Coherent updating on finite spaces

Enrique Miranda
Dep. of Statistics and Operations Research
Rey Juan Carlos University, Spain
enrique.miranda@urjc.es

Abstract

We compare the different notions of conditional coherence within the behavioural theory of imprecise probabilities when all the spaces are finite. We show that the differences between the notions are due to conditioning on sets of (lower, and in some cases upper) probability zero. Next, we characterise the range of coherent extensions, proving that the greatest coherent extensions can always be calculated using the notion of regular extension.

Keywords: Lower previsions, avoiding partial loss, weak and strong coherence, regular extension, natural extension.

1 Introduction

This paper is devoted to the study of the different notions of coherence within the theory of conditional lower previsions. This theory, established mainly in [3], provides a behavioural interpretation of probability in terms of acceptable buying and selling prices for gambles. It includes as particular cases most of the other uncertainty models present in the literature.

In spite of these virtues, one of the reasons why this theory is not so popular is the difficulty in the verification of the coherence of a number of assessments. The goal of this paper is to provide more manageable expressions for the notions of weak and strong coherence of a number of conditional lower previsions.

We restrict ourselves here to finite spaces. In that case, the different notions of coherence can be simplified a bit, and conditional lower previsions can be seen as models for the imprecise knowledge of conditional linear previsions (conditional expectations with respect to finitely additive probabilities). The finite case is also interesting for a number of applications, for instance in the context of credal networks.

The paper is organised as follows: in Section 2, we give a brief introduction to the behavioural theory of conditional lower previsions; in section 3, we compare the notions of weak and strong coherence, and avoiding partial loss; in section 4 we provide the smallest and the greatest conditional lower previsions which are coherent with some joint; and in section 5 we give some further comments on the subject.

2 Coherence notions on finite spaces

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 [3] for an in-depth study of these and other properties.

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

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 55–62
Torremolinos (Málaga), June 22–27, 2008

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 any $\epsilon > 0$ and any f in $\mathcal{K}$ the subject is disposed to accept the uncertain reward $f - \underline{P}(f) + \epsilon$. Given a lower prevision $\underline{P}$, $\overline{P}$ will denote its conjugate *upper prevision*, given by $\overline{P}(f) = -\underline{P}(-f)$ for any gamble f. $\overline{P}(f)$ represents the infimum acceptable selling price for the gamble f for our subject.

We can also consider the supremum buying prices for a gamble, *conditional* on a subset of Ω. Given such a set B and a gamble f on Ω, the lower prevision $\underline{P}(f|B)$ represents the subject's supremum acceptable buying price for the gamble f, updated after coming to know that the unknown value ω belongs to B, and nothing else. If we consider a partition $\mathcal{B}$ of Ω (for instance a set of categories), then we shall represent by $\underline{P}(f|\mathcal{B})$ the gamble on Ω that takes the value $\underline{P}(f|B)$ if and only if $\omega \in B$. The functional $\underline{P}(\cdot|\mathcal{B})$ that maps any gamble f on its domain into the gamble $\underline{P}(f|\mathcal{B})$ is called a *conditional lower prevision*.

Let us now re-formulate the above concepts in terms of random variables, which are the focus of our attention in this paper. Consider random variables $X_1, \dots, X_n$, taking values in respective *finite* sets $\mathcal{X}_1, \dots, \mathcal{X}_n$. For any subset $J \subseteq \{1, \dots, n\}$ we shall denote by X_J the (new) random variable

$$X_J := (X_j)_{j \in J},$$

which takes values in the product space

$$\mathcal{X}_J := \times_{j \in J} \mathcal{X}_j.$$

We shall also use the notation $\mathcal{X}^n$ for $\mathcal{X}_{\{1,\dots,n\}}$. In the current formulation made by random variables, $\mathcal{X}^n$ is just the definition of the possibility space Ω.

Definition 1. Let J be a subset of $\{1, \dots, n\}$, 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 any $x, y \in \mathcal{X}^n$, $\pi_J(x) = \pi_J(y)$ 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 subsets O, I of $\{1, \dots, n\}$. $\underline{P}(X_O|X_I)$ represents a subject's behavioural dispositions about the gambles that depend on the outcome of the variables $\{X_k, k \in O\}$, after coming to know the outcome of the variables $\{X_k, k \in I\}$. As such, it is defined on the set of gambles that depend on the values of the variables in $O \cup I$ only, i.e., on 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 $x \in \mathcal{X}_I$, $\underline{P}(f|X_I = x)$ represents a subject's supremum acceptable buying price for the gamble f, if he came to know that the variable X_I took the value x (and nothing else). Under the notation we gave above for lower previsions conditional on events and partitions, this would be $\underline{P}(f|B)$, where $B := \pi_I^{-1}(x)$. When there is no possible confusion about the variables involved in the lower prevision, we shall use the notation $\underline{P}(f|x)$ for $\underline{P}(f|X_I = x)$. The sets $\{\pi_I^{-1}(x) : x \in \mathcal{X}_I\}$ form a partition of $\mathcal{X}^n$. Hence, we can define the gamble $\underline{P}(f|X_I)$, which takes the value $\underline{P}(f|x)$ on $x \in \mathcal{X}_I$. This is a conditional lower prevision.

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 $\mathcal{K}_{O\cup I}$ should be *separately coherent*. In this case, where the domain is a linear set of gambles, separate coherence holds if and only if the following conditions are satisfied for any $x \in \mathcal{X}_I$, $f, g \in \mathcal{K}_{O\cup I}$, and $\lambda > 0$:

$$\underline{P}(f|x) \geq \min_{\omega \in \pi_I^{-1}(x)} f(\omega). \tag{SC1}$$

$$\underline{P}(\lambda f|x) = \lambda \underline{P}(f|x). \tag{SC2}$$

$$\underline{P}(f + g|x) \geq \underline{P}(f|x) + \underline{P}(g|x). \tag{SC3}$$

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 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 the following three conditions hold for any $f, g \in \mathcal{K}_O$, and $\lambda > 0$:

$$\underline{P}(f) \geq \min f. \tag{C1}$$

$$\underline{P}(\lambda f) = \lambda \underline{P}(f). \tag{C2}$$

$$\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g). \tag{C3}$$

In general, separate coherence is not enough to guarantee the consistency of the lower previsions: conditional lower previsions can be conditional on the values of many different variables, and still we should verify that the assessments they provide are consistent not only separately, but also with each other. Formally, we are going to consider what we shall call *collections* of conditional lower previsions.

Definition 2. Consider conditional previsions $\{\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})\}$ with respective domains $\mathcal{K}^1, \dots, \mathcal{K}^m \subseteq \mathcal{L}(\mathcal{X}^n)$, where $\mathcal{K}^j$ is the set of $\mathcal{X}_{O_j \cup I_j}$-measurable gambles,[1] for $j = 1, \dots, m$. This is called a *collection* on X^n when for each $j_1 \neq j_2$ in $\{1, \dots, m\}$, either $O_{j_1} \neq O_{j_2}$ or $I_{j_1} \neq I_{j_2}$.

This means that we do not have two different conditional lower previsions giving information about the same set of variables X_O, conditional on the same set of variables X_I. Given a collection $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$ of conditional lower previsions, there are different ways in which we can guarantee their consistency[2]. The first one is called avoiding partial loss.

The $\mathcal{X}_I$-*support* $S(f)$ of a gamble f in $\mathcal{K}_{O \cup I}$ is given by

$$S(f) := \{\pi_I^{-1}(x) : x \in \mathcal{X}_I, f\mathbb{I}_{\pi_I^{-1}(x)} \neq 0\}, \tag{1}$$

i.e., it is the set of conditioning events for which the restriction of f is not identically zero. We shall also use the notations

$$G(f|x) = \mathbb{I}_x(f - \underline{P}(f|x)),$$
$$G(f|X_I) = \sum_{x \in \mathcal{X}_I} G(f|x) = f - \underline{P}(f|X_I)$$

for any $f \in \mathcal{K}_{O \cup I}$ and any $x \in \mathcal{X}_I$.

Definition 3. Consider separately coherent $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$. We say that they *avoid partial loss* when for any $f_j \in \mathcal{K}^j$, $j = 1, \dots, m$,

$$\max_{\omega \in A_{f_1,\dots,f_m}} \left[\sum_{j=1}^m G_j(f_j|X_{I_j}) \right](\omega) \geq 0,$$

where $A_{f_1,\dots,f_m}$ is the set of elements that belong to some $B \in S_j(f_j)$ for some $j = 1, \dots, m$.

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.

Definition 4. Consider separately coherent $\underline{P}_1(X_{O_1}|X_{I_1}), \dots, \underline{P}_m(X_{O_m}|X_{I_m})$. We say that they are *weakly coherent* when for any $f_j \in \mathcal{K}^j$, $j = 1, \dots, m$, $j_0 \in \{1, \dots, m\}$, $f_0 \in \mathcal{K}^{j_0}, x_0 \in \mathcal{X}_{I_{j_0}}$,

$$\max_{\omega \in \mathcal{X}^n} \left[\sum_{j=1}^m G_j(f_j|X_{I_j}) - G_{j_0}(f_0|x_0) \right](\omega) \geq 0.$$

With this condition we require that our subject should not be able to raise his supremum acceptable buying price $\underline{P}_{j_0}(f_0|x_0)$ for a gamble f_0 contingent on x_0 by taking into account other conditional assessments. However, under the behavioural interpretation, a number of weakly coherent conditional lower previsions can still present some forms of inconsistency with each other; see [3, Example 7.3.5] for an example and [3, Chapter 7] and [4] for some discussion. On the other hand, weak coherence neither implies or is implied by the notion of avoiding partial loss. Because of

[1] 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.

[2] We give the particular definitions of these notions for finite spaces. See [1, 3] for the general definitions of these notions on infinite spaces and non-linear domains.

these two facts, we consider another notion which is stronger than both, and which is called (*joint* or *strong*) coherence:[3]

Definition 5. Consider separately coherent $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$. We say that they are *coherent* when for every $f_j \in \mathcal{K}^j$, $j=1,\dots,m$, $j_0 \in \{1,\dots,m\}$, $f_0 \in \mathcal{K}^{j_0}, x_0 \in \mathcal{X}_{I_{j_0}}$,

$$\left[\sum_{j=1}^{m} G_j(f_j|X_{I_j}) - G_{j_0}(f_0|x_0)\right](\omega) \geq 0 \quad (2)$$

for some $\omega \in \pi^{-1}_{I_{j_0}}(x_0) \cup A_{f_1,\dots,f_m}$.

Because we are dealing with finite spaces, this notion coincides with the one given by Williams in [5]. 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.

It is important at this point to introduce a particular case of conditional lower previsions that will be of special interest for us: that of *conditional linear previsions*. We say that a conditional lower prevision $\underline{P}(X_O|X_I)$ on the set $\mathcal{K}_{O\cup I}$ is linear if and only if it is separately coherent and moreover $\underline{P}(f+g|x)=\underline{P}(f|x)+\underline{P}(g|x)$ for any $x \in \mathcal{X}_I$ 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 (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 use the notation $P(X_O)$.

One interesting particular case is that where we are given only an unconditional lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}^n)$ and a conditional lower prevision $\underline{P}(X_O|X_I)$ on $\mathcal{K}_{O\cup I}$. Then weak and strong coherence are equivalent, and they both hold if and only if, for any $\mathcal{X}_{O\cup I}$-measurable f and any $x \in \mathcal{X}_I$,

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

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

If P and $P(X_O|X_I)$ are linear, they are coherent if and only if for any $\mathcal{X}_{O\cup I}$-measurable f, $P(f) = P(P(f|X_I))$. This is equivalent to requiring that $P(f|x) = \frac{P(f\mathbb{I}_x)}{P(x)}$ for all $f \in \mathcal{K}_{O\cup I}$ and all $x \in \mathcal{X}_I$ with $P(x) > 0$.

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

3 Relationships between weak and strong coherence

Let us study in more detail the notions of avoiding sure loss, weak coherence and strong coherence. We start by recalling a recent characterisation of weak coherence:

Theorem 1. *[2, Theorem 1] $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$ are weakly coherent if and only if there is a lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}^n)$ that is pairwise coherent with each conditional lower prevision $\underline{P}_j(X_{O_j}|X_{I_j})$. In particular, given linear conditional previsions $P_j(X_{O_j}|X_{I_j})$ for $j=1,\dots,m$, they are weakly coherent if and only if there is a linear prevision P which is coherent with each $P_j(X_{O_j}|X_{I_j})$.*

This theorem shows one of the differences between weak and strong coherence: weak coherence is equivalent to the existence of a joint which is coherent with each of the assessments, considered separately; coherence on the other hand is equivalent to the existence of a joint which is coherent with all the assessments, taken together.

Weakly coherent conditional previsions can be given a sensitivity analysis interpretation as lower envelopes of precise models; a similar result for coherence can be found in [3, Theorem 8.1.9].

Theorem 2. *Any weakly coherent $\underline{P}_1(X_{O_1}|X_{I_1}),\dots,\underline{P}_m(X_{O_m}|X_{I_m})$ are the lower envelope of a family of weakly coherent conditional linear previsions.*

From Theorem 1, weakly coherent conditional lower previsions always have a compatible joint $\underline{P}$. Our following result establishes the smallest such joint:

Theorem 3. *Consider weakly coherent $\underline{P}(X_{O_1}|X_{I_1}),\dots,\underline{P}(X_{O_m}|X_{I_m})$, and let $\underline{E}$ be given on $\mathcal{L}(\mathcal{X}^n)$ by*

$$\underline{E}(f) := \sup\{\alpha : \exists f_j \in \mathcal{K}^j, j = 1,\dots,m, s.t. \max_{\omega\in\mathcal{X}^n}[\sum_{j=1}^{m} G(f_j|X_{I_j}) - (f-\alpha)](\omega) < 0\}. \quad (3)$$

$\underline{E}$ is the smallest coherent lower prevision which is coherent with each $\underline{P}(X_{O_j}|X_{I_j})$.

Using this result and Theorem 2, we can also give a sensitivity analysis interpretation to $\underline{E}$ in the precise case.

Corollary 1. *Given weakly coherent $P(X_{O_1}|X_{I_1}),\dots,P(X_{O_m}|X_{I_m})$, the lower prevision $\underline{E}$ defined in* (3) *is the lower envelope of the set $\mathcal{M}$ of linear previsions which are coherent with each $P(X_{O_j}|X_{I_j}), j = 1,\dots,m$.*

Let us focus now on the relationship between weak and strong coherence and avoiding partial loss. We start by considering this problem in the precise case. In this case coherence is equivalent to avoiding partial loss, and is in general greater than weak coherence; see [3, Example 7.3.5] for an example of weakly coherent conditional previsions that incur a sure loss. We are going to show next that when a number of conditional previsions are weakly coherent but not coherent, this is due to the definition of the conditional previsions on some sets of probability zero.

Theorem 4. *Consider weakly coherent conditional linear previsions $P(X_{O_1}|X_{I_1}),\dots,P(X_{O_m}|X_{I_m})$ with respective domains $\mathcal{K}^1,\dots,\mathcal{K}^m$, and let $\overline{E}$ be the conjugate of the functional $\underline{E}$ defined in* (3)*. They are coherent if and only if for all gambles $f_j \in \mathcal{K}^j$, $j = 1,\dots,m$ with $\overline{E}(A_{f_1,\dots,f_m}) = 0$, $\max_{\omega\in A_{f_1,\dots,f_m}} \sum_{j=1}^{m}[f_j - P(X_{O_j}|X_{I_j})](\omega) \geq 0$.*

Taking into account this theorem and the envelope result established in Theorem 2, we can characterise the difference between weak coherence and avoiding partial loss for conditional lower previsions:

Corollary 2. *Consider weakly coherent $\underline{P}(X_{O_1}|X_{I_1}),\dots,\underline{P}(X_{O_m}|X_{I_m})$. They avoid partial loss if and only if for all $f_j \in \mathcal{K}^j$, $j = 1,\dots,m$ with $\overline{E}(A_{f_1,\dots,f_m}) = 0$, $\max_{\omega\in A_{f_1,\dots,f_m}} \sum_{j=1}^{m}[f_j - \underline{P}(X_{O_j}|X_{I_j})](\omega) \geq 0$, where $\overline{E}$ is the conjugate of the functional defined in* (3)*.*

Hence, if a number of weakly coherent lower previsions incur sure loss, this incoherent behaviour is due to the definition of the conditional previsions on some sets of zero upper probability. It may be argued, specially since we are dealing with finite spaces, that we may modify the definition of these conditional lower previsions on these sets in order to avoid partial loss without further consequences, in the sense that this will not affect their weak coherence: they will still be weakly coherent with the same unconditional $\underline{P}$.

So let us consider a number of weakly coherent conditional lower previsions that avoid partial loss. Our next example shows that, unlike for precise previsions, this is not sufficient for coherence. Hence, Theorem 4 does not extend to the imprecise case. This is because the condition equivalent to avoiding partial loss in Corollary 2 is not equivalent to coherence, in the sense that the union of the supports of a number of gambles producing incoherence may have positive upper probability:

Example 1. Consider two random variables X_1, X_2 taking values in the finite space $\mathcal{X} := \{1,2,3\}$, and let us define conditional lower previsions $\underline{P}(X_2|X_1)$ and $\underline{P}(X_1|X_2)$ by

$$\begin{aligned}
&\underline{P}(f|X_1 = 1) = f(1,1)\\
&\underline{P}(f|X_1 = 2) = f(2,3)\\
&\underline{P}(f|X_1 = 3) = \min\{f(3,2), f(3,3)\}\\
&\underline{P}(f|X_2 = 1) = f(2,1)\\
&\underline{P}(f|X_2 = 2) = \min\{f(1,2), f(2,2), f(3,2)\}\\
&\underline{P}(f|X_2 = 3) = \min\{f(1,3), f(2,3), f(3,3)\},
\end{aligned}$$

for any gamble f in $\mathcal{L}(\mathcal{X}^2)$.

Let us consider the unconditional lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}^2)$ given by $\underline{P}(f) = \min\{f(3,2), f(3,3)\}$. Using Theorem 1, we can see that $\underline{P}, \underline{P}(X_1|X_2)$ and $\underline{P}(X_2|X_1)$ are weakly coherent.

To see that $\underline{P}(X_1|X_2)$ and $\underline{P}(X_2|X_1)$ avoid partial loss, we apply Corollary 2 and consider any $f_1, f_2 \in \mathcal{L}(\mathcal{X}^2)$ such that $\overline{P}(A_{f_1,f_2}) = 0$. Let us prove that

$$\max_{\omega \in A_{f_1,f_2}} [G(f_1|X_2) + G(f_2|X_1)](\omega) \geq 0. \quad (4)$$

Assume $f_1 \neq 0 \neq f_2$; the other cases are similar (and easier). Since $\overline{P}(A_{f_1,f_2}) = 0$ for any coherent lower prevision that is weakly coherent with $\underline{P}(X_1|X_2)$ and $\underline{P}(X_2|X_1)$, we deduce that neither $(3,2)$ nor $(3,3)$ belong to A_{f_1,f_2}, and consequently $f_1(x,2) = f_1(x,3) = 0$ for $x = 1,2,3$. If $(X_1 = 2) \in S_1(f_2)$, then $[G(f_1|X_2) + G(f_2|X_1)](2,3) = 0 + 0 = 0$, and therefore Equation (4) holds. If $(X_1 = 2) \notin S_1(f_2)$, then $[G(f_1|X_2) + G(f_2|X_1)](2,1) = 0 + 0 = 0$.

Let us prove finally that $\underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ are not coherent. Let $f_1 = -I_{\{(1,1),(3,1)\}}$, $f_2 = -I_{\{(1,2),(1,3),(2,1),(2,2)\}}$ and $f_3 = I_{\{(2,3),(3,3)\}}$, and let us show that

$$[G(f_1|X_2) + G(f_2|X_1) - G(f_3|X_2 = 3)](\omega) < 0$$

for all $\omega \in B := \pi_2^{-1}(3) \cup A_{f_1,f_2}$. $S_2(f_1) = \{X_2 = 1\}$ and $S_1(f_2) = \{X_1 = 1, X_1 = 2\}$, whence $B = S_2(f_1) \cup S_1(f_2) \cup \{X_2 = 3\} = \mathcal{X}^2 \setminus \{(3,2)\}$. On the other hand, the gamble $g := G(f_1|X_2) + G(f_2|X_1) - G(f_3|X_2 = 3)$ satisfies $g(\omega) = -1$ for all $\omega \in B$. This shows that $\underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ are not coherent. However, $\overline{E}(B) = 1$ because $(3,3) \in B$. ♦

Hence, when a number of conditional lower previsions are weakly coherent but not coherent, the behaviour causing a contradiction can be caused by conditioning on sets of positive upper probability.

It is interesting to look for conditions under which it suffices to check the weak coherence of a number of previsions to be able to deduce their coherence. One such condition was established, in a different context, in [2].

In the case of conditional linear previsions, Theorem 4 allows us to derive immediately the following result:

Lemma 1. *Consider weakly coherent $P_1(X_{O_1}|X_{I_1}), \ldots, P_m(X_{O_m}|X_{I_m})$, and let P be a coherent prevision such that $P, P_j(X_{O_j}|X_{I_j})$ are coherent for $j = 1, \ldots, m$. If $P(x) > 0$ for any $x \in \mathcal{X}_{I_j}$, $j = 1, \ldots, m$, then $P_1(X_{O_1}|X_{I_1}), \ldots, P_m(X_{O_m}|X_{I_m})$ are coherent.*

From this result, we can easily derive a similar condition for conditional lower previsions.

Theorem 5. *Consider weakly coherent $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$, and let $\underline{P}$ be a coherent prevision such that $\underline{P}, \underline{P}_j(X_{O_j}|X_{I_j})$ are coherent for $j = 1, \ldots, m$. If $\underline{P}(x) > 0$ for all $x \in \mathcal{X}_{I_j}$ and all $j = 1, \ldots, m$, then $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ are coherent.*

We can deduce from the proof of this theorem that if a number of weakly coherent conditional lower previsions avoid partial loss but are not coherent, for any gambles $f_0, \ldots, f_m$ violating Eq. (2) it must be $\underline{E}(\pi_{I_{j_0}}^{-1}(x_0) \cup A_{f_1,\ldots,f_m}) = 0$ (although, as Example 1 shows, it can be $\overline{E}(\pi_{I_{j_0}}^{-1}(x_0) \cup A_{f_1,\ldots,f_m}) > 0$).

Note that when the conditioning events have all positive lower probability, the conditional lower previsions are uniquely determined by the joint $\underline{P}$ and by (GBR). Hence, in that case $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ are the only conditional previsions which are coherent with $\underline{P}$.

4 Coherent updating

Although our last result is interesting, it is fairly common in situations of imprecise information to be conditioning on some sets of lower probability zero and positive upper probability. In that case, there is an infinite number of conditional lower previsions which are coherent with the unconditional $\underline{P}$. In this section, we characterise them by determining the smallest and the greatest coherent extensions.

4.1 Updating with the regular extension

The first updating rule we are going to consider is called the *regular extension.* Consider an unconditional lower prevision $\underline{P}$ and disjoint O, I in $\{1, \ldots, n\}$. The conditional lower prevision $\underline{R}(X_O|X_I)$ defined by regular extension is given, for any $f \in \mathcal{K}_{O\cup I}$ and any $x \in \mathcal{X}_I$, by

$$\underline{R}(f|x) := \inf\left\{\frac{P(f\mathbb{I}_x)}{P(x)} : P \geq \underline{P}, P(x) > 0\right\}.$$

For this definition to be applicable, we need that $\overline{P}(x) > 0$ for any $x \in \mathcal{X}_I$. The regular extension is the lower envelope of the updated linear previsions using Bayes's rule.

Lemma 2. *Let $\underline{P}, \underline{P}(X_O|X_I)$ be coherent unconditional and conditional previsions, with $\mathcal{X}_I$ finite. Assume that $\overline{P}(x) > 0$ for all $x \in \mathcal{X}_I$, and define $\underline{R}(X_O|X_I)$ from $\underline{P}$ using regular extension. Then:*

1. *$\underline{P}, \underline{R}(X_O|X_I)$ are coherent.*
2. *$\underline{R}(X_O|X_I) \geq \underline{P}(X_O|X_I)$.*
3. *For any $P \geq \underline{P}$, there exists some $P(X_O|X_I)$ which is coherent with P and dominates $\underline{P}(X_O|X_I)$.*

From this lemma, we can deduce that if we use regular extension to define conditional lower previsions $\underline{R}_1(X_{O_1}|X_{I_1}), \ldots, \underline{R}_m(X_{O_m}|X_{I_m})$ from an unconditional $\underline{P}$, then $\underline{P}, \underline{R}_1(X_{O_1}|X_{I_1}), \ldots, \underline{R}_m(X_{O_m}|X_{I_m})$ are weakly coherent. Moreover, if we consider any other weakly coherent conditional lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$, it must hold that $\underline{R}_j(X_{O_j}|X_{I_j}) \geq \underline{P}_j(X_{O_j}|X_{I_j})$ for $j = 1, \ldots, m$. Hence, the procedure of regular extension provides the greatest, or more informative, updated lower previsions that are weakly coherent. In the following theorem we prove that they are also coherent.

Theorem 6. *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\mathcal{X}^n)$, and consider disjoint O_j, I_j for $j = 1, \ldots, m$. Assume that $\overline{P}(x) > 0$ for all $x \in \mathcal{X}_{I_j}$, and let us define $\underline{R}_j(X_{O_j}|X_{I_j})$ using regular extension for $j = 1, \ldots, m$. Then $\underline{P}, \underline{R}_1(X_{O_1}|X_{I_1}), \ldots, \underline{R}_m(X_{O_m}|X_{I_m})$ are coherent.*

When $\overline{P}(x) = 0$ for some $x \in \mathcal{X}_{I_j}$, $j = 1, \ldots, m$, we cannot use regular extension to define $\underline{R}_j(X_{O_j}|x)$. It can be checked that in that case any separately coherent conditional lower prevision is weakly coherence with $\underline{P}$. However, we cannot guarantee the strong coherence:

Example 2. Let $\mathcal{X}_1 = \mathcal{X}_2 = \{1, 2, 3\}$, and $P(X_1), P(X_2|X_1)$ determined by $P(X_1 = 3) = 1$, and $P(X_2 = x|X_1 = x) = 1$ for $x = 1, 2, 3$. From [3, Theorem 6.7.2], $P(X_1), P(X_2|X_1)$ are coherent. However, if we define arbitrarily $P(X_1|X_2 = x)$ when $P(X_2 = x) = 0$ (that is, for $x = 1, 2$), then $P(X_1|X_2)$ and $P(X_2|X_1)$ may not be coherent: make it for instance $P(X_1 = 1|X_2 = 2) = 1 = P(X_1 = 2|X_2 = 1) = P(X_1 = 3|X_2 = 3)$. Then [3, Example 7.3.5] shows that $P(X_1|X_2)$ and $P(X_2|X_1)$ are not coherent. ♦

4.2 Updating with the natural extension

Next, we introduce the notion of *natural extension.* Consider conditional lower previsions $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{O_m})$ defined on linear spaces $\mathcal{H}^1, \ldots, \mathcal{H}^m$ and avoiding partial loss. Given $j_0 \in \{1, \ldots, m\}$, a gamble f on $\mathcal{X}^n$ and an element x_0 of $\mathcal{X}_{I_{j_0}}$, the natural extension $\underline{E}_{j_0}(f|x_0)$ is defined as the supremum α for which there are $f_j \in \mathcal{H}^j, j = 1, \ldots, m$ such that

$$\left[\sum_{j=1}^{m} G(f_j|X_{I_j}) - I_{\pi^{-1}_{I_{j_0}}(x_0)}(f - \alpha)\right](\omega) < 0$$

for all $\omega \in \pi^{-1}_{I_{j_0}} \cup A_{f_1,\ldots,f_m}$. It is proven in [3, Theorem 8.1.9] that the lower previsions $\underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ obtained in this way are the smallest coherent conditional previsions that dominate $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ on their domains.

Given disjoint O_j, I_j for $j = 1, \ldots, m$, we can define separately coherent $\underline{P}_j(X_{O_j}|X_{I_j})$ on the linear set of constant gambles by $\underline{P}_j(\mu|x) = \mu$ for all $x \in \mathcal{X}_j$,

$j = 1, \ldots, m$. Then, given any coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}^n)$ the lower previsions $\underline{P}, \underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ avoid partial loss. We can then consider their natural extensions $\underline{P}, \underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ using the above definition.

Theorem 7. *If we use the above procedure,* $\underline{P}, \underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ *are coherent.*

The natural extension provides the smallest conditional lower previsions which are coherent together with $\underline{P}$. The conditional lower previsions $\underline{E}_j(X_{O_j}|X_{I_j})$ are uniquely determined by the (GBR) when $\underline{P}(x) > 0$ and are vacuous when $\underline{P}(x) = 0$, being then defined by $\underline{E}_j(f|x) = \min_{\omega \in \pi^{-1}_{I_j}(x)} f(\omega)$ for any $f \in \mathcal{K}^j$. Hence, in that respect the natural extensions can be calculated more easily than the regular extensions.

We showed before that the conditional previsions defined by regular extension were also the greatest conditional lower previsions that are weakly coherent with the unconditional lower prevision $\underline{P}$. Using Theorem 1 and the results in [3, Chapter 6], it is not difficult to show that the natural extensions are the smallest weakly coherent extensions:

Theorem 8. *Let* $\underline{P}$ *be coherent on* $\mathcal{L}(\mathcal{X}^n)$*, and define* $\underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ *using natural extension. Then* $\underline{P}, \underline{E}_1(X_{O_1}|X_{I_1}), \ldots, \underline{E}_m(X_{O_m}|X_{I_m})$ *are weakly coherent and for any other conditional previsions* $\underline{P}_1(X_{O_1}|X_{I_1}), \ldots, \underline{P}_m(X_{O_m}|X_{I_m})$ *which are weakly coherent with* $\underline{P}$*, it holds that* $\underline{P}_j(X_{O_j}|X_{I_j}) \geq \underline{E}_j(X_{O_j}|X_{I_j})$ *for* $j = 1, \ldots, m$.

5 Conclusions

In this paper we have studied the difference between weak and strong coherence in the case of finite spaces, and established the smallest and greatest updated previsions. Although weak and strong coherence are not equivalent, the smallest and greatest weakly coherent updated previsions coincide with the smallest and greatest coherent updated previsions, and are given by the natural and regular extensions, respectively.

It is important to remark that most of the properties established in this paper do not extend to conditional previsions on infinite spaces: for instance, weakly coherent conditional lower previsions are not necessarily lower envelopes of sets of weakly coherent conditional linear previsions. Similarly, the regular extensions are not necessarily coherent with the unconditional prevision, and they will only provide an upper bound for the greatest coherent updated previsions. The study of these properties in the infinite case is the main open problem we point out for the future.

Acknowledgements

This research was started while I was making a stay at IDSIA with Marco Zaffalon. His encouragement and helpful comments are gratefully acknowledged. I also thank the financial support of the project TSI2007-66706-C04-01.

References

[1] E. Miranda and G. de Cooman (2005). Coherence and independence in non-linear spaces. Technical Report 05/10, Rey Juan Carlos University.

[2] E. Miranda and M. Zaffalon (2007). Coherence graphs. In *Proceedings of ISIPTA'07*, pages 297-307, Prague, Czech Republic, July 2007.

[3] P. Walley (1991). *Statistical reasoning with imprecise probabilities.* Chapman and Hall, London.

[4] P. Walley, R. Pelessoni, and P. Vicig (2004). Direct algorithms for checking consistecy and making inferences for conditional probability assessments. *Journal of Statistical Planning and Inference*, 126(1):119–151.

[5] P. Williams (1975). Notes on conditional previsions. Technical report, University of Sussex. Reprinted in a revised form in the International Journal of Approximate Reasoning, 44(3):366-383, 2007.