MARGINAL EXTENSION IN THE THEORY OF COHERENT LOWER PREVISIONS

ENRIQUE MIRANDA AND GERT DE COOMAN

ABSTRACT. We generalise Walley's Marginal Extension Theorem to the case of any finite number of conditional lower previsions. Unlike the procedure of natural extension, our marginal extension always provides the smallest (most conservative) coherent extensions. We show that they can also be calculated as lower envelopes of marginal extensions of conditional linear (precise) previsions. Finally, we use our version of the theorem to study the so-called forward irrelevant product and forward irrelevant natural extension of a number of marginal lower previsions.

1. INTRODUCTION

To sketch the context for this paper, let us consider a simple example. Suppose we have two random variables[1] X_1 and X_2 assuming values in the respective finite sets $\mathscr{X}_1$ and $\mathscr{X}_2$.[2] We have a *marginal* probability mass function p_1 for the first variable: for each x_1 in $\mathscr{X}_1$, $p_1(x_1)$ is the probability that the first random variable X_1 assumes the value x_1, irrespective of the value that the second random variable X_2 assumes in $\mathscr{X}_2$.

For the second variable, we have a *conditional* probability mass function: for all x_1 in $\mathscr{X}_1$ and x_2 in $\mathscr{X}_2$, $p_2(x_2|x_1)$ is the conditional probability that X_2 assumes the value x_2, given that X_1 assumes the value x_1.

We can then ask for the *joint* probability mass function m of X_1 and X_2. It is a consequence of Bayes' rule that

$$m(x_1,x_2) := p_1(x_1)p_2(x_2|x_1) \tag{1}$$

is the probability that the random variable (X_1,X_2) assumes the value (x_1,x_2) in $\mathscr{X}_1 \times \mathscr{X}_2$. It is instructive to rewrite this formula in terms of expectations, or, to use de Finetti's language, previsions ([6]). Consider a real-valued map h on $\mathscr{X}_1 \times \mathscr{X}_2$; we shall call such maps *gambles* because they can be interpreted as uncertain, or random, rewards. Call

$$M(h) := \sum_{(x_1,x_2)\in\mathscr{X}_1\times\mathscr{X}_2} m(x_1,x_2)h(x_1,x_2)$$

the *prevision* of h. Then it follows from Eq. (1) that

$$M(h) = \sum_{x_1\in\mathscr{X}_1} p(x_1)\left(\sum_{x_2\in\mathscr{X}_2} p(x_2|x_1)h(x_1,x_2)\right) = P_1(P_2(h|X_1)). \tag{2}$$

Key words and phrases. Imprecise probabilities, lower previsions, coherence, natural extension, marginal extension, epistemic irrelevance, forward irrelevance, forward irrelevant natural extension, forward irrelevant product.

[1]By a *random variable*, we mean a variable whose value is unknown.

[2]We assume finiteness of the sets $\mathscr{X}_1$ and $\mathscr{X}_2$ here only to make this introductory discussion as simple as possible. We shall consider more general situations further on.

Here we have denoted by $P_2(h|X_1)$ the real-valued map (gamble) on $\mathscr{X}_1$ that assumes the value

$$P_2(h|x_1) := P_2(h(x_1,\cdot)|x_1) = \sum_{x_2\in\mathscr{X}_2} h(x_1,x_2)p_2(x_2|x_1)$$

in $x_1 \in \mathscr{X}_1$, where $h(x_1,\cdot)$ denotes the gamble on $\mathscr{X}_2$ given by $h(x_1,\cdot)(x_2) = h(x_1,x_2)$. We see, then, that $P_2(h|x_1)$ is the conditional prevision of h given that X_1 assumes the value x_1. Similarly, we have denoted by $P_1(f) := \sum_{x_1\in\mathscr{X}_1} f(x_1)p_1(x_1)$ the (marginal) prevision of a gamble f on $\mathscr{X}_1$. The *concatenation formula* (2) is equivalent to Bayes' rule (1), and it provides a neat expression $M = P_1(P_2(\cdot|X_1))$ for the *joint prevision* M in terms of the *marginal prevision* P_1 and the *conditional prevision* $P_2(\cdot|X_1)$. This concatenation rule is also sometimes called the *conglomerative property* ([6, Sections 4.7 and 4.19]), or the *law of total probability*. In the language of this paper, we shall call M the *marginal extension* of the unconditional prevision P_1 on $\mathscr{X}_1$ and the conditional prevision $P_2(\cdot|X_1)$ on $\mathscr{X}_2$ to a joint linear prevision on $\mathscr{X}_1 \times \mathscr{X}_2$. Since it was derived using only Bayes' rule and the linear character of (conditional) previsions, it is a necessary consequence of *coherence* in de Finetti's account of (subjective) probability. In fact, marginal extension is the only coherent way to obtain a joint prevision from P_1 and $P_2(\cdot|X_1)$.

In his important work ([10]) on the behavioural theory of imprecise probabilities, Walley has generalised this result to the case where the uncertainty about the values of the random variables is not modelled by previsions, but rather by *lower and upper previsions*. On this approach, a subject's lower prevision $\underline{P}(f)$ for a gamble (or uncertain reward) f is the supremum price he is disposed to pay for buying the gamble, or in other words, the supremum s such that the subject accepts the gamble $f-s$. His upper prevision $\overline{P}(f)$ is the infimum price he is disposed to receive for selling the gamble, or in other words, the infimum s such that he accepts the gamble $s-f$. When the lower and upper prevision for a gamble happen to coincide, the common value is called the subject's *fair price*, or *(precise) prevision*, $P(f)$ for the gamble f. But in contrast with de Finetti's approach, it is not required that a subject should be disposed to always specify supremum buying and infimum selling prices that are equal to each other. What the two approaches to modelling uncertainty have in common, however, is that they impose certain rationality, or *coherence*, requirements on a subject's behavioural dispositions as summarised by his (lower or upper) previsions. Walley's behavioural theory of imprecise probabilities subsumes existing models of upper and lower expectations ([1, 7]), sets of probability measures ([8]), upper and lower previsions, sets of desirable gambles, and preference orderings ([10]). We give a reasonably detailed introduction to this theory in Section 2.

Walley's Marginal Extension Theorem (MET, [10, Section 6.7]) essentially states, then, that if we have a marginal lower prevision $\underline{P}_1$ for the first random variable X_1, and a conditional lower prevision $\underline{P}_2(\cdot|X_1)$ for the second random variable X_2, then their *marginal extension* $\underline{M}$, given by the concatenation $\underline{M} := \underline{P}_1(\underline{P}_2(\cdot|X_1))$, is the point-wise smallest (i.e., the behaviourally most conservative or least committal) *coherent* joint lower prevision for (X_1,X_2). Walley also shows that the marginal extension $\underline{M}$ is *uniquely coherent* (i.e., it is the only coherent extension) whenever the conditional lower prevision $\underline{P}_2(\cdot|X_1)$ is precise, i.e., a conditional prevision $P_2(\cdot|X_1)$, no matter whether $\underline{P}_1$ is precise or not.

What we shall do in this paper, and in particular in Sections 3 (for conditioning on partitions) and 5 (for conditioning on random variables), is generalise Walley's result to any finite number of partitions or random variables. This generalisation is not in any way obvious or immediate, as a comparison of our proofs with Walley's will show. This is because the general coherence requirements become considerably simpler to work with in the case of a single conditioning random variable (or partition, Walley's special case); see Sections 6.5 and 7.1, and in particular the Reduction Theorem 7.1.5 in [10] for more details.

In Section 4, we show how the marginal extension can be obtained as a lower envelope of a set of marginal extensions of precise (conditional) previsions. This allows us to provide our Marginal Extension Theorem with an additional *Bayesian sensitivity analysis* interpretation.

In Walley's theory of coherent lower previsions, there is a powerful notion of *natural extension*, which seeks to infer the (conditional) lower prevision for a gamble by making finite combinations of given (conditional) lower prevision assessments for other gambles; see Section 2 and [10] for more details. In Section 6, we compare natural and marginal extension, and we show by means of a counterexample that these extensions need not generally coincide. This provides another interesting example of the well-established fact (see [10, Section 8.1]) that natural extension may be incoherent when applied to conditional lower previsions.

Finally, in Section 7 we show that marginal extension provides us with a natural and interesting way of forming a 'product' of a number of given marginal lower previsions, based on a special type of 'independence' assessment. To see how this comes about, let us go back to the example in the beginning of this Introduction. Suppose we only have a marginal (as opposed to a conditional) probability mass function p_2 for the second random variable X_2. Then we can still use marginal extension to calculate the joint mass function, provided we can make the following independence assessment[3]

$$p_2(x_2|x_1) = p_2(x_2) \tag{3}$$

for all x_1 in $\mathscr{X}_1$ and x_2 in $\mathscr{X}_2$. In that case we find for the joint probability mass function that $m(x_1,x_2) = p_1(x_1)p_2(x_2)$, or in terms of previsions,

$$M(h) = \sum_{(x_1,x_2)\in\mathscr{X}_1\times\mathscr{X}_2} p_1(x_1)p_2(x_2)h(x_1,x_2) = P_1(P_2(h)),$$

where we let $P_2(h)$ be the gamble on $\mathscr{X}_1$ that assumes the value

$$P_2(h(x_1,\cdot)) := \sum_{x_2\in\mathscr{X}_2} p_2(x_2)h(x_1,x_2) = \sum_{x_2\in\mathscr{X}_2} p_2(x_2|x_1)h(x_1,x_2) = P_2(h(x_1,\cdot)|x_1)$$

in $x_1 \in \mathscr{X}_1$. For any gamble g on $\mathscr{X}_2$, this tells us that the marginal prevision $P_2(g)$ of g is equal to the conditional prevision $P_2(g|x_1)$ for all x_1 in $\mathscr{X}_1$, which is an equivalent way of formulating the independence assessment (3). The linear prevision $M = P_1(P_2(\cdot))$, obtained using marginal extension and the independence assessment, is called the (independent) *product* of the previsions P_1 and P_2. Observe that also[4] $M = P_2(P_1(\cdot))$ and that for any gambles f on $\mathscr{X}_1$ and g on $\mathscr{X}_2$, we find that $M(fg) = P_1(f)P_2(g)$, which is sometimes referred to as the *Product Rule*, and which provides an alternative (and more directly symmetrical) way to define independence.

Generalising this to marginal lower previsions $\underline{P}_1$ for X_1 and $\underline{P}_2$ for X_2 seems straightforward, but there is a catch. Indeed, we can now use Walley's Marginal Extension Theorem

---

[3]Nothing essential changes if we impose this requirement only for those x_1 in $\mathscr{X}_1$ for which $p(x_1) > 0$. This restriction is often made to ensure that the symmetrical counterpart '$p_1(x_1|x_2) = p_1(x_1)$ when $p_2(x_2) > 0$' is implied by Condition (3), which turns independence into a *symmetrical* notion. This symmetry is not immediately apparent in (3), and is actually broken when we generalise (3) to lower previsions. See further on.

[4]When $\mathscr{X}_1$ and $\mathscr{X}_2$ may be infinite, this symmetry is no longer guaranteed if we take de Finetti's position of only requiring the previsions P_1 and P_2 to be *finitely* additive on events (see [6, Section 3.11] for a discussion of finite additivity).

to find a point-wise smallest coherent joint lower prevision, provided we make the following assessment:

$$\underline{P}_2(g|x_1) = \underline{P}_2(g) \tag{4}$$

for all gambles g on $\mathscr{X}_2$ and all x_1 in $\mathscr{X}_1$. In that case $\underline{M}(h) = \underline{P}_1(\underline{P}_2(h))$, where, in a similar vein as before, $\underline{P}_2(h)$ is the gamble on $\mathscr{X}_1$ that assumes the value $\underline{P}_2(h(x_1,\cdot))$ in $x_1 \in \mathscr{X}_1$. A subject who makes assessment (4) effectively models that knowing what value X_1 assumes in $\mathscr{X}_1$ does not affect his beliefs about the value that X_2 assumes in $\mathscr{X}_2$. In Walley's terminology ([10, Chapter 9], see also [5]), X_1 is then said to be *epistemically irrelevant* to X_2. We shall call the lower prevision $\underline{M} = \underline{P}_1(\underline{P}_2(\cdot))$ the *forward irrelevant product* of the marginals $\underline{P}_1$ and $\underline{P}_2$.

Reversing the roles of X_1 and X_2 in this reasoning leads to the *backward irrelevant product* $\underline{M}' = \underline{P}_2(\underline{P}_1(\cdot))$, based on the assessment that X_2 is epistemically irrelevant to X_1. Interestingly, and in contrast with what we have just seen for precise previsions, it does not generally hold that $\underline{M} = \underline{M}'$, i.e., that $\underline{P}_1(\underline{P}_2(\cdot)) = \underline{P}_2(\underline{P}_1(\cdot))$; see [3] for more details and a counterexample. This means that epistemic irrelevance is an asymmetrical notion. To assert that X_1 and X_2 are *epistemically independent*, we have to require that X_1 is epistemically irrelevant to X_2 *and* that X_2 is epistemically irrelevant to X_1; see also [5] for further discussion. The *independent product* ([10, Section 9.3] and [5]) of the marginals $\underline{P}_1$ and $\underline{P}_2$ is then defined as the point-wise smallest coherent joint lower prevision with these marginals that expresses such epistemic independence. Such an independent product does not always exist,[5] and may, if it exists, be quite difficult to compute (see [10, Section 9.3.2]). In contrast, the forward/backward irrelevant products $\underline{M}$ and $\underline{M}'$ always exist (are always coherent), and are, as we have seen, very easy to compute.[6]

In Section 7 then, we use our more general version of the Marginal Extension Theorem to generalise the notion of a forward irrelevant product to any finite number of marginals. We also prove a number of interesting properties for this type of product, such as a generalised (but weaker) version of the above-mentioned Product Rule.

Why do we believe our results to be relevant? Our generalised version of the Marginal Extension Theorem allows us to make most-conservative (least-committal) and coherent inferences in a straight-forward manner in a number of interesting situations where the available information has an hierarchical structure, namely, when it is characterised by conditioning on increasingly finer partitions, or by nested collections of conditioning variables. And even though, obviously, not all probabilistic or statistical reasoning falls within the scope of marginal extension, it does so in a number of theoretically interesting as well as practically useful situations. Let us end this Introduction with two examples. In [4], one of us, in co-operation with M. Zaffalon, has used an earlier, less general version of the MET with three variables to justify using a so-called *conservative updating rule* for dealing with missing data in probabilistic expert systems based on Bayesian networks. And in [2], we use, amongst other things, the MET in conjunction with a forward irrelevance assessment, stating that 'we do not learn about the future by observing the past' to derive quite powerful weak and strong laws of large numbers that subsume most of the related results in the literature, and to weaken considerably the conditions under which such laws can be shown to hold.

[5] It always exists when $\mathscr{X}_1$ or $\mathscr{X}_2$ is finite, but there may be problems in case both $\mathscr{X}_1$ and $\mathscr{X}_2$ are infinite; again, see [10, Section 9.3].

[6] The independent product, if it exists, dominates both $\underline{M}$ and $\underline{M}'$ on all gambles.

2. Coherent lower previsions

In order to make this paper reasonably self-contained, we discuss here the relevant main ideas in the behavioural theory of coherent lower previsions, as formulated in much more detail and depth by Walley in [10].

2.1. **Basic notions and notation.** We consider a subject who is uncertain about something, say, the outcome of an experiment. Let Ω be the set of all possible outcomes, then a bounded real-valued function on Ω is called a *gamble* on Ω. The set of all gambles on Ω is denoted by $\mathscr{L}(\Omega)$. A gamble f is interpreted as an uncertain reward: if the outcome of the experiment turns out to be $\omega \in \Omega$, then the corresponding (positive or negative) reward will be $f(\omega)$, expressed in units of some (predetermined) linear utility.

As we have already announced in the Introduction, a subject's *lower prevision* $\underline{P}(f)$ for a gamble f is defined as his supremum acceptable price for buying f, i.e., the highest price μ such that the subject will accept to buy the uncertain reward f for all prices strictly smaller than μ (buying f for a price x is the same thing as accepting the uncertain reward $f-x$). Similarly, a subject's *upper prevision* $\overline{P}(f)$ for f is his infimum acceptable selling price for f, so he accepts the uncertain reward $\mu - f$ for all prices $\mu > \overline{P}(f)$. Clearly, $\overline{P}(f) = -\underline{P}(-f)$ since selling f for a price x is the same thing as buying $-f$ for the price $-x$. This *conjugacy relation* allows us to restrict our attention to lower previsions.

For any subset A of Ω, also called an *event*, its *lower probability* $\underline{P}(A)$ is defined as the lower prevision $\underline{P}(I_A)$ of its *indicator* I_A, where I_A denotes the gamble on Ω that assumes the value one on A and zero elsewhere. Similarly for its *upper* probability, $\overline{P}(A) = \overline{P}(I_A)$. $\underline{P}(A)$ can be interpreted as the supremum rate for betting on the occurence of the event A.

2.2. **Rationality criteria.** Assume that the subject has given lower prevision assessments $\underline{P}(f)$ for all gambles f in some set of gambles $\mathscr{K} \subseteq \mathscr{L}(\Omega)$, which need not have any predefined structure. Since these assessments represent commitments of the subject to act in certain ways, they are subject to a number of rationality requirements. The strongest such requirement that we shall consider here, is that $\underline{P}$ should be *coherent* ([10, Definition 2.5.1]). This is the case if for any $n, m \geq 0$ in $\mathbb{N}$, and $f_0, \ldots, f_n$ in $\mathscr{K}$:

$$\sup_{\omega\in\Omega}\left[\sum_{k=1}^{n}[f_k(\omega)-\underline{P}(f_k)]-m[f_0(\omega)-\underline{P}(f_0)]\right]\geq 0. \tag{5}$$

Assume that this condition fails for some $n, m > 0$, $f_0, \ldots, f_n \in \mathscr{K}$. Then, there would be some $\varepsilon > 0$ such that $m[f_0 - [\underline{P}(f_0)+\varepsilon]]$ point-wise dominates the acceptable combination of buying transactions $\sum_{k=1}^{n}[f_k - \underline{P}(f_k)+\varepsilon]$, and is therefore acceptable as well.[7] This would mean that by combining acceptable transactions derived from his assessments, the subject can be effectively forced to buy f_0 at the price $\underline{P}(f_0)+\varepsilon$, which is strictly higher than the supremum acceptable buying price $\underline{P}(f_0)$ that he has specified for it.

It also follows from Eq. (5), for the case where $m = 0$, that the subject's assessments *avoid sure loss* ([10, Definition 2.4.1]): for any n in the set of positive natural numbers $\mathbb{N}$ and for any $f_1, \ldots, f_n$ in $\mathscr{K}$ we require that

$$\sup_{\omega\in\Omega}\left[\sum_{k=1}^{n}[f_k(\omega)-\underline{P}(f_k)]\right]\geq 0.$$

[7] The underlying assumption, or axiom of rationality, here is that a finite non-negative linear combination of acceptable gambles is acceptable. This assumption is closely linked with the linearity of the chosen utility scale.

Otherwise, there would be some $\varepsilon > 0$ such that for all ω in Ω:

$$\sum_{k=1}^{n} [f_k(\omega) - \underline{P}(f_k) + \varepsilon] \leq -\varepsilon,$$

i.e., the net reward of buying the gambles f_k for the acceptable prices $\underline{P}(f_k) - \varepsilon$ is sure to lead to a loss of at least ε, whatever the outcome of the experiment!

And finally, if the coherence condition fails for some $n = 0, m > 0, f_0 \in \mathscr{K}$, we deduce that we can raise $\underline{P}(f_0)$ in some positive quantity ε, contradicting its interpretation as a supremum acceptable buying price. All of these are inconsistencies that should be avoided.

We list a few interesting consequences of coherence here, as they will turn out useful in later proofs. Any coherent lower prevision $\underline{P}$ is monotone: $f \leq g \Rightarrow \underline{P}(f) \leq \underline{P}(g)$; super-additive: $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$; positively homogeneous: $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ for all real $\lambda \geq 0$; constant additive: $\underline{P}(f+\mu) = \underline{P}(f) + \mu$, and also $\underline{P}(\mu) = \mu$ for all real numbers μ; and it satisfies $\underline{P}(f) \geq \inf f$. (When a gamble appears as an argument of $\underline{P}$ in the above expressions, it is of course assumed to be in the domain of $\underline{P}$.)

When the domain $\mathscr{K}$ of the lower prevision $\underline{P}$ is a linear space, i.e., closed under the point-wise addition of gambles, and the scalar (point-wise) multiplication of gambles with real numbers, the form of the coherence requirement simplifies considerably. Indeed, then $\underline{P}$ is coherent if and only if ([10, Section 2.3.3])

1. $\underline{P}(f) \geq \inf f$ [accepting sure gains; positivity];
2. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ [positive homogeneity];
3. $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$ [super-additivity]

for all f and g in $\mathscr{K}$ and non-negative real λ.

2.3. **Natural extension.** A lower prevision $\underline{P}$ defined on an arbitrary set of gambles $\mathscr{K}$ can, provided it avoids sure loss, always be corrected to a coherent lower prevision $\underline{E}$ on the set of all gambles $\mathscr{L}(\Omega)$, through a procedure called *natural extension*. The natural extension $\underline{E}$ of $\underline{P}$ is the smallest coherent lower prevision on $\mathscr{L}(\Omega)$ that point-wise dominates $\underline{P}$ on $\mathscr{K}$. It is given for all gambles f on Ω by ([10, Sections 3.1.1 and 3.1.3])

$$\underline{E}(f) = \sup_{\substack{f_1,\dots,f_n \in \mathscr{K} \\ \mu_1,\dots,\mu_n \geq 0, n \geq 0}} \inf_{\omega \in \Omega} \left[f(\omega) - \sum_{k=1}^{n} \mu_k [f_k(\omega) - \underline{P}(f_k)] \right]. \tag{6}$$

The natural extension summarises the behavioural implications of the assessments present in $\underline{P}$: $\underline{E}(f)$ is the supremum buying price for f that can be derived from the lower prevision $\underline{P}$ by arguments of coherence alone: we see from its definition above that it is the supremum of all prices that the subject can be effectively forced to buy the gamble f for, by combining finite numbers of buying transactions implicit in his lower prevision assessments $\underline{P}$.

The concept of natural extension can also be used to characterise the coherence of a lower prevision $\underline{P}$ on $\mathscr{K}$: a lower prevision $\underline{P}$ that avoids sure loss is coherent if and only if it coincides with its natural extension on its domain $\mathscr{K}$, i.e., if $\underline{E}$ is indeed an extension of $\underline{P}$. Observe that if $\underline{P}$ is coherent, then $\underline{E}$ will not be in general the unique coherent extension of $\underline{P}$ to $\mathscr{L}(\Omega)$; but any other coherent extension of $\underline{P}$ will dominate $\underline{E}$ on all gambles, and will therefore represent behavioural dispositions not present in $\underline{P}$.

We shall see further on in Section 6 that this notion of natural extension can be generalised from the unconditional lower previsions considered here to the conditional lower previsions to be introduced and studied later. In order to distinguish between the two types of natural extension, we shall sometimes refer to the present notion as *unconditional* natural extension.

2.4. **In terms of linear previsions.** When $\underline{P}(f) = \overline{P}(f)$, the subject's supremum buying price for the gamble f coincides with his infimum selling price, and this common value is then a *prevision* or *fair price* for the gamble f, in the sense of de Finetti ([6]). When a lower prevision $\underline{P}$ is defined on a negation invariant set of gambles $\mathscr{K}$, meaning that $-\mathscr{K} = \{-f : f \in \mathscr{K}\} = \mathscr{K}$, we can define the *conjugate upper prevision* $\overline{P}$ on $\mathscr{K}$ by $\overline{P}(f) = -\underline{P}(-f)$, $f \in \mathscr{K}$; and we call $\underline{P}$ *self-conjugate* if $\underline{P} = \overline{P}$ on $\mathscr{K}$, i.e., when our subject can determine fair prices for all the gambles in this domain.

We then define a *linear prevision* P on the set of all gambles $\mathscr{L}(\Omega)$ ([10, Section 2.8]) as a self-conjugate coherent lower prevision, or equivalently, as a real-valued linear functional on $\mathscr{L}(\Omega)$ that is positive (if $f \geq 0$ then $P(f) \geq 0$) and has unit norm ($P(I_\Omega) = 1$). Its restriction to events is a finitely additive probability. Let us denote by $\mathbb{P}(\Omega)$ the set of all linear previsions on $\mathscr{L}(\Omega)$. A real-valued functional P defined on some domain $\mathscr{K}$ is called a linear prevision if it can be extended to a linear prevision on all gambles.

The notions of avoiding sure loss, coherence, and natural extension can also be characterised in terms of sets of linear previsions. Consider a lower prevision $\underline{P}$ defined on a set of gambles $\mathscr{K}$. Its set of dominating linear previsions is given by

$$\mathscr{M}(\underline{P}) := \{P \in \mathbb{P}(\Omega) : (\forall f \in \mathscr{K})(P(f) \geq \underline{P}(f))\}.$$

Then $\underline{P}$ avoids sure loss if and only if $\mathscr{M}(\underline{P}) \neq \emptyset$, i.e., if $\underline{P}$ it has some dominating linear prevision. $\underline{P}$ is coherent if and only if $\underline{P}(f) = \min\{P(f) : P \in \mathscr{M}(\underline{P})\}$ for all f in $\mathscr{K}$, i.e., if $\underline{P}$ is the *lower envelope* of $\mathscr{M}(\underline{P})$. And the natural extension $\underline{E}$ of $\underline{P}$ is given by $\underline{E}(f) = \min\{P(f) : P \in \mathscr{M}(\underline{P})\}$ for all f in $\mathscr{L}(\Omega)$. The natural extension of a coherent lower prevision $\underline{P}$ can therefore be computed as the lower envelope of the linear previsions that dominate $\underline{P}$ on its domain. We deduce from this fact that the procedure of natural extension is transitive: if we consider the natural extension $\underline{E}_1$ of $\underline{P}$ to some domain $\mathscr{K}_1 \supseteq \mathscr{K}$ and then the natural extension of $\underline{E}_1$ to all gambles, this extension will agree with the natural extension of $\underline{P}$ to all gambles.

These properties allow us to give coherent lower previsions a *Bayesian sensitivity analysis interpretation*, in contrast with the direct behavioural one considered so far: we might assume the existence of some ideal but unknown linear prevision P modelling the behavioural dispositions of our subject, and model that we know P only imperfectly by stating that P belongs to some (compact and convex) set $\mathscr{M}$ of possible candidate linear previsions. Then this set of (precise) linear previsions is mathematically equivalent to its lower envelope $\underline{P}$, which is a coherent lower prevision, and $P \in \mathscr{M}$ is equivalent to $P \geq \underline{P}$.

2.5. **Conditioning.** Let $\mathscr{B}$ be a partition of Ω, i.e., a set of mutually disjoint events whose union is Ω. Then we can consider for every $B \in \mathscr{B}$ and any gamble f on Ω a subject's *conditional lower prevision* $\underline{P}(f|B)$ *of* f *given* B, defined as the supremum price he would currently be willing to pay for f if he came to know subsequently that the outcome of the experiment took a value in B (and nothing else). Alternatively, it could be defined as the subject's supremum buying price for the so-called *contingent* gamble $I_B f$, which is called-off when B doesn't occur.

If we assume that the conditional lower previsions $\underline{P}(\cdot|B)$ are defined on the same domain $\mathscr{H}$ for all $B \in \mathscr{B}$,[8] then we can summarise all these conditional lower previsions through the mathematical device $\underline{P}(\cdot|\mathscr{B})$ – a two-place function –, where for all $f \in \mathscr{H}$

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

[8]This is no essential restriction; for more details see [10, Section 6.2.4].

We shall also call the object $\underline{P}(\cdot|\mathscr{B})$ a *conditional lower prevision*. Interestingly, and quite importantly, $\underline{P}(f|\mathscr{B})$ can be regarded as a gamble on Ω that is *$\mathscr{B}$-measurable*, that is, constant on the elements of $\mathscr{B}$: it assumes the value $\underline{P}(f|B)$ on any $x \in B$. We shall frequently use the notations

$$G(f|\mathscr{B}) := f - \underline{P}(f|\mathscr{B}) = \sum_{B\in\mathscr{B}} G(f|B) = \sum_{B\in\mathscr{B}} I_B[f - \underline{P}(f|B)]. \tag{7}$$

$G(f|B)$ is called the *marginal gamble* on f *contingent on* B: it is the gamble that is called off unless B occurs, and where the subject pays the price $\underline{P}(f|B)$ for f, when B does occur.

Conditional lower previsions $\underline{P}(\cdot|\mathscr{B})$ are also subject to a number of rationality criteria, which we now turn to.

Separate coherence. The first requirement we consider is that of separate coherence:

Definition 1. [10, Section 6.2.2] Let $\underline{P}(\cdot|\mathscr{B})$ be a conditional lower prevision with domain $\mathscr{H}$. It is called *separately coherent* when the following two conditions hold:

(SC1) For each B in $\mathscr{B}$, $\underline{P}(\cdot|B)$ is a coherent lower prevision on $\mathscr{H}$.
(SC2) $I_B \in \mathscr{H}$ and $\underline{P}(B|B) = 1$ for each $B \in \mathscr{B}$.

It is an immediate but rather important consequence of separate coherence that

$$I_B f = I_B g \Rightarrow \underline{P}(f|B) = \underline{P}(g|B) \tag{8}$$

for all gambles f and g in the domain of $\underline{P}(\cdot|\mathscr{B})$ and all B in $\mathscr{B}$; see [10, Lemma 6.2.4]. It also follows from separate coherence that for any gamble f, there is a linear prevision $P(\cdot|B)$ that dominates $\underline{P}(\cdot|B)$ on its domain and satisfies $P(f|B) = \underline{P}(f|B)$ and $P(B|B) = 1$.

Now if f is a *$\mathscr{B}$-measurable* gamble, i.e., constant on the elements B of the partition $\mathscr{B}$, we can write $f = \sum_{B\in\mathscr{B}} f(B)I_B$, where $f(B)$ denotes the constant value of f on B. So we see that $fI_B = f(B)I_B$, and coherence of the lower prevision $\underline{P}(\cdot|B)$ requires for the constant gamble $f(B)$ that $\underline{P}(f(B)|B) = f(B)$. It then follows from separate coherence that if the $\mathscr{B}$-measurable gamble f belongs to the domain of $\underline{P}(\cdot|\mathscr{B})$, then $\underline{P}(f|B) = f(B)$ for all $B \in \mathscr{B}$, or equivalently, $\underline{P}(f|\mathscr{B}) = f$.[9] We may therefore always assume, without loss of generality, that the domain of a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ contains all $\mathscr{B}$-measurable gambles, because for such gambles f the value $\underline{P}(f|\mathscr{B}) = f$ is uniquely determined by separate coherence. We can go still further. Indeed, choose, for any B in $\mathscr{B}$, a gamble f_B in the domain $\mathscr{H}$, and consider the real-valued map $f = \sum_{B\in\mathscr{B}} I_B f_B$, which we shall assume to be bounded (a gamble). Then $fI_B = f_B I_B$, so separate coherence requires that if $f \in \mathscr{H}$, then $\underline{P}(f|B) = \underline{P}(f_B|B)$ for all B in $\mathscr{B}$. If f doesn't belong to $\mathscr{H}$, then this tells us that we can extend $\underline{P}(\cdot|\mathscr{B})$ *uniquely* to f, by separate coherence. We may therefore also always assume, without loss of generality, that the domain $\mathscr{H}$ of $\underline{P}(\cdot|\mathscr{B})$ is *$\mathscr{B}$-closed*, meaning that if $f_B \in \mathscr{H}$ for all $B \in \mathscr{B}$, then if $\sum_{B\in\mathscr{B}} f_B I_B$ is bounded, it belongs to $\mathscr{H}$ as well.

If we have a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$, defined on some domain $\mathscr{H}$, then for each $B \in \mathscr{B}$, we can consider the (unconditional) natural extension $\underline{E}(\cdot|B)$ of the coherent lower prevision $\underline{P}(\cdot|B)$ to all gambles, given by

$$\underline{E}(f|B) = \sup_{\substack{f_i\in\mathscr{H},\lambda_i\geq 0\\ i=1,\dots,m, m\geq 0}} \inf_{\omega\in\Omega} \left[f(\omega) - \sum_{i=1}^m \lambda_i \left[f_i(\omega) - \underline{P}(f_i|B) \right] \right]$$

[9] It follows from these comments that if we have two separately coherent conditional lower previsions $\underline{P}_1(\cdot|\mathscr{B}_1)$ and $\underline{P}_2(\cdot|\mathscr{B}_2)$, and a set B belonging to both partitions $\mathscr{B}_1$ and $\mathscr{B}_2$, then $\underline{P}_1(B|\mathscr{B}_1) = \underline{P}_2(B|\mathscr{B}_2)$ is equal to I_B, the gamble which takes the value 1 on B and 0 on B^c.

for all gambles f on Ω. Note that, as we did in Section 2.4, we can define the set of linear previsions

$$\mathscr{M}(\underline{P}(\cdot|B)) := \{P(\cdot|B) \in \mathbb{P}(\Omega)\colon (\forall f \in \mathscr{H})(P(f|B) \geq \underline{P}(f|B))\}.$$

Then $\underline{P}(\cdot|B)$ will be coherent if and only if it is the lower envelope of $\mathscr{M}(\underline{P}(\cdot|B))$, and the natural extension $\underline{E}(\cdot|B)$ will be the lower envelope of $\mathscr{M}(\underline{P}(\cdot|B))$. This leads to a conditional lower prevision $\underline{E}(\cdot|\mathscr{B})$ on $\mathscr{L}(\Omega)$, called the (unconditional) *natural extension* of $\underline{P}(\cdot|\mathscr{B})$, which clearly is separately coherent as well.

Finally, when the domain $\mathscr{H}$ of the conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ is a linear space that contains all I_B, $B \in \mathscr{B}$, then $\underline{P}(\cdot|\mathscr{B})$ is separately coherent if and only if ([10, Theorem 6.2.7])

1. $\underline{P}(f|B) \geq \inf_{\omega\in B} f(\omega)$;
2. $\underline{P}(\lambda f|\mathscr{B}) = \lambda \underline{P}(f|\mathscr{B})$;
3. $\underline{P}(f+g|\mathscr{B}) \geq \underline{P}(f|\mathscr{B}) + \underline{P}(g|\mathscr{B})$;

for all f and g in $\mathscr{H}$, $\lambda \geq 0$, and $B \in \mathscr{B}$. This should be compared with the characterisation of coherence near the end of Section 2.2.

Joint coherence. If besides the conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ we have other coherent conditional or unconditional lower previsions, we should require, besides separate coherence, that the assessments of all these (conditional) lower previsions should be consistent with one another. This leads to the requirement of *joint coherence*.[10] It is easier to formulate it for the case where we have a finite number of conditional lower previsions, $\underline{P}_1(\cdot|\mathscr{B}_1), \dots, \underline{P}_N(\cdot|\mathscr{B}_N)$, since unconditional previsions correspond to the particular case where $\mathscr{B} = \{\Omega\}$.

Definition 2. [10, Definition 7.1.4] Let $\underline{P}_1(\cdot|\mathscr{B}_1)$, ..., $\underline{P}_N(\cdot|\mathscr{B}_N)$ be separately coherent conditional lower previsions with respective linear domains $\mathscr{H}_1, \dots, \mathscr{H}_N \subseteq \mathscr{L}(\Omega)$.[11] They are called *jointly coherent* if for any $f_j \in \mathscr{H}_j$ where $j = 1, \dots, N$, and for any i in $\{1, \dots, N\}$, $f_0 \in \mathscr{H}_i$ and $B_0 \in \mathscr{B}_i$, there is some event B in $\{B_0\} \cup \bigcup_{j=1}^N S_j(f_j)$ such that

$$\sup_{\omega\in B}\left[\sum_{j=1}^N G(f_j|\mathscr{B}_j) - G(f_0|B_0)\right](\omega) \geq 0, \tag{9}$$

where the $\mathscr{B}_j$*-support* $S_j(f_j)$ of the gamble f_j is defined as the set of events

$$S_j(f_j) := \left\{B_j \in \mathscr{B}_j\colon I_{B_j} f_j \neq 0\right\}. \tag{10}$$

Similarly to the condition (5) for the coherence of an unconditional lower prevision, the condition (9) means that our subject's supremum acceptable buying price for a gamble f conditional on B_0 cannot be raised by considering the implications of the behavioural dispositions expressed through the other assessments.

It follows from the definition above that if $\underline{P}_1(\cdot|\mathscr{B}_1), \dots, \underline{P}_N(\cdot|\mathscr{B}_N)$ are jointly coherent, then $\underline{P}_i(\cdot|\mathscr{B}_i)$ is separately coherent for each $i = 1, \dots, N$. If in particular $\mathscr{B}_i = \{\Omega\}$ for some i (i.e., if we have an unconditional lower prevision), we deduce that $\underline{P}_i(\cdot|\mathscr{B}_i)$ is a coherent lower prevision. Note moreover (and this is one of the things that renders our

[10]This requirement is called simply '*coherence*' in [10]. We have preferred to use the terminology '*joint* coherence' in order to emphasise the distinction with '*separate* coherence'.

[11]We assume here that the domains are linear spaces only for the sake of simplicity. Nothing essential changes if we drop this assumption; we only need to replace the gambles $G(f_j|\mathscr{B}_j)$ in the condition (9) by finite non-negative linear combinations $\sum_{k=1}^{n_j} \lambda_j^k G(f_j^k|\mathscr{B}_j)$. For more details, see [9].

task in this paper difficult) that the notion of joint coherence is not transitive: given conditional lower previsions $\underline{P}_1(\cdot|\mathscr{B}_1)$, $\underline{P}_2(\cdot|\mathscr{B}_2)$ and $\underline{P}_3(\cdot|\mathscr{B}_3)$, the joint coherence of $\underline{P}_1(\cdot|\mathscr{B}_1)$ with $\underline{P}_2(\cdot|\mathscr{B}_2)$ and of $\underline{P}_2(\cdot|\mathscr{B}_2)$ with $\underline{P}_3(\cdot|\mathscr{B}_3)$ does not imply that $\underline{P}_1(\cdot|\mathscr{B}_1)$, $\underline{P}_2(\cdot|\mathscr{B}_2)$ and $\underline{P}_3(\cdot|\mathscr{B}_3)$ are jointly coherent.

In the particular case where we only have an unconditional coherent lower prevision $\underline{P}$ defined on $\mathscr{H}_1$ and a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ defined on $\mathscr{H}_2$,[12] the joint coherence condition (9) simplifies to ([10, Section 6.3.2])

$$\sup_{\omega\in\Omega}[G(f)+G(g|\mathscr{B})-G(h)](\omega)\geq 0$$
$$\sup_{\omega\in\Omega}[G(f)+G(g|\mathscr{B})-G(w|B)](\omega)\geq 0$$

for any $f,h\in\mathscr{H}_1$, $g,w\in\mathscr{H}_2$ and $B\in\mathscr{B}$, where similarly to Eq. (7), $G(f):=f-\underline{P}(f)$.

It is instructive to look at the special case that $\mathscr{H}_1$ includes $\mathscr{H}_2$, so $\underline{P}(f)$ is defined for all gambles f for which $\underline{P}(f|\mathscr{B})$ is defined. Then the coherence conditions above become

$$\underline{P}(G(f|\mathscr{B}))\geq 0 \tag{CP}$$
$$\underline{P}(G(f|B))=0, \tag{GBR}$$

for all f in $\mathscr{H}_1$ and B in $\mathscr{B}$. When both $\underline{P}$ and $\underline{P}(\cdot|\mathscr{B})$ are linear previsions P and $P(\cdot|\mathscr{B})$, these conditions turn into $P(f)=P(P(f|\mathscr{B}))$ and $P(f|B)P(B)=P(fI_B)$, respectively. The second condition is of course Bayes' Rule, which is why its counterpart (GBR) for lower previsions is called the *Generalised Bayes Rule*. This shows that, in the case of linear previsions, Bayes' Rule is necessary for coherence, but not sufficient in general, because the first condition, which is the *Conglomerative Property* (see the Introduction), also has to hold. In this respect, Walley's approach to coherence is even more demanding than de Finetti's [6] (or Williams' [11]), because de Finetti specifically does not require the Conglomerative Property to hold when the partition $\mathscr{B}$ is infinite.

It follows from the coherence of the lower prevision $\underline{P}$ that (CP) implies that

$$\underline{P}(\overline{P}(f|\mathscr{B}))\geq\underline{P}(f)\geq\underline{P}(\underline{P}(f|\mathscr{B})), \tag{11}$$

which is, of course, a necessary condition for joint coherence. When, in particular, $\underline{P}(\cdot|\mathscr{B})$ is a conditional *linear* prevision $P(\cdot|\mathscr{B})$, these inequalities turn into the equality $\underline{P}(f)=\underline{P}(P(f|\mathscr{B}))$.

2.6. **The Marginal Extension Theorem.** To complete this introduction to coherent lower previsions, we turn to the precise formulation of Walley's Marginal Extension Theorem, already mentioned in the Introduction.

Consider a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ defined on a set of gambles $\mathscr{H}\subseteq\mathscr{L}(\Omega)$, as well as a coherent unconditional lower prevision $\underline{P}$ defined on another set of gambles $\mathscr{K}\subseteq\mathscr{L}(\Omega)$. Assume in addition, *and this is crucial*, that $\mathscr{K}$ consists only of gambles that are *$\mathscr{B}$-measurable*, i.e., constant on the elements of the partition $\mathscr{B}$.

Now consider the (unconditional) natural extension $\underline{E}$ of $\underline{P}$ to the set of all $\mathscr{B}$-measurable gambles, and for each B in $\mathscr{B}$, denote by $\underline{M}(\cdot|B)$ the (unconditional) natural extension of the coherent lower prevision $\underline{P}(\cdot|B)$ to the set of all gambles $\mathscr{L}(\Omega)$. This leads to a new, separately coherent, conditional lower prevision $\underline{M}(\cdot|\mathscr{B})$.

Interestingly and surprisingly, when all elements of $\mathscr{K}$ are $\mathscr{B}$-measurable, requiring the *separate coherence* of $\underline{P}$ and $\underline{P}(\cdot|\mathscr{B})$ is enough to guarantee that they are also *jointly*

[12] There are some additional technical and essentially unrestrictive requirements on the domains $\mathscr{H}_1$ and $\mathscr{H}_2$; see [10, Section 6.3.1] for more details.

coherent. The (unconditional) natural extensions $\underline{E}$ and $\underline{M}(\cdot|\mathscr{B})$ also have a part in characterising their jointly coherent extensions, as the following theorem states. For a remarkably simple proof, we refer to [10, Section 6.7.2].

Theorem 1 (Marginal Extension Theorem)**.** *Suppose that (i) $\underline{P}$ is a coherent lower prevision on a domain $\mathscr{K}$ where (ii) all gambles in $\mathscr{K}$ are $\mathscr{B}$-measurable, and (iii) $\underline{P}(\cdot|\mathscr{B})$ is separately coherent conditional lower prevision on an arbitrary domain $\mathscr{H}$. Then*

1. *$\underline{P}$ and $\underline{P}(\cdot|\mathscr{B})$ are jointly coherent, and they have jointly coherent extensions to all of $\mathscr{L}(\Omega)$;*
2. *The point-wise smallest jointly coherent extensions of $\underline{P}$ and $\underline{P}(\cdot|\mathscr{B})$ to $\mathscr{L}(\Omega)$ are $\underline{M}$ and $\underline{M}(\cdot|\mathscr{B})$ where $\underline{M}$ is the lower prevision defined by*

$$\underline{M}(f) = \underline{E}(\underline{M}(f|\mathscr{B})).$$

The marginal extension $\underline{M}$ is not necessarily equal to the (unconditional) natural extension $\underline{E}$ of $\underline{P}$ alone, as it also has to take into account the behavioural consequences of the assessments that are present in $\underline{P}(\cdot|\mathscr{B})$! But since it follows from separate coherence that for any $\mathscr{B}$-measurable gamble f, $\underline{M}(f|\mathscr{B}) = f$, we see that $\underline{M}$ and $\underline{E}$ coincide at least on all $\mathscr{B}$-measurable gambles.

In general, $\underline{M}$ will not be the only extension of $\underline{P}$ that is jointly coherent with $\underline{P}(\cdot|\mathscr{B})$; but any other coherent extension will dominate $\underline{M}$ and will therefore represent behavioural dispositions not present in $\underline{P}$ and $\underline{P}(\cdot|\mathscr{B})$.

Finally, observe that for $\underline{M}$ and $\underline{M}(\cdot|\mathscr{B})$, the equality is reached in the second of the inequalities (11), because, as we have seen above, $\underline{M}$ and $\underline{E}$ coincide on $\mathscr{B}$-measurable gambles such as $\underline{M}(f|\mathscr{B})$. These inequalities also tell us that the extension $\underline{M}(f)$ is *uniquely jointly coherent* whenever $\underline{P}(f|\mathscr{B}) = \overline{P}(f|\mathscr{B})$ is precise and belongs to $\mathscr{K}$. In particular, when $P(\cdot|\mathscr{B})$ is precise and defined on all gambles, and when $\underline{P}$ is defined on all $\mathscr{B}$-measurable gambles, this tells us that $\underline{P}$ and $P(\cdot|\mathscr{B})$ have unique jointly coherent extensions $\underline{M} = \underline{P}(P(\cdot|\mathscr{B}))$ and $\underline{M}(\cdot|\mathscr{B}) = P(\cdot|\mathscr{B})$ to all gambles.

The idea behind the requirement of $\mathscr{B}$-measurability for the gambles in $\mathscr{K}$ is to have some sort of 'concatenation', or hierarchy, in the model. That is, we have some marginal information about the occurrence of the elements of the partition $\mathscr{B}$, in the form of a coherent lower prevision $\underline{P}$ defined only on $\mathscr{B}$-measurable gambles, and a lower prevision $\underline{P}(\cdot|\mathscr{B})$ conditional on $\mathscr{B}$. The marginal extension theorem allows us to combine these two lower previsions into a least-committal jointly coherent pair $\underline{M}$, $\underline{M}(\cdot|\mathscr{B})$.

It is perhaps easier to see this if we reformulate the Marginal Extension Theorem in terms of random variables, i.e., the way it is discussed in the Introduction. Consider two random variables X_1 and X_2 taking values in the respective sets $\mathscr{X}_1$ and $\mathscr{X}_2$. We may consider an unconditional (marginal) lower prevision $\underline{P}_1$ on $\mathscr{H}_1 \subseteq \mathscr{L}(\mathscr{X}_1)$, modelling our information about X_1 and a conditional lower prevision $\underline{P}_2(\cdot|X_1)$ on $\mathscr{H}_2 \subseteq \mathscr{L}(\mathscr{X}_2)$ modelling beliefs about the value of X_2 conditional on what value X_1 assumes.

If we let $\Omega = \mathscr{X}_1 \times \mathscr{X}_2$, then we can identify any gamble on $\mathscr{X}_1$ with a gamble on Ω that only depends on the first coordinate x_1 in $\omega = (x_1, x_2)$, i.e., which is $\mathscr{X}_1$-*variable*. If we consider the partition $\mathscr{B} = \{\{x_1\} \times \mathscr{X}_2 \colon x_1 \in \mathscr{X}_1\}$ of Ω, then we see that we can identify $\underline{P}_1$ with a lower prevision $\underline{P}$ on gambles on Ω that are $\mathscr{B}$-measurable, and its natural extension $\underline{E}$ to all $\mathscr{B}$-measurable gambles can be identified with the natural extension $\underline{E}_1$ of $\underline{P}_1$ to $\mathscr{L}(\mathscr{X}_1)$.

Similarly, we can associate the statement '$X_1 = x_1$' with the element $\{x_1\} \times \mathscr{X}_2$ of the partition $\mathscr{B}$. We can therefore identify the conditional lower prevision $\underline{P}_2(\cdot|X_1)$ with a conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ defined on a set of gambles on Ω, and the natural extension $\underline{M}(\cdot|\mathscr{B})$ can be identified with the natural extension $\underline{E}_2(\cdot|X_1)$ of $\underline{P}_2(\cdot|X_1)$. Indeed, it can be shown (see Lemma 6 further on) that

$$\underline{M}(h|\{x_1\}\times\mathscr{X}_2)=\underline{E}_2(h(x_1,\cdot)|x_1)$$

for all gambles h on Ω and all x_1 in $\mathscr{X}$, where $\underline{E}_2(\cdot|x_1)$ is the natural extension to $\mathscr{L}(\mathscr{X}_2)$ of the coherent lower prevision $\underline{P}_2(\cdot|x_1)$. So the Marginal Extension Theorem tells us, as we already essentially announced in the Introduction, that the point-wise smallest jointly coherent extensions of $\underline{P}_1$ and $\underline{P}_2(\cdot|X_1)$ to all gambles on $\mathscr{X}_1\times\mathscr{X}_2$ are given by $\underline{M}=\underline{E}_1(\underline{E}_2(\cdot|X_1))$ and $\underline{E}_2(\cdot|X_1)$. $\underline{M}$ models the 'information' about the value that the joint random variable (X_1,X_2) assumes in $\mathscr{X}_1\times\mathscr{X}_2$. $\underline{M}(h)=\underline{E}_1(\underline{E}_2(h|X_1))$ is the smallest (most conservative) supremum acceptable buying price for a gamble h on $\mathscr{X}_1\times\mathscr{X}_2$ that can be derived from $\underline{P}_1$ and $\underline{P}_2(\cdot|X_1)$ using (only) coherence.

In the rest of this paper, we shall generalise these results to a more general setting. But before we start doing that, it will be convenient to derive a number of (new) additional results about coherent conditioning.

2.7. **Further properties of coherent conditional lower previsions.** The first result deals with the notion of separate coherence, and tells us that, unsurprisingly, it leads to conditions very similar to the coherence condition (5) for unconditional previsions, but with a suitably restricted supremum.

Proposition 1. *Let $\mathscr{B}$ be a partition of Ω, and consider a conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ defined on a set of gambles $\mathscr{H}$ that is $\mathscr{B}$-closed and contains all $\mathscr{B}$-measurable gambles. Then $\underline{P}(\cdot|\mathscr{B})$ is separately coherent if and only if for all B in $\mathscr{B}$, all $n,m\geq 0$ and all f_0, f_1, ..., f_n in $\mathscr{H}$,*

$$\sup_{\omega\in B}\left[\sum_{k=1}^{n}[f_k(\omega)-\underline{P}(f_k|B)]-m[f_0(\omega)-\underline{P}(f_0|B)]\right]\geq 0. \tag{12}$$

Proof. First, assume that $\underline{P}(\cdot|\mathscr{B})$ is separately coherent. Fix natural numbers m and n, an element B of $\mathscr{B}$, as well as gambles f_0, f_1, ..., f_n in $\mathscr{H}$. Note that we may assume without loss of generality that $n>0$: when $n=0$, the separate coherence of $\underline{P}(\cdot|B)$ implies that $\underline{P}(f_0|B)\geq\inf_{\omega\in B}f_0(\omega)$, and Eq. (12) holds. Consider the gambles $g_0:=f_0I_B$ and $g_k:=f_kI_B-\mu I_{B^c}$ for $k=1,\dots,n$, where μ is an arbitrary real number, and B^c the set-theoretic complement of B. Since $\mathscr{H}$ is assumed to be $\mathscr{B}$-closed, all these gambles belong to $\mathscr{H}$. Since moreover $I_Bg_k=I_Bf_k$, it follows from (8) that $\underline{P}(f_k|B)=\underline{P}(g_k|B)$, for $k=0,1,\dots,n$. It also follows from the coherence of the lower prevision $\underline{P}(\cdot|B)$ that the maximum of

$$\left\{\sup_{\omega\in B}\left[\sum_{k=1}^{n}[f_k(\omega)-\underline{P}(f_k|B)]-m[f_0(\omega)-\underline{P}(f_0|B)]\right],-n\mu+m\underline{P}(f_0|B)-\sum_{k=1}^{n}\underline{P}(f_k|B)\right\},$$

being equal to

$$\sup_{\omega\in\Omega}\left[\sum_{k=1}^{n}[g_k(\omega)-\underline{P}(g_k|B)]-m[g_0(\omega)-\underline{P}(g_0|B)]\right],$$

is non-negative. Since this must hold for any μ, we find in particular that if we let

$$\mu>\frac{m\underline{P}(f_0|B)-\sum_{k=1}^{n}\underline{P}(f_k|B)}{n},$$

then the desired inequality indeed follows.

To prove the converse implication, consider any B in $\mathscr{B}$. Then we must show that $\underline{P}(\cdot|B)$ is a coherent lower prevision on $\mathscr{H}$, and that $\underline{P}(B|B)=1$ (recall that I_B is $\mathscr{B}$-measurable, and therefore belongs to $\mathscr{H}$.) It follows from the condition (12) that for all $n,m \geq 0$ and all $f_0, f_1, \dots, f_n$ in $\mathscr{H}$,

$$\sup_{\omega\in\Omega}\left[\sum_{k=1}^{n}[f_k(\omega)-\underline{P}(f_k|B)]-m[f_0(\omega)-\underline{P}(f_0|B)]\right] \geq \sup_{\omega\in B}\left[\sum_{k=1}^{n}[f_k(\omega)-\underline{P}(f_k|B)]-m[f_0(\omega)-\underline{P}(f_0|B)]\right] \geq 0,$$

so $\underline{P}(\cdot|B)$ is indeed coherent. To prove that $\underline{P}(B|B)=1$, let $n=1$, $f_0=f_1=I_B$ and m arbitrary, and apply (12) to find that $(1-m)\geq(1-m)\underline{P}(B|B)$. Choosing $m=0$ and $m=2$ leads to the desired equality. □

The second result is a generalisation of the coherence condition (CP) and (one of) the inequalities (11).

Proposition 2. *Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be two partitions of Ω, such that $\mathscr{B}_2$ is finer*[13] *than $\mathscr{B}_1$, and let $\underline{P}_1(\cdot|\mathscr{B}_1)$ and $\underline{P}_2(\cdot|\mathscr{B}_2)$ be two separately coherent conditional lower previsions defined on $\mathscr{L}(\Omega)$ that are also jointly coherent. Then for any gamble f on Ω, $\underline{P}_1(G(f|\mathscr{B}_2)|\mathscr{B}_1)\geq 0$ and $\underline{P}_1(f|\mathscr{B}_1)\geq\underline{P}_1(\underline{P}_2(f|\mathscr{B}_2)|\mathscr{B}_1)$.*

Proof. Consider f in $\mathscr{L}(\Omega)$ and B_1 in $\mathscr{B}_1$. Let $g=I_{B_1}f$ and $h=G(f|\mathscr{B}_2)$. Then the joint coherence of $\underline{P}_1(\cdot|\mathscr{B}_1)$ and $\underline{P}_2(\cdot|\mathscr{B}_2)$ implies in particular that there is some B in $\{B_1\}\cup S_2(g)$ such that

$$\begin{aligned}0 &\leq \sup_{\omega\in B}[G(0|\mathscr{B}_1)+G(g|\mathscr{B}_2)(\omega)-G(h|B_1)(\omega)]\\ &= \sup_{\omega\in B}[G(I_{B_1}f|\mathscr{B}_2)(\omega)-I_{B_1}(\omega)[G(f|\mathscr{B}_2)(\omega)-\underline{P}_1(G(f|\mathscr{B}_2)|B_1)]].\end{aligned}$$

Now we take into account that any element of $S_2(g)$ is included in B_1, since g is zero outside B_1. This means that the supremum over any B in $\{B_1\}\cup S_2(g)$ is dominated by the supremum over B_1, which leads to

$$\begin{aligned}0 &\leq \sup_{\omega\in B_1}[G(I_{B_1}f|\mathscr{B}_2)(\omega)-[G(f|\mathscr{B}_2)(\omega)-\underline{P}_1(G(f|\mathscr{B}_2)|B_1)]]\\ &= \underline{P}_1(G(f|\mathscr{B}_2)|B_1)+\sup_{\omega\in B_1}[G(I_{B_1}f|\mathscr{B}_2)(\omega)-G(f|\mathscr{B}_2)(\omega)]\\ &= \underline{P}_1(G(f|\mathscr{B}_2)|B_1)+\sup_{\omega\in B_1}\sum_{B_2\in\mathscr{B}_2}I_{B_2}(\omega)[I_{B_1}(\omega)f(\omega)-f(\omega)+\underline{P}_2(f|B_2)-\underline{P}_2(I_{B_1}f|B_2)]\\ &= \underline{P}_1(G(f|\mathscr{B}_2)|B_1)\\ &= \underline{P}_1(f-\underline{P}_2(f|\mathscr{B}_2)|B_1),\end{aligned}$$

where the third equality holds because either $B_2\subseteq B_1$ or $B_1\cap B_2=\emptyset$, and if $B_2\subseteq B_1$ then $I_{B_1}I_{B_2}=I_{B_2}$ and $I_{B_2}(I_{B_1}f)=I_{B_2}f$, so the separate coherence of $\underline{P}_2(\cdot|\mathscr{B}_2)$ ensures that $\underline{P}_2(I_{B_1}f|B_2)=\underline{P}_2(f|B_2)$; see (8). Since this holds for all B_1 in $\mathscr{B}_1$, we see that, indeed, $\underline{P}_1(G(f|\mathscr{B}_2)|\mathscr{B}_1)\geq 0$.

Now the coherence of the lower prevision $\underline{P}_1(\cdot|B_1)$ implies that

$$\underline{P}_1(f-\underline{P}(f|\mathscr{B}_2)|B_1)\leq\underline{P}_1(f|B_1)+\overline{P}_1(-\underline{P}_2(f|\mathscr{B}_2)|B_1)=\underline{P}_1(f|B_1)-\underline{P}_1(\underline{P}_2(f|\mathscr{B}_2)|B_1),$$

whence $\underline{P}_1(f|B_1)\geq\underline{P}_2(\underline{P}(f|\mathscr{B}_2)|B_1)$. Since this holds for all B_1 in $\mathscr{B}_1$, we indeed get $\underline{P}_1(f|\mathscr{B}_1)\geq\underline{P}_1(\underline{P}_2(f|\mathscr{B}_2)|\mathscr{B}_1)$. □

[13]This means that the elements of $\mathscr{B}_1$ are unions of elements of $\mathscr{B}_2$.

3. THE MARGINAL EXTENSION THEOREM

We now proceed to formulate and prove our generalised version of the Marginal Extension Theorem.

Let us consider $N > 0$ partitions $\mathscr{B}_1, \dots, \mathscr{B}_N$ of Ω that are increasingly finer, i.e., $\mathscr{B}_{i+1}$ is finer than $\mathscr{B}_i$ for $i = 1, \dots, N-1$. For each partition $\mathscr{B}_i$, we consider a separately coherent conditional lower prevision $\underline{P}_i(\cdot|\mathscr{B}_i)$, defined on a set of gambles $\mathscr{H}_i \subseteq \mathscr{L}(\Omega)$. We make the *crucial* additional assumption that all gambles in $\mathscr{H}_i$ are $\mathscr{B}_{i+1}$-measurable, i.e., constant on the elements of $\mathscr{B}_{i+1}$, for $i = 1, \dots, N-1$. Then $\underline{P}_i(\cdot|\mathscr{B}_i)$ can be regarded as 'marginal information' about the occurrence of the elements of $\mathscr{B}_{i+1}$.

In turns out that the separate coherence of these conditional lower previsions is enough to guarantee that they are also jointly coherent, and that they have jointly coherent extensions to all of $\mathscr{L}(\Omega)$. We characterise the smallest such extensions in the following theorem. Note that it extends Walley's Marginal Extension Theorem even in the case that $N = 2$, because we do not require that one of the lower previsions should be an unconditional one. Our proof (and in particular Lemma 3, which contains the crux of the argument) is inspired by ideas first expressed by De Cooman and Zaffalon in [4].

Theorem 2 (Marginal Extension Theorem; general version for partitions)**.** *Let $\underline{P}_1(\cdot|\mathscr{B}_1), \dots, \underline{P}_N(\cdot|\mathscr{B}_N)$ be separately coherent lower previsions with respective domains $\mathscr{H}_1, \dots, \mathscr{H}_N$. Assume that, for any $i = 2, \dots, N$, the partition $\mathscr{B}_i$ is finer than $\mathscr{B}_{i-1}$, and that moreover any gamble in $\mathscr{H}_{i-1}$ is $\mathscr{B}_i$-measurable. Then*

1. *$\underline{P}_1(\cdot|\mathscr{B}_1), \dots, \underline{P}_N(\cdot|\mathscr{B}_N)$ are jointly coherent and they have separately and jointly coherent extensions to all of $\mathscr{L}(\Omega)$;*
2. *the point-wise smallest separately and jointly coherent extensions are $\underline{M}_1(\cdot|\mathscr{B}_1), \dots, \underline{M}_N(\cdot|\mathscr{B}_N)$, where*
$$\underline{M}_i(\cdot|\mathscr{B}_i) = \underline{E}_i(\underline{E}_{i+1}(\dots(\underline{E}_N(\cdot|\mathscr{B}_N))\dots|\mathscr{B}_{i+1})|\mathscr{B}_i), \tag{13}$$
where for each B_N in $\mathscr{B}_N$, $\underline{E}_N(\cdot|B_N)$ is the (unconditional) natural extension of $\underline{P}_N(\cdot|B_N)$ to $\mathscr{L}(\Omega)$ and for each B_j in $\mathscr{B}_j$, $j = 1, \dots, N-1$, $\underline{E}_j(\cdot|B_j)$ is the (unconditional) natural extension of $\underline{P}_j(\cdot|B_j)$ to the set of all $\mathscr{B}_{j+1}$-measurable gambles.

Before proving the theorem, we wish to point out that the so-called *marginal extensions* $\underline{M}_i(\cdot|\mathscr{B}_i)$ can be obtained using the following 'backward' recursion formula
$$\underline{M}_i(\cdot|\mathscr{B}_i) = \underline{E}_i(\underline{M}_{i+1}(\cdot|\mathscr{B}_{i+1})|\mathscr{B}_i), \quad i = 1, \dots, N-1 \tag{14}$$
with 'initial condition'
$$\underline{M}_N(\cdot|\mathscr{B}_N) = \underline{E}_N(\cdot|\mathscr{B}_N). \tag{15}$$

Proof. For any $i = 1, \dots, N$, $\underline{E}_i(\cdot|\mathscr{B}_i)$ is the (unconditional) natural extension of the separately coherent lower prevision $\underline{P}_i(\cdot|\mathscr{B}_i)$, and is therefore also separately coherent. Then, we prove in Lemma 1 that the conditional lower previsions $\underline{M}_i(\cdot|\mathscr{B}_i)$ are separately coherent, and in Lemma 3 that they are jointly coherent. In Lemma 2 we show that $\underline{M}_i(\cdot|\mathscr{B}_i)$ extends $\underline{P}_i(\cdot|\mathscr{B}_i)$ to $\mathscr{L}(\Omega)$ for all $i = 1, \dots, N$, which shows that the latter indeed have separately and jointly coherent extensions to $\mathscr{L}(\Omega)$, and are therefore also jointly coherent. Finally, in Lemma 4 we show that the $\underline{M}_i(\cdot|\mathscr{B}_i)$ are the smallest jointly coherent extensions. □

Lemma 1. *The conditional lower previsions $\underline{M}_i(\cdot|\mathscr{B}_i)$ are separately coherent, for $i = 1, \dots, N$.*

Proof. We give a proof by induction. First of all, using Eq. (15), we see that $\underline{M}_N(\cdot|\mathscr{B}_N) = \underline{E}_N(\cdot|\mathscr{B}_N)$ is separately coherent. We complete the proof by showing that if $\underline{M}_{i+1}(\cdot|\mathscr{B}_{i+1})$ is separately coherent, then so is $\underline{M}_i(\cdot|\mathscr{B}_i)$, for $i = 1,\dots,N-1$.

We shall verify that $\underline{M}_i(\cdot|\mathscr{B}_i)$ satisfies the separate coherence axioms we mentioned in Section 2.5 for conditional lower previsions with linear domains. First of all, using Eq. (14), we get for any f in $\mathscr{L}(\Omega)$ and any B_i in $\mathscr{B}_i$,

$$\begin{aligned}\underline{M}_i(f|B_i) &= \underline{E}_i(\underline{M}_{i+1}(f|\mathscr{B}_{i+1})|B_i)\\ &\geq \inf_{\omega\in B_i}\underline{M}_{i+1}(f|\mathscr{B}_{i+1})(\omega) = \inf_{B_{i+1}\subseteq B_i}\underline{M}_{i+1}(f|B_{i+1})\\ &\geq \inf_{B_{i+1}\subseteq B_i}\inf_{\omega\in B_{i+1}} f(\omega) = \inf_{\omega\in B_i} f(\omega),\end{aligned}$$

where the first inequality follows from the separate coherence of $\underline{E}_i(\cdot|\mathscr{B}_i)$, the second inequality from the separate coherence of $\underline{M}_{i+1}(\cdot|\mathscr{B}_{i+1})$ [induction hypothesis], and where we have also used the fact that the partitions are increasingly finer.

Next, for any f in $\mathscr{L}(\Omega)$ and $\lambda \geq 0$, we have, using Eq. (14) and the separate coherence of $\underline{M}_{i+1}(\cdot|\mathscr{B}_{i+1})$ [induction hypothesis] and $\underline{E}_i(\cdot|\mathscr{B}_i)$,

$$\begin{aligned}\underline{M}_i(\lambda f|\mathscr{B}_i) &= \underline{E}_i(\underline{M}_{i+1}(\lambda f|\mathscr{B}_{i+1})|\mathscr{B}_i)\\ &= \underline{E}_i(\lambda\underline{M}_{i+1}(f|\mathscr{B}_{i+1})|\mathscr{B}_i)\\ &= \lambda\underline{E}_i(\underline{M}_{i+1}(f|\mathscr{B}_{i+1})|\mathscr{B}_i) = \lambda\underline{M}_i(f|\mathscr{B}_i).\end{aligned}$$

Finally, given any f and g in $\mathscr{L}(\Omega)$, we have, again using Eq. (14) and the separate coherence of $\underline{M}_{i+1}(\cdot|\mathscr{B}_{i+1})$ [induction hypothesis] and $\underline{E}_i(\cdot|\mathscr{B}_i)$,

$$\begin{aligned}\underline{M}_i(f+g|\mathscr{B}_i) &= \underline{E}_i(\underline{M}_{i+1}(f+g|\mathscr{B}_{i+1})|\mathscr{B}_i)\\ &\geq \underline{E}_i(\underline{M}_{i+1}(f|\mathscr{B}_{i+1})+\underline{M}_{i+1}(g|\mathscr{B}_{i+1})|\mathscr{B}_i)\\ &\geq \underline{E}_i(\underline{M}_{i+1}(f|\mathscr{B}_{i+1})|\mathscr{B}_i)+\underline{E}_i(\underline{M}_{i+1}(g|\mathscr{B}_{i+1})|\mathscr{B}_i)\\ &= \underline{M}_i(f|\mathscr{B}_i)+\underline{M}_i(g|\mathscr{B}_i).\end{aligned}$$ □

Lemma 2. *$\underline{M}_i(\cdot|\mathscr{B}_i)$ is an extension of $\underline{P}_i(\cdot|\mathscr{B}_i)$, for $i = 1,\dots,N$.*

Proof. For $f \in \mathscr{H}_N$ we have $\underline{M}_N(f|\mathscr{B}_N) = \underline{E}_N(f|\mathscr{B}_N) = \underline{P}_N(f|\mathscr{B}_N)$, because $\underline{P}_N(\cdot|\mathscr{B}_N)$ is separately coherent and therefore coincides with its natural extension on its domain. Next, consider $f \in \mathscr{H}_i$ for some $1 \leq i < N$. Then f is $\mathscr{B}_{i+1}$-measurable, and, since $\underline{M}_{i+1}(\cdot|\mathscr{B}_{i+1})$ is separately coherent, it follows [see Section 2.5] that $\underline{M}_{i+1}(f|\mathscr{B}_{i+1}) = f$. Hence, we get, using Eq. (14), that

$$\underline{M}_i(f|\mathscr{B}_i) = \underline{E}_i(\underline{M}_{i+1}(f|\mathscr{B}_{i+1})|\mathscr{B}_i) = \underline{E}_i(f|\mathscr{B}_i) = \underline{P}_i(f|\mathscr{B}_i),$$

where the last equality follows because $\underline{P}_i(\cdot|\mathscr{B}_i)$ is separately coherent and therefore coincides with its natural extension on its domain $\mathscr{H}_i$. □

Lemma 3. *The conditional lower previsions $\underline{M}_1(\cdot|\mathscr{B}_1)$, ..., $\underline{M}_N(\cdot|\mathscr{B}_N)$ are jointly coherent.*

Proof. Fix arbitrary f_0, f_1, ..., f_N in $\mathscr{L}(\Omega)$, i in $\{1,\dots,N\}$ and $B_i \in \mathscr{B}_i$. We must show that there is some event B in $\{B_i\}\cup\bigcup_{j=1}^N S_j(f_j)$ such that

$$\sup_{\omega\in B}\left[\sum_{j=1}^N G(f_j|\mathscr{B}_j) - G(f_0|B_i)\right](\omega) \geq 0. \tag{16}$$

Let us introduce the notations $g := \sum_{j=1}^{N} G(f_j|\mathscr{B}_j) - G(f_0|B_i)$, $h_j^\ell := \underline{M}_\ell(f_j|\mathscr{B}_\ell)$ for $\ell = j,\dots,N$ and $h_j^{N+1} := f_j$ for $j = 0,\dots,N$. Since $I_{B_i}[f_0 - \underline{M}_i(f_0|B_i)] = I_{B_i}[f_0 - \underline{M}_i(f_0|\mathscr{B}_i)]$, we then see that

$$
\begin{aligned}
g &= \sum_{j=1}^{N} \left[f_j - \underline{M}_j(f_j|\mathscr{B}_j)\right] - I_{B_i}\left[f_0 - \underline{M}_i(f_0|\mathscr{B}_i)\right] = \sum_{j=1}^{N} \left[h_j^{N+1} - h_j^j\right] - I_{B_i}\left[h_0^{N+1} - h_0^i\right] \\
&= \sum_{j=1}^{N} \sum_{\ell=j}^{N} \left[h_j^{\ell+1} - h_j^\ell\right] - I_{B_i} \sum_{\ell=i}^{N} \left[h_0^{\ell+1} - h_0^\ell\right] = \sum_{\ell=1}^{N} \sum_{j=1}^{\ell} \left[h_j^{\ell+1} - h_j^\ell\right] - I_{B_i} \sum_{\ell=i}^{N} \left[h_0^{\ell+1} - h_0^\ell\right].
\end{aligned}
$$

If we also define, for $\ell = 1,\dots,N$,

$$
g^\ell := \begin{cases} \sum_{j=1}^{\ell} [h_j^{\ell+1} - h_j^\ell] - I_{B_i}[h_0^{\ell+1} - h_0^\ell] & \text{if } \ell \geq i \\ \sum_{j=1}^{\ell} [h_j^{\ell+1} - h_j^\ell] & \text{otherwise,} \end{cases}
$$

then, clearly, $g = \sum_{\ell=1}^{N} g^\ell$. Also observe that for any $\ell = 1,\dots,N-1$, the gamble g_ℓ is $\mathscr{B}_{\ell+1}$-measurable. We first prove that for all $\ell = 1,\dots,N$ and all $C_\ell \in \mathscr{B}_\ell$,

$$
\sup_{\omega \in C_\ell} g^\ell(\omega) \geq 0. \tag{17}
$$

We shall distinguish between three possible cases. The first one is that $i \leq \ell < N$. Then we have for any $C_\ell \in \mathscr{B}_\ell$

$$
\begin{aligned}
&\sup_{\omega \in C_\ell} g^\ell(\omega) \\
&= \sup_{\omega \in C_\ell} \left[\sum_{j=1}^{\ell} [h_j^{\ell+1} - h_j^\ell] - I_{B_i}[h_0^{\ell+1} - h_0^\ell]\right](\omega) \\
&= \sup_{\omega \in C_\ell} \left[\sum_{j=1}^{\ell} [\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1}) - \underline{M}_\ell(f_j|\mathscr{B}_\ell)] - I_{B_i}[\underline{M}_{\ell+1}(f_0|\mathscr{B}_{\ell+1}) - \underline{M}_\ell(f_0|\mathscr{B}_\ell)]\right](\omega) \\
&= \sup_{\omega \in C_\ell} \sum_{j=1}^{\ell} [\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1})(\omega) - \underline{M}_\ell(f_j|C_\ell)] \\
&\qquad - I_{B_i}(\omega)[\underline{M}_{\ell+1}(f_0|\mathscr{B}_{\ell+1})(\omega) - \underline{M}_\ell(f_0|C_\ell)] \\
&= \sup_{\omega \in C_\ell} \sum_{j=1}^{\ell} [\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1})(\omega) - \underline{E}_\ell(\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1})|C_\ell)] \\
&\qquad - I_{B_i}(\omega)[\underline{M}_{\ell+1}(f_0|\mathscr{B}_{\ell+1})(\omega) - \underline{E}_\ell(\underline{M}_{\ell+1}(f_0|\mathscr{B}_{\ell+1})|C_\ell)] \geq 0,
\end{aligned}
$$

where the last inequality follows from the separate coherence of $\underline{E}_\ell(\cdot|\mathscr{B}_\ell)$, Proposition 1, and the fact that either $C_\ell \subseteq B_i$ or $C_\ell \cap B_i = \emptyset$.

The second case is that $i \leq \ell = N$. Then for any C_N in $\mathscr{B}_N$

$$
\sup_{\omega \in C_N} g^N(\omega) = \sup_{\omega \in C_N} \left[\sum_{j=1}^{N} [f_j(\omega) - \underline{E}_N(f_j|C_N)] - I_{B_i}(\omega)[f_0(\omega) - \underline{E}_N(f_0|C_N)]\right] \geq 0,
$$

again taking into account that $\underline{E}_N(\cdot|\mathscr{B}_N)$ is separately coherent, Proposition 1, and the fact that either $C_N \subseteq B_i$ or $C_N \cap B_i = \emptyset$.

The final case is that $\ell < i \leq N$. Then we get for any C_ℓ in $\mathscr{B}_\ell$ that

$$\begin{aligned}\sup_{\omega\in C_\ell} g^\ell(\omega) &= \sup_{\omega\in C_\ell}\sum_{j=1}^{\ell}[\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1})(\omega)-\underline{M}_\ell(f_j|C_\ell)]\\ &= \sup_{\omega\in C_\ell}\sum_{j=1}^{\ell}[\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1})(\omega)-\underline{E}_\ell(\underline{M}_{\ell+1}(f_j|\mathscr{B}_{\ell+1})|C_\ell)]\geq 0,\end{aligned}$$

where the last inequality follows yet again from the separate coherence of the lower prevision $\underline{E}_\ell(\cdot|\mathscr{B}_\ell)$ and Proposition 1. This proves that the inequality (17) indeed holds.

Next, we prove that

$$\sup_{\omega\in C_\ell}\sum_{k=\ell}^{N} g^k(\omega)\geq 0, \tag{18}$$

for $\ell = 1,\dots,N$ and for all $C_\ell \in \mathscr{B}_\ell$. We give a proof by induction. Recall that g^ℓ is $\mathscr{B}_{\ell+1}$-measurable for all $\ell = 1,\dots,N-1$, and denote by $g^\ell(C_{\ell+1})$ the constant value that g^ℓ attains on the element $C_{\ell+1}$ of the partition $\mathscr{B}_{\ell+1}$. It is obvious by applying the inequality (17) for $\ell = N$ that the desired inequality (18) holds for $\ell = N$. Assume now that the equality (18) holds for $\ell = n$, where $2 \leq n \leq N$ [this is the induction hypothesis]. Then for any C_{n-1} in $\mathscr{B}_{n-1}$,

$$\begin{aligned}\sup_{\omega\in C_{n-1}}\sum_{k=n-1}^{N} g^k(\omega) &= \sup_{C_n\in\mathscr{B}_n, C_n\subseteq C_{n-1}}\ \sup_{\omega\in C_n}\sum_{k=n-1}^{N} g^k(\omega)\\ &= \sup_{C_n\in\mathscr{B}_n, C_n\subseteq C_{n-1}}[g^{n-1}(C_n)+\sup_{\omega\in C_n}\sum_{k=n}^{N} g^k(\omega)]\\ &\geq \sup_{C_n\in\mathscr{B}_n, C_n\subseteq C_{n-1}} g^{n-1}(C_n) = \sup_{\omega\in C_{n-1}} g^{n-1}(\omega)\geq 0,\end{aligned}$$

where the first inequality follows from the induction hypothesis and the second by applying the inequality (17) for $\ell = n-1$. This proves that the desired inequality (18) also holds for $\ell = n-1$, and consequently it holds for all $\ell = 1,\dots,N$.

We are now ready to prove joint coherence. Let j be the smallest integer such that $f_j \neq 0$. Then, $h_j^\ell = 0$ for all $k = 1,\dots,j-1, \ell \geq k$, whence $g^k = 0$ for all $k < \min\{j,i\}$ and $g = \sum_{k=\min\{j,i\}}^{n} g^k$. If $j \leq i$, we consider $D_j \in S_j(f_j)$. If we invoke the inequality (18) for $\ell = j$ and $C_\ell = D_j$, we find that

$$\sup_{\omega\in D_j} g(\omega) = \sup_{\omega\in D_j}\sum_{k=j}^{N} g^k(\omega)\geq 0.$$

On the other hand, if $j > i$, we may again invoke the inequality (18) for $\ell = i$ and $C_\ell = B_i$ to find that

$$\sup_{\omega\in B_i} g(\omega) = \sup_{\omega\in B_i}\sum_{k=i}^{N} g^k(\omega)\geq 0.$$

In any of the two cases, there is some B in $\{B_i\}\cup\bigcup_{j=1}^{N} S_j(f_j)$ such that (16) holds, and we conclude that the conditional lower previsions $\underline{M}_1(\cdot|\mathscr{B}_1), \dots, \underline{M}_N(\cdot|\mathscr{B}_N)$ are indeed jointly coherent. □

Lemma 4. *$\underline{M}_1(\cdot|\mathscr{B}_1)$, …, $\underline{M}_N(\cdot|\mathscr{B}_N)$ are the smallest jointly coherent extensions of the lower previsions $\underline{P}_1(\cdot|\mathscr{B}_1)$, …, $\underline{P}_N(\cdot|\mathscr{B}_N)$.*

Proof. We give a proof by induction on N. We first show that the result holds for $N=2$. We deduce from Lemmas 1, 2 and 3 that the separately coherent conditional lower previsions $\underline{M}_1(\cdot|\mathscr{B}_1)$ and $\underline{M}_2(\cdot|\mathscr{B}_2)$ are jointly coherent extensions of $\underline{P}_1(\cdot|\mathscr{B}_1)$ and $\underline{P}_2(\cdot|\mathscr{B}_2)$ to $\mathscr{L}(\Omega)$, respectively. Consider two other separately and jointly coherent extensions $\underline{M}'_1(\cdot|\mathscr{B}_1)$ and $\underline{M}'_2(\cdot|\mathscr{B}_2)$, and any gamble f on Ω. Then taking into account that, by construction, $\underline{M}_2(f|\mathscr{B}_2)=\underline{E}_2(f|\mathscr{B}_2)$, it is clear that $\underline{M}'_2(f|\mathscr{B}_2)\geq\underline{M}_2(f|\mathscr{B}_2)$, because for each B_2 in $\mathscr{B}_2$, the coherent lower prevision $\underline{M}'_2(\cdot|B_2)$ is an extension to $\mathscr{L}(\Omega)$ of the coherent lower prevision $\underline{P}_2(\cdot|B_2)$ and it therefore dominates its (unconditional) natural extension $\underline{E}_2(\cdot|B_2)$ of $\underline{P}_2(\cdot|B_2)$ to all gambles. At the same time,

$$\begin{aligned}\underline{M}'_1(f|\mathscr{B}_1)\geq\underline{M}'_1(\underline{M}'_2(f|\mathscr{B}_2)|\mathscr{B}_1)\geq\underline{E}_1(\underline{M}'_2(f|\mathscr{B}_2)|\mathscr{B}_1)&\\ \geq\underline{E}_1(\underline{E}_2(f|\mathscr{B}_2)|\mathscr{B}_1)=\underline{M}_1(f|\mathscr{B}_1),&\end{aligned}$$

where the first inequality follows from Proposition 2, and the second inequality holds because $\underline{M}'_2(\cdot|\mathscr{B}_2)$ is $\mathscr{B}_2$-measurable and $\underline{M}'_1(\cdot|\mathscr{B}_1)$ is a coherent extension of $\underline{P}_1(\cdot|\mathscr{B}_1)$, which therefore dominates the smallest coherent extension $\underline{E}_1(\cdot|\mathscr{B}_1)$ of $\underline{P}_1(\cdot|\mathscr{B}_1)$ to $\mathscr{B}_2$-measurable gambles. Hence $\underline{M}_1(\cdot|\mathscr{B}_1)$ and $\underline{M}_2(\cdot|\mathscr{B}_2)$ are the smallest jointly coherent extensions.

We now prove the result for $N>2$. Let us introduce the notations

$$\underline{M}^n_i(\cdot|\mathscr{B}_i)=\underline{E}_i(\underline{E}_{i+1}(\dots(\underline{E}_n(\cdot|\mathscr{B}_n))\dots|\mathscr{B}_{i+1})|\mathscr{B}_i),$$

for $i=1,\dots,n$ and $n\geq1$. Then the induction hypothesis, namely that the result holds for $N=n-1$, amounts to stating that the separately coherent conditional lower previsions $\underline{M}^{n-1}_1(\cdot|\mathscr{B}_1)$, …, $\underline{M}^{n-1}_{n-1}(\cdot|\mathscr{B}_{n-1})$ are the smallest jointly coherent extensions of $\underline{P}_1(\cdot|\mathscr{B}_1)$, …, $\underline{P}_{n-1}(\cdot|\mathscr{B}_{n-1})$ to $\mathscr{L}(\Omega)$. We want to prove that the result holds for $N=n$. It is easy to see that $\underline{M}^n_i(f|\mathscr{B}_i)=\underline{M}^{n-1}_i(\underline{E}_n(f|\mathscr{B}_n)|\mathscr{B}_i)$ for $i=1,\dots,n-1$ and all f in $\mathscr{L}(\Omega)$. Now, given other jointly coherent extensions $\underline{M}'_1(\cdot|\mathscr{B}_1)$, …, $\underline{M}'_n(\cdot|\mathscr{B}_n)$ of $\underline{P}_1(\cdot|\mathscr{B}_1)$, …, $\underline{P}_n(\cdot|\mathscr{B}_n)$ to $\mathscr{L}(\Omega)$, we have by a similar course of reasoning as above that $\underline{M}'_n(f|\mathscr{B}_n)\geq\underline{E}_n(f|\mathscr{B}_n)=\underline{M}^n_n(f|\mathscr{B}_n)$ for any $f\in\mathscr{L}(\Omega)$, and moreover for any $i\in\{1,\dots,n-1\}$,

$$\underline{M}'_i(f|\mathscr{B}_i)\geq\underline{M}'_i(\underline{M}'_n(f|\mathscr{B}_n)|\mathscr{B}_i)\geq\underline{M}'_i(\underline{E}_n(f|\mathscr{B}_n)|\mathscr{B}_i),$$

where the first inequality follows from Proposition 2. Now, $\underline{M}'_1(\cdot|\mathscr{B}_1)$, …, $\underline{M}'_{n-1}(\cdot|\mathscr{B}_{n-1})$ are also jointly coherent extensions of $\underline{P}_1(\cdot|\mathscr{B}_1)$, …, $\underline{P}_{n-1}(\cdot|\mathscr{B}_{n-1})$ and they therefore dominate the smallest jointly coherent extensions, which by the induction hypothesis are $\underline{M}^{n-1}_1(\cdot|\mathscr{B}_1)$, …, $\underline{M}^{n-1}_{n-1}(\cdot|\mathscr{B}_{n-1})$. Hence,

$$\underline{M}'_i(\underline{E}_n(f|\mathscr{B}_n)|\mathscr{B}_i)\geq\underline{M}^{n-1}_i(\underline{E}_n(f|\mathscr{B}_n)|\mathscr{B}_i)=\underline{M}^n_i(f|\mathscr{B}_i).$$

This proves that the result holds for $N=n$. We conclude that $\underline{M}_1(\cdot|\mathscr{B}_1)$, …, $\underline{M}_N(\cdot|\mathscr{B}_N)$ are indeed the smallest jointly coherent extensions of $\underline{P}_1(\cdot|\mathscr{B}_1)$, …, $\underline{P}_N(\cdot|\mathscr{B}_N)$. □

As was the case with Theorem 1, our version of the Marginal Extension Theorem states that, if we want to combine all these conditional lower previsions and extend them in a coherent way to $\mathscr{L}(\Omega)$, we must use a two-step procedure: first, we must extend the marginal information $\underline{P}_{i-1}(\cdot|\mathscr{B}_{i-1})$ about each partition $\mathscr{B}_i$ to the set of all $\mathscr{B}_i$-measurable gambles, using (unconditional) natural extension; and secondly, we must concatenate this marginal information by means of Eq. (13).

4. THE MARGINAL EXTENSION THEOREM IN TERMS OF SETS OF LINEAR PREVISIONS

The marginal extensions can also be calculated as lower envelopes of jointly coherent conditional linear previsions, obtained by applying the Marginal Extension Theorem to the conditional linear previsions in the sets

$$\mathscr{M}(\underline{P}_k(\cdot|\mathscr{B}_k)) := \{P_k(\cdot|\mathscr{B}_k) \colon (\forall B_k \in \mathscr{B}_k)(P_k(\cdot|B_k) \in \mathscr{M}(\underline{P}(\cdot|B_k)))\}.$$

This is proven in the following theorem, which is a generalisation of Theorem 6.7.4 in [10].

Theorem 3 (Lower envelope theorem)**.** *Let $\underline{P}_1(\cdot|\mathscr{B}_1)$, ..., $\underline{P}_N(\cdot|\mathscr{B}_N)$ be separately coherent lower previsions with respective domains $\mathscr{H}_1$, ..., $\mathscr{H}_N$. Assume that, for any $i = 2,\dots,N$, the partition $\mathscr{B}_i$ is finer than $\mathscr{B}_{i-1}$, and that moreover any gamble in $\mathscr{H}_{i-1}$ is $\mathscr{B}_i$-measurable. For any $1 \leq k \leq N$ and any $B_k \in \mathscr{B}_k$, let $P_k(\cdot|B_k)$ be any element of $\mathscr{M}(\underline{P}_k(\cdot|B_k))$. Define, for any gamble f on Ω, $P_k(f|\mathscr{B}_k)$ as the $\mathscr{B}_k$-measurable gamble that assumes the value $P_k(f|B_k)$ on B_k. Let, for any gamble f on Ω,*

$$M_k(f|\mathscr{B}_k) = P_k(P_{k+1}(\dots(P_N(f|\mathscr{B}_N))\dots|\mathscr{B}_{k+1})|\mathscr{B}_k)$$

for $k = 1,\dots,N$. Then the marginal extensions $M_k(\cdot|\mathscr{B}_k)$ constructed in this way are jointly (and separately) coherent conditional (linear) previsions on $\mathscr{L}(\Omega)$. Moreover, $\underline{M}_k(\cdot|\mathscr{B}_k)$ is the lower envelope of all such conditional linear previsions $M_k(\cdot|\mathscr{B}_k)$, and for any gamble f on Ω there is such a conditional linear prevision that coincides on f with $\underline{M}_k(\cdot|\mathscr{B}_k)$.

Proof. It is easy to see that the $M_k(\cdot|\mathscr{B}_k)$ are (separately coherent) conditional linear previsions, and they are jointly coherent by the Marginal Extension Theorem (Theorem 2), so we concentrate on the rest of the proof. We shall prove the result for $\underline{M}_1(\cdot|\mathscr{B}_1)$, since the proof we give essentially contains the proofs for $\underline{M}_k(\cdot|\mathscr{B}_k)$ for any $k = 2,\dots,N$. Consider any such $M_1(\cdot|\mathscr{B}_1)$. Then, for any gamble f on Ω and any B_N in $\mathscr{B}_N$ we have that

$$P_N(f|B_N) \geq \underline{E}_N(f|B_N),$$

since by construction $P_N(\cdot|B_N)$ belongs to $\mathscr{M}(\underline{P}_N(\cdot|B_N))$. Therefore,

$$M_N(f|\mathscr{B}_N) = P_N(f|\mathscr{B}_N) \geq \underline{E}_N(f|\mathscr{B}_N) = \underline{M}_N(f|\mathscr{B}_N).$$

Consequently, for any B_{N-1} in $\mathscr{B}_{N-1}$ we get in a similar way that

$$P_{N-1}(P_N(f|\mathscr{B}_N)|B_{N-1}) \geq \underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|B_{N-1}),$$

whence

$$\begin{aligned} M_{N-1}(f|\mathscr{B}_{N-1}) &= P_{N-1}(P_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}) \\ &\geq \underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}) = \underline{M}_{N-1}(f|\mathscr{B}_{N-1}). \end{aligned}$$

If we continue this process, we eventually get to

$$\begin{aligned} M_1(f|\mathscr{B}_1) &= P_1(\dots(P_{N-1}(P_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}))\dots|\mathscr{B}_1) \\ &\geq \underline{E}_1(\dots(\underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}))\dots|\mathscr{B}_1) = \underline{M}_1(f|\mathscr{B}_1), \end{aligned}$$

also using Eq. (13). This proves that $M_1(\cdot|\mathscr{B}_1)$ dominates $\underline{M}_1(\cdot|\mathscr{B}_1)$.

To complete the proof, fix a gamble f on Ω. Then we know that for any B_N in $\mathscr{B}_N$, there is some $Q_{B_N}(\cdot|B_N)$ in $\mathscr{M}(\underline{P}_N(\cdot|B_N))$ such that [see the discussion of (unconditional) natural extension in Section 2.5]

$$Q_{B_N}(f|B_N) = \underline{E}_N(f|B_N).$$

This can be done for all B_N in $\mathscr{B}_N$ and we can use this to define a conditional linear prevision $Q_N(\cdot|\mathscr{B}_N)$ in $\mathscr{M}(\underline{P}_N(\cdot|\mathscr{B}_N))$ that satisfies, by construction,

$$Q_N(f|\mathscr{B}_N) = \underline{E}_N(f|\mathscr{B}_N). \tag{19}$$

Now $\underline{E}_N(f|\mathscr{B}_N)$ is a $\mathscr{B}_N$-measurable gamble, and we know that for any B_{N-1} in $\mathscr{B}_{N-1}$, there is some $Q_{B_{N-1}}(\cdot|B_{N_1})$ in $\mathscr{M}(\underline{P}_{N-1}(\cdot|B_{N-1}))$ such that

$$Q_{B_{N-1}}(\underline{E}_N(f|\mathscr{B}_N)|B_{N-1}) = \underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|B_{N-1}).$$

This can be done for all B_{N-1} in $\mathscr{B}_{N-1}$ and we can use this to define a conditional linear prevision $Q_{N-1}(\cdot|\mathscr{B}_{N-1})$ in $\mathscr{M}(\underline{P}_{N-1}(\cdot|\mathscr{B}_{N-1}))$ that satisfies, by construction,

$$Q_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}) = \underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}),$$

and using Eq. (19), this leads to

$$Q_{N-1}(Q_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}) = \underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}) = \underline{M}_{N-1}(f|\mathscr{B}_{N-1}).$$

If we follow this process, we eventually obtain a conditional linear prevision $Q_1(\cdot|\mathscr{B}_1)$ in $\mathscr{M}(\underline{P}_1(\cdot|\mathscr{B}_1))$ in the way described above, such that

$$\begin{aligned} Q_1(\dots(Q_{N-1}(Q_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}))\dots|\mathscr{B}_1) &= \underline{E}_1(\dots(\underline{E}_{N-1}(\underline{E}_N(f|\mathscr{B}_N)|\mathscr{B}_{N-1}))\dots|\mathscr{B}_1) \\ &= \underline{M}_1(f|\mathscr{B}_1). \end{aligned} \qquad \square$$

This theorem allows us to give our Marginal Extension Theorem a sensitivity analysis interpretation: we might assume the existence of precise (but unknown) conditional linear previsions $Q_k(\cdot|\mathscr{B}_k)$, and we may model the 'available information' about $Q_k(\cdot|\mathscr{B}_k)$ using the separately coherent conditional lower previsions $\underline{P}_k(\cdot|\mathscr{B}_k)$, or equivalently, by a set of candidate conditional linear previsions $\mathscr{M}(\underline{P}_k(\cdot|\mathscr{B}_k))$. Then the combination of these separate pieces of information should be done by selecting candidate conditional linear previsions $P_k(\cdot|\mathscr{B}_k)$ in these sets, and combining them using marginal extension (Bayes' rule). This leads to sets of jointly coherent marginal extensions, whose lower envelopes are precisely the marginal extensions of the conditional lower previsions $\underline{P}_k(\cdot|\mathscr{B}_k)$, as Theorem 3 guarantees.

It is an immediate consequence of Theorem 3 that in the particular case where we have linear marginals $P_k(\cdot|\mathscr{B}_k)$ defined on the classes of *all* $\mathscr{B}_{k+1}$-measurable gambles, $k = 1,\dots,N$, the marginal extensions are their *unique* coherent extensions to $\mathscr{L}(\Omega)$. This can also be seen using Theorem 2: the marginal extensions are simultaneously the smallest and the largest coherent extensions, and they are therefore unique. Note that in the general case of coherent lower previsions the marginal extensions are only the smallest coherent extensions, but there can be other coherent extensions of $\underline{P}_1(\cdot|\mathscr{B}_1),\dots,\underline{P}_k(\cdot|\mathscr{B}_k)$ to $\mathscr{L}(\Omega)$.

Finally, let us remark that Theorem 3 still holds if we replace in the sets $\mathscr{M}(\underline{P}_k(\cdot|B_k))$ by their sets of extreme points, because for any gamble f on $\mathscr{X}^k$ the value $\underline{P}_k(f|B_k)$ is attained on one of the extreme points of $\mathscr{M}(\underline{P}_k(\cdot|B_k))$. As a consequence, we deduce that the extreme points of $\underline{M}_1(\cdot|\mathscr{B}_1)$ are concatenations of extreme points of $\underline{P}_k(\cdot|\mathscr{B}_k)$ for $k = 1,\dots,N$, in the manner described in the theorem.

We also deduce from this theorem that the marginal extension theorem we have proven can be seen as a finite number of iterations of the marginal extension theorem for two partitions. To see this more clearly, assume for instance that $N = 3$. If we consider $\underline{M}_2^3(\cdot|\mathscr{B}_2)$ the marginal extension of the conditional lower previsions $\underline{P}_2(\cdot|\mathscr{B}_2)$ and $\underline{P}_3(\cdot|\mathscr{B}_3)$ and then the marginal extension of $\underline{P}_1(\cdot|\mathscr{B}_1)$ and $\underline{M}_2^3(\cdot|\mathscr{B}_2)$, we obtain again the marginal extension $\underline{M}_3(\cdot|\mathscr{B}_3)$ of $\underline{P}_1(\cdot|\mathscr{B}_1),\underline{P}_2(\cdot|\mathscr{B}_2)$ and $\underline{P}_3(\cdot|\mathscr{B}_3)$. It suffices to see that the linear previsions that dominate $\underline{M}_3(\cdot|\mathscr{B}_3)$ are precisely the combinations of the linear previsions that dominate $\underline{P}_1(\cdot|\mathscr{B}_1)$ and those that dominate $\underline{M}_2^3(\cdot|\mathscr{B}_2)$. In particular, when one of the lower previsions is unconditional, the marginal extension can be seen as a finite replication of Walley's marginal extension theorem.

5. The Marginal Extension Theorem in terms of random variables

The Marginal Extension Theorem we have just proven can perhaps be better understood if we consider a sequence of random variables instead of a sequence of increasingly finer partitions. Let $X_1, \dots, X_N$ be $N > 0$ random variables taking values in the respective non-empty sets $\mathscr{X}_1, \dots, \mathscr{X}_N$. We could interpret the index k of the random variable X_k as a 'time', in which case it seems natural to consider the case where we observe the values of $X_1, \dots, X_k$, and use these observations to infer something about the as yet unobserved random variables $X_{k+1}, \dots, X_N$. This is the general problem of predictive inference.

In order to be able to study this problem in more detail, let us introduce the following definition.

Definition 3. For any k in $\{1,\dots,N\}$, we define the product random variable

$$X^k := \prod_{i=1}^{k} X_i, \tag{20}$$

that assumes values in some subset of the product space

$$\mathscr{X}^k := \prod_{i=1}^{k} \mathscr{X}_i. \tag{21}$$

Now let us consider the special case that our subject models his beliefs about the value of the k-th random variable X_k conditional on the observation $(x_1,\dots,x_{k-1})$ of the previous $k-1$ variables $X_1,\dots,X_{k-1}$ in the form of a coherent lower prevision on some subset $\mathscr{H}_k$ of $\mathscr{L}(\mathscr{X}_k)$, which we denote as $\underline{P}_k(\cdot|x_1,\dots,x_{k-1})$. Suppose he does this for all $(x_1,\dots,x_{k-1})$ in $\mathscr{X}^{k-1}$ (and all $k = 1,\dots,N$).[14] Also suppose that the domain of $\underline{P}_k(\cdot|x_1,\dots,x_{k-1})$ is the same set $\mathscr{H}_k$ for all $(x_1,\dots,x_{k-1})$ in $\mathscr{X}^{k-1}$. We shall assume, in addition, that each domain $\mathscr{H}_k$ contains all constant gambles λ on $\mathscr{X}_k$, and that $\underline{P}_k(\lambda|x_1,\dots,x_{k-1}) = \lambda$ for all $\lambda \in \mathbb{R}$. We can always make such an assumption without loss of generality, because coherence requires that no other value than λ can be assigned to $\underline{P}_k(\lambda|x_1,\dots,x_{k-1})$, and assigning such a value does not affect the coherence in any way. This simple cosmetic trick will make life much easier for us further on, however.

We then construct the 'conditional lower prevision' $\underline{P}_k(\cdot|X^{k-1})$ as a two-place function that summarises the available assessments as follows: for any gamble f_k in $\mathscr{H}_k$, $\underline{P}_k(f_k|X^{k-1})$ is a gamble on $\mathscr{X}^{k-1}$ that assumes the value $\underline{P}_k(f_k|x_1,\dots,x_{k-1})$ in any element $(x_1,\dots,x_{k-1})$ of $\mathscr{X}^{k-1}$. We should be careful, however, in using the term 'conditional lower prevision' for $\underline{P}_k(f_k|X^{k-1})$, because so far, we have only defined conditional lower previsions with respect to *partitions*. We can, however, easily reinterpret $\underline{P}_k(f_k|X^{k-1})$ as a conditional lower prevision, as we next proceed to show. We first make suitable transformations on the domains. For this, we introduce the following definition:

Definition 4. Take $I \subseteq \{1,\dots,N\}$. Then, a gamble $f \in \mathscr{X}^N$ is called *$\mathscr{X}_I$-variable* when for every $x,y \in \mathscr{X}^N$ such that $x_i = y_i$ for all $i \in I$, we have that $f(x) = f(y)$.

[14] Actually, there is some abuse of notation here, as for $k = 1$, no observation has yet been made, and we denote the corresponding (unconditional) coherent lower prevision by $\underline{P}_1$.

Step 1: There is a one-to-one correspondence between the set of gambles on $\mathscr{X}^N$ which are $\mathscr{X}_I$-variable and the gambles on $\prod_{i\in I}\mathscr{X}_i$. In particular, a gamble f_k on $\mathscr{X}_k$ can be uniquely associated with a $\mathscr{X}_k$-variable gamble $\hat{f}_k$ on $\mathscr{X}^N$, given by $\hat{f}_k(x_1,\dots,x_N) = f_k(x_k)$ for all $(x_1,\dots,x_N)\in\mathscr{X}^N$. Let us define $\hat{\mathscr{H}}_k := \{\hat{f}_k : f_k \in \mathscr{H}_k\} \subseteq \mathscr{L}(\mathscr{X}^N)$ for all $k=1,\dots,N$.

Step 2: Next, we define suitable partitions for our purposes. For any $k=2,\dots,N$ and any $(x_1,\dots,x_{k-1})$, let us define the set $B_{(x_1,\dots,x_{k-1})} := \{(x_1,\dots,x_{k-1})\}\times\times_{\ell=k}^N\mathscr{X}_\ell$. Consider the partitions of $\mathscr{X}^N$ given by $\mathscr{B}_1 = \{\mathscr{X}^N\}$, $\mathscr{B}_k := \left\{B_z\colon z\in\mathscr{X}^{k-1}\right\}$, for $k=2,\dots,N$, and $\mathscr{B}_{N+1} := \left\{\{(x_1,\dots,x_N)\}\colon (x_1,\dots,x_N)\in\mathscr{X}^N\right\}$.

Step 3: We define a conditional lower prevision $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ on $\hat{\mathscr{H}}_k$ by

$$\underline{\hat{P}}_k(\hat{f}_k|B_{(x_1,\dots,x_{k-1})}) := \underline{P}_k(f_k|x_1,\dots,x_{k-1})$$

for all $f_k\in\mathscr{H}_k$ and $(x_1,\dots,x_{k-1})\in\mathscr{X}^{k-1}$. This lower prevision is only defined on some $\mathscr{X}_k$-variable gambles. But since we want $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ to be separately coherent, we are going to use some of the consequences of this property (see Section 2) to considerably enlarge its domain.

Step 4: Consider the set

$$\mathscr{H}^k := \left\{g\in\mathscr{L}(\mathscr{X}^k)\colon (\forall(x_1,\dots,x_{k-1})\in\mathscr{X}^{k-1})(g(x_1,\dots,x_{k-1},\cdot)\in\mathscr{H}_k)\right\}$$

of gambles on $\mathscr{X}^k$. Observe that we have defined $\mathscr{H}^k$ as a set of gambles on $\mathscr{X}^k$, but it can equally well (and actually should) be considered a set of gambles on $\mathscr{X}^N$ that are $\mathscr{X}^k$*-variable*. We shall henceforth leave all such trivial identifications implicit.

The set $\mathscr{H}^k$ is easily seen to be $\mathscr{B}_k$-closed. In fact, it is the smallest $\mathscr{B}_k$-closed set of gambles that contains all the $\hat{f}_k$ for $f_k\in\mathscr{H}_k$. It also contains all $\mathscr{B}_k$-measurable gambles, simply because we took care to assume from the outset that $\mathscr{H}_k$ contains all constant gambles. Now consider any g in $\mathscr{H}^k$, then we have that $g=\sum_{z\in\mathscr{X}^{k-1}}g(z,\cdot)I_{B_z}$, where $g(z,\cdot)\in\mathscr{H}_k$ for all $z\in\mathscr{X}^{k-1}$, and therefore separate coherence [see the discussion in Section 2] leaves us with no other choice but to let

$$\underline{\hat{P}}_k(g|B_z) = \underline{\hat{P}}_k(g(z,\cdot)|B_z) = \underline{P}_k(g(z,\cdot)|z), \tag{22}$$

for all z in $\mathscr{X}^{k-1}$. In particular, we see that $B_y\in\mathscr{H}^k$ and that $\underline{\hat{P}}_k(B_y|B_z)=\delta_{z,y}$ (Kronecker delta) for all $y,z\in\mathscr{X}^{k-1}$.

The construction above provides us with conditional lower previsions $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ with domains $\mathscr{H}^k$, for $k=1,\dots,N$. The following lemma shows that these previsions are separately coherent.

Lemma 5. *The conditional lower previsions $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ on $\mathscr{H}^k$ are separately coherent, for $k=1,\dots,N$.*

Proof. Since the domains $\mathscr{H}^k$ are $\mathscr{B}_k$-closed and contain all $\mathscr{B}_k$-measurable gambles, we may invoke Proposition 1. Consider therefore any $z\in\mathscr{X}^{k-1}$, any $m,n\geq 0$, and any gambles g_0, g_1, …, g_n in $\mathscr{H}^k$. For this fixed value z, consider the gambles $f_\ell = g_\ell(z,\cdot)$,

$\ell = 0,1,\dots,n$, which belong to $\mathscr{H}_k$, by construction. Then, using Eq. (22),

$$\sup_{x\in B_z}\left[\sum_{\ell=1}^{n}[g_\ell(x)-\underline{\hat{P}}_k(g_\ell|B_z)]-m[g_0(x)-\underline{\hat{P}}_k(g_0|B_z)]\right]$$
$$=\sup_{x_k\in\mathscr{X}_k}\left[\sum_{\ell=1}^{n}[g_\ell(z,x_k)-\underline{\hat{P}}_k(g_\ell(z,\cdot)|B_z)]-m[g_0(z,x_k)-\underline{\hat{P}}_k(g_0(z,\cdot)|B_z)]\right]$$
$$=\sup_{x_k\in\mathscr{X}_k}\left[\sum_{\ell=1}^{n}[f_\ell(x_k)-\underline{P}_k(f_\ell|z)]-m[f_0(x_k)-\underline{P}_k(f_0|z)]\right]\geq 0,$$

where the inequality follows from the coherence of the lower prevision $\underline{P}_k(\cdot|z)$ on $\mathscr{H}_k$. □

Hence, we are able to identify conditional lower previsions with respect to variables as conditional lower previsions with respect to partitions, and this identification allows us to impose and interpret requirements of joint coherence on the former.

Next, observe that all gambles in the domain $\mathscr{H}^k$ of the separately coherent conditional lower prevision $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ are not only $\mathscr{B}_k$- but also $\mathscr{B}_{k+1}$-measurable, i.e., $\mathscr{X}^k$-variable, for $k=1,\dots,N-1$. If we also remark that the partitions $\mathscr{B}_k$ are increasingly finer, we see that we are in a position to apply our Marginal Extension Theorem. We shall see that this theorem takes on a remarkably intuitive and simple form when written in terms of random variables, rather than partitions. But before we can appreciate its full power and elegance, we need to take one more (small) step, which is related to the identification of conditional lower previsions with respect to variables with their counterparts with respect to partitions.

Indeed, in order to apply the theorem, we need to find, for each $B_z\in\mathscr{B}_k$, i.e., for each $z\in\mathscr{X}^{k-1}$, the (unconditional) natural extension $\underline{\hat{E}}_k(\cdot|B_z)$ of the coherent lower prevision $\underline{\hat{P}}_k(\cdot|B_z)$ from its domain $\mathscr{H}^k$ to the set $\mathscr{L}(\mathscr{X}^k)$ (essentially) of all $\mathscr{B}_{k+1}$-measurable gambles, for $k=1,\dots,N$.[15] This leads to the separately coherent conditional lower prevision $\underline{\hat{E}}_k(\cdot|\mathscr{B}_k)$ defined on the set $\mathscr{L}(\mathscr{X}^k)$ of all $\mathscr{B}_{k+1}$-measurable gambles.

But let us consider, instead, for each $z=(z_1,\dots,z_{k-1})$ in $\mathscr{X}^{k-1}$, the natural extension $\underline{E}_k(\cdot|z)$ of the coherent lower prevision $\underline{P}_k(\cdot|z)$ from the subset $\mathscr{H}_k$ of $\mathscr{L}(\mathscr{X}_k)$ to all gambles on $\mathscr{X}_k$. This leads to the conditional lower prevision $\underline{E}_k(\cdot|X^{k-1})$ defined on $\mathscr{L}(\mathscr{X}_k)$. The following lemma tells us that there is an interesting, and perhaps at this point unsurprising, relationship between $\underline{\hat{E}}_k(\cdot|\mathscr{B}_k)$ and $\underline{E}_k(\cdot|X^{k-1})$, which extends in a natural way the relation (22) between $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ and $\underline{P}_k(\cdot|X^{k-1})$.

Lemma 6. *Let $k\in\{1,\dots,N\}$. Then for all $\mathscr{B}_{k+1}$-measurable gambles g on $\mathscr{X}^N$, or in other words for all g in $\mathscr{L}(\mathscr{X}^k)$, and for all z in $\mathscr{X}^{k-1}$, it holds that*

$$\underline{\hat{E}}_k(g|B_z)=\underline{E}_k(g(z,\cdot)|z).$$

Proof. First of all, it is easy to see that because $\underline{\hat{P}}_k(\cdot|\mathscr{B}_k)$ is separately coherent on $\mathscr{H}^k$, its (unconditional) natural extension $\underline{\hat{E}}_k(\cdot|\mathscr{B}_k)$ to $\mathscr{L}(\mathscr{X}^k)$ is separately coherent as well. As a result [see Eq. (8)], we find that

$$\underline{\hat{E}}_k(f|B_z)=\underline{\hat{E}}_k(f(z,\cdot)|B_z)$$

for any $\mathscr{B}_{k+1}$-measurable, i.e., $\mathscr{X}^k$-variable, gamble f and any z in $\mathscr{X}^{k-1}$. Let, then, g be any gamble on $\mathscr{X}_k$ and z any element of $\mathscr{X}^{k-1}$. Clearly, it now only remains to show that

$$\underline{\hat{E}}_k(g|B_z)=\underline{E}_k(g|z).$$

[15] The set of $\mathscr{B}_{N+1}$-measurable gambles is simply $\mathscr{L}(\mathscr{X}^N)$, and the reason why we introduced $\mathscr{B}_{N+1}$ as well, is precisely to be able to treat the case $k=N$ in one sweep with the other cases.

By definition of (unconditional) natural extension,

$$\begin{aligned}\underline{\hat{E}}_k(g|B_z) &= \sup_{\substack{g_i\in\mathscr{H}^k,\lambda_i\geq 0\\ i=1,\dots,m,m\geq 0}} \inf_{x\in\mathscr{X}^k}\left[g(x_k)-\sum_{i=1}^m \lambda_i\left[g_i(x)-\underline{\hat{P}}_k(g_i|B_z)\right]\right]\\ &= \sup_{\substack{g_i\in\mathscr{H}^k,\lambda_i\geq 0\\ i=1,\dots,m,m\geq 0}} \inf_{(y,x_k)\in\mathscr{X}^{k-1}\times\mathscr{X}_k}\left[g(x_k)-\sum_{i=1}^m \lambda_i\left[g_i(y,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right]\right].\end{aligned}$$

If we now recall that $B_z\in\mathscr{H}^k$, and that $\underline{\hat{P}}_k(B_z|B_z)=\underline{P}_k(\mathscr{X}_k|z)=1$, we see that the right-hand side can be rewritten as

$$\sup_{\substack{g_i\in\mathscr{H}^k,\lambda_i\geq 0\\ i=1,\dots,m,m\geq 0\\ \mu\geq 0}} \inf_{x_k\in\mathscr{X}_k}\inf_{y\in\mathscr{X}^{k-1}}\left[g(x_k)-\mu[I_{B_z}(y,x_k)-1]-\sum_{i=1}^m \lambda_i\left[g_i(y,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right]\right].$$

Now consider the last infimum in this expression. It is equal to

$$\min\Bigg\{g(x_k)-\sum_{i=1}^m\lambda_i\left[g_i(z,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right], \\ \mu+g(x_k)-\sup_{y\neq z}\sum_{i=1}^m\lambda_i\left[g_i(y,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right]\Bigg\}.$$

Now for any choice of the g_i and λ_i we can always choose $\mu\geq 0$ such that

$$\begin{aligned}\mu &\geq \sup_{y\neq z}\sum_{i=1}^m\lambda_i\left[g_i(y,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right]-\sum_{i=1}^m\lambda_i\left[g_i(z,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right]\\ &= \sup_{y\neq z}\sum_{i=1}^m\lambda_i\left[g_i(y,x_k)-g_i(z,x_k)\right]\end{aligned}$$

for all x_k in $\mathscr{X}_k$, and therefore we find that

$$\begin{aligned}\underline{\hat{E}}_k(g|B_z) &= \sup_{\substack{g_i\in\mathscr{H}^k,\lambda_i\geq 0\\ i=1,\dots,m,m\geq 0}} \inf_{x_k\in\mathscr{X}_k}\left[g(x_k)-\sum_{i=1}^m \lambda_i\left[g_i(z,x_k)-\underline{P}_k(g_i(z,\cdot)|z)\right]\right]\\ &= \sup_{\substack{f_i\in\mathscr{H}_k,\lambda_i\geq 0\\ i=1,\dots,m,m\geq 0}} \inf_{x_k\in\mathscr{X}_k}\left[g(x_k)-\sum_{i=1}^m \lambda_i\left[f_i(x_k)-\underline{P}_k(f_i|z)\right]\right]=\underline{E}_k(g|z).\end{aligned}$$ □

We are now ready to prove the following:

Theorem 4 (Marginal Extension Theorem for variables)**.** *Let us consider the separately coherent conditional lower previsions $\underline{P}_1$, $\underline{P}_2(\cdot|X^1)$, …, $\underline{P}_N(\cdot|X^{N-1})$, with respective domains $\mathscr{H}_k\subseteq\mathscr{L}(\mathscr{X}_k)$, $k=1,\dots,N$. Then*

1. *$\underline{P}_1$, $\underline{P}_2(\cdot|X^1)$, …, $\underline{P}_N(\cdot|X^{N-1})$ are jointly coherent and have separately and jointly coherent extensions to all of $\mathscr{L}(\mathscr{X}^N)$;*
2. *the point-wise smallest separately and jointly coherent extensions to $\mathscr{L}(\mathscr{X}^N)$ are $\underline{M}_1$, $\underline{M}_2(\cdot|X^1)$, …, $\underline{M}_N(\cdot|X^{N-1})$, where*

$$\underline{M}_1=\underline{E}_1(\underline{E}_2(\dots(\underline{E}_N(\cdot|X^{N-1}))\dots|X^1)),$$

and

$$\underline{M}_i(\cdot|X^{i-1})=\underline{E}_i(\underline{E}_{i+1}(\dots(\underline{E}_N(\cdot|X^{N-1}))\dots|X^i)|X^{i-1}),$$

for $i = 2, \dots, N$. In this expression, for each z in $\mathscr{X}^{k-1}$, $\underline{E}_k(\cdot|z)$ is the (unconditional) natural extension of the coherent lower prevision $\underline{P}_k(\cdot|z)$ to $\mathscr{L}(\mathscr{X}_k)$, which can be extended uniquely by separate coherence to all $\mathscr{X}^k$-variable gambles f by $\underline{E}_k(f|z) := \underline{E}_k(f(z,\cdot)|z)$, for $k = 1, \dots, N$.

Proof. The proof is simple now, with all the preparatory scaffolding in place. We know that partitions $\mathscr{B}_k$ are increasingly finer, that the conditional lower previsions $\hat{\underline{P}}_k(\cdot|\mathscr{B}_k)$ are separately coherent, and defined on domains $\mathscr{H}^k$ that are $\mathscr{B}_{k+1}$-measurable. So we can apply the Marginal Extension Theorem to find that they are jointly coherent and have coherent extensions to $\mathscr{L}(\mathscr{X}^N)$. This translates back to the first statement of the present theorem. But we may in addition infer that the pointwise smallest jointly and separately coherent extensions to $\mathscr{L}(\mathscr{X}^N)$ are actually given by $\hat{\underline{M}}_1(\cdot|\mathscr{B}_1)$, …, $\hat{\underline{M}}_N(\cdot|\mathscr{B}_N)$, where, with the notations established above,

$$\hat{\underline{M}}_k(\cdot|\mathscr{B}_k) = \hat{\underline{E}}_k(\hat{\underline{E}}_{k+1}(\dots(\hat{\underline{E}}_N(\cdot|\mathscr{B}_N))\dots|\mathscr{B}_{k+1})|\mathscr{B}_k).$$

Taking into account Lemma 6, this translates back to the second statement of the present theorem. □

The idea behind this theorem is the following: since the natural extension $\underline{E}_1$ of $\underline{P}_1$ represents marginal information about the value that X_1 assumes, we must concatenate it with $\underline{E}_2(\cdot|X^1)$ in order to have information about the value assumed by (X_1, X_2). Then we concatenate this with the natural extension $\underline{E}_3(\cdot|X^2)$ of $\underline{P}_3(\cdot|X^2)$, and, if we follow this process, we get to the lower prevision $\underline{M}$ on $\mathscr{L}(\mathscr{X}^N)$ which models the information about the value that X^N assumes. In particular, the conditional lower previsions $\underline{M}(\cdot|X^k)$ model the 'information' about the value that $(X_{k+1}, \dots, X_N)$ assumes *conditional* on the value taken by the first k variables in the process, X_1, …, X_k.

6. Marginal versus natural extension

It behoves us to compare the procedure of marginal extension we have studied in the previous sections to the notion of natural extension for conditional lower previsions developed by Walley in [10, Section 8.1].

Consider conditional lower previsions $\underline{P}_k(\cdot|\mathscr{B}_k)$ defined on *linear* spaces $\mathscr{H}_k \subseteq \mathscr{L}(\Omega)$ for $1 \leq k \leq N$, that are separately and jointly coherent. Then Walley ([10, Section 8.1.1]) defines their *natural extensions* $\underline{F}_1(\cdot|\mathscr{B}_1)$, …, $\underline{F}_N(\cdot|\mathscr{B}_N)$ to $\mathscr{L}(\Omega)$ in the following way: for each $f \in \mathscr{L}(\Omega)$ and each $B_0 \in \mathscr{B}_k$, $\underline{F}_k(f|B_0)$ [this is the value of $\underline{F}_k(f|\mathscr{B}_k)$ on B_0] is defined as the supremum value of α for which there are $f_i \in \mathscr{H}_i$, $i = 1, \dots, N$ such that

$$\sup_{\omega \in B} \left[\sum_{i=1}^{N} G(f_i|\mathscr{B}_i)(\omega) - I_{B_0}(\omega)[f(\omega) - \alpha] \right] < 0 \quad \text{for all } B \in \bigcup_{i=1}^{N} S_i(f_i) \cup \{B_0\}, \tag{23}$$

where the supports $S_i(f_i)$ are defined by Eq.(10). See [9] for an extension of this notion to the case of conditional lower previsions defined on domains that are not necessarily linear spaces.

In the particular case where we only have an unconditional lower prevision $\underline{P}$, and nothing else, this notion of natural extension $\underline{F}$ agrees with the one ($\underline{E}$) given by Eq. (6). But we should be very careful in more general situations. Indeed, in the previous sections, we have sometimes considered the (unconditional) natural extensions $\underline{E}_k(\cdot|B_k)$ of the lower previsions $\underline{P}_k(\cdot|B_k)$ for all B_k in $\mathscr{B}_k$, leading to the conditional lower prevision $\underline{E}_k(\cdot|\mathscr{B}_k)$. This conditional lower prevision, obtained through (unconditional) natural extension, will in general be different from, and will actually be dominated by, the (conditional) natural extension $\underline{F}_k(\cdot|\mathscr{B}_k)$. The reason is, of course, that in constructing $\underline{E}_k(\cdot|\mathscr{B}_k)$, we only consider the assessments present in the $\underline{P}_k(\cdot|\mathscr{B}_k)$, but not the ones incorporated in the other conditional lower previsions $\underline{P}_\ell(\cdot|\mathscr{B}_\ell)$ with $\ell \neq k$.

Also, the natural extensions given by Eq. (23) may not possess some of the properties of the unconditional natural extension: their most important drawback is that they may not be the point-wise smallest jointly coherent extensions. This may happen for instance because there simply are no jointly coherent extensions to all of $\mathscr{L}(\Omega)$ of the given jointly coherent conditional lower previsions. But even if there were, the natural extensions generally only provide a lower bound for the smallest jointly coherent extensions, and they are the point-wise smallest jointly coherent extensions if and only if they are jointly coherent themselves (see [10, Theorem 8.1.2 and Example 8.1.3]).

It is interesting, therefore, to study the properties of the (conditional) natural extensions in the specific situation considered in this paper, where we have separately coherent lower previsions $\underline{P}_k(\cdot|\mathscr{B}_k)$ conditional on a sequence of increasingly finer partitions $\mathscr{B}_k$, such that moreover $\underline{P}_k(\cdot|\mathscr{B}_k)$ is defined on a set of $\mathscr{B}_{k+1}$-measurable gambles for $k = 1,\dots,N-1$. We know from Theorem 2 that these conditional lower previsions are jointly coherent, and have jointly coherent extensions to all of $\mathscr{L}(\Omega)$. We have even characterised the point-wise smallest such extensions $\underline{M}_k(\cdot|\mathscr{B}_k)$. From our discussion above, the natural extensions $\underline{F}_k(\cdot|\mathscr{B}_k)$ provide only a lower bound for the smallest coherent extensions, and therefore $\underline{F}_k(\cdot|\mathscr{B}_k) \leq \underline{M}_k(\cdot|\mathscr{B}_k)$ for all $k = 1,\dots,N$. It remains to be seen whether these two extensions agree in general. Only in that case will these natural extensions be jointly coherent! The answer to this question is negative, as the following counterexample shows. It deals with a single partition, and it is therefore already relevant even in the context of Walley's simpler version of the Marginal Extension Theorem (Theorem 1, discussed in Section 2). We therefore use the notations established there.

Example 1. Let us consider the possibility space $\Omega = \mathscr{X}_1 \times \mathscr{X}_2$, where $\mathscr{X}_1 = \mathscr{X}_2 = [0,1]$, and the partition $\mathscr{B} = \{B_{x_1} : x_1 \in \mathscr{X}_1\}$, where $B_{x_1} := \{x_1\} \times \mathscr{X}_2$. Also consider the subsets $\mathscr{K} = \{\lambda\pi_1 : \lambda \in \mathbb{R}\}$ and $\mathscr{H} = \{g\pi_2 : g \in \mathscr{L}(\mathscr{X}_1)\}$ of $\mathscr{L}(\Omega)$, where the gamble $\lambda\pi_1$ is defined by $\lambda\pi_1(x_1,x_2) = \lambda x_1$, and the gamble $g\pi_2$ by $g\pi_2(x_1,x_2) = g(x_1)x_2$. Then it is easy to verify that $\mathscr{K}$ and $\mathscr{H}$ are linear subspaces of $\mathscr{L}(\Omega)$. Let us define the linear (and therefore coherent lower) prevision P on $\mathscr{K}$ by

$$P(\lambda\pi_1) = \lambda,$$

and the conditional linear prevision $P(\cdot|\mathscr{B})$ on $\mathscr{H}$ by

$$P(g\pi_2|B_{x_1}) = g(x_1)$$

for any x_1 in $\mathscr{X}_1$. Since $P(\cdot|B_{x_1})$ is a linear (and therefore coherent lower) prevision on $\mathscr{H}$ for all x_1, $P(\cdot|\mathscr{B})$ is separately coherent.[16] Moreover, any gamble in $\mathscr{K}$ is $\mathscr{B}$-measurable. Hence, we may apply Walley's Marginal Extension Theorem (Theorem 1) and conclude that the point-wise smallest jointly coherent extensions to $\mathscr{L}(\Omega)$ are $\underline{M}$ and $\underline{M}(\cdot|\mathscr{B})$, where for each x_1 in $\mathscr{X}_1$, $\underline{M}(\cdot|B_{x_1})$ is the (unconditional) natural extension of $P(\cdot|B_{x_1})$ to $\mathscr{L}(\Omega)$, $\underline{E}$ is the unconditional natural extension of P to the set of $\mathscr{B}$-measurable gambles, and $\underline{M}$ is defined on $\mathscr{L}(\Omega)$ by $\underline{M} = \underline{E}(\underline{M}(\cdot|\mathscr{B}))$.

[16] Since the B_{x_1} do not belong to $\mathscr{H}$, this statement may seem surprising. But it is easily verified that the natural extension $\underline{E}(\cdot|B_{x_1})$ of the coherent lower prevision $P(\cdot|B_{x_1})$ assumes the value one in B_{x_1}, so adding the assessments $P(B_{x_1}|B_{x_1}) = 1$ in no way affects the coherence and natural extensions of $P(\cdot|B_{x_1})$. In this sense, $P(\cdot|\mathscr{B})$ is indeed separately coherent.

We calculate the value of $\underline{M}$ in the gamble f on Ω, given by

$$f(x_1,x_2)=\begin{cases}0 & \text{if } (x_1,x_2)=(1-\frac{1}{n},1-\frac{1}{n}) \text{ for some } n>0\\ 1 & \text{otherwise.}\end{cases}$$

We first calculate $\underline{M}(f|B_{x_1})=\underline{M}(f(x_1,\cdot)|B_{x_1})=\underline{E}(f(x_1,\cdot)|B_{x_1})$ [the equalities follow from the fact that $\underline{M}(\cdot|\mathscr{B})$ is separately coherent, and Eq. (8)]. We get, using the formula for (unconditional) natural extension in Section 2,

$$\begin{aligned}\underline{E}(f(x_1,\cdot)|B_{x_1}) &= \sup_{h\in\mathscr{H}}\inf_{(y_1,y_2)\in\Omega}[f(x_1,y_2)-[h(y_1,y_2)-P(h|B_{x_1})]]\\ &= \sup_{g\in\mathscr{L}(\mathscr{X}_1)}\inf_{(y_1,y_2)\in\Omega}[f(x_1,y_2)-[g(y_1)y_2-g(x_1)]].\end{aligned}$$

First, let $x_1=1-\frac{1}{m}$ for some $m>0$, and consider the gamble g on $\mathscr{X}_1$ given by $g(y_1)=m$ for all y_1 in $\mathscr{X}_1$. Then

$$\begin{aligned}&\inf_{(y_1,y_2)\in\Omega}[f(x_1,y_2)-[g(y_1)y_2-g(x_1)]]\\ &\qquad=\min\left\{0-[m(1-\frac{1}{m})-m],\inf_{y_2\neq 1-1/m}[1-[m(y_2-1)]\right\}=1.\end{aligned}$$

This tells us that $\underline{E}(f(x_1,\cdot)|B_{x_1})\geq 1$, and since $f\leq 1$, it follows from the coherence of the lower prevision $\underline{E}(\cdot|B_{x_1})$ that $\underline{E}(f(x_1,\cdot)|B_{x_1})\leq 1$. So we may conclude that $\underline{M}(f|B_{x_1})=1$ when $x_1=1-\frac{1}{m}$ for some $m>0$. Now, for any x_1 in $\mathscr{X}_1$ that differs from $1-\frac{1}{m}$ for all $m>0$, the gamble $f(x_1,\cdot)$ on $\mathscr{X}_2$ is identically 1, and the coherence of the lower prevision $\underline{E}(\cdot|B_{x_1})$ then implies that $\underline{E}(f(x_1,\cdot)|B_{x_1})=1$. This implies that the gamble $\underline{M}(f|\mathscr{B})$ is identically 1, whence $\underline{M}(f)=\underline{E}(\underline{M}(f|\mathscr{B}))=\underline{E}(1)=1$, using the coherence of the (unconditional) natural extension $\underline{E}$ of P.

Let us now study the natural extension $\underline{F}$ of the pair P, $P(\cdot|\mathscr{B})$ to a coherent lower prevision on $\mathscr{L}(\Omega)$, and in particular its value $\underline{F}(f)$ in the gamble f. We use Eq. (23), with $N=2$ and $B_0=\Omega$, $\mathscr{B}_1=\{\Omega\}$, $\mathscr{B}_2=\mathscr{B}$, $G(f_1|\mathscr{B}_1)=f_1-P(f_1)=\lambda\pi_1-\lambda$ for all $f_1=\lambda\pi_1$ in $\mathscr{H}_1=\mathscr{K}$, and $G(f_2|\mathscr{B}_2)=g\pi_2-g$ for all $f_2=g\pi_2$ in $\mathscr{H}_2=\mathscr{H}$. Since $B_0=\Omega$, we see that the supremum over all B in $\bigcup_{i=1}^{2}S_i(f_i)\cup\{B_0\}$ in Eq. (23) will be negative if and only if the supremum over Ω is negative, so we get

$$\begin{aligned}\underline{F}(f)&=\sup\{\alpha\colon(\exists f_i\in\mathscr{H}_i,i=1,2)(\sup[G(f_1|\mathscr{B}_1)+G(f_2|\mathscr{B}_2)-(f-\alpha)]<0)\}\\ &=\sup_{f_i\in\mathscr{H}_i}\inf[f-[G(f_1|\mathscr{B}_1)+G(f_2|\mathscr{B}_2)]]\\ &=\sup_{\lambda\in\mathbb{R},g\in\mathscr{L}(\mathscr{X}_1)}\inf_{(y_1,y_2)\in\Omega}[f(y_1,y_2)-[\lambda(y_1-1)+g(y_1)(y_2-1)]]\\ &\leq\sup_{\lambda\in\mathbb{R},g\in\mathscr{L}(\mathscr{X}_1)}\inf_{n\in\mathbb{N}}\left[0-\left[-\frac{\lambda}{n}-\frac{g(1-\frac{1}{n})}{n}\right]\right]\\ &=\sup_{\lambda\in\mathbb{R},g\in\mathscr{L}(\mathscr{X}_1)}\inf_{n\in\mathbb{N}}\frac{\lambda+g(1-\frac{1}{n})}{n}=0,\end{aligned}$$

where the last equality holds because all g are bounded. Hence, $\underline{F}(f)\leq 0$, and since $f\geq 0$ and $\underline{F}$ is a coherent lower prevision, we also have that $\underline{F}(f)\geq 0$, whence $\underline{F}(f)=0<1=\underline{M}(f)$: the natural and the marginal extensions do not coincide on the gamble f. ♦

This example shows that, in the situations considered in this paper, the procedure of natural extension may fail to lead to jointly coherent (conditional) lower previsions, and the marginal extensions must be considered instead. Interestingly, Walley proves in [10, Theorem 8.1.8.] that, when one of the lower previsions $\underline{P}_1(\cdot|\mathscr{B}_1) = \underline{P}_1$ is an unconditional one, as is the case in our counterexample, and in the context we discussed for marginal extension for variables, the natural extension $\underline{E}_1$ is the point-wise smallest extension of $\underline{P}_1$ to $\mathscr{L}(\Omega)$ that is coherent with the original coherent lower previsions $\underline{P}_1$, $\underline{P}_2(\cdot|\mathscr{B}_2)$, ..., $\underline{P}_N(\cdot|\mathscr{B}_N)$. However, as we deduce from the previous counterexample, *this does not imply that* $\underline{E}_1$ *is still jointly coherent with the extensions* $\underline{E}_2(\cdot|\mathscr{B}_2)$*, ...,* $\underline{E}_N(\cdot|\mathscr{B}_N)$*!* Indeed, it can be checked that in Example 1 the natural extensions $\underline{E}$ and $\underline{E}(\cdot|\mathscr{B})$ are not jointly coherent. Therefore, if we really want to exploit the behavioural dispositions present in these (conditional) lower previsions to extend *all of them* to the set of all gambles, we must consider their marginal extensions instead of their natural extensions.

It would be interesting to study under which conditions marginal extension and natural extension agree on $\mathscr{L}(\Omega)$; from our discussion above, we see that it would suffice to check when the natural extensions defined through Eq. (23) are jointly coherent. It is claimed in [10, Theorem 8.1.9] that this is the case as soon as all the partitions $\mathscr{B}_k$, $k = 1,\dots,N$ are finite (and, in the case of variables, when these take values in finite sets). Although Walley's general theory of natural extension assumes the linearity of the domains $\mathscr{H}_1,\dots,\mathscr{H}_N$, nothing essential changes if we consider conditional lower previsions defined on arbitrary sets of gambles; the only difference would be that we would replace each $G(f_i|\mathscr{B}_i)$ in Eq. (23) by a finite linear combination $\sum_{k=1}^{n_i} \lambda_{ik} G(f_i^k|\mathscr{B}_i)$, with $\lambda_{ik} \geq 0$ and $f_i^k \in \mathscr{H}_i$, for all $k = 1,\dots,n_i$ where $i = 1,\dots,N$ ([9]). See also [3] for some additional comments on the relationship between the marginal and natural extensions in the case of variables.

In the case of a conditional lower prevision $\underline{P}(\cdot|\mathscr{B})$ and an unconditional lower prevision $\underline{P}$ with respective domains $\mathscr{H}$ and $\mathscr{K}$, and where all the gambles in $\mathscr{K}$ are $\mathscr{B}$-measurable (as in our example), we deduce from the comments above that the marginal and natural extensions coincide when $\mathscr{H} = \mathscr{L}(\Omega)$. If moreover $\underline{P}(\cdot|\mathscr{B})$ is a linear prevision, then they are the only extensions that preserve joint coherence.

7. The forward irrelevant product

Let us apply, as a final step in this development, the Marginal Extension Theorem for variables to a particular case. Consider, as before, $N > 0$ random variables X_1, ..., X_N taking values in the respective non-empty sets $\mathscr{X}_1, \dots, \mathscr{X}_N$. For each variable X_k, a subject has information about the value it assumes in $\mathscr{X}_k$, which he models in terms of a (marginal) coherent lower prevision $\underline{P}_k$ defined on a set of gambles $\mathscr{H}_k \subseteq \mathscr{L}(\mathscr{X}_k)$. As we argued before, we can assume without loss of generality that all these domains include the constant gambles.[17] Let X^k and $\mathscr{X}^k$ denote the product variables and spaces as in Eqs. (20) and (21).

Our subject now assesses that his beliefs about the value that the random variable X_k assumes in $\mathscr{X}_k$ will not change after observing the values of the 'previous' variables X_1, ..., X_{k-1}. This is an assessment of so-called *forward epistemic irrelevance*, and it can be expressed by means of conditional lower previsions $\underline{P}_k(\cdot|X^{k-1})$, $k = 2,\dots,N$, where

$$\underline{P}_k(f|x_1,\dots,x_{k-1}) = \underline{P}_k(f) \tag{24}$$

[17] We shall also assume here that the domains $\mathscr{H}_k$ of the marginal lower previsions $\underline{P}_k$ are *cones*, i.e., closed under multiplication with non-negative real numbers. This will considerably simplify the proof of Proposition 5 further on. We can make this assumption without loss of generality, because by coherence, if $\underline{P}(f)$ is given, then the uniquely coherent value for $\underline{P}(\lambda f)$ is $\lambda \underline{P}(f)$ for all real $\lambda \geq 0$. As a consequence, the domains $\mathscr{H}^k$ considered further on in this section are cones as well.

for any $(x_1,\dots,x_{k-1})$ in $\mathscr{X}^{k-1}$ and all $f \in \mathscr{H}_k$.

In summary, we have the following assessments: a unconditional (marginal) lower prevision $\underline{P}_1$ defined on $\mathscr{H}_1$, and conditional lower previsions $\underline{P}_k(\cdot|X^{k-1})$ defined on $\mathscr{H}_k$, which are derived from the marginals $\underline{P}_k$ and the forward epistemic irrelevance assessment (24), for $2 \le k \le N$.

So we see that we have landed squarely in the domain where our Marginal Extension Theorem for variables (Theorem 4) can be applied to conclude that the point-wise smallest jointly coherent extensions of $\underline{P}_1$, $\underline{P}_2(\cdot|X^1)$, ..., $\underline{P}_N(\cdot|X^{N-1})$ to $\mathscr{L}(\mathscr{X}^N)$ are given by $\underline{M}_1$, $\underline{M}_2(\cdot|X^1)$, ..., $\underline{M}_N(\cdot|X^{N-1})$, where

$$\underline{M}_1 = \underline{E}_1(\underline{E}_2(\dots(\underline{E}_N(\cdot|X^{N-1}))\dots|X^1)),$$

and

$$\underline{M}_i(\cdot|X^{i-1}) = \underline{E}_i(\underline{E}_{i+1}(\dots(\underline{E}_N(\cdot|X^{N-1}))\dots|X^i)|X^{i-1}),$$

for $i = 2,\dots,N$, using the notations established in Theorem 4. Let us take a close look at these expressions, using what we have learned in Section 5. Let f be any gamble on $\mathscr{X}^N$. We may then apply Lemma 6 to find that for any $(x_1,\dots,x_{N-1})$ in $\mathscr{X}^{N-1}$,

$$\underline{M}_N(f|x_1,\dots,x_{N-1}) = \underline{E}_N(f|x_1,\dots,x_{N-1}) = \underline{E}_N(f(x_1,\dots,x_{N-1},\cdot))$$

where $\underline{E}_N$ is the (unconditional) natural extension of the marginal $\underline{P}_N$ to all gambles on $\mathscr{X}_N$. It will be convenient to let $\underline{E}_N(f)$ denote the gamble on $\mathscr{X}^{N-1}$ that assumes the value $\underline{E}_N(f(x_1,\dots,x_{N-1},\cdot)) = \underline{E}_N(f|x_1,\dots,x_{N-1})$ in the element $(x_1,\dots,x_{N-1})$ of $\mathscr{X}^{N-1}$. More generally, if h is a gamble on $\mathscr{X}^k$, we shall denote by $\underline{E}_k(h)$ the gamble on $\mathscr{X}^{k-1}$ that assumes the value $\underline{E}_k(f(x_1,\dots,x_{k-1},\cdot)) = \underline{E}_k(f|x_1,\dots,x_{k-1})$ in the element $(x_1,\dots,x_{k-1})$ of $\mathscr{X}^{k-1}$, where $\underline{E}_k$ is the (unconditional) natural extension of the marginal $\underline{P}_k$ to all gambles on $\mathscr{X}_k$. If we now apply Lemma 6 again for $k = N-1$ to the gamble $h = \underline{E}_N(f)$ on $\mathscr{X}^{N-1}$, we find that for any $(x_1,\dots,x_{N-2})$ in $\mathscr{X}^{N-2}$,

$$\begin{aligned}\underline{M}_{N-1}(f|x_1,\dots,x_{N-2}) &= \underline{E}_{N-1}(h|x_1,\dots,x_{N-2}) = \underline{E}_{N-1}(h)(x_1,\dots,x_{N-2})\\ &= \underline{E}_{N-1}(h(x_1,\dots,x_{N-2},\cdot)) = \underline{E}_{N-1}(\underline{E}_N(f)(x_1,\dots,x_{N-2},\cdot))\\ &= \underline{E}_{N-1}(\underline{E}_N(f(x_1,\dots,x_{N-2},\cdot,\cdot)),\end{aligned}$$

where $\underline{E}_{N-1}$ is the (unconditional) natural extension of the marginal $\underline{P}_{N-1}$ to all gambles on $\mathscr{X}_{N-1}$. Again, we denote by $\underline{E}_{N-1}(\underline{E}_N(f))$ the gamble on $\mathscr{X}^{N-2}$ that assumes the value $\underline{E}_{N-1}(\underline{E}_N(f(x_1,\dots,x_{N-2},\cdot,\cdot))$ for any $(x_1,\dots,x_{N-2})$ in $\mathscr{X}^{N-2}$. The pattern that is emerging should by now be quite clear:

$$\underline{M}_k(f|x_1,\dots,x_{k-1}) = \underline{E}_k(\underline{E}_{k+1}(\dots(\underline{E}_N(f))\dots))$$

and in particular we find for the unconditional (joint) lower prevision $\underline{M}_1$ on $\mathscr{X}^N$

$$\underline{M}_1(f) = \underline{E}_1(\underline{E}_2(\dots(\underline{E}_{k+1}(\dots(\underline{E}_N(f))\dots)\dots))).$$

$\underline{M}_1$ is called the *forward irrelevant product* of the marginal lower previsions $\underline{P}_1$, ..., $\underline{P}_N$. It is obtained by first using $\underline{E}_N$ to 'integrate out' the last variable, then using $\underline{E}_{N-1}$ to 'integrate out' the next but last variable, ..., and finally using $\underline{E}_1$ to 'integrate out' the only remaining, first, variable.

Let us denote by $\underline{M}^k$ the forward irrelevant product of the first k marginals $\underline{P}_1$, ..., $\underline{P}_k$. It is easy to derive the following interesting recursion formula from the discussion above: for all gambles f on $\mathscr{X}^k$:

$$\underline{M}^k(f) = \underline{M}^{k-1}(\underline{E}_k(f)). \tag{25}$$

The lower previsions $\underline{M}_1$, $\underline{M}_k(\cdot|X^{k-1})$, $k = 2,\dots,N$ are the point-wise smallest jointly coherent extensions of $\underline{P}_1$, $\underline{P}_k(\cdot|X^{k-1})$ to the set of all gambles on $\mathscr{X}^N$, or in other words, the point-wise smallest jointly coherent extensions of the marginals $\underline{P}_1, \dots, \underline{P}_N$ together with the forward epistemic irrelevance assessment (24).

If we are only interested in extending $\underline{P}_1$ and preserving joint coherence with $\underline{P}_2(\cdot|X^1)$, $\dots$, $\underline{P}_N(\cdot|X^{N-1})$, without extending the latter, then the natural extension $\underline{F}$ can be used. This natural extension will be called in this context the *forward irrelevant natural extension* of the marginals $\underline{P}_1, \dots, \underline{P}_N$. It is dominated by $\underline{M}_1$ on $\mathscr{L}(\mathscr{X}^N)$, because in general the natural extension is dominated by all jointly coherent extensions. $\underline{F}$ is the point-wise smallest extension of $\underline{P}_1$ to $\mathscr{L}(\mathscr{X}^N)$ that is jointly coherent with $\underline{P}_2(\cdot|X^1), \dots, \underline{P}_N(\cdot|X^{N-1})$, but it need not in general be jointly coherent with the corresponding extensions $\underline{M}_2(\cdot|X^1)$,$\dots$, $\underline{M}_N(\cdot|X^{N-1})$. $\underline{M}_1$ is, however, and it can be very easily calculated.

To see that these two extensions $\underline{M}_1$ and $\underline{F}$ do not coincide in general, it suffices to check that what we have done in Example 1 is nothing but calculate the marginal and the natural extensions of the coherent marginal lower previsions $\underline{P}_1$ and $\underline{P}_2$ defined on respective domains $\mathscr{H}_1 := \{\lambda\pi_1 \colon \lambda \in \mathbb{R}\} \subseteq \mathscr{L}(\mathscr{X}_1)$ and $\mathscr{H}_2 := \{\lambda\pi_2 \colon \lambda \in \mathbb{R}\} \subseteq \mathscr{L}(\mathscr{X}_2)$ by $\underline{P}_1(\lambda\pi_1) := \lambda$ and $\underline{P}_2(\lambda\pi_2) := \lambda$ for any $\lambda \in \mathbb{R}$, under the additional assumption of epistemic irrelevance.

In the rest of this section, we take a closer look at a number of properties of both the forward irrelevant product $\underline{M}_1$ and the forward irrelevant natural extension $\underline{F}$. Before doing this, we need some further preparation. First of all, we derive an explicit expression for the forward irrelevant natural extension $\underline{F}$.

It is clear, recalling the discussion and the notations leading to Theorem 4 in Section 5, that $\underline{F}$ is the natural extension of the conditional lower previsions $\hat{\underline{P}}_k(\cdot|\mathscr{B}_k)$ defined by

$$\hat{\underline{P}}_k(g|B_z) := \underline{P}_k(g(z,\cdot))$$

for all $z \in \mathscr{X}^{k-1}$, on domains $\mathscr{H}^k$ given by

$$\mathscr{H}^k := \left\{g \in \mathscr{L}(\mathscr{X}^k) \colon (\forall(x_1,\dots,x_{k-1}) \in \mathscr{X}^{k-1})(g(x_1,\dots,x_{k-1},\cdot) \in \mathscr{H}_k)\right\}.$$

Now consider any gamble f on $\mathscr{X}^N$. Taking into account that the domains $\mathscr{H}^k$ are cones but not linear spaces, $\underline{F}(f)$ is equal to (see Section 6 and [9, Definition 6 and Theorem 12]) the supremum value of α for which there are $g_k^j \in \mathscr{H}^k$, $j = 1,\dots,n_i$, $i = 1,\dots,N$ such that

$$\sup_{x\in\mathscr{X}^N}\left[\sum_{k=1}^N\sum_{j=1}^{n_k}\hat{G}(g_k^j|\mathscr{B}_k)(x) - [f(x)-\alpha]\right] < 0.$$

Now clearly

$$\hat{G}(g_k^j|\mathscr{B}_k) = \sum_{z\in\mathscr{X}^{k-1}} I_{B_z}[g_k^j - \underline{P}_k(g_k^j(z,\cdot))],$$

so we get

$$\underline{F}(f) = \sup_{\substack{g_k^j\in\mathscr{H}^k,\\ j=1,\dots,n_k,n_k\geq 0\\ k=1,\dots,N}} \inf_{x\in\mathscr{X}^N}\left[f(x) - \sum_{k=1}^N\sum_{j=1}^{n_k}[g_k^j(x_1,\dots,x_k) - \underline{P}_k(g_k^j(x_1,\dots,x_{k-1},\cdot))]\right]. \quad (26)$$

Finally, let $\underline{F}^k$ denote the forward irrelevant natural extension of the first k marginals $\underline{P}_1$, $\dots$, $\underline{P}_k$.

Proposition 3 (External linearity)**.** *Let f_k be any gamble on $\mathscr{X}_k$ for $1 \leq k \leq N$. Then*

$$\underline{F}^N\left(\sum_{k=1}^N f_k\right) = \underline{M}^N\left(\sum_{k=1}^N f_k\right) = \sum_{k=1}^N \underline{E}_k(f_k).$$

Proof. We give a proof by induction on N, where $N \geq 1$. It is obvious that the result holds for $N = 1$, since $\underline{M}^1(f_1) = \underline{F}^1(f_1) = \underline{E}_1(f_1)$. Assume therefore that the result holds for $N = n-1$, then we prove that it also holds for $N = n$, where $n \geq 2$. Denote by s_ℓ the gamble $\sum_{k=1}^{\ell} f_k$ on $\mathscr{X}^\ell$, then for any $(x_1,\dots,x_{n-1})$ in $\mathscr{X}^{n-1}$ we have that $s_n(x_1,\dots,x_{n-1},\cdot) = s_{n-1}(x_1,\dots,x_{n-1}) + f_n$, so it follows from the coherence of the lower prevision $\underline{E}_{n-1}$ that

$$\underline{E}_n(s_n(x_1,\dots,x_{n-1},\cdot)) = \underline{E}_n(s_{n-1}(x_1,\dots,x_{n-1}) + f_n) = s_{n-1}(x_1,\dots,x_{n-1}) + \underline{E}_n(f_n),$$

and from Eq. (25) and the coherence of the lower prevision $\underline{M}^n$ that

$$\begin{aligned}\underline{M}^n(s_n) &= \underline{M}^{n-1}(\underline{E}_n(s_n)) = \underline{M}^{n-1}(s_{n-1} + \underline{E}_n(f_n)) \\ &= \underline{M}^{n-1}(s_{n-1}) + \underline{E}_n(f_n) = \sum_{k=1}^{n-1} \underline{E}_k(f_k) + \underline{E}_n(f_n),\end{aligned}$$

where last equality follows from the induction hypothesis. We deduce that $\underline{F}^n(\sum_{k=1}^n f_k) \leq \underline{M}^n(\sum_{k=1}^n f_k) = \sum_{k=1}^n \underline{E}_k(f_k)$. Let us prove the converse inequality.

First of all, consider a gamble g_k on $\mathscr{H}_k$, for $k \in \{1,\dots,n\}$. Then, we deduce from Eq. (26) that $\underline{F}^n(g_k) \geq \underline{P}_k(g_k)$: it suffices to take $g_k^j = g_k, n_j = 1$ and $n_i = 0$ for all $i \neq j$. As a consequence, we deduce that $\underline{F}^n(g_k) \geq \underline{E}_k(g_k)$ for all gambles f_k on $\mathscr{X}_k$, and for all $k = 1,\dots,n$: $\underline{E}_k$ is the smallest coherent extension of $\underline{P}_k$ to $\mathscr{L}(\mathscr{X}_k)$, and is therefore dominated by all coherent lower previsions on $\mathscr{L}(\mathscr{X}_k)$ that dominate $\underline{P}_k$ on its domain. The super-additivity of $\underline{F}^n$ implies then that $\underline{F}^n(\sum_{k=1}^n f_k) \geq \sum_{k=1}^n \underline{F}^n(f_k) \geq \sum_{k=1}^n \underline{E}_k(f_k)$. □

We can generalise this result and prove the additivity of $\underline{M}^N$ and $\underline{F}^N$ on sums of gambles that depend on different variables. This means that if we consider two gambles f and g whose values depend on the outcome of different (disjoint) parts of the sequence $(X_1,\dots,X_N)$, then our supremum betting rate on the gamble $f+g$ should be the sum of our supremum betting rate on f and our supremum betting rate on g. However, and in contradistinction with the previous result, we shall not have in general the equality between $\underline{M}^N$ and $\underline{F}^N$ on these sums.

The forward irrelevant natural extension and the forward irrelevant product are indeed *products*: $\underline{F}^N$ and $\underline{M}^N$ are extensions of the marginals $\underline{P}_k$.

Proposition 4. *Let f_k be any gamble on $\mathscr{X}_k$. Then $\underline{F}^N(f_k) = \underline{M}^N(f_k) = \underline{E}_k(f_k)$. If in particular f_k belongs to $\mathscr{H}_k$, then $\underline{F}^N(f_k) = \underline{M}^N(f_k) = \underline{P}_k(f_k)$, for all $1 \leq k \leq N$.*

Proof. Immediately from Proposition 3 and the coherence of the marginal lower previsions $\underline{P}_k$, which tells us that $\underline{E}_k(f_k) = \underline{P}_k(f_k)$ for f_k in $\mathscr{H}_k$ and $k = 1,\dots,N$. □

The forward irrelevant natural extension and product also satisfy a (restricted) product rule.

Proposition 5 (Product Rule)**.** *Let f_k be a non-negative gamble on $\mathscr{X}_k$ for $1 \leq k \leq N$. Then*

$$\begin{aligned}\underline{F}^N(f_1 \dots f_N) &= \underline{M}^N(f_1 \dots f_N) = \underline{E}_1(f_1)\dots\underline{E}_N(f_N) \\ \overline{F}^N(f_1 \dots f_N) &= \overline{M}^N(f_1 \dots f_N) = \overline{E}_1(f_1)\dots\overline{E}_N(f_N).\end{aligned}$$

In particular, let A_k be any subset of $\mathscr{X}_k$ for $1 \leq k \leq N$. Then

$$\begin{aligned}\underline{F}^N(A_1 \times \dots \times A_N) &= \underline{M}^N(A_1 \times \dots \times A_N) = \underline{E}_1(A_1)\dots\underline{E}_N(A_N) \\ \overline{F}^N(A_1 \times \dots \times A_N) &= \overline{M}^N(A_1 \times \dots \times A_N) = \overline{E}_1(A_1)\dots\overline{E}_N(A_N).\end{aligned}$$

Proof. We shall prove the result for the lower previsions. The proof for the conjugate upper previsions is similar. We first prove the equality for the forward irrelevant product and then for the forward irrelevant natural extension. We apply induction on N. It is obvious that the result holds for $N=1$. Assume therefore that the result holds for $N=\ell-1$ (where $\ell\geq 2$), then we prove that the result holds for $N=\ell$ as well. Let f_k be a non-negative gamble on $\mathscr{X}_k$ for $1\leq k\leq \ell$. Then for any $(x_1,\dots,x_{\ell-1})\in\mathscr{X}^{\ell-1}$,

$$\begin{aligned}\underline{E}_\ell(f_1\dots f_\ell)(x_1,\dots,x_{\ell-1})&=\underline{E}_\ell(f_1(x_1)\dots f_{\ell-1}(x_{\ell-1})f_\ell)\\&=f_1(x_1)\dots f_{\ell-1}(x_{\ell-1})\underline{E}_\ell(f_\ell),\end{aligned}$$

since all gambles f_k are non-negative, and $\underline{E}_\ell$ is a coherent lower prevision. Then, using Eq. (25),

$$\begin{aligned}\underline{M}^\ell(f_1\dots f_\ell)&=\underline{M}^{\ell-1}(\underline{E}_\ell(f_1\dots f_\ell))\\&=\underline{M}^{\ell-1}(f_1\dots f_{\ell-1}\underline{E}_\ell(f_\ell))\\&=\underline{M}^{\ell-1}(f_1\dots f_{\ell-1})\underline{E}_\ell(f_\ell)\\&=\underline{E}_1(f_1)\dots\underline{E}_{\ell-1}(f_{\ell-1})\underline{E}_\ell(f_\ell),\end{aligned}$$

taking into account the coherence of the lower prevision $\underline{M}^{\ell-1}$ and the fact that the coherence of the lower prevision $\underline{E}_\ell$ implies that $\underline{E}_\ell(f_\ell)\geq\inf f_\ell\geq 0$. The last equality follows from the induction hypothesis. Since we already know that $\underline{F}^\ell(f_1\dots f_\ell)\leq\underline{M}^\ell(f_1\dots f_\ell)=\underline{E}_1(f_1)\dots\underline{E}_\ell(f_\ell)$, we now set out to prove the converse inequality. If $\underline{E}_\ell(f_\ell)=0$, the coherence of the lower prevision $\underline{F}^\ell$ implies that

$$\underline{F}^\ell(f_1\dots f_\ell)\geq 0=\underline{E}_1(f_1)\dots\underline{E}_\ell(f_\ell)=\underline{M}^\ell(f_1\dots f_\ell).$$

Assume therefore that $\underline{E}_\ell(f_\ell)>0$, and consider $0<\varepsilon<\underline{E}_\ell(f_\ell)$. Then it follows from the definition of the (unconditional) natural extension $\underline{E}_\ell$ [see Eq. (6) and use the fact that the domains $\mathscr{H}_k$ are assumed to be cones] that there are $n_\ell\geq 0$ and gambles g_ℓ^j in $\mathscr{H}_\ell$ for $j=1,\dots,n_\ell$ such that

$$\inf_{x_\ell\in\mathscr{X}_\ell}\left[f_\ell(x_\ell)-\sum_{j=1}^{n_\ell}\left[g_\ell^j(x_\ell)-\underline{P}_\ell(g_\ell^j)\right]\right]\geq\underline{E}_\ell(f_\ell)-\varepsilon. \tag{27}$$

Define the gambles h_ℓ^j on $\mathscr{X}^\ell$ by $h_\ell^j:=f_1\dots f_{\ell-1}g_\ell^j$. All these gambles[18] belong to $\mathscr{H}^\ell$. Now, using Eq. (26) for $N=\ell$ and $f=f_1\dots f_\ell$, we get that $\underline{F}^\ell(f_1\dots f_\ell)$ is greater than or equal to

$$\sup_{\substack{g_i^j\in\mathscr{H}^i,j=1,\dots,n_i,\\ i=1,\dots,n-1}}\inf_{x\in\mathscr{X}^\ell}\left[f_1(x_1)\dots f_\ell(x_\ell)-\sum_{i=1}^{\ell-1}\sum_{j=1}^{n_i}\hat{G}(g_i^j|\mathscr{B}_i)\right.$$
$$\left.-\sum_{j=1}^{n_\ell}\left[h_\ell^j(x_1,\dots,x_\ell)-\underline{P}_\ell(h_\ell^j(x_1,\dots,x_{\ell-1},\cdot))\right]\right],$$

[18]This holds because we assumed that the domains $\mathscr{H}_k$ of the marginal lower previsions $\underline{P}_k$ are cones. See footnote 17.

and after some manipulations, using the coherence of $\underline{P}_\ell$ and the fact that all the f_k are non-negative, this can be rewritten as

$$\sup_{\substack{g_i^j \in \mathscr{H}^i, j=1,\dots,n_i,\\ i=1,\dots,n-1}} \inf_{x \in \mathscr{X}^{\ell-1}} \Bigg[-\sum_{i=1}^{\ell-1} \sum_{j=1}^{n_i} \hat{G}(g_i^j|\mathscr{B}_i) + f_1(x_1)\dots f_{\ell-1}(x_{\ell-1}) \inf_{x_\ell \in \mathscr{X}_\ell} \Bigg[f_\ell(x_\ell) - \sum_{j=1}^{n_\ell} \Big[g_\ell^j(x_\ell) - \underline{P}_\ell(g_\ell^j) \Big] \Bigg] \Bigg].$$

Now if we use the inequality (27), we see that this is greater than or equal to

$$\sup_{\substack{g_i^j \in \mathscr{H}^i, j=1,\dots,n_i,\\ i=1,\dots,n-1}} \inf_{x \in \mathscr{X}^{\ell-1}} \Bigg[-\sum_{i=1}^{\ell-1} \sum_{j=1}^{n_i} \hat{G}(g_i^j|\mathscr{B}_i) + f_1(x_1)\dots f_{\ell-1}(x_{\ell-1})[\underline{E}_\ell(f_\ell) - \varepsilon] \Bigg],$$

Summing all this up, and using Eq. (26) for $N = \ell - 1$ and $f = f_1 \dots f_{\ell-1}[\underline{E}_\ell(f_\ell) - \varepsilon]$, we get that

$$\begin{aligned} \underline{F}^\ell(f_1 \dots f_\ell) &\geq \underline{F}^{\ell-1}(f_1 \dots f_{\ell-1}[\underline{E}_\ell(f_\ell) - \varepsilon]) \\ &= [\underline{E}_\ell(f_\ell) - \varepsilon]\underline{F}^{\ell-1}(f_1 \dots f_{\ell-1}) \\ &= [\underline{E}_\ell(f_\ell) - \varepsilon]\underline{E}_1(f_1)\dots\underline{E}_{\ell-1}(f_{\ell-1}), \end{aligned}$$

where the first equality follows from the coherence of the lower prevision $\underline{F}^{\ell-1}$ and the fact that $\underline{E}_\ell(f_\ell) - \varepsilon > 0$, and the second equality from the induction hypothesis. Since this happens for ε arbitrarily close to 0, we deduce that indeed

$$\underline{F}^\ell(f_1 \dots f_\ell) \geq \underline{E}_1(f_1)\dots\underline{E}_\ell(f_\ell).$$

The second part of the proposition follows immediately from the first. □

These and other properties of the forward irrelevant product allows us to establish laws of large numbers for coherent lower previsions, see [2] for more information.

8. Conclusions

The problem of coherently extending a number of assessments is one of the most important in the theory of subjective probability. When these assessments are represented by means of an unconditional lower prevision, the way to do so is by means of Walley's notion of natural extension. This extension has a clear behavioural interpretation and can also be given a Bayesian sensitivity one, as a lower envelope of linear extensions.

The task becomes more involved when the assessments are of a more complicated nature, and are represented by means of conditional lower previsions. In that case, Walley's notion of natural extension does not generally yield the smallest coherent extensions, but only a lower bound for them. This can't be helped in some cases, because there may not be any coherent extensions at all; but in other cases there *are* coherent extensions, but the procedure of natural extension may fail to produce any of them. This leads us to search for the smallest coherent extensions, which will reflect the minimal behavioural consequences of the given assessments.

Walley has proved that when we have an unconditional and a conditional lower prevision with some properties, the smallest coherent extensions are obtained through the procedure of marginal extension. In this paper, we have extended his result to the case where we have a finite number of lower previsions conditional on increasingly finer partitions.

The marginal extensions then provide the smallest coherent extensions, and moreover have a sensitivity analysis representation as lower envelopes of linear previsions. As such, they prove to be superior of the ones obtained through natural extension.

In essence, what our results tell us is that if we have hierarchical assessments (which is the idea behind the increasingly finer partitions, and more clearly behind the representation for variables), the way to extend these assessments to all gambles is to use unconditional natural extension at each hierarchical level, and then use concatenation. Moreover, this concatenation is equivalent to Bayes's rule in the case of linear conditional and unconditional prevision.

As topics for further research, we suggest the study of the smallest coherent extensions under other conditions, and the investigation of the equality between natural and marginal extensions. We'd also like to mention our suspicion that it may be possible to find a simpler, or perhaps more directly intuitive, proof for our Marginal Extension Theorem.

Acknowledgements

We gratefully acknowledge the financial support provided by MCYT, under the projects MTM2004-01269, TSI2004-06801-C04-01, and by research grant G.0139.01 of the Flemish Fund for Scientific Research (FWO). We also wish to acknowledge Marco Zaffalon and two anonymous referees for some insightful comments.

References

[1] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier. Universit de Grenoble*, 5:131–295, 1953.
[2] G. de Cooman and E. Miranda. Weak and strong laws of large numbers for coherent lower previsions. 2006. Submitted for publication.
[3] G. de Cooman and E. Miranda. Forward epistemic irrelevance. 2006. Submitted for publication.
[4] G. de Cooman and M. Zaffalon. Updating beliefs with incomplete observations. *Artificial Intelligence*, 159(1–2):75–125, 2004.
[5] I. Couso, S. Moral and P. Walley. Examples of independence with imprecise probabilities. *Risk, Decision and Policy*, 5:165–181, 2000.
[6] B. de Finetti. *Theory of probability*. John Wiley and Sons, Chichester, 1974–1975. English translation of *Teoria delle Probabilità*.
[7] D. Denneberg. *Non-additive measure and integral*. Kluwer Academic Publishers, Dordrecht, 1994.
[8] I. Levi. *The enterprise of knowledge*. MIT Press, Cambridge, 1980.
[9] E. Miranda and G. de Cooman. Coherence and independence in non-linear spaces. *Technical Reports of Statistics and Decision Sciences,* (http://bayes.escet.urjc.es/∼publicaciones.php) TR05/10, 2005.
[10] P. Walley. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London, 1991.
[11] P. Williams. Notes on conditional lower previsions. *Technical Report, University of Sussex*, 1975.

REY JUAN CARLOS UNIVERSITY, DEPT. OF STATISTICS AND OPERATIONS RESEARCH. C-TULIPÁN, S/N 28933, MÓSTOLES, SPAIN
E-mail address: enrique.miranda@urjc.es

GHENT UNIVERSITY, SYSTEMS RESEARCH GROUP, TECHNOLOGIEPARK – ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM
E-mail address: gert.decooman@ugent.be