Natural extension as a limit of regular extensions

Enrique Miranda
University of Oviedo (Spain)
mirandaenrique@uniovi.es

Marco Zaffalon
IDSIA, Lugano (Switzerland)
zaffalon@idsia.ch

Abstract

This paper is devoted to the extension of conditional assessments that satisfy some consistency criteria, such as weak or strong coherence, to further domains. In particular, we characterise the natural extension of a number of conditional lower previsions on finite spaces, by showing that it can be calculated as the limit of a sequence of conditional lower previsions defined by regular extension. Our results are valid for conditional lower previsions with non-linear domains, and allow us to give an equivalent formulation of the notion of coherence in terms of credal sets.

Keywords. Coherent lower previsions, weak and strong coherence, natural extension, regular extension, desirable gambles.

1 Introduction

A distinctive feature of subjective (or personal) probability is its being founded on a notion of self-consistency, which is often called *coherence*. Loosely speaking, coherence requires that the logical implications of any part of the assessments made cannot force a change in the remaining assessments. Since de Finetti [4], coherence is at the heart of precise personal probability, such as the Bayesian theory; later work by Williams [18] and Walley [14] has made of it the central notion also for imprecisely specified probabilities. Nowadays coherence is largely used in imprecise probability to guide research in *coherent lower previsions*.

A coherent lower prevision formalises a subject's beliefs about *gambles*, which represent uncertain rewards. In this it implements a 'direct' approach to belief assessment. The more traditional approach made of probability measures, can be regarded as dual to the former: in fact, a coherent lower prevision is a model equivalent to a closed convex set of probability measures, also called *credal set* after Levi [8].

Despite this equivalence, coherence is used almost exclusively together with coherent lower previsions instead of with sets of probability measures. The reason is that coherence has been, somewhat naturally, formulated only in terms of gambles and lower previsions. This is unfortunate as it prevents coherent modelling to be easily carried over to traditional probability, which is the framework much more commonly used and understood.

With this paper we make a step in the direction of expressing coherence in a dual form. We focus in particular on Walley's notion of *strong* (or *joint*) coherence [14, Section 7.1.4]. We work with variables $X_1, \dots, X_n$ that are assumed to take finitely many values, and furthermore assume to be given m coherent lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ that express beliefs about them.

What we show, loosely speaking, is that $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are jointly coherent if and only if there is a sequence of unconditional lower previsions $\underline{P}_\epsilon(X_1, \dots, X_n)$, $\epsilon \in \mathbb{R}^+$, such that $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are the limit, when ϵ goes to zero, of conditional assessments derived from $\underline{P}_\epsilon(X_1, \dots, X_n)$. This means that by applying Bayes' rule whenever possible to the mass functions in the set equivalent to $\underline{P}_\epsilon(X_1, \dots, X_n)$, we recover the original conditional lower previsions in the limit.

This result relates coherence to the existence of a sequence of joint unconditional credal sets for $X_1, \dots, X_n$. This is interesting because traditionally in precise probability self-consistency is often intended as the existence of a global model: a joint mass function for $X_1, \dots, X_n$. In a sense our results confirm that having a global model is essential for coherence, but also that we need more than that. This is related to the existence of events which are assigned lower probability zero through the original assessments: in fact, a single global model cannot detect in general the inconsistencies that may arise on top of zero probabilities (see [12, Theorem 1], [10]); the sequence, instead, can.

But the sequence does more than that: any least-committal coherent inference that logically follows from $\underline{P}_1(X_{O_1}|X_{I_1}),\ldots,\underline{P}_m(X_{O_m}|X_{I_m})$ alone, can be equivalently done again applying Bayes' rule to the elements of the sequence: in other words, the so-called *natural extension* of the original assessments to a new lower prevision $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ is nothing else but the application of Bayes' rule to $\underline{P}_\epsilon(X_1,\ldots,X_n)$ with $\epsilon \to 0$. This appears to give coherent inference a very accessible interpretation from the dual perspective of traditional probability.

We should mention that ours is not the first work in this direction. A very interesting paper by Walley, Pelessoni and Vicig [17] has introduced the same ideas we consider in Section 5 while restricting the attention to events (rather than gambles) and therefore to finitely many probabilistic assessments. Our work builds upon those ideas, while generalising them so that the only actual restriction now is the finiteness of the spaces.

In particular, we are not limited to what Walley calls *finitely generated* sets of gambles [14, Section 4.2]: we consider also credal sets that cannot be summarised by any finite set of mass functions, or equivalently, that have infinitely many *extreme points* (remember that credal sets are convex). This infinitary dimension has required us to use technical tools other than those in [17], and this has made the technical development somewhat more involved.

We begin by recalling some introductory notions about coherent lower previsions in Section 2. In Section 3 we give new characterisations of avoiding uniform and partial loss, while in Section 4 we deal with weak coherence. In this case, we focus on extending weakly coherent lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}),\ldots,\underline{P}_m(X_{O_m}|X_{I_m})$ to new ones, and an interesting result here is that this extension can be made through conditioning the smallest unconditional prevision $\underline{P}(X_1,\ldots,X_n)$ that is weakly coherent with them. In doing so, we give a number of side results that generalise previous work to domains made of arbitrary sets of gambles. In Section 5 we are finally able to address the main problems described above. Moreover, we relate the need of the sequence $\underline{P}_\epsilon(X_1,\ldots,X_n)$, $\epsilon \in \mathbb{R}^+$, to the existence of events of lower probability equal to zero. This shows also that the natural extension of a number of strongly coherent lower previsions cannot be done, as in the case of weak coherence, through the smallest unconditional joint lower prevision that is coherent with them.

2 Coherence notions on finite spaces

2.1 The behavioural interpretation

Let us give a short introduction to the concepts and results from the behavioural theory of imprecise probabilities that we shall use in the rest of the paper. We refer to [14] for an in-depth study of these and other properties, and to [9] for a brief survey.

Given a possibility space Ω, a *gamble* is a bounded real-valued function on Ω. This function represents a random reward $f(\omega)$, which depends on the a priori unknown value ω of Ω. We shall denote by $\mathcal{L}(\Omega)$ the set of all gambles on Ω. A *lower prevision* $\underline{P}$ is a real functional defined on some set of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$. It is used to represent a subject's supremum acceptable buying prices for these gambles, in the sense that for all $\epsilon > 0$ and all f in $\mathcal{K}$ the subject is disposed to accept the uncertain reward $f - \underline{P}(f) + \epsilon$.

From any lower prevision $\underline{P}$ we can define an upper prevision $\overline{P}$ using conjugacy: $\overline{P}(f) = -\underline{P}(-f)$ for any gamble f. $\overline{P}(f)$ can be interpreted as the infimum acceptable selling price for the gamble f. Because of this relationship, it will suffice for the purposes of this paper to concentrate on lower previsions.

Consider variables $X_1,\ldots,X_n$, taking values in respective *finite* sets $\mathcal{X}_1,\ldots,\mathcal{X}_n$. For any non-empty subset $J \subseteq \{1,\ldots,n\}$ we shall denote by X_J the (new) variable $X_J := (X_j)_{j\in J}$, which takes values in the product space $\mathcal{X}_J := \times_{j\in J}\mathcal{X}_j$. This means that X_J is made of variables that are *logically independent*. We shall also use the notation $\mathcal{X}^n$ for $\mathcal{X}_{\{1,\ldots,n\}}$. In the current formulation made by variables, $\mathcal{X}^n$ is just the definition of the possibility space Ω.

Definition 1. Let J be a subset of $\{1,\ldots,n\}$, and let $\pi_J : \mathcal{X}^n \to \mathcal{X}_J$ be the so-called *projection operator*, i.e., the operator that drops the elements of a vector in $\mathcal{X}^n$ that do not correspond to indexes in J. A gamble f on $\mathcal{X}^n$ is called $\mathcal{X}_J$*-measurable* when for all $x,y \in \mathcal{X}^n$, $\pi_J(x) = \pi_J(y)$ implies that $f(x) = f(y)$.

There is a one-to-one correspondence between the gambles on $\mathcal{X}^n$ that are $\mathcal{X}_J$-measurable and the gambles on $\mathcal{X}_J$. We shall denote by $\mathcal{K}_J$ the set of $\mathcal{X}_J$-measurable gambles.

Consider two disjoint[1] subsets O, I of $\{1,\ldots,n\}$, with $O \neq \emptyset$. $\underline{P}(X_O|X_I)$ represents a subject's behavioural dispositions about the gambles that depend on the outcome of the variables $\{X_j, j \in O\}$, after coming to know the outcome of the variables $\{X_j, j \in I\}$. As such, it is defined at most on gambles that depend on the values of the variables in $O \cup I$ only, i.e., on

[1]That they are taken disjoint is not restrictive. This can be shown using *separate coherence*, given in Definition 2.

the set $\mathcal{K}_{O\cup I}$ of the $\mathcal{X}_{O\cup I}$-measurable gambles on $\mathcal{X}^n$. Given such a gamble f and $z \in \mathcal{X}_I$, $\underline{P}(f|X_I = z)$ represents a subject's supremum acceptable buying price for the gamble f, provided he later comes to know that the variable X_I took the value z (and nothing else). When there is no possible confusion about the variables involved in the lower prevision, we shall use the notation $\underline{P}(f|z)$ for $\underline{P}(f|X_I = z)$. We can define the gamble $\underline{P}(f|X_I)$, which takes the value $\underline{P}(f|z)$ on the elements of $\pi_I^{-1}(z)$ for every $z \in \mathcal{X}_I$. This is a *conditional lower prevision*.

We shall also use the notations

$$G(f|z) := \pi_I^{-1}(z)(f - \underline{P}(f|z))$$
$$G(f|X_I) := \sum_{z\in\mathcal{X}_I} G(f|z) = f - \underline{P}(f|X_I)$$

for all $f \in \mathcal{K}_{O\cup I}$ and all $z \in \mathcal{X}_I$. In the case of an unconditional lower prevision $\underline{P}$, we shall denote $G(f) := f - \underline{P}(f)$ for any gamble f in its domain. Here, and in the rest of the paper, we shall use A to denote both a set A and its indicator function.

The gambles $G(f|z)$ and $G(f|X_I)$ are *almost-desirable*, in the sense that for every $\epsilon > 0$, the gambles $G(f|z) + \epsilon\pi_I^{-1}(z)$ and $G(f|X_I) + \epsilon$ should be desirable for our subject.

2.2 Consistency notions

These assessments can be made for any disjoint subsets O, I of $\{1, \dots, n\}$, and therefore it is not uncommon to model a subject's beliefs using a finite number of different conditional previsions. We should verify then that all the assessments modelled by these conditional previsions are coherent with each other. The first requirement we make is that for any disjoint $O, I \subseteq \{1, \dots, n\}$, the conditional lower prevision $\underline{P}(X_O|X_I)$ defined on a subset $\mathcal{H}_{O\cup I}$ of $\mathcal{K}_{O\cup I}$ should be separately coherent.

Definition 2. A conditional lower prevision $\underline{P}(X_O|X_I)$ with domain $\mathcal{H}_{O\cup I}$ is *separately coherent* if for every $z \in \mathcal{X}_I$, the gamble $\pi_I^{-1}(z)$ belongs to $\mathcal{H}_{O\cup I}$ and $\underline{P}(\pi_I^{-1}(z)|z) = 1$, and moreover

$$\max_{x\in\pi_I^{-1}(z)} \left[\sum_{j=1}^{n} \lambda_j G(f_j|z) - G(f_0|z)\right](x) \geq 0$$

for every $n \in \mathbb{N}, f_j \in \mathcal{H}_{O\cup I}, \lambda_j \geq 0, j = 1, \dots, n, f_0 \in \mathcal{H}_{O\cup I}$.

It is also useful for this paper to consider the particular case where $I = \emptyset$, that is, when we have unconditional information about the variables X_O. We have then an *(unconditional) lower prevision* $\underline{P}(X_O)$ on a subset $\mathcal{H}_O$ of the set $\mathcal{K}_O$ of $\mathcal{X}_O$-measurable gambles. Separate coherence is called then simply *coherence*, and it holds if and only if

$$\max_{x\in\mathcal{X}^n} \left[\sum_{j=1}^{n} \lambda_j G(f_j) - G(f_0)\right](x) \geq 0 \qquad (1)$$

for every $n \in \mathbb{N}, f_0, f_1, \dots, f_n \in \mathcal{H}_O, \lambda_1, \dots, \lambda_n \geq 0$.

Consider now separately coherent conditional lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ with respective domains $\mathcal{H}^1, \dots, \mathcal{H}^m \subseteq \mathcal{L}(\mathcal{X}^n)$, where $\mathcal{H}^j$ is a subset of the set $\mathcal{K}^j$ of $\mathcal{X}_{O_j\cup I_j}$-measurable gambles,[2] for $j = 1, \dots, m$. There are different ways in which we can guarantee their consistency.

Definition 3. $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ *avoid uniform sure loss* if for every $f_j^k \in \mathcal{H}^j$ and every $\lambda_j^k \geq 0, j = 1, \dots, m, k = 1, \dots, n_j$,

$$\max_{x\in\mathcal{X}^n} \left[\sum_{j=1}^{m}\sum_{k=1}^{n_j} \lambda_j^k G_j(f_j^k|X_{I_j})\right](x) \geq 0.$$

A slightly stronger notion is called avoiding partial loss. For this, we define the *$\mathcal{X}_I$-support* $S(f)$ of a gamble f in $\mathcal{K}_{O\cup I}$ as

$$S(f) := \{\pi_I^{-1}(z) : z \in \mathcal{X}_I, f\pi_I^{-1}(z) \neq 0\};$$

i.e., it is the set of conditioning events for which the restriction of f is not identically zero.

Definition 4. $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ *avoid partial loss* if for every $f_j^k \in \mathcal{H}^j$ and every $\lambda_j^k \geq 0$, $j = 1, \dots, m, k = 1, \dots, n_j$ such that not all the $\lambda_j^k f_j^k$ are zero gambles,

$$\max_{x\in\bigcup\cup_{j=1}^{m}\cup_{k=1}^{n_j} S_j(\lambda_j^k f_j^k)} \left[\sum_{j=1}^{m}\sum_{k=1}^{n_j} \lambda_j^k G_j(f_j^k|X_{I_j})\right](x) \geq 0,$$

where by $\bigcup\cup_{j=1}^{m}\cup_{k=1}^{n_j} S_j(\lambda_j^k f_j^k)$ we mean the set of elements that belong to some set in $S_j(\lambda_j^k f_j^k)$ for some $j \in \{1, \dots, m\}, k \in \{1, \dots, n_j\}$.

The idea behind this notion is that a combination of transactions that are acceptable for our subject should not make him lose utiles. It is based on the rationality requirement that a gamble $f \leq 0$ such that $f < 0$ on some set A should not be desirable.

We next give two notions that generalise the concept of coherence in Eq. (1) to the conditional case:

Definition 5. $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are *weakly coherent* if for every $f_j^k \in \mathcal{H}^j, \lambda_j^k \geq$

[2] We use $\mathcal{K}^j$ instead of $\mathcal{K}_{O_j\cup I_j}$ in order to alleviate the notation when no confusion is possible about the variables involved.

$0, j = 1, \dots, m, k = 1, \dots, n_j$, and for every $j_0 \in \{1, \dots, m\}, f_0 \in \mathcal{H}^{j_0}, z_{j_0} \in \mathcal{X}_{I_{j_0}}$,

$$\max_{x \in \mathcal{X}^n} \left[\sum_{j=1}^{m} \sum_{k=1}^{n_j} \lambda_j^k G_j(f_j^k | X_{I_j}) - G_{j_0}(f_0 | z_{j_0}) \right](x) \geq 0.$$

With this condition we require that our subject should not be able to raise his supremum acceptable buying price $\underline{P}_{j_0}(f_{j_0}|z_{j_0})$ for a gamble f_{j_0} contingent on z_{j_0} by taking into account other conditional assessments. However, a number of weakly coherent conditional lower previsions can still present some forms of inconsistency with each other. See [14, Chapter 7], [10] and [17] for some discussion and [14, Sect. 7.3.5] and [10, Examples 4 and 7] for examples of weakly coherent conditionals. On the other hand, weak coherence neither implies nor is implied by the notion of avoiding partial loss. Because of these two facts, we consider another notion which is stronger than both, and which is called (*joint* or *strong*) *coherence*:[3]

Definition 6. $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are *coherent* when for every $f_j^k \in \mathcal{H}^j, \lambda_j^k \geq 0, j = 1, \dots, m, k = 1, \dots, n_j$, and for every $j_0 \in \{1, \dots, m\}, f_{j_0} \in \mathcal{H}^{j_0}, z_{j_0} \in \mathcal{X}_{I_{j_0}}$,

$$\left[\sum_{j=1}^{m} \sum_{k=1}^{n_j} \lambda_j^k G_j(f_j^k | X_{I_j}) - G_{j_0}(f_{j_0} | z_{j_0}) \right](x) \geq 0$$

for some $x \in \bigcup \pi_{I_{j_0}}^{-1}(z_{j_0}) \cup \bigcup_{j=1}^{n} \bigcup_{k=1}^{n_j} S_j(\lambda_j^k f_j^k)$.

Because we are dealing with finite spaces, this notion coincides in the case of linear domains with the one given by Williams in [18]. The coherence of a collection of conditional lower previsions implies their weak coherence; although the converse does not hold in general, it does in the particular case when we only have a conditional and an unconditional lower prevision $\underline{P}(X_O|X_I), \underline{P}$ with domains $\mathcal{H}_{O\cup I}, \mathcal{H}$. If in particular $\mathcal{H}_{O\cup I} = \mathcal{K}_{O\cup I}$ and $\mathcal{H} = \mathcal{L}(\mathcal{X}^n)$, coherence holds if and only if, for all $\mathcal{X}_{O\cup I}$-measurable f and all $z \in \mathcal{X}_I$,

$$\underline{P}(G(f|z)) = 0. \qquad \text{(GBR)}$$

This is called the Generalised Bayes Rule (GBR). When $\underline{P}(z) > 0$, GBR can be used to determine the value $\underline{P}(f|z)$: it is then the *unique* value for which $\underline{P}(G(f|z)) = \underline{P}(\pi_I^{-1}(z)(f - \underline{P}(f|z))) = 0$ holds.

2.3 Linear previsions and envelope theorems

We say that a conditional lower prevision $\underline{P}(X_O|X_I)$ on the set $\mathcal{K}_{O\cup I}$[4] is *linear* if and only if it is separately coherent and moreover $\underline{P}(f+g|z) = \underline{P}(f|z) + \underline{P}(g|z)$ for all $z \in \mathcal{X}_I$ and $f, g \in \mathcal{K}_{O\cup I}$. Conditional linear previsions correspond to the case where a subject's supremum acceptable buying price (lower prevision) coincides with his infimum acceptable selling price (or upper prevision) for any gamble on the domain. When a separately coherent conditional lower prevision $\underline{P}(X_O|X_I)$ is linear we shall denote it by $P(X_O|X_I)$; in the unconditional case, we shall denote it by P and assume that its domain is the set $\mathcal{L}(\mathcal{X}^n)$ of all gambles. The definition of linear prevision implies that in the unconditional case it is just a coherent prevision in de Finetti's sense. In the conditional case, this still holds but it is required that in addition $P(\pi_I^{-1}(z)|z) = 1$ for all $z \in \mathcal{X}_I$. In other words, conditional linear previsions correspond to conditional expectations with respect to a probability. In particular, an unconditional linear prevision P is the expectation with respect to the probability which is the restriction of P to events.

A number of conditional linear previsions are coherent if and only if they avoid partial loss. They are weakly coherent if and only if they avoid uniform sure loss.

Given an unconditional lower prevision $\underline{P}$ with domain $\mathcal{H}$, we shall denote the set of *dominating* linear previsions by $\mathcal{M}(\underline{P}) := \{P : P(f) \geq \underline{P}(f)\ \forall f \in \mathcal{H}\}$. Similarly, for a conditional lower prevision $\underline{P}(X_O|X_I)$ with domain $\mathcal{H}_{O\cup I}$, we define $\mathcal{M}(\underline{P}(X_O|X_I))$ as the set of linear previsions $P(X_O|X_I)$ such that

$$P(f|z) \geq \underline{P}(f|z)\ \forall f \in \mathcal{H}_{O\cup I}, z \in \mathcal{X}_I.$$

Then $\underline{P}$ is coherent if and only if it is the lower envelope of $\mathcal{M}(\underline{P})$, and $\underline{P}(X_O|X_I)$ is separately coherent if and only if it is the lower envelope of $\mathcal{M}(\underline{P}(X_O|X_I))$.

The situation is more complicated when we have more than one conditional lower prevision, as the previous results essentially hold for finite spaces. In [14] Walley proved that when the referential spaces are finite and the domains are linear spaces, coherent $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are always the envelope of a set $\{P_1^\lambda(X_{O_1}|X_{I_1}), \dots, P_m^\lambda(X_{O_m}|X_{I_m}) : \lambda \in \Lambda\}$ of dominating coherent conditional linear previsions. In [10], a similar property was established for weak coherence. In Section 4 we shall generalise this second property to arbitrary domains.

2.4 Extensions to further domains

Let $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ be separately coherent conditional lower previsions with domains $\mathcal{H}^i \subseteq \mathcal{K}^i$ for $i = 1, \dots, m$ and avoiding partial loss.

Their *natural extensions* to the sets $\mathcal{K}^1, \dots, \mathcal{K}^m$ are

[3] The distinction with the unconditional notion of coherence mentioned above will always be clear from the context.

[4] We shall always assume in this paper that the domain of a conditional linear prevision $P(X_O|X_I)$ is the whole set $\mathcal{K}_{O\cup I}$ of $\mathcal{X}_{O\cup I}$-measurable gambles.

defined, [5] for every $f \in \mathcal{K}^j$ and every $z_j \in \mathcal{X}_{I_j}$, by

$$\underline{E}_j(f|z_j) = \sup\{\alpha : \exists f_j^k \in \mathcal{H}^j, \lambda_j^k \geq 0, \text{ s.t.} \\ [\sum_{j=1}^{m}\sum_{k=1}^{n_j} \lambda_j^k G_j(f_j^k|X_{I_j}) - \pi_{I_j}^{-1}(z_j)(f-\alpha)] < 0 \\ \text{on } \bigcup \cup_{j=1}^m \cup_{k=1}^{n_j} S_j(\lambda_j^k f_j^k) \cup \pi_{I_j}^{-1}(z_j)\}. \quad (2)$$

In the context of this paper, where all the conditioning spaces are finite, the natural extensions are the smallest conditional lower previsions which are coherent and dominate $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$. Moreover, they coincide with the initial assessments if and only if $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ are themselves coherent. Otherwise, they 'correct' the initial assessments taking into account the implications of the notions of coherence [11, Prop. 11]. In the rest of the paper we shall consider at some point also the natural extension $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$, for arbitrary disjoint subsets O_{m+1}, I_{m+1} of $\{1, \ldots, n\}$. Doing so amounts to implicitly include in the original set of lower previsions, an additional one $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ defined on a trivial domain (such as the constant gambles), and then to take the natural extension.

In this paper, we shall also define conditional lower previsions coherently by using the *regular extension*. Given a credal set $\mathcal{M}$ and disjoint O, I, the regular extension $\underline{R}(X_O|X_I)$ is given by

$$\underline{R}(f|z) := \inf\left\{\frac{P(f\pi_I^{-1}(z))}{P(z)} : P \in \mathcal{M}, P(z) > 0\right\}$$

for every $z \in \mathcal{X}_I, f \in \mathcal{K}_{O\cup I}$. This amounts to applying Bayes' rule to the linear previsions in $\mathcal{M}$ whenever possible. The regular extension has been proposed and used a number of times in the literature as an updating rule [2, 3, 5, 6, 14, 15]. See [10] for a comparison with natural extension in the finite case.

3 Characterising avoiding uniform sure loss and avoiding partial loss

Let $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ be separately coherent conditional lower previsions with respective domains $\mathcal{H}^1, \ldots, \mathcal{H}^m$, where $\mathcal{H}^j$ is a (not necessarily linear) subset of the class $\mathcal{K}^j$ of $\mathcal{X}_{O_j\cup I_j}$-measurable gambles.

Our first result is an extension of [12, Prop. 5] to arbitrary domains. It uses the following lemma:

[5] We do not extend $\underline{P}_j(X_{O_j}|X_{I_j})$ beyond the set $\mathcal{K}^j$ of $\mathcal{X}_{O_j\cup I_j}$-measurable gambles as that would not be compatible with the interpretation we have given of $\underline{P}_j(X_{O_j}|X_{I_j})$; yet, it is possible to extend it to $\mathcal{L}(\mathcal{X}^n)$ by considering $\underline{P}(X_{I^c}|X_I)$ instead of $\underline{P}(X_O|X_I)$, and with the same initial domain.

Lemma 1. *Let $\underline{P}, \underline{P}(X_O|X_I)$ be coherent lower previsions with respective domains $\mathcal{L}(\mathcal{X}^n), \mathcal{H}_{O\cup I}$. For every $P \in \mathcal{M}(\underline{P})$ there is some conditional linear prevision $P(X_O|X_I)$ in $\mathcal{M}(\underline{P}(X_O|X_I))$ such that $P, P(X_O|X_I)$ are coherent. Moreover,*

$$\underline{P}(G(f|z)) = 0, \ \underline{P}(G(f|X_I)) \geq 0$$

for every gamble $f \in \mathcal{H}_{O\cup I}$ and every $z \in \mathcal{X}_I$.

Proposition 1. $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ *avoid uniform sure loss if and only if there are dominating weakly coherent conditional linear previsions with domains $\mathcal{K}^1, \ldots, \mathcal{K}^m$.*

This result will be interesting in Section 4 when we study the smallest dominating weakly coherent lower previsions. It follows that avoiding uniform sure loss is a necessary and sufficient condition for the existence of such lower previsions. Since moreover we shall prove in Theorem 1 that when all the referential spaces are finite weak coherence is preserved by taking lower envelopes, we deduce that a way of computing the smallest dominating weakly coherent lower previsions is to take the lower envelopes of the (non-empty) sets of weakly coherent dominating conditional linear previsions.

On the other hand, it follows from [14, Sec. 8.1] that when all the referential spaces are finite and the domains are linear spaces, the notion of avoiding partial loss is equivalent to the existence of dominating coherent linear conditional previsions. We here generalise the result to non-linear domains.

Lemma 2. *Assume that the conditional lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ avoid partial loss, and let $\underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ be their natural extensions to $\mathcal{K}^1, \ldots, \mathcal{K}^m$. Then $\underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ are coherent.*

Proposition 2. $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ *avoid partial loss if and only if there are dominating coherent conditional linear previsions with domains $\mathcal{K}^1, \ldots, \mathcal{K}^m$.*

The notions of avoiding partial and uniform sure loss constitute a generalisation, to conditional assessments, of a consistency notion for unconditional lower previsions, called *avoiding sure loss*. It is established in [14, Thm. 3.3.3] that avoiding sure loss is equivalent to the existence of a dominating coherent linear prevision, and therefore can be seen as a minimal consistency requirement.

When we move towards conditional lower previsions, we have seen in Section 2 that there are two ways of extending the notion of coherence of lower previsions, called weak and (strong) coherence. What we have proved by means of Propositions 1 and 2 is that avoiding uniform sure and partial loss are the respective counterparts of avoiding sure loss for each of these two extensions.

We conclude the section with another characterisation of avoiding partial loss, where we can find some of the ideas we shall use in our approximation of the natural extension in Section 5.

Proposition 3. $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ *avoid partial loss if and only if for all* $\epsilon > 0$, $f_j^k \in \mathcal{H}^j$, $\lambda_j^k \geq 0$, $j = 1, \ldots, m$, $k = 1, \ldots, n_j$ *such that not all products* $\lambda_j^k f_j^k$ *are zero gambles, it holds that*

$$\max_{x \in \mathcal{X}^n} \left[\sum_{j=1}^{m} \sum_{k=1}^{n_j} \lambda_j^k (G_j(f_j^k|X_{I_j}) + \epsilon S_j(f_j^k)) \right](x) > 0.$$

Hence, by introducing these ϵ-terms, we can replace the maximum on the union of the supports with a maximum on $\mathcal{X}^n$. We shall relate this later to the weak coherence of some approximations of our conditional lower previsions.

4 Extensions of weakly coherent conditionals

We focus next on the notion of weak coherence of a number of conditional lower previsions. We begin by giving a characterisation of weak coherence and determining the smallest (unconditional) coherent lower prevision which is weakly coherent with a number of conditionals. This extends [10, Thms. 2 and 3] to arbitrary domains:

Theorem 1. *Let* $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ *be separately coherent conditional lower previsions with domains* $\mathcal{H}^1, \ldots, \mathcal{H}^m$. *The following are equivalent:*

(WC1) They are weakly coherent.

(WC2) They are the lower envelopes of a class of weakly coherent conditional linear previsions, $\{P_1^\lambda(X_{O_1}|X_{I_1}), \ldots, P_m^\lambda(X_{O_m}|X_{I_m}) : \lambda \in \Lambda\}$.

(WC3) There is a coherent lower prevision $\underline{P}$ *on* $\mathcal{L}(\mathcal{X}^n)$ *which is weakly coherent with them.*

(WC4) There is a coherent lower prevision $\underline{P}$ *on* $\mathcal{L}(\mathcal{X}^n)$ *which is pairwise coherent with them.*

Moreover, the smallest coherent lower prevision in (WC3) and (WC4) is given, for any gamble f *on* $\mathcal{X}^n$, *by*

$$\underline{P}(f) = \sup\{\alpha : \exists f_j^k \in \mathcal{H}^j, \lambda_j^k \geq 0, \ s.t.$$
$$\max_{x \in \mathcal{X}^n} [\sum_{j=1}^{m} \sum_{k=1}^{n_j} \lambda_j^k G(f_j^k|X_{I_j}) - (f - \alpha)](x) < 0\}. \quad (3)$$

We summarise the relationships between the different consistency conditions when all the referential spaces are finite in the following figure.

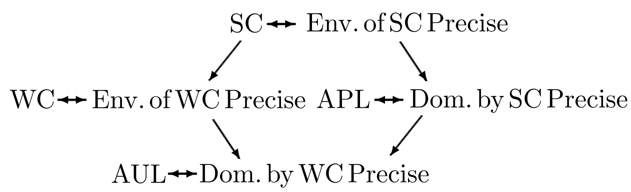

Figure 1: Equivalences and implications between consistency concepts analysed in the paper. Keys: SC = strongly coherent; WC = weakly coherent; AUL = avoiding uniform sure loss; APL = avoiding partial loss; Env. = envelope; Dom. = dominated.

Under some conditions, the functional we just defined is also the natural extension of a number of conditional lower previsions:

Corollary 1. $\underline{P}$ *is the smallest coherent lower prevision which is coherent with* $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ *if and only if these conditional previsions are coherent.*

It is useful at this point to compare the functional $\underline{P}$ defined in Eq. (3) with the unconditional natural extension $\underline{E}$ that we should define using Eq. (2). In order to do this, we should consider $O_{m+1} = \{1, \ldots, n\}, I_{m+1} = \emptyset$ and add $\underline{P}(X_{O_{m+1}})$ to our set of gambles with the trivial domain given by the constant gambles. For this discussion to make sense, we are going to assume also that $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ avoid partial loss and are weakly coherent.

We see from [11, Theorem 12] that in that case the functionals $\underline{P}$ and $\underline{E}$ coincide. Hence, the unconditional natural extension $\underline{E}$ is the smallest unconditional lower prevision which is weakly coherent with $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$; and as we have proven in Corollary 1, it is coherent with them if and only if the initial assessments are coherent. A sufficient condition for the coherence of $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ when their domains are $\mathcal{K}^1, \ldots, \mathcal{K}^m$ is that $\underline{P}(z_j) > 0$ for all $z_j \in \mathcal{X}_{I_j}$ and for all $j = 1, \ldots, m$ [10, Thm. 11]. On the other hand, in [10, Example 2] we can find an example of assessments which avoid partial loss and are weakly coherent, but are not coherent.

Assume now that we have weakly coherent $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$, and that given disjoint O_{m+1}, I_{m+1}, we want to determine the smallest conditional lower prevision $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ which is weakly coherent with the rest. Our next result shows that it suffices to go through the unconditional lower prevision $\underline{P}$ given by Eq. (3):

Theorem 2. *The smallest conditional lower prevision $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ with domain $\mathcal{K}^{m+1}$ which is weakly coherent with $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$ is given, for every $f\in\mathcal{K}^{m+1}, z_{m+1}\in\mathcal{X}_{I_{m+1}}$, by $\underline{P}_{m+1}(f|z_{m+1}) :=$*

$$\begin{cases}\min_{x\in\pi^{-1}_{I_{m+1}}(z_{m+1})} f(x) & \text{if } \underline{P}(z_{m+1})=0\\ \min\{P(f|z_{m+1}) : P\geq\underline{P}\} & \text{otherwise,}\end{cases} \tag{4}$$

where $\underline{P}$ is given by Eq. (3).

This stresses once more the fact that the informative content of a number of weakly coherent lower previsions is preserved by summarising them with an unconditional lower prevision. Note moreover that if $\underline{P}_{m+1}(z_{m+1})>0$ the conditional lower prevision $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ is uniquely determined from $\underline{P}$ by the Generalised Bayes Rule.

Our final result in this section shows that we can use the definition of natural extension to obtain a conditional lower prevision which is weakly coherent with a number of assessments.

Proposition 4. *Consider weakly coherent $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$ with domains $\mathcal{H}^1,\dots,\mathcal{H}^m$, and let $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ be defined on $\mathcal{K}^{m+1}$ by Eq. (2). Then $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ is weakly coherent with $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$.*

In particular, if $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$ are coherent, it follows from the results in [11] that $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ is the smallest conditional lower prevision that is coherent with them. It may be strictly greater than the conditional lower prevision $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ derived in Eq. (4). An instance of such a situation can be found in [16, Example 8]; it can be checked that the smallest weakly coherent conditional lower prevision derived from the assessments in the example is vacuous.

This shows on the one hand that the notion of weak coherence is indeed too weak to fully capture the behavioural implications of our assessments, and on the other that the natural extension cannot be derived in general from the unconditional lower prevision $\underline{P}$. In the following section, we get around this problem by showing: (i) that we can instead derive it using a *sequence* of unconditional lower previsions that converges to $\underline{P}$ and (ii) that in some cases it coincides with the weakly coherent natural extension.

5 Natural extension as a limit of regular extensions

Let $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$ now be separately coherent conditional lower previsions with domains $\mathcal{H}^j\subseteq\mathcal{K}^j$ for $j=1,\dots,m$. We shall assume that they are weakly coherent and avoid partial loss, but they are not necessarily coherent. Our goal in this section is to characterise their natural extension $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ given by Eq. (2).

Although in general we shall assume that the index $m+1$ does not belong to $\{1,\dots,m\}$ (and then we have to include among the original assessments a conditional lower prevision $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ defined on the set of constant gambles), the results are still valid if what we study is the natural extension of one of our assessments $\underline{P}_j(X_{O_j}|X_{I_j})$ to $\mathcal{K}^j$.

We shall prove later (in Theorem 3) that this natural extension can be computed as a limit of regular extensions. In order to do this, we are going to consider a sequence of credal sets which are compatible with conditional lower previsions which converge point-wise to $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$. For every $\epsilon>0$, let $\mathcal{M}(\epsilon)$ be the set of linear previsions satisfying

$$P(f_j\pi^{-1}_{I_j}(z_j))\geq P(z_j)(\underline{P}_j(f_j|z_j)-\epsilon R(f_j)) \tag{5}$$

for every $f_j\in\mathcal{H}^j, z_j\in\mathcal{X}_{I_j}, j=1,\dots,m$, where $R(f_j)=\max f_j-\min f_j$ is the range of the gamble f_j. Let us also consider the set of gambles

$$\mathcal{V}_\epsilon := \{f\geq\sum_{j=1}^m\sum_{k=1}^{n_j}\lambda_j^k(G_j(f_j^k|X_{I_j})+\epsilon R(f_j^k)S_j(f_j^k)) \text{ for some } f_j^k\in\mathcal{H}^j,\lambda_j^k\geq 0\}, \tag{6}$$

where, with a certain abuse of notation, $S_j(f_j^k)$ is used to denote the indicator function of the set of elements which belong to some set in $S_j(f_j^k)$.

For $\epsilon=0$ we obtain the set $\mathcal{M}(0)$ of linear previsions P such that

$$P(f_j\pi^{-1}_{I_j}(z_j))\geq P(z_j)\underline{P}_j(f_j|z_j) \tag{7}$$

for all $f_j\in\mathcal{H}^j, z_j\in\mathcal{X}_{I_j}, j=1,\dots,m$, and the set of gambles

$$\mathcal{V} := \{f\geq\sum_{j=1}^m\sum_{k=1}^{n_j}\lambda_j^kG_j(f_j^k|X_{I_j}) \text{ for some } f_j^k\in\mathcal{H}^j,\lambda_j^k\geq 0\}. \tag{8}$$

It follows from their definition that $\mathcal{V}_\epsilon \subseteq \mathcal{V}$ and $\mathcal{M}(0) \subseteq \mathcal{M}(\epsilon)$ for any $\epsilon > 0$. Since the gamble constant on 0 belongs to $\mathcal{V}_\epsilon$ for all $\epsilon \geq 0$, we deduce that these sets of gambles are non-empty. On the other hand, it follows that $\mathcal{M}(\epsilon)$ are convex sets of linear previsions for all $\epsilon > 0$. $\mathcal{M}(0)$ (and therefore also $\mathcal{M}(\epsilon)$ for all $\epsilon > 0$) is non-empty because the conditional lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are weakly coherent. This follows from the following proposition. Let $\underline{P}_\epsilon$ denote the lower envelope of the credal set $\mathcal{M}(\epsilon)$, and $\underline{P}_0$ the lower envelope of $\mathcal{M}(0)$.

Proposition 5. *Consider weakly coherent $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ with respective domains $\mathcal{H}^1, \dots, \mathcal{H}^m$ and avoiding partial loss.*

1. *For any $\epsilon \geq 0$, $\mathcal{M}(\epsilon) = \{P : P(f) \geq 0\ \forall f \in \mathcal{V}_\epsilon\}$, and $\{f : P(f) \geq 0\ \forall P \in \mathcal{M}(\epsilon)\} = \overline{\mathcal{V}}_\epsilon$, where the closure is taken in the topology of uniform convergence.*
2. *$\mathcal{M}(0) = \cap_{\epsilon>0}\mathcal{M}(\epsilon) = \mathcal{M}(\underline{P})$, where $\underline{P}$ is the coherent lower prevision given by Eq. (3).*
3. *$\underline{P}_0 = \sup_{\epsilon>0} \underline{P}_\epsilon = \underline{P}$.*

In the particular case of precise assessments (i.e., conditional linear previsions) we can go a bit further. In this case, and in analogy with the situation in the unconditional case, we can show that events provide all the information we need. Note also that in the linear case the notion of avoiding partial loss is equivalent to coherence (and implies therefore weak coherence).

Proposition 6. *Consider coherent $P_1(X_{O_1}|X_{I_1}), \dots, P_m(X_{O_m}|X_{I_m})$ with domains $\mathcal{K}^1, \dots, \mathcal{K}^m$. Let $\mathcal{V}_\epsilon$ be the set of gambles given by Eq. (6), and $\mathcal{M}(\epsilon)$ be the credal set given by Eq. (5). Let us denote moreover by $\mathcal{V}^A_\epsilon, \mathcal{M}^A_\epsilon$ the corresponding sets determined by the restrictions to events of $P_1(X_{O_1}|X_{I_1}), \dots, P_m(X_{O_m}|X_{I_m})$.*

1. *For every $\epsilon > 0$, $\overline{\mathcal{V}}_\epsilon \subseteq \mathcal{V}^A_{\epsilon_1}$, where $\epsilon_1 = \frac{\epsilon}{\max_j |\mathcal{X}_{O_j}|}$, and as a consequence $\cup_\epsilon \mathcal{V}_\epsilon = \cup_\epsilon \mathcal{V}^A_\epsilon = \cup_\epsilon \overline{\mathcal{V}}_\epsilon$.*
2. *$\mathcal{M}(\epsilon) \supseteq \mathcal{M}^A_{\epsilon_1}$, whence $\cap_\epsilon \mathcal{M}(\epsilon) = \cap_\epsilon \mathcal{M}^A_\epsilon$.*

This result will be very useful for us because it allows us to connect our results with the ones established in [17] for the particular case of conditional lower previsions defined on events. The case of events is also interesting because the sets of desirable gambles we use are finitely generated, and this makes it easier to apply separation results.

Now that we have clarified a bit the structure of the sets $\mathcal{M}(\epsilon), \mathcal{V}_\epsilon$, we explore how they can be used to characterise the conditional natural extension.

Proposition 7. *Consider $f \in \mathcal{K}^{m+1}$ and $z_{m+1} \in \mathcal{X}_{I_{m+1}}$. Then $\sup\{\mu : \pi^{-1}_{I_{m+1}}(z_{m+1})(f-\mu) \in \cup_\epsilon \mathcal{V}_\epsilon\} = \underline{E}_{m+1}(f|z_{m+1}) \leq \sup\{\mu : \pi^{-1}_{I_{m+1}}(z_{m+1})(f-\mu) \in \mathcal{V}\}$, where $\mathcal{V}$ is given by Eq. (8).*

For every $\epsilon > 0$, let us define $\underline{R}^\epsilon_{m+1}(f|z_{m+1})$ from $\mathcal{M}(\epsilon)$ by regular extension, i.e., let it be given by

$$\inf\{P(f|z_{m+1}) : P \in \mathcal{M}(\epsilon), P(z_{m+1}) > 0\}. \quad (9)$$

The first thing we have to prove is that this definition makes sense.

Proposition 8. *For every $z_{m+1} \in \mathcal{X}_{I_{m+1}}$ and every $\epsilon > 0$, there is some $P \in \mathcal{M}(\epsilon)$ s.t. $P(z_{m+1}) > 0$.*

Since the credal set $\mathcal{M}(\epsilon)$ does not increase as ϵ converges to zero, we deduce that the conditional lower previsions $\underline{R}^\epsilon_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ given by Eq. (9) do not decrease as ϵ goes to zero. We can thus consider

$$\underline{F}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}}) := \lim_{\epsilon \to 0} \underline{R}^\epsilon_{m+1}(X_{O_{m+1}}|X_{I_{m+1}}),$$

the limit of these conditional lower previsions.

In analogy with Proposition 7, we can characterise $\underline{F}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ in terms of desirable gambles:

Lemma 3. *For every $f \in \mathcal{K}^{m+1}, z_{m+1} \in \mathcal{X}_{I_{m+1}}$, $\underline{F}_{m+1}(f|z_{m+1}) = \sup\{\mu : \pi^{-1}_{I_{m+1}}(z_{m+1})(f-\mu) \in \cup_\epsilon \overline{\mathcal{V}}_\epsilon\}$. As a consequence, $\underline{F}(f|z_{m+1}) \geq \underline{E}(f|z_{m+1})$.*

Since the sets $\mathcal{V}_\epsilon$ are not necessarily closed, we may wonder if the functional $\underline{F}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ defined as a limit of regular extensions is actually more precise that the natural extension $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$. In our next result, we show that this is not the case. The proof is based on using Proposition 6 to obtain the result for linear previsions, and then apply envelope results. It is a generalisation of a result established in [17] for conditional lower probabilities:

Theorem 3. *Assume that $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are weakly coherent and avoid partial loss. Then $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}}) = \underline{F}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$.*

Of course, the result is valid in particular if $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ are coherent. We can also determine, as a corollary, that the conditional lower prevision derived from an unconditional by natural extension is also the limit of conditional lower previsions obtained by regular extension. Note that in this particular case $\mathcal{M}(\epsilon), \mathcal{M}(0)$ would be

$$\mathcal{M}(\epsilon) = \{P : P(f) \geq \underline{P}(f) - \epsilon R(f)\ \forall f \in \mathcal{H}\}, \quad (10)$$

and $\mathcal{M}(0) = \mathcal{M}(\underline{P})$. Another interesting point is that in this particular case where we have a conditional and an unconditional lower prevision only, weak and strong coherence are equivalent:

Corollary 2. *Let $\underline{P}$ be a coherent lower prevision with domain $\mathcal{H}$, and consider disjoint O, I. For every $\epsilon > 0$, let $\underline{R}^{\epsilon}(X_O|X_I)$ be the conditional lower prevision defined from $\mathcal{M}(\epsilon)$ using regular extension, where $\mathcal{M}(\epsilon)$ is given by Eq. (10). Then $\lim_{\epsilon \to 0} \underline{R}^{\epsilon}(X_O|X_I)$ coincides with the conditional natural extension $\underline{E}(X_O|X_I)$.*

At this point we may still be wondering if going through the sets $\mathcal{M}(\epsilon)$ is really necessary, or if we could have applied regular extension on the credal set $\mathcal{M}(0)$ given by Eq. (7) and use it to approximate $\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$. This is not possible in general, because Proposition 8 does not necessarily hold for $\epsilon = 0$, i.e., there may not be any $P \in \mathcal{M}(0)$ such that $P(z_{m+1}) > 0$, and therefore we may not be able to use the regular extension in that case; this is easy to see with precise assessments. Moreover, even if we can apply regular extension to $\mathcal{M}(0)$, we do not necessarily have the equality $\underline{E}_{m+1}(f|z_{m+1}) = \inf\{P(f|z_{m+1}) : P \in \mathcal{M}(0), P(z_{m+1}) > 0\}$. This is discussed for the particular case of lower probabilities in [17, Sects. 3.7,3.8], and some illustrative examples are provided.

Hence, the inequality given in Proposition 7 is not necessarily an equality. In the following result, we show that a sufficient condition for the equality to hold is that the lower probability of the conditioning event is positive; see also [14, Thm. 8.1.4]:

Proposition 9. *Consider weakly coherent $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ that avoid partial loss. Let $\underline{P}$ be their unconditional natural extension, given by Eq. (3), and let $\underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}})$ be given by Eq. (4). If $\underline{P}(z_{m+1}) > 0$, then for all $f \in \mathcal{K}^{m+1}$, $\underline{E}_{m+1}(f|z_{m+1}) = \underline{P}_{m+1}(f|z_{m+1}) = \sup\{\mu : \pi^{-1}_{I_{m+1}}(z_{m+1})(f - \mu) \in \mathcal{V}\}$.*

Hence, we also show that in this case the natural extension is also the smallest conditional lower prevision that is weakly coherent with the initial assessments. In particular, if $\underline{P}(z_{m+1}) > 0$ for all $z_{m+1} \in \mathcal{X}_{I_{m+1}}$, we should deduce that

$$\underline{E}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}}) = \underline{P}_{m+1}(X_{O_{m+1}}|X_{I_{m+1}}).$$

The intuition here is that in that case $\underline{R}^{\epsilon}(z_{m+1}) > 0$ for all $z_{m+1} \in \mathcal{X}_{I_{m+1}}$ and for ϵ small enough, and then the regular extension from $\mathcal{M}(\epsilon)$ coincides with the natural extension. From here it suffices then to apply a limit result.

Finally, we are going to show that our results allow to derive a characterisation of the notion of coherence for conditional lower previsions on finite spaces.

Lemma 4. *Consider a sequence of conditional lower previsions $\{\underline{P}^k_1(X_{O_1}|X_{I_1}), \dots, \underline{P}^k_m(X_{O_m}|X_{I_m})\}_{k \in \mathbb{N}}$ with respective domains $\mathcal{H}^1, \dots, \mathcal{H}^m$. Assume their point-wise limits $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ exist. If $\underline{P}^k_1(X_{O_1}|X_{I_1}), \dots, \underline{P}^k_m(X_{O_m}|X_{I_m})$ are weakly coherent (resp., coherent) for all k, then so are $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$.*

Using this lemma, we can derive the following:

Theorem 4. *Let $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ be separately coherent conditional lower previsions. They are coherent if and only if they are the point-wise limits of a sequence of coherent conditional lower previsions defined by regular extension.*

Hence, in the case of finite spaces the notion of coherence, which, as we have argued, is the central (and in a way the unique) consistency notion in Walley's theory, is equivalent to the approximation by means of regular extensions.

6 Conclusions

In this paper we have focused on providing a dual view of Walley's strong coherence and natural extension in the case of finite spaces. Our main result shows that there is an equivalent model made of a sequence of unconditional credal sets. By this sequence we can recreate the original conditional lower previsions using Bayes' rule; moreover, we can use this rule to compute any natural extension. This shows, in a sense, that the essence of coherence within finite spaces is just Bayes' rule. But it also suggests that the basic modelling unit in a traditional theory of (coherent) probability, even a precise one, should be a sequence of unconditional credal sets rather than a single unconditional model. This might give a new perspective on probabilistic modelling; and it might make coherence and natural extension accessible and usable concepts without notions of coherent lower previsions.

In developing the main results we have given a number of new results more strictly related to coherent lower previsions. We have given new characterisations of the notions of avoiding partial and uniform sure loss. We have shown that there is an extension of weakly coherent lower previsions that we could call *weak natural extension* and that it can be characterised through conditioning the smallest unconditional lower prevision that is weakly coherent with the former ones. Finally, we have discussed some key differences between the weak natural extension and the natural extension. All of this seems to be interesting in its own as it shows, for example, that what some applications of credal sets do is to make weakly coherent inferences rather than computing natural extensions, and therefore points to possible improvements of those approaches.

With respect to future work, we should like to point out three avenues: one is the obvious possibility to try to extend the results presented here to the case of infinite spaces. We envisage that most of them will not be immediately extendable because in our proofs we have used a number of separation theorems and envelope results that do not apply directly to the infinite case. Another aspect worth investigating is whether the equivalence mentioned initially between conditional lower previsions and the sequence remains valid also when structural judgments are introduced in a model. Finally, the idea of using a certain sequence to check coherence and compute extensions is present also in other works [1, 13] which have a common root in the work of Krauss [7]. The relationship between the sequences used here and those used in the mentioned works should also be investigated.

Acknowledgements

Work partially supported by the projects TIN2008-06796-C04-01, MTM2007-61193 and by the Swiss NSF grants n. 200020-116674/1 and 200020-121785/1.

References

[1] G. Coletti and R. Scozzafava. *Probabilistic logic in a coherent setting*. Kluwer, 2002.

[2] L. M. de Campos, M. T. Lamata, and S. Moral. The concept of conditional fuzzy measures. *International Journal of Intelligent Systems*, 5:237–246, 1990.

[3] G. de Cooman and M. Zaffalon. Updating beliefs with incomplete observations. *Artificial Intelligence*, 159:75–125, 2004.

[4] B. de Finetti. *Theory of Probability*. John Wiley & Sons, Chichester, 1974–1975. English Translation of *Teoria delle Probabilità* (Einaudi, Turin, 1970), two volumes.

[5] R. Fagin and J. 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.

[6] J.-Y. Jaffray. Bayesian updating and belief functions. *IEEE Transactions on Systems, Man and Cybernetics*, 22:1144–1152, 1992.

[7] P. H. Krauss. Representation of conditional probability measures on boolean algebras. *Acta Math. Acad. Sci. Hungar*, 19:229–241, 1968.

[8] I. Levi. Potential surprise: its role in inference and decision making. In L. J. Cohen and M. Hesse, editors, *Applications of Inductive Logic*, pages 1–27. Clarendon Press, Oxford, 1980.

[9] E. Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.

[10] E. Miranda. Updating coherent previsions on finite spaces. *Fuzzy Sets and Systems*, 160(9):1286–1307, 2009.

[11] E. Miranda and G. de Cooman. Coherence and independence in non-linear spaces. Technical report, 2005. Downloadable at `http://bellman.ciencias.uniovi.es/~emiranda/wp05-10.pdf`.

[12] E. Miranda and M. Zaffalon. Coherence graphs. *Artificial Intelligence*, 173(1):104–144, 2009.

[13] R. Pelessoni and P. Vicig. A consistency problem for imprecise conditional probability assessments. In *Proceedings of the Seventh International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 1998)*, pages 1478–1485. EDK, Paris, 1998.

[14] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[15] P. 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.

[16] P. Walley, R. Pelessoni, and P. Vicig. Direct algorithms for checking coherence and making inferences for conditional probability assessments. Technical report, University of Trieste, 1999. Downloadable at `http://www2.units.it/~renatop/`.

[17] P. Walley, R. Pelessoni, and P. Vicig. Direct algorithms for checking consistecy and making inferences for conditional probability assessments. *Journal of Statistical Planning and Inference*, 126:119–151, 2004.

[18] P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Reprinted in *International Journal of Approximate Reasoning*, 44(3):366–383, 2007.