\pi(x,y)}{\pi(x,y)+\pi(y)+1-\max\{\pi(x,y),\Pi(\{y\}^c)\}} & \text{if } \pi(y) > 0 \\ 1 & \text{if } \pi(y) = 0. \end{cases}$$

However, this rule may be dominated by the regular extension, that produces the conditional possibility measure

$$\pi_{RE}(x|y) = \begin{cases} \dfrac{\pi(x,y)}{\pi(x,y)+1-\max\{\pi(x,y),\Pi(\{y\}^c)\}} & \text{if } \pi(y^c) < 1 \\ 0 & \text{if } \Pi(\{y\}^c) = 1, \pi(y) > \pi(x,y) = 0 \\ 1 & \text{otherwise,} \end{cases} \tag{11}$$

and as a consequence it is not a valid updating rule if we are working with upper previsions instead of upper probabilities. Consider the following example:

Example 6. Consider $\mathcal{X} = \{x_1, x_2\}, \mathcal{Y} = \{y_1, y_2\}$ and let Π be the possibility measure associated with the possibility distribution

$$\pi(x_1, y_1) = 0.3,\ \pi(x_1, y_2) = 1,\ \pi(x_2, y_1) = 0.5,\ \pi(x_2, y_2) = 0.2.$$

Then the conditional possibility distributions determined by regular extension and the harmonic mean rule are given by the following table:

	$\pi(x_1\|y_1)$	$\pi(x_2\|y_1)$	$\pi(x_1\|y_2)$	$\pi(x_2\|y_2)$
Dempster's rule	0.6	1	1	0.2
Natural extension	1	1	1	0.285
Harmonic mean	0.75	1	1	0.235
Regular extension	1	1	1	0.285

This shows that the conditional possibility measure determined by the harmonic mean is dominated by the one produced by regular extension, and as a consequence the conditional upper prevision determined by means of the Choquet integral with respect to $\Pi(\cdot|Y)$ is not coherent with the unconditional upper prevision associated with Π. ♦

Note, however, that the Choquet integral with respect to the necessity measure associated with the possibility distribution in this example, while ∞-monotone, is

coherence provided in [52, Lemma 3], that can be verified for Dempster's rule using that possibility measures are maximum-preserving. See nonetheless the next footnote.

[9]Namely, that Dempster's rule cannot guarantee coherence when we want to determine two conditional possibilities $\Pi(\cdot|\mathcal{X}), \Pi(\cdot|\mathcal{Y})$ at the same time; see [52, Section 6] for details.

not minimum-preserving: this follows from the results in [12], because the possibility distribution in Example 6 is not $\{0,1\}$-valued. In fact, for minimum-preserving lower previsions we can establish the following result:

Proposition 12. *Let $\overline{P}$ be a maximum-preserving coherent lower prevision on $\mathcal{L}(\mathcal{X}\times\mathcal{Y})$, and let π be the possibility distribution associated with its restriction Π to events.*

(1) $\pi_{RE}(x|y) = \pi_{DE}(x|y) = \begin{cases} 0 & \text{if } \pi(x,y)=0<1=\pi(y) \\ 1 & \text{otherwise.} \end{cases}$

(2) Given a conditional possibility measure $\Pi_1(\cdot|Y)$ and the separately coherent conditional upper prevision $\overline{P}_1(\cdot|Y)$ given by its Choquet integral, it holds that

$$\begin{aligned} \overline{P}, \overline{P}_1(\cdot|Y) \text{ coherent } &\Leftrightarrow \pi_1(x|y) \in [\pi_{RE}(x|y), \pi_{NE}(x|y)] \ \forall y \text{ s.t. } \pi(y)>0, \forall x \\ &\Leftrightarrow \Pi, \Pi_1(\cdot|Y) \text{ coherent.} \end{aligned}$$

Proof. (1) This follows from Equations (10) and (11).

(2) The first equivalence follows from Proposition 2, taking into account that Eq. (2) is trivial when $\pi(y)=0$, and also that, given y s.t. $\pi(y)>0$, it holds that

$$\begin{aligned} &\pi_1(x|y) \in [\pi_{RE}(x|y), \pi_{NE}(x|y)] \ \forall x \\ &\qquad \Leftrightarrow (C)\int f d\pi_1(\cdot|y) \in \left[(C)\int f d\pi_{RE}(\cdot|y), (C)\int f d\pi_{NE}(\cdot|y)\right] \ \forall f. \end{aligned}$$

The second equivalence follows from [52, Theorem 4] and the first statement. □

We see then that, even though the problem of coherently updating 2-monotone lower previsions cannot be simplified to that of coherently updating 2-monotone lower probabilities, both problems are equivalent in the particular case when the lower prevision is minimum-preserving.

6. Conclusions

In this work we have considered the problem of conditioning a coherent lower prevision while preserving the property of coherence. This problem has a simple solution when the conditioning event has a positive lower probability, as shown by Walley in [48]: it suffices to apply Generalised Bayes Rule. However, when the conditioning event has zero lower probability and positive upper probability, there may be an infinite number of coherent updated models. These were characterised in [37] for the particular case of finite referential spaces, and it was proven that a conditional lower prevision is coherent with the unconditional model if and only it lies between the conditional lower previsions determined by the procedures of natural and regular extension.

When there is an infinite number of coherent conditional models, it becomes necessary to determine a rule to elicit the appropriate one for the problem at hand. From our discussion in this paper, the following differences can be established between the natural and the regular extension:

- The natural extension is the most conservative model that satisfies the property of coherence with the unconditional lower prevision, while the regular extension is the most informative.
- From the behavioural point of view, regular extension should be used if we assume that a gamble with zero lower prevision (i.e., that we would accept if they offer it to us in addition with a positive but arbitrarily small amount of utility) and positive upper prevision (meaning that there is a positive utility that we would not accept in exchange for selling the gamble) should be desirable to us.
- If we interpret coherent lower previsions as models for the imprecise knowledge of a precise probability, natural extension amounts to working with full conditional measures when some of the possible precise models give zero probability to the conditioning event, while regular extension entails avoiding full conditional measures whenever possible.

Here, we have studied in which cases the matter can be avoided, because the procedures of natural and regular extension give rise to the same updated model. We have considered the particular case when our unconditional model satisfies the property of 2-monotonicity, which guarantees that the lower prevision is the Choquet integral of the coherent lower probability that is its restriction to events, and we have obtained necessary and sufficient conditions for the equality between the natural and regular extensions.

As particular cases, we have also considered the updating problem for completely monotone lower previsions, random sets and minimum-preserving previsions. We have obtained necessary and sufficient conditions in terms of the focal elements of the associated belief function, and we have shown that the procedure of regular extension is compatible with the epistemic interpretation of random sets. As a summary of our results, the natural and regular extensions $\underline{E}(\cdot|B), \underline{R}(\cdot|B)$ coincide if the conditioning event B has zero upper probability of positive lower probability. When $\underline{P}(B) = 0 < \overline{P}(B)$, the natural and the regular extensions coincide, and there is only one way of coherently updating $\underline{P}$, under the following conditions:

Properties of the lower prevision $\underline{P}$	Equivalent conditions for $\underline{E}(\cdot|B) = \underline{R}(\cdot|B)$
2-monotone	$\overline{P}(\omega) > 0 \ \forall \omega \in B$
∞-monotone	$B \subseteq F$, where F is the union of the focal elements
minimum-preserving	$B \subseteq F$, where F is the only focal element

Another interesting property we have explored in this paper is to which extent n-monotonicity is satisfied by the updated model. In this sense, we have to consider three cases:

- When $\overline{P}(B) = 0$, both the natural and the regular extensions coincide, and they are the vacuous lower prevision on B, that is minimum-preserving (note that this holds irrespective of the properties of the unconditional model).
- When $\underline{P}(B) = 0 < \overline{P}(B)$, the natural and the regular extensions may not coincide; the natural extension $\underline{E}(\cdot|B)$ is the vacuous lower prevision

on B, which is minimum-preserving; if $\underline{P}$ is 2-monotone, then the regular extension $\underline{R}(\cdot|B)$ is also minimum-preserving (although it may be a vacuous prevision on a proper subset of B).

- Finally, when $\underline{P}(B) > 0$, then again the natural and regular extensions coincide; they are the only updated model $\underline{P}(\cdot|B)$ satisfying coherence, that is determined by the Generalized Bayes Rule. Its properties are the following:

Properties of the lower prevision $\underline{P}$	$\underline{P}(\cdot\|B)$ on events	$\underline{P}(\cdot\|B)$ on gambles
2-monotone	2-monotone	NOT 2-monotone
∞-monotone	∞-monotone	NOT 2-monotone
minimum-preserving	minimum-preserving	minimum-preserving

Recall on the other hand that the properties of the natural and the regular extension are not shared in general by all the conditional models that are coherent with the unconditional one, as we can see from Example 5.

Finally, let us stress once again that, even if the property of 2-monotonicity means that the lower prevision is uniquely determined by its lower probability, the problem of coherently updating 2-monotone lower probabilities is not equivalent to that of updating 2-monotone lower previsions; this can be seen from the results in Section 5.1.

With respect to the open problems arising from this work, in our opinion the most important one from the technical point of view would be the extension of our results to infinite spaces. Although some work in this direction was already carried out in [37], we expect the problem to be much more difficult; one of the reasons is that the coherence condition between the unconditional and conditional lower previsions must take into account the property of conglomerability. See [48, Chapter 6] and [40] for more details. Another interesting line of research may be the extension of our work to the case where we want to update our model by more than one partition. In that case, we should distinguish between the notions of weak and strong coherence studied by Walley in [48, Chapter 7]. Some results in this sense have been obtained in [37], and in [52] in the context of possibility measures. Finally, another open problem is the study of the coherent updating of other types of non-additive measures.

On the other hand, the interest of our results is restricted to the case where the problem of conditioning is regarded as a problem of querying, or focusing, because this is the interpretation that in our opinion underlies Walley's notion of coherence for conditional and unconditional lower previsions. As discussed by Dubois et al. [7, 21, 27], the situation is different if we consider a problem of belief revision, because in that case it is acceptable that our coherent lower prevision is not coherent with the conditional one, since we are in the possession of new knowledge. See [48, Section 6.1.2] for some comments in this respect. It would be interesting to conduct a similar study to the one carried out in this paper to compare the different possibilities of belief revision in the imprecise case.

Similarly, it would be interesting to study the extension of the results we have considered here to a context of *dynamic coherence*, in the manner discussed by

Skyrms [45] or with the work by Zaffalon and Miranda in [57], which is closer to Walley's spirit.

Acknowledgements

The research in this paper was originated by some discussions with Didier Dubois during the 5th SIPTA School on Imprecise Probabilities. His insight and useful suggestions are much appreciated. We would also like to thank the financial support by project MTM2010-17844 and by the Science and Education Ministry FPU grant AP2009-1034, as well as the reviewers for some helpful comments.

References

[1] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.
[2] Andrea Capotorti, L. Galli, and Barbara Vantaggi. Locally strong coherence and inference with lower-upper probabilities. *Soft Computing*, 7:280–287, 2003.
[3] Andrea Capotorti and Barbara Vantaggi. Locally strong coherence in inference processes. *Annals of Mathematics and Artificial Intelligence*, 35:125–149, 2002.
[4] Gustave Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953–1954.
[5] Giulanella Coletti and Romano Scozzafava. The role of coherence in eliciting and handling imprecise probabilities and its application to medical diagnosis. *Information Sciences*, 130:41–65, 2000.
[6] Giulanella Coletti and Romano Scozzafava. *Probabilistic logic in a coherent setting*. Kluwer, 2002.
[7] Inés Couso and Didier Dubois. Statistical reasoning with set-valued information: Ontic vs. epistemic views. *International Journal of Approximate Reasoning*, 2013. In press.
[8] Milan Daniel. Introduction to an algebra of belief functions on three-element frame of discernment: A Quasi Bayesian case. In Salvatore Greco, Bernadette Bouchon-Meunier, Giulianella Coletti, Mario Fedrizzi, Benedetto Matarazzo, and Ronald R. Yager, editors, *Advances in Computational Intelligence- Proceedings of IPMU'2012*, pages 532–542. Springer-Verlag, 2012.
[9] Luis M. de Campos, María T. Lamata, and Serafín Moral. The concept of conditional fuzzy measure. *International Journal of Intelligent Systems*, 5:237–246, 1990.
[10] Gert de Cooman. Possibility theory II: conditional possibility. *International Journal of General Systems*, 25:325–351, 1997.
[11] Gert de Cooman. Integration and conditioning in numerical possibility theory. *Annals of Mathematics and Artificial Intelligence*, 32:87–123, 2001.
[12] Gert de Cooman and Dirk Aeyels. Supremum preserving upper probabilities. *Information Sciences*, 118:173–212, 1999.
[13] Gert de Cooman and Enrique Miranda. Lower previsions induced by filter maps. *Journal of Mathematical Analysis and Applications*, 410(1):101–116, 2014.
[14] Gert de Cooman, Matthias C. M. Troffaes, and Enrique Miranda. n-Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347:143–156, 2008.
[15] Bruno de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.
[16] Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester, 1974. English translation of [15].
[17] Arthur P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
[18] Dieter Denneberg. *Non-Additive Measure and Integral*. Kluwer Academic, Dordrecht, 1994.
[19] Dieter Denneberg. Conditional expectation for monotone measures, the discrete case. *Journal of Mathematical Economics*, 37:105–121, 2002.
[20] Lester E. Dubins. Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability*, 3:88–99, 1975.
[21] Didier Dubois. Three scenarios for the revision of epistemic states. *Journal of Logic and Computation*, 18(5):721–738, 2008.

[22] Didier Dubois and Thierry Denoeux. Conditioning in Dempster-Shafer theory: Prediction vs. revision. *Advances in Intelligent and Soft Computing*, 164:385–392, 2012.
[23] Didier Dubois, Hung T. Nguyen, and Henri Prade. Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy Sets*, Handbooks of Fuzzy Sets Series, pages 343–438. Kluwer, 2000.
[24] Didier Dubois and Henri Prade. *Possibility Theory*. Plenum Press, New York, 1988.
[25] Didier Dubois and Henri Prade. Bayesian conditioning in possibility theory. *Fuzzy Sets and Systems*, 92:223–240, 1997.
[26] Didier Dubois and Henri Prade. Formal representations of uncertainty. In Denis Bouyssou, Didier Dubois, Marc Pirlot, and Henri Prade, editors, *Decision making processes: concepts and methods*, pages 85–156. ISTE London & Wiley, 2010.
[27] Didier Dubois, Henri Prade, and Philippe Smets. Representing partial ignorance. *IEEE Transactions on Systems, Man and Cybernetics*, 26:361–377, 1996.
[28] Jürgen Eichberger, Simon Grant, and David Kelsey. Updating Choquet beliefs. *Journal of Mathematical Economics*, 43:888–899, 2007.
[29] Ronald Fagin and Joseph Y. Halpern. A new approach to updating beliefs. In P. P. Bonissone, M. Henrion, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence*, volume 6, pages 347–374. North-Holland, Amsterdam, 1991.
[30] Peter Gärdenfors. *Knowledge in Flux – Modeling the Dynamics of Epistemic States*. The MIT Press, Cambridge, MA, 1988.
[31] Jean-Yves Jaffray. Bayesian updating and belief functions. *IEEE Transactions on Systems, Man and Cybernetics*, 22:1144–1152, 1992.
[32] Peter H. Krauss. Representation of conditional probability measures on Boolean algebras. *Acta Mathematica Academiae Scientiarum Hungaricae*, 19:229–241, 1968.
[33] Rudolf Kruse and Klaus D. Meyer. *Statistics with vague data*. D. Reidel Publishing Company, Dordrecht, 1987.
[34] Vladimir P. Kuznetsov. *Interval Statistical Methods*. Radio i Svyaz Publ., 1991. (in Russian).
[35] Isaac Levi. *The enterprise of knowledge*. MIT Press, Cambridge, 1980.
[36] Enrique Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.
[37] Enrique Miranda. Updating coherent lower previsions on finite spaces. *Fuzzy Sets and Systems*, 160(9):1286–1307, 2009.
[38] Enrique Miranda, Inés Couso, and Pedro Gil. Relationships between possibility measures and nested random sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(1):1–15, 2002.
[39] Enrique Miranda, Inés Couso, and Pedro Gil. Upper probabilities and selectors of random sets. In P. Grzegorzewski, O. Hryniewicz, and M. A. Gil, editors, *Soft methods in probability, statistics and data analysis*, pages 126–133. Physica-Verlag, Heidelberg, 2002.
[40] Enrique Miranda, Marco Zaffalon, and Gert de Cooman. Conglomerable natural extension. *International Journal of Approximate Reasoning*, 53(8):1200–1227, 2012.
[41] Hung T. Nguyen. *An introduction to random sets*. Chapman and Hall, 2006.
[42] Cesaltina Pires. A rule for updating ambigous beliefs. *Theory and Decision*, 53(2):137–152, 2002.
[43] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
[44] Lloyd S. Shapley. A value for n-person games. In H. Kuhn and W. Tucker, editors, *Contributions to the theory of Games, II*. Princeton University Press, 1953.
[45] Brian Skyrms. Dynamic coherence and probability kinematics. *Philosophy of Science*, 54:1–20, 1987.
[46] Carl Sundberg and Carl Wagner. Generalized finite differences and Bayesian conditioning of Choquet capacities. *Advances in Applied Mathematics*, 13(3):262–272, 1992.
[47] Peter Walley. Coherent lower (and upper) probabilities. Technical Report 22, University of Warwick, Coventry, 1981.
[48] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
[49] Peter Walley. Inferences from multinomial data: learning about a bag of marbles. *Journal of the Royal Statistical Society, Series B*, 58:3–57, 1996. With discussion.

[50] Peter Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83(1):1–58, May 1996.
[51] Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.
[52] Peter Walley and Gert de Cooman. Coherence of rules for defining conditional possibility. *International Journal of Approximate Reasoning*, 21:63–107, 1999.
[53] Peter M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Reprinted in [55].
[54] Peter M. Williams. On an new theory of epistemic probability. *British Journal for the Philosophy of Science*, 29:375–387, 1978.
[55] Peter M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of [53].
[56] Lofti A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.
[57] Marco Zaffalon and Enrique Miranda. Probability and time. *Artificial Intelligence*, 198:1–51, 2013.

Dep. of Statistics and Operations Research, University of Oviedo.
E-mail address: (mirandaenrique,imontes)@uniovi.es