Independent natural extension for infinite spaces ☆

Jasper De Bock

Ghent University, Department of Electronics and Information Systems, Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde, Belgium

A R T I C L E   I N F O

A B S T R A C T

We define and study the independent natural extension of two local uncertainty models for the general case of infinite spaces, using the frameworks of sets of desirable gambles and conditional lower previsions. In contrast to Miranda and Zaffalon [16], we adopt Williams-coherence instead of Walley-coherence. We show that our notion of independent natural extension always exists—whereas theirs does not—and that it satisfies various convenient properties, including factorisation and external additivity. The strength of these properties depends on the specific type of epistemic independence that is adopted. In particular, epistemic event-independence is shown to outperform epistemic atom-independence. Finally, the cases of lower expectations, expectations, lower probabilities and probabilities are obtained as special instances of our general definition. By applying our results to these instances, we demonstrate that epistemic independence is indeed epistemic, and that it includes the conventional notion of independence as a special case.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

When probabilities are imprecise, in the sense that they are only partially specified, it is no longer clear what it means for two variables to be independent [1]. One approach is to apply the standard notion of independence to every element of some set of probability measures. The alternative, called epistemic independence, is to define independence as mutual irrelevance, in the sense that receiving information about one of the variables will not affect our uncertainty model for the other. The advantage of this intuitive alternative is that it has a much wider scope: since epistemic independence is expressed in terms of uncertainty models instead of probabilities, it can easily be applied to a variety of such models, including non-probabilistic ones. We here focus on sets of desirable gambles and conditional lower previsions, the latter of which includes lower expectations, expectations, lower probabilities and probabilities as special cases.

When an assessment of epistemic independence is combined with local uncertainty models, it leads to a unique corresponding joint uncertainty model that is called the independent natural extension. If the variables involved can take only a finite number of values, this independent natural extension always exists, and it then satisfies various convenient properties that allow for the design of efficient algorithms [9,10]. If the variables involved take values in an infinite set, the situation becomes more complicated. On the one hand, for the specific case of lower probabilities, Vicig [21] managed to obtain results that resemble those of the finite case. On the other hand, for the more general case of lower previsions, Miranda and Zaffalon [16] recently found that the independent natural extension may not even exist.

☆ This paper is part of the Virtual special issue on Tenth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA '17), Edited by Alessandro Antonucci, Giorgio Corani, Inés Couso and Sébastien Destercke.

E-mail address: jasper.debock@ugent.be.

Our main contribution consists in generalising the results of Vicig [21] to the case of conditional lower previsions, using sets of desirable gambles as an intermediate step. The key technical difference with [16] is that we use Williams-coherence [24,25] instead of Walley-coherence [22]. This difference turns out to be crucial because our notion of independent natural extension always exists. Furthermore, as we will see, it satisfies the same convenient properties that are known to hold in the finite case, including factorisation and external additivity.

An essential feature of our approach is that we adopt a very general notion of epistemic independence, where the choice of conditioning events is initially left open. Important examples such as epistemic atom-independence and epistemic event-independence are then obtained as special cases. In this way, we are able to study the effect of different types of epistemic independence on the resulting notion of independent natural extension.

The rest of the paper is structured as follows. We start in Section 2 by introducing some basic technical concepts and notation that will be needed further on. Next, in Section 3, we explain how uncertainty can be modelled with sets of desirable gambles and conditional lower previsions. For readers that are unfamiliar with these frameworks, this section serves as a stand-alone introduction. With all these preliminaries in place, the paper then moves on to its main contributions, which we report on in Sections 4–7. Section 4 introduces our general notion of epistemic independence and explains how it subsumes epistemic atom- and event-independence as special cases. Next, Section 5 introduces our central object of interest, which is the independent natural extension, and Section 6 discusses the extent to which it depends on the specific type of epistemic independence that is adopted. Crucially, we find that in our approach, regardless of the chosen type of epistemic independence, the independent natural extension always exists. Our external additivity and factorisation results are reported on in Section 7; the strength of these results does depend on the chosen type of epistemic independence. In the last part of the paper, which consists of Sections 8 and 9, we reinterpret our results in terms of lower expectations, expectations, lower probabilities and probabilities. Section 8 recalls how all of these models are special cases of conditional lower previsions, which then enables us to apply our results to them in Section 9. We also use this connection to explain why epistemic independence is indeed epistemic, and how it includes the conventional notion of independence as a special case. We end the paper in Section 10 with a brief summary of our main results and findings. The proofs of our results are gathered in an appendix.

Finally, I would like to add that this paper is an extended version of an earlier conference version [4]. The most substantial additions are the results in Sections 8 and 9 and the proofs in Appendix A. We have also added several examples.

2. Preliminaries and notation

We use $\mathbb{N}$ to denote the natural numbers without zero and let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. $\mathbb{R}$ is the set of real numbers and $\mathbb{Q}$ is the set of rational numbers. Sign restrictions are imposed with subscripts. For example, we let $\mathbb{R}_{>0}$ be the set of positive real numbers and let $\mathbb{Q}_{\geq 0}$ be the set of non-negative rational numbers. The extended real numbers are denoted by $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.

For any non-empty set $\mathcal{X}$, the power set of $\mathcal{X}$—the set of all subsets of $\mathcal{X}$—is denoted by $\mathcal{P}(\mathcal{X})$, and we let $\mathcal{P}_\emptyset(\mathcal{X}) := \mathcal{P}(\mathcal{X}) \setminus \{\emptyset\}$ be the set of all non-empty subsets of $\mathcal{X}$. Elements of $\mathcal{P}(\mathcal{X})$ are called events. A set of events $\mathcal{B} \subseteq \mathcal{P}(\mathcal{X})$ is called a field if it is non-empty and closed with respect to complements and finite intersections and unions. If it is also closed with respect to countable intersections and unions, it is called a sigma field. A partition of $\mathcal{X}$ is a set $\mathcal{B} \subseteq \mathcal{P}_\emptyset(\mathcal{X})$ of pairwise disjoint non-empty subsets of $\mathcal{X}$ whose union is equal to $\mathcal{X}$. We also adopt the notational trick of identifying $\mathcal{X}$ with the set of atoms $\{\{x\}: x \in \mathcal{X}\}$, which allows us to regard $\mathcal{X}$ as a partition of $\mathcal{X}$.

A bounded real-valued function on $\mathcal{X}$ will be called a gamble on $\mathcal{X}$. The set of all gambles on $\mathcal{X}$ is denoted by $\mathcal{G}(\mathcal{X})$, the set of all non-negative gambles on $\mathcal{X}$ is denoted by $\mathcal{G}_{\geq 0}(\mathcal{X})$, and we let $\mathcal{G}_{>0}(\mathcal{X}) := \mathcal{G}_{\geq 0}(\mathcal{X}) \setminus \{0\}$ be the set of all non-negative non-zero gambles. For any set of gambles $\mathcal{A} \subseteq \mathcal{G}(\mathcal{X})$, we let

$$\mathrm{posi}(\mathcal{A}) := \left\{ \sum_{i=1}^n \lambda_i f_i : n \in \mathbb{N}, \lambda_i \in \mathbb{R}_{>0}, f_i \in \mathcal{A} \right\} \tag{1}$$

and

$$\mathcal{E}(\mathcal{A}) := \mathrm{posi}\left(\mathcal{A} \cup \mathcal{G}_{>0}(\mathcal{X})\right). \tag{2}$$

Indicators are a particular type of gamble. For any $A \in \mathcal{P}(\mathcal{X})$, the corresponding indicator $\mathbb{I}_A$ of A is a gamble in $\mathcal{G}_{\geq 0}(\mathcal{X})$, defined for all $x \in \mathcal{X}$ by $\mathbb{I}_A(x) := 1$ if $x \in A$ and $\mathbb{I}_A(x) := 0$ otherwise.

Finally, for any—possibly empty—$\mathcal{B} \subseteq \mathcal{P}_\emptyset(\mathcal{X})$, we will also require the notion of a non-negative $\mathcal{B}$-measurable gamble, which we define as a uniform limit of simple $\mathcal{B}$-measurable gambles.

Definition 2.1. Let $\mathcal{B} \subseteq \mathcal{P}_\emptyset(\mathcal{X})$. We call $g \in \mathcal{G}_{\geq 0}(\mathcal{X})$ a simple $\mathcal{B}$-measurable gamble if there are $c_0 \in \mathbb{R}_{\geq 0}$, $n \in \mathbb{N}_0$ and, for all $i \in \{1, \ldots, n\}$, $c_i \in \mathbb{R}_{\geq 0}$ and $B_i \in \mathcal{B}$, such that $g = c_0 + \sum_{i=1}^n c_i \mathbb{I}_{B_i}$.

Definition 2.2. Let $\mathcal{B} \subseteq \mathcal{P}_\emptyset(\mathcal{X})$. A gamble $g \in \mathcal{G}_{\geq 0}(\mathcal{X})$ is $\mathcal{B}$-measurable if it is a uniform limit of non-negative simple $\mathcal{B}$-measurable gambles, in the sense that there is a sequence $\{g_n\}_{n \in \mathbb{N}}$ of simple $\mathcal{B}$-measurable gambles in $\mathcal{G}_{\geq 0}(\mathcal{X})$ such that $\lim_{n \to +\infty} \sup |g - g_n| = 0$.

Readers that are familiar with the concepts of simple and measurable functions that are common in measure theory will observe some similarities. However, there are some important differences as well. On the one hand, our definitions are more restrictive: we only consider bounded non-negative functions, Definition 2.1 requires that the coefficients c_i are non-negative, and Definition 2.2 considers uniform limits instead of pointwise limits. On the other hand, our definitions are more general because we allow for $\mathscr{B}$ to be any subset of $\mathscr{P}_\emptyset(\mathscr{X})$. Nevertheless, if $\mathscr{B} \cup \{\emptyset\}$ is a sigma field, we have the following equivalence.

Proposition 2.1. *Consider any $\mathscr{B} \subseteq \mathscr{P}_\emptyset(\mathscr{X})$ such that $\mathscr{B}^* := \mathscr{B} \cup \{\emptyset\}$ is a sigma field. Then for any $g \in \mathscr{G}_{\geq 0}(\mathscr{X})$, g is $\mathscr{B}^*$-measurable in the measure-theoretic sense [17, Definition 10.1] if and only if it is $\mathscr{B}$-measurable in the sense of Definition 2.2.*

The proof of this result is based on the following sufficient condition for $\mathscr{B}$-measurability, which provides a convenient tool for establishing the $\mathscr{B}$-measurability of a given function. In particular, it implies that every non-negative gamble is $\mathscr{P}_\emptyset(\mathscr{X})$-measurable.

Proposition 2.2. *Let $\mathscr{B} \subseteq \mathscr{P}_\emptyset(\mathscr{X})$ and $g \in \mathscr{G}_{\geq 0}(\mathscr{X})$. If, for all $r \in \mathbb{Q}_{\geq 0}$, the set $\{x \in \mathscr{X} : g(x) \geq r\}$ is a finite union of pairwise disjoint events in $\mathscr{B} \cup \{\mathscr{X}, \emptyset\}$, then g is $\mathscr{B}$-measurable.*

Corollary 2.3. *Every $g \in \mathscr{G}_{\geq 0}(\mathscr{X})$ is $\mathscr{P}_\emptyset(\mathscr{X})$-measurable.*

The following three examples provide these abstract concepts and results with some intuition, by studying the (non-)$\mathscr{B}$-measurability of various functions, for different choices of $\mathscr{B}$. They also demonstrate that $\mathscr{B}$-measurability is easier to achieve if $\mathscr{B}$ contains more events.

Example 2.1. Let $\mathscr{X} = \mathbb{N}$ and let $g \in \mathscr{G}_{\geq 0}(\mathscr{X})$ be defined by $g(x) := \frac{1}{x}$ for all $x \in \mathbb{N}$. Corollary 2.3 then trivially implies that g is $\mathscr{P}_\emptyset(\mathbb{N})$-measurable. However, g is also $\mathscr{B}$-measurable for some strict subsets $\mathscr{B}$ of $\mathscr{P}_\emptyset(\mathbb{N})$. For example, Proposition 2.2 implies that g is $\mathbb{N}$-measurable, because for every $r \in \mathbb{Q}_{\geq 0}$, the set

$$\{x \in \mathbb{N} : g(x) \geq r\} = \{x \in \mathbb{N} : \frac{1}{x} \geq r\} = \begin{cases} \emptyset & \text{if } r > 1 \\ \{1, \ldots, \lfloor \frac{1}{r} \rfloor\} & \text{if } 0 < r \leq 1 \\ \mathbb{N} & \text{if } r = 0 \end{cases}$$

is clearly a finite union of pairwise disjoint events in $\mathscr{B} \cup \{\mathscr{X}, \emptyset\} = \mathbb{N} \cup \{\mathbb{N}, \emptyset\}$. The set $\mathscr{B}$ cannot be too small though. For example, in the extreme case where $\mathscr{B} = \emptyset$, g is no longer $\mathscr{B}$-measurable. The easiest way to see this is to infer from Definition 2.1 that simple $\emptyset$-measurable gambles are constant. Therefore, since uniform limits of constant gambles are constant, Definition 2.2 implies that all $\emptyset$-measurable gambles are constant. Hence, since g is not constant, it is not $\emptyset$-measurable. ◇

Example 2.2. Let $\mathscr{X} = \mathbb{N}$ and let $\mathbb{I}_{\text{odd}} \in \mathscr{G}_{\geq 0}(\mathbb{N})$ be the indicator of the odd numbers, defined for all $x \in \mathbb{N}$ by $\mathbb{I}_{\text{odd}}(x) = 1$ if x is odd and $\mathbb{I}_{\text{odd}}(x) = 0$ otherwise. Similarly, let $\mathbb{I}_{\text{even}} \in \mathscr{G}_{\geq 0}(\mathbb{N})$ be the indicator of the even numbers, defined by $\mathbb{I}_{\text{even}}(x) := 1$ if x is even and $\mathbb{I}_{\text{even}}(x) := 0$ otherwise. Here too, it follows from Corollary 2.3 that $\mathbb{I}_{\text{odd}}$ and $\mathbb{I}_{\text{even}}$ are $\mathscr{P}_\emptyset(\mathbb{N})$-measurable. However, as we are about to prove, and in contrast with the previous example, $\mathbb{I}_{\text{odd}}$ and $\mathbb{I}_{\text{even}}$ are not $\mathbb{N}$-measurable. We focus on $\mathbb{I}_{\text{odd}}$; the argument for $\mathbb{I}_{\text{even}}$ is completely analogous.

Assume *ex absurdo* that $\mathbb{I}_{\text{odd}}$ is $\mathbb{N}$-measurable. It then follows from Definition 2.2 that there is some simple $\mathbb{N}$-measurable function $g \in \mathscr{G}_{\geq 0}$ such that $|\mathbb{I}_{\text{odd}}(x) - g(x)| < \frac{1}{2}$ for all $x \in \mathbb{N}$. Since $\mathscr{B} = \mathbb{N}$, Definition 2.1 then implies that there is some $c_0 \in \mathbb{R}_{\geq 0}$ and some finite set A such that $g(x) = c_0$ for all $x \in \mathbb{N} \setminus A$. Since $\mathbb{N}$ contains infinitely many odd and even numbers, this in turn implies that there is some (and in fact infinitely many) odd x_{odd} and even x_{even} in $\mathbb{N} \setminus A$ for which $g(x_{\text{odd}}) = g(x_{\text{even}}) = c_0$. Hence, we find that

$$1 = |\mathbb{I}_{\text{odd}}(x_{\text{odd}}) - \mathbb{I}_{\text{odd}}(x_{\text{even}})| \leq |\mathbb{I}_{\text{odd}}(x_{\text{odd}}) - c_0| + |c_0 - \mathbb{I}_{\text{odd}}(x_{\text{even}})|$$

$$= |\mathbb{I}_{\text{odd}}(x_{\text{odd}}) - g(x_{\text{odd}})| + |g(x_{\text{even}}) - \mathbb{I}_{\text{odd}}(x_{\text{even}})| < \frac{1}{2} + \frac{1}{2} = 1,$$

a contradiction. ◇

Example 2.3. Let $\mathscr{X} = \mathbb{R}$ and let $\mathscr{B} := \mathscr{B}^* \setminus \{\emptyset\}$, with $\mathscr{B}^*$ the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}$. Let $g \in \mathscr{G}_{\geq 0}(\mathbb{R})$ be the indicator of the non-negative reals, defined for all $x \in \mathbb{R}$ by $g(x) := 1$ if $x \geq 0$ and $g(x) := 0$ otherwise, and let $h \in \mathscr{G}(\mathbb{R})$ be defined by $h(x) := x^3$ for all $x \in \mathbb{R}$. Then g is Lebesgue measurable because it is a step function and h is Lebesgue measurable because it is continuous. Therefore, it follows from Proposition 2.1 that g and h are $\mathscr{B}$-measurable. ◇

3. Modelling uncertainty

A subject's uncertainty about a variable X that takes values x in some non-empty set $\mathscr{X}$ can be mathematically represented in various ways. The most popular such method is probability theory, but it is by no means the only one, nor is it the most general one. We here adopt the more general frameworks of sets of desirable gambles and conditional lower previsions.

The main aim of this section is to provide an overview of the basic technical aspects of these frameworks, because they will be essential in the rest of the paper. Notably, we do not impose any constraints on the cardinality of $\mathscr{X}$: it may be finite, countably infinite or uncountably infinite. Connections with other—perhaps better known—models for uncertainty, including probability theory, will be briefly touched upon at the end of this section; detailed connections will be established in Section 8.

The basic idea behind *sets of desirable gambles* is to model a subject's uncertainty about X by considering her attitude towards gambles—bets—on $\mathscr{X}$. In particular, we consider the gambles $f \in \mathscr{G}(\mathscr{X})$ that she finds *desirable*, in the sense that she is willing to engage in a transaction where, once the actual value $x \in \mathscr{X}$ of X is known, she will receive a—possibly negative—reward $f(x)$ in some linear utility scale. Even more so, she prefers these desirable gambles over the status quo, that is, over not conducting any transaction at all. A set of desirable gambles is called coherent if it satisfies the following rationality requirements.

Definition 3.1. A coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}$ is a subset of $\mathscr{G}(\mathscr{X})$ such that, for any two gambles $f, g \in \mathscr{G}(\mathscr{X})$ and any positive real number $\lambda \in \mathbb{R}_{>0}$:

D1: if $f \geq 0$ and $f \neq 0$, then $f \in \mathscr{D}$;
D2: if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$;
D3: if $f, g \in \mathscr{D}$, then $f + g \in \mathscr{D}$;
D4: if $f \leq 0$, then $f \notin \mathscr{D}$.

The first axiom states that the possibility of a positive reward without risking a negative reward should always be desirable, whereas the fourth axiom states that gambles that offer no positive rewards should never be desirable. The other two axioms are immediate consequences of the linearity of the utility scale. Despite their simplicity, sets of desirable gambles offer a surprisingly powerful framework for modelling uncertainty; see for example References [20,23]. For our present purposes though, all we need for now is Definition 3.1.

Conditional lower previsions also model a subject's uncertainty about X by considering her attitude towards gambles on $\mathscr{X}$. However, in this case, instead of considering sets of gambles, we consider the prices at which a subject is willing to buy these gambles. Let

$$\mathscr{C}(\mathscr{X}) := \mathscr{G}(\mathscr{X}) \times \mathscr{P}_\emptyset(\mathscr{X})$$

be the set of all pairs (f, B), where f is a gamble on $\mathscr{X}$ and B is a non-empty subset of $\mathscr{X}$—an event. A conditional lower prevision is then defined as follows.

Definition 3.2. A conditional lower prevision $\underline{P}$ on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is a map

$$\underline{P} : \mathscr{C} \to \overline{\mathbb{R}} : (f, B) \to \underline{P}(f|B).$$

For any (f, B) in the domain $\mathscr{C}$, the lower prevision $\underline{P}(f|B)$ of f conditional on B is interpreted as a subject's supremum price μ for buying f, under the condition that the transaction is called off when B does not happen—if $x \notin B$. In other words, $\underline{P}(f|B)$ is the supremum value of μ for which she is willing to engage in a transaction where she receives $f(x) - \mu$ if $x \in B$ and zero otherwise, and furthermore prefers this transaction to the status quo. If $B = \mathscr{X}$, we adopt the shorthand notation $\underline{P}(f) := \underline{P}(f|\mathscr{X})$ and then call $\underline{P}(f)$ the lower prevision of f. If $B = \mathscr{X}$ for all $(f, B) \in \mathscr{C}$, meaning that there is some $\mathscr{G} \subseteq \mathscr{G}(\mathscr{X})$ such that $\mathscr{C} = \{(f, \mathscr{X}) : f \in \mathscr{G}\}$, this convention allows us to regard $\underline{P}$ as an operator on $\mathscr{G}$, and we then say that $\underline{P}$ is a *lower prevision*. In this sense, lower previsions are a special case of conditional lower previsions.

It is also possible to consider *conditional upper previsions* $\overline{P}(f|B)$, which are interpreted as infimum selling prices. However, since selling f for μ is equivalent to buying $-f$ for $-\mu$, we have that $\overline{P}(f|B) = -\underline{P}(-f|B)$. For that reason, we will mainly focus on conditional lower previsions. A similar remark applies to (unconditional) *upper previsions*.

Because of their interpretation in terms of buying prices for gambles, a particularly intuitive way to obtain a conditional lower prevision $\underline{P}$ is to derive it from a set of gambles $\mathscr{D}$. Specifically, for every $\mathscr{D} \subseteq \mathscr{G}(\mathscr{X})$, we let

$$\underline{P}_{\mathscr{D}}(f|B) := \sup\{\mu \in \mathbb{R} : [f - \mu]\mathbb{I}_B \in \mathscr{D}\} \text{ for all } (f, B) \in \mathscr{C}(\mathscr{X}). \tag{3}$$

A conditional lower prevision is then called coherent if can be derived from a coherent set of desirable gambles in this way.

Definition 3.3. A conditional lower prevision $\underline{P}$ on a domain $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is coherent if there is a coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}$ such that $\underline{P}$ coincides with $\underline{P}_{\mathscr{D}}$ on $\mathscr{C}$.

This definition of coherence is heavily inspired by the work of Williams [24,25]. The only two minor differences are that our rationality axioms on $\mathscr{D}$ are slightly different from his—for example, he allows for $\mathscr{D}$ to include the zero gamble—and that we do not impose any structure on the domain $\mathscr{C}$. Nevertheless, when the domain $\mathscr{C}$ satisfies the structural constraints in [25], Definition 3.3 is equivalent to that of Williams. More generally, as the following result establishes, it is equivalent to the structure-free notion of Williams-coherence that was developed by Pelessoni and Vicig [18].

Proposition 3.1. *A conditional lower prevision $\underline{P}$ on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is coherent if and only if it is real-valued and, for all $n \in \mathbb{N}_0$ and all choices of $\lambda_0, \ldots, \lambda_n \in \mathbb{R}_{\geq 0}$ and $(f_0, B_0), \ldots, (f_n, B_n) \in \mathscr{C}$:*

$$\sup_{x \in B} \left(\sum_{i=1}^{n} \lambda_i \mathbb{I}_{B_i}(x)[f_i(x) - \underline{P}(f_i|B_i)] - \lambda_0 \mathbb{I}_{B_0}(x)[f_0(x) - \underline{P}(f_0|B_0)] \right) \geq 0, \tag{4}$$

where we let $B := \cup_{i=0}^{n} B_i$.

The advantage of this alternative characterisation is that it is expressed directly in terms of lower previsions. Nevertheless, we consider Equation (4) to be less intuitive than Definition 3.3, which is why we prefer the latter.

From a mathematical point of view, Definition 3.3 also has the advantage that it allows for simple and elegant proofs of some well-known results. For example, it follows trivially from our definition of coherence that the domain of a coherent conditional lower prevision can be arbitrarily extended while preserving coherence, whereas deriving this result directly from Equation (4) is substantially more involved; see for example the proof of [18, Proposition 1]. Furthermore, our definition also allows for a very natural derivation of the so-called *natural extension* of $\underline{P}$, that is, the most conservative extension of $\underline{P}$ to $\mathscr{C}(\mathscr{X})$. In particular, instead of having to derive this natural extension directly, Definition 3.3 allows us to rephrase this problem into a closely related yet simpler question: what is the smallest coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}$ such that $\underline{P}_{\mathscr{D}}$ coincides with $\underline{P}$ on $\mathscr{C}$? The answer turns out to be surprisingly simple.

Proposition 3.2. *Consider a coherent conditional lower prevision $\underline{P}$ on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ and let*

$$\mathscr{A}_{\underline{P}} := \left\{ [f - \mu]\mathbb{I}_B \colon (f, B) \in \mathscr{C}, \mu < \underline{P}(f|B) \right\} \quad \text{and} \quad \mathscr{E}(\underline{P}) := \mathscr{E}(\mathscr{A}_{\underline{P}}). \tag{5}$$

Then $\mathscr{E}(\underline{P})$ is a coherent set of desirable gambles on $\mathscr{X}$ and $\underline{P}_{\mathscr{E}(\underline{P})}$ coincides with $\underline{P}$ on $\mathscr{C}$. Furthermore, for any other coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}$ such that $\underline{P}_{\mathscr{D}}$ coincides with $\underline{P}$ on $\mathscr{C}$, we have that $\mathscr{E}(\underline{P}) \subseteq \mathscr{D}$.

Abstracting away some technical details, the reason why this result holds should be intuitively clear. First, since conditional lower previsions are interpreted as called-off supremum buying prices, we see that the gambles in $\mathscr{A}_{\underline{P}}$ should be desirable. Combined with D1–D3, the desirability of the gambles in $\mathscr{E}(\underline{P})$ then follows.

Since smaller sets of desirable gambles lead to more conservative—pointwise smaller—lower previsions, we conclude that the natural extension of $\underline{P}$ is given by

$$\underline{E}(f|B) := \underline{P}_{\mathscr{E}(\underline{P})}(f|B) \text{ for all } (f, B) \in \mathscr{C}(\mathscr{X}). \tag{6}$$

The following proposition provides a formal statement of this result.

Proposition 3.3. *Let $\underline{P}$ be a coherent conditional lower prevision on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$. Then $\underline{E}$, as defined by Equation (6), is the pointwise smallest coherent conditional lower prevision on $\mathscr{C}(\mathscr{X})$ that coincides with $\underline{P}$ on $\mathscr{C}$.*

All in all, we conclude that Definition 3.3 provides an intuitive as well as mathematically convenient characterisation of Williams-coherence that is furthermore equivalent to the structure-free version of Pelessoni and Vicig [18]. From a technical point of view, this equivalence will not be essential further on, because most of our arguments will be based on the connection with sets of desirable gambles. From a practical point of view though, this equivalence is highly important, because the Williams-coherent conditional lower previsions that are considered in [18] are well-known to include as special cases a variety of other uncertainty models, including expectations, lower expectations, probabilities and lower probabilities. For that reason, our results can be applied to—and interpreted in terms of—these special cases as well. A detailed discussion of this point is deferred to Sections 8 and 9; for now, we focus on sets of desirable gambles and conditional lower previsions. We end this section by listing some well-known properties of the latter; see for example References [18,22,25].

Proposition 3.4. *Let $\underline{P}$ be a coherent conditional lower prevision on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$. Then for any two gambles $f, g \in \mathscr{G}(\mathscr{X})$, any two events $A, B \in \mathscr{P}_{\emptyset}(\mathscr{X})$, any real number $\lambda \in \mathbb{R}$ and any sequence of gambles $\{f_n\}_{n \in \mathbb{N}} \subseteq \mathscr{G}(\mathscr{X})$, whenever the involved conditional lower and upper previsions are well-defined—that is, if the arguments belong to their domain—we have that*

LP1: $\underline{P}(f|B) \geq \inf_{x \in B} f(x)$ *[boundedness]*
LP2: $\underline{P}(\lambda f|B) = \lambda \underline{P}(f|B)$ *if* $\lambda \geq 0$ *[non-negative homogeneity]*
LP3: $\underline{P}(f+g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$ *[superadditivity]*
LP4: $\underline{P}(\mathbb{I}_B[f - \underline{P}(f|A \cap B)]|A) = 0$ *if* $A \cap B \neq \emptyset$ *[generalised Bayes rule]*
LP5: $\lim_{n \to +\infty} \sup |f - f_n| = 0 \Rightarrow \lim_{n \to +\infty} \underline{P}(f_n|B) = \underline{P}(f|B)$ *[uniform continuity]*
LP6: $\underline{P}(f+\lambda|B) = \underline{P}(f|B) + \lambda$ *[constant additivity]*
LP7: $f(x) \geq g(x)$ *for all* $x \in B \Rightarrow \underline{P}(f|B) \geq \underline{P}(g|B)$ *[monotonicity]*
LP8: $\underline{P}(f|B) \leq -\underline{P}(-f|B) = \overline{P}(f|B)$

Proposition 3.5. *Consider a set of events* $\mathscr{B} \subseteq \mathscr{P}_\emptyset(\mathscr{X})$ *that is closed under finite unions and let* $\mathscr{F} \subseteq \mathscr{G}(\mathscr{X})$ *be a linear space of gambles such that* $\mathbb{I}_B f \in \mathscr{F}$ *and* $\mathbb{I}_B \in \mathscr{F}$ *for every* $f \in \mathscr{F}$ *and* $B \in \mathscr{B}$. *Now let* $\mathscr{C} := \{(f, B) \colon f \in \mathscr{F}, B \in \mathscr{B}\}$. *Then a conditional lower prevision* $\underline{P}$ *on* $\mathscr{C}$ *is coherent if and only if it is real-valued and satisfies LP1–LP4.*

Corollary 3.6. *A conditional lower prevision* $\underline{P}$ *on* $\mathscr{C}(\mathscr{X})$ *is coherent if and only if it is real-valued and satisfies LP1–LP4.*

Corollary 3.7. *Consider a linear space of gambles* $\mathscr{G} \subseteq \mathscr{G}(\mathscr{X})$ *that includes the constant gamble* 1. *A lower prevision* $\underline{P}$ *on* $\mathscr{G}$ *is then coherent if and only if it is real-valued and satisfies LP1–LP3.*

4. Epistemic independence

Having introduced our main tools for modelling uncertainty, the next step towards developing a notion of independent natural extension is to agree on what we mean by independence.

The approach that we adopt here is to define it as an assessment of mutual irrelevance. In particular, we say that X_1 and X_2 are independent if our uncertainty model for X_1 is not affected by conditioning on information about X_2, and vice versa. As we will see in Section 9, this definition can be applied to a probability measure, and then yields the usual notion of independence. However, and that is what makes this approach powerful and intuitive, it can just as easily be applied to lower previsions, sets of desirable gambles, or any other type of uncertainty model. This type of independence is usually referred to as epistemic independence. The aim of this section is to formalize this concept for the case of two variables, in terms of sets of desirable gambles and conditional lower previsions.

Consider two variables X_1 and X_2 where, for every $i \in \{1, 2\}$, X_i takes values x_i in a non-empty set $\mathscr{X}_i$ that may be uncountably infinite. We assume that X_1 and X_2 are logically independent, meaning that $X_1 = x_1$ and $X_2 = x_2$ are jointly possible, for all $x_1 \in \mathscr{X}_1$ and $x_2 \in \mathscr{X}_2$. The corresponding joint variable $X := (X_1, X_2)$ therefore takes values $x := (x_1, x_2)$ in $\mathscr{X}_1 \times \mathscr{X}_2$. In this context, whenever convenient, we will identify $B_1 \in \mathscr{P}_\emptyset(\mathscr{X}_1)$ with $B_1 \times \mathscr{X}_2$ and $B_2 \in \mathscr{P}_\emptyset(\mathscr{X}_2)$ with $\mathscr{X}_1 \times B_2$. For any two events $B_1 \in \mathscr{P}_\emptyset(\mathscr{X}_1)$ and $B_2 \in \mathscr{P}_\emptyset(\mathscr{X}_2)$, this allows us to use $B_1 \cap B_2$ as an intuitive alternative notation for $B_1 \times B_2$. Similarly, for any $i \in \{1, 2\}$, we identify $f \in \mathscr{G}(\mathscr{X}_i)$ with its cylindrical extension to $\mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$, defined by

$$f(x_1, x_2) := f(x_i) \text{ for all } x = (x_1, x_2) \in \mathscr{X}_1 \times \mathscr{X}_2.$$

In order to make this explicit, we will then often denote this cylindrical extension by $f(X_i)$. In this way, for example, for any $f \in \mathscr{G}(\mathscr{X}_2)$ and $B \in \mathscr{P}(\mathscr{X}_1)$, we can write $f(X_2)\mathbb{I}_B(X_1)$ to denote a gamble in $\mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$ whose value in (x_1, x_2) is equal to $f(x_2)$ if $x_1 \in B$ and equal to zero otherwise. Using these conventions, for any set of gambles $\mathscr{D}$ on $\mathscr{X}_1 \times \mathscr{X}_2$, we define the marginal models

$$\text{marg}_1(\mathscr{D}) := \{f \in \mathscr{G}(\mathscr{X}_1) \colon f(X_1) \in \mathscr{D}\} \text{ and } \text{marg}_2(\mathscr{D}) := \{f \in \mathscr{G}(\mathscr{X}_2) \colon f(X_2) \in \mathscr{D}\}$$

and, for any events $B_1 \in \mathscr{P}_\emptyset(\mathscr{X}_1)$ and $B_2 \in \mathscr{P}_\emptyset(\mathscr{X}_2)$, the conditional models

$$\text{marg}_1(\mathscr{D}|B_2) := \{f \in \mathscr{G}(\mathscr{X}_1) \colon f(X_1)\mathbb{I}_{B_2}(X_2) \in \mathscr{D}\} \text{ and } \text{marg}_2(\mathscr{D}|B_1) := \{f \in \mathscr{G}(\mathscr{X}_2) \colon f(X_2)\mathbb{I}_{B_1}(X_1) \in \mathscr{D}\}.$$

Conditioning and marginalisation both preserve coherence: if $\mathscr{D}$ is a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$, then $\text{marg}_1(\mathscr{D})$ and $\text{marg}_1(\mathscr{D}|B_2)$ are coherent sets of desirable gambles on $\mathscr{X}_1$, and $\text{marg}_2(\mathscr{D})$ and $\text{marg}_2(\mathscr{D}|B_1)$ are coherent sets of desirable gambles on $\mathscr{X}_2$.

That said, let us now recall our informal definition of epistemic independence, which was that the uncertainty model for X_1 is not affected by conditioning on information about X_2, and vice versa. In the context of sets of desirable gambles, this can now be formalized as follows:

$$\text{marg}_1(\mathscr{D}|B_2) = \text{marg}_1(\mathscr{D}) \text{ and } \text{marg}_2(\mathscr{D}|B_1) = \text{marg}_2(\mathscr{D}).$$

The only thing that is left to specify are the conditioning events B_1 and B_2 for which we want this condition to hold. We think that the most intuitive approach is to impose this for every $B_1 \in \mathscr{P}_\emptyset(\mathscr{X}_1)$ and $B_2 \in \mathscr{P}_\emptyset(\mathscr{X}_2)$, and will call this epistemic event-independence. However, this is not what is usually done. The conventional approach, which we will refer

to as epistemic atom-independence, is to focus on singleton events of the type $B_1 = \{x_1\}$ and $B_2 = \{x_2\}$; see for example References [9,22].[1]

The main reason why epistemic atom-independence is the conventional go-to definition for epistemic independence is that Walley adopted it in his seminal book [22, Section 9.2]. Walley seems to take this choice for granted; we assume that this is a consequence of his focus on conditional lower previsions whose conditioning events belong to a (finite number of) partition(s). The advantage of using such a partition—and a set of atoms in particular—is that it can be regarded as representing the possible outcomes of an experiment, allowing for a quite natural study of statistical inference and updating. Partitions are also essential if one wich to impose the controversial property of conglomerability [15]. Nevertheless, we consider this focus on partitions—which is inherent in Walley's approach—to be overly restrictive; we prefer having the option to condition on any possible event, especially since one can always zoom in on a particular partition whenever needed or convenient. This serves as a first reason why we prefer epistemic event-independence over epistemic atom-independence.

Other than that, as we will see further on, epistemic event-independence also has several technical advantages; in fact, this will be one of the main conclusions of this contribution. For now, however, we postpone this debate between event- and atom-independence by adopting a very general approach that subsumes the former two as special cases. In particular, for every $i \in \{1, 2\}$, we simply fix a generic set of conditioning events $\mathscr{B}_i \subseteq \mathscr{P}_\emptyset(\mathscr{X}_i)$. Epistemic atom-independence corresponds to choosing $\mathscr{B}_i = \mathscr{X}_i$, whereas epistemic event-independence corresponds to choosing $\mathscr{B}_i = \mathscr{P}_\emptyset(\mathscr{X}_i)$.

For sets of desirable gambles, this leads us to the following definition.

Definition 4.1. Let $\mathscr{D}$ be a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$. Then $\mathscr{D}$ is epistemically independent if, for any i and j such that $\{i, j\} = \{1, 2\}$:

$$\text{marg}_i(\mathscr{D}|B_j) = \text{marg}_i(\mathscr{D}) \text{ for all } B_j \in \mathscr{B}_j.$$

For coherent lower previsions, as a prerequisite for defining epistemic independence, we require that the domain $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$ is independent, by which we mean that for any i and j such that $\{i, j\} = \{1, 2\}$, any pair $(f_i, B_i) \in \mathscr{C}(\mathscr{X}_i)$ and any event $B_j \in \mathscr{B}_j$:

$$(f_i, B_i) \in \mathscr{C} \Leftrightarrow (f_i, B_i \cap B_j) \in \mathscr{C}. \tag{7}$$

Other than that, we impose no restrictions on $\mathscr{C}$; its elements $(f, B) \in \mathscr{C}$ are for example not restricted to the types that appear in Equation (7). As a result, the following definition of epistemic independence is applicable beyond the context of lower previsions. For example, by restricting the domain to indicators, we obtain a notion of epistemic independence that applies to conditional lower probabilities. A detailed discussion of these special cases, however, is deferred to Section 8.

Definition 4.2. Let $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$ be an independent domain. A coherent conditional lower prevision $\underline{P}$ on $\mathscr{C}$ is then epistemically independent if, for any i and j such that $\{i, j\} = \{1, 2\}$:

$$\underline{P}(f_i|B_i) = \underline{P}(f_i|B_i \cap B_j) \text{ for all } (f_i, B_i) \in \mathscr{C}(\mathscr{X}_i) \cap \mathscr{C} \text{ and } B_j \in \mathscr{B}_j.$$

Another important feature of this definition is that B_j is not only irrelevant to the unconditional lower previsions of local gambles f_i—in the sense that $\underline{P}(f_i) = \underline{P}(f_i|B_j)$—but also to their conditional local lower previsions—in the sense that $\underline{P}(f_i|B_i) = \underline{P}(f_i|B_i \cap B_j)$. This type of irrelevance is called h-irrelevance; see References [2,3]. Note however that this feature is optional within our framework; it only appears when $\mathscr{C}$ is sufficiently large. If instead $B_i = \mathscr{X}_i$ for all $(f_i, B_i) \in \mathscr{C}(\mathscr{X}_i) \cap \mathscr{C}$, then our definition reduces to the more simple requirement that $\underline{P}(f_i) = \underline{P}(f_i|B_j)$. The following example illustrates this subtle feature and also demonstrates the difference between epistemic atom- and event-independence.

Example 4.1. Let $\mathscr{X}_1 = \mathbb{R}$ and $\mathscr{X}_2 = \mathbb{N}$ and consider any coherent conditional lower prevision $\underline{P}$ on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$ that is epistemically independent, with $\mathscr{C}$ an independent domain. For any $(f_1, B_1) \in \mathscr{C}(\mathscr{X}_1) \cap \mathscr{C}$ and $B_2 \in \mathscr{B}_2$, it then follows from Definition 4.2 that $\underline{P}(f_1|B_1) = \underline{P}(f_1|B_1 \cap B_2)$. To make this more concrete, we now consider several examples.

We first consider the most powerful case, where the domain $\mathscr{C}$ is equal to $\mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$—the largest possible domain—and where the type of independence that is considered is event-independence. Now let $f_1 := \sin(X_1)$ and let $B_2 := \{2n : n \in \mathbb{N}\}$ be the event that X_2 is even. For $B_1 = \mathscr{X}_1$, we then find that $\underline{P}(\sin(X_1)) = \underline{P}(\sin(X_1)|X_2 \text{ even})$, meaning that conditioning on the event that X_2 is even has no effect on the lower prevision of $\sin(X_1)$. In much the same way, for $B_1 = \mathbb{R}_{\geq 0}$, we find that $\underline{P}(\sin(X_1)|X_1 \geq 0) = \underline{P}(\sin(X_1)|X_1 \geq 0 \text{ and } X_2 \text{ even})$, which means that the conditional lower prevision of $\sin(X_1)$

---

[1] Readers that are familiar with some of my previous work [3,4] may notice that I have changed terminology: what I now call epistemic event- and atom-independence, I previously referred to as epistemic subset- and value-independence. This new terminology was suggested to me by an anonymous reviewer, and I could not but agree that it indeed better reflects the meaning of the respective concepts.

given $X_1 \geq 0$ does not change if we additionally condition on the event that X_2 is even. This second example provides a nice illustration of the fact that our definition of independence imposes mutual h-irrelevance rather than mutual irrelevance.

If we shrink the domain $\mathscr{C}$ sufficiently, to the extent that every $(f_1, B_1) \in \mathscr{C}(\mathscr{X}_1) \cap \mathscr{C}$ is of the form $(f_1, \mathscr{X}_1)$, then the added value of h-irrelevance disappears because our definition then only imposes assessments of the type $\underline{P}(f_1) = \underline{P}(f_1|B_2)$, such as, for example, the assessement $\underline{P}(\sin(X_1)) = \underline{P}(\sin(X_1)|X_2 \text{ even})$ that we have seen before.

The effect of replacing event-independence with atom-independence is also quite substantial. In particular, since $\mathscr{B}_2$ then changes from $\mathscr{P}_\emptyset(\mathbb{N})$ to $\mathbb{N}$, we can then no longer condition on the event $B_2 := \{2n: n \in \mathbb{N}\}$ that X_2 is even, because that event is not a singleton. Instead, for atom-indepence, we can only condition on events B_2 of the form $\{x_2\}$, such as, for example, the assessement $\underline{P}(\sin(X_1)) = \underline{P}(\sin(X_1)|X_2 = 5)$. $\diamond$

5. The independent natural extension

All of that said, we are now finally ready to introduce our central object of interest, which is the *independent natural extension*. Basically, the question to which this concept provides an answer is: given two local uncertainty models and an assessment of epistemic independence, what then should be the corresponding joint model? The answer depends on the specific framework that is being considered.

Within the framework of sets of desirable gambles, the local uncertainty models are coherent sets of desirable gambles. In particular, for each $i \in \{1, 2\}$, we are given a coherent set of desirable gambles $\mathscr{D}_i$ on $\mathscr{X}_i$. The aim is to combine these local models with an assessment of epistemic independence to obtain a coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}_1 \times \mathscr{X}_2$. The first requirement on $\mathscr{D}$, therefore, is that it should have $\mathscr{D}_1$ and $\mathscr{D}_2$ as its marginals, in the sense that $\text{marg}_i(\mathscr{D}) = \mathscr{D}_i$ for all $i \in \{1, 2\}$. The second is that $\mathscr{D}$ should be epistemically independent. If both requirements are met, $\mathscr{D}$ is called an independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$. The most conservative among these independent products is called the independent natural extension.

Definition 5.1. An independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$ is an epistemically independent coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}_1 \times \mathscr{X}_2$ that has $\mathscr{D}_1$ and $\mathscr{D}_2$ as its marginals.

Definition 5.2. The independent natural extension of $\mathscr{D}_1$ and $\mathscr{D}_2$ is the smallest independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$.

If all we know is that $\mathscr{D}$ is epistemically independent and has $\mathscr{D}_1$ and $\mathscr{D}_2$ as its marginal models, then the safest choice for $\mathscr{D}$—the only choice that does not require any additional assessments—is their independent natural extension, provided of course that it exists. In order to show that it always does, we let

$$\mathscr{D}_1 \otimes \mathscr{D}_2 := \mathscr{E}\left(\mathscr{A}_{1\to 2} \cup \mathscr{A}_{2\to 1}\right), \tag{8}$$

with

$$\mathscr{A}_{1\to 2} := \left\{ f_2(X_2)\mathbb{I}_{B_1}(X_1): f_2 \in \mathscr{D}_2, B_1 \in \mathscr{B}_1 \cup \{\mathscr{X}_1\} \right\} \tag{9}$$

and

$$\mathscr{A}_{2\to 1} := \left\{ f_1(X_1)\mathbb{I}_{B_2}(X_2): f_1 \in \mathscr{D}_1, B_2 \in \mathscr{B}_2 \cup \{\mathscr{X}_2\} \right\}. \tag{10}$$

The following result establishes that $\mathscr{D}_1 \otimes \mathscr{D}_2$ is the independent natural extension of $\mathscr{D}_1$ and $\mathscr{D}_2$.

Theorem 5.1. $\mathscr{D}_1 \otimes \mathscr{D}_2$ *is the independent natural extension of* $\mathscr{D}_1$ *and* $\mathscr{D}_2$.

Similar concepts can be defined for conditional lower previsions as well. In that case, the local uncertainty models are coherent conditional lower previsions. In particular, for every $i \in \{1, 2\}$, we are given a coherent conditional lower prevision $\underline{P}_i$ on some freely chosen local domain $\mathscr{C}_i \subseteq \mathscr{C}(\mathscr{X}_i)$. Note that this freedom implies that $\underline{P}_i$ can also be an unconditional; this corresponds to choosing $\mathscr{C}_i := \{(f_i, \mathscr{X}_i): f_i \in \mathscr{G}_i\}$ for some $\mathscr{G}_i \subseteq \mathscr{G}(\mathscr{X}_i)$. In any case, the aim is now to construct an epistemically independent coherent conditional lower prevision $\underline{P}$ on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$ that has $\underline{P}_1$ and $\underline{P}_2$ as its marginals, in the sense that $\underline{P}$ coincides with $\underline{P}_1$ and $\underline{P}_2$ on their local domain: $\underline{P}(f_i|B_i) = \underline{P}_i(f_i|B_i)$ for all $i \in \{1, 2\}$ and $(f_i, B_i) \in \mathscr{C}_i$. As before, a model that meets these criteria is then called an independent product, and the most conservative among them is called the independent natural extension. Clearly, in order for these notions to make sense, the global domain $\mathscr{C}$ must at least include the local domains $\mathscr{C}_1$ and $\mathscr{C}_2$ and must furthermore be independent in the sense of Equation (7). The definitions and results below take this for granted.

Definition 5.3. An independent product of $\underline{P}_1$ and $\underline{P}_2$ is an epistemically independent coherent conditional lower prevision on $\mathscr{C}$ that has $\underline{P}_1$ and $\underline{P}_2$ as its marginals.

Definition 5.4. The independent natural extension of $\underline{P}_1$ and $\underline{P}_2$ is the point-wise smallest independent product of $\underline{P}_1$ and $\underline{P}_2$.

Here too, if all we know is that $\underline{P}$ is epistemically independent and has $\underline{P}_1$ and $\underline{P}_2$ as its marginal models, then the safest choice for $\underline{P}$—the only choice that does not require any additional assessments—is the independent natural extension, provided that it exists. The following result establishes that it does, by showing that it is a restriction of the operator $\underline{P}_1 \otimes \underline{P}_2$, defined by

$$(\underline{P}_1 \otimes \underline{P}_2)(f|B) := \underline{P}_{\mathscr{D}}(f|B) \text{ for all } (f, B) \in \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2), \text{ with } \mathscr{D} = \mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2). \tag{11}$$

Theorem 5.2. *The independent natural extension of $\underline{P}_1$ and $\underline{P}_2$ is the restriction of $\underline{P}_1 \otimes \underline{P}_2$ to $\mathscr{C}$.*

Interestingly, as can be seen from this result, the choice of the joint domain $\mathscr{C}$ does not affect the resulting independent natural extension, in the sense that any $\mathscr{C}$ that includes (f, B) will lead to the same value of $(\underline{P}_1 \otimes \underline{P}_2)(f|B)$. For that reason, we will henceforth assume without loss of generality that $\mathscr{C} = \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$.

6. On the choice of conditioning events

The fact that the existence results in the previous section are valid regardless of the choice of $\mathscr{B}_1$ and $\mathscr{B}_2$ should not be taken to mean that this choice does not affect the model. In some cases, it most definitely does. In the remainder of this contribution, we will study the extent to which it does, and how it affects the properties of the resulting notion of independent natural extension.

As a first observation, we note that larger sets of conditioning events correspond to stronger assessments of epistemic independence, and therefore lead to more informative joint models. For example, as can be seen from Equations (8)–(10), adding events to $\mathscr{B}_1$ and $\mathscr{B}_2$ leads to a larger—more informative—set of desirable gambles $\mathscr{D}_1 \otimes \mathscr{D}_2$. Similarly, as can be seen from Equation (11), it leads to a joint lower prevision that is higher—and therefore again more informative. There is one important exception to this observation though, which occurs when we add conditioning events that are a finite disjoint union of other conditioning events. In that case, the resulting notion of independent natural extension does not change.

Proposition 6.1. *For each $i \in \{1, 2\}$, let $\mathscr{B}'_i$ be a superset of $\mathscr{B}_i$ that consists of finite disjoint unions of events in $\mathscr{B}_i$. Replacing $\mathscr{B}_1$ by $\mathscr{B}'_1$ and $\mathscr{B}_2$ by $\mathscr{B}'_2$ then has no effect on the resulting independent natural extension $\mathscr{D}_1 \otimes \mathscr{D}_2$ or $\underline{P}_1 \otimes \underline{P}_2$.*

As a particular case of this result, it follows that if $\mathscr{B}_i$ is a finite partition of $\mathscr{X}_i$, we can replace it by the generated algebra—minus the empty event. As an even more particular case, if $\mathscr{X}_1$ and $\mathscr{X}_2$ are finite, we find that epistemic atom- and event-independence lead to the same notion of independent natural extension. For that reason, in the finite case, it does not really matter which of these two types of epistemic independence is adopted.

In the infinite case though, we will see that the difference does matter, which requires one to choose between epistemic atom- and event-independence. For lower previsions, Miranda and Zaffalon [16] recently adopted epistemic atom-independence in combination with Walley-coherence. Unfortunately, they found that the corresponding notion of independent natural extension does not always exist. They also considered the combination of epistemic atom-independence with Williams-coherence, and argued that the resulting model was too weak. For the case of lower probabilities, Vicig [21] adopted epistemic event-independence in combination with Williams-coherence, showed that the corresponding independent natural extension always exists, and proved that it satisfies factorisation properties. Our results so far can be regarded as a generalisation of the existence results of Vicig [21]. As we are about to show, his factorisation results can be generalised as well.

7. Factorisation and external additivity

When $\mathscr{X}_1$ and $\mathscr{X}_2$ are finite, the independent natural extension of two lower previsions $\underline{P}_1$ and $\underline{P}_2$ is well-known to satisfy the properties of factorisation and external additivity [10]. Factorisation, on the one hand, states that

$$(\underline{P}_1 \otimes \underline{P}_2)(gh) = \underline{P}_1(g\underline{P}_2(h)) = \begin{cases} \underline{P}_1(g)\underline{P}_2(h) & \text{if } \underline{P}_2(h) \geq 0 \\ \overline{P}_1(g)\underline{P}_2(h) & \text{if } \underline{P}_2(h) \leq 0, \end{cases} \tag{12}$$

where g is a non-negative gamble on $\mathscr{X}_1$, h is a gamble on $\mathscr{X}_2$ and $\overline{P}_1(g) := -\underline{P}_1(-g)$. By symmetry, the role of 1 and 2 can of course be reversed. External additivity, on the other hand, states that

$$(\underline{P}_1 \otimes \underline{P}_2)(f + h) = \underline{P}_1(f) + \underline{P}_2(h) \tag{13}$$

where f and h are gambles on $\mathscr{X}_1$ and $\mathscr{X}_2$, respectively.

Compared to the properties that are satisfied by the joint expectation of a product measure of two probability measures, these notions of factorisation and external additivity are rather weak. For example, for a product measure, additivity is not 'external', in the sense that f and h do not have to be defined on separate variables, nor does factorisation require g to be non-negative. Nevertheless, even in this weaker form, these properties remain of crucial practical importance. For example, as explained in Reference [10], factorisation properties such as Equation (12)—when applied to more than two variables—are sufficient in order to establish laws of large numbers for lower previsions [8]. As another example, in the context of credal networks, which are Bayesian networks whose local models are partially specified, properties such as Equations (12) and (13) turned out to be the key to the development of efficient inference algorithms; see for example References [3,6,7]. Any notion of independent natural extension that aims to extend such algorithms to infinite spaces, therefore, should preserve some suitable version of Equations (12) and (13).

The aim of this section is to study the extent to which these equations are satisfied by the notion of independent natural extension that was developed in this paper. As we will see, the answer ends up being surprisingly positive.

For all $i \in \{1, 2\}$, let $\underline{P}_i$ be a coherent conditional lower prevision on $\mathscr{C}_i \subseteq \mathscr{C}(\mathscr{X}_i)$, let $\underline{E}_i$ be its natural extension to $\mathscr{C}(\mathscr{X}_i)$, and let $\mathscr{B}_i$ be a subset of $\mathscr{P}_\emptyset(\mathscr{X}_i)$. The independent natural extension of $\underline{P}_1$ and $\underline{P}_2$ then satisfies the following three properties, the first of which implies the other two as special cases.

Theorem 7.1. *Let $\{i, j\} = \{1, 2\}$. For any $f \in \mathscr{G}(\mathscr{X}_i)$, $h \in \mathscr{G}(\mathscr{X}_j)$ and $\mathscr{B}_i$-measurable $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_i)$, we then have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) = \underline{E}_i\big(f + g\underline{E}_j(h)\big).$$

Corollary 7.2 *(Factorisation). Let $\{i, j\} = \{1, 2\}$. For any $h \in \mathscr{G}(\mathscr{X}_j)$ and any $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_i)$ that is $\mathscr{B}_i$-measurable, we then have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(gh) = \underline{E}_i\big(g\underline{E}_j(h)\big) = \begin{cases} \underline{E}_i(g)\underline{E}_j(h) & \text{if } \underline{E}_j(h) \geq 0; \\ \overline{E}_i(g)\underline{E}_j(h) & \text{if } \underline{E}_j(h) \leq 0. \end{cases}$$

Corollary 7.3 *(External additivity). For any $f \in \mathscr{G}(\mathscr{X}_1)$ and $h \in \mathscr{G}(\mathscr{X}_2)$, we have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(f + h) = \underline{E}_1(f) + \underline{E}_2(h).$$

In each of these results, if the local domains $\mathscr{C}_1$ and $\mathscr{C}_2$ are sufficiently large—that is, if they include the gambles that appear in the statement of the results—it follows from Proposition 3.3 that $\underline{E}_i$ and $\underline{E}_j$ can be replaced by $\underline{P}_i$ and $\underline{P}_j$, respectively, and similarly for $\overline{E}_i$ and $\overline{P}_i$.

That said, let us now go back to the question of whether or not Equations (12) and (13) can be generalised to the case of infinite spaces. For the case of external additivity, it clearly follows from Corollary 7.3 that the answer is fully positive. Furthermore, this conclusion holds regardless of our choice for $\mathscr{B}_1$ and $\mathscr{B}_2$; they can even be empty. For factorisation, the answer does depend on $\mathscr{B}_1$ and $\mathscr{B}_2$. If we adopt epistemic event-independence—that is, if we choose $\mathscr{B}_1 = \mathscr{P}_\emptyset(\mathscr{X}_1)$ and $\mathscr{B}_2 = \mathscr{P}_\emptyset(\mathscr{X}_2)$—it follows from Corollaries 2.3 and 7.2 that the answer is again fully positive, because $\mathscr{P}_\emptyset(\mathscr{X}_i)$-measurability then holds trivially. If $\mathscr{B}_1 \cup \{\emptyset\}$ and $\mathscr{B}_2 \cup \{\emptyset\}$ are sigma fields, the answer remains fairly positive as well, because Proposition 2.1 then implies that it suffices for g to be measurable in the usual, measure-theoretic sense.

Example 7.1. Let $\mathscr{X}_1 = \mathscr{X}_2 = \mathbb{R}$ and let $\mathscr{B}_1 = \mathscr{B}_2 = \mathscr{B}$, with $\mathscr{B} = \mathscr{B}^* \setminus \{\emptyset\}$ and with $\mathscr{B}^*$ the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}$. Furthermore, let $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$ be the indicator of the nonnegative reals, defined for all $x_1 \in \mathbb{R}$ by $g(x_1) := 1$ if $x_1 \geq 0$ and $g(x_1) := 0$ otherwise, and let $h \in \mathscr{G}(\mathscr{X}_2)$ be defined by $h(x_2) := x_2^3$ for all $x_2 \in \mathbb{R}$. We then know from Example 2.3 that g and h are both Lebesgue-measurable and therefore also $\mathscr{B}$-measurable. Therefore, for any two coherent conditional lower previsions $\underline{P}_1$ on $\mathscr{C}(\mathscr{X}_1)$ and $\underline{P}_2$ on $\mathscr{C}(\mathscr{X}_2)$, Corollary 7.2 implies that

$$(\underline{P}_1 \otimes \underline{P}_2)(gh) = \begin{cases} \underline{P}_1(g)\underline{P}_2(h) & \text{if } \underline{P}_2(h) \geq 0; \\ \overline{P}_1(g)\underline{P}_2(h) & \text{if } \underline{P}_2(h) \leq 0. \end{cases} \quad \lozenge$$

However, these positive conclusions do not apply if we adopt epistemic atom-independence—that is, if we choose $\mathscr{B}_1 = \mathscr{X}_1$ and $\mathscr{B}_2 = \mathscr{X}_2$—because our factorisation result then requires g to be $\mathscr{X}_i$-measurable, which, as we know from Example 2.2, is a rather strong requirement that easily fails. Nevertheless, the factorisation properties that do hold for atom-independence are stronger than what is suggested in Reference [16]. In order to illustrate this, we consider the following adaptation of their Example 4.

Example 7.2. Let $\mathscr{X}_1 = \mathscr{X}_2 = \mathbb{N}$ and let $\underline{P}_1$ and $\underline{P}_2$ be coherent lower previsions on $\mathscr{G}(\mathscr{X}_1)$ and $\mathscr{G}(\mathscr{X}_2)$, respectively. Let $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$ and $h \in \mathscr{G}_{\geq 0}(\mathscr{X}_2)$ be defined by $g(x_1) := \frac{1}{x_1}$ and $h(x_2) := \frac{1}{x_2}$ for all $x_1 \in \mathscr{X}_1$ and $x_2 \in \mathscr{X}_2$. We then know from Example 2.1 that the gambles g and h are both $\mathscr{P}_\emptyset(\mathbb{N})$-measurable and $\mathbb{N}$-measurable. Furthermore, since g and h are nonnegative, LP1 implies that $\underline{P}_1(g) \geq 0$ and $\underline{P}_2(h) \geq 0$. Hence, regardless of whether we adopt epistemic event- or atom-independence, we can apply Corollary 7.2 to find that $(\underline{P}_1 \otimes \underline{P}_2)(gh) = \underline{P}_1(g)\underline{P}_2(h)$. $\quad \lozenge$

For readers that are familiar with the work of Miranda and Zaffalon [16], the final conclusion of this example may seem surprising at first, because in their Example 4, Miranda and Zaffalon show that for the same two gambles g and h, for atom-independence and Williams-coherence, the independent natural extension assigns lower prevision zero to gh, regardless of the chosen local models $\underline{P}_1$ and $\underline{P}_2$. In contrast, our example above concludes that this lower prevision is equal to the product of $\underline{P}_1(g)$ and $\underline{P}_2(h)$. This apparent contradiction is a consequence of the fact that Miranda and Zaffalon do not require the independent natural extension of $\underline{P}_1$ and $\underline{P}_2$ to have $\underline{P}_1$ and $\underline{P}_2$ as its marginals. Instead—using their notation and terminology—they only require weak coherence with $\underline{P}_1(\cdot|X_2)$ and $\underline{P}_2(\cdot|X_1)$; see Reference [16] for more information. As we can see here, this leads to a notion of independent natural extension that satisfies fewer factorisation properties. Our Definitions 5.3 and 5.4 avoid this, by explicitly imposing that the independent natural extension of $\underline{P}_1$ and $\underline{P}_2$ should have $\underline{P}_1$ and $\underline{P}_2$ as its marginals.

Still, even with our strengthened definition, the issue remains that for atom-independence, our factorisation result requires the stringent assumption that g is $\mathscr{X}_i$-measurable. This issue is fundamental because, as our next example demonstrates, it is not just a feature of Corollary 7.2 but rather an inherent property of atom-independence: epistemic atom-independence indeed leads to weaker factorisation properties.

Example 7.3. Let $\mathscr{X}_1 = \mathscr{X}_2 = \mathbb{N}$ and let $g = \mathbb{I}_{\text{odd}} \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$ and $h = \mathbb{I}_{\text{even}} \in \mathscr{G}_{\geq 0}(\mathscr{X}_2)$, with $\mathbb{I}_{\text{odd}}$ and $\mathbb{I}_{\text{even}}$ defined as in Example 2.2. Furthermore, let $\underline{P}_1$ be a coherent lower prevision on $\mathscr{G}(\mathscr{X}_1)$ and let $\underline{P}_2$ be a coherent lower prevision on $\mathscr{G}(\mathscr{X}_2)$. Since g and h are nonnegative, LP1 then implies that $\underline{P}_1(g) \geq 0$ and $\underline{P}_2(h) \geq 0$. Therefore, and because we know from Example 2.2 that g and h are $\mathscr{P}_\emptyset(\mathbb{N})$-measurable, we can apply Corollary 7.2 to find that for event-independence: $(\underline{P}_1 \otimes \underline{P}_2)(gh) = \underline{P}_1(g)\underline{P}_2(h)$. However, unfortunately, this corollary cannot be applied for atom-independence, because we know from Example 2.2 that g and h are not $\mathbb{N}$-measurable.

Of course, one could still believe that factorisation can be established in some other way, and that it is simply Corollary 7.2 that is lacking in power, rather than the concept of atom-independence itself. This is however not the case: for atom-independence, as we will demonstrate in Example 9.1, $(\underline{P}_1 \otimes \underline{P}_2)(\mathbb{I}_{\text{odd}}(X_1)\mathbb{I}_{\text{even}}(X_2))$ can be strictly smaller than $\underline{P}_1(\mathbb{I}_{\text{odd}}(X_1))\underline{P}_2(\mathbb{I}_{\text{even}}(X_2))$ $\Diamond$

Because of these weak factorisation properties, we think that for the case of infinite spaces, when it comes to choosing between epistemic atom- and event-independence, the latter should be preferred over the former. That is not the only reason though. There is also a second, closely related reason, which is that event-independence leads to much more informative inferences. However, in order to explain and demonstrate that, we first need to establish a connection between conditional lower previsions, probabilities and expectations, which is what we now set out to do.

8. Connecting lower previsions with expectations and probabilities

The key to understanding the connection between lower previsions, expectations and probabilities is to consider conditional lower previsions that are self-conjugate, in the sense that they coincide with their corresponding upper prevision. In that case, we simply refer to them as *conditional previsions* and denote them by P instead of $\underline{P}$.

Definition 8.1 (*Conditional prevision*). A conditional prevision P on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is a conditional lower prevision on $\mathscr{C}$ that is self-conjugate, in the sense that

$$(-f, B) \in \mathscr{C} \text{ and } P(f|B) = -P(-f|B) \text{ for all } (f, B) \in \mathscr{C}. \tag{14}$$

Unconditional previsions correspond to a special case. First, if $B = \mathscr{X}$, then similarly to what we did for conditional lower previsions, we adopt the shorthand notation $P(f) := P(f|\mathscr{X})$ and call $P(f)$ the prevision of f. Second, if $B = \mathscr{X}$ for all $(f, B) \in \mathscr{C}$, meaning that there is some $\mathscr{G} \subseteq \mathscr{G}(\mathscr{X})$ such that $\mathscr{C} = \{(f, \mathscr{X}): f \in \mathscr{G}\}$, we regard P as an operator on $\mathscr{G}$ and then call P a(n unconditional) *prevision*.

If a conditional prevision is coherent, we refer to it as a *conditional linear prevision*. Similarly, coherent (unconditional) previsions are called *linear previsions*.

Definition 8.2 (*Conditional linear prevision*). A conditional linear prevision P on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is a coherent conditional prevision on $\mathscr{C}$.

The reason for this terminology, quite obviously, is that conditional linear previsions can be shown to be linear operators. In fact, they satisfy various other convenient properties as well, which, for the sake of completeness, are listed below.

Proposition 8.1. *Let P be a conditional linear prevision on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$. Then for any two gambles $f, g \in \mathscr{G}(\mathscr{X})$, any two events $A, B \in \mathscr{P}_\emptyset(\mathscr{X})$, any real number $\lambda \in \mathbb{R}$ and any sequence of gambles $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathscr{G}(\mathscr{X})$, whenever the involved conditional previsions are well-defined—that is, if the arguments belong to their domain—we have that*

P1: $P(f|B) \geq \inf_{x \in B} f(x)$ *[boundedness]*

P2: $P(\lambda f|B) = \lambda P(f|B)$ *[homogeneity]*

P3: $P(f + g|B) = P(f|B) + P(g|B)$ *[additivity]*

P4: $P(\mathbb{I}_B f|A) = P(f|A \cap B)P(B|A)$ if $A \cap B \neq \emptyset$ *[Bayes rule]*

P5: $\lim_{n \to +\infty} \sup|f - f_n| = 0 \Rightarrow \lim_{n \to +\infty} P(f_n|B) = P(f|B)$ *[uniform continuity]*

P6: $P(f + \lambda|B) = P(f|B) + \lambda$ *[constant additivity]*

P7: $f(x) \geq g(x)$ for all $x \in B \Rightarrow P(f|B) \geq P(g|B)$ *[monotonicity]*

Proposition 8.2. *Consider a set of events* $\mathcal{B} \subseteq \mathscr{P}_{\emptyset}(\mathcal{X})$ *that is closed under finite unions and let* $\mathcal{F} \subseteq \mathcal{G}(\mathcal{X})$ *be a linear space of gambles such that* $\mathbb{I}_B f \in \mathcal{F}$ *and* $\mathbb{I}_B \in \mathcal{F}$ *for every* $f \in \mathcal{F}$ *and* $B \in \mathcal{B}$. *Now let* $\mathcal{C} := \{(f, B) \colon f \in \mathcal{F}, B \in \mathcal{B}\}$. *Then a conditional prevision* P *on* $\mathcal{C}$ *is a conditional linear prevision on* $\mathcal{C}$ *if and only if it is real-valued and satisfies P1–P4.*

Corollary 8.3. *A conditional prevision* P *on* $\mathcal{C}(\mathcal{X})$ *is linear if and only if it is real-valued and satisfies P1–P4.*

Corollary 8.4. *Consider a linear space of gambles* $\mathcal{G} \subseteq \mathcal{G}(\mathcal{X})$ *that includes the constant gamble* 1. *A prevision* P *on* $\mathcal{G}$ *is then linear if and only if it is real-valued and satisfies P1–P3.*

By comparing these properties with the ones in Section 3, we see that the linearity of conditional linear previsions—the fact that they satisfy P2 and P3—is their most important property, in the sense that it distinguishes them from general coherent conditional lower previsions. Furthermore, this property is also what allows us to establish a connection with expectations. In particular, if we allow ourselves a small leap of faith here, then since expectations are well known to be linear, the fact that conditional linear previsions are also linear suggests that we can simply interpret them as conditional expectations.

In order to clarify why this is more than just intuition, it is instrumental to restrict the domain of P to elements that are of the form $(\mathbb{I}_A, B)$, where $\mathbb{I}_A$ is the indicator of an event A, and to then follow de Finetti [11] in adopting the alternative notation $P(A|B) := P(\mathbb{I}_A|B)$. As this notation already suggests, $P(A|B)$ can then be interpreted as the probability of A conditional on B. This interpretation is furthermore mathematically sound, because the obtained objects $P(A|B)$ can be shown to satisfy all the essential properties of conditional probabilities, including finite—but not necessarily countable—additivity and, if $B \neq \emptyset$, Bayes's rule. Hence, by restricting the domain of a conditional linear prevision P to elements of the form $(\mathbb{I}_A, B)$, we obtain a conditional probability measure. The original unrestricted conditional linear prevision P is then the conditional expectation operator that corresponds to this conditional probability measure. Here too, this connection is not merely intuitive, but can be made mathematically rigorous. A detailed account of the mathematics behind these connections, however, is beyond the scope of this contribution. For more information about (finitely additive) conditional probability measures, the interested reader is referred to the work of Dubins [13].

For our present purposes, it suffices to know that conditional linear previsions can indeed be interpreted as conditional expectation operators and that conditional probabilities are conditional linear previsions whose domain contains only—or is restricted to—elements of the form $(\mathbb{I}_A, B)$. A similar observation applies to unconditional expectation operators and probability measures, with the role of the conditional linear prevision now taken up by an unconditional one.

Since conditional linear previsions are themselves a special (self-conjugate) case of coherent conditional lower previsions, we conclude that conditional expectations and conditional probability measures can both be regarded as special cases of conditional coherent lower previsions. Similarly, unconditional expectations and probability measures are special cases of coherent lower previsions. However, the connection goes much further, because conditional linear previsions are not just a special case of coherent conditional lower previsions: they can also be used to characterise them.

Proposition 8.5. *A conditional lower prevision* $\underline{P}$ *on* $\mathcal{C} \subseteq \mathcal{C}(\mathcal{X})$ *is coherent if and only if there is a non-empty set* $\mathbb{P}^*$ *of conditional linear previsions on* $\mathcal{C}(\mathcal{X})$ *such that*

$$\underline{P}(f|B) = \inf\{P(f|B) \colon P \in \mathbb{P}^*\} \text{ for all } (f, B) \in \mathcal{C}. \tag{15}$$

The same is true if the infimum in this expression is replaced by a minimum.

This well-known result is essentially due to Williams [24,25] The result is fundamental, because it provides coherent conditional lower previsions with a second, alternative interpretation. Indeed, because of Proposition 8.5, a conditional lower prevision is not only a supreming buying price; alternatively, it can also be regarded as an infimum of conditional previsions. Since—as we have just seen—conditional previsions can themselves be interpreted as conditional expectations, this implies that coherent conditional lower previsions can be interpreted as lower envelopes of expectations, often referred to as lower expectations.

This interpretation can also be used to develop an alternative characterisation for the natural extension $\underline{E}$ of a coherent conditional lower prevision $\underline{P}$ on $\mathcal{C} \subseteq \mathcal{C}(\mathcal{X})$. In order to do that, we let $\mathbb{P}$ be the set of all conditional linear previsions on $\mathcal{C}(\mathcal{X})$ and then let

$$\mathbb{P}_{\underline{P}} := \{P \in \mathbb{P} \colon P(f|B) \geq \underline{P}(f|B) \text{ for all } (f, B) \in \mathscr{C}\} \tag{16}$$

be the subset that dominates $\underline{P}$. The natural extension $\underline{E}$ is then the lower envelope of $\mathbb{P}_{\underline{P}}$ and, similarly, $\overline{E}$ is its upper envelope.

Proposition 8.6. *Let $\underline{P}$ be a coherent conditional lower prevision $\underline{P}$ on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$, let $\underline{E}$ be its natural extension to $\mathscr{C}(\mathscr{X})$ and let $\overline{E}$ be the corresponding conditional upper prevision on $\mathscr{C}(\mathscr{X})$. Then $\mathbb{P}_{\underline{P}} \neq \emptyset$ and, for all $(f, B) \in \mathscr{C}(\mathscr{X})$:*

$$\underline{E}(f|B) = \min\{P(f|B) \colon P \in \mathbb{P}_{\underline{P}}\} \text{ and } \overline{E}(f|B) = \max\{P(f|B) \colon P \in \mathbb{P}_{\underline{P}}\}. \tag{17}$$

Also, for any $(f, B) \in \mathscr{C}(\mathscr{X})$ and $\alpha \in [\underline{E}(f|B), \overline{E}(f|B)]$, there is some $P \in \mathbb{P}_{\underline{P}}$ such that $P(f|B) = \alpha$.

The final connection that remains to be discussed is that between conditional lower previsions and conditional lower probabilities. However, since conditional lower previsions are lower envelopes of conditional linear previsions, and since conditional probability measures are conditional linear previsions whose domain is restricted to elements of the form $(\mathbb{I}_A, B)$, this connection is immediate: conditional lower probabilities are simply conditional lower previsions whose domain is restricted to elements of the form $(\mathbb{I}_A, B)$. In that case, in order to emphasize this, we adopt $\underline{P}(A|B)$ as an intuitive alternative notation for $\underline{P}(\mathbb{I}_A|B)$.

9. Special cases of the independent natural extension

Now that we have established that expectations, lower expectations, probabilities and lower probabilities are indeed all special cases of lower previsions, we can come back to our claim at the end of Section 3, which was that our results can be applied to—and interpreted in terms of—these special cases as well.

Applying our results to the case of lower expectations is straightforward. Mathematically, nothing changes. The only difference is that the local conditional lower previsions that we start from are now interpreted—or defined—as lower bounds on expectations, and similarly for the independent natural extension that is derived from them. For lower probabilities, it suffices to restrict the domain of the local models $\underline{P}_1$ and $\underline{P}_2$ to elements of the form $(\mathbb{I}_{A_1}, B_1)$ and $(\mathbb{I}_{A_2}, B_2)$, respectively, and to similarly restrict the domain $\mathscr{C}$ of $\underline{P}_1 \otimes \underline{P}_2$ to elements of the form $(\mathbb{I}_A, B)$. Other than that, here too, the only difference is the interpretation. For results that are tailored to this specific case, we refer the interested reader to the work of Vicig [21], who focused on the independent natural extension for lower probabilities, but within a more general context that allows for more than two variables. As explained before, our results are basically a generalisation of his, extending them from lower probabilities to conditional lower previsions.

Once we interpret the independent natural extension as a lower expectation or a lower probability, it makes sense to consider the set of conditional linear previsions $\mathbb{P}_{\underline{P}_1 \otimes \underline{P}_2}$ that dominates the independent natural extension $\underline{P}_1 \otimes \underline{P}_2$, and to then interpret the latter in terms of the elements of the former. An essential observation here is that the elements of $\mathbb{P}_{\underline{P}_1 \otimes \underline{P}_2}$ need not be independent themselves, nor is this the case for its extreme points. Consequently, the independent natural extension is not in general a lower envelope of precise independent models [22, Section 9.3.4]. For example, for any $f_1 \in \mathscr{G}(\mathscr{X}_1)$ and $B_2 \in \mathscr{B}_2$, since $\underline{P}_1 \otimes \underline{P}_2$ is an independent product of $\underline{P}_1$ and $\underline{P}_2$, it follows from Definitions 5.3 and 4.2 that

$$(\underline{P}_1 \otimes \underline{P}_2)(f_1|B_2) = (\underline{P}_1 \otimes \underline{P}_2)(f_1) = \underline{P}_1(f_1).$$

However, for $P \in \mathbb{P}_{\underline{P}_1 \otimes \underline{P}_2}$, this does not necessarily imply that $P(f_1|B_2) = P(f_1)$. Instead, the only constraint that is imposed on $P(f_1|B_2)$ and $P(f_1)$ is that they both belong to $[\underline{P}_1(f_1), \overline{P}_1(f_1)]$:

$$\underline{P}_1(f_1) \leq P(f_1|B_2) \leq \overline{P}_1(f_1) \text{ and } \underline{P}_1(f_1) \leq P(f_1) \leq \overline{P}_1(f_1).$$

This feature is an essential aspect of epistemic independence: it imposes independence on the uncertainty model itself. If this uncertainty model is a set of conditional expectations, then epistemic independence imposes constraints on this set—in this case, on the resulting lower and upper expectations—but not on the individual expectations themselves. Similarly, for lower probabilities, epistemic independence does not require $P(A_1|B_2)$ and $P(A_1)$ to be equal, but only requires that

$$\underline{P}_1(A_1) \leq P(A_1|B_2) \leq \overline{P}_1(A_1) \text{ and } \underline{P}_1(A_1) \leq P(A_1) \leq \overline{P}_1(A_1). \tag{18}$$

In this sense, epistemic independence requires that our knowledge about $P(A_1|B_2)$ and $P(A_1)$ is identical: conditioning on B_2 should have no effect on our bounds for $P(A_1)$. This also explains the prefix epistemic: epistemic independence imposes a constraint on our knowledge. In other words, and at the risk of oversimplifying it, one could say that an assessment of epistemic independence does not entail a belief of independence—because, for a given $P \in \mathbb{P}_{\underline{P}_1 \otimes \underline{P}_2}$, X_1 and X_2 may very well be correlated—but rather an independence of beliefs.

Nevertheless, the standard notion of independence does correspond to a special case of epistemic independence: it suffices to use—restrictions of—two linear previsions P_1 on $\mathscr{G}(\mathscr{X}_1)$ and P_2 on $\mathscr{G}(\mathscr{X}_2)$ as our local models. Indeed, in that case, since the linearity of P_1 implies that $\underline{P}_1(A_1) = \overline{P}_1(A_1) = P_1(A_1)$, Equation (18) implies that for all $P \in \mathscr{P}_{P_1 \otimes P_2}$:

$$P(A_1 | B_2) = P(A_1) \text{ for all } A_1 \in \mathscr{P}(\mathscr{X}_1) \text{ and } B_2 \in \mathscr{B}_2,$$

and similarly if the indexes 1 and 2 are reversed. Furthermore, it then follows from Corollary 7.2 and Property P1 that $P(A_1 \cap B_2) = P(A_1)P(B_2)$, which is the conventional and well known defining factorisation property of independence.

We end by taking an even closer look at this specific case. So consider a linear prevision P_1 on $\mathscr{G}(\mathscr{X}_1)$ and a linear prevision P_2 on $\mathscr{G}(\mathscr{X}_2)$ or—in case we want to consider probability measures instead of expectations—their restrictions to a suitable set of indicators. Using the notation that we have adopted so far, the independent natural extension of these local models is then denoted by $P_1 \otimes P_2$. In this case however, this notation is a bit unfortunate, because it suggests that $P_1 \otimes P_2$ is a (conditional) linear prevision itself, which, as we will see, may not be the case. Therefore, we will adopt $P_1 \underline{\otimes} P_2$ as an alternative—less suggestive—notation for $P_1 \otimes P_2$ and will then use $P_1 \overline{\otimes} P_2$ to denote its corresponding conditional upper prevision, defined by

$$(P_1 \overline{\otimes} P_2)(f | B) := -(P_1 \otimes P_2)(-f | B) \text{ for all } (f, B) \in \mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2).$$

For atom-independence, the difference between $(P_1 \overline{\otimes} P_2)(f | B)$ and $(P_1 \underline{\otimes} P_2)(f | B)$ can be surprisingly large. For example, if the local models assign probability zero to all the singletons, then as already pointed out in the work of Miranda and Zaffalon [16], any joint linear prevision that marginalises to these local models will dominate the independent natural extension.

Proposition 9.1. *Consider a linear prevision P_1 on $\mathscr{G}(\mathscr{X}_1)$ and a linear prevision P_2 on $\mathscr{G}(\mathscr{X}_2)$ such that $P_1(\mathbb{I}_{x_1}) = 0$ and $P_2(\mathbb{I}_{x_2}) = 0$ for all $x_1 \in \mathscr{X}_1$ and $x_2 \in \mathscr{X}_2$. Let P_{12} be a linear prevision on $\mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$ that has P_1 and P_2 as its marginals. For atom-independence, we then have that for all $f \in \mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$:*

$$(P_1 \underline{\otimes} P_2)(f) \leq P_{12}(f) \leq (P_1 \overline{\otimes} P_2)(f).$$

We now use this result to obtain the following example. It demonstrates that $(P_1 \overline{\otimes} P_2)(f | B)$ and $(P_1 \underline{\otimes} P_2)(f | B)$ may be—substantially—different, and therefore, that $\mathbb{P}_{P_1 \otimes P_2}$ may have more than one element. Furthermore, and perhaps most importantly, it proves the claim that we made at the end of Example 7.3.

Example 9.1. Let $\mathscr{X}_1 = \mathscr{X}_2 = \mathbb{N}$, and let $g = \mathbb{I}_{\text{odd}} \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$ and $h = \mathbb{I}_{\text{even}} \in \mathscr{G}_{\geq 0}(\mathscr{X}_2)$, with $\mathbb{I}_{\text{odd}}$ and $\mathbb{I}_{\text{even}}$ defined as in Example 2.2. Consider now two linear previsions P_{odd} and P_{even} on $\mathscr{G}(\mathbb{N})$ such that $P_{\text{odd}}(\mathbb{I}_n) = 0$ and $P_{\text{even}}(\mathbb{I}_n) = 0$ for all $n \in \mathbb{N}$ and such that $P_{\text{odd}}(\mathbb{I}_{\text{odd}}) = 1 = P_{\text{even}}(\mathbb{I}_{\text{even}})$, and let $P := \frac{1}{2}(P_{\text{odd}} + P_{\text{even}})$. It then follows from Corollary 8.4 that P is a linear prevision on $\mathscr{G}(\mathbb{N})$. Furthermore, we find that $P(\mathbb{I}_n) = \frac{1}{2}(P_{\text{odd}}(\mathbb{I}_n) + P_{\text{even}}(\mathbb{I}_n)) = 0$ for all $n \in \mathbb{N}$ and, in the same way, that $P(\mathbb{I}_{\text{odd}}) = P(\mathbb{I}_{\text{even}}) = \frac{1}{2}$.

Now let $P_1 = P_2 = P$. Then as demonstrated in [16, Example 5], it is possible to construct a linear prevision P_{12} on $\mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$ that has P_1 and P_2 as its marginals and for which $P_{12}(X_1 \text{ odd and } X_2 \text{ even}) = P_{12}(gh) = 0$. Therefore, for atom-independence, Proposition 9.1 implies that $(P_1 \underline{\otimes} P_2)(gh) \leq 0$. Since the converse inequality follows from Property P1 and the non-negativity of fg, this implies that for atom-independence, $(P_1 \underline{\otimes} P_2)(gh) = 0$. Since $P_1(g)P_2(h) = P(\mathbb{I}_{\text{odd}})P(\mathbb{I}_{\text{even}}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, this shows that atom-independence does not lead to factorisation here, and proves our claim from Example 7.3.

For event-independence however, since linear previsions are a special case of coherent lower previsions, we already know from Example 7.3 that we do have factorisation here, and therefore, that $(P_1 \underline{\otimes} P_2)(gh) = \frac{1}{4}$.

Finally, we note that it follows from Property LP8 that, regardless of whether we adopt epistemic subset- or atom-indepence, $(P_1 \underline{\otimes} P_2)(gh) \leq (P_1 \overline{\otimes} P_2)(gh)$. Therefore, and because we know from Section 6 that epistemic event-independence leads to more informative joint models than epistemic atom-independence, our results in this example imply that for epistemic atom-independence:

$$(P_1 \overline{\otimes} P_2)(gh) \geq \frac{1}{4} > 0 = (P_1 \underline{\otimes} P_2)(gh). \qquad \Diamond$$

As explained in Section 7, the failure of factorisation that we observe in this example is a first important reason why we prefer epistemic event-independence over epistemic atom-independence. The second reason is that epistemic atom-independence leads to an independent natural extension that may be too uninformative, in the sense that $P_1 \underline{\otimes} P_2$ can be excessively small. This too was illustrated in the example above: for the same local models, epistemic event-independence gave rise to substantially higher joint lower previsions.

That said, even for event-independence, $(P_1 \overline{\otimes} P_2)(f | B)$ and $(P_1 \underline{\otimes} P_2)(f | B)$ may still be different. In that case, however, the reason is more subtle, and can be partially attributed to the fact that for infinite spaces and finitely additive probability measures, Fubini's theorem may not hold. We demonstrate this in our final example, which relies heavily on the following proposition.

Proposition 9.2. *Consider a linear prevision P_1 on $\mathscr{G}(\mathscr{X}_1)$ and a linear prevision P_2 on $\mathscr{G}(\mathscr{X}_2)$. Then for all $f \in \mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$, we have that*

$$(P_1 \underline{\otimes} P_2)(f) \leq P_1(P_2(f)) \leq (P_1 \overline{\otimes} P_2)(f),$$

with $P_2(f)$ a gamble on $\mathscr{X}_1$, defined by $P_2(f)(x_1) := P_2(f(x_1, X_2))$ for all $x_1 \in \mathscr{X}_1$.

Example 9.2. Consider a linear prevision P_1 on $\mathscr{G}(\mathscr{X}_1)$, a linear prevision P_2 on $\mathscr{G}(\mathscr{X}_2)$ and any $f \in \mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$ such that $P_1(P_2(f)) \neq P_2(P_1(f))$, where $P_2(f)$ is defined as in Proposition 9.2 and where, similarly, $P_1(f)$ is a gamble on $\mathscr{X}_2$ that is defined by $P_1(f)(x_2) := P_1(f(X_1, x_2))$ for all $x_2 \in \mathscr{X}_2$. It then follows from Proposition 9.2 and symmetry—reversing the role of P_1 and P_2—that

$$(P_1 \underline{\otimes} P_2)(f) \leq \min\{P_1(P_2(f)), P_2(P_1(f))\} < \max\{P_1(P_2(f)), P_2(P_1(f))\} \leq (P_1 \overline{\otimes} P_2)(f).$$

Concrete examples where this situation occurs—that is, where $P_1(P_2(f)) \neq P_2(P_1(f))$—can be found in [16, Example 1] and [19, Example 4].   ◇

10. Conclusions and future work

The main conclusion of this work is that by combining Williams-coherence with epistemic event-independence, we obtain a notion of independent natural extension that always exists, and that furthermore satisfies factorisation and external additivity. For weaker types of epistemic independence, including epistemic atom-independence, the existence result and the external additivity property remain valid, but factorisation then requires measurability conditions and the resulting inferences become less informative. For that reason, I think that when it comes to choosing between epistemic event-independence and epistemic atom-independence, the former should be preferred over the latter. In fact, I would advocate that from now on, and contrary to the current convention, epistemic independence should be taken to mean epistemic event-independence.

As far as future research is concerned, a first important step would be to extend our results from the case of two variables to that of any finite number of variables. Based on our own preliminary exploration of this topic, we expect that our proofs can be easily extended to that case. However, care will have to be taken when considering concepts such as independence, factorisation and external additivity, because for multiple variables, these have several variations; for finite state spaces, Reference [10] provides an excellent starting point.

Next, these extended versions of our results could then be used to develop efficient algorithms for credal networks whose variables take values in infinite spaces, by suitably adapting existing algorithms for the finite case. See for example the work in References [3,6,7].

On the more technical side, it would be useful to see whether our results can be extended from gambles—which are taken to be bounded—to the more general case of unbounded functions; in that case, establishing factorisation could prove to be tricky, because our proof for Corollary 7.2 relies rather heavily on the fact that gambles are bounded.

Finally, for variables that take values in Euclidean space, I would suggest to take a closer look at the case where $\mathscr{B}_1$ and $\mathscr{B}_2$ are restricted to the Lebesgue measurable events. Combined with a suitably chosen assessment of continuity, I think that this might lead to the development of a notion of independent natural extension that includes sigma additive product measures as a special case.

Appendix A. Proofs and additional material

In order to avoid forward referencing—and the associated risk of circular reasoning—the ordering of the proofs in this appendix sometimes differs from the order in which the corresponding results appear in the main text. Most importantly, the additional material for Section 8 presented immediately after of Section 3. Furthermore, since the proof of Proposition 2.1 relies on Proposition 2.2, the order of the proofs of these two results is reversed.

A.1. Proofs and additional material for Section 2

Proof of Proposition 2.2. Since $g \geq 0$ is a gamble and therefore by definition bounded, there is some $\alpha \in \mathbb{Q}_{>0}$ such that $0 \leq g < \alpha$. Fix any $n \in \mathbb{N}$ and let $g_n \in \mathscr{G}(\mathscr{X})$ be defined by

$$g_n := \frac{1}{n}\alpha \sum_{k=1}^{n-1} \mathbb{I}_{A_k}, \text{ where, for all } k \in \{1, \ldots, n-1\}, A_k := \left\{ x \in \mathscr{X} : g(x) \geq \frac{k}{n}\alpha \right\}.$$

For all $x \in \mathscr{X}$, we then find that

$$g_n(x) = \frac{k_x}{n}\alpha \leq g(x) \leq \frac{k_x+1}{n}\alpha, \text{ where we let } k_x := \max\{k \in \{0, \ldots, n-1\} : g(x) \geq \frac{k}{n}\alpha\},$$

which implies that $|g(x) - g_n(x)| \leq \frac{\alpha}{n}$. Since this is true for every $x \in \mathscr{X}$, this allows us to infer that $\sup |g - g_n| \leq \frac{\alpha}{n}$.

Consider now any $k \in \{1, \ldots, n-1\}$. Since $\frac{k}{n}\alpha \in \mathbb{Q}_{\geq 0}$, it follows from our assumptions on g that A_k is a finite union of pairwise disjoint events in $\mathscr{B} \cup \{\mathscr{X}, \emptyset\}$. Therefore, there is some $m_k \in \mathbb{N}$ and, for all $i \in \{1, \ldots, m_k\}$, some $B_{k,i} \in \mathscr{B} \cup \{\mathscr{X}, \emptyset\}$ such that $\mathbb{I}_{A_k} = \sum_{i=1}^{m_k} \mathbb{I}_{B_{k,i}}$. Since this is true for every $k \in \{1, \ldots, n-1\}$, it follows that $g_n = \frac{\alpha}{n} \sum_{k=1}^{n-1} \sum_{i=1}^{m_k} \mathbb{I}_{B_{k,i}}$. Since g_n is clearly non-negative, and because $\mathbb{I}_{\mathscr{X}} = 1$ and $\mathbb{I}_{\emptyset} = 0$, it now follows from Definition 2.1 that $g_n \in \mathscr{G}_{\geq 0}(\mathscr{X})$ is a simple $\mathscr{B}$-measurable gamble.

So, in summary then, for any fixed $n \in \mathbb{N}$, we know that we can construct a simple $\mathscr{B}$-measurable gamble $g_n \in \mathscr{G}_{\geq 0}(\mathscr{X})$ such that $\sup |g - g_n| \leq \frac{\alpha}{n}$. Definition 2.2 therefore clearly implies that g is $\mathscr{B}$-measurable. $\square$

Proof of Proposition 2.1. Consider any $\mathscr{B} \subseteq \mathscr{P}_{\emptyset}(\mathscr{X})$ such that $\mathscr{B}^* := \mathscr{B} \cup \{\emptyset\}$ is a sigma field and fix some $g \in \mathscr{G}_{\geq 0}(\mathscr{X})$.

We first prove the 'only if' part of the statement. So assume that g is $\mathscr{B}^*$-measurable in the measure-theoretic sense [17, Definition 10.1]. It then follows from [17, Corollary 10.5] that $\{x \in \mathscr{X} : g(x) \geq r\} \in \mathscr{B}^* = \mathscr{B} \cup \{\emptyset\}$ for all $r \in \mathbb{Q}_{\geq 0}$. Therefore, it follows from Proposition 2.2 that g is $\mathscr{B}$-measurable in the sense of Definition 2.2.

We end by proving the 'if' part of the statement. So assume that g is $\mathscr{B}$-measurable in the sense of Definition 2.2. This means that there is a sequence $\{g_n\}_{n \in \mathbb{N}}$ of simple $\mathscr{B}$-measurable gambles in $\mathscr{G}_{\geq 0}(\mathscr{X})$ such that $\lim_{n \to +\infty} \sup |g - g_n| = 0$. Then on the one hand, since $\lim_{n \to +\infty} \sup |g - g_n| = 0$ implies that $\lim_{n \to +\infty} |g(x) - g_n(x)| = 0$ for all $x \in \mathscr{X}$, we know that $\{g_n\}_{n \in \mathbb{N}}$ converges pointwise to g on $\mathscr{X}$. On the other hand, for any $n \in \mathbb{N}$, we know from Definition 2.1 that there are $c_0 \in \mathbb{R}_{\geq 0}$, $m \in \mathbb{N}_0$ and, for all $i \in \{1, \ldots, m\}$, $c_i \in \mathbb{R}_{\geq 0}$ and $B_i \in \mathscr{B}$, such that $g = c_0 + \sum_{i=1}^{m} c_i \mathbb{I}_{B_i}$. Let $B_0 = \mathscr{X}$. Since $\mathbb{I}_{\mathscr{X}} = 1$, and because $\mathscr{B}^*$ is a sigma field and therefore includes $\mathscr{X}$, we then find that $g = \sum_{i=0}^{m} c_i \mathbb{I}_{B_i}$, where, for all $i \in \{0, \ldots, n\}$, $B_i \in \mathscr{B}^*$. [17, Example 10.2] therefore implies that g_n is a $\mathscr{B}^*$-measurable function in the measure-theoretic sense. Since this is true for every $n \in \mathbb{N}$, and because $\{g_n\}_{n \in \mathbb{N}}$ converges pointwise to g on $\mathscr{X}$, it now follows from [17, Corollary 10.11(a)] that g is $\mathscr{B}^*$-measurable in the measure-theoretic sense. $\square$

Proof of Corollary 2.3. Immediate consequence of Proposition 2.2. $\square$

A.2. Additional material for Section 3

Many of the results in Section 3 and this corresponding part of the appendix are essentially well-known. Historically, most of them date back to Williams [24,25]. Our versions are basically just minor variations of his results, expressed in terms of lower previsions—instead of upper previsions—and without imposing structural constraints on the domain. Similar results can also be found in Reference [18] and, for the case of Walley-coherence, in Reference [14]. For that reason, and in order not to dilute the core novel part of our work, we state these results without proof. For those interested, explicit proofs are available in an online arXiv version of this contribution [5].

Lemma A.1. *For any $\mathscr{A} \subseteq \mathscr{G}(\mathscr{X})$, $\mathscr{E}(\mathscr{A})$ is a coherent set of desirable gambles on $\mathscr{X}$ if and only if it satisfies D4.*

Lemma A.2. *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be two subsets of $\mathscr{G}(\mathscr{X})$ such that $\mathscr{A}_1 \subseteq \mathscr{A}_2$. Then*

$$\text{posi}(\mathscr{A}_1) \subseteq \text{posi}(\mathscr{A}_2) \ \text{ and } \ \mathscr{E}(\mathscr{A}_1) \subseteq \mathscr{E}(\mathscr{A}_2).$$

Lemma A.3. *Let $\mathscr{D}$ be a coherent set of desirable gambles on $\mathscr{X}$. Then $\mathscr{E}(\mathscr{D}) = \mathscr{D}$.*

Lemma A.4. *Let $\mathscr{D}$ be a coherent set of desirable gambles on $\mathscr{X}$. If $f \in \mathscr{G}(\mathscr{X})$ and $f \notin \mathscr{D} \cup \{0\}$, then $\mathscr{E}(\mathscr{D} \cup \{-f\})$ is a coherent set of desirable gambles on $\mathscr{X}$.*

Proposition A.5. *Let $\underline{P}$ be a coherent conditional lower prevision on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$. Then for any $\mathscr{C}' \subseteq \mathscr{C}(\mathscr{X})$ such that $\mathscr{C} \subseteq \mathscr{C}'$, the restriction of $\underline{E}$ to $\mathscr{C}'$ is the pointwise smallest coherent conditional lower prevision on $\mathscr{C}'$ that coincides with $\underline{P}$ on $\mathscr{C}$.*

Lemma A.6. *Let $\underline{P}$ be a coherent conditional lower prevision on $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$. Then for any $(f, B) \in \mathscr{C}$ such that $f \mathbb{I}_B \in \mathscr{E}(\underline{P})$, we have that $\underline{P}(f|B) \geq 0$.*

Proof of Lemma A.6. It suffices to notice that

$$\underline{P}(f|B) = \underline{E}(f|B) = \underline{P}_{\mathscr{E}(\underline{P})}(f|B) = \sup\{\mu \in \mathbb{R} \colon [f - \mu]\mathbb{I}_B \in \mathscr{E}(\underline{P})\} \geq 0,$$

where the equalities follow from Proposition 3.3, Equation (6) and Equation (3), respectively, and where the final inequality follows from the fact that $f \mathbb{I}_B \in \mathscr{E}(\underline{P})$. □

Lemma A.7. *Let $\underline{P}$ be a coherent lower prevision on $\mathscr{G} \subseteq \mathscr{G}(\mathscr{X})$. Then for any $f \in \mathscr{G} \cap \mathscr{E}(\underline{P})$, we have that $\underline{P}(f) \geq 0$.*

Proof of Lemma A.7. Immediate consequence of Lemma A.6, for $\mathscr{C} = \{(f, \mathscr{X}) \colon f \in \mathscr{G}\}$. □

A.3. Additional material for Section 8

Here too, as in Appendix A.2, it should be noted that the results in Section 8 and this corresponding part of the appendix are essentially well-known; they are basically just minor variations of the results of Williams [24,25] and Pelessoni and Vicig [18]. As before, we therefore state these results without proof. For those interested, explicit proofs and intermediate results are available in an online arXiv version of this contribution [5].

Proposition A.8. *Let $\underline{P}$ be a coherent conditional lower prevision on $\mathscr{C}(\mathscr{X})$ and let $\overline{P}$ be the corresponding conditional upper prevision on $\mathscr{C}(\mathscr{X})$. Then there is a non-empty set $\mathbb{P}^*$ of conditional linear previsions on $\mathscr{C}(\mathscr{X})$ such that*

$$\underline{P}(f|B) = \min\{P(f|B) \colon P \in \mathbb{P}^*\} \text{ and } \overline{P}(f|B) = \max\{P(f|B) \colon P \in \mathbb{P}^*\}$$

for all $(f, B) \in \mathscr{C}(\mathscr{X})$. Furthermore, for any $(f, B) \in \mathscr{C}(\mathscr{X})$ and $\alpha \in [\underline{P}(f|B), \overline{P}(f|B)]$, there is some $P \in \mathbb{P}^*$ such that $P(f|B) = \alpha$.

A.4. Proofs and additional material for Section 5

A.4.1. The sets of desirable gambles part

Proposition A.9. *$\mathscr{D}_1 \otimes \mathscr{D}_2$ is a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$.*

Proof. Because of Lemma A.1, it suffices to prove D4. So consider any $f \in \mathscr{D}_1 \otimes \mathscr{D}_2$ and assume *ex absurdo* that $f \leq 0$. We will prove that this leads to a contradiction.

Since $\mathscr{D}_1$ and $\mathscr{D}_2$ are coherent, they are closed with respect to positive scaling and finite sums. Therefore, and because $f \in \mathscr{D}_1 \otimes \mathscr{D}_2 = \mathscr{E}(\mathscr{A}_{1 \to 2} \cup \mathscr{A}_{2 \to 1})$, it follows from Equations (9) and (10) that

$$f = \sum_{i \in I} \mathbb{I}_{B_{1,i}}(X_1) f_{2,i}(X_2) + \sum_{j \in J} \mathbb{I}_{B_{2,j}}(X_2) f_{1,j}(X_1) + g, \tag{A.1}$$

with I and J finite—possibly empty—index sets, with $B_{1,i} \in \mathscr{P}_\emptyset(\mathscr{X}_1)$ and $f_{2,i} \in \mathscr{D}_2$ for all $i \in I$, with $B_{2,j} \in \mathscr{P}_\emptyset(\mathscr{X}_2)$ and $f_{1,j} \in \mathscr{D}_1$ for all $j \in J$, with $g \geq 0$, and where $g = 0$ is only possible if $|I| + |J| > 0$.

Let us assume *ex absurdo* that $|I| + |J| = 0$. Then on the one hand, since we know that $g = 0$ is only possible if $|I| + |J| > 0$, it follows that $g \neq 0$. On the other hand, $|I| + |J| = 0$ also implies that $I = J = \emptyset$, and therefore, due to Equation (A.1), that $f = g$. Since $g \geq 0$ and $f \leq 0$, this in turn implies that $g = 0$, thereby contradicting the fact that $g \neq 0$. Hence, it follows that at least one of the two *ex absurdo* assumptions that we have so far made must be wrong. If $f \nleq 0$, then the proof is finished. For that reason, in the remainder of the proof, we can assume that $|I| + |J| \neq 0$, and therefore, that $|I| + |J| > 0$. The only *ex absurdo* assumption that still remains is that $f \leq 0$.

Now let $\{B_{1,k}\}_{k \in K}$ be the set consisting of those atoms of the algebra generated by $\{B_{1,i}\}_{i \in I}$ that belong to $\cup_{i \in I} B_{1,i}$ and, for all $k \in K$, let $f_{2,k} := \sum_{i \in I \colon B_{1,k} \subseteq B_{1,i}} f_{2,i}$. The following properties are then easily verified. First, since I is finite, K is also finite. Secondly, $|K| = 0$ if and only if $|I| = 0$. Thirdly, for all $k \in K$, we have that $B_{1,k} \in \mathscr{P}_\emptyset(\mathscr{X}_1)$ and—since $\mathscr{D}_1$ is coherent and therefore satisfies D3—that $f_{2,k} \in \mathscr{D}_1$. Fourthly, $\sum_{k \in K} \mathbb{I}_{B_{1,k}}(X_1) f_{2,k}(X_2)$ is equal to $\sum_{i \in I} \mathbb{I}_{B_{1,i}}(X_1) f_{2,i}(X_2)$. Fifthly, the events in $\{B_{1,k}\}_{k \in K}$ are pairwise disjoint. For this reason, without loss of generality, we can assume the events $\{B_{1,i}\}_{i \in I}$

in Equation (A.1) to be pairwise disjoint. A completely similar argument leads us to conclude that the events $\{B_{2,j}\}_{j \in J}$ in Equation (A.1) can be assumed to be pairwise disjoint, again without loss of generality.

If $\{B_{1,i}\}_{i \in I}$ is a partition of $\mathscr{X}_1$, then we let $\mathscr{Y}_1 := I$. Otherwise, we let $\mathscr{Y}_1 := I \cup \{i^*\}$ and define $B_{1,i^*} := \mathscr{X}_1 \setminus \cup_{i \in I} B_{1,i}$. Similarly, we let $\mathscr{Y}_2 := J$ if $\{B_{2,j}\}_{j \in J}$ is a partition of $\mathscr{X}_2$, and let $\mathscr{Y}_2 := J \cup \{j^*\}$ and $B_{1,j^*} := \mathscr{X}_2 \setminus \cup_{j \in J} B_{2,j}$ otherwise. Next, for every $i \in I$, we let $h_{2,i}$ be a gamble on $\mathscr{Y}_2$, defined by

$$h_{2,i}(y_2) := \sup\{f_{2,i}(x_2) : x_2 \in B_{2,y_2}\} \text{ for all } y_2 \in \mathscr{Y}_2. \tag{A.2}$$

Similarly, for every $j \in J$, we let $h_{1,j}$ be a gamble on $\mathscr{Y}_1$, defined by

$$h_{1,j}(y_1) := \sup\{f_{1,j}(x_1) : x_1 \in B_{1,y_1}\} \text{ for all } y_1 \in \mathscr{Y}_1. \tag{A.3}$$

Using these gambles on $\mathscr{Y}_1$ and $\mathscr{Y}_2$, we now construct a real-valued function h on $\mathscr{Y}_1 \times \mathscr{Y}_2$, defined by

$$h(y_1, y_2) := \sum_{i \in I} \mathbb{I}_i(y_1) h_{2,i}(y_2) + \sum_{j \in J} \mathbb{I}_j(y_2) h_{1,j}(y_1) \text{ for all } y_1 \in \mathscr{Y}_1 \text{ and } y_2 \in \mathscr{Y}_2 \tag{A.4}$$

This function is non-positive, in the sense that $h \le 0$. In order to prove that, let us fix any $y_1 \in \mathscr{Y}_1$ and $y_2 \in \mathscr{Y}_2$. It then follows from Equations (A.2) and (A.3) that

$$h(y_1, y_2) = \sum_{i \in I} \mathbb{I}_i(y_1) \sup_{x_2 \in B_{2,y_2}} f_{2,i}(x_2) + \sum_{j \in J} \mathbb{I}_j(y_2) \sup_{x_1 \in B_{1,y_1}} f_{1,j}(x_1).$$

Since $\mathbb{I}_i(y_1)$ can be non-zero for at most one $i \in I$ and $\mathbb{I}_j(y_2)$ can be non-zero for at most one $j \in J$, we know that each of the two summations on the right hand side contains at most one non-zero term. The suprema can therefore be moved outside of the summations, yielding

$$h(y_1, y_2) = \sup_{x_1 \in B_{1,y_1}} \sup_{x_2 \in B_{2,y_2}} \left(\sum_{i \in I} \mathbb{I}_i(y_1) f_{2,i}(x_2) + \sum_{j \in J} \mathbb{I}_j(y_2) f_{1,j}(x_1) \right).$$

For the next step, we start by observing the following. For any $x_1 \in B_{1,y_1}$ and any $i \in I$, since the sets $\{B_{1,i}\}_{i \in I}$ are pairwise disjoint, we know that $x_1 \in B_{1,i}$ if and only if $y_1 = i$, which implies that $\mathbb{I}_i(y_1) = \mathbb{I}_{B_{1,i}}(x_1)$. Similarly, for any $x_2 \in B_{2,y_2}$ and any $j \in J$, since the sets $\{B_{2,j}\}_{j \in J}$ are pairwise disjoint, we know that $x_2 \in B_{2,j}$ if and only if $y_2 = j$, which implies that $\mathbb{I}_j(y_2) = \mathbb{I}_{B_{2,j}}(x_2)$. As an immediate consequence, it follows that

$$h(y_1, y_2) = \sup_{x_1 \in B_{1,y_1}} \sup_{x_2 \in B_{2,y_2}} \left(\sum_{i \in I} \mathbb{I}_{B_{1,i}}(x_1) f_{2,i}(x_2) + \sum_{j \in J} \mathbb{I}_{B_{2,j}}(x_2) f_{1,j}(x_1) \right).$$

Finally, in combination with Equation (A.1), this implies that

$$h(y_1, y_2) = \sup_{x_1 \in B_{1,y_1}} \sup_{x_2 \in B_{2,y_2}} (f(x_1, x_2) - g(x_1, x_2)) \le 0,$$

where, for the last inequality, we use the fact that $f \le 0$ and $g \ge 0$. Since this true for every $y_1 \in \mathscr{Y}_1$ and $y_2 \in \mathscr{Y}_2$, it follows that $h \le 0$.

Now let $\mathscr{A}_1 := \{h_{1,j} : j \in J\}$ and $\mathscr{A}_2 := \{h_{2,i} : i \in I\}$ and assume *ex absurdo* that $\mathscr{H}_1 := \mathscr{E}(\mathscr{A}_1)$ and $\mathscr{H}_2 := \mathscr{E}(\mathscr{A}_2)$ are coherent sets of desirable gambles on $\mathscr{Y}_1$ and $\mathscr{Y}_2$, respectively. We will prove that this is impossible, by constructing a probability mass function p on $\mathscr{X}_1 \times \mathscr{X}_2$ such that the corresponding expectation of h is both non-positive *and* positive, thereby obtaining a contradiction. In order to do that, we borrow an argument of De Cooman and Miranda [9, Proof of Proposition 15] that is based on a very useful lemma of them, which, in order to make this paper self-contained, is restated here in Lemma A.10.

Since $\mathscr{H}_1$ is a coherent set of desirable gambles on $\mathscr{Y}_1$, it follows from Definition 3.1—and D4 in particular—that $0 \notin \mathscr{H}_1 = \mathscr{E}(\mathscr{A}_1)$. Therefore, and because $\mathscr{Y}_1$ and J—and hence also $\mathscr{A}_1$—are finite, it follows from Lemma A.10 that there is a probability mass function p_1 on $\mathscr{Y}_1$ such that $p_1(y_1) > 0$ for all $y_1 \in \mathscr{Y}_1$ and $\sum_{y_1 \in \mathscr{Y}_1} p_1(y_1) h_{1,j}(y_1) > 0$ for all $j \in J$. Using a completely analogous argument, we also infer that there is a probability mass function p_2 on $\mathscr{Y}_2$ such that $p_2(y_2) > 0$ for all $y_2 \in \mathscr{Y}_2$ and $\sum_{y_2 \in \mathscr{Y}_2} p_2(y_2) h_{2,i}(y_2) > 0$ for all $i \in I$.

We now let p be the probability mass function on $\mathscr{Y}_1 \times \mathscr{Y}_2$ that is defined by $p(y_1, y_2) := p_1(y_1) p_2(y_2)$ for all $y_1 \in \mathscr{Y}_1$ and $y_2 \in \mathscr{Y}_2$, and we let $E_p(h)$ be the expectation of h with respect to p, as defined by

$$E_p(h) := \sum_{y_1 \in \mathscr{Y}_1} \sum_{y_2 \in \mathscr{Y}_2} p(y_1, y_2) h(y_1, y_2) = \sum_{y_1 \in \mathscr{Y}_1} \sum_{y_2 \in \mathscr{Y}_2} p_1(y_1) p_2(y_2) h(y_1, y_2).$$

Then on the one hand, since $h \le 0$, we have that $E_p(h) \le 0$. On the other hand, however, it follows from Equation (A.4) that

$$
\begin{aligned}
E_p(h) &= \sum_{y_1 \in \mathscr{Y}_1} \sum_{y_2 \in \mathscr{Y}_2} p_1(y_1) p_2(y_2) \left(\sum_{i \in I} \mathbb{I}_i(y_1) h_{2,i}(y_2) + \sum_{j \in J} \mathbb{I}_j(y_2) h_{1,j}(y_1) \right) \\
&= \sum_{y_1 \in \mathscr{Y}_1} \sum_{y_2 \in \mathscr{Y}_2} p_1(y_1) p_2(y_2) \sum_{i \in I} \mathbb{I}_i(y_1) h_{2,i}(y_2) + \sum_{y_1 \in \mathscr{Y}_1} \sum_{y_2 \in \mathscr{Y}_2} p_1(y_1) p_2(y_2) \sum_{j \in J} \mathbb{I}_j(y_2) h_{1,j}(y_1) \\
&= \sum_{i \in I} \sum_{y_1 \in \mathscr{Y}_1} p_1(y_1) \mathbb{I}_i(y_1) \sum_{y_2 \in \mathscr{Y}_2} p_2(y_2) h_{2,i}(y_2) + \sum_{j \in J} \sum_{y_2 \in \mathscr{Y}_2} p_2(y_2) \mathbb{I}_j(y_2) \sum_{y_1 \in \mathscr{Y}_1} p_1(y_1) h_{1,j}(y_1) \\
&= \sum_{i \in I} p_1(i) \sum_{y_2 \in \mathscr{Y}_2} p_2(y_2) h_{2,i}(y_2) + \sum_{j \in J} p_2(j) \sum_{y_1 \in \mathscr{Y}_1} p_1(y_1) h_{1,j}(y_1).
\end{aligned}
$$

For every $i \in I$, it follows from the properties of p_1 and p_2 that the corresponding term in this summation is positive. Similarly, for every $j \in J$, it follows from the properties of p_1 and p_2 that the corresponding term in this summation is positive. Since $|I| + |J| > 0$, this implies that $E_p(h) > 0$, thereby contradicting the fact that $E_p(h) \leq 0$. Hence, it follows that one of the two remaining *ex absurdo* assumptions is wrong. If $f \leq 0$, then the proof is finished. Therefore, in the remainder of the proof, we can assume that there is at least one $i \in \{1, 2\}$ for which $\mathscr{H}_i$ is incoherent. Without loss of generality, symmetry allows us to assume that $i = 1$, that is, that $\mathscr{H}_1$ is incoherent. The only *ex absurdo* assumption that still remains is that $f \leq 0$.

Since $\mathscr{H}_1$ is incoherent, it follows from Lemma A.1 that there is some $h^* \in \mathscr{H}_1$ such that $h^* \leq 0$. Furthermore, since $h^* \in \mathscr{H}_1$, Equation (2) implies that $h^* = \lambda g^* + \sum_{j \in J} \lambda_j h_{1,j}$, for some $\lambda \in \mathbb{R}_{\geq 0}$ and $g^* \in \mathscr{G}_{>0}(\mathscr{Y}_1)$ and, for all $j \in J$, some $\lambda_j \in \mathbb{R}_{\geq 0}$, with $\lambda + \sum_{j \in J} \lambda_j > 0$. If $\lambda_j = 0$ for all $j \in J$, then $\lambda > 0$ and $g^* = \frac{1}{\lambda} h^* \leq 0$, which is impossible because $g^* \in \mathscr{G}_{>0}(\mathscr{Y}_1)$. Therefore, we know that there is at least one $j \in J$ such that $\lambda_j > 0$.

Now let $f_1 := \sum_{j \in J} \lambda_j f_{1,j}$ and fix any $x_1^* \in \mathscr{X}_1$. Since the events in $\{B_{1,y_1}\}_{y_1 \in \mathscr{Y}_1}$ are pairwise disjoint, there will then be a unique $y_1^* \in \mathscr{Y}_1$ such that $x_1^* \in B_{1,y_1^*}$. For this particular choice of y_1^*, we then find that

$$
f_1(x_1^*) = \sum_{j \in J} \lambda_j f_{1,j}(x_1^*) \leq \sum_{j \in J} \lambda_j \sup_{x_1 \in B_{1,y_1^*}} f_{1,j}(x_1) = \sum_{j \in J} \lambda_j h_{1,j}(y_1^*) = h^*(y_1^*) - \lambda g^*(y_1^*) \leq 0,
$$

where the first equality follows from Equation (A.3) and the second inequality follows from the fact that $h^* \leq 0$, $\lambda \geq 0$ and $g^* \in \mathscr{G}_{>0}(\mathscr{Y}_1)$. Since this is true for every $x_1^* \in \mathscr{X}_1$, we infer that $f_1 \leq 0$. However, on the other hand, since there is at least one $j \in J$ such that $\lambda_j > 0$, and because $f_{1,j} \in \mathscr{D}_1$ for all $j \in J$, the coherence of $\mathscr{D}_1$ implies that $f_1 \in \mathscr{D}_1$ and therefore, because of D4, that $f_1 \nleq 0$. From this contradiction, it follows that one of our *ex absurdo* assumptions must be false. Since the only remaining *ex absurdo* assumption is that $f \leq 0$, this concludes the proof. $\square$

Lemma A.10. *[9, Lemma 2] Let Ω be a finite set and consider some finite subset $\mathscr{A}$ of $\mathscr{G}(\Omega)$. Then $0 \notin \mathscr{E}(\mathscr{A})$ if and only if there is a probability mass function p on Ω such that $p(\omega) > 0$ for all $\omega \in \Omega$ and $\sum_{\omega \in \Omega} p(\omega) f(\omega) > 0$ for all $f \in \mathscr{A}$.*

Proposition A.11. *$\mathscr{D}_1 \otimes \mathscr{D}_2$ is an independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$.*

Proof. For ease of notation, let $\mathscr{D} := \mathscr{D}_1 \otimes \mathscr{D}_2$. Because of symmetry, it clearly suffices to prove that

$$
(\forall B_2 \in \mathscr{B}_2) \quad \mathscr{D}_1 = \mathrm{marg}_1(\mathscr{D}) = \mathrm{marg}_1(\mathscr{D} | B_2),
$$

which, since $\mathrm{marg}_1(\mathscr{D}) = \mathrm{marg}_1(\mathscr{D} | \mathscr{X}_2)$, is equivalent to proving that, for all $f_1 \in \mathscr{G}(\mathscr{X}_1)$ and $B_2 \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$,

$$
f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{D} \quad \Leftrightarrow \quad f_1 \in \mathscr{D}_1.
$$

Since $f_1 \in \mathscr{D}_1$ implies that $f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{A}_{2 \to 1} \subseteq \mathscr{D}$ for all $B_1 \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$, the converse implication holds trivially. So consider any $f_1 \in \mathscr{G}(\mathscr{X}_1)$ and $B_2 \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$ such that $f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{D}$. Since we know from Proposition A.9 that $\mathscr{D}$ is coherent, this implies that $f_1 \neq 0$. It remains to prove that $f_1 \in \mathscr{D}_1$.

Assume *ex absurdo* that $f_1 \notin \mathscr{D}_1$. Then since $f_1 \neq 0$, $\mathscr{D}_1^\bullet := \mathscr{E}(\mathscr{D}_1 \cup \{-f_1\})$ is a coherent set of desirable gambles on $\mathscr{X}_1$ because of Lemma A.4, and therefore, if we let

$$
\mathscr{A}_{2 \to 1}^\bullet := \left\{ f_1'(X_1) \mathbb{I}_{B_2'}(X_2) \colon f_1' \in \mathscr{D}_1^\bullet, B_2' \in \mathscr{B}_2 \cup \{\mathscr{X}_2\} \right\}, \tag{A.5}
$$

it follows from Proposition A.9 that $\mathscr{D}_1^\bullet \otimes \mathscr{D}_2 := \mathscr{E}(\mathscr{A}_{1 \to 2} \cup \mathscr{A}_{2 \to 1}^\bullet)$ is a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$. Now on the one hand, since $-f_1 \in \mathscr{D}_1^\bullet$, it follows from Equation (A.5) that $-f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{A}_{2 \to 1}^\bullet \subseteq \mathscr{D}_1^\bullet \otimes \mathscr{D}_2$. On the other hand, since $\mathscr{D}_1 \subseteq \mathscr{D}_1^\bullet$ implies that $\mathscr{D}_1 \otimes \mathscr{D}_2 \subseteq \mathscr{D}_1^\bullet \otimes \mathscr{D}_2$, we infer from $f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{D}$ that $f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{D}_1^\bullet \otimes \mathscr{D}_2$. Since $\mathscr{D}_1^\bullet \otimes \mathscr{D}_2$ is coherent, this implies that

$$
0 = f_1(X_1) \mathbb{I}_{B_2}(X_2) - f_1(X_1) \mathbb{I}_{B_2}(X_2) \in \mathscr{D}_1^\bullet \otimes \mathscr{D}_2,
$$

which contradicts D4. $\square$

Proof of Theorem 5.1. Since we know from Proposition A.11 that $\mathscr{D}_1 \otimes \mathscr{D}_2$ is an independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$, it suffices to prove that any other such independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$ is a superset of $\mathscr{D}_1 \otimes \mathscr{D}_2$.

So let $\mathscr{D}$ be any independent product of $\mathscr{D}_1$ and $\mathscr{D}_2$. Definition 5.1 then implies that $\mathscr{D}$ is coherent and that $\mathscr{A}_{1 \to 2} \cup \mathscr{A}_{2 \to 1} \subseteq \mathscr{D}$. Hence, we find that

$$\mathscr{D}_1 \otimes \mathscr{D}_2 = \mathscr{E}\left(\mathscr{A}_{1 \to 2} \cup \mathscr{A}_{2 \to 1}\right) \subseteq \mathscr{E}(\mathscr{D}) = \mathscr{D},$$

where the inclusion follows from Lemma A.2 and the final equality from Lemma A.3. $\square$

A.4.2. The conditional lower previsions part

Proposition A.12. $\underline{P}_1 \otimes \underline{P}_2$ *is a coherent conditional probability on* $\mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$.

Proof. For all $i \in \{1, 2\}$, since $\underline{P}_i$ is a coherent conditional lower prevision on $\mathscr{C}_i$, it follows from Proposition 3.2 that $\mathscr{E}(\underline{P}_i)$ is a coherent set of desirable gambles on $\mathscr{X}_i$. Therefore, Proposition A.9 implies that $\mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2)$ is a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$. The result now follows from Definition 3.3. $\square$

Proposition A.13. *Consider two indexes i and j such that $\{i, j\} = \{1, 2\}$. Then for any $f_i \in \mathscr{G}(\mathscr{X}_i)$ and $B_i \in \mathscr{P}_\emptyset(\mathscr{X}_i)$ and any $B_j \in \mathscr{B}_j$, we have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(f_i | B_i \cap B_j) = (\underline{P}_1 \otimes \underline{P}_2)(f_i | B_i) = \underline{E}_i(f_i | B_i). \tag{A.6}$$

Proof. For all $i \in \{1, 2\}$, since $\underline{P}_i$ is a coherent conditional lower prevision on $\mathscr{C}_i$, it follows from Proposition 3.2 that $\mathscr{E}(\underline{P}_i)$ is a coherent set of desirable gambles on $\mathscr{X}_i$. Therefore, we infer from Proposition A.11 that $\mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2)$ is an independent product of $\mathscr{E}(\underline{P}_1)$ and $\mathscr{E}(\underline{P}_2)$. For ease of notation, we now let $\underline{P} := \underline{P}_1 \otimes \underline{P}_2$ and $\mathscr{D} := \mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2)$. As we know from Equation (11), $\underline{P}$ is then equal to $\underline{P}_{\mathscr{D}}$. Furthermore, since $\mathscr{D}$ is an independent product of $\mathscr{E}(\underline{P}_1)$ and $\mathscr{E}(\underline{P}_2)$, we know that $\mathscr{D}$ is epistemically independent and that it has $\mathscr{E}(\underline{P}_1)$ and $\mathscr{E}(\underline{P}_2)$ as its marginals.

We are now ready to prove Equation (A.6). In order to do that, we fix any two indexes i and j such that $\{i, j\} = \{1, 2\}$, any $f_i \in \mathscr{G}(\mathscr{X}_i)$ and $B_i \in \mathscr{P}_\emptyset(\mathscr{X}_i)$ and any $B_j \in \mathscr{B}_j$. We start by proving the first equality. Since $\mathscr{D}$ is epistemically independent, we know that

$$[f_i - \mu]\mathbb{I}_{B_i} \in \mathscr{D} \Leftrightarrow [f_i - \mu]\mathbb{I}_{B_i} \in \mathrm{marg}_i(\mathscr{D})$$
$$\Leftrightarrow [f_i - \mu]\mathbb{I}_{B_i} \in \mathrm{marg}_i(\mathscr{D}|B_j) \Leftrightarrow [f_i - \mu]\mathbb{I}_{B_i}\mathbb{I}_{B_j} \in \mathscr{D} \Leftrightarrow [f_i - \mu]\mathbb{I}_{B_i \cap B_j} \in \mathscr{D}$$

for all $\mu \in \mathbb{R}$, and therefore, we find that

$$\underline{P}(f_i | B_i) = \sup\left\{\mu \in \mathbb{R} : [f_i - \mu]\mathbb{I}_{B_i} \in \mathscr{D}\right\} = \sup\left\{\mu \in \mathbb{R} : [f_i - \mu]\mathbb{I}_{B_i \cap B_j} \in \mathscr{D}\right\} = \underline{P}(f_i | B_i \cap B_j).$$

Next, we prove the second equality of Equation (A.6). Since $\mathscr{D}$ has $\mathscr{E}(\underline{P}_1)$ and $\mathscr{E}(\underline{P}_2)$ as its marginals, we know that

$$[f_i - \mu]\mathbb{I}_{B_i} \in \mathscr{D} \Leftrightarrow [f_i - \mu]\mathbb{I}_{B_i} \in \mathrm{marg}_i(\mathscr{D}) \Leftrightarrow [f_i - \mu]\mathbb{I}_{B_i} \in \mathscr{E}(\underline{P}_i)$$

for all $\mu \in \mathbb{R}$, and therefore, we find that

$$\underline{P}(f_i | B_i) = \sup\left\{\mu \in \mathbb{R} : [f_i - \mu]\mathbb{I}_{B_i} \in \mathscr{D}\right\} = \sup\left\{\mu \in \mathbb{R} : [f_i - \mu]\mathbb{I}_{B_i} \in \mathscr{E}(\underline{P}_i)\right\} = \underline{E}_i(f_i | B_i),$$

using Equation 6 to establish the last equality. $\square$

Proposition A.14. *The restriction of $\underline{P}_1 \otimes \underline{P}_2$ to $\mathscr{C}$ is an independent product of $\underline{P}_1$ and $\underline{P}_2$.*

Proof. Since we know from Proposition A.12 that $\underline{P}_1 \otimes \underline{P}_2$ is a coherent lower prevision on $\mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$, it follows from Definition 3.3 that its restriction to $\mathscr{C}$ is coherent as well. Due to Definition 5.3, it remains to show that this restriction of $\underline{P}_1 \otimes \underline{P}_2$ to $\mathscr{C}$ is epistemically independent and that it coincides with $\underline{P}_1$ and $\underline{P}_2$ on their domain. Epistemic independence follows trivially from Definition 4.2 and Proposition A.13. Hence, it remains to prove that the restriction of $\underline{P}_1 \otimes \underline{P}_2$ to $\mathscr{C}$ coincides with $\underline{P}_1$ and $\underline{P}_2$ on their domain, or equivalently, that

$$(\underline{P}_1 \otimes \underline{P}_2)(f_i | B_i) = \underline{P}_i(f_i | B_i) \text{ for all } i \in \{1, 2\} \text{ and } (f_i, B_i) \in \mathscr{C}_i.$$

So fix any $i \in \{1, 2\}$ and $(f_i, B_i) \in \mathscr{C}_i$. We then find that indeed, as desired,

$$(\underline{P}_1 \otimes \underline{P}_2)(f_i | B_i) = \underline{E}_i(f_i | B_i) = \underline{P}_i(f_i | B_i),$$

where the first equality follows from Proposition A.13 and the second equality follows from Equation (6) and Proposition 3.2. $\square$

Proof of Theorem 5.2. Since we know from Proposition A.14 that the restriction of $\underline{P}_1 \otimes \underline{P}_2$ to $\mathscr{C}$ is an independent product of $\underline{P}_1$ and $\underline{P}_2$, it suffices to prove that any other such independent product of $\underline{P}_1$ and $\underline{P}_2$ dominates $\underline{P}_1 \otimes \underline{P}_2$ on $\mathscr{C}$.

So let $\underline{P}$ be any independent product of $\underline{P}_1$ and $\underline{P}_2$. Definition 5.3 then implies that $\underline{P}$ is an epistemically independent coherent conditional lower prevision on $\mathscr{C}$ that coincides with $\underline{P}_1$ and $\underline{P}_2$ on their domain. Let $\mathscr{A}_P$ be the corresponding set of gambles, as defined by Equation (5), and let $\mathscr{D} := \mathscr{E}(\underline{P}) = \mathscr{E}(\mathscr{A}_P)$. We then know from Proposition 3.2 that $\mathscr{D}$ is a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$ and that $\underline{P}_{\mathscr{D}}$ coincides with $\underline{P}$ on $\mathscr{C}$. In the remainder of this proof, we will show that $\mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2) \subseteq \mathscr{D}$. Because of Equation (11), this clearly implies that $\underline{P}_{\mathscr{D}}(f|B) \geq (\underline{P}_1 \otimes \underline{P}_2)(f|B)$ for all $(f, B) \in \mathscr{C}$. Since $\underline{P}_{\mathscr{D}}$ coincides with $\underline{P}$ on $\mathscr{C}$, this implies that $\underline{P}$ dominates $\underline{P}_1 \otimes \underline{P}_2$ on $\mathscr{C}$, thereby concluding the proof.

Let $\mathscr{D}_1 := \mathscr{E}(\underline{P}_1)$ and let $\mathscr{A}_{2 \to 1}$ be the corresponding set of gambles on $\mathscr{X}_1 \times \mathscr{X}_2$, as defined by Equation (10). We will now prove that $\mathscr{A}_{2 \to 1} \subseteq \mathscr{D}$. So consider any $f_1 \in \mathscr{D}_1$ and any $B_2 \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$. We need to prove that $f_1(X_1)\mathbb{I}_{B_2}(X_2) \in \mathscr{D}$. Since $f_1 \in \mathscr{D}_1 = \mathscr{E}(\underline{P}_1) = \mathrm{posi}(\mathscr{A}_{P_1} \cup \mathscr{G}_{>0}(\mathscr{X}_1))$, it follows from Equation (1) that there are $n \in \mathbb{N}$ and, for all $i \in \{1, \ldots, n\}$, $\lambda_i \in \mathbb{R}_{>0}$ and $g_i \in \mathscr{A}_{P_1} \cup \mathscr{G}_{>0}(\mathscr{X}_1)$ such that $f_1 = \sum_{i=1}^n \lambda_i g_i$.

For any $i \in \{1, \ldots, n\}$, we now let $h_i(X_1, X_2) := g_i(X_1)\mathbb{I}_{B_2}(X_2) \in \mathscr{G}(\mathscr{X}_1 \times \mathscr{X}_2)$. As we will show, this gamble h_i belongs to $\mathscr{D}$. We consider two cases: $g_i \in \mathscr{G}_{>0}(\mathscr{X}_1)$ and $g_i \notin \mathscr{G}_{>0}(\mathscr{X}_1)$. If $g_i \in \mathscr{G}_{>0}(\mathscr{X}_1)$, then $h_i \in \mathscr{G}_{>0}(\mathscr{X}_1 \times \mathscr{X}_2)$, which, since $\mathscr{D}$ is a coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$, implies that $h_i \in \mathscr{D}$. If $g_i \notin \mathscr{G}_{>0}$, then since $g_i \in \mathscr{A}_{P_1} \cup \mathscr{G}_{>0}(\mathscr{X}_1)$, it follows that $g_i \in \mathscr{A}_{P_1}$, which implies that there are $(f_1', B_1) \in \mathscr{C}_1$ and $\mu < \underline{P}_1(f_1'|B_1)$ such that $g_i = [f_1' - \mu]\mathbb{I}_{B_1}$. Furthermore, since $\underline{P}$ coincides with $\underline{P}_1$ on its domain, we also know that $\underline{P}_1(f_1'|B_1) = \underline{P}(f_1'|B_1)$. If $B_2 = \mathscr{X}_2$, Equation (5) therefore implies that $h_i \in \mathscr{A}_P \subseteq \mathscr{D}$ because $\mathbb{I}_{B_2} = 1$. If $B_2 \neq \mathscr{X}_2$, then $B_2 \in \mathscr{B}_2$. Since $\underline{P}$ is epistemically independent, this implies that $\underline{P}(f_1'|B_1) = \underline{P}(f_1'|B_1 \cap B_2)$. Hence, here too, Equation (5) implies that $h_i \in \mathscr{A}_P \subseteq \mathscr{D}$—because $\mathbb{I}_{B_1 \cap B_2} = \mathbb{I}_{B_1}\mathbb{I}_{B_2}$.

In summary then, we have found that $h_i \in \mathscr{D}$ for all $i \in \{1, \ldots, n\}$. Since $f_1 = \sum_{i=1}^n \lambda_i g_i$, this implies that

$$f_1(X_1)\mathbb{I}_{B_2}(X_2) = \left(\sum_{i=1}^n \lambda_i g_i(X_1)\right)\mathbb{I}_{B_2}(X_2) = \sum_{i=1}^n \lambda_i g_i(X_1)\mathbb{I}_{B_2}(X_2) = \sum_{i=1}^n \lambda_i h_i(X_1, X_2) \in \mathscr{D},$$

where the inclusion holds because $\mathscr{D}$ is coherent. Since this is true for every $f_1 \in \mathscr{D}_1$ and every $B_2 \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$, it follows that $\mathscr{A}_{2 \to 1} \subseteq \mathscr{D}$. Using a completely analogous argument, it also follows that $\mathscr{A}_{1 \to 2} \subseteq \mathscr{D}$, with $\mathscr{A}_{1 \to 2}$ defined by Equation (9) for $\mathscr{D}_2 := \mathscr{E}(\underline{P}_2)$. Hence, we find that $\mathscr{A}_{1 \to 2} \cup \mathscr{A}_{2 \to 1} \subseteq \mathscr{D}$, and therefore, that

$$\mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2) = \mathscr{D}_1 \otimes \mathscr{D}_2 = \mathscr{E}(\mathscr{A}_{1 \to 2} \cup \mathscr{A}_{2 \to 1}) \subseteq \mathscr{E}(\mathscr{D}) = \mathscr{D},$$

where the second equality follows from Equation (8), the inclusion follows from Lemma A.2, and the last equality follows from Lemma A.3. $\square$

A.5. Proofs and additional material for Section 6

Proof of Proposition 6.1. We only prove the result for $\mathscr{D}_1 \otimes \mathscr{D}_2$. The result for $\underline{P}_1 \otimes \underline{P}_2$ then follows trivially from Equation (11).

Let $\mathscr{D}_1 \otimes \mathscr{D}_2$ be the independent natural extension that corresponds to $\mathscr{B}_1$ and $\mathscr{B}_2$, as defined by Equations (8)–(10), and let $\mathscr{D}_1 \otimes' \mathscr{D}_2$ be the independent natural extension that corresponds to $\mathscr{B}_1'$ and $\mathscr{B}_2'$, defined by

$$\mathscr{D}_1 \otimes' \mathscr{D}_2 := \mathscr{E}\left(\mathscr{A}_{1 \to 2}' \cup \mathscr{A}_{2 \to 1}'\right),$$

with

$$\mathscr{A}_{1 \to 2}' := \left\{f_2(X_2)\mathbb{I}_{B_1'}(X_1) \colon f_2 \in \mathscr{D}_2, B_1' \in \mathscr{B}_1' \cup \{\mathscr{X}_1\}\right\}$$

and

$$\mathscr{A}_{2 \to 1}' := \left\{f_1(X_1)\mathbb{I}_{B_2'}(X_2) \colon f_1 \in \mathscr{D}_1, B_2' \in \mathscr{B}_2' \cup \{\mathscr{X}_2\}\right\}.$$

Then as explained in the main text, in the paragraph that precedes Proposition 6.1, we have that $\mathscr{D}_1 \otimes \mathscr{D}_2 \subseteq \mathscr{D}_1 \otimes' \mathscr{D}_2$. It remains to prove that $\mathscr{D}_1 \otimes' \mathscr{D}_2 \subseteq \mathscr{D}_1 \otimes \mathscr{D}_2$.

Fix any $f_2 \in \mathscr{D}_2$ and $B_1' \in \mathscr{B}_1' \cup \{\mathscr{X}_1\}$. We will prove that $f_2(X_2)\mathbb{I}_{B_1'}(X_1) \in \mathscr{D}_1 \otimes \mathscr{D}_2$. If $B_1' = \mathscr{X}_1$, this follows trivially from Equations (8) and (9). Otherwise, it follows from our assumptions that there is some $m \in \mathbb{N}$ and, for all $k \in \{1, \ldots, m\}$, some $B_{1,k} \in \mathscr{B}_1$ such that B_1' is a finite disjoint union of the events $\{B_{1,k}\}_{1 \leq k \leq m}$, which implies that $\mathbb{I}_{B_1'} = \sum_{k=1}^m \mathbb{I}_{B_{1,k}}$ and therefore also that $f_2(X_2)\mathbb{I}_{B_1'}(X_1) = \sum_{k=1}^m f_2(X_2)\mathbb{I}_{B_1}(X_1)$. Hence, Equations (8) and (9) again imply that $f_2(X_2)\mathbb{I}_{B_1'}(X_1) \in \mathscr{D}_1 \otimes \mathscr{D}_2$. Since this is true for every $f_2 \in \mathscr{D}_2$ and $B_1' \in \mathscr{B}_1' \cup \{\mathscr{X}_1\}$, it follows that $\mathscr{A}_{1 \to 2}' \subseteq \mathscr{D}_1 \otimes \mathscr{D}_2$. Using a completely analogous argument, we also infer that $\mathscr{A}_{2 \to 1}' \subseteq \mathscr{D}_1 \otimes \mathscr{D}_2$. The result now follows because $\mathscr{A}_{1 \to 2}' \cup \mathscr{A}_{2 \to 1}' \subseteq \mathscr{D}_1 \otimes \mathscr{D}_2$ implies that

$$\mathscr{D}_1 \otimes' \mathscr{D}_2 = \mathscr{E}\left(\mathscr{A}_{1 \to 2}' \cup \mathscr{A}_{2 \to 1}'\right) \subseteq \mathscr{E}(\mathscr{D}_1 \otimes \mathscr{D}_2) = \mathscr{D}_1 \otimes \mathscr{D}_2,$$

using Lemma A.2 for the inclusion and Lemma A.3 and Proposition A.9 for the last equality. $\square$

A.6. Proofs and additional material for Section 7

Lemma A.15. *For any $f \in \mathscr{G}(\mathscr{X}_1)$ and $h \in \mathscr{G}(\mathscr{X}_2)$ and any simple $\mathscr{B}_1$-measurable $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$, we have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) \geq \underline{E}_1\big(f + g\underline{E}_2(h)\big).$$

Proof. Since $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$ is a simple $\mathscr{B}$-measurable gamble, we know from Definition 2.1 that there are $c_0 \in \mathbb{R}_{\geq 0}$, $n \in \mathbb{N}_0$ and, for all $i \in \{1, \ldots, n\}$, $c_i \in \mathbb{R}_{\geq 0}$ and $B_i \in \mathscr{B}_1$, such that $g = c_0 + \sum_{i=1}^{n} c_i \mathbb{I}_{B_i}$. Furthermore, since we know from Proposition A.12 that $\underline{P}_1 \otimes \underline{P}_2$ is coherent, it follows from Proposition 3.4 that $\underline{P}_1 \otimes \underline{P}_2$ satisfies LP2, LP3 and LP4. Finally, since $\underline{E}_2$ is coherent, we know from Proposition 3.4 that it satisfies LP6. Therefore, we find that

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) = (\underline{P}_1 \otimes \underline{P}_2)\Big(f + g\underline{E}_2(h) + \big(c_0 + \sum_{i=1}^{n} c_i \mathbb{I}_{B_i}\big)[h - \underline{E}_2(h)]\Big)$$

$$\geq (\underline{P}_1 \otimes \underline{P}_2)\big(f + g\underline{E}_2(h)\big) + c_0(\underline{P}_1 \otimes \underline{P}_2)\big(h - \underline{E}_2(h)\big) + \sum_{i=1}^{n} c_i (\underline{P}_1 \otimes \underline{P}_2)\big(\mathbb{I}_{B_i}[h - \underline{E}_2(h)]\big)$$

$$= \underline{E}_1\big(f + g\underline{E}_2(h)\big) + c_0 \underline{E}_2\big(h - \underline{E}_2(h)\big) + \sum_{i=1}^{n} c_i (\underline{P}_1 \otimes \underline{P}_2)\big(\mathbb{I}_{B_i}[h - (\underline{P}_1 \otimes \underline{P}_2)(h|B_i)]\big)$$

$$= \underline{E}_1\big(f + g\underline{E}_2(h)\big) + c_0\big(\underline{E}_2(h) - \underline{E}_2(h)\big) = \underline{E}_1\big(f + g\underline{E}_2(h)\big),$$

where the first equality follows because $g = c_0 + \sum_{i=1}^{n} c_i \mathbb{I}_{B_i}$, where the first inequality follows because $\underline{P}_1 \otimes \underline{P}_2$ satisfies LP3 and LP2, where the second equality follows from Proposition A.13, and where the third equality follows because $\underline{E}_2$ satisfies LP6 and $\underline{P}_1 \otimes \underline{P}_2$ satisfies LP4. $\square$

Lemma A.16. *For any $f \in \mathscr{G}(\mathscr{X}_1)$ and $h \in \mathscr{G}(\mathscr{X}_2)$ and any simple $\mathscr{B}_1$-measurable $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$, we have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) \leq \underline{E}_1\big(f + g\underline{E}_2(h)\big).$$

Proof. Since $\underline{E}_2$ is a coherent conditional lower prevision on $\mathscr{C}(\mathscr{X}_2)$, we know from Proposition A.8 that there is a conditional linear prevision P_2 on $\mathscr{C}(\mathscr{X}_2)$ such that $P_2(h) = \underline{E}_2(h)$ and $P_2 \geq \underline{E}_2$. Similarly, since $\underline{E}_1$ is a coherent conditional lower prevision on $\mathscr{C}(\mathscr{X}_1)$, we know from Proposition A.8 that there is a conditional linear prevision P_1 on $\mathscr{C}(\mathscr{X}_1)$ such that $P_1\big(f + g\underline{E}_2(h)\big) = \underline{E}_1\big(f + g\underline{E}_2(h)\big)$ and $P_1 \geq \underline{E}_1$.

Consider now any $i \in \{1, 2\}$. We then know from Proposition A.5 that $\underline{E}_i$ coincides with $\underline{P}_i$ on $\mathscr{C}_i$. Therefore, and because $P_i \geq \underline{E}_i$, we also know that P_i dominates $\underline{P}_i$ on $\mathscr{C}_i$. Due to Equation (5), this implies that $\mathscr{A}_{\underline{P}_i} \subseteq \mathscr{A}_{P_i}$ and therefore, using Lemma A.2, also that $\mathscr{E}(\underline{P}_i) \subseteq \mathscr{E}(P_i)$. Since this is true for every $i \in \{1, 2\}$, it follows from Equation (8) and Lemma A.2 that $\mathscr{E}(\underline{P}_1) \otimes \mathscr{E}(\underline{P}_2) \subseteq \mathscr{E}(P_1) \otimes \mathscr{E}(P_2)$, and therefore, because of Equation (11), that $\underline{P}_1 \otimes \underline{P}_2 \leq P_1 \otimes P_2$.

The result can now be proved as follows. First, since $\underline{P}_1 \otimes \underline{P}_2 \leq P_1 \otimes P_2$, we find that

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) \leq (P_1 \otimes P_2)(f + gh). \tag{A.7}$$

Secondly, since we know from Proposition A.12 that $(P_1 \otimes P_2)$ is coherent, it follows from Proposition 3.4 that $(P_1 \otimes P_2)$ satisfies LP8, which implies that

$$(P_1 \otimes P_2)(f + gh) \leq -(P_1 \otimes P_2)(-f - gh) \leq -P_1\big(-f + gP_2(-h)\big), \tag{A.8}$$

using Lemma A.15 for the second inequality. Finally, we also know that

$$-P_1\big(-f + gP_2(-h)\big) = P_1\big(f + gP_2(h)\big) = \underline{E}_1\big(f + g\underline{E}_2(h)\big), \tag{A.9}$$

where the first equality follows from Definitions 8.1 and 8.2 because P_1 and P_2 are conditional linear previsions, and where the second equality follows because $P_2(h) = \underline{E}_2(h)$ and $P_1\big(f + g\underline{E}_2(h)\big) = \underline{E}_1\big(f + g\underline{E}_2(h)\big)$. By combining Equations (A.7)–(A.9), the result is now immediate. $\square$

Proposition A.17. *For any $f \in \mathscr{G}(\mathscr{X}_1)$ and $h \in \mathscr{G}(\mathscr{X}_2)$ and any simple $\mathscr{B}_1$-measurable $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$, we have that*

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) = \underline{E}_1\big(f + g\underline{E}_2(h)\big).$$

Proof. Immediate consequence of Lemmas A.15 and A.16. $\square$

Proof of Theorem 7.1. Since $g \in \mathscr{G}_{\geq 0}(\mathscr{X}_1)$ is $\mathscr{B}_1$-measurable, we know from Definition 2.2 that there is a sequence $\{g_n\}_{n \in \mathbb{N}}$ of simple $\mathscr{B}_1$-measurable gambles in $\mathscr{G}_{\geq 0}(\mathscr{X}_1)$ such that g_n converges uniformly to g. This also implies that $f + g_n \underline{E}_2(h)$ converges uniformly to $f + g \underline{E}_2(h)$ and, since h is a gamble and therefore by definition bounded, that $f + g_n h$ converges uniformly to $f + gh$. The result now follows from the following series of equalities:

$$(\underline{P}_1 \otimes \underline{P}_2)(f + gh) = \lim_{n \to +\infty} (\underline{P}_1 \otimes \underline{P}_2)(f + g_n h) = \lim_{n \to +\infty} \underline{E}_1 \big(f + g_n \underline{E}_2(h) \big) = \underline{E}_1 \big(f + g \underline{E}_2(h) \big).$$

The first of these equalities holds because it follows from Propositions A.12 and 3.4 that $\underline{P}_1 \otimes \underline{P}_2$ satisfies LP5. The second equality follows from Proposition A.17. The third equality holds because the coherence of $\underline{E}_1$ allows us to infer from Proposition 3.4 that $\underline{E}_1$ satisfies LP5. $\square$

Proof of Corollary 7.2. Let $f := 0 \in \mathscr{G}(\mathscr{X}_i)$. We then know from Theorem 7.1 that

$$(\underline{P}_1 \otimes \underline{P}_2)(gh) = (\underline{P}_1 \otimes \underline{P}_2)(f + gh) = \underline{E}_i \big(f + g \underline{E}_j(h) \big) = \underline{E}_i \big(g \underline{E}_j(h) \big).$$

The result can now be inferred from the non-negative homogeneity—LP2—of $\underline{E}_i$ that is implied by its coherence. If $\underline{E}_j(h) \geq 0$, we simply apply the non-negative homogeneity for $\lambda := \underline{E}_j(h)$. If $\underline{E}_j(h) \leq 0$, we apply it for $\lambda := -\underline{E}_j(h)$ and combine this with the fact that $\overline{E}_i(g) := -\underline{E}_i(-g)$. $\square$

Proof of Corollary 7.3. Let $g := 1$. Then g belongs to $\mathscr{G}_{\geq 0}(\mathscr{X}_1)$ and is $\mathscr{B}_1$-measurable. Therefore, we know from Theorem 7.1 that

$$(\underline{P}_1 \otimes \underline{P}_2)(f + h) = (\underline{P}_1 \otimes \underline{P}_2)(f + gh) = \underline{E}_1 \big(f + g \underline{E}_2(h) \big) = \underline{E}_1 \big(f + \underline{E}_2(h) \big).$$

The result now follows from the constant additivity—LP6—of $\underline{E}_1$ that is implied by its coherence. $\square$

A.7. Proofs and additional material for Section 9

Proof of Proposition 9.1. Fix any $\epsilon > 0$. It then follows from Equations (11) and (3) that there is some $\mu \in \mathbb{R}$ such that $\mu > (P_1 \underline{\otimes} P_2)(f) - \epsilon$ and $f - \mu \in \mathscr{E}(P_1) \otimes \mathscr{E}(P_2)$. Consider any such μ. Since $f - \mu \in \mathscr{E}(P_1) \otimes \mathscr{E}(P_2)$, it follows from Equations (8), (9), (10), (2) and (1) that

$$f - \mu = g + \sum_{i=1}^{n} \lambda_i f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1) + \sum_{j=1}^{m} \lambda_j f_{1,j}(X_1) \mathbb{I}_{B_{2,j}}(X_2)$$

with $g \in \mathscr{G}_{\geq 0}$ and $n, m \in \mathbb{N}_0$ and, for all $i \in \{1, \ldots, n\}$, $\lambda_i \in \mathbb{R}_{>0}$, $f_{2,i} \in \mathscr{E}(P_2)$ and $B_{1,i} \in \mathscr{B}_1 \cup \{\mathscr{X}_1\}$ and, for all $j \in \{1, \ldots, m\}$, $\lambda_j \in \mathbb{R}_{>0}$, $f_{1,j} \in \mathscr{E}(P_1)$ and $B_{2,j} \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$. Since P_{12} is a linear prevision—and hence satisfies P1–P3—this implies that

$$P_{12}(f) = P_{12}(\mu) + P_{12}(g) + \sum_{i=1}^{n} \lambda_i P_{12}(f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1)) + \sum_{j=1}^{m} \lambda_j P_{12}(f_{1,j}(X_1) \mathbb{I}_{B_{2,j}}(X_2))$$

$$\geq \mu + \sum_{i=1}^{n} \lambda_i P_{12}(f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1)) + \sum_{j=1}^{m} \lambda_j P_{12}(f_{1,j}(X_1) \mathbb{I}_{B_{2,j}}(X_2))$$

Furthermore, for any $i \in \{1, \ldots, n\}$, there are two cases. If $B_{1,i} = \mathscr{X}_1$, then

$$P_{12}(f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1)) = P_{12}(f_{2,i}(X_2)) = P_2(f_{2,i}) \geq 0,$$

where the second equality follows from our assumptions on P_{12} and the third equality follows from Lemma A.7. If $B_{1,i} \in \mathscr{B}_1$, then

$$P_{12}(f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1)) \geq P_{12}(\inf_{x_2 \in \mathscr{X}_2} f_{2,i} \mathbb{I}_{B_{1,i}}(X_1)) = \inf_{x_2 \in \mathscr{X}_2} f_{2,i} P_{12}(\mathbb{I}_{B_{1,i}}(X_1)) = \inf_{x_2 \in \mathscr{X}_2} f_{2,i} P_1(\mathbb{I}_{B_{1,i}}) = 0,$$

where the inequality follows from LP7, the first equality follows from P2 and the last two equalities follow from our assumptions on P_{12} and P_1 and from the fact that—since we consider atom-independence—$B_{1,i} \in \mathscr{B}_1$ implies that there is some $x_1 \in \mathscr{X}_1$ such that $B_{1,i} = \{x_1\}$ and hence also $P_1(\mathbb{I}_{B_{1,i}}) = P_1(\mathbb{I}_{x_1}) = 0$. In both cases, we find that $P_{12}(f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1)) \geq 0$. Similarly, for any $j \in \{1, \ldots, m\}$, we find that $P_{12}(f_{1,j}(X_1) \mathbb{I}_{B_{2,j}}(X_2)) \geq 0$. Hence, it follows that $P_{12}(f) \geq \mu$. Since $\mu > (P_1 \underline{\otimes} P_2)(f) - \epsilon$, this implies that $P_{12}(f) > (P_1 \underline{\otimes} P_2)(f) - \epsilon$, and, since $\epsilon > 0$ is arbitrary, this in turn implies that $P_{12}(f) \geq (P_1 \underline{\otimes} P_2)(f)$. This already establishes the first inequality of the statement. For the second equality of the statement, it suffices to apply the first equality to $-f$, yielding $(P_1 \underline{\otimes} P_2)(-f) \leq P_{12}(-f)$. Indeed, since P_{12} is linear, this trivially implies that

$$(P_1 \overline{\otimes} P_2)(f) = -(P_1 \underline{\otimes} P_2)(-f) \geq -P_{12}(-f) = P_{12}(f),$$

thereby establishing the second inequality of the statement. $\square$

Proof of Proposition 9.2. Fix any $\epsilon > 0$. It then follows from Equations (11) and (3) that there is some $\mu \in \mathbb{R}$ such that $\mu > (P_1 \underline{\otimes} P_2)(f) - \epsilon$ and $f - \mu \in \mathscr{E}(P_1) \otimes \mathscr{E}(P_2)$. Consider any such μ. Since $f - \mu \in \mathscr{E}(P_1) \otimes \mathscr{E}(P_2)$, it follows from Equations (8), (9), (10), (2) and (1) that

$$f - \mu = g + \sum_{i=1}^{n} \lambda_i f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1) + \sum_{j=1}^{m} \lambda_j f_{1,j}(X_1) \mathbb{I}_{B_{2,j}}(X_2)$$

with $g \in \mathscr{G}_{\geq 0}$ and $n, m \in \mathbb{N}_0$ and, for all $i \in \{1, \ldots, n\}$, $\lambda_i \in \mathbb{R}_{>0}$, $f_{2,i} \in \mathscr{E}(P_2)$ and $B_{1,i} \in \mathscr{B}_1 \cup \{\mathscr{X}_1\}$ and, for all $j \in \{1, \ldots, m\}$, $\lambda_j \in \mathbb{R}_{>0}$, $f_{1,j} \in \mathscr{E}(P_1)$ and $B_{2,j} \in \mathscr{B}_2 \cup \{\mathscr{X}_2\}$. Since P_1 and P_2 are linear previsions—and hence satisfy P1–P3—this implies that

$$P_1(P_2(f)) = P_1(P_2(\mu) + P_1(P_2(g)) + \sum_{i=1}^{n} \lambda_i P_1(P_2(f_{2,i}(X_2) \mathbb{I}_{B_{1,i}}(X_1))) + \sum_{j=1}^{m} \lambda_j P_1(P_2(f_{1,j}(X_1) \mathbb{I}_{B_{2,j}}(X_2)))$$

$$\geq \mu + \sum_{i=1}^{n} \lambda_i P_1(P_2(f_{2,i}(X_2)) \mathbb{I}_{B_{1,i}}(X_1)) + \sum_{j=1}^{m} \lambda_j P_1(f_{1,j}(X_1) P_2(\mathbb{I}_{B_{2,j}}(X_2)))$$

$$= \mu + \sum_{i=1}^{n} \lambda_i P_1(\mathbb{I}_{B_{1,i}}(X_1)) P_2(f_{2,i}(X_2)) + \sum_{j=1}^{m} \lambda_j P_1(f_{1,j}(X_1)) P_2(\mathbb{I}_{B_{2,j}}(X_2)).$$

Furthermore, for any $i \in \{1, \ldots, n\}$, it follows from Lemma A.7 that $P_2(f_{2,i}(X_2)) \geq 0$ and from P1 that $P_1(\mathbb{I}_{B_{1,i}}(X_1)) \geq 0$, which implies that $\lambda_i P_1(\mathbb{I}_{B_{1,i}}(X_1)) P_2(f_{2,i}(X_2)) \geq 0$. Using a completely analogous argument, for any $j \in \{1, \ldots, m\}$, we find that $\lambda_j P_1(f_{1,j}(X_1)) P_2(\mathbb{I}_{B_{2,j}}(X_2)) \geq 0$. Hence, it follows that $P_1(P_2(f)) \geq \mu$. Since $\mu > (P_1 \underline{\otimes} P_2)(f) - \epsilon$, this implies that $P_1(P_2(f)) > (P_1 \underline{\otimes} P_2)(f) - \epsilon$, and, since $\epsilon > 0$ is arbitrary, this in turn implies that $P_1(P_2(f)) \geq (P_1 \underline{\otimes} P_2)(f)$. This already establishes the first inequality of the statement. For the second equality of the statement, it suffices to apply the first equality to $-f$, yielding $(P_1 \underline{\otimes} P_2)(-f) \leq P_1(P_2(-f))$. Indeed, since P_1 and P_2 are linear, this trivially implies that

$$(P_1 \overline{\otimes} P_2)(f) = -(P_1 \underline{\otimes} P_2)(-f) \geq -P_1(P_2(-f)) = P_1(P_2(f)),$$

thereby establishing the second inequality of the statement. $\square$

References

[1] I. Couso, S. Moral, P. Walley, Examples of independence for imprecise probabilities, in: ISIPTA '99: Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications, 1999, pp. 121–130.
[2] F.G. Cozman, Independence for sets of full conditional probabilities, sets of lexicographic probabilities, and sets of desirable gambles, in: ISIPTA '13: Proceedings of the Eighth International Symposium on Imprecise Probability: Theory and Applications, 2013, pp. 87–97.
[3] J. De Bock, Credal Networks Under Epistemic Irrelevance: Theory and Algorithms, PhD thesis, Ghent University, 2015.
[4] J. De Bock, Independent natural extension for infinite spaces: Williams-coherence to the rescue, in: PMLR: Proceedings of Machine Learning Research, 62 (Proceedings of ISIPTA 2017), 2017, pp. 121–132.
[5] J. De Bock, Independent natural extension for infinite spaces, arXiv:1805.01139 [math.PR], 2018.
[6] J. De Bock, G. de Cooman, An efficient algorithm for estimating state sequences in imprecise hidden Markov models, J. Artif. Intell. Res. 50 (2014) 189–233.
[7] G. de Cooman, F. Hermans, A. Antonucci, M. Zaffalon, Epistemic irrelevance in credal nets: the case of imprecise Markov trees, Int. J. Approx. Reason. 51 (9) (2010) 1029–1052.
[8] G. de Cooman, E. Miranda, Weak and strong laws of large numbers for coherent lower previsions, J. Stat. Plan. Inference 138 (8) (2008) 2409–2432.
[9] G. de Cooman, E. Miranda, Irrelevant and independent natural extension for sets of desirable gambles, J. Artif. Intell. Res. 45 (2012) 601–640.
[10] G. de Cooman, E. Miranda, M. Zaffalon, Independent natural extension, Artif. Intell. 175 (12) (2011) 1911–1950.
[11] B. de Finetti, Teoria delle probabilità, Einaudi, Turin, 1970, English translation: [12].
[12] B. de Finetti, Theory of Probability, John Wiley & Sons, New York, 1974.
[13] L.E. Dubins, Finitely additive conditional probabilities, conglomerability and disintegrations, Ann. Appl. Probab. 3 (1) (1975) 89–99.
[14] E. Miranda, M. Zaffalon, Notes on desirability and conditional lower previsions, Ann. Math. Artif. Intell. 60 (3–4) (2010) 251–309.
[15] E. Miranda, M. Zaffalon, Conglomerable coherence, Int. J. Approx. Reason. 54 (9) (2013) 1322–1350.
[16] E. Miranda, M. Zaffalon, Independent products in infinite spaces, J. Math. Anal. Appl. 425 (1) (2015) 460–488.
[17] O.A. Nielsen, An Introduction to Integration and Measure Theory, Wiley, 1997.
[18] R. Pelessoni, P. Vicig, Williams coherence and beyond, Int. J. Approx. Reason. 50 (4) (2009) 612–626.
[19] M.J. Schervish, T. Seidenfeld, R. Stern, J.B. Kadane, What finite-additivity can add to decision theory, Submitted to the Ann. Stat.
[20] E. Quaeghebeur, Desirability, in: T. Augustin, F.P.A. Coolen, G. de Cooman, M.C.M.Troffaes (Eds.), Introduction to Imprecise Probabilities, John Wiley & Sons, 2014, pp. 1–27.
[21] P. Vicig, Epistemic independence for imprecise probabilities, Int. J. Approx. Reason. 24 (2–3) (2000) 235–250.
[22] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.
[23] P. Walley, Towards a unified theory of imprecise probability, Int. J. Approx. Reason. 24 (2–3) (2000) 125–148.
[24] P.M. Williams, Notes on Conditional Previsions, Technical report, School of Mathematical and Physical Science, University of Sussex, 1975.
[25] P.M. Williams, Notes on conditional previsions, Int. J. Approx. Reason. 44 (3) (2007) 366–383.