INDEPENDENT NATURAL EXTENSION

GERT DE COOMAN, ENRIQUE MIRANDA, AND MARCO ZAFFALON

ABSTRACT. There is no unique extension of the standard notion of probabilistic independence to the case where probabilities are indeterminate or imprecisely specified. *Epistemic independence* is an extension that formalises the intuitive idea of mutual irrelevance between different sources of information. This gives epistemic independence very wide scope as well as appeal: this interpretation of independence is often taken as natural also in precise-probabilistic contexts. Nevertheless, epistemic independence has received little attention so far. This paper develops the foundations of this notion for variables assuming values in finite spaces. We define (epistemically) independent products of marginals (or possibly conditionals) and show that there always is a unique least-committal such independent product, which we call the *independent natural extension*. We supply an explicit formula for it, and study some of its properties, such as associativity, marginalisation and external additivity, which are basic tools to work with the independent natural extension. Additionally, we consider a number of ways in which the standard factorisation formula for independence can be generalised to an imprecise-probabilistic context. We show, under some mild conditions, that when the focus is on least-committal models, using the independent natural extension is equivalent to imposing a so-called strong factorisation property. This is an important outcome for applications as it gives a simple tool to make sure that inferences are consistent with epistemic independence judgements. We discuss the potential of our results for applications in Artificial Intelligence by recalling recent work by some of us, where the independent natural extension was applied to graphical models. It has allowed, for the first time, the development of an exact linear-time algorithm for the imprecise probability updating of credal trees.

1. INTRODUCTION

Background and motivation. This is a paper on the notion of independence in probability theory. Anyone interested in or familiar with uncertain reasoning or statistics knows how fundamental this notion is. But what is independence?

Most of us have been taught that two variables X_1 and X_2 are independent when their joint probability distribution $P_{\{1,2\}}$ factorises as the product of its marginals P_1 and P_2. This is the *formalist* route that defines independence through a mathematical property of the joint, and that has its roots in the Kolmogorovian, measure- and integral-theoretic formalisation of probability theory.

In Artificial Intelligence (AI)—thanks to Judea Pearl in particular [27]—, but also in the tradition of subjective probability—due to a large extent to Bruno de Finetti [17]—, independence has much more often an *epistemic* flavour: it is a subject who regards two variables as independent, because she judges that learning about the value of any one of them will not affect her beliefs about the other. This means that the subject assesses that her conditional beliefs equal her marginal ones: $P_1(\cdot|X_2) = P_1$ and $P_2(\cdot|X_1) = P_2$, in more mathematical parlance.

That the epistemic approach has become so popular, should not be all that surprising. The formalist approach comes with the idea that independence is something given, which might hold or not: it is just a property of the joint. On the epistemic view, however, independence

Key words and phrases. Epistemic irrelevance, epistemic independence, independent natural extension, strong product, factorisation, coherent lower previsions.

Preliminary work on the subject of this paper has appeared in the conferences IPMU 2010 [15] and SMPS 2010 [14].

is something we are (to some extent) in control of. And this control is essential in order to aggregate simple, independent components into complex multivariate models.

It might be argued that the difference between the two approaches is mostly philosophical: in fact, the two routes are known to be formally equivalent.[1] But it turns out that we lose this formal equivalence as soon as we consider probabilities that may be imprecisely specified, meaning that the available information is conveniently expressed through sets of probabilities (sets of mass functions). In this case the two routes diverge also mathematically, as we shall see further on. This is exemplified by the existence of the different notions of *strong* and *epistemic independence*, respectively.[2] Of these two, strong independence has been most thoroughly investigated in the literature. Studies of epistemic independence are confined to a relatively small number of papers [6, 10, 26, 29] inspired by Peter Walley's [30, Section 9.3] seminal ideas. We mention in particular Paolo Vicig's interesting study [29], for the case of coherent lower *probabilities* (which may be defined on infinite spaces), of some of the notions considered in this paper as well.

This situation is somewhat unfortunate as the scope of strong independence is relatively narrow: in fact, its justification seems to rely on a *sensitivity analysis interpretation* of imprecise probabilities. On this interpretation, one assumes that there exists some (kind of 'ideal' or 'true') precise probability $P^T_{\{1,2\}}$ for the variables X_1 and X_2 that satisfies stochastic independence, and that, due to the lack of time or other resources, can only be partially specified or assessed. Then one considers all the precise-probabilistic models $P_{\{1,2\}}$ that are consistent with the partial assessments and that satisfy stochastic independence. Taken together, they constitute the set of probabilities for the problem under consideration. This set models a subject's (partial) ignorance about the true model $P^T_{\{1,2\}}$.

It is questionable that this sensitivity analysis interpretation is broadly applicable, for the simple reason that it hinges on the assumption of the existence of the underlying 'true' probability $P^T_{\{1,2\}}$. Consider the situation where we wish to model an expert's beliefs: the expert usually does not know much about ideal probabilities, and what she tells us is simply that information about one variable does not influence her beliefs about the other. Moreover, we could well argue that expert knowledge is *inherently* imprecise to some extent, no matter the resources that we employ to capture it.[3] Therefore, why not take the expert at her word and model *only* the information she provides us about the mutual irrelevance of the two variables under consideration? After all, forcing a sensitivity analysis interpretation here would amount to adding unwarranted assumptions, which may lead us to draw stronger conclusions than those the expert herself might be prepared to get to.

In order to model such mutual irrelevance, we need a different understanding of imprecise probability models that does not (necessarily) rely on precise probability as a more primitive notion: Walley's behavioural theory of imprecise probability [30], which models beliefs by looking at a subject's buying and selling prices for gambles. The perceived mutual irrelevance of two sources of information can be formalised easily in this framework: we state that the subject is not going to change her prices for gambles that depend on one variable, when the other variable is observed. This turns out to be still equivalent to modelling the problem through a set of precise probabilities $P_{\{1,2\}}$ but, in contradistinction with the case of sensitivity analysis, not all those probabilities satisfy stochastic independence in general. The reason for this is that epistemic independence is a property of the set of probabilities that cannot be explained through the properties of the precise probabilities that make up the set. This point is not without importance, as it shows that buying and selling prices for gambles are actually a more primitive and fundamental notion in a theory of personal probability.

[1] There may be subtleties, however, related to events of probability zero. See Refs. [6, Notes 5 and 6 in Section 3], [30, Sections 6.5 and 6.10] and [1] for more information.

[2] Other possible ways to define independence under imprecise probability are given in Ref. [2].

[3] See [30, Chapter 5] for a detailed exposition of this view.

This illustrates that a behavioural theory of probability and the notion of epistemic independence fit nicely together. It also indicates that epistemic independence has a very wide scope, as it needs to meet fewer requirements than strong independence in order to be employed. That being so, why has strong independence been studied and applied much more extensively than its epistemic counterpart, even in work based on Walley's approach? This is probably due to a number of concurring factors: (i) a tendency in the literature to extend stochastic independence, perhaps somewhat uncritically, in a straightforward way to imprecise probabilities; (ii) the fact that epistemic independence does not appear to be as well-behaved as strong independence, for instance with respect to the graphoid axioms [6];[4] and, perhaps more importantly, (iii) the lack of formal tools for handling epistemic independence assessments. To give a telling illustration of this last point: epistemically independent products have so far been given a formal definition [30, Section 9.3] only for the case of two variables X_1 and X_2.

Aims and contributions. In the present paper, then, we intend to address and remedy this lack of formal tools by providing a firm foundation for, and a thorough mathematical discussion of, epistemic independence in the case of a finite number of variables X_n taking values in finite sets $\mathscr{X}_n$, $n \in N$. Perhaps surprisingly, this will also shed positive new light on the second of the above-mentioned factors: it will allow us to show that, despite the apparently negative results in Ref. [6], epistemic independence can actually be used effectively in at least some types of graphical models. We will come back to this issue later in this Introduction.

We ground our analysis in the conceptual and formal framework of *coherent lower previsions*, which are lower expectation functionals equivalent to closed convex sets of probability mass functions. In the case of precise probabilities, we refer to an expectation functional as a *linear prevision*. Section 2 gives a brief introduction to coherent lower previsions and reports basic results that will be used in the rest of the paper. It should make the paper as self-contained as is reasonably achievable within the scope of a research paper.

The real work starts in Section 3, where we introduce and discuss several generalisations to coherent lower previsions of the standard notion of factorisation: *productivity*, which was used by some of us in Ref. [12] to derive a very general law of large numbers, *factorisation* and *strong factorisation*, which we needed in our research on *credal networks* (an imprecise-probabilistic graphical model) [11], the *Kuznetsov* and *strong Kuznetsov* properties, originating in the work of the Russian mathematician Vladimir Kuznetsov [19], and also studied by Fabio Cozman [3]. It is useful to keep in mind that the 'strong' versions of these properties involve factorisations over any subsets of variables, while the 'plain' ones are the special case obtained when some of the subsets are singletons. For linear previsions—the precise-probability models—all these properties coincide with the classical notion of stochastic independence. For the more general lower previsions, we investigate how these notions are related, and we show that the *strong product*—the product that arises through strong independence—is strongly factorising.

In Section 4, we go over to the epistemic side. We introduce two notions: *many-to-many independence*, where a subject judges that learning about the value that any subset of the variables $\{X_n \colon n \in N\}$ assumes will not affect her beliefs about the values of any other disjoint subset; and the weaker notion of *many-to-one independence*, where she judges that learning about the value that any subset of the variables assumes will not affect her beliefs about the value of any remaining single variable. This leads to the definition of two corresponding types of *independent products*. We prove some useful associativity and marginalisation properties for these, which form a basis for building them recursively, and prove a very useful theorem that immediately allows all these notions, as well as the results in the rest of the paper, to be extended to the case of conditional independence.

[4] But also see Ref. [26] for a discussion with less negative conclusions.

Moreover, we show that the strong product is one particular many-to-many (and therefore many-to-one) independent product, and that it is the only such independent product when the given marginals are linear previsions.

There is no such uniqueness in the more general case of marginal lower previsions: the strong product is only one of the generally infinitely many possible independent products. In Section 5, we focus on the pointwise smallest of all these: the least-committal many-to-many, and the least-committal many-to-one, independent products of given marginals. It is an important, and quite involved, result of our analysis that these two smallest independent products turn out to always exist, and to coincide. We call this common smallest independent product the *independent natural extension* of the given marginals. It generalises, to any finite number of variables, a definition given by Walley for two variables [30, Section 9.3].[5] We then go on to derive an explicit and constructive formula for the independent natural extension, and we prove that it too satisfies useful associativity and marginalisation properties, and that it is externally additive. We work out interesting particular cases in some detail.

The relation with the more formal factorisation properties considered in Section 3 comes to the fore in our important next result: that the independent natural extension is strongly factorising. We go somewhat further in Section 7, where we show that, quite naturally, any factorising lower prevision must be a many-to-one independent product. Under some mild conditions, we also show that any strongly factorising lower prevision must be a many-to-many independent product. And since we already know that the smallest many-to-one independent product is the independent natural extension, we deduce that when looking for least-committal models, it is immaterial whether we focus on factorisation or on being an independent product. This outcome might be very important in applications, as it allows one to work with the independent natural extension simply by imposing a (strong) factorisation property while searching for least-committal models. On the more theoretical side, it constitutes a solid bridge between the formalist and epistemic approaches to independence.

We believe these results give epistemic independence an opportunity to become a competitive alternative to more consolidated notions of independence. But how can epistemic independence be used in specific examples, and are there advantages to doing so? We present an interesting case study in Section 8, where we survey and discuss some of our recent work on credal trees [11]. These constitute a special case of credal networks [4], which in turn extend Bayesian nets to deal with imprecise probabilities. Traditionally, the extension is achieved by replacing the precise-probabilistic parameters of a Bayesian net with imprecise ones, and by re-interpreting the Markov condition through strong rather than stochastic independence. But we have already argued that this might not be the best choice in all cases. For this reason, the work in Ref. [11] imposes an epistemic Markov condition on directed trees. We discuss this condition and provide some examples, showing that it makes certain variables in the tree become epistemically many-to-many independent. Moreover, we show how the independent natural extension, and its properties proved here, are crucial stepping stones that allow us to construct the least-committal joint model over the tree that arises out of the parameters through the epistemic Markov condition. This particular type of joint allows for the development of an exact linear-time message-passing algorithm that performs imprecise-probabilistic updating of the epistemic tree. That this is at all possible, is rather surprising because of the above-mentioned perceived incompatibilities between epistemic independence and the graphoid axioms. It shows that epistemic independence has a significant role to play in probabilistic-graphical models.

We summarise our views on the results of this paper in Section 9. Appendices A and B respectively collect the proofs of all results, and the counter-examples needed to explore the relations between the many notions we introduce and study.

5In the simple case of two variables, there is no need to distinguish between many-to-one and many-to-many independence.

2. Coherent lower previsions

Let us give a brief overview of the concepts and results from the theory of coherent lower previsions that we use in this paper. We refer to Ref. [30] for an in-depth study, and to Ref. [21] for a survey.

2.1. **Lower and upper previsions.**

Consider a variable X taking values in some possibility space $\mathscr{X}$, which we assume in this paper to be finite. The theory of coherent lower previsions aims to model uncertainty about the value of X by means of lower and upper previsions of gambles. A *gamble* is a real-valued function on $\mathscr{X}$, and we denote by $\mathscr{L}(\mathscr{X})$ the set of all gambles on $\mathscr{X}$. This set is a linear space under pointwise addition of gambles, and pointwise multiplication of gambles with real numbers. For any subset $\mathscr{A}$ of $\mathscr{L}(\mathscr{X})$, we denote by $\operatorname{posi}(\mathscr{A})$ the set of all positive linear combinations of gambles in $\mathscr{A}$:

$$\operatorname{posi}(\mathscr{A}) := \left\{ \sum_{k=1}^{n} \lambda_k f_k \colon f_k \in \mathscr{A}, \lambda_k > 0, n > 0 \right\}.$$

We call $\mathscr{A}$ a *convex cone* if it is closed under positive linear combinations, meaning that $\operatorname{posi}(\mathscr{A}) = \mathscr{A}$.

For any two gambles f and g on a set $\mathscr{X}$, we write '$f \geq g$' if $(\forall x \in \mathscr{X}) f(x) \geq g(x)$, and '$f > g$' if $f \geq g$ and $f \neq g$. A gamble $f > 0$ is called *positive*. A gamble $g \leq 0$ is called *non-positive*. $\mathscr{L}(\mathscr{X})_{\neq 0}$ denotes the set of all non-zero gambles, $\mathscr{L}(\mathscr{X})_{>0}$ the convex cone of all positive gambles, and $\mathscr{L}(\mathscr{X})_{\leq 0}$ the convex cone of all non-positive gambles on $\mathscr{X}$.

A lower prevision is a real-valued functional $\underline{P}$ defined on $\mathscr{L}(\mathscr{X})$. The lower prevision $\underline{P}$ is said to be *coherent* when it satisfies the following three conditions:

- C1. $\underline{P}(f) \geq \min f$ for all $f \in \mathscr{L}(\mathscr{X})$;
- C2. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ for all $f \in \mathscr{L}(\mathscr{X})$ and real $\lambda \geq 0$;
- C3. $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$ for all $f, g \in \mathscr{L}(\mathscr{X})$.

The conjugate of a lower prevision $\underline{P}$ is called an *upper prevision*. It is denoted by $\overline{P}$, and defined by $\overline{P}(f) := -\underline{P}(-f)$ for any gamble f on $\mathscr{X}$.

One interesting particular case of lower previsions are the *vacuous* ones. Given a non-empty subset A of $\mathscr{X}$, the vacuous lower prevision $\underline{P}_A$ relative to A is given by $\underline{P}_A(f) = \min_{x \in A} f(x)$. It serves as an appropriate model for those situations where the only information we have about X is that it takes a value in the set A.

2.2. **Linear previsions and envelope theorems.**

A coherent lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X})$ satisfying $\underline{P}(f+g) = \underline{P}(f) + \underline{P}(g)$ for all $f, g \in \mathscr{L}(\mathscr{X})$ is called a *linear prevision*, and is usually denoted by P. It corresponds to an expectation operator associated with the additive probability that is its restriction to events. We denote the set of all linear previsions on $\mathscr{L}(\mathscr{X})$ by $\mathbb{P}(\mathscr{X})$. For any linear prevision P on $\mathscr{L}(\mathscr{X})$, the corresponding *mass function* p is defined by $p(x) := P(\mathbb{I}_{\{x\}})$, $x \in \mathscr{X}$, where $\mathbb{I}_{\{x\}}$ denotes the indicator of the singleton $\{x\}$. Then of course $P(f) = \sum_{x \in \mathscr{X}} f(x) p(x)$.

Linear previsions can also be used to characterise the notion of coherence for lower previsions: a lower prevision $\underline{P}$ is coherent if and only if it is the lower envelope of the closed convex set of dominating linear previsions

$$\mathscr{M}(\underline{P}) := \{P \in \mathbb{P}(\mathscr{X}) \colon (\forall f \in \mathscr{L}(\mathscr{X})) P(f) \geq \underline{P}(f)\},$$

so we have

$$\underline{P}(f) = \min\{P(f) \colon P \in \mathscr{M}(\underline{P})\}.$$

This is also equivalent to requiring that $\underline{P}$ should be the lower envelope of the set of *extreme points* of $\mathscr{M}(\underline{P})$, which we denote by $\operatorname{ext}(\mathscr{M}(\underline{P}))$:

$$\underline{P}(f) = \min\{P(f) \colon P \in \operatorname{ext}(\mathscr{M}(\underline{P}))\},$$

where P is an extreme point of $\mathscr{M}(\underline{P})$ when it cannot be written as a non-trivial convex combination of two different elements of $\mathscr{M}(\underline{P})$.

2.3. **Conditional lower previsions.** Next, consider a number of variables X_n, $n \in N$, taking values in the respective finite sets $\mathscr{X}_n$. Here N is some finite index set.

For every subset R of N, we denote by X_R the tuple of variables (with one component for each $r \in R$) that takes values in the Cartesian product $\mathscr{X}_R := \times_{r\in R}\mathscr{X}_r$. This Cartesian product is the set of all maps x_R from R to $\bigcup_{r\in R}\mathscr{X}_r$ such that $x_r := x_R(r) \in \mathscr{X}_r$ for all $r \in R$. The elements of $\mathscr{X}_R$ are generically denoted by x_R or z_R, with corresponding components $x_r := x_R(r)$ or $z_r := z_R(r)$, $r \in R$. We will always assume that the variables X_r are *logically independent*, which means that for each non-empty subset R of N, X_R may assume all values in $\mathscr{X}_R$.

We must pay particular attention to the case $R = \emptyset$. By definition, $\mathscr{X}_\emptyset$ is the set of all maps from $\emptyset$ to $\bigcup_{r\in\emptyset}\mathscr{X}_r = \emptyset$. It contains only one element $x_\emptyset$: the empty map. This means that there is no uncertainty about the value of the variable $X_\emptyset$: it can assume only one value (the empty map). Moreover $\mathbb{I}_{\{x_\emptyset\}} = 1$.

We also denote by $\mathscr{L}(\mathscr{X}_R)$ the set of gambles defined on $\mathscr{X}_R$. We will frequently use the simplifying device of identifying a gamble f_R on $\mathscr{X}_R$ with its *cylindrical extension* to $\mathscr{X}_N$, which is the gamble $\tilde{f}_R$ defined by $\tilde{f}_R(x_N) := f_R(x_R)$ for all $x_N \in \mathscr{X}_N$, where x_R is the restriction (i.e., the projection) of x_N to $\mathscr{X}_R$. To give an example, if $\mathscr{K} \subseteq \mathscr{L}(\mathscr{X}_N)$, this trick allows us to consider $\mathscr{K} \cap \mathscr{L}(\mathscr{X}_R)$ as the set of those gambles in $\mathscr{K}$ that depend only (at most) on the variable X_R. As another example, this device allows us to identify the gambles $\mathbb{I}_{\{x_R\}}$ and $\mathbb{I}_{\{x_R\}\times\mathscr{X}_{N\setminus R}}$, and therefore also the events $\{x_R\}$ and $\{x_R\} \times \mathscr{X}_{N\setminus R}$. More generally, for any event $A \subseteq \mathscr{X}_R$, we can identify the gambles $\mathbb{I}_A$ and $\mathbb{I}_{A\times\mathscr{X}_{N\setminus R}}$, and therefore also the events A and $A \times \mathscr{X}_{N\setminus R}$. In the same spirit, a lower prevision on all gambles in $\mathscr{L}(\mathscr{X}_R)$ can be identified with a lower prevision defined on the set of corresponding gambles on $\mathscr{X}_N$, a subset of $\mathscr{L}(\mathscr{X}_N)$. If in particular R is the empty-set, then $\mathscr{L}(\mathscr{X}_\emptyset)$ corresponds to the set of real numbers, which we can also identify with the set of constant gambles on $\mathscr{X}_N$.

If $\underline{P}_N$ is a coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$, then for any non-empty subset R of N we can consider its *$\mathscr{X}_R$-marginal* $\underline{P}_R$ as the coherent lower prevision on $\mathscr{L}(\mathscr{X}_R)$ defined by $\underline{P}_R(f) := \underline{P}_N(f)$ for all gambles f on $\mathscr{X}_R$: the restriction of $\underline{P}_N$ to gambles that depend only (at most) on X_R.

Given two disjoint subsets O and I of N, we define a *conditional lower prevision* $\underline{P}_{O\cup I}(\cdot|X_I)$ as a special two-place function. For any $x_I \in \mathscr{X}_I$, $\underline{P}_{O\cup I}(\cdot|x_I)$ is a real functional on the set $\mathscr{L}(\mathscr{X}_{O\cup I})$ of all gambles on $\mathscr{X}_{O\cup I}$. For any gamble f on $\mathscr{X}_{O\cup I}$, $\underline{P}_{O\cup I}(f|x_I)$ is the *lower prevision of f, conditional on $X_I = x_I$*. Moreover, the object $\underline{P}_{O\cup I}(f|X_I)$ is considered as the gamble on $\mathscr{X}_I$ that assumes the value $\underline{P}_{O\cup I}(f|x_I)$ in x_I.

We are allowing for I and O to be empty, mainly for the sake of generality and elegance in mathematical formulation and proofs. If $I = \emptyset$, then $X_I = X_\emptyset$ assumes its only possible value (the empty map $x_\emptyset$) with certainty, so conditioning on $X_\emptyset = x_\emptyset$ amounts to not conditioning at all, and $\underline{P}_{O\cup I}(f|X_I)$ is then essentially the same thing as an unconditional lower prevision $\underline{P}_O$. We will come back to the other case $O = \emptyset$ shortly.

We now turn to the most important rationality criteria for such conditional lower previsions. The conditional lower prevision $\underline{P}_{O\cup I}(\cdot|X_I)$ is called *separately coherent* when it satisfies the following three conditions for all $x_I \in \mathscr{X}_I$, non-negative λ and gambles $f, g \in \mathscr{L}(\mathscr{X}_{O\cup I})$:

SC1. $\underline{P}_{O\cup I}(f|x_I) \geq \min_{x_O\in\mathscr{X}_O} f(x_O, x_I)$;
SC2. $\underline{P}_{O\cup I}(\lambda f|x_I) = \lambda\underline{P}_{O\cup I}(f|x_I)$;
SC3. $\underline{P}_{O\cup I}(f+g|x_I) \geq \underline{P}_{O\cup I}(f|x_I) + \underline{P}_{O\cup I}(g|x_I)$.

When SC3 is satisfied with equality for all $f, g \in \mathscr{L}(\mathscr{X}_{O\cup I})$, then $\underline{P}_{O\cup I}(\cdot|X_I)$ is called a *conditional linear prevision*, and usually denoted by $P_{O\cup I}(\cdot|X_I)$. It is an expectation operator with respect to a conditional probability (or mass function).

An important consequence of separate coherence is that

$$\underline{P}_{O\cup I}(g|x_I) = \underline{P}_{O\cup I}(g(\cdot, x_I)|x_I) \text{ and } \underline{P}_{O\cup I}(fg|X_I) = f\underline{P}_{O\cup I}(g|X_I) \tag{1}$$

for all $x_I \in \mathscr{X}_I$, all non-negative gambles f on $\mathscr{X}_I$ and all gambles g on $\mathscr{X}_{O\cup I}$ [30, Theorems 6.2.4 and 6.2.6(l)]. The first equality tells us that $\underline{P}_{O\cup I}(\cdot|x_I)$ is completely determined by its behaviour on $\mathscr{L}(\mathscr{X}_O)$, and we will therefore often identify $\underline{P}_{O\cup I}(\cdot|x_I)$ with a lower prevision on $\mathscr{L}(\mathscr{X}_O)$. To prove (1), observe that if $h_1(\cdot,x_I) = h_2(\cdot,x_I)$ then it follows from SC1 that $\underline{P}_{O\cup I}(h_1-h_2|x_I) = \underline{P}_{O\cup I}(h_2-h_1|x_I) = 0$, and therefore from SC3 that $\underline{P}_{O\cup I}(h_1|x_I) = \underline{P}_{O\cup I}(h_2|x_I)$. The first equality then follows by letting $h_1 := g$ and $h_2 := g(\cdot,x_I)$.

It is clear from SC1–SC3 that $\underline{P}_{O\cup I}(\cdot|X_I)$ is separately coherent if and only if for all $x_I \in \mathscr{X}_I$, $\underline{P}_{O\cup I}(\cdot|x_I)$ is a coherent lower prevision on $\mathscr{L}(\mathscr{X}_O)$ and moreover Condition (1) holds [this second condition turns out to be equivalent to requiring that $\underline{P}_{O\cup I}(\{x_I\}|x_I) = 1$ for every $x_I \in \mathscr{X}_I$].

In the degenerate case that $O=\emptyset$, separate coherence guarantees that $\underline{P}_{O\cup I}(f|x_I) = f(x_I)$ and therefore $\underline{P}_{O\cup I}(f|X_I) = f$ for all $f \in \mathscr{L}(\mathscr{X}_I)$. When $I=\emptyset$, separate coherence of $\underline{P}_{O\cup I}(\cdot|X_I) = \underline{P}_O$ reduces to coherence of the unconditional lower prevision $\underline{P}_O$.

2.4. **The behavioural interpretation.** The coherence concepts introduced above may be better understood in terms of the behavioural interpretation of (conditional) lower previsions.[6] If we see a gamble f as an uncertain reward, then the lower prevision $\underline{P}(f)$ can be interpreted as a subject's supremum acceptable price for buying the gamble f, in the sense that it is the supremum real μ such that she considers the transaction $f-\mu$, which is equivalent to buying f for a price μ, to be desirable. It follows that she considers it desirable to buy f for any price $\underline{P}(f)-\varepsilon$, $\varepsilon>0$. Similarly, we can regard her upper prevision $\overline{P}(f)$ for f her infimum acceptable selling price for f, in the sense that it is the infimum real μ such that she considers the transaction $\mu-f$, which is equivalent to selling f for a price μ, to be desirable. In particular, she considers it desirable to sell f for any price $\overline{P}(f)+\varepsilon$, $\varepsilon>0$. And a linear prevision $P(f)$ corresponds to the case where her supremum acceptable buying price for the gamble f coincides with her infimum acceptable selling price, meaning that she expresses a preference between buying and selling the gamble f for a price μ for almost all prices μ.

If we follow this interpretation, we can similarly interpret a subject's conditional lower prevision for a gamble f conditional on a value x_I as her current supremum acceptable buying price for f if she were to find out (at some later point) that $X_I = x_I$.

A coherent unconditional lower prevision $\underline{P}$ is one for which we cannot raise the supremum acceptable buying price $\underline{P}(f)$ for any gamble f by taking into account the implications of other desirable transactions. A separately coherent conditional lower prevision $\underline{P}_{O\cup I}(\cdot|X_I)$ is one where a similar requirement holds for every conditioning event $X_I = x_I$, and where moreover our subject is currently disposed to betting at all odds on the event $X_I = x_I$ if she were to observe it at some later point.

2.5. **Coherence and weak coherence.** We now turn from separate to joint coherence. For any gamble f on $\mathscr{X}_{O\cup I}$ and any $x_I \in \mathscr{X}_I$, we define

$$G_{O\cup I}(f|x_I) := \mathbb{I}_{\{x_I\}}[f-\underline{P}_{O\cup I}(f|X_I)] = \mathbb{I}_{\{x_I\}}[f(\cdot,x_I)-\underline{P}_{O\cup I}(f(\cdot,x_I)|x_I)]$$

and

$$G_{O\cup I}(f|X_I) := f-\underline{P}_{O\cup I}(f|X_I) = \sum_{x_I\in\mathscr{X}_I} G_{O\cup I}(f|x_I) = \sum_{x_I\in\mathscr{X}_I} \mathbb{I}_{\{x_I\}}[f(\cdot,x_I)-\underline{P}_{O\cup I}(f(\cdot,x_I)|x_I)].$$

Taking into account the behavioural interpretation of conditional lower previsions summarised in Section 2.4, we may regard $G_{O\cup I}(f|x_I)$ as an *almost-desirable* gamble, in the sense that for every $\varepsilon>0$, the gamble $G_{O\cup I}(f|x_I)+\varepsilon\mathbb{I}_{\{x_I\}}$ corresponds to buying f for a price $\underline{P}_{O\cup I}(f|x_I)+\varepsilon$, contingent on the event $\{x_I\}$. And since taking finite sums of almost-desirable gambles produces almost-desirable gambles, so should $G_{O\cup I}(f|X_I)$ be.

[6] See Refs. [30, 32] for more details.

Observe that $G_{O\cup I}(f|X_I)$ is always equal to 0 when $O=\emptyset$. We also define, for any gamble f on $\mathscr{X}_{O\cup I}$, the *$\mathscr{X}_I$-support* $\operatorname{supp}_I(f)$ of f as the set of elements of $\mathscr{X}_I$ where the partial gamble $f(\cdot,x_I)$ is non-zero:

$$\operatorname{supp}_I(f) := \left\{x_I\in\mathscr{X}_I\colon \mathbb{I}_{\{x_I\}}f\neq 0\right\} = \{x_I\in\mathscr{X}_I\colon f(\cdot,x_I)\neq 0\}.$$

This support $\operatorname{supp}_I(f)$ is a subset of $\mathscr{X}_I$, but as we already mentioned before, it will be convenient to identify it with the subset $\operatorname{supp}_I(f)\times\mathscr{X}_{N\setminus I}$ of $\mathscr{X}_N$.

Consider disjoint subsets O_j and I_j of N. A collection of (separately coherent) conditional *linear* previsions $P_{O_j\cup I_j}(\cdot|X_{I_j})$ defined on the sets of gambles $\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$, $j\in\{1,\dots,m\}$ is called *(strongly) coherent* if for all $f_j\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$, $j\in\{1,\dots,m\}$, there is some $z_N\in\bigcup_{j=1}^m\operatorname{supp}_{I_j}(f_j)$ such that:

$$\left[\sum_{j=1}^m G_{O_j\cup I_j}(f_j|X_{I_j})\right](z_N)\geq 0. \tag{2}$$

The (separately coherent) conditional *lower* previsions $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$ defined on the sets of gambles $\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$, $j\in\{1,\dots,m\}$ are called *coherent* if and only if they are lower envelopes of a collection $\left\{P^{\lambda}_{O_j\cup I_j}(\cdot|X_{I_j})\colon \lambda\in\Lambda\right\}$ of coherent conditional linear previsions. This is equivalent to requiring that for all $f_j\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$ where $j\in\{1,\dots,m\}$, all $k\in\{1,\dots,m\}$, all $x_{I_k}\in\mathscr{X}_{I_k}$ and all $g\in\mathscr{L}(\mathscr{X}_{O_k\cup I_k})$, there is some $z_N\in\{x_{I_k}\}\cup\bigcup_{j=1}^m\operatorname{supp}_{I_j}(f_j)$ such that:

$$\left[\sum_{j=1}^m G_{O_j\cup I_j}(f_j|X_{I_j})-G_{O_k\cup I_k}(g|x_{I_k})\right](z_N)\geq 0. \tag{3}$$

We say that the conditional lower previsions $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$ are *weakly coherent* if for all $f_j\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$ where $j\in\{1,\dots,m\}$, all $k\in\{1,\dots,m\}$, all $x_{I_k}\in\mathscr{X}_{I_k}$ and all $g\in\mathscr{L}(\mathscr{X}_{O_k\cup I_k})$, there is some $z_N\in\mathscr{X}_N$ such that:

$$\left[\sum_{j=1}^m G_{O_j\cup I_j}(f_j|X_{I_j})-G_{O_k\cup I_k}(g|x_{I_k})\right](z_N)\geq 0.$$

This condition requires that our subject should not be able to raise her supremum acceptable buying price $\underline{P}_{O_k\cup I_k}(g|x_{I_k})$ for a gamble g contingent on $\{x_{I_k}\}$ by taking into account the implications of other conditional assessments. However, under the behavioural interpretation, a collection of weakly coherent conditional lower previsions can still present some forms of inconsistency with one another. See Refs. [30, Chapter 7], [22] and [31] for discussion. These inconsistencies are eliminated by the stronger notion of coherence given by Eq. (3), where we focus only on the elements in the supports of the gambles.

If the conditional lower previsions $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$ are coherent, then they are clearly also weakly coherent. The following characterisation of weak coherence will be useful. The equivalence between the first two statements was proved in Ref. [24, Theorem 1], while the equivalence between the second and third statements is a consequence of Ref. [30, Section 6.5.4].

Theorem 1. *The conditional lower previsions $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$, $j=1,\dots,m$, are weakly coherent if and only if there is some coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ satisfying any (and hence all) of the following equivalent conditions:*

(a) *$\underline{P}_N$ and $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$, $j=1,\dots,m$ are weakly coherent;*
(b) *For all $j=1,\dots,m$, $\underline{P}_N$ and $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$ are pairwise coherent;*
(c) *For all $j=1,\dots,m$, all $x_{I_j}\in\mathscr{X}_{I_j}$ and all gambles f on $\mathscr{X}_{O_j\cup I_j}$:*

$$\underline{P}_N(G_{O_j\cup I_j}(f|x_{I_j}))=0. \tag{GBR}$$

The last condition in this theorem is called the *Generalised Bayes Rule*, and reduces to Bayes's rule in the case of linear conditional and unconditional previsions. A consequence of (GBR) in our context is that

$$\underline{P}_N(\overline{P}_{O_j\cup I_j}(f|X_{I_j})) \geq \underline{P}_N(f) \geq \underline{P}_N(\underline{P}_{O_j\cup I_j}(f|X_{I_j})) \text{ for every gamble } f \text{ on } \mathscr{X}_{O_j\cup I_j}. \quad (4)$$

We will also use the following result in our argumentation.

Theorem 2 (Reduction Theorem [30, Theorem 7.1.5])**.** *Let $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$ be separately coherent conditional lower previsions defined on the sets of gambles $\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$ with $I_j \neq \emptyset$ for all $j = 1,\dots,m$, and let $\underline{P}_N$ be a coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$. Then $\underline{P}_N$ and $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$, $j = 1,\dots,m$ are coherent if and only if the following two conditions hold:*

(a) *$\underline{P}_N$ and $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$, $j = 1,\dots,m$ are weakly coherent;*
(b) *$\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$, $j = 1,\dots,m$ are coherent.*

2.6. **Natural and regular extension.** Let $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$ be coherent conditional lower previsions defined on the sets of gambles $\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$, $j = 1,\dots,m$. If we now consider any disjoint subsets O and I of N, the natural extension $\underline{P}_{O\cup I}(\cdot|X_I)$ of these conditional lower previsions is defined on $\mathscr{L}(\mathscr{X}_{O\cup I})$ by

$$\underline{E}_{O\cup I}(f|x_I) := \sup\Bigg\{\mu\colon \Bigg[\sum_{j=1}^m G_{O_j\cup I_j}(g_j|X_{I_j}) - \mathbb{I}_{\{x_I\}}(f-\mu)\Bigg] < 0 \\ \text{on } \{x_I\} \cup \bigcup_{j=1}^m \operatorname{supp}_{I_j}(g_j) \text{ for some } g_j \in \mathscr{L}(\mathscr{X}_{O_j\cup I_j})\Bigg\}$$

for all $f \in \mathscr{L}(\mathscr{X}_{O\cup I})$ and all $x_I \in \mathscr{X}_I$. $\underline{E}_{O\cup I}(f|x_I)$ represents the supremum acceptable buying price for a gamble f contingent on $\{x_I\}$ that can be derived from the assessments in $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j}), j = 1,\dots,m$ using arguments of coherence. See Refs. [30, Chapter 8], [22] and [31] for additional information.

In particular, the *unconditional natural extension* of $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j}), j = 1,\dots,m$ is given by

$$\underline{E}_N(f) := \sup\Bigg\{\min\Bigg[f - \sum_{j=1}^m G_{O_j\cup I_j}(g_j|X_{I_j})\Bigg] : g_j \in \mathscr{L}(\mathscr{X}_{O_j\cup I_j})\Bigg\}, \quad (5)$$

for all gambles f on $\mathscr{X}_N$, and it is the pointwise smallest coherent lower prevision that is coherent with the $\underline{P}_{O_j\cup I_j}(\cdot|X_{I_j})$, $j = 1,\dots,m$.

Another particular case of interest is when we want to derive conditional lower previsions from unconditional ones. Consider a subset I of N, and a coherent lower prevision $\underline{P}_N$ on $\mathscr{X}_N$. The natural extension $\underline{E}_N(\cdot|X_I)$ of $\underline{P}_N$ to a lower prevision on $\mathscr{L}(\mathscr{X}_N)$ conditional on X_I is given by

$$\underline{E}_N(f|x_I) = \begin{cases} \max\left\{\mu \in \mathbb{R}\colon \underline{P}_N(\mathbb{I}_{\{x_I\}}[f-\mu]) \geq 0\right\} & \text{if } \underline{P}_N(\{x_I\}) > 0 \\ \min_{z_{N\setminus I}\in\mathscr{X}_{N\setminus I}} f(z_{N\setminus I}, x_I) & \text{if } \underline{P}_N(\{x_I\}) = 0. \end{cases}$$

It defines a separately coherent conditional lower prevision that is also coherent with $\underline{P}_N$, and it is indeed the smallest such conditional lower prevision.

On the other hand, the *regular extension* $\underline{R}(\cdot|X_I)$ of $\underline{P}_N$ to a lower prevision on $\mathscr{L}(\mathscr{X}_N)$ conditional on X_I is given by

$$\underline{R}_N(f|x_I) := \begin{cases} \max\left\{\mu \in \mathbb{R}\colon \underline{P}_N(\mathbb{I}_{\{x_I\}}[f-\mu]) \geq 0\right\} & \text{if } \overline{P}_N(\{x_I\}) > 0 \\ \min_{z_{N\setminus I}\in\mathscr{X}_{N\setminus I}} f(z_{N\setminus I}, x_I) & \text{if } \overline{P}_N(\{x_I\}) = 0. \end{cases}$$

The natural and regular extensions coincide unless $\overline{P}_N(\{x_I\}) > \underline{P}_N(\{x_I\}) = 0$, in which case we may have $\underline{R}_N(f|x_I) > \underline{E}_N(f|x_I)$. In fact, when $\underline{P}_N(\{x_I\}) > 0$ there is a unique value of $\underline{P}_N(f|x_I)$ satisfying (GBR) with respect to $\underline{P}_N$, but this is no longer true if $\underline{P}_N(\{x_I\}) = 0$. See Refs. [22, 25] for additional information.

The regular extension defines a separately coherent conditional lower prevision that is also coherent with $\underline{P}_N$, and we will have occasion to use it as a tool for deriving conditional lower previsions from unconditional ones. The following result, which follows from Theorem 6 in Ref. [22], makes regular extension especially useful in this respect:

Theorem 3. *Let $\underline{P}_N$ be a coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$. Consider disjoint O_j and I_j for $j = 1,\dots,m$. Assume that $\overline{P}_N(\{x_{I_j}\}) > 0$ for all $x_{I_j} \in \mathscr{X}_{I_j}$, and define $\underline{P}_{O_j \cup I_j}(\cdot|X_{I_j})$ using regular extension for $j = 1,\dots,m$. Then $\underline{P}_N$ is coherent with the conditional lower previsions $\underline{P}_{O_1\cup I_1}(\cdot|X_{I_1})$, ..., $\underline{P}_{O_m\cup I_m}(\cdot|X_{I_m})$.*

3. The formal approach to independence

3.1. **Basic definitions.** Consider a number of (logically independent) variables X_n, $n \in N$, assuming values in the respective finite sets $\mathscr{X}_n$. Here N is some finite index set. We assume that for each of these variables X_n, we have an uncertainty model for the values that it assumes in $\mathscr{X}_n$, in the form of a coherent lower prevision $\underline{P}_n$ on the set $\mathscr{L}(\mathscr{X}_n)$ of all gambles (real-valued maps) on $\mathscr{X}_n$.

We begin our discussion of independence by following the formalist route: we introduce a number of interesting generalisations of the notion of an independent product of linear previsions.

The first is a stronger, symmetrised version of the notion of 'forward factorisation' that was introduced elsewhere [12].

Definition 1 (Productivity)**.** Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. We call this lower prevision *productive* if for all disjoint subsets[7] I and O of N, all $g \in \mathscr{L}(\mathscr{X}_O)$ and all non-negative $f \in \mathscr{L}(\mathscr{X}_I)$, $\underline{P}_N(f[g-\underline{P}_N(g)]) \geq 0$.

The intuition behind this definition is that a coherent lower prevision $\underline{P}_N$ is productive if multiplying an almost-desirable gamble on X_O (the gamble $g-\underline{P}_N(g)$, which has lower prevision zero) with any non-negative gamble f that depends on a different variable X_I, preserves its almost-desirability, in the sense that the lower prevision of the product is non-negative. In other words, if we construct a gamble on $\mathscr{X}_{O\cup I}$ by piecing together almost-desirable gambles from $\mathscr{L}(\mathscr{X}_O)$, we obtain a gamble that is still almost-desirable.

A lower envelope $\underline{P}_N$ of productive coherent lower previsions $\underline{P}_\lambda$, $\lambda \in \Lambda$ is again productive: for all disjoint subsets I and O of N, all $g \in \mathscr{L}(\mathscr{X}_O)$ and all non-negative $f \in \mathscr{L}(\mathscr{X}_I)$, we have that $\underline{P}_\lambda(f[g-\underline{P}_N(g)]) \geq \underline{P}_\lambda(f[g-\underline{P}_\lambda(g)]) \geq 0$, and therefore indeed $\underline{P}_N(f[g-\underline{P}_N(g)]) \geq 0$.

In a paper [12] on laws of large numbers for coherent lower previsions, which generalises and subsumes most known versions in the literature, we have proved that the condition for forward irrelevance[8] (which is implied by the present productivity condition) is sufficient for a weak law of large numbers to hold. So we are led to the following immediate conclusion.

Theorem 4 (Weak law of large numbers [12, Theorem 2])**.** *Let the coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ be productive. Let $\varepsilon > 0$ and consider arbitrary gambles h_n on $\mathscr{X}_n$, $n \in N$. Let B be a common bound for the ranges of these gambles and let* $\min h_n \leq m_n \leq \underline{P}_N(h_n) \leq \overline{P}_N(h_n) \leq M_n \leq \max h_n$ *for all $n \in N$. Then*

$$\underline{P}_N\left(\left\{x_N \in \mathscr{X}_N \colon \sum_{n\in N}\frac{m_n}{|N|} - \varepsilon \leq \sum_{n\in N}\frac{h_n(x_n)}{|N|} \leq \sum_{n\in N}\frac{M_n}{|N|} + \varepsilon\right\}\right) \geq 1 - 2\exp\left(-\frac{|N|\,\varepsilon^2}{4B^2}\right).$$

Next comes a version of a condition that has proved quite useful in the context of research on credal networks [11], which we shall discuss in Section 8.

---

[7]The coherence of $\underline{P}_N$ guarantees that for empty I or O the corresponding condition is trivially satisfied.

[8]This condition is a particular instance of the epistemic irrelevance we discuss in Section 4: if we consider n variables X_1, ..., X_n, then for all $k = 2,\dots,n$ the variables X_1, ..., X_{k-1} are epistemically irrelevant to X_k. See Ref. [12, 13] for more information.

Definition 2 (Factorisation)**.** Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. We call this lower prevision

(i) *factorising*[9] if for all $o \in N$ and all $I \subseteq N \setminus \{o\}$,[10] all $f_o \in \mathscr{L}(\mathscr{X}_o)$ and all non-negative $f_i \in \mathscr{L}(\mathscr{X}_i)$, $i \in I$, $\underline{P}_N(f_I f_o) = \underline{P}_N(f_I \underline{P}_N(f_o))$, where $f_I := \prod_{i\in I} f_i$;

(ii) *strongly factorising* if $\underline{P}_N(fg) = \underline{P}_N(f\underline{P}_N(g))$ for all $g \in \mathscr{L}(\mathscr{X}_O)$ and non-negative $f \in \mathscr{L}(\mathscr{X}_I)$, where I and O are any [11] disjoint subsets of N.

It will at this point be useful to introduce the following notation. Consider a real interval $\underline{\overline{a}} := [\underline{a}, \overline{a}]$ and a real number b, then

$$\underline{\overline{a}} \odot b := \begin{cases} \underline{a}b & \text{if } b \geq 0 \\ \overline{a}b & \text{if } b \leq 0. \end{cases}$$

Also, we denote by $\underline{\overline{P}}(f)$ the interval $[\underline{P}(f), \overline{P}(f)]$.

It follows from the coherence of $\underline{P}_N$ that given a factorising (respectively strongly factorising) coherent lower prevision we also get $\underline{P}_N(f_I \underline{P}_N(f_o)) = \underline{\overline{P}}_N(f_I) \odot \underline{P}_N(f_o)$ (respectively $\underline{P}_N(f\underline{P}_N(g)) = \underline{\overline{P}}_N(f) \odot \underline{P}_N(g)$) in Definition 2. We then have the following characterisations of factorising lower previsions, based on their coherence and conjugacy properties.

Proposition 5. *A coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ is factorising if and only if for all $o \in N$, all $f_o \in \mathscr{L}(\mathscr{X}_o)$ and all non-negative $f_i \in \mathscr{L}(\mathscr{X}_i)$, $i \in N \setminus \{o\}$, any (and hence all) of the following equivalent conditions holds:*

(i) $\underline{P}_N(\prod_{n\in N} f_n) = \underline{P}_N(\underline{P}_N(f_o) \prod_{i\in N\setminus\{o\}} f_i) = \underline{\overline{P}}_N(\prod_{i\in N\setminus\{o\}} f_i) \odot \underline{P}_N(f_o)$;

(ii) $\underline{P}_N(\prod_{n\in N} f_n) = \begin{cases} \underline{P}_N(f_o) \prod_{i\in N\setminus\{o\}} \underline{P}_N(f_i) & \textit{if } \underline{P}_N(f_o) \geq 0 \\ \underline{P}_N(f_o) \prod_{i\in N\setminus\{o\}} \overline{P}_N(f_i) & \textit{if } \underline{P}_N(f_o) \leq 0. \end{cases}$

The difference between factorisation and strong factorisation lies in the types of products considered: in the first case, we only consider gambles that are products of non-negative gambles each depending on a single variable, while in the second case the non-negative gamble considered need not be such a product.

Finally, there is the property that the late Russian mathematician Vladimir Kuznetsov first drew attention to [19]. In order to define it, we use $\boxtimes$ to denote the (commutative and associative) *interval product* operator defined by:

$$\begin{aligned} [a,b] \boxtimes [c,d] &:= \{xy \colon x \in [a,b] \text{ and } y \in [c,d]\} \\ &= [\min\{ac,ad,bc,bd\}, \max\{ac,ad,bc,bd\}] \text{ for all } a \leq b \text{ and } c \leq d \text{ in } \mathbb{R}. \end{aligned}$$

Definition 3 (Kuznetsov product)**.** Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. We call this lower prevision

(i) a *Kuznetsov product*, or simply, *Kuznetsov*, if $\underline{\overline{P}}_N(\prod_{n\in N} f_n) = \boxtimes_{n\in N} \underline{\overline{P}}_N(f_n)$ for all $f_n \in \mathscr{L}(\mathscr{X}_n)$, $n \in N$.

(ii) a *strong Kuznetsov product*, or simply, *strongly Kuznetsov*, if $\underline{\overline{P}}_N(fg) = \underline{\overline{P}}_N(f) \boxtimes \underline{\overline{P}}_N(g)$ for all $g \in \mathscr{L}(\mathscr{X}_O)$ and all $f \in \mathscr{L}(\mathscr{X}_I)$, where I and O are any disjoint subsets of N.[11]

These two properties are based on the sensitivity analysis interpretation of coherent lower previsions: if we consider a product of gambles and for each of gamble we have an interval of possible values for its expectation, the Kuznetsov property implies that the interval of possible values for the expectation of the product coincides with the product of the intervals of the expectations of the different gambles. As before, the difference between being a

[9] The present notion of factorisation when restricted to lower probabilities and events, is called *strict factorisation* in Ref. [29].

[10] The coherence of $\underline{P}_N$ guarantees that for empty I the corresponding condition is trivially satisfied.

[11] The coherence of $\underline{P}_N$ guarantees that for empty I or O the corresponding condition is trivially satisfied.

Kuznetsov product and being a strong Kuznetsov product resides in whether the gambles we are multiplying depend on one variable only or on several variables at once.

We will show later that all the properties introduced above generalise the factorisation property of precise probabilities to the imprecise case. Here are the general relationships between them:

Proposition 6. *Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. Then*

$$\begin{array}{ccccc} \underline{P}_N \text{ is strongly Kuznetsov} & \Rightarrow & \underline{P}_N \text{ is strongly factorising} & \Rightarrow & \underline{P}_N \text{ is productive} \\ \Downarrow & & \Downarrow & & \\ \underline{P}_N \text{ is Kuznetsov} & \Rightarrow & \underline{P}_N \text{ is factorising.} & & \end{array}$$

The intuition behind these implications is the following (see Appendix A for a detailed proof). On the one hand, being strongly Kuznetsov clearly implies being Kuznetsov, and strong factorisation implies factorisation, because we allow for more involved products of gambles in the definition of the former. On the other hand, the factorisation conditions focus on the lower prevision only, whereas the Kuznetsov ones involve both lower and the upper previsions, and therefore give rise to stronger conditions. And finally, it follows from its definition that a strongly factorising coherent lower prevision satisfies the condition of productivity with equality instead of with inequality.

In Appendix B we present examples showing that the converses of the implications in this proposition do not hold in general. Specifically, in Example 3 we give a coherent lower prevision that satisfies productivity but none of the other properties; in Example 5 we give a coherent lower prevision that is strongly factorising (and as a consequence also factorising and productive) but not Kuznetsov (and therefore not strongly Kuznetsov); and in Example 6 we have a factorising coherent lower prevision that satisfies none of the other properties. The only related open problem at this point is whether being Kuznetsov is generally equivalent to being strongly Kuznetsov.

We will show that the independent natural extension from Section 5 is strongly factorising but not Kuznetsov in general. In the rest of this section, we look at a number of special cases that will prove instrumental in what follows. In particular, the strong product we will study in Section 3.3 satisfies all the properties we have introduced here.

3.2. **The product of linear previsions.** If we have linear previsions P_n on $\mathscr{L}(\mathscr{X}_n)$ with corresponding mass functions p_n, then their *product* $S_N := \times_{n \in N} P_n$ is the linear prevision on $\mathscr{L}(\mathscr{X}_N)$ defined as

$$S_N(f) = \sum_{x_N \in \mathscr{X}_N} f(x_N) \prod_{n \in N} p_n(x_n) \text{ for all } f \in \mathscr{L}(\mathscr{X}_N). \tag{6}$$

For any non-empty subset R of N, we also denote by $S_R := \times_{r \in R} P_r$ the product of the linear previsions P_r, $r \in R$.

Useful, and immediate, are the following marginalisation and associativity properties of the product of linear previsions. They imply that for linear previsions all the properties introduced in Section 3.1 coincide.

Proposition 7. *Consider arbitrary linear previsions P_n on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$.*

(i) *For any non-empty subset R of N, S_R is the $\mathscr{X}_R$-marginal of S_N: $S_N(g) = S_R(g)$ for all gambles g on $\mathscr{X}_R$;*
(ii) *For any partition N_1 and N_2 of N, $\times_{n \in N_1 \cup N_2} P_n = (\times_{n \in N_1} P_n) \times (\times_{n \in N_2} P_n)$, or in other words, $S_N = S_{N_1} \times S_{N_2}$.*

Moreover, for any linear prevision P_N on $\mathscr{L}(\mathscr{X}_N)$, the following statements are equivalent:

(a) *$P_N = \times_{n \in N} P_n$ is the product of its marginals P_n;*
(b) *$P_N(\prod_{n \in N} f_n) = \prod_{n \in N} P_n(f_n)$ for all $f_n \in \mathscr{L}(\mathscr{X}_n)$, $n \in N$;*
(c) *P_N is strongly Kuznetsov;*

(d) P_N *is Kuznetsov;*
(e) P_N *is strongly factorising;*
(f) P_N *is factorising;*
(g) P_N *is productive.*

3.3. **The strong product of coherent lower previsions.** In a similar vein, if we have coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, then their so-called *strong product* [30, Section 9.3.5][12] $\underline{S}_N \coloneqq \times_{n\in N}\underline{P}_n$ is defined as the coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$ that is the lower envelope of the set of independent products $\{\times_{n\in N}P_n\colon (\forall n\in N)P_n\in\mathscr{M}(\underline{P}_n)\}$, or equivalently, of the set $\{\times_{n\in N}P_n\colon (\forall n\in N)P_n\in\operatorname{ext}(\mathscr{M}(\underline{P}_n))\}$. So for every $f\in\mathscr{L}(\mathscr{X}_N)$:

$$\begin{aligned}\underline{S}_N(f) &= \inf\{\times_{n\in N}P_n(f)\colon (\forall n\in N)P_n\in\mathscr{M}(\underline{P}_n)\} &(7)\\ &= \inf\{\times_{n\in N}P_n(f)\colon (\forall n\in N)P_n\in\operatorname{ext}(\mathscr{M}(\underline{P}_n))\}. &(8)\end{aligned}$$

For any non-empty subset R of N, we also denote by $\underline{S}_R \coloneqq \times_{r\in R}\underline{P}_r$ the strong product of the coherent lower previsions $\underline{P}_r$, $r\in R$. Like the product of linear previsions, the strong product of lower previsions satisfies the following marginalisation and associativity properties, and also all the properties defined in Section 3.1:[13]

Proposition 8. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n\in N$.*
(i) *For any non-empty subset R of N, $\underline{S}_R$ is the $\mathscr{X}_R$-marginal of $\underline{S}_N$: $\underline{S}_N(g)=\underline{S}_R(g)$ for all gambles g on $\mathscr{X}_R$;*
(ii) $\operatorname{ext}(\mathscr{M}(\underline{S}_N))=\{\times_{n\in N}P_n\colon (\forall n\in N)P_n\in\operatorname{ext}(\mathscr{M}(\underline{P}_n))\}$*;*
(iii) *For any partition N_1 and N_2 of N, $\times_{n\in N_1\cup N_2}\underline{P}_n=(\times_{n\in N_1}\underline{P}_n)\times(\times_{n\in N_2}\underline{P}_n)$, or in other words, $\underline{S}_N=\underline{S}_{N_1}\times\underline{S}_{N_2}$.*
(iv) *The strong product $\underline{S}_N$ is strongly Kuznetsov, and therefore also Kuznetsov, strongly factorising, factorising and productive.*

The nice characterisation of the set $\operatorname{ext}(\mathscr{M}(\underline{S}_N))$ in the second statement guarantees that the infima in Eqs. (7) and (8) are actually minima. This set of extreme points may be infinite even when dealing with finite $\mathscr{X}_n$, $n\in N$, and therefore hence the second statement is not immediate. In addition, this result guarantees, amongst other things, that the strong product satisfies the weak law of large numbers of Theorem 4.

This marks a preliminary end to our formalist discussion of independence for coherent lower previsions. In the next section, we turn to the treatment of independence following an epistemic and coherentist approach, where independence is considered to be an assessment a subject makes. The more formalist thread will be taken up again in Section 6.

4. Epistemic irrelevance and independence

Consider two disjoint subsets I and O of N. We say that a subject judges that X_I *is epistemically irrelevant to* X_O when she assumes that learning which value X_I assumes in $\mathscr{X}_I$ will not affect her beliefs about X_O. Taking into account the behavioural interpretation of coherent lower previsions summarised in Section 2.4, this means that the subject's supremum acceptable buying price for a gamble f contingent on the event that $X_I=x_I$ coincides with her supremum acceptable buying price for f, irrespective of the value $x_I\in\mathscr{X}_I$ that is observed.

Now assume that our subject has a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. If she assesses that X_I is epistemically irrelevant to X_O, this implies that she can infer from her joint model $\underline{P}_N$ the following conditional model $\underline{P}_{O\cup I}(\cdot|X_I)$ on the set $\mathscr{L}(\mathscr{X}_{O\cup I})$:

$$\underline{P}_{O\cup I}(h|x_I)\coloneqq\underline{P}_N(h(\cdot,x_I))\text{ for all gambles }h\text{ on }\mathscr{X}_{O\cup I}\text{ and all }x_I\in\mathscr{X}_I.$$

[12] Walley [30, Section 9.3.5] calls this lower prevision the *type-1 product*. The term 'strong product' seems to go back to Cozman [3].

[13] For the case of two variables ($N=\{1,2\}$), Cozman [3] was the first to prove that the strong product is Kuznetsov.

So an assessment of epistemic irrelevance is useful because it allows us to derive conditional lower previsions from unconditional ones. It follows from the comments in Section 2 that such an assessment is trivial when either O or I is the empty set. Similarly, if the gamble h depends only on X_O, then the conditional lower prevision for h conditional on some value $x_I \in \mathscr{X}_I$ coincides with the unconditional lower prevision of h.

It should be clear from the previous discussion that epistemic irrelevance is an asymmetric notion: we only require that X_I is epistemically irrelevant to X_O; and this does not necessarily imply that X_O should be epistemically irrelevant to X_I. Refs. [6, 13] include thorough discussions of this point, along with examples illustrating how such asymmetries may appear in concrete cases. Because of this, it becomes necessary to pay special attention to the mutual irrelevance of two or more variables. We then talk about epistemic independence, a suitably symmetrised version of epistemic irrelevance.

4.1. **Epistemic many-to-many independence.** We say that a subject judges the variables X_n, $n \in N$, to be *epistemically many-to-many independent* when she assumes that learning the value of any number of these variables will not affect her beliefs about the others. In other words, if she judges for any disjoint subsets I and O of N that X_I is epistemically irrelevant to X_O.

Again, if our subject has a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$, and she assesses that the variables X_n, $n \in N$, are *epistemically many-to-many independent*, then she can infer from her joint model $\underline{P}_N$ a family of conditional models

$$\mathscr{I}(\underline{P}_N) := \{\underline{P}_{O\cup I}(\cdot|X_I) \colon I \text{ and } O \text{ disjoint subsets of } N\},$$

where $\underline{P}_{O\cup I}(\cdot|X_I)$ is the coherent lower prevision on $\mathscr{L}(\mathscr{X}_{O\cup I})$ given by:

$$\underline{P}_{O\cup I}(h|x_I) := \underline{P}_N(h(\cdot,x_I)) \text{ for all gambles } h \text{ on } \mathscr{X}_{O\cup I} \text{ and all } x_I \in \mathscr{X}_I.$$

The crucial idea that the arguments in this paper hinge on, is that a coherent lower prevision $\underline{P}_N$ *expresses independence* when it does not lead to incoherence when combined with an assessment of epistemic independence: the lower prevision $\underline{P}_N$ and the conditional supremum acceptable buying prices derived from it using epistemic independence should not violate the consistency conditions introduced in Section 2.

Definition 4 (Many-to-many independence)**.** A coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ is called *many-to-many independent* if it is coherent with the family of conditional lower previsions $\mathscr{I}(\underline{P}_N)$. For a collection of coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, any many-to-many independent coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ that coincides with the $\underline{P}_n$ on their domains $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, is called a *many-to-many independent product* of these marginals.

As we show in the following example, this requirement of coherence of the unconditional and resulting conditional models is by no means trivial: not all coherent lower previsions $\underline{P}_N$ express—are compatible with an assessment of—epistemic independence.

Example 1 (Independence is not trivial)**.** Consider $\mathscr{X}_1 = \mathscr{X}_2 = \{0,1\}$, and let $P_{\{1,2\}}$ be the linear prevision determined by $P_{\{1,2\}}(\{(0,0)\}) = P_{\{1,2\}}(\{(1,1)\}) = 1/2$. Consider the gamble $f := \mathbb{I}_{\{1\}}$ on $\mathscr{X}_1$, so $P_{\{1,2\}}(f) = 1/2$. On the other hand, with $x_2 = 1$, we get

$$P_{\{1,2\}}(\mathbb{I}_{\{x_2\}}[f - P_{\{1,2\}}(f)]) = \frac{1}{2} - \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4} \neq 0.$$

This shows that $P_{\{1,2\}}$ is not coherent with the conditional prevision $P_{\{1,2\}}(\cdot|X_2)$, and as a consequence it is not many-to-many independent. ⧫

That a coherent lower prevision is not a many-to-many independent product need not mean that it cannot be coherently updated: in the case of Example 1, we can always do so by applying Bayes's rule. What it does mean is that it cannot be coherently updated in such a way that at the same time the epistemic independence conditions are satisfied.

We shall see examples further on that seem to suggest that being a many-to-many independent product might be too weak a requirement in certain situations: Example 3 in Appendix B establishes the existence of a non-vacuous many-to-many independent product with vacuous marginals, and intuition might suggest that an independent product of vacuous marginals should be vacuous. The reason why such examples exist, is that coherence with the conditional lower previsions induced by the marginals can sometimes be very easy to satisfy, as Proposition 26 further on will show. Because of this, it might be thought useful to require, in addition, some of the factorisation conditions from Section 3. We come back to this issue in the Conclusions.

4.2. **Epistemic many-to-one independence.** There is a weaker notion of independence that we will consider here, which is mainly useful as a kind of catalyst, facilitating our search for the many-to-many independent products we are really after. We will coin the term 'epistemic many-to-one independence' to identify it. We say that a subject judges the variables X_n, $n \in N$, to be *epistemically many-to-one independent* when she assumes that learning the value of any number of these variables will not affect her beliefs about any *single* other. In other words, if she judges for any $o \in N$ and any subset I of $N \setminus \{o\}$ that X_I is epistemically irrelevant to X_o.

Once again, if our subject has a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$, and she assesses that the variables X_n, $n \in N$, are *epistemically many-to-one independent*, then she can infer from her joint model $\underline{P}_N$ a family of conditional models

$$\mathscr{N}(\underline{P}_n, n \in N) := \left\{ \underline{P}_{\{o\}\cup I}(\cdot|X_I) \colon o \in N \text{ and } I \subseteq N \setminus \{o\} \right\},$$

where $\underline{P}_{\{o\}\cup I}(\cdot|X_I)$ is the coherent lower prevision on $\mathscr{L}(\mathscr{X}_{\{o\}\cup I})$ given by:

$$\underline{P}_{\{o\}\cup I}(h|x_I) := \underline{P}_N(h(\cdot, x_I)) = \underline{P}_o(h(\cdot, x_I)) \text{ for all } h \in \mathscr{L}(\mathscr{X}_{\{o\}\cup I}) \text{ and } x_I \in \mathscr{X}_I, \tag{9}$$

where of course $\underline{P}_o$ is the $\mathscr{X}_o$-marginal lower prevision of $\underline{P}_N$. In the set $\mathscr{N}(\underline{P}_n, n \in N)$ we are also allowing for empty I, in which case the conditional lower prevision $\underline{P}_{\{o\}\cup I}(\cdot|X_I)$ reduces to the marginal $\underline{P}_o$. So we see that the family of conditional lower previsions $\mathscr{N}(\underline{P}_n, n \in N)$ only depends on the joint model $\underline{P}_N$ through its $\mathscr{X}_n$-marginals $\underline{P}_n$, $n \in N$ (which, of course, explains our notation for it). This allows us, in particular, to define the family $\mathscr{N}(\underline{P}_n, n \in N)$ also starting from coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, rather than from a joint $\underline{P}_N$. The distinction between the two cases will be clear from the context. We use this observation in the definition of the many-to-one independent product of given marginals:

Definition 5. A coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ with marginals $\underline{P}_n, n \in N$, is called *many-to-one independent* if it is coherent with the family $\mathscr{N}(\underline{P}_n, n \in N)$. For a collection of coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, any coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ that is coherent with the family $\mathscr{N}(\underline{P}_n, n \in N)$ is called a *many-to-one independent product* of these lower previsions $\underline{P}_n$.

If a coherent lower prevision $\underline{P}_N$ is many-to-many independent, then it is also many-to-one independent: if $\underline{P}_N$ is coherent with the family $\mathscr{I}(\underline{P}_N)$, it is certainly also coherent with the subfamily $\mathscr{N}(\underline{P}_n, n \in N)$ of $\mathscr{I}(\underline{P}_N)$. Trivially, both conditions are equivalent when $N = \{1, 2\}$, and as a consequence we can also use Example 1 to conclude that also many-to-one independence is not trivial. However, many-to-many and many-to-one independence are generally not equivalent when the set N has more than two elements: an explicit example is provided in Example 6 in Appendix B.

Any many-to-one independent product of the coherent lower previsions $\underline{P}_n$, $n \in N$, must have these lower previsions as its marginals. This follows by applying the coherence condition in particular to the pairs $\underline{P}_N$ and $\underline{P}_{\{n\}\cup\emptyset}(\cdot|X_\emptyset)$, as will be made explicit in Corollary 15.

4.3. **Useful basic properties.** A basic coherence result [30, Theorem 7.1.6] states that taking lower envelopes of a family of coherent conditional lower previsions produces coherent conditional lower previsions. As a consequence, if we consider a family $\underline{P}_N^\lambda$, $\lambda \in \Lambda$, of many-to-one independent products, then for each λ the lower prevision $\underline{P}_N^\lambda$ is coherent with the family of conditional lower previsions $\mathscr{N}(\underline{P}_n^\lambda, n \in N)$. By taking lower envelopes, we deduce that $\underline{P}_N := \inf_{\lambda\in\Lambda} \underline{P}_N^\lambda$ is coherent with the lower envelopes of $\mathscr{N}(\underline{P}_n^\lambda, n \in N)$, which are precisely the family $\mathscr{N}(\underline{P}_n, n \in N)$ of conditional lower previsions that can be derived from $\underline{P}_N$ using epistemic irrelevance: the marginal lower previsions of $\underline{P}_N$ are the lower envelopes of the marginal lower previsions of $\underline{P}_N^\lambda$. Hence, $\underline{P}_N$ is also a many-to-one independent product. A similar argument shows that many-to-many independence is preserved by taking lower envelopes.

We also have the following marginalisation and associativity properties.

Proposition 9. *Consider arbitrary coherent lower previsions $\underline{P}_n$, $n \in N$. Let $\underline{P}_N$ be any many-to-one independent product and $\underline{Q}_N$ any many-to-many independent product of the marginals $\underline{P}_n$, $n \in N$. Let R and S be any subsets of N.*

(i) *The $\mathscr{X}_R$-marginal $\underline{P}_R$ of $\underline{P}_N$ is a many-to-one independent product of its marginals $\underline{P}_r$, $r \in R$;*
(ii) *The $\mathscr{X}_R$-marginal $\underline{Q}_R$ of $\underline{Q}_N$ is a many-to-many independent product of its marginals $\underline{P}_r$, $r \in R$;*
(iii) *If R and S constitute a partition of N, then $\underline{Q}_N$ is a many-to-many independent product of its $\mathscr{X}_R$-marginal $\underline{Q}_R$ and its $\mathscr{X}_S$-marginal $\underline{Q}_S$.*

The associativity property in (iii) follows immediately from the definition of a many-to-many independent product, and means that the joint model still satisfies many-to-many independence with respect to the marginals $\underline{Q}_R$ and $\underline{Q}_S$; we have established a similar property for strong independence in Proposition 8. To see that an analogous property does not generally hold for many-to-one products, consider the coherent lower prevision $\underline{Q}_N$ in Example 6 in Appendix B.

To conclude, we consider the case where all the lower previsions we want to combine into an independent joint model are actually linear previsions. The following result shows that our definitions of many-to-one and many-to-many independent products extend the existing ones for linear previsions. It will also provide the basis for Proposition 12, which will allow us to kick-start the discussion in Section 5.

Proposition 10. *Any linear previsions P_n on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, have a unique many-to-many independent product and a unique many-to-one independent product, and both are equal to the (strong) product $S_N := \times_{n\in N} P_n$.*

4.4. **Conditional independence.** Besides the variables X_n, $n \in N$, considered so far, we now consider another variable Y assuming values in a finite set $\mathscr{Y}$. We assume the variables X_N and Y to be logically independent: the variable (X_N, Y) can assume all values in the Cartesian product $\mathscr{X}_N \times \mathscr{Y}$.

We also consider separately coherent conditional lower previsions $\underline{P}_{O\cup I}(\cdot|X_I, Y)$ on $\mathscr{L}(\mathscr{X}_{O\cup I} \times \mathscr{Y})$ where I and O are disjoint subsets of N. *It is important to realise that in all these conditional lower previsions, the variable Y consistently appears as a conditioning variable.*

We can use this set-up to generalise the notions of epistemic irrelevance and independence to those of *conditional* epistemic irrelevance and independence. As an example, consider two disjoint subsets I and O of N. We say that a subject judges that X_I *is epistemically irrelevant to X_O, conditional on Y*, when she assumes that when knowing the value of Y, learning in addition which value X_I assumes in $\mathscr{X}_I$ will not affect her beliefs about X_O.

Assume that our subject has a separately coherent conditional lower prevision $\underline{P}_N(\cdot|Y)$ on $\mathscr{L}(\mathscr{X}_N \times \mathscr{Y})$. If she assesses that X_I is epistemically irrelevant to X_O conditional on Y, this implies that she can infer from her model $\underline{P}_N(\cdot|Y)$ a conditional model $\underline{P}_{O\cup I}(\cdot|X_I,Y)$ on the set $\mathscr{L}(\mathscr{X}_{O\cup I}\times\mathscr{Y})$ given by

$$\underline{P}_{O\cup I}(h|x_I,y) := \underline{P}_N(h(\cdot,x_I,y)|y) \text{ for all } h\in\mathscr{L}(\mathscr{X}_{O\cup I}\times\mathscr{Y}) \text{ and all } (x_I,y)\in\mathscr{X}_I\times\mathscr{Y}.$$

We can now extend the entire discussion in this and the next section to the conditional case. We will, however, refrain from doing so explicitly, or in any detail. Rather, we present a result that simply reduces problems of conditional epistemic irrelevance and independence, as formulated above, to a collection of problems of epistemic irrelevance and independence. This will allow us to immediately and automatically extend all the results in this and the next section from the case of independence to that of independence conditional on the additional variable Y.

With any conditional lower prevision $\underline{P}_{O\cup I}(\cdot|X_I,Y)$ on $\mathscr{L}(\mathscr{X}_{O\cup I}\times\mathscr{Y})$, we associate a collection of (separately coherent) conditional lower previsions $\underline{Q}^y_{O\cup I}(\cdot|X_I)$ on $\mathscr{L}(\mathscr{X}_{O\cup I})$, one for each y in $\mathscr{Y}$, defined by

$$\underline{Q}^y_{O\cup I}(f|x_I) := \underline{P}_{O\cup I}(f|x_I,y) = \underline{P}_{O\cup I}(f(\cdot,x_I)|x_I,y) \text{ for all } f\in\mathscr{L}(\mathscr{X}_{O\cup I}) \text{ and } x_I\in\mathscr{X}_I. \tag{10}$$

We can reduce the problem of checking the coherence of a collection of conditional lower previsions $\underline{P}_{O_k\cup I_k}(\cdot|X_{I_k},Y)$, $k=1,\dots,m$ to a number $|\mathscr{Y}|$ of coherence problems that are simpler, in the sense that the conditioning variable Y disappears from them; we have, for each $y\in\mathscr{Y}$, to check the coherence of the collection $\underline{Q}^y_{O_k\cup I_k}(\cdot|X_{I_k})$, $k=1,\dots,m$.

Theorem 11 (Elimination of common conditioning variables)**.** *Consider m arbitrary but different pairs of disjoint subsets O_k and I_k of N, $k=1,\dots,m$. Consider separately coherent conditional lower previsions $\underline{P}_{O_k\cup I_k}(\cdot|X_{I_k},Y)$ on $\mathscr{L}(\mathscr{X}_{O_k\cup I_k}\times\mathscr{Y})$, $k=1,\dots,m$, and for each $y\in\mathscr{Y}$, the corresponding separately coherent $\underline{Q}^y_{O_k\cup I_k}(\cdot|X_{I_k})$ on $\mathscr{L}(\mathscr{X}_{O_k\cup I_k})$, $k=1,\dots,m$. Then the following statements are equivalent:*

(i) *The collection $\underline{P}_{O_k\cup I_k}(\cdot|X_{I_k},Y)$, $k=1,\dots,m$ is coherent;*
(ii) *For each y in $\mathscr{Y}$, the collection $\underline{Q}^y_{O_k\cup I_k}(\cdot|X_{I_k})$, $k=1,\dots,m$ is coherent.*

5. Independent natural extension

A number of examples in Appendix B show that, when we leave linear for lower previsions, many-to-one and many-to-many independent products are generally speaking not unique.

What we want to do in this section, then, is to show that any collection of coherent marginals always has a pointwise smallest many-to-one, and a pointwise smallest many-to-many, independent product, and that these products coincide.

We begin by observing that there always is at least one many-to-many (and therefore also many-to-one) independent product:

Proposition 12. *Consider coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n\in N$. Then their strong product $\times_{n\in N}\underline{P}_n$ is a many-to-many and many-to-one independent product of the marginals $\underline{P}_n$.*

5.1. Definition of the many-to-one independent natural extension.

Although the notion of epistemic many-to-many independence seems to be the more intuitively appealing and useful of the two, it turns out to be easier to approach the study with many-to-one independent products. So, for the time being, we concentrate on the latter notion of independence, which is related to the collection of conditional lower previsions:

$$\mathscr{N}(\underline{P}_n, n\in N) = \left\{\underline{P}_{\{o\}\cup I}(\cdot|X_I)\colon o\in N \text{ and } I\subseteq N\setminus\{o\}\right\},$$

where the conditional lower previsions are given by Eq. (9). Proposition 12 is instrumental in establishing the following crucial observation.

Proposition 13. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Then the collection $\mathscr{N}(\underline{P}_n, n \in N)$ of conditional lower previsions $\underline{P}_{\{o\}\cup I}(\cdot|X_I)$ is coherent.*

As an immediate consequence, we see that checking whether a joint lower prevision is a many-to-one independent product is a fairly straightforward matter.

Corollary 14. *Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$, and coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Then $\underline{P}_N$ is a many-to-one independent product of the $\underline{P}_n$, $n \in N$, if and only if any (and hence all) of the following equivalent conditions is satisfied:*

(i) *$\underline{P}_N$ is weakly coherent with the collection $\mathscr{N}(\underline{P}_n, n \in N)$ of conditional lower previsions $\underline{P}_{\{o\}\cup I}(\cdot|X_I)$;*
(ii) *$\underline{P}_N(\mathbb{I}_{\{x_I\}}[g - \underline{P}_o(g)]) = 0$ for all $o \in N$, all $I \subseteq N \setminus \{o\}$, all gambles g on $\mathscr{X}_o$ and all $x_I \in \mathscr{X}_I$.*

If we apply Condition (ii) in this corollary to the special case where I is the empty set, we deduce immediately that any (many-to-one) independent product of a number of lower previsions must have these lower previsions as its marginals:

Corollary 15. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. If the coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ is a many-to-one independent product of these lower previsions $\underline{P}_n$, then for all $n \in N$, $\underline{P}_n$ is the $\mathscr{X}_n$-marginal of $\underline{P}_N$: $\underline{P}_N(g) = \underline{P}_n(g)$ for all gambles g on $\mathscr{X}_n$.*

Because all the sets $\mathscr{X}_n$ are finite, we can invoke Walley's Finite Extension Theorem [30, Theorem 8.1.9] to derive from Proposition 13 that there always is a *pointwise smallest* joint coherent lower prevision $\underline{E}_N$ that is coherent with the coherent family $\mathscr{N}(\underline{P}_n, n \in N)$. This leads to the following definition.

Definition 6 (Many-to-one independent natural extension)**.** Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. We call the pointwise smallest coherent lower prevision that is coherent with the family of conditional lower previsions $\mathscr{N}(\underline{P}_n, n \in N)$ the *many-to-one independent natural extension* of the marginals $\underline{P}_n$, and we denote it by $\otimes_{n\in N}\underline{P}_n$, or alternatively by $\underline{E}_N$ when it is clear from the context what the marginals are.

Alternatively, and equivalently, the many-to-one independent natural extension of the marginals $\underline{P}_n$ is the pointwise smallest many-to-one independent coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$ whose marginals coincide with the given $\underline{P}_n$. We gather from Corollary 14 that $\otimes_{n\in N}\underline{P}_n$ is also the smallest coherent lower prevision that is weakly coherent with the conditional lower previsions in $\mathscr{N}(\underline{P}_n, n \in N)$.

Since the strong product $\times_{n\in N}\underline{P}_n$ is a many-to-one independent product of the marginals $\underline{P}_n$, $n \in N$, by Proposition 12, it has to dominate the many-to-one independent natural extension $\otimes_{n\in N}\underline{P}_n$: $\times_{n\in N}\underline{P}_n \geq \otimes_{n\in N}\underline{P}_n$. These products do not coincide in general: Walley [30, Section 9.3.4] discusses an example where the many-to-one independent natural extension is not a lower envelope of independent linear products, and as a consequence cannot coincide with the strong product.

On the other hand, the strong product is not generally the greatest many-to-one independent product of given marginals, as we show in Example 3 in Appendix B.

5.2. **Immediate properties of the many-to-one independent natural extension.** It will pay to study this many-to-one independent natural extension in greater detail. We begin by deriving a workable expression for it. By definition, $\underline{E}_N$ is the smallest joint coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$ that is coherent with the family of conditional lower previsions $\mathscr{N}(\underline{P}_n, n \in N)$, so we can infer from Walley's Finite Extension Theorem [30, Theorem 8.1.9] that $\underline{E}_N$ is the *natural extension* of the coherent collection $\mathscr{N}(\underline{P}_n, n \in N)$ to an unconditional (joint) lower prevision. Using Eq. (5), we find that it is given by:

$$\underline{E}_N(f) = \sup\left\{\min\left[f - \sum_{o\in N, I\subseteq N\setminus\{o\}} G_{\{o\}\cup I}(g_{I,o}|X_I)\right] : g_{I,o} \in \mathscr{L}(\mathscr{X}_{\{o\}\cup I}), o \in N, I \subseteq N \setminus \{o\}\right\},$$

for all gambles f on $\mathscr{X}_N$, where we let

$$G_{\{o\}\cup I}(g_{I,o}|X_I) \coloneqq \sum_{x_I\in\mathscr{X}_I}\mathbb{I}_{\{x_I\}}[g_{I,o}-\underline{P}_{\{o\}\cup I}(g_{i,o}|X_I)] = \sum_{x_I\in\mathscr{X}_I}\mathbb{I}_{\{x_I\}}[g_{I,o}(\cdot,x_I)-\underline{P}_o(g_{I,o}(\cdot,x_I))].$$

Therefore we find that:

$$\underline{E}_N(f) = \sup_{\substack{g_{I,o}\in\mathscr{L}(\mathscr{X}_{\{o\}\cup I})\\ o\in N, I\subseteq N\setminus\{o\}}} \min_{z_N\in\mathscr{X}_N}\left[f(z_N) - \sum_{o\in N, I\subseteq N\setminus\{o\}}[g_{I,o}(z_o,z_I)-\underline{P}_o(g_{I,o}(\cdot,z_I))]\right]. \quad (11)$$

We first show that we can simplify this expression, by restricting the I in the supremum to their largest possible values.

Theorem 16. *Consider coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n\in N$. Then for all gambles f on $\mathscr{X}_N$:*

$$\underline{E}_N(f) = \sup_{\substack{h_n\in\mathscr{L}(\mathscr{X}_N)\\ n\in N}} \min_{z_N\in\mathscr{X}_N}\left[f(z_N) - \sum_{n\in N}[h_n(z_N)-\underline{P}_n(h_n(\cdot,z_{N\setminus\{n\}}))]\right]. \quad (12)$$

The coherent lower prevision in Eq. (12) is actually the natural extension of the following coherent family of conditional lower previsions:

$$\mathscr{N}_{\mathrm{ext}}(\underline{P}_n, n\in N) \coloneqq \left\{\underline{P}_{\{n\}\cup N\setminus\{n\}}(\cdot|X_{N\setminus\{n\}})\colon n\in N\right\} \subseteq \mathscr{N}(\underline{P}_n, n\in N).$$

Relying on this expression for the independent natural extension, it is fairly straightforward to show that it has the following monotonicity property.

Proposition 17. *Let $\underline{P}_n$ and $\underline{Q}_n$ be coherent lower previsions on $\mathscr{L}(\mathscr{X}_n)$ such that $\underline{P}_n\leq\underline{Q}_n$, $n\in N$. Then $\otimes_{n\in N}\underline{P}_n \leq \otimes_{n\in N}\underline{Q}_n$.*

5.3. **Marginalising and conditioning the many-to-one independent natural extension.** Let us consider the coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n\in N$. For any non-empty subset R of N, we denote the independent natural extension of the marginal lower previsions $\underline{P}_r$, $r\in R$, by $\underline{E}_R = \otimes_{r\in R}\underline{P}_r$. This $\underline{E}_R$ turns out to be the $\mathscr{X}_R$-marginal of $\underline{E}_N = \otimes_{n\in N}\underline{P}_n$.

Theorem 18. *Consider a non-empty subset R of N. Then $\underline{E}_N(f) = \underline{E}_R(f)$ for all gambles f on $\mathscr{X}_R$.*

The argumentation leading to this marginalisation property also allows us to prove the following result.

Proposition 19. *$\underline{E}_N$ is productive. Moreover, $\underline{E}_N(\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]) = 0$ for all disjoint subsets I and O of N, all $x_I\in\mathscr{X}_I$ and all $g\in\mathscr{L}(\mathscr{X}_O)$.*

This implies that the independent natural extension $\underline{E}_N$ satisfies the law of large numbers of Theorem 4. *So does, therefore, any (many-to-one or many-to-many) independent product of these marginals*, as it must dominate $\underline{E}_N$, even though, as we show in Appendix B, not all such independent products are productive!

Let us now define, for any disjoint subsets I and O of N, the conditional lower previsions $\underline{E}_{O\cup I}(\cdot|X_I)$ on the set $\mathscr{L}(\mathscr{X}_{O\cup I})$ as follows:

$$\underline{E}_{O\cup I}(h|x_I) \coloneqq \underline{E}_N(h(\cdot,x_I)) = \underline{E}_O(h(\cdot,x_I)) \text{ for all } h\in\mathscr{L}(\mathscr{X}_{O\cup I}) \text{ and } x_I\in\mathscr{X}_I, \quad (13)$$

where the last equality follows from Theorem 18. This allows us to infer from the many-to-one independent natural extension $\underline{E}_N$ a family of conditional models

$$\mathscr{I}(\underline{E}_N) \coloneqq \{\underline{E}_{O\cup I}(\cdot|X_I)\colon I \text{ and } O \text{ disjoint subsets of } N\},$$

Interestingly, $\underline{E}_N$ is coherent with the family $\mathscr{I}(\underline{E}_N)$, and therefore:

Theorem 20. *$\underline{E}_N$ is a many-to-many independent product of the coherent lower previsions $\underline{P}_n$, $n\in N$.*

Since any many-to-many independent product is in particular also a many-to-one independent product, we are led to the following immediate conclusion:

Theorem 21 (Independent natural extension)**.** *The many-to-one independent natural extension $\underline{E}_N = \otimes_{n\in N}\underline{P}_n$ of the coherent lower previsions $\underline{P}_n$, $n \in N$, is also their pointwise smallest many-to-many independent product. We can therefore simply call it the* independent natural extension *of the marginals $\underline{P}_n$.*

The independent natural extension is not only a many-to-one and many-to-many independent product of its marginals: it is also a factorising lower prevision.

Theorem 22. *Consider coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Then their independent natural extension $\otimes_{n\in N}\underline{P}_n$ is factorising.*

To show that it is also strongly factorising, it seems easiest to first show that it has an associativity property that is similar to the one discussed in Propositions 8 and 9.

5.4. **Associativity of the independent natural extension.** Assume that N is the union of two disjoint sets N_1 and N_2. Then it is natural to ask whether taking the independent natural extension is an associative operation, i.e., whether

$$\otimes_{n\in N_1\cup N_2}\underline{P}_n = (\otimes_{n\in N_1}\underline{P}_n)\otimes(\otimes_{n\in N_2}\underline{P}_n)?$$

Let us look at this formulation from a slightly different angle. We can consider the tuple X_{N_1} as a variable assuming values in the set $\mathscr{X}_{N_1}$, and $\underline{E}_{N_1} := \otimes_{n\in N_1}\underline{P}_n$ as the corresponding 'marginal' lower prevision on $\mathscr{L}(\mathscr{X}_{N_1})$. Similarly, we can consider the tuple X_{N_2} as a variable assuming values in $\mathscr{X}_{N_2}$, and $\underline{E}_{N_2} := \otimes_{n\in N_2}\underline{P}_n$ as the corresponding 'marginal' lower prevision on $\mathscr{L}(\mathscr{X}_{N_2})$. We now consider the joint variable $X_{\{N_1,N_2\}}$ assuming values in $\mathscr{X}_{\{N_1,N_2\}}$, and the independent natural extension $\underline{E}_{\{N_1,N_2\}} := \underline{E}_{N_1}\otimes\underline{E}_{N_2} = (\otimes_{n\in N_1}\underline{P}_n)\otimes(\otimes_{n\in N_2}\underline{P}_n)$ of these two 'marginals'. Since the variable $X_{N_1\cup N_2}$ is (essentially) the same as the variable $X_{\{N_1,N_2\}}$, the natural question to ask is, whether $\underline{E}_{N_1\cup N_2} = \underline{E}_{\{N_1,N_2\}}$?

Theorem 23. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Consider a partition N_1 and N_2 of N, then $\otimes_{n\in N_1\cup N_2}\underline{P}_n = (\otimes_{n\in N_1}\underline{P}_n)\otimes(\otimes_{n\in N_2}\underline{P}_n)$.*

One important and fairly immediate consequence of this associativity is that it allows us to derive from the factorising character of the independent natural extension that it is also strongly factorising:

Theorem 24. *Consider coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Then their independent natural extension $\otimes_{n\in N}\underline{P}_n$ is strongly factorising.*

5.5. **Interesting special cases.** When some of the marginals are linear or vacuous, the expression for the independent natural extension in Eq. (12) simplifies to a great extent. Because of the associativity result in Theorem 23, it suffices to consider the case of two variables X_1 and X_2, so we let $N = \{1,2\}$.

When one of the marginals is linear, all independent products coincide:

Proposition 25. *Let P_1 be any linear prevision on $\mathscr{L}(\mathscr{X}_1)$, and let $\underline{P}_2$ be any coherent lower prevision on $\mathscr{L}(\mathscr{X}_2)$. Let $\underline{P}_{\{1,2\}}$ be any independent product of P_1 and $\underline{P}_2$. Then for all gambles f on $\mathscr{X}_1\times\mathscr{X}_2$:*

$$\underline{P}_{\{1,2\}}(f) = (P_1\times\underline{P}_2)(f) = (P_1\otimes\underline{P}_2)(f) = \underline{P}_2(P_1(f)),$$

where $P_1(f)$ is the gamble on $\mathscr{X}_2$ defined by $P_1(f)(x_2) := P_1(f(\cdot,x_2))$ for all $x_2 \in \mathscr{X}_2$.

On the other hand, when one of the marginals is vacuous, then the strong product and the independent natural extension are also guaranteed to coincide:

Proposition 26. *Let $\underline{P}_1^{A_1}$ be the vacuous lower prevision on $\mathscr{L}(\mathscr{X}_1)$ relative to the non-empty set $A_1 \subseteq \mathscr{X}_1$, and let $\underline{P}_2$ be any coherent lower prevision on $\mathscr{L}(\mathscr{X}_2)$. For all gambles f on $\mathscr{X}_1\times\mathscr{X}_2$:*

(i) $(\underline{P}_1^{A_1} \times \underline{P}_2)(f) = (\underline{P}_1^{A_1} \otimes \underline{P}_2)(f) = \min_{x_1 \in A_1} \underline{P}_2(f(x_1, \cdot))$.

(ii) *If $\underline{P}$ is a factorising coherent lower prevision with these marginals, then*

$$\underline{P}(f) = (\underline{P}_1^{A_1} \times \underline{P}_2)(f) = (\underline{P}_1^{A_1} \otimes \underline{P}_2)(f) = \min_{x_1 \in A_1} \underline{P}_2(f(x_1, \cdot)).$$

(iii) *If $\underline{P}_2 = \underline{P}_2^{A_2}$ is the vacuous lower prevision on $\mathscr{L}(\mathscr{X}_2)$ relative to the non-empty set $A_2 \subseteq \mathscr{X}_2$, then any coherent lower prevision with these marginals $\underline{P}$ is an independent product of $\underline{P}_1^{A_1}$ and $\underline{P}_2^{A_2}$.*

In contrast with what we might come to expect from Proposition 25, when one of the marginals is vacuous, we cannot guarantee that all independent products coincide: see Example 3 in Appendix B. This implies that the second statement of Proposition 26 cannot be extended from factorising coherent lower previsions to independent products.[14]

6. External additivity

We can now bring what we have learnt about the independent natural extension to bear on our discussion of the more formalist approaches to independence. In the following section, we investigate the connections between epistemic independence and factorisation. Here, we discuss weakened versions for lower previsions of the additivity property that all linear previsions have. Vicig [29] discusses a related but weaker notion of external n-monotonicity for the case of coherent lower *probabilities*.

Definition 7 (External additivity)**.** Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. We call this lower prevision:

(i) *externally additive* if for all non-empty $R \subseteq N$ and all gambles f_r on $\mathscr{X}_r$, $r \in R$, $\underline{P}_N(\sum_{r \in R} f_r) = \sum_{r \in R} \underline{P}_N(f_r)$;

(ii) *strongly externally additive* if $\underline{P}_N(f+g) = \underline{P}_N(f) + \underline{P}_N(g)$ for all $f \in \mathscr{L}(\mathscr{X}_I)$ and $g \in \mathscr{L}(\mathscr{X}_O)$, where I and O are any disjoint subsets[15] of N.

Clearly, strong external additivity implies external additivity. The latter is called *summation independence* by Cozman [3], who also gives, for the case $N = \{1,2\}$, a proof for the external additivity of the strong product [3, Theorem 1]. We generalise his result by proving that both the strong product and the epistemic natural extension are generally (strongly) externally additive.

Proposition 27. *Consider arbitrary coherent lower previsions $\underline{P}_n$, $n \in N$. Then both their strong product $\underline{S}_N$ and their independent natural extension $\underline{E}_N$ are strongly externally additive, and therefore also externally additive.*

It follows from the definition that any convex combination of (strongly) externally additive coherent lower previsions is again (strongly) externally additive. In fact, looking at the proof of this result in Appendix A, we see that *any many-to-one independent product of the given marginals that is dominated by the strong product is also externally additive.* To see that this does not extend to all many-to-one independent products, consider Example 3 in Appendix B.

On the other hand, Example 4 in the same appendix shows that the properties of external additivity and strong external additivity are not equivalent. It also shows that not all many-to-one independent products between the independent natural extension and the strong product are strongly externally additive.

---

[14] It also explains why the lower prevision in Example 3 of Appendix B cannot be factorising.

[15] The coherence of $\underline{P}_N$ guarantees that for empty I or O the corresponding condition is trivially satisfied.

7. FACTORISATION AND INDEPENDENCE

Since we know from Proposition 8(iv) that the strong product is factorising, we wonder if we can use factorising lower previsions as many-to-one independent products. That this is indeed the case is proved in the following theorem:

Theorem 28. *Consider an arbitrary coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. If it is factorising, then it is a many-to-one independent product of its marginals $\underline{P}_n$, $n \in N$.*

It is therefore natural to ask whether, by extension, all strongly factorising lower previsions are many-to-many independent. While we have not been able to answer this question in its full generality, we are able to show that this is indeed the case under fairly weak additional positivity conditions.

Recall that for a factorising (or indeed any) coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ to be an independent product, it must be *strongly* coherent with the family $\mathscr{I}(\underline{P}_N)$ of conditional lower previsions. It turns out that at least the *weak* coherence is never an issue.

Theorem 29. *Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. If it is strongly factorising, then it is weakly coherent with the family $\mathscr{I}(\underline{P}_N)$.*

Next, we turn to deriving a sufficient condition for a strongly factorising lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ to be also many-to-many independent, so strongly coherent with the family $\mathscr{I}(\underline{P}_N)$. For any non-empty subset I of N, we see at once that

$$\underline{P}_N(\times_{i\in I}A_i) = \prod_{i\in I}\underline{P}_N(A_i) \text{ and } \overline{P}_N(\times_{i\in I}A_i) = \prod_{i\in I}\overline{P}_N(A_i),$$

where $A_i \subseteq \mathscr{X}_i$ for all $i \in I$. Now suppose we want to condition $\underline{P}_N$ on an observation $X_I = x_I$, where I is some subset of N. To this end, we calculate the regular extension, as discussed in Section 2.6:

$$\underline{R}(h|x_I) := \max\left\{\mu \in \mathbb{R} \colon \underline{P}_N(\mathbb{I}_{\{x_I\}}[h-\mu]) \geq 0\right\},$$

where h is any gamble on $\mathscr{X}_O$ and O is any subset of $N \setminus I$. If $\underline{P}_N$ is strongly factorising, we see that

$$\begin{aligned}\underline{P}_N(\mathbb{I}_{\{x_I\}}[h-\mu]) &= \underline{P}_N(\mathbb{I}_{\{x_I\}}\underline{P}_N(h-\mu)) = \underline{P}_N(\mathbb{I}_{\{x_I\}}(\underline{P}_N(h)-\mu))\\ &= \begin{cases}\underline{P}_N(\{x_I\})(\underline{P}_N(h)-\mu) & \text{if } \underline{P}_N(h) \geq \mu\\ \overline{P}_N(\{x_I\})(\underline{P}_N(h)-\mu) & \text{if } \underline{P}_N(h) \leq \mu,\end{cases}\end{aligned}$$

so we conclude that

$$\underline{R}(h|x_I) = \underline{P}_N(h) \text{ as soon as } \overline{P}_N(\{x_I\}) > 0.$$

This means that, heuristically speaking and under some positivity assumptions, the conditional lower previsions that are found by conditioning a strongly factorising joint lower prevision using regular extension, reflect the irrelevance conditions that are involved in the definition of many-to-many independence. Taking into account that, from Theorem 3, the conditional lower previsions derived by regular extension are strongly coherent, we deduce the following:

Theorem 30. *Let $\underline{P}_N$ be a strongly factorising coherent lower prevision. If $\overline{P}_N(\{x_I\}) > 0$ for every $\{x_I\} \in \mathscr{X}_I$, then $\underline{P}_N$ is many-to-many independent.*

It is an open problem at this point whether this positivity condition is really necessary.

Since the independent natural extension is the pointwise smallest many-to-one independent product of given marginals, and since we have shown in Proposition 8(iv) that the strong product is in particular Kuznetsov, we deduce that the smallest many-to-one independent product that is still Kuznetsov lies between the independent natural extension and the strong product. For the case $N = \{1,2\}$, this was also established by Cozman [3].

Concerning the other conditions introduced in Section 3, we point out the following:

Proposition 31. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, and let $\underline{Q}_1$ and $\underline{Q}_2$ be coherent lower previsions on $\mathscr{L}(\mathscr{X}_N)$ with these marginals $\underline{P}_n$. Let $\underline{Q}_3$ be any coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$ such that $\underline{Q}_1 \leq \underline{Q}_3 \leq \underline{Q}_2$. Then the following statements hold:*

(i) *If $\underline{Q}_1$ and $\underline{Q}_2$ are many-to-one independent products, then so is $\underline{Q}_3$;*
(ii) *If $\underline{Q}_1$ and $\underline{Q}_2$ are factorising, then so is $\underline{Q}_3$;*
(iii) *If $\underline{Q}_1$ and $\underline{Q}_2$ are Kuznetsov, then so is $\underline{Q}_3$;*
(iv) *If $\underline{Q}_2$ is externally additive, then so is $\underline{Q}_3$.*

We deduce that a convex combination of many-to-one independent products of the same given marginals is again a many-to-one independent product of these marginals. A similar result holds for factorising or Kuznetsov lower previsions.

It is also interesting to investigate the connections between independent products and the notion of productivity.

Proposition 32. *Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. Then the following statements are equivalent:*

(i) *$\underline{P}_N$ is weakly coherent with $\mathscr{I}(\underline{P}_N)$;*
(ii) *$\underline{P}_N$ is productive.*

In addition, any many-to-many independent lower prevision is productive, and any productive coherent lower prevision is many-to-one independent. Moreover, if $\underline{P}_N(\{x_I\}) > 0$ for every $x_I \in \mathscr{X}_I$ and every subset I of N, then $\underline{P}_N$ is many-to-many independent if and only if it is productive.

We conclude that productivity, which is sufficient for the law of large numbers in Theorem 4 to hold, is intermediate between being a many-to-one and being a many-to-many independent product, and is equivalent to many-to-many independence when all the conditioning events have positive lower probability. Example 6 in Appendix B shows that not all many-to-one independent lower previsions are productive, and that many-to-one and many-to-many independence are not equivalent even if the conditioning events all have positive lower probability.

8. An application: probabilistic inference in imprecise Markov trees

Independence is at the very heart of much research done in Artificial Intelligence. In this section we show how the independent natural extension affects a very traditional domain of AI: probabilistic graphical models. We do so by reviewing and discussing a model and algorithm recently introduced elsewhere by some of us [11]. The coherence results that are at the core of that algorithm rely quite heavily on the properties of the independent natural extension proved in this paper[16], such as its being the smallest many-to-many independent product, its marginalisation and associativity properties, and its being strongly factorising.

8.1. Background notions and notation.

As is widely known, a graphical model consists of a graph enriched with specific probabilistic information. One such model is a *Bayesian network* [27]. Here a directed graph represents variables through nodes, and independence statements between them by arcs. Conditional probabilities are associated with the nodes of the graph.

Let us make this more precise by introducing some notation. We tailor the notation to the case of tree topologies, which are the focus of our attention: every node of the graph has exactly one parent, with the exception of one node, called *root*, which has no parents.

We call T the set of the nodes s of the tree, and we denote the root node by $\Box$. For any node s, we denote its *mother node* by $m(s)$. The root $\Box$ has no mother node, and we use the convention $m(\Box) = \emptyset$. For each node s, we denote the set of its *children* by $C(s)$. If $C(s) = \emptyset$,

[16] The paper was written jointly with the present one, but published earlier due to circumstance.

then we call s a *leaf*, or *terminal node*. Moreover, $D(s)$ denotes the set of *descendants* of s (its successors in the graph). We also use the notation $\downarrow s := D(s) \cup \{s\}$ for the sub-tree with root s. Similarly, we let $\downarrow S := \bigcup\{\downarrow s\colon s \in S\}$ for any subset $S \subseteq T$. For any node s, its set of non-parent non-descendants is then given by $\overline{s} := T \setminus (\{m(s)\} \cup \downarrow s)$.

With each node s of the tree, there is associated a variable X_s. We adopt the usual notation and assumptions for the variables in the tree. In particular, variable X_s takes values in the corresponding non-empty finite set $\mathscr{X}_s$, and we assume all variables to be logically independent.

At this point we can define the probabilistic information in a Bayesian net more precisely: a generic node s is equipped with a mass function for X_s conditional on $X_{m(s)} = z_{m(s)}$, for all $z_{m(s)} \in \mathscr{X}_{m(s)}$. We can rephrase this in the language of previsions by saying that each node s has a (separately coherent) conditional linear prevision $Q_s(\cdot|X_{m(s)})$ on $\mathscr{L}(\mathscr{X}_s)$: for each possible value $z_{m(s)}$ of the variable $X_{m(s)}$ associated with its mother node $m(s)$, we have a linear prevision $Q_s(\cdot|z_{m(s)})$ for the value of X_s, conditional on $X_{m(s)} = z_{m(s)}$.[17] We call $Q_s(\cdot|X_{m(s)})$ a *local uncertainty model*.

The tree together with the local uncertainty models provides a compact way to define a joint probability mass function over X_T. This follows from the *Markov condition*, which is also what provides the graphical model with a probabilistic semantics: conditional on its mother variable $X_{m(s)}$, variable X_s is assumed to be independent of its non-parent non-descendant variables $X_{\overline{s}}$. Again, in terms of lower previsions, this means that the graphical model is an equivalent representation of a global uncertainty model P_T, that is, a linear prevision defined on $\mathscr{L}(\mathscr{X}_T)$.

The global uncertainty model is then used for inference, which most often amounts to computing posterior beliefs (i.e., probabilities or expectations) for a query variable X_s conditional on $X_E = x_E$, where E is a non-empty subset of T whose variables are in the known state x_E. This task is called *updating*. Updating is performed by applying Bayes's rule to the global uncertainty model while exploiting the structure of the graph in order to perform the computation as efficiently as possible. In fact, the computation of updating on a tree-shaped Bayesian network is an easy task, which is solved exactly in time linear in the size of the tree.

8.2. **Epistemic trees.** Bayesian networks have been extended to deal with imprecisely specified probabilities. The resulting models are called *credal networks* [4]. The extension is achieved primarily by replacing the local uncertainty models of Bayesian networks with imprecise ones: in the most common case, this means that each mass function required to specify a Bayesian net is replaced by a closed convex set of mass functions. In the language of previsions, in a credal network each node s is equipped with a (separately coherent) conditional *lower* prevision $\underline{Q}_s(\cdot|X_{m(s)})$ on $\mathscr{L}(\mathscr{X}_s)$: for each possible value $z_{m(s)}$ of the variable $X_{m(s)}$ associated with its mother node $m(s)$, we have a coherent *lower* prevision $\underline{Q}_s(\cdot|z_{m(s)})$ for the value of X_s, conditional on $X_{m(s)} = z_{m(s)}$.

As in Bayesian networks, these local uncertainty models need to be combined into a global uncertainty model that is later used for (imprecise-probabilistic) inference. The Markov condition still plays the leading role in this process, but has to be modified to take into account the specific notion of independence (or irrelevance) that the graph is assumed to represent.

The traditional approach in the literature focuses on strong independence: in this case, conditional on its mother variable $X_{m(s)}$, a variable X_s is assumed to be strongly independent of its non-parent non-descendant variables $X_{\overline{s}}$. The global uncertainty model $\underline{P}_T$ obtained through this so-called *strong Markov condition*, is called the *strong extension*. The strong extension comes with a sensitivity analysis interpretation: in fact, each of its extreme points

[17] In the root, this corresponds to having an unconditional local uncertainty model $Q_\Box$ for $X_\Box$: a linear prevision on $\mathscr{L}(\mathscr{X}_\Box)$.

can be regarded as arising from a Bayesian network with the same graph as the credal net, and whose local uncertainty models dominate the local uncertainty models of the credal network. In other words, a credal network under strong independence can be regarded as the set of all the Bayesian networks that are compatible with the probabilities that have been imprecisely specified in the design of the credal net.

But the sensitivity analysis interpretation of an imprecise probability model is not always applicable, as we have discussed in the Introduction with reference to modelling expert knowledge. In this case an assumption of epistemic irrelevance may allow one to represent more faithfully an expert's beliefs.[18] This point is especially important because modelling expert knowledge is among the main motivations for using credal networks (as well as Bayesian nets).

This lead has been followed in Ref. [11] by looking at the following type of Markov condition based on epistemic irrelevance:

CI. Consider any node s in the tree T, any subset S of its set of children $C(s)$, and the set $\overline{S} := \bigcap_{c\in S}\overline{c}$ of their common non-parent non-descendants. Then *conditional on the variable X_s, the non-parent non-descendant variables $X_{\overline{S}}$ are assumed to be epistemically irrelevant to the variables $X_{\downarrow S}$ associated with the children in S and their descendants.*

This condition turns T into a *credal tree under epistemic irrelevance*, which we call *imprecise Markov tree*.

Before we proceed, let us briefly address a technical question: the form of CI might look unusual when compared to the common statement of the Markov condition. In fact, CI seems to impose a wider set of irrelevances, focusing as it does on sets of children of s, and on the related subtrees, rather than on a single child. But this difference is only apparent, because the strong Markov condition, as well as the precise-probabilistic Markov condition, can be reformulated *equivalently* in a completely similar way: the apparent additional irrelevances (or independencies) are actually implied by those in the standard Markov condition [use d-separation to see that s blocks all the relevant paths]. Whether or not this is the case when we use epistemic irrelevance is not obvious to us at present; this is due to epistemic irrelevance being a relatively weak notion for imprecise probability models. For this reason, CI is formulated by imposing all the additional epistemic irrelevances explicitly.

We now shed more light on two immediate consequences of CI.

First, consider some non-terminal node s different from $\square$, and its mother variable $X_{m(s)}$. We infer from CI that this mother variable $X_{m(s)}$ is epistemically irrelevant to the variable $X_{\downarrow C(s)}$ conditional on X_s:

$$\begin{tikzcd} & X_{m(s)} \arrow[d] & \\ & X_s \arrow[dl] \arrow[d] \arrow[dr] & \\ X_{\downarrow c_1} & \dots & X_{\downarrow c_n} \end{tikzcd} \qquad \text{or equivalently,} \qquad \begin{tikzcd} X_{m(s)} \arrow[d] \\ X_s \arrow[d] \\ X_{\downarrow C(s)} \end{tikzcd}$$

It is worth stressing that such is not necessarily the case when we reason in the opposite way: CI does *not* imply that $X_{\downarrow C(s)}$ is epistemically irrelevant to $X_{m(s)}$ in case we observe X_s. This kind of symmetrised irrelevance (that is, independence) can be imposed too, but its treatment is much more involved and problematical from the algorithmic side, because it complicates the construction of the global model tremendously. In addition, we would argue that the irrelevances imposed by CI are more natural than their symmetrised counterparts for *directed* graphical models. We will come back to this point in Section 8.4.

---

[18] Obviously, there may be cases where strong independence is justified in order to model an expert's knowledge. Moreover, strong independence could provide a good approximation to more accurate models, even when it is not entirely appropriate.

Next, consider some node s. Then CI tells us that for any two children $c_1, c_2 \in C(s)$ of s, the variable $X_{\downarrow c_1}$ is epistemically irrelevant to the variable $X_{\downarrow c_2}$, conditional on X_s.

$$\begin{tikzcd} & X_s \arrow[dl] \arrow[d] \arrow[dr] & \\ X_{\downarrow c_1} & \dots & X_{\downarrow c_2} \end{tikzcd}$$

It even tells us that for any two disjoint non-empty sets $S_1 \subseteq C(s)$ and $S_2 \subseteq C(s)$ of children of s, the variable $X_{\downarrow S_1}$ is epistemically irrelevant to $X_{\downarrow S_2}$, conditional on X_s. In contradistinction with the previous example, here we see that the symmetrised irrelevance conditions (here between children) originate spontaneously from CI: we conclude that, conditional on a node, its children c (or rather, the variables associated with their sub-trees $\downarrow c$) are *epistemically many-to-many independent*.

This is the specific point where our present work on the independent natural extension meets credal trees. If we want to obtain the most conservative global model that arises through coherence from the local models and the statements of conditional irrelevance implied by CI, then we need to compute the independent natural extension in order to summarise the information carried by the variables associated with the sub-trees $\downarrow c$, with $c \in C(s)$.

8.3. **Constructing the most conservative global model.** Let us illustrate how to construct the most conservative global model for the variables in the tree that extends the local models and expresses all conditional irrelevancies encoded by the imprecise Markov tree through CI. Consider the following fragment of the tree.

$$\begin{tikzcd} & & X_{m(s)} \arrow[d] & & & \\ & & X_s \arrow[dll] \arrow[dl] \arrow[d] \arrow[dr] \arrow[rr, dotted] & & \underline{Q}_s(\cdot|X_{m(s)}) & \\ X_{\downarrow c_1} & X_{\downarrow c_2} & \dots & X_{\downarrow c_n} \arrow[r, dotted] & \underline{P}_{\downarrow c_k}(\cdot|X_s) & \end{tikzcd}$$

Here we denote by $\underline{P}_{\downarrow c_k}(\cdot|X_s)$ the lower prevision on $\mathscr{L}(\mathscr{X}_{\downarrow c_k})$ that is the most conservative global model for $X_{\downarrow c_k}$ constructed from CI and the local models *in the subtree with root* c_k. This is a *conditional* global model as it depends on the value of X_s. For the time being, we assume that such conditional global models related to the children of s exist and have already been computed. Since we know from the foregoing discussion that the $X_{\downarrow c_1}, \dots,$ $X_{\downarrow c_n}$ are many-to-many independent conditional on X_s, we can compute their independent natural extension $\otimes_{c \in C(s)} \underline{P}_{\downarrow c}(\cdot|X_s)$, which is a conditional lower prevision $\underline{P}_{\downarrow C(s)}(\cdot|X_s)$ on $\mathscr{L}(\mathscr{X}_{\downarrow C(s)})$. We can reorganise the graph accordingly by clustering all the children into a single node.

$$\begin{tikzcd} X_{m(s)} \arrow[d] & \\ X_s \arrow[d] \arrow[r, dotted] & \underline{Q}_s(\cdot|X_{m(s)}) \\ X_{\downarrow C(s)} \arrow[r, dotted] & \otimes_{c \in C(s)} \underline{P}_{\downarrow c}(\cdot|X_s) =: \underline{P}_{\downarrow C(s)}(\cdot|X_s) \end{tikzcd}$$

At this point, the local model $\underline{Q}_s(\cdot|X_{m(s)})$ must be combined with $\underline{P}_{\downarrow C(s)}(\cdot|X_s)$ into a least-committal (pointwise smallest) global conditional model about $X_{\downarrow s}$. This is achieved by taking their *marginal extension*[19]

$$\underline{Q}_s(\underline{P}_{\downarrow C(s)}(\cdot|X_{\{m(s),s\}})|X_{m(s)}) = \underline{Q}_s(\underline{P}_{\downarrow C(s)}(\cdot|X_s)|X_{m(s)});$$

[19]Marginal extension is, in the special case of precise probability models, also known as the law of total probability, or the law or iterated expectations.

see Refs. [23] and [30, Section 6.7.2] for more details. Graphically:

$$\begin{array}{l} X_{m(s)} \\ \downarrow \\ X_{\downarrow s} \dashrightarrow \underline{Q}_s(\underline{P}_{\downarrow C(s)}(\cdot|X_s)|X_{m(s)}) =: \underline{P}_{\downarrow s}(\cdot|X_{m(s)}) \end{array}$$

It is clear that this process can be iterated by starting from the leaves of the tree and letting $\underline{P}_{\downarrow t}(\cdot|X_{m(t)}) := \underline{Q}_t(\cdot|X_{m(t)})$ for all leaves t, and working our way recursively up to the root. When we eventually get to the root, this process yields a global model $\underline{P}_{\downarrow\Box}(\cdot|X_{m(\Box)}) =: \underline{P}_T$.

A central theorem in Ref. [11] then guarantees that, under mild positivity conditions on the local upper previsions, $\underline{P}_T$ is the most conservative (pointwise smallest) joint lower prevision that coherently extends the local models and that expresses all the conditional irrelevance statements implied by CI. This remarkable result relies quite heavily on the independent natural extension and its properties, as introduced and studied in this paper. The particular properties that are crucial in establishing it are: that it is the smallest many-to-many independent product (Theorem 21), that it is strongly factorising (Theorem 24), that it satisfies some marginalisation and associativity properties (Theorems 18 and 23), and that all of this can be extended to the conditional setup (Theorem 11).

These results have important implications for imprecise-probabilistic graphical models under epistemic irrelevance. They show that epistemic irrelevance can be used in the context of at least some graphical models in much the same way as stochastic independence or strong independence. In fact, not only does the work in Ref. [11] show by construction that the 'epistemic extension' $\underline{P}_T$ generally exists, but it goes further by using it to derive an efficient algorithm for updating in imprecise Markov trees—another name for credal trees. Similarly to the traditional algorithms for Bayesian nets, it works in a distributed fashion by passing (imprecise-)probabilistic messages along the nodes of the tree until it converges and thus yields the (exact) updated lower previsions of interest. Moreover, it works in time linear in the size of the tree, as is the case with the more traditional algorithms for precise probabilities.

8.4. **Some remarks.** The results obtained with imprecise Markov trees are particularly interesting because it had been uncertain until quite recently whether or not epistemic irrelevance could be used to design efficient algorithms in graphical models. This is related to epistemic irrelevance being not as well-behaved as other independence notions with respect to the graphoid axioms; see in particular Ref. [6], but also Ref. [26] for a more positive view. Part of the interest in these results is due to complexity reasons. Computation in credal nets based on strong independence is substantially harder than the case of Bayesian nets: it is an NP-hard task even on polytrees[20] [9]. What complexity updating credal *trees* under strong independence has, is still an open problem, but preliminary analyses indicate that this task could be NP-hard too [8]. If this were to be confirmed, we should have a clear example where epistemic irrelevance leads to simpler models of computation than strong independence.

On the other hand, we should take into account that these positive results have been obtained in the case of tree topologies. The expressivity of trees should not be overlooked because, for example, updating problems in Bayesian networks can be solved through the well-known join tree structure (this is an undirected tree, but it can be as well represented in a directed way with minor changes).[21] And yet, it does not seem likely that something similar can be done under epistemic irrelevance. The crucial point here is that when we convert a polytree into a (directed) join tree, we shall be obliged to invert the direction of some arcs. This is easy to see by considering a V-shaped graph made of two chains that

[20] These are directed graphs that become (undirected) trees after dropping the orientations of the arcs.

[21] This was observed by Pearl already at the time of the original proposal by Lauritzen and Spiegelhalter, see Ref. [20, p. 211].

converge into a node. The problem is that the epistemic irrelevances coded by the original V-shaped polytree will not imply in general those related to the inverted arcs in the directed tree, simply because epistemic irrelevance is an asymmetric notion. Therefore, imprecise Markov trees do not seem suited to be exploited as tools for solving updating problems in more general credal networks under epistemic irrelevance. This situation could perhaps change if condition CI were reformulated to code symmetric irrelevances (that is, epistemic independencies): in this case inverting the arcs should cause no problems. However, as we already suggested above, the development of efficient algorithms for trees under such a modified (strengthened) Markov condition appears to be quite problematic, because the condition complicates the computation of the joint model. All these considerations lead us to conjecture that the extension of the existing results from trees to polytrees might require a substantial increase in computational complexity.

Still, there is another direction, relying on tree topologies, that appears to be particularly promising for addressing interesting problems under epistemic irrelevance. It is based on the observation that discrete-time sequential processes are very often represented by *hidden Markov models* [28], which are special trees in the language of graphical models. Hidden Markov models have a number of applications, often related to some kind of recognition: speech, gesture, or word recognition. Technically, this is achieved by computing sequences of hidden variables (states) from the observation of sequences of other, manifest, variables (outputs). In fact, in order to turn credal trees into workable imprecise hidden Markov models, the additional complication that needs to be faced is that of querying the tree for the *joint* value of the hidden variables, rather than for a single one. Recent work [7] has shown that, fortunately, such a task can be solved again exactly and with a complexity that is essentially the same as that required for precise-probabilistic hidden Markov models. Once again, the key to this result is the use of the independent natural extension and its properties, as developed here.

9. CONCLUSIONS

We have worked out the foundations of epistemic independence, a generalisation of stochastic independence to imprecise probabilities. This has led to a definition of independent products, and in particular to the especially interesting and useful notion of *independent natural extension*. Like the strong product, or any other independent product, for that matter, it captures the idea of mutual irrelevance between sources of information. But it is the most conservative independent product to do so, which indicates that it is the only one that is based *solely* on this mutual irrelevance (and coherence, of course).[22] We see that, because it encompasses all these types of independent product, the notion of epistemic independence has very wide scope.

We have carried out our study by focusing on variables assuming values in finite spaces, and have followed two different routes. On the one hand, we have considered generalisations to imprecise probabilities of the factorisation formula: productivity, (strong) factorisation, being (strongly) Kuznetsov. On the other, we have defined many-to-one and many-to-many independent products of marginals, or, possibly, conditionals, based on a behavioural notion of symmetrised epistemic irrelevance, or in other words, epistemic independence. We have shown that the two routes are tightly interwoven, as factorisation implies many-to-one independence and strong factorisation implies many-to-many independence (under weak positivity assumptions). The most important notion of this paper, the independent natural extension, has been shown to be the smallest many-to-many (and many-to-one) independent product. It also satisfies useful basic properties related to marginalisation, associativity, and external additivity.

[22] The 'extra information' entering the strong product seems to be the underlying assumption of an ideal precise model. In order to make the strong product the smallest coherent independent product, the notion of coherence would have to be strengthened.

What the independent natural extension embodies, in other words, is a way to coherently extend marginal imprecise probability models into a joint model using symmetrised epistemic irrelevance judgements alone. As far as practical applications are concerned, this task is simplified by our next important result: under mild conditions, using the independent natural extension is equivalent to imposing strong factorisation when looking for least-committal models. In fact, this is the crucial result that has allowed the independent natural extension to be used successfully in graphical models in Ref. [11], as we have discussed above.

All these results appear here, at this level of generality, for the first time. It is natural to wonder what they might eventually lead us to.

As far as applications are concerned, we see much scope for epistemic independence. As already indicated, the domain of graphical models is particularly worth of consideration in this respect. Future research could, as mentioned in Section 8.4, try to extend the work in Ref. [11] to more general graphs, such as polytrees. The recent development of an efficient algorithm for the exact computation of state sequences in imprecise hidden Markov models [7] (which are special imprecise Markov trees) should favour the emergence of new interesting applications exploiting epistemic independence. Another potential domain of application could be probabilistic logic, and in particular the extensions that have been proposed to embed independence [5]. This would lead to approaches to probabilistic logic allowing both for imprecise probabilities and epistemic independence judgements.

As regards more technical questions arising from this paper, we summarise the main problems that remain open at this stage.

Primo, in Proposition 6, we have established relationships—in fact, implications—between the different factorisation conditions introduced in Section 3. In Appendix B, we give a number of counterexamples that show that none of the converse implications hold in general, except for one of them: we still do not know whether the Kuznetsov and strong Kuznetsov conditions are generally equivalent.

Secundo, in Section 4 we have introduced many-to-one and many-to-many independent products, and we have shown in Appendix B that the second of these notions is more restrictive. But let us look at the connections between these epistemic notions and the more formal conditions introduced in Section 3. We have proved that a factorising coherent lower prevision (and as a consequence also one that is strongly factorising, Kuznetsov or strongly Kuznetsov) is many-to-one independent, although in Appendix B we can see that the converse is not true in general. We also show in Appendix B that not every many-to-many independent product is factorising (and therefore it need not be Kuznetsov, strongly factorising or strongly Kuznetsov). But we do not know whether a strongly factorising coherent lower prevision is necessarily many-to-many independent.

Tertio, we have established in Section 5 that the independent natural extension is the smallest many-to-one independent product of given marginals, and that it also is the smallest product that is factorising, strongly factorising, or many-to-many independent. We show in Appendix B that it is neither the smallest Kuznetsov nor the smallest strongly Kuznetsov product. Another example of this is given in Ref. [3], where it is shown that the least-committal Kuznetsov product of given marginals may be different from the strong product.

Quarto, and related to this, we show in Example 3 that there are many-to-many independent products that dominate the strong product, but our next conjecture, that the strong product could be the greatest factorising product of given marginals, still remains to be proved or disproved. If the conjecture were to hold, then the strong product would also be the greatest strongly factorising, Kuznetsov and strongly Kuznetsov product. Related to this, it might be argued that the coherent lower prevision in Example 3 represents a somewhat pathological situation, and that it points to the fact that considering only many-to-many independence as our main requirement could be too weak. One possibility would be to restrict ourselves to many-to-many independent products which satisfy one of the factorisation conditions we have introduced in this paper; this seems to be in accordance with the results in Section 5.5, which show that the coherent lower prevision in Example 3 is not factorising, and that for those marginals the only factorising product is the intuitive one. Hence, we may focus for instance on factorising coherent lower previsions, or on many-to-many independent products which are at the same time strongly factorising.

Finally, we do not know whether all factorising coherent lower previsions are externally additive, nor whether all strongly factorising coherent lower previsions are. These open problems could be related to our conjecture about the strong product being the greatest factorising lower prevision with given marginals.

Acknowledgements. This work has been supported by SBO project 060043 of the IWT-Vlaanderen, projects TIN2008-06796-C04-01, MTM2010-17844, by Swiss NSF grants n. 200020_134759 / 1, 200020-121785 / 1, and by the Hasler foundation grant n. 10030. We would also like to thank the reviewers for their helpful comments.

REFERENCES

[1] Giulanella Coletti and Romano Scozzafava. *Probabilistic logic in a coherent setting*. Kluwer, 2002.
[2] Inés Couso, Serafín Moral, and Peter Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5:165–181, 2000.
[3] Fabio G. Cozman. Constructing sets of probability measures through Kuznetsov's independence condition. In G. de Cooman, T. L. Fine, and T. Seidenfeld, editors, *ISIPTA '01 – Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, pages 104–111. Shaker Publishing, Maastricht, 2000.
[4] Fabio G. Cozman. Credal networks. *Artificial Intelligence*, 120:199–233, 2000.
[5] Fabio G. Cozman, Cassio P. de Campos, and José C. Ferreira da Rocha. Probabilistic logic with independence. *International Journal of Approximate Reasoning*, 49(1):3–17, 2008.
[6] Fabio G. Cozman and Peter Walley. Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence*, 45(1–2):173–195, 2005.
[7] Jasper De Bock and Gert de Cooman. State sequence prediction in imprecise hidden Markov models. 2011. Submitted for publication.
[8] Cassio P. de Campos. Personal communication, 2010.
[9] Cassio P. de Campos and Fabio G. Cozman. The inferential complexity of Bayesian and credal networks. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1313–1318, Edinburgh, 2005.
[10] Cassio P. de Campos and Fabio G. Cozman. Computing lower and upper expectations under epistemic independence. *International Journal of Approximate Reasoning*, 44(3):244–260, 2007.
[11] Gert de Cooman, Filip Hermans, Alessandro Antonucci, and Marco Zaffalon. Epistemic irrelevance in credal nets: the case of imprecise Markov trees. *International Journal of Approximate Reasoning*, 51(9):1029–1052, 2010.
[12] Gert de Cooman and Enrique Miranda. Weak and strong laws of large numbers for coherent lower previsions. *Journal of Statistical Planning and Inference*, 138(8):2409–2432, 2008.
[13] Gert de Cooman and Enrique Miranda. Forward irrelevance. *Journal of Statistical Planning and Inference*, 139(2):256–276, 2009.
[14] Gert de Cooman, Enrique Miranda, and Marco Zaffalon. Factorisation properties of the strong product. In *Combining Soft Computing and Statistical Methods in Data Analysis, Proceedings of the 5th International Conference on Soft Methods in Probability and Statistics, SMPS 2010, Oviedo and Mieres (Asturias), Spain, September 28 – October 1, 2010*, volume 77 of *Advances in Soft Computing*, pages 139–147, 2010.
[15] Gert de Cooman, Enrique Miranda, and Marco Zaffalon. Independent natural extension. In *Computational Intelligence for Knowledge-Based Systems Design, 13th International Conference on Information Processing and Management of Uncertainty, IPMU 2010, Dortmund, Germany, June 28 – July 2, 2010. Proceedings*, volume 6178 of *Lecture Notes in Computer Science*, pages 737–746. Springer, 2010.
[16] Bruno de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.
[17] Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester, 1974–1975. English translation of [16], two volumes.
[18] Richard Holmes. *Geometric Functional Analysis and Its Applications*. Springer-Verlag, New York, 1975.
[19] Vladimir Kuznetsov. *Interval Statistical Methods*. Radio i Svyaz Publ., 1991. (In Russian). Downloadable at `http://www.sipta.org/index.php?id=res#kuz`.
[20] Steffen L. Lauritzen and David J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems (with discussion). *Journal of the Royal Statistical Society, Series B*, 50:157–224, 1988.

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

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

[23] Enrique Miranda and Gert de Cooman. Marginal extension in the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 46(1):188–225, 2007.

[24] Enrique Miranda and Marco Zaffalon. Coherence graphs. *Artificial Intelligence*, 173:104–144, 2009.

[25] Enrique Miranda and Marco Zaffalon. Natural extension as a limit of regular extensions. *Journal of Statistical Planning and Inference*, 140(7):1805–1833, 2010.

[26] Serafín Moral. Epistemic irrelevance on sets of desirable gambles. *Annals of Mathematics and Artificial Intelligence*, 45:197–214, 2005.

[27] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA, 1988.

[28] Lawrence R. Rabiner. A tutorial on HMM and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, February 1989.

[29] Paolo Vicig. Epistemic independence for imprecise probabilities. *International Journal of Approximate Reasoning*, 24(3):235–250, 2000.

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

[31] Peter Walley, Renato Pelessoni, and Paolo Vicig. Direct algorithms for checking consistency and making inferences from conditional probability assessments. *Journal of Statistical Planning and Inference*, 126:119–151, 2004.

[32] Peter M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Revised journal version: [33].

[33] Peter M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of [32].

APPENDIX A. PROOFS OF RESULTS

Proof of Proposition 5. We first show that $\underline{P}_N$ is factorising if and only if (i) holds. Since the direct implication is trivial, it suffices to prove the converse. Consider $o \in N$, $f_o \in \mathscr{L}(\mathscr{X}_o)$ and non-negative $f_i \in \mathscr{L}(\mathscr{X}_i)$ for $i \in I$, where I is any subset of N that does not include o. Let f_j be the gamble with the constant value 1, for every $j \notin I \cup \{o\}$. We deduce from (i) that

$$\underline{P}_N(f_o \prod_{i \in I} f_i) = \underline{P}_N(f_o \prod_{i \neq o} f_i) = \underline{P}_N(\underline{P}_N(f_o) \prod_{i \neq o} f_i) = \underline{P}_N(\underline{P}_N(f_o) \prod_{i \in I} f_i),$$

so $\underline{P}_N$ is factorising.

Next, we prove that (i) and (ii) are equivalent. Let, for notational simplicity, $f_R := \prod_{r \in R} f_r$ for any subset R of N.

(i)$\Rightarrow$(ii). If $\underline{P}_N(f_o) \geq 0$, then the coherence of $\underline{P}_N$ tells us that $\underline{P}_N(\underline{P}_N(f_o) f_{N \setminus \{o\}}) = \underline{P}_N(f_o)\underline{P}_N(f_{N \setminus \{o\}})$. Similarly, if $\underline{P}_N(f_o) \leq 0$, then

$$\underline{P}_N(\underline{P}_N(f_o) f_{N \setminus \{o\}}) = -\overline{P}_N(-\underline{P}_N(f_o) f_{N \setminus \{o\}}) = \underline{P}_N(f_o)\overline{P}_N(f_{N \setminus \{o\}}).$$

It now suffices to establish the equalities $\underline{P}_N(f_{N \setminus \{o\}}) = \prod_{i \in N \setminus \{o\}} \underline{P}_N(f_i)$ and $\overline{P}_N(f_{N \setminus \{o\}}) = \prod_{i \in N \setminus \{o\}} \overline{P}_N(f_i)$. These follow easily by applying induction on the number of the elements in the product.

(ii)$\Rightarrow$(i). By letting $f_o := 1$, we infer from (ii) that $\underline{P}_N(f_{N \setminus \{o\}}) = \prod_{i \in N \setminus \{o\}} \underline{P}_N(f_i)$, and by letting $f_o := -1$, that $\overline{P}_N(f_{N \setminus \{o\}}) = \prod_{i \in N \setminus \{o\}} \overline{P}_N(f_i)$. Going back to general f_o, we now derive from (ii) and the coherence of $\underline{P}_N$ that, when $\underline{P}_N(f_o) \geq 0$:

$$\underline{P}_N(f_o) \prod_{i \in N \setminus \{o\}} \underline{P}_N(f_i) = \underline{P}_N(f_o)\underline{P}_N(f_{N \setminus \{o\}}) = \underline{P}_N(\underline{P}_N(f_o) f_{N \setminus \{o\}}),$$

and that when $\underline{P}_N(f_o) \leq 0$:

$$\underline{P}_N(f_o) \prod_{i \in N \setminus \{o\}} \overline{P}_N(f_i) = \underline{P}_N(f_o)\overline{P}_N(f_{N \setminus \{o\}}) = -\overline{P}_N(-\underline{P}_N(f_o) f_{N \setminus \{o\}}) = \underline{P}_N(\underline{P}_N(f_o) f_{N \setminus \{o\}}). \quad \square$$

Proof of Proposition 6. First, assume that $\underline{P}_N$ is strongly Kuznetsov. We show that $\underline{P}_N$ is strongly factorising. Consider disjoint subsets I and O of N, a gamble g on $\mathscr{X}_O$ and a non-negative gamble f on $\mathscr{X}_I$. Since $\overline{P}_N(f) \geq \underline{P}_N(f) \geq 0$ and $\overline{P}_N(g) \geq \underline{P}_N(g)$, we infer that

$$\begin{aligned}\underline{\overline{P}}_N(fg) &= \underline{\overline{P}}_N(f) \boxtimes \underline{\overline{P}}_N(g)\\ &= [\min\{\underline{P}_N(f)\underline{P}_N(g), \overline{P}_N(f)\underline{P}_N(g)\}, \max\{\underline{P}_N(f)\overline{P}_N(g), \overline{P}_N(f)\overline{P}_N(g)\}]\\ &= \begin{cases}[\underline{P}_N(f)\underline{P}_N(g), \overline{P}_N(f)\overline{P}_N(g)] & \text{if } \underline{P}_N(g) \geq 0\\ [\overline{P}_N(f)\underline{P}_N(g), \overline{P}_N(f)\overline{P}_N(g)] & \text{if } \underline{P}_N(g) \leq 0 \leq \overline{P}_N(g)\\ [\overline{P}_N(f)\underline{P}_N(g), \underline{P}_N(f)\overline{P}_N(g)] & \text{if } \overline{P}_N(g) \leq 0,\end{cases}\end{aligned}$$

and by considering the lower interval bounds, we see that $\underline{P}_N$ is indeed strongly factorising.

Next, assume that $\underline{P}_N$ is strongly factorising. We show that $\underline{P}_N$ is productive. Consider disjoint subsets I and O of N, a gamble g on $\mathscr{X}_O$ and a non-negative gamble f on $\mathscr{X}_I$. Then it follows from the fact that $\underline{P}_N$ is strongly factorising and coherent that

$$\underline{P}_N(f[g-\underline{P}_N(g)]) = \underline{P}_N(f\underline{P}_N(g-\underline{P}_N(g))) = \underline{P}_N(f[\underline{P}_N(g)-\underline{P}_N(g)]) = 0,$$

so $\underline{P}_N$ is indeed productive.

The proofs of the remaining implications are either similar, or trivial. □

Proof of Proposition 7. The first part of the proposition is immediate. For the second, use the first to see that (c) and (d) are equivalent, and that (e) and (f) are equivalent. For the other equivalences, use the self-conjugacy and coherence of P_N. □

Proof of Proposition 8. The first statement is an immediate consequence of Eq. (8) and Proposition 7(i).

We turn to the proof of the second statement. We first show that the set $\operatorname{ext}(\mathscr{M}(\underline{S}_N))$ is included in $\{\times_{n\in N}P_n \colon (\forall n \in N)P_n \in \operatorname{ext}(\mathscr{M}(\underline{P}_n))\}$. The sets $\mathscr{M}(\underline{P}_n)$, $n \in N$, are closed [in the topology of pointwise convergence, or equivalently, in the Euclidean topology, because we are working in finite-dimensional linear spaces]. So the Cartesian product $\times_{n\in N}\mathscr{M}(\underline{P}_n)$ is closed in the product topology. Since it is clear from Eq. (6) that taking a product of linear previsions is a continuous operation with respect to these topologies, it follows that the set of linear previsions $\mathscr{M} := \{\times_{n\in N}P_n \colon (\forall n \in N)P_n \in \mathscr{M}(\underline{P}_n)\}$ is closed. It follows from Walley's weak* compactness theorem [30, Theorem 3.6.1] that since $\underline{S}_N$ is the lower envelope of $\mathscr{M}$, the convex compact set $\mathscr{M}(\underline{S}_N)$ is equal to the closed convex hull of its subset $\mathscr{M}$. It then follows from the extended form of the Krein–Milman Theorem in Ref. [18, p. 74] that $\operatorname{ext}(\mathscr{M}(\underline{S}_N)) \subseteq \mathscr{M}$ (because $\mathscr{M}$ is closed). Suppose *ex absurdo* that some extreme point $S_N = \times_{n\in N}Q_n$ of $\mathscr{M}(\underline{S}_N)$ does not belong to $\{\times_{n\in N}P_n \colon (\forall n \in N)P_n \in \operatorname{ext}(\mathscr{M}(\underline{P}_n))\}$. Then there must be some $r \in N$ such that Q_r is not an extreme point of $\mathscr{M}(\underline{P}_r)$, so there are different Q_r^1 and Q_r^2 in $\mathscr{M}(\underline{P}_r)$ and $\alpha \in (0,1)$ such that $P_r = \alpha Q_r^1 + (1-\alpha)Q_r^2$. But then $S_N = \alpha Q_N^1 + (1-\alpha)Q_N^2$, where $Q_N^1 := (\times_{n\neq r}P_n) \times Q_r^1 \in \mathscr{M}$ and $Q_N^2 := (\times_{n\neq r}P_n) \times Q_r^2 \in \mathscr{M}$. Since $\mathscr{M} \subseteq \mathscr{M}(\underline{S}_N)$, this contradicts that S_N is an extreme point of $\mathscr{M}(\underline{S}_N)$. We deduce that indeed $\operatorname{ext}(\mathscr{M}(\underline{S}_N)) \subseteq \{\times_{n\in N}P_n \colon (\forall n \in N)P_n \in \operatorname{ext}(\mathscr{M}(\underline{P}_n))\}$.

On to the converse inequality $\{\times_{n\in N}P_n \colon (\forall n \in N)P_n \in \operatorname{ext}(\mathscr{M}(\underline{P}_n))\} \subseteq \operatorname{ext}(\mathscr{M}(\underline{S}_N))$. Consider arbitrary $P_n \in \operatorname{ext}(\mathscr{M}(\underline{P}_n))$ for all $n \in N$. Then $S_N := \times_{n\in N}P_n \in \mathscr{M}(\underline{S}_N)$ by Eq. (8). It follows from Minkowski's theorem (or Krein–Milman Theorem in finite dimensions) that S_N is a convex combination of elements of $\operatorname{ext}(\mathscr{M}(\underline{S}_N))$: there are $m \geq 1$, non-negative real $\alpha_1, \ldots, \alpha_m$ such that $\sum_{k=1}^m \alpha_k = 1$ and $Q_N^1, \ldots, Q_N^m$ in $\operatorname{ext}(\mathscr{M}(\underline{S}_N))$ such that $S_N = \sum_{k=1}^m \alpha_k Q_N^k$. If $m = 1$ then clearly $S_N \in \operatorname{ext}(\mathscr{M}(\underline{S}_N))$, so we may assume without loss of generality that $m > 1$.

Consider any $n \in N$, then it follows by marginalisation that $P_n = \sum_{k=1}^m \alpha_k Q_n^k$, where Q_n^k is the $\mathscr{X}_n$-marginal of Q_N^k, $k = 1, \ldots, m$. [That the $\mathscr{X}_n$-marginal of S_N is P_n follows from Proposition 7(i).] Since $Q_N^k \in \mathscr{M}(\underline{S}_N)$ we find that $Q_N^k(g) \geq \underline{S}_N(g) = \underline{P}_n(g)$ for any $g \in \mathscr{L}(\mathscr{X}_n)$, where the equality follows from (i). This implies that $Q_n^k \in \mathscr{M}(\underline{P}_n)$, $k =$

$1,\dots,m$. But since we assumed that $P_n \in \operatorname{ext}(\mathscr{M}(\underline{P}_n))$ and $m > 1$, we must have that $Q_n^1 = \dots = Q_n^m = P_n$.

Since this holds for all $n \in N$, and since we already know from the argumentation above [$\operatorname{ext}(\mathscr{M}(\underline{S}_N)) \subseteq \mathscr{M}$] that $Q_N^k = \times_{n\in N} Q_n^k$ for $k = 1,\dots,m$, we see that $Q_N^1 = \dots = Q_N^m = S_N$ and therefore $S_N \in \operatorname{ext}(\mathscr{M}(\underline{S}_N))$.

The third statement is an immediate consequence of (ii) and Proposition 7(ii).

Let us finally prove the fourth statement. Consider arbitrary disjoint proper subsets I and O of N, gambles $f \in \mathscr{L}(\mathscr{X}_I)$ and $g \in \mathscr{L}(\mathscr{X}_O)$. It follows from the third statement, Eqs. (7)&(8), Proposition 7 and conjugacy that

$$\underline{S}_N(fg) = \min\{P_I(f)P_O(g)\colon P_I \in \mathscr{M}(\underline{S}_I) \text{ and } P_O \in \mathscr{M}(\underline{S}_O)\}$$
$$\overline{S}_N(fg) = \max\{P_I(f)P_O(g)\colon P_I \in \mathscr{M}(\underline{S}_I) \text{ and } P_O \in \mathscr{M}(\underline{S}_O)\}.$$

This clearly implies that

$$\underline{S}_N(fg) = \min\left\{ab\colon a \in \overline{\underline{S}}_I(f) \text{ and } b \in \overline{\underline{S}}_O(g)\right\}$$
$$\overline{S}_N(fg) = \max\left\{ab\colon a \in \overline{\underline{S}}_I(f) \text{ and } b \in \overline{\underline{S}}_O(g)\right\},$$

or in other words $\overline{\underline{S}}_N(fg) = \overline{\underline{S}}_I(f) \boxtimes \overline{\underline{S}}_O(g)$. Now use the first statement and conjugacy to find that $\overline{\underline{S}}_I(f) = \overline{\underline{S}}_N(f)$ and $\overline{\underline{S}}_O(g) = \overline{\underline{S}}_N(g)$.

The rest of the proof now follows from Proposition 6. $\square$

Proof of Proposition 9. The proof of (i) is an easy consequence of the coherence condition: if $\underline{P}_N$ is coherent with the family of conditional lower previsions $\mathscr{N}(\underline{P}_n, n \in N)$, then its restriction to $\mathscr{L}(\mathscr{X}_R)$ is coherent with the subfamily $\mathscr{N}(\underline{P}_r, r \in R)$ containing restrictions of certain elements of $\mathscr{N}(\underline{P}_n, n \in N)$.

The argumentation for (ii) is similar.

For (iii), we have to prove that $\underline{Q}_N$, $\underline{Q}_N(\cdot|X_R)$ and $\underline{Q}_N(\cdot|X_S)$ are coherent, where for instance

$$\underline{Q}_N(f|x_R) := \underline{Q}_N(f(\cdot,x_R)) \text{ for all } x_R \in \mathscr{X}_R \text{ and } f \in \mathscr{L}(\mathscr{X}_N).$$

Since both $\underline{Q}_N(\cdot|X_R)$ and $\underline{Q}_N(\cdot|X_S)$ belong to $\mathscr{I}(\underline{Q}_N)$, this follows from the coherence of $\underline{Q}_N$ with $\mathscr{I}(\underline{Q}_N)$. $\square$

Proof of Proposition 10. We first show that S_N is a many-to-many independent product of its marginals. It will then automatically follow that S_N is a many-to-one independent product of its marginals as well. This will establish that arbitrary linear previsions P_n on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, always have many-to-many and many-to-one independent products.

To show that S_N is a many-to-many independent product of its marginals, we use Theorem 2. This means that we need to prove that (a) the family of conditional lower previsions $\mathscr{I}(S_N)$ is coherent; and (b) that the joint model S_N is weakly coherent with the family $\mathscr{I}(S_N)$.

We begin with (a). Because we are dealing with linear previsions, coherence is equivalent to the condition in Equation (2). Assume *ex absurdo* that there are $f_{O,I} \in \mathscr{L}(\mathscr{X}_{O\cup I})$ for all disjoint subsets I and O of N, and $\delta > 0$ such that

$$\sum_{O,I} G_{O\cup I}(f_{O,I}|X_I) \leq -\delta \mathbb{I}_A \tag{14}$$

and $A := \bigcup_{O,I} \operatorname{supp}_I(f_{O,I}) \neq \emptyset$.

There are two possibilities. The first is that $P_n(\{x_n\}) > 0$ for all $x_n \in \mathscr{X}_n$ and all $n \in N$. Since S_N is linear and strongly factorising, it follows that

$$S_N(G_{O\cup I}(f_{O,I}|X_I)) = \sum_{x_I \in \mathscr{X}_I} S_N\left(\mathbb{I}_{\{x_I\}}[f_{O,I}(\cdot,x_I) - S_N(f_{O,I}(\cdot,x_I))]\right) = \sum_{x_I \in \mathscr{X}_I} S_N(\{x_I\}) \cdot 0 = 0.$$

Then, if we apply S_N to both sides of the inequality in Eq. (14), we get that $S_N(A) = 0$, which contradicts the assumption that $P_n(\{x_n\}) > 0$ for all $x_n \in \mathscr{X}_n$ and all $n \in N$.

The second possibility is that there are $n \in N$ and $x_n \in \mathscr{X}_n$ such that $P_n(\{x_n\}) = 0$. Consider any $\varepsilon \in (0,1)$. For all $n \in N$, let $A_n := \{x_n \in \mathscr{X}_n \colon P_n(\{x_n\}) = 0\}$, and let P_n^ε be the linear prevision on $\mathscr{L}(\mathscr{X}_n)$ defined by

$$P_n^\varepsilon(\{x_n\}) := \begin{cases} \varepsilon/|A_n| & \text{if } x_n \in A_n \\ (1-\varepsilon)P_n(\{x_n\}) & \text{otherwise} \end{cases}$$

when A_n is non-empty, and $P_n^\varepsilon := P_n$ when A_n is empty. Let S_N^ε be the product of the linear previsions P_n^ε, $n \in N$. Then for every I, O and $x_I \in \mathscr{X}_I$, it holds that, with obvious notations:

$$\begin{aligned} |G_{O\cup I}(f_{O,I}|x_I) - G^\varepsilon_{O\cup I}(f_{O,I}|x_I)| &= \mathbb{I}_{\{x_I\}}\,|S_N(f_{I,O}(\cdot,x_I)) - S_N^\varepsilon(f_{O,I}(\cdot,x_I))| \\ &\leq \mathbb{I}_{\{x_I\}}\varepsilon\,|\mathscr{X}_N|\max|f_{O,I}|, \end{aligned}$$

whence

$$\begin{aligned} |G_{O\cup I}(f_{O,I}|X_I) - G^\varepsilon_{O\cup I}(f_{O,I}|X_I)| &\leq \sum_{x_I\in\operatorname{supp}_I(f_{O,I})} |G_{O\cup I}(f_{O,I}|x_I) - G^\varepsilon_{O\cup I}(f_{O,I}|x_I)| \\ &\leq \mathbb{I}_A\varepsilon\,|\mathscr{X}_N|\max|f_{O,I}|, \end{aligned}$$

recalling that $A = \bigcup_{O,I}\operatorname{supp}_I(f_{O,I})$. By summing over all disjoint subsets O,I of N, we obtain that

$$\left|\sum_{O,I} G_{O\cup I}(f_{O,I}|X_I) - G^\varepsilon_{O\cup I}(f_{O,I}|X_I)\right| \leq \mathbb{I}_A\varepsilon\,|\mathscr{X}_N|\sum_{O,I}\max|f_{O,I}| =: \mathbb{I}_A\varepsilon K.$$

If we let $0 < \varepsilon < \min\{\delta/2K, 1\}$, then we infer that

$$\sum_{O,I} G^\varepsilon_{O\cup I}(f_{O,I}|X_I) \leq -\frac{\delta}{2}\mathbb{I}_A.$$

As before, applying S_N^ε to both sides of this inequality leads to $S_N^\varepsilon(A) = 0$, a contradiction.

Next, we turn to (b). By Theorem 1, we must establish the coherence of S_N with each conditional linear prevision $S_{O\cup I}(\cdot|X_I)$ (taken separately) for each pair of disjoint subsets I and O of N, or equivalently, that $S_N(\mathbb{I}_{\{x_I\}}[f - S_{O\cup I}(f|x_I)]) = 0$ for all $f \in \mathscr{L}(\mathscr{X}_O)$ and all $x_I \in \mathscr{X}_I$. We see that indeed:

$$S_N(\mathbb{I}_{\{x_I\}}[f - S_O(f|x_I)]) = S_N(\mathbb{I}_{\{x_I\}}[f - S_N(f)]) = S_N(\{x_I\})S_N(f - S_N(f)) = 0,$$

where the first equality follows from the definition of the conditional linear prevision $S_{O\cup I}(\cdot|X_I)$, and the second one because S_N is linear and strongly factorising.

To complete the proof, we show that S_N is the only joint coherent lower prevision that is a many-to-one independent product of the marginals P_n, $n \in N$. Since any many-to-many independent product of the marginals P_n, $n \in N$, is in particular also a many-to-one independent product of these marginals, it will then also follow that S_N is the only joint coherent lower prevision that is a many-to-many independent product of these marginals.

Consider any *linear* prevision P_N that is a many-to-one independent product of the marginals P_n, $n \in N$. Fix $o \in N$ and $I \subseteq N \setminus \{o\}$, $g \in \mathscr{L}(\mathscr{X}_o)$ and non-negative $f_i \in \mathscr{L}(\mathscr{X}_i)$, $i \in I$. Let $f_I := \prod_{i\in I} f_i$. By assumption P_N and $P_{\{o\}\cup I}(\cdot|X_I)$ are coherent, and therefore we infer from (GBR) that $P_N(\mathbb{I}_{\{x_I\}}[g - P_N(g)]) = 0$ for every $x_I \in \mathscr{X}_I$. Hence also $P_N(f_I[g - P_N(g)]) = \sum_{x_i\in\mathscr{X}_I} f_I(x_I)P_N(\mathbb{I}_{\{x_I\}}[g - P_N(g)]) = 0$, where the first equality is due to the linearity of P_N. It follows that P_N is factorising, and applying Proposition 7, we deduce that it coincides with the product S_N.

Finally, assume that the *lower* prevision $\underline{P}_N$ is a many-to-one independent product of the marginals P_n, $n \in N$. So $\underline{P}_N$ is coherent with the family of conditional linear previsions $\mathscr{N}(P_n, n \in N)$, so it must be a lower envelope of linear previsions P_N coherent with $\mathscr{N}(P_n, n \in N)$ by the lower envelope theorem [30, Theorem 8.1.10], given that the spaces $\mathscr{X}_n$, $n \in N$ are finite. Since we have seen above that there is a unique linear prevision S_N coherent with $\mathscr{N}(P_n, n \in N)$, we deduce that $\underline{P}_N = S_N$. □

Proof of Theorem 11. (i)⇒(ii). Fix y in $\mathscr{Y}$. We show that the conditional lower previsions $\underline{Q}^y_{O_k\cup I_k}(\cdot|X_{I_k})$, $k=1,\dots,m$ are coherent. Consider arbitrary gambles f_k on $\mathscr{X}_{O_k\cup I_k}$, $k=1,\dots,m$, and arbitrary $j\in\{1,\dots,m\}$, $x_{I_j}\in\mathscr{X}_{I_j}$ and $f'\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$. Let $g_k:=\mathbb{I}_{\{y\}}f_k\in\mathscr{L}(\mathscr{X}_{O_k\cup I_k}\times\mathscr{Y})$ and $g':=\mathbb{I}_{\{y\}}f'\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j}\times\mathscr{Y})$. Then it follows from Eq. (10) that, with obvious notations, for all z_N in $\mathscr{X}_N$:

$$\begin{aligned}G^y_{O_k\cup I_k}(f_k|X_{I_k})(z_N)&=f_k(z_{O_k},z_{I_k})-\underline{Q}^y_{O_k\cup I_k}(f_k(\cdot,z_{I_k})|z_{I_k})\\&=g_k(z_{O_k},z_{I_k},y)-\underline{P}_{O_k\cup I_k}(g_k(\cdot,z_{I_k},y)|z_{I_k},y)=G_{O_k\cup I_k}(g_k|X_{I_k},Y)(z_N,y),\end{aligned}$$

and similarly,

$$\begin{aligned}G^y_{O_j\cup I_j}(f'|x_{I_j})(z_N)&=\mathbb{I}_{\{x_{I_j}\}}(z_{I_j})[f'(z_{O_j},x_{I_j})-\underline{Q}^y_{O_j\cup I_j}(f'(\cdot,x_{I_j})|x_{I_j})]\\&=\mathbb{I}_{\{x_{I_j}\}}(z_{I_j})\mathbb{I}_{\{y\}}(y)[g'(z_{O_j},x_{I_j},y)-\underline{P}_{O_j\cup I_j}(g'(\cdot,x_{I_j},y)|x_{I_j},y)]\\&=G_{O_j\cup I_j}(g'|x_{I_j},y)(z_N,y),\end{aligned}$$

and therefore

$$\begin{aligned}&\left[\sum_{k=1}^m G^y_{O_k\cup I_k}(f_k|X_{I_k})-G^y_{O_j\cup I_j}(f'|x_{I_j})\right](z_N)\\&\qquad=\left[\sum_{k=1}^m G_{O_k\cup I_k}(g_k|X_{I_k},Y)-G_{O_j\cup I_j}(g'|x_{I_j},y)\right](z_N,y).\end{aligned}$$

Moreover, it follows, again with obvious notations, that

$$\operatorname{supp}_{I_k}(g_k):=\left\{(x_{I_k},u)\in\mathscr{X}_{I_k}\times\mathscr{Y}:\,g_k(\cdot,x_{I_k},u)\neq 0\right\}=\operatorname{supp}_{I_k}(f_k)\times\{y\}$$

and therefore

$$\{(x_{I_j},y)\}\cup\bigcup_{k=1}^m\operatorname{supp}_{I_k}(g_k)=\left(\{x_{I_j}\}\cup\bigcup_{k=1}^m\operatorname{supp}_{I_k}(f_k)\right)\times\{y\}.$$

Using this equality and (i), we find that there is some $z_N\in\{x_{I_j}\}\cup\bigcup_{k=1}^m\operatorname{supp}_{I_k}(f_k)$ such that $[\sum_{k=1}^m G_{O_k\cup I_k}(g_k|X_{I_k},Y)-G_{O_j\cup I_j}(g'|x_{I_j},y)](z_N,y)\geq 0$, and therefore $[\sum_{k=1}^m G^y_{O_k\cup I_k}(f_k|X_{I_k})-G^y_{O_j\cup I_j}(f'|x_{I_j})](z_N)\geq 0$. This implies that the collection $\underline{Q}^y_{O_k\cup I_k}(\cdot|X_{I_k})$, $k=1,\dots,m$ is coherent. Since y is arbitrary, (ii) holds.

(ii)⇒(i). Consider arbitrary $g_k\in\mathscr{L}(\mathscr{X}_{O_k\cup I_k}\times\mathscr{Y})$, $k=1,\dots,m$, as well as arbitrary $j\in\{1,\dots,m\}$, $(x_{I_j},y)\in\mathscr{X}_{I_j}\times\mathscr{Y}$ and $g'\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j}\times\mathscr{Y})$. We have to prove that there is some (z_N,u) in $\{(x_{I_j},y)\}\cup\bigcup_{k=1}^m\operatorname{supp}_{I_k}(g_k)$ such that

$$\left[\sum_{k=1}^m G_{O_k\cup I_k}(g_k|X_{I_k},Y)-G_{O_j\cup I_j}(g'|x_{I_j},y)\right](z_N,u)\geq 0.$$

Let $f_k:=g_k(\cdot,y)\in\mathscr{L}(\mathscr{X}_{O_k\cup I_k})$ and $f':=g'(\cdot,y)\in\mathscr{L}(\mathscr{X}_{O_j\cup I_j})$. Then it holds for all $z_N\in\mathscr{X}_N$ that

$$\begin{aligned}G^y_{O_k\cup I_k}(f_k|X_{I_k})(z_N)&=f_k(z_{O_k},z_{I_k})-\underline{Q}^y_{O_k\cup I_k}(f_k(\cdot,z_{I_k})|z_{I_k})\\&=g_k(z_{O_k},z_{I_k},y)-\underline{P}_{O_k\cup I_k}(g_k(\cdot,z_{I_k},y)|z_{I_k},y)=G_{O_k\cup I_k}(g_k|X_{I_k},Y)(z_N,y),\end{aligned}$$

and similarly,

$$\begin{aligned}G^y_{O_j\cup I_j}(f'|x_{I_j})(z_N)&=\mathbb{I}_{\{x_{I_j}\}}(z_{I_j})[f'(z_{O_j},x_{I_j})-\underline{Q}^y_{O_j\cup I_j}(f'(\cdot,x_{I_j})|x_{I_j})]\\&=\mathbb{I}_{\{x_{I_j}\}}(z_{I_j})\mathbb{I}_{\{y\}}(y)[g'(z_{O_j},x_{I_j},y)-\underline{P}_{O_j\cup I_j}(g'(\cdot,x_{I_j},y)|x_{I_j},y)]\\&=G_{O_j\cup I_j}(g'|x_{I_j},y)(z_N,y),\end{aligned}$$

and therefore

$$\left[\sum_{k=1}^{m} G^{y}_{O_k\cup I_k}(f_k|X_{I_k}) - G^{y}_{O_j\cup I_j}(f'|x_{I_j})\right](z_N) = \left[\sum_{k=1}^{m} G_{O_k\cup I_k}(g_k|X_{I_k},Y) - G_{O_j\cup I_j}(g'|x_{I_j},y)\right](z_N,y).$$

Since the collection $\underline{Q}^{y}_{O_k\cup I_k}(\cdot|X_{I_k})$, $k=1,\dots,m$ is coherent, we infer that there is some $z_N \in \{x_{I_j}\}\cup\bigcup_{k=1}^{m}\operatorname{supp}_{I_k}(f_k)$ such that $[\sum_{k=1}^{m} G_{O_k\cup I_k}(g_k|X_{I_k},Y) - G_{O_j\cup I_j}(g'|x_{I_j},y)](z_N,y)\geq 0$. It therefore suffices to prove that

$$(z_N,y)\in\{(x_{I_j},y)\}\cup\bigcup_{k=1}^{m}\operatorname{supp}_{I_k}(g_k).$$

This certainly holds if $z_{I_j}=x_{I_j}$. If not, then there must be some $k\in\{1,\dots,m\}$ such that $z_N\in\operatorname{supp}_{I_k}(f_k)$, meaning that $0\neq f_k(\cdot,z_{I_k})=g_k(\cdot,z_{I_k},y)$, so indeed $(z_N,y)\in\operatorname{supp}_{I_k}(g_k)$. $\square$

Proof of Proposition 12. It follows from its definition that the strong product is a lower envelope of product linear previsions. By Proposition 10, it is therefore a lower envelope of many-to-many independent (respectively many-to-one independent) products. Since both of these two properties are preserved by taking lower envelopes, $\times_{n\in N}\underline{P}_n$ is also a many-to-many and many-to-one independent product. $\square$

Proof of Proposition 13. We infer from Proposition 12 that the strong product $\times_{n\in N}\underline{P}_n$ is coherent with the collection $\mathscr{N}(\underline{P}_n,n\in N)$. This implies in particular that $\mathscr{N}(\underline{P}_n,n\in N)$ is itself coherent. $\square$

Proof of Corollary 14. By Proposition 13, the collection $\mathscr{N}(\underline{P}_n,n\in N)$ is coherent. Theorem 2 then tells us that $\underline{P}_N$ is coherent with $\mathscr{N}(\underline{P}_n,n\in N)$ if and only if it is weakly coherent with $\mathscr{N}(\underline{P}_n,n\in N)$. Taking into account Theorem 1, this holds if and only if for every $o\in N$ and $I\subseteq N\setminus\{o\}$,

$$\underline{P}_N(\mathbb{I}_{\{x_I\}}[f-\underline{P}_{\{o\}\cup I}(f|x_I)])=0 \text{ for all } f\in\mathscr{L}(\mathscr{X}_{\{o\}\cup I}) \text{ and all } x_I\in\mathscr{X}_I.$$

Now use Eq. (9) to find that

$$\mathbb{I}_{\{x_I\}}[f-\underline{P}_{\{o\}\cup I}(f|x_I)]=\mathbb{I}_{\{x_I\}}[f(\cdot,x_I)-\underline{P}_o(f(\cdot,x_I))]. \qquad \square$$

Proof of Theorem 16. Denote the right-hand side in Eq. (12) by $\underline{Q}_N(f)$. It follows easily from Eq. (11) that $\underline{E}_N(f)\geq\underline{Q}_N(f)$, so we concentrate on the converse inequality. Consider any real $\alpha<\underline{E}_N(f)$, then there are gambles $g_{I,o}$ in $\mathscr{L}(\mathscr{X}_{\{o\}\cup I})$ for all $o\in N$ and $I\subseteq N\setminus\{o\}$ such that

$$\min_{z_N\in\mathscr{X}_N}\left[f(z_N)-\sum_{o\in N,I\subseteq N\setminus\{o\}}[g_{I,o}(z_o,z_I)-\underline{P}_o(g_{I,o}(\cdot,z_I))]\right]\geq\alpha. \tag{15}$$

For every n in N, define the gamble h_n on $\mathscr{X}_N$ by $h_n:=\sum_{I\subseteq N\setminus\{n\}}g_{I,n}$. Then for all $z_N\in\mathscr{X}_N$, $h_n(z_N)=\sum_{I\subseteq N\setminus\{n\}}g_{I,n}(z_n,z_I)$ and

$$\underline{P}_n(h_n(\cdot,z_{N\setminus\{n\}}))=\underline{P}_n\left(\sum_{I\subseteq N\setminus\{n\}}g_{I,n}(\cdot,z_I)\right)\geq\sum_{I\subseteq N\setminus\{n\}}\underline{P}_n(g_{I,n}(\cdot,z_I)),$$

where the inequality follows from the coherence of $\underline{P}_n$. We then infer from Eq. (15) and the definition of $\underline{Q}_N(f)$ that

$$\underline{Q}_N(f)\geq\min_{z_N\in\mathscr{X}_N}\left[f(z_N)-\sum_{n\in N}[h_n(z_N)-\underline{P}_n(h_n(\cdot,z_{N\setminus\{n\}}))]\right]\geq\alpha.$$

Since this inequality holds for all real $\alpha<\underline{E}_N(f)$, we see that indeed $\underline{Q}_N(f)\geq\underline{E}_N(f)$. $\square$

Next, we turn to the proof of Theorem 18. In order to do this, it will be very helpful to work with sets of so-called strictly desirable gambles [30]. For every $n \in N$, consider the following subset of $\mathscr{L}(\mathscr{X}_n)$:

$$\mathscr{A}_n := \{f \in \mathscr{L}(\mathscr{X}_n) \colon f > 0 \text{ or } \underline{P}_n(f) > 0\};$$

we use these sets to define the following subsets of $\mathscr{L}(\mathscr{X}_R)$, where R is any non-empty subset of N:

$$\mathscr{A}_r^R := \left\{f \in \mathscr{L}(\mathscr{X}_R)_{\neq 0} \colon (\forall x_{R\setminus\{r\}} \in \mathscr{X}_{R\setminus\{r\}}) f(\cdot, x_{R\setminus\{r\}}) \in \mathscr{A}_r \cup \{0\}\right\}, \quad r \in R.$$

We also define, for any subset S of N:

$$\mathscr{E}_S := \operatorname{posi}\left(\mathscr{L}(\mathscr{X}_S)_{>0} \cup \bigcup_{s\in S} \mathscr{A}_s^S\right). \tag{16}$$

For $S = \emptyset$, this leads to $\mathscr{E}_\emptyset = \mathscr{L}(\mathscr{X}_\emptyset)_{>0}$, which we have identified with the set of positive real numbers.

We begin by proving a crucial property of all these sets $\mathscr{A}_n$, $\mathscr{A}_n^N$ and $\mathscr{E}_N$ in Lemma 35. In order to prove this result, and a few more involved ones further on, we need the following lemmas, one of which is a convenient version of the separating hyperplane theorem:

Lemma 33. *Consider a finite subset $\mathscr{A}$ of $\mathscr{L}(\mathscr{X})$. Then $0 \notin \operatorname{posi}(\mathscr{L}(\mathscr{X})_{>0} \cup \mathscr{A})$ if and only if there is some linear prevision P on $\mathscr{L}(\mathscr{X})$ with mass function p such that $P(f) = \sum_{x\in\mathscr{X}} p(x)f(x) > 0$ for all $f \in \mathscr{A}$ and $p(x) > 0$ for all $x \in \mathscr{X}$.*

Proof of Lemma 33. It clearly suffices to prove necessity. Since $0 \notin \operatorname{posi}(\mathscr{L}(\mathscr{X})_{>0} \cup \mathscr{A})$, we infer applying a version of the separating hyperplane theorem in finite-dimensional spaces[23] that there is a linear functional Λ on $\mathscr{L}(\mathscr{X})$ such that

$$(\forall x \in \mathscr{X})\Lambda(\mathbb{I}_{\{x\}}) > 0 \text{ and } (\forall f \in \mathscr{A})\Lambda(f) > 0.$$

Then $\Lambda(\mathscr{X}) = \sum_{x\in\mathscr{X})} \Lambda(\mathbb{I}_{\{x\}}) > 0$, and if we let $P := \Lambda/\Lambda(\mathscr{X})$ then P is clearly a linear prevision on $\mathscr{L}(\mathscr{X})$ for which $P(f) > 0$ for all $f \in \mathscr{A}$. Moreover, for any $x \in \mathscr{X}$, $p(x) = P(\mathbb{I}_{\{x\}}) = \Lambda(\mathbb{I}_{\{x\}})/\Lambda(\mathscr{X}) > 0$. □

Lemma 34. *Consider a convex cone $\mathscr{A}$ of gambles on $\mathscr{X}$ such that $\max f > 0$ for all $f \in \mathscr{A}$. Consider any non-zero gamble g on $\mathscr{X}$. If $g \notin \mathscr{A}$ then $0 \notin \operatorname{posi}(\mathscr{A} \cup \{-g\})$.*

Proof. Consider a non-zero gamble $g \notin \mathscr{A}$, and assume *ex absurdo* that $0 \in \operatorname{posi}(\mathscr{A} \cup \{-g\})$. Then it follows from the assumptions that there are $m > 0$, $\lambda_k > 0$, $f_k \in \mathscr{A}$, $k = 1, \dots, m$ and $\mu > 0$ such that $0 = \sum_{k=1}^m \lambda_k f_k + \mu(-g)$. Hence $g \in \operatorname{posi}(\mathscr{A}) = \mathscr{A}$, a contradiction. □

Lemma 35. *Let S be any subset of N, and let R be any non-empty subset of N. Consider any $n \in N$ and $r \in R$.*

(i) *$\mathscr{A}_n$ is a convex cone such that $\mathscr{L}(\mathscr{X}_n)_{>0} \subseteq \mathscr{A}_n$ and $\mathscr{L}(\mathscr{X}_n)_{\leq 0} \cap \mathscr{A}_n = \emptyset$:*
(ii) *$\mathscr{A}_r^R$ is a convex cone such that $\mathscr{L}(\mathscr{X}_R)_{>0} \subseteq \mathscr{A}_r^R$ and $\mathscr{L}(\mathscr{X}_R)_{\leq 0} \cap \mathscr{A}_r^R = \emptyset$;*
(iii) *$\mathscr{E}_S$ is a convex cone such that $\mathscr{L}(\mathscr{X}_S)_{>0} \subseteq \mathscr{E}_S$ and $\mathscr{L}(\mathscr{X}_S)_{\leq 0} \cap \mathscr{E}_S = \emptyset$.*

Proof. (i). Immediate: use the coherence of the lower prevision $\underline{P}_n$.

(ii). Consider any $f \in \mathscr{L}(\mathscr{X}_R)_{>0}$. Then for all $z_{R\setminus\{r\}} \in \mathscr{X}_{R\setminus\{r\}}$, $f(\cdot, z_{R\setminus\{r\}}) \geq 0$ and therefore $f(\cdot, z_{R\setminus\{r\}}) \in \mathscr{A}_n \cup \{0\}$. Since moreover $f \neq 0$, it follows that $f \in \mathscr{A}_r^R$. This shows that $\mathscr{L}(\mathscr{X}_R)_{>0} \subseteq \mathscr{A}_r^R$.

To prove that $\mathscr{A}_r^R$ is a convex cone, it clearly suffices to show that $0 \notin \operatorname{posi}(\mathscr{A}_r^R)$. Assume *ex absurdo* that there are $m > 0$, $\lambda_k > 0$ and $f_k \in \mathscr{A}_r^R$ such that $0 = \sum_{k=1}^m \lambda_k f_k$. Fix any $z_{R\setminus\{r\}} \in \mathscr{X}_{R\setminus\{r\}}$. Then $0 = \sum_{k=1}^m \lambda_k f_k(\cdot, z_{N\setminus\{n\}})$, and it follows from (i) that this can only

[23]We use the version in Appendix E.1 of Ref. [30], for the special choice $\mathscr{V} := \mathscr{A} \cup \{\mathbb{I}_{\{x\}} : x \in \mathscr{X}\}$ and $\mathscr{W} := \{0\}$.

happen if all $f_k(\cdot, z_{R\setminus\{r\}}) = 0$. But since this holds for all $z_{R\setminus\{r\}} \in \mathscr{X}_{R\setminus\{r\}}$, we infer that all $f_k = 0$, a contradiction. Hence $\mathscr{A}_n^N$ is indeed a convex cone.

Finally, consider $f \in \mathscr{L}(\mathscr{X}_R)_{\leq 0}$ and assume that $f \in \mathscr{A}_r^R$. It already follows from the reasoning above that $f < 0$, and therefore also $-f \in \mathscr{A}_r^R$. Therefore $0 = f + (-f) \in \mathscr{A}_r^R$ since $\mathscr{A}_r^R$ is a convex cone [closed under addition], a contradiction. Hence indeed $\mathscr{L}(\mathscr{X}_R)_{\leq 0} \cap \mathscr{A}_r^R = \emptyset$.

(iii). It clearly suffices to prove that $\mathscr{L}(\mathscr{X}_S)_{\leq 0} \cap \mathscr{E}_S = \emptyset$. This is trivially so if $S = \emptyset$, so let us assume that S is non-empty. Let $f \in \mathscr{L}(\mathscr{X}_S)_{\leq 0}$ and assume *ex absurdo* that $f \in \mathscr{E}_S$. Then there are $\lambda_s \geq 0$, $\mu \geq 0$, $f_s \in \mathscr{A}_s^S$ and $g > 0$ such that $f = \mu g + \sum_{s\in S} \lambda_s f_s$ and $\max\{\mu, \max_{s\in S} \lambda_s\} > 0$.

Fix any $s \in S$. Let $\mathscr{A} := \left\{f_s(\cdot, z_{S\setminus\{s\}}) \colon z_{S\setminus\{s\}} \in \mathscr{X}_{S\setminus\{s\}}, f_s(\cdot, z_{S\setminus\{s\}}) \neq 0\right\}$, then it follows from the assumptions that $\mathscr{A}$ is a finite non-empty subset of $\mathscr{A}_s$, and therefore it follows from (i) and Lemma 33 that there is a linear prevision P_s on $\mathscr{L}(\mathscr{X}_s)$ with mass function p_s such that

$$\begin{cases} (\forall x_s \in \mathscr{X}_s) p_s(x_s) > 0 \\ (\forall z_{S\setminus\{s\}} \in \mathscr{X}_{S\setminus\{s\}})(f_s(\cdot, z_{S\setminus\{s\}}) \neq 0 \Rightarrow P_s(f_s(\cdot, z_{S\setminus\{s\}})) > 0). \end{cases}$$

We conclude that if we define the gamble g_s on $\mathscr{X}_{S\setminus\{s\}}$ by $g_s(z_{S\setminus\{s\}}) := P_s(f_s(\cdot, z_{S\setminus\{s\}}))$ for all $z_{S\setminus\{s\}}$ in $\mathscr{X}_{S\setminus\{s\}}$, then $g_s > 0$.

Since we can do this for all $s \in S$, we can define a mass function p_S on $\mathscr{X}_S$ by letting $p_S(z_S) := \prod_{s\in S} p_s(z_s) > 0$ for all $z_S \in \mathscr{X}_S$. The corresponding linear prevision P_S is of course the product linear prevision $\times_{s\in S} P_s$ of the marginal linear previsions P_s. But then it follows from the reasoning and assumptions above that

$$P_S(f) = \mu P_S(g) + \sum_{s\in S} \lambda_s P_S(f_s) = \mu P_S(g) + \sum_{s\in S} \lambda_s P_S(g_s) > 0,$$

because $P_S(g) > 0$ and all $P_S(g_s) > 0$, whereas $f \leq 0$ leads us to conclude that $P_S(f) \leq 0$, a contradiction. □

It turns out that there is a very close relationship between the set of gambles $\mathscr{E}_N$ and the many-to-one independent natural extension $\underline{E}_N$:

Lemma 36. *For all $f \in \mathscr{L}(\mathscr{X}_N)$, $\underline{E}_N(f) = \sup\{\alpha \colon f - \alpha \in \mathscr{E}_N\}$.*

Proof. Let, for the sake of notational simplicity, $\underline{Q}_N(f) := \sup\{\alpha \colon f - \alpha \in \mathscr{E}_N\}$, then we have to prove that $\underline{E}_N(f) = \underline{Q}_N(f)$.

Consider any real α such that $f - \alpha \in \mathscr{E}_N$. Then there are non-negative μ and λ_n, $g > 0$ and $f_n \in \mathscr{A}_n^N$ such that $f - \alpha = \mu g + \sum_{n\in N} \lambda_n f_n$ with $\max\{\mu, \max_{n\in N} \lambda_n\} > 0$. We infer from $f_n \in \mathscr{A}_n^N$ that $f_n(\cdot, z_{N\setminus\{n\}}) \in \mathscr{A}_n \cup \{0\}$ for all $z_{N\setminus\{n\}} \in \mathscr{X}_{N\setminus\{n\}}$, so we conclude by considering $g_n := \lambda_n f_n \in \mathscr{L}(\mathscr{X}_N)$ that $\lambda_n f_n(\cdot, z_{N\setminus\{n\}}) \geq g_n(\cdot, z_{N\setminus\{n\}}) - \underline{P}_n(g_n(\cdot, z_{N\setminus\{n\}}))$ for all $z_{N\setminus\{n\}} \in \mathscr{X}_{N\setminus\{n\}}$, and therefore

$$\lambda_n f_n(z_N) \geq g_n(z_N) - \underline{P}_n(g_n(\cdot, z_{N\setminus\{n\}})) \text{ for all } z_N \in \mathscr{X}_N \text{ and all } n \in N.$$

So we find that

$$f(z_N) - \alpha \geq \sum_{n\in N} [g_n(z_N) - \underline{P}_n(g_n(\cdot, z_{N\setminus\{n\}}))] \text{ for all } z_N \in \mathscr{X}_N,$$

and therefore $\alpha \leq \underline{E}_N(f)$. Hence $\underline{Q}_N(f) \leq \underline{E}_N(f)$.

Conversely, let $\alpha < \underline{E}_N(f)$, then there are $\varepsilon > 0$ and $g_n \in \mathscr{L}(\mathscr{X}_N)$, $n \in N$, such that

$$\alpha + \varepsilon\,|N| \leq f(z_N) - \sum_{n\in N} [g_n(z_N) - \underline{P}_n(g_n(\cdot, z_{N\setminus\{n\}}))] \text{ for all } z_N \in \mathscr{X}_N,$$

or in other words $f - \alpha \geq \sum_{n\in N} h_n$, where we let $h_n(z_N) := g_n(z_N) - \underline{P}_n(g_n(\cdot, z_{N\setminus\{n\}})) + \varepsilon$ for all $z_N \in \mathscr{X}_N$. This implies that $h_n(\cdot, z_{N\setminus\{n\}}) \in \mathscr{A}_n$ for all $z_{N\setminus\{n\}} \in \mathscr{X}_{N\setminus\{n\}}$, whence $h_n \in \mathscr{A}_n^N$, and therefore $f - \alpha \in \mathscr{E}_N$. We infer that $\alpha \leq \underline{Q}_N(f)$ and therefore also that $\underline{E}_N(f) \leq \underline{Q}_N(f)$. □

We now show that for any subset R of N, $\mathscr{E}_R$ is the set of those gambles in $\mathscr{E}_N$ that depend at most on the variables X_R, and not on $X_{N\setminus R}$.

Lemma 37. *For every subset R of N, $\mathscr{E}_R = \mathscr{E}_N \cap \mathscr{L}(\mathscr{X}_R)$.*

Proof. The result is trivial for $R = \emptyset$ and $R = N$. Let us therefore assume that both R and $N \setminus R$ are proper subsets of N. Recall from Section 2 that we are interpreting gambles on $\mathscr{X}_R$ as special gambles on $\mathscr{X}_N$. Keeping this in mind, it is obvious that $\mathscr{A}_r^R \subseteq \mathscr{A}_r^N$ for all $r \in R$, and therefore $\mathscr{E}_R \subseteq \mathscr{E}_N$. So we already find that $\mathscr{E}_R \subseteq \mathscr{E}_N \cap \mathscr{L}(\mathscr{X}_R)$.

We prove the converse inequality. Let $f \in \mathscr{E}_N \cap \mathscr{L}(\mathscr{X}_R)$ and assume *ex absurdo* that $f \notin \mathscr{E}_R$. It follows from Lemma 35(iii) that $f \neq 0$. Since $f \in \mathscr{E}_N$, there are $S \subseteq N$, $f_s \in \mathscr{A}_s^N$, $s \in S$ and $g \in \mathscr{L}(\mathscr{X}_N)$ with $g \geq 0$ such that $f = g + \sum_{s\in S} f_s$. Clearly $S \setminus R \neq \emptyset$, because $S \setminus R = \emptyset$ would imply that, with $x_{N\setminus R}$ any element of $\mathscr{X}_{N\setminus R}$, $f = f(\cdot, x_{N\setminus R}) = g(\cdot, x_{N\setminus R}) + \sum_{s\in S\cap R} f_s(\cdot, x_{N\setminus R}) \in \mathscr{E}_R$, since we infer from Lemma 38 below that $f_s(\cdot, x_{N\setminus R}) \in \mathscr{A}_s^R \cup \{0\}$ for all $s \in S \cap R$.

It follows from Lemma 35(iii) and Lemma 34 and $f \notin \mathscr{E}_R$ that $0 \notin \operatorname{posi}(\mathscr{E}_R \cup \{-f\})$. Let $\mathscr{A} := \left\{ f_s(\cdot, z_{N\setminus R}) \colon s \in S \cap R, z_{N\setminus R} \in \mathscr{X}_{N\setminus R}, f_s(\cdot, z_{N\setminus R}) \neq 0 \right\}$. Then $\mathscr{A}$ is clearly a finite subset of $\mathscr{E}_R$ [to see this, use a similar argument as above, involving Lemma 38], so we deduce from Lemma 35(iii) and Lemma 33 that there is some linear prevision P_R on $\mathscr{L}(\mathscr{X}_R)$ with mass function p_R such that

$$\begin{cases} (\forall x_R \in \mathscr{X}_R) p_R(x_R) > 0 \\ (\forall s \in S \cap R)(\forall z_{N\setminus R} \in \mathscr{X}_{N\setminus R}) P_R(f_s(\cdot, z_{N\setminus R})) \geq 0 \\ P_R(f) < 0. \end{cases}$$

We then infer from $f = f(\cdot, z_{N\setminus R}) = g(\cdot, z_{N\setminus R}) + \sum_{s\in S\cap R} f_s(\cdot, z_{N\setminus R}) + \sum_{s\in S\setminus R} f_s(\cdot, z_{N\setminus R})$ that for all $z_{N\setminus R}$ in $\mathscr{X}_{N\setminus R}$:

$$\begin{aligned} 0 > P_R(f) - P_R(g(\cdot, z_{N\setminus R})) - \sum_{s\in S\cap R} P_R(f_s(\cdot, z_{N\setminus R})) &= \sum_{s\in S\setminus R} P_R(f_s(\cdot, z_{N\setminus R})) \\ &= \sum_{s\in S\setminus R} \sum_{x_R\in\mathscr{X}_R} p_R(x_R) f_s(x_R, z_{N\setminus R})). \end{aligned}$$

The gambles $f_s(x_R, \cdot)$ on $\mathscr{X}_{N\setminus R}$ [where $x_R \in \mathscr{X}_R$ and $s \in S \setminus R$] can clearly not all be zero, and the non-zero ones belong to $\mathscr{E}_{N\setminus R}$ by Lemma 38. Since $\mathscr{E}_{N\setminus R}$ is closed under positive linear combinations by Lemma 35, the gamble $h := \sum_{s\in N\setminus R} \sum_{x_R\in\mathscr{X}_R} p_R(x_R) f_s(x_R, \cdot)$ is an element of $\mathscr{E}_{N\setminus R}$ that is everywhere strictly negative. But on the other hand we should have that $\max h > 0$ by Lemma 35(iii), a contradiction. We may therefore conclude that indeed $f \in \mathscr{E}_R$. □

Lemma 38. *Consider an arbitrary non-empty subset R of N and let $r \in R$. Then $f(\cdot, x_{N\setminus R}) \in \mathscr{A}_r^R \cup \{0\}$ for all $f \in \mathscr{A}_r^N$ and all $x_{N\setminus R} \in \mathscr{X}_{N\setminus R}$.*

Proof. Fix $f \in \mathscr{A}_r^N$ and $x_{N\setminus R} \in \mathscr{X}_{N\setminus R}$ and consider the gamble $g := f(\cdot, x_{N\setminus R})$ on $\mathscr{X}_R$. Then it follows from the assumptions that for all $x_{R\setminus\{r\}} \in \mathscr{X}_{R\setminus\{r\}}$:

$$g(\cdot, x_{R\setminus\{r\}}) = f(\cdot, x_{R\setminus\{r\}}, x_{N\setminus R}) = f(\cdot, x_{N\setminus\{r\}}) \in \mathscr{A}_r \cup \{0\},$$

whence indeed $g \in \mathscr{A}_r^R \cup \{0\}$. □

Proof of Theorem 18. Use Lemma 36 and the fact that $f - \alpha \in \mathscr{E}_N \Leftrightarrow f - \alpha \in \mathscr{E}_R$ for all $f \in \mathscr{L}(\mathscr{X}_R)$, by Lemma 37. □

Next, consider an arbitrary subset R of N and consider, for each $x_R \in \mathscr{X}_R$, the set of gambles:

$$\mathscr{E}_N \rfloor x_R := \left\{ f \in \mathscr{L}(\mathscr{X}_{N\setminus R}) \colon \mathbb{I}_{\{x_R\}} f \in \mathscr{E}_N \right\}.$$

Note that if R is the empty set then we obtain trivially that $\mathscr{E}_N \rfloor x_\emptyset = \mathscr{E}_N$.

Then the set $\mathscr{E}_N$ satisfies the following interesting epistemic irrelevance condition:

Lemma 39. *For every subset R of N and every $x_{N\setminus R} \in \mathscr{X}_{N\setminus R}$, $\mathscr{E}_N|x_{N\setminus R} = \mathscr{E}_R$.*

Proof. The proof is similar to that of Lemma 37. Again, it is trivial for $R = \emptyset$ or $R = N$, so we turn to the case where both R and $N \setminus R$ are proper subsets of N. We first show that $\mathscr{E}_N|x_{N\setminus R} \supseteq \mathscr{E}_R$. Consider any gamble $f \in \mathscr{E}_R$, so there are non-negative λ_r and μ, $f_r \in \mathscr{A}_r^R$ for all $r \in R$ and $g \in \mathscr{L}(\mathscr{X}_R)_{>0}$ such that $f = \mu g + \sum_{r\in R} \lambda_r f_r$, with $\max\{\mu, \max_{r\in R} \lambda_r\} > 0$. Fix $r \in R$ and let $f_r' := \mathbb{I}_{\{x_{N\setminus R}\}} f_r \in \mathscr{L}(\mathscr{X}_N)$. Then $f_r' \neq 0$, and for all $z_{N\setminus\{r\}} \in \mathscr{X}_{N\setminus\{r\}}$, it follows from the definition of $\mathscr{A}_r^R$ that

$$f_r'(\cdot, z_{N\setminus\{r\}}) = \mathbb{I}_{\{x_{N\setminus R}\}}(z_{N\setminus R}) f_r(\cdot, z_{R\setminus\{r\}}) \in \mathscr{A}_r \cup \{0\},$$

and therefore the definition of $\mathscr{A}_r^N$ tells us that $f_r' \in \mathscr{A}_r^N$. Similarly, if we let $g' := \mathbb{I}_{\{x_{N\setminus R}\}} g \in \mathscr{L}(\mathscr{X}_N)$, then $g' > 0$. Hence $\mathbb{I}_{\{x_{N\setminus R}\}} f = \mu g' + \sum_{r\in R} \lambda_r f_r'$, and therefore $\mathbb{I}_{\{x_{N\setminus R}\}} f \in \mathscr{E}_N$.

We now turn to the converse inequality $\mathscr{E}_N|x_{N\setminus R} \subseteq \mathscr{E}_R$. Consider any gamble $f \in \mathscr{L}(\mathscr{X}_R)$ such that $\mathbb{I}_{\{x_{N\setminus R}\}} f$ belongs to $\mathscr{E}_N$ and assume *ex absurdo* that $f \notin \mathscr{E}_R$. Since $\mathbb{I}_{\{x_{N\setminus R}\}} f \in \mathscr{E}_N$, there are $S \subseteq N$, $f_s \in \mathscr{A}_s^N$, $s \in S$ and $g \in \mathscr{L}(\mathscr{X}_N)$ with $g \geq 0$ such that $\mathbb{I}_{\{x_{N\setminus R}\}} f = g + \sum_{s\in S} f_s$. Clearly $S \setminus R \neq \emptyset$, because $S \setminus R = \emptyset$ would imply that $f = g(\cdot, x_{N\setminus R}) + \sum_{s\in S\cap R} f_s(\cdot, x_{N\setminus R}) \in \mathscr{E}_R$, since Lemma 38 shows that $f_s(\cdot, x_{N\setminus R}) \in \mathscr{A}_s^R$ for all $s \in S \cap R$.

It follows from Lemma 35(iii), Lemma 34 and $f \notin \mathscr{E}_R$ that $0 \notin \operatorname{posi}(\mathscr{E}_R \cup \{-f\})$. Let $\mathscr{A} := \{f_s(\cdot, x_{N\setminus R}) \colon s \in S \cap R, f_s(\cdot, x_{N\setminus R}) \neq 0\}$. Then $\mathscr{A}$ is clearly a finite subset of $\mathscr{E}_R$ [to see this, use a similar argument as above, involving Lemma 38], so we deduce from Lemma 33 that there is some linear prevision P_R on $\mathscr{L}(\mathscr{X}_R)$ with mass function p_R such that

$$\begin{cases} (\forall x_R \in \mathscr{X}_R) p_R(x_R) > 0 \\ (\forall s \in S \cap R) P_R(f_s(\cdot, x_{N\setminus R})) \geq 0 \\ P_R(f) < 0. \end{cases}$$

We then infer from $f = g(\cdot, x_{N\setminus R}) + \sum_{s\in S\cap R} f_s(\cdot, x_{N\setminus R}) + \sum_{s\in S\setminus R} f_s(\cdot, x_{N\setminus R})$ that:

$$\begin{aligned} 0 > P_R(f) - P_R(g(\cdot, x_{N\setminus R})) - \sum_{s\in S\cap R} P_R(f_s(\cdot, x_{N\setminus R})) &= \sum_{s\in S\setminus R} P_R(f_s(\cdot, x_{N\setminus R})) \\ &= \sum_{s\in S\setminus R} \sum_{x_R\in\mathscr{X}_R} p_R(x_R) f_s(x_R, x_{N\setminus R})). \end{aligned}$$

Similarly, since $0 = g(\cdot, z_{N\setminus R}) + \sum_{s\in S\cap R} f_s(\cdot, z_{N\setminus R}) + \sum_{s\in S\setminus R} f_s(\cdot, z_{N\setminus R})$ for all $z_{N\setminus R} \in \mathscr{X}_{N\setminus R} \setminus \{x_{N\setminus R}\}$, we infer that:

$$\begin{aligned} 0 \geq -P_R(g(\cdot, z_{N\setminus R})) - \sum_{s\in S\cap R} P_R(f_s(\cdot, z_{N\setminus R})) &= \sum_{s\in S\setminus R} P_R(f_s(\cdot, z_{N\setminus R})) \\ &= \sum_{s\in S\setminus R} \sum_{x_R\in\mathscr{X}_R} p_R(x_R) f_s(x_R, z_{N\setminus R})). \end{aligned}$$

Hence

$$h := \sum_{s\in S\setminus R} \sum_{x_R\in\mathscr{X}_R} p_R(x_R) f_s(x_R, \cdot) < 0.$$

The gambles $f_s(x_R, \cdot)$ on $\mathscr{X}_{N\setminus R}$ [where $x_R \in \mathscr{X}_R$ and $s \in S \setminus R$] can clearly not all be zero, and the non-zero ones belong to $\mathscr{E}_{N\setminus R}$ by Lemma 38. Since $\mathscr{E}_{N\setminus R}$ is closed under positive linear combinations by Lemma 35, the gamble $h < 0$ is an element of $\mathscr{E}_{N\setminus R}$. But on the other hand we should have that $\max h > 0$ by Lemma 35(iii), a contradiction. We may therefore conclude that indeed $f \in \mathscr{E}_R$. □

From this lemma, we deduce the following:

Lemma 40. *Consider any subset R of N. Then $fg \in \mathscr{E}_N$ for all $f \in \mathscr{L}(\mathscr{X}_{N\setminus R})_{>0}$ and all $g \in \mathscr{E}_R$.*

Proof. Consider any $g \in \mathscr{E}_R$ and any $f \in \mathscr{L}(\mathscr{X}_{N\setminus R})_{>0}$. Then it follows from Lemma 39 that $\mathbb{I}_{\{x_{N\setminus R}\}}g \in \mathscr{E}_N$ for all $x_{N\setminus R} \in \mathscr{X}_{N\setminus R}$. But $\mathscr{E}_N$ is closed under positive linear combinations by Lemma 35, and there is at least one $x_{N\setminus R} \in \mathscr{X}_{N\setminus R}$ for which $f(x_{N\setminus R}) > 0$, so we deduce that $fg = \sum_{x_{N\setminus R}\in\mathscr{X}_{N\setminus R}} f(x_{N\setminus R})\mathbb{I}_{\{x_{N\setminus R}\}}g$ is essentially a positive linear combination of gambles in $\mathscr{E}_N$, and therefore also belongs to $\mathscr{E}_N$. □

Proof of Proposition 19. We begin by showing that $\underline{E}_N$ is productive. First, consider arbitrary disjoint *proper* subsets I and O of N, $x_I \in \mathscr{X}_I$, $g \in \mathscr{L}(\mathscr{X}_O)$ and non-negative $f \in \mathscr{L}(\mathscr{X}_I)$, and let us prove that $\underline{E}_N(f[g-\underline{E}_N(g)]) \geq 0$.

Fix $\alpha > 0$ and $\beta > 0$, then $f+\beta > 0$ and $g-\underline{E}_O(g)+\beta \in \mathscr{E}_O$, by Lemma 36 [where we replace the set N with O]. Hence $(f+\alpha)[g-\underline{E}_O(g)+\beta] \in \mathscr{E}_{O\cup I} \subseteq \mathscr{E}_N$, by Lemmas 40 [where we replace N with $O\cup I$ and R with O] and 37. Then Lemma 36 tells us that $\underline{E}_N((f+\alpha)[g-\underline{E}_O(g)+\beta]) \geq 0$. If we now invoke the coherence of the lower prevision $\underline{E}_N$, we see that

$$\begin{aligned} 0 &\leq \underline{E}_N((f+\alpha)[g-\underline{E}_O(g)+\beta]) \\ &\leq \underline{E}_N(f[g-\underline{E}_O(g)]) + \beta\overline{E}_N(f) + \alpha[\overline{E}_N(g) - \underline{E}_O(g)+\beta]. \end{aligned}$$

Since this holds for all $\alpha > 0$ and all $\beta > 0$, we infer that $\underline{E}_N(f[g-\underline{E}_N(g)]) = \underline{E}_N(f[g-\underline{E}_O(g)]) \geq 0$, where the equality follows from Theorem 18. Hence $\underline{E}_N$ is indeed productive.

Next let us consider the special cases where I or O are empty. If $I = \emptyset$, then we obtain $\underline{E}_N(f[g-\underline{E}_N(g)]) = f(\underline{E}_N(g)-\underline{E}_N(g)) = 0$, since f is then a non-negative real number. If on the other hand $O = \emptyset$, then $\underline{E}_N(f[g-\underline{E}_N(g)]) = \underline{E}_N(f\cdot 0) = 0$, since in this case g is a real number.

This implies in particular that $\underline{E}_N(\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]) \geq 0$ for arbitrary disjoint subsets I and O of N. Assume *ex absurdo* that $\underline{E}_N(\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]) > 0$. By Lemma 36, there is some $\alpha > 0$ such that $\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]-\alpha \in \mathscr{E}_N$. Since $\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)-\alpha] \geq \mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]-\alpha$, this implies that $\mathbb{I}_{\{x_I\}}[g-\underline{E}_N(g)-\alpha] \in \mathscr{E}_N$. By Lemmas 39 and 37, this implies that $g-\underline{E}_O(g)-\alpha \in \mathscr{E}_O$. But then Lemma 36 implies that $-\alpha = \underline{E}_O(g-\underline{E}_O(g)-\alpha) \geq 0$, a contradiction. Hence indeed $\underline{E}_N(\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]) = 0$. □

Proof of Theorem 20. By Theorem 2, it suffices to prove that (a) $\underline{E}_N$ is weakly coherent with the family $\mathscr{I}(\underline{E}_N)$; and that (b) the family $\mathscr{I}(\underline{E}_N)$ is coherent.

We begin by showing that (a) $\underline{E}_N$ is weakly coherent with the family $\mathscr{I}(\underline{E}_N)$. Consider any disjoint subsets I and O of N. Taking into account Theorem 1, it suffices to show that $\underline{E}_N(\mathbb{I}_{\{x_I\}}[f-\underline{E}_{O\cup I}(f|x_I)]) = 0$ for all $x_I \in \mathscr{X}_I$ and all $f \in \mathscr{L}(\mathscr{X}_{O\cup I})$. If we look at Eq. (13), we see that this amounts to proving that $\underline{E}_N(\mathbb{I}_{\{x_I\}}[g-\underline{E}_O(g)]) = 0$ for all $x_I \in \mathscr{X}_I$ and all $g \in \mathscr{L}(\mathscr{X}_O)$. Now use Proposition 19.

To finish the proof, we show that (b) the family $\mathscr{I}(\underline{E}_N)$ is coherent. Assume *ex absurdo* that there are $f_{O,I} \in \mathscr{L}(\mathscr{X}_{O\cup I})$ for all disjoint subsets I and O of N, disjoint subsets I^* and O^* of N, $g \in \mathscr{L}(\mathscr{X}_{O^*\cup I^*})$, $x_{I^*} \in \mathscr{X}_{I^*}$ and $\delta > 0$ such that

$$\sum_{O,I} G_{O\cup I}(f_{O,I}|X_I) - G_{O^*\cup I^*}(g|x_{I^*}) \leq -\delta\mathbb{I}_A,$$

where $A := \{x_{I^*}\} \cup \bigcup_{O,I} \operatorname{supp}_I(f_{O,I})$. But $K\mathbb{I}_A \geq \mathbb{I}_{\{x_{I^*}\}} + \sum_{O,I}\mathbb{I}_{\operatorname{supp}_I(f_{O,I})}$ for some natural number $K > 0$, so there is some $\varepsilon := \delta/K > 0$ such that $-\delta\mathbb{I}_A \leq -\varepsilon\mathbb{I}_{\{x_{I^*}\}} - \sum_{O,I}\varepsilon\mathbb{I}_{\operatorname{supp}_I(f_{O,I})}$ and therefore also [see Theorem 18]:

$$\mathbb{I}_{\{x_{I^*}\}}[g(\cdot,x_{I^*}) - \underline{E}_{O^*}(g(\cdot,x_{I^*})) - \varepsilon] \geq \sum_{O,I}[G_{O\cup I}(f_{O,I}|X_I) + \varepsilon\mathbb{I}_{\operatorname{supp}_I(f_{O,I})}]. \tag{17}$$

For arbitrary disjoint I and O, it follows from the definition of $\operatorname{supp}_I(f_{O,I})$ and Theorem 18 that

$$G_{O\cup I}(f_{O,I}|X_I) + \varepsilon\mathbb{I}_{\operatorname{supp}_I(f_{O,I})} = \sum_{x_I\in\operatorname{supp}_I(f_{O,I})} \mathbb{I}_{\{x_I\}}[f_{O,I}(\cdot,x_I) - \underline{E}_O(f_{O,I}(\cdot,x_I)) + \varepsilon],.$$

so we infer from Lemmas 36 and 40 that the gamble $G_{O\cup I}(f_{O,I}|X_I)+\varepsilon\mathbb{I}_{\mathrm{supp}_I(f_{O,I})}$ belongs to the convex cone $\mathscr{E}_N$. So does, therefore, the right-hand side in Eq. (17). As a consequence, $\mathbb{I}_{\{x_{I^*}\}}[g(\cdot,x_{I^*})-\underline{E}_{O^*}(g(\cdot,x_{I^*}))-\varepsilon]\in\mathscr{E}_N$, so we infer from Lemmas 39 and 37 that $g(\cdot,x_{I^*})-\underline{E}_{O^*}(g(\cdot,x_{I^*}))-\varepsilon\in\mathscr{E}_{O^*}$. But then Lemma 36 and the coherence of $\underline{E}_{O^*}$ lead to $-\varepsilon=\underline{E}_{O^*}(g(\cdot,x_{I^*})-\underline{E}_{O^*}(g(\cdot,x_{I^*}))-\varepsilon)\geq 0$, a contradiction. □

Proof of Theorem 22. Denote the independent natural extension $\otimes_{n\in N}\underline{P}_n$ by $\underline{E}_N$, and the strong product $\times_{n\in N}\underline{P}_n$ by $\underline{S}_N$. We use the characterisation of factorisation in Proposition 5, and we take into account that for all $n\in N$ and $f_n\in\mathscr{L}(\mathscr{X}_n)$, $\underline{E}_N(f_n)=\underline{S}_N(f_n)=\underline{P}_n(f_n)$.

We begin by proving an auxiliary result, namely that

$$\underline{E}_N\left(\prod_{i\in I}f_i\right)=\prod_{i\in I}\underline{P}_i(f_i)\text{ for all non-negative }f_i\in\mathscr{L}(\mathscr{X}_i),\ i\in I,\qquad (H_I^1)$$

for all subsets I of N. We give a proof by induction. It is clear from the coherence of $\underline{E}_N$ that the statement (H_I^1) holds for $I=\emptyset$. Next, assume that (H_I^1) holds for $I=M$, where M is some subset of N that does not coincide with N [this is the induction hypothesis]. Then the statement is proved if we can show that (H_I^1) holds for $I=M':=M\cup\{n\}$, where n is any element of the non-empty set $N\setminus M$.

So consider any non-negative $f_i\in\mathscr{L}(\mathscr{X}_i)$, $i\in M'$. Let, for ease of notation, $f_M:=\prod_{i\in M}f_i$, so $\prod_{i\in M'}f_i=f_Mf_n$. Because of the induction hypothesis, $\underline{E}_N(f_M)=\prod_{i\in M}\underline{P}_i(f_i)$, so we have to show that $\underline{E}_N(f_Mf_n)=\underline{E}_N(f_M)\underline{P}_n(f_n)$. Because the strong product is factorising [see Proposition 8 and Proposition 5], we also have that $\underline{S}_N(f_M)=\prod_{i\in M}\underline{P}_i(f_i)$, and therefore $\underline{E}_N(f_M)=\underline{S}_N(f_M)$.

Since $\underline{E}_N$ is dominated by the (many-to-one independent) strong product $\underline{S}_N$ [see Definition 6 and Proposition 12], and because the strong product is factorising [see Proposition8 and Proposition 5], we see that

$$\underline{E}_N(f_Mf_n)\leq\underline{S}_N(f_Mf_n)=\underline{S}_N(f_M)\underline{P}_n(f_n)=\underline{E}_N(f_M)\underline{P}_n(f_n).$$

We now prove the converse inequality. Recall from its definition that $\underline{E}_N$ is coherent with the conditional lower prevision $\underline{P}_{\{n\}\cup M}(\cdot|X_M)$ on $\mathscr{L}(\mathscr{X}_{\{n\}\cup M})$ defined by

$$\underline{P}_{\{n\}\cup M}(h|x_M):=\underline{P}_n(h(\cdot,x_M))\text{ for all }h\in\mathscr{L}(\mathscr{X}_{\{n\}\cup M})\text{ and all }x_M\in\mathscr{X}_M,\qquad (18)$$

so it follows that:

$$\underline{E}_N(f_Mf_n)\geq\underline{E}_N(\underline{P}_{\{n\}\cup M}(f_Mf_n|X_M))=\underline{E}_N(f_M\underline{P}_n(f_n))=\underline{E}_N(f_M)\underline{P}_n(f_n).$$

Here, the inequality follows from the coherence of $\underline{E}_N$ with $\underline{P}_n(\cdot|X_M)$ and Eq. (4), the first equality from the fact that $\underline{P}_n(f_Mf_n|X_M)=f_M\underline{P}_n(f_n|X_M)$ by Eq. (1) and Eq. (18), and the last equality from the coherence of $\underline{E}_N$ and the fact that $\underline{P}_n(f_n)\geq 0$. This completes the proof of the auxiliary result.

The proof that $\underline{E}_N$ is factorising goes along similar lines. Fix any o in N and any $I\subseteq N\setminus\{o\}$, any $f_o\in\mathscr{L}(\mathscr{X}_o)$, and non-negative $f_i\in\mathscr{L}(\mathscr{X}_i)$, $i\in I$. Let, for ease of notation, $f_I:=\prod_{i\in I}f_i$. Then we already know from the argumentation above that $\underline{E}_N(f_I)=\underline{S}_N(f_I)=\prod_{i\in I}\underline{P}_i(f_i)$. We have to show that $\underline{E}_N(f_of_I)=\underline{E}_N(f_I\underline{P}_o(f_o))$.

As before, since $\underline{E}_N$ is dominated by the (many-to-one independent) strong product $\underline{S}_N$, and because the strong product is factorising, we see that

$$\underline{E}_N(f_of_I)\leq\underline{S}_N(f_of_I)=\overline{S}_N(f_I)\odot\underline{P}_o(f_o)=\overline{E}_N(f_I)\odot\underline{P}_o(f_o)=\underline{E}_N(f_I\underline{P}_o(f_o)).$$

We now prove the converse inequality. Recall from its definition that $\underline{E}_N$ is coherent with the conditional lower prevision $\underline{P}_{\{o\}\cup I}(\cdot|X_I)$ on $\mathscr{L}(\mathscr{X}_o)$ defined by

$$\underline{P}_o(h|x_I):=\underline{P}_o(h(\cdot,x_I))\text{ for all }h\in\mathscr{L}(\mathscr{X}_{\{o\}\cup I})\text{ and all }x_I\in\mathscr{X}_I,\qquad (19)$$

so it follows that:

$$\underline{E}_N(f_of_I)\geq\underline{E}_N(\underline{P}_{\{o\}\cup I}(f_of_I|X_I))=\underline{E}_N(f_I\underline{P}_{\{o\}\cup I}(f_o|X_I))=\underline{E}_N(f_I\underline{P}_o(f_o)).$$

Here, similarly as before, the inequality follows from the coherence of $\underline{E}_N$ with $\underline{P}_o(\cdot|X_I)$ and Eq. (4), the first equality from the fact that $\underline{P}_{\{o\}\cup I}(f_o f_I|X_I) = f_I \underline{P}_{\{o\}\cup I}(f_o|X_I)$ by Eq. (1), and the second equality from Eq. (19). □

Proof of Theorem 23. We construct the set $\mathscr{E}_{\{N_1,N_2\}}$ after the fashion of Eq. (16). We let

$$\mathscr{A}_{N_1} := \left\{f \in \mathscr{L}(\mathscr{X}_{N_1})\colon f > 0 \text{ or } \underline{E}_{N_1}(f) > 0\right\}$$

and

$$\mathscr{A}_{N_1}^{\{N_1,N_2\}} := \left\{f \in \mathscr{L}(\mathscr{X}_{\{N_1,N_2\}})_{\neq 0}\colon (\forall z_{N_2} \in \mathscr{X}_{N_2}) f(\cdot, z_{N_2}) \in \mathscr{A}_{N_1} \cup \{0\}\right\}$$

and similarly for $\mathscr{A}_{N_2}$ and $\mathscr{A}_{N_2}^{\{N_1,N_2\}}$. Then

$$\mathscr{E}_{\{N_1,N_2\}} := \operatorname{posi}\left(\mathscr{L}(\mathscr{X}_{\{N_1,N_2\}})_{>0} \cup \mathscr{A}_{N_1}^{\{N_1,N_2\}} \cup \mathscr{A}_{N_2}^{\{N_1,N_2\}}\right). \tag{20}$$

We infer from Lemma 36 that $\mathscr{A}_{N_1} \subseteq \mathscr{E}_{N_1}$ and $\mathscr{A}_{N_2} \subseteq \mathscr{E}_{N_2}$. Now let $f_1 \in \mathscr{A}_{N_1}^{\{N_1,N_2\}}$ and consider any $z_{N_2} \in \mathscr{X}_{N_2}$. Then $f_1(\cdot, z_{N_2}) \in \mathscr{E}_{N_1} \cup \{0\}$, and therefore[24] $\mathbb{I}_{\{z_{N_2}\}} f_1(\cdot, z_{N_2}) \in \mathscr{E}_{N_1\cup N_2} \cup \{0\}$ by Lemma 40. Since $f_1 \neq 0$ and $\mathscr{E}_{N_1\cup N_2}$ is a convex cone by Lemma 35(iii), it follows that $f_1 = \sum_{z_{N_2}\in\mathscr{X}_{N_2}} \mathbb{I}_{\{z_{N_2}\}} f_1(\cdot, z_{N_2}) \in \mathscr{E}_{N_1\cup N_2}$. Hence $\mathscr{A}_{N_1}^{\{N_1,N_2\}} \subseteq \mathscr{E}_{N_1\cup N_2}$ and similarly, $\mathscr{A}_{N_2}^{\{N_1,N_2\}} \subseteq \mathscr{E}_{N_1\cup N_2}$, so we infer from Eq. (20) that $\mathscr{E}_{\{N_1,N_2\}} \subseteq \mathscr{E}_{N_1\cup N_2}$. Hence $\underline{E}_{\{N_1,N_2\}} \leq \underline{E}_{N_1\cup N_2}$, by Lemma 36.

For the converse inequality, it suffices to prove that $\underline{E}_{\{N_1,N_2\}}$ is an independent many-to-one product of the marginals $\underline{P}_n$, $n \in N_1 \cup N_2$. We use Corollary 14(ii). Consider any $o \in N_1 \cup N_2$, any $I \subseteq (N_1 \cup N_2) \setminus \{o\}$, any $g \in \mathscr{L}(\mathscr{X}_o)$, and any $x_I \in \mathscr{X}_I$. We have to prove that $\underline{E}_{\{N_1,N_2\}}(\mathbb{I}_{\{x_I\}}[g - \underline{P}_o(g)]) = 0$. Let $I_1 := I \cap N_1$ and $I_2 := I \cap N_2$. We may assume without loss of generality that $o \in N_2$. Since the independent natural extension is factorising [by Theorem 22], we find that indeed

$$\begin{aligned}\underline{E}_{\{N_1,N_2\}}(\mathbb{I}_{\{x_I\}}[g - \underline{P}_o(g)]) &= \underline{E}_{\{N_1,N_2\}}(\mathbb{I}_{\{x_{I_1}\}}\mathbb{I}_{\{x_{I_2}\}}[g - \underline{P}_o(g)]) \\ &= \overline{E}_{N_1}(\mathbb{I}_{\{x_{I_1}\}}) \odot \underline{E}_{N_2}(\mathbb{I}_{\{x_{I_2}\}}[g - \underline{P}_o(g)]) = \overline{E}_{N_1}(\mathbb{I}_{\{x_{I_1}\}}) \odot 0 = 0,\end{aligned}$$

where the third equality follows from Corollary 14(ii). □

Proof of Theorem 24. Consider arbitrary disjoint subsets I and O of N, an arbitrary gamble g on $\mathscr{X}_O$ and an arbitrary non-negative gamble f on $\mathscr{X}_I$. We have to show that $\underline{E}_N(fg) = \underline{E}_N(f\underline{E}_N(g))$. Consider any partition N_1 and N_2 of N such that $I \subseteq N_1$ and $O \subseteq N_2$. Since the independent natural extension $\underline{E}_{\{N_1,N_2\}} = \underline{E}_{N_1} \otimes \underline{E}_{N_2}$ of $\underline{E}_{N_1}$ and $\underline{E}_{N_2}$ is factorising by Theorem 22, we see that $\underline{E}_{\{N_1,N_2\}}(fg) = \underline{E}_{\{N_1,N_2\}}(f\underline{E}_{\{N_1,N_2\}}(g))$. Now use Theorem 23 to find that this implies that indeed $\underline{E}_N(fg) = \underline{E}_N(f\underline{E}_N(g))$. □

Proof of Proposition 25. Let $\underline{P}_{\{1,2\}}$ be any (many-to-one) independent product of P_1 and $\underline{P}_2$, and consider the conditional linear prevision $P_{\{1,2\}}(\cdot|X_2)$ on $\mathscr{L}(\mathscr{X}_{\{1,2\}})$ defined by

$$P_{\{1,2\}}(f|x_2) := P_1(f(\cdot, x_2)) \text{ for all } f \in \mathscr{L}(\mathscr{X}_{\{1,2\}}) \text{ and } x_2 \in \mathscr{X}_2.$$

Then $\underline{P}_{\{1,2\}}$ is in particular coherent with $P_{\{1,2\}}(\cdot|X_2)$. Fix f in $\mathscr{L}(\mathscr{X}_{\{1,2\}})$. It follows from Eq. (4) and the self-conjugacy of $P_{\{1,2\}}(\cdot|X_2)$ that $\underline{P}_{\{1,2\}}(f) = \underline{P}_{\{1,2\}}(P_{\{1,2\}}(f|X_2))$. Moreover, it follows from Corollary 15 that $\underline{P}_2$ is the $\mathscr{X}_2$-marginal of $\underline{P}_{\{1,2\}}$, so $\underline{P}_{\{1,2\}}(f) = \underline{P}_2(P_{\{1,2\}}(f|X_2)) = \underline{P}_2(P_1(f))$. This holds in particular for $\underline{P}_N = (P_1 \times \underline{P}_2)$ and $\underline{P}_N = (P_1 \otimes \underline{P}_2)$. □

[24] We can identify a gamble on $\mathscr{X}_{\{N_1,N_2\}}$ in a trivial way with a unique corresponding gamble on $\mathscr{X}_{N_1\cup N_2}$.

Proof of Proposition 26. Let $\underline{P}_{\{1,2\}}(\cdot|X_1)$ be the conditional lower prevision on $\mathscr{L}(\mathscr{X}_{\{1,2\}})$ defined by

$$\underline{P}_{\{1,2\}}(f|x_1) := \underline{P}_2(f(x_1,\cdot)) \text{ for all } f \in \mathscr{L}(\mathscr{X}_{\{1,2\}}) \text{ and } x_1 \in \mathscr{X}_1.$$

Then the independent natural extension $\underline{P}_1^{A_1} \otimes \underline{P}_2$ is coherent with $\underline{P}_{\{1,2\}}(\cdot|X_1)$, so we infer from Eq. (4) that $\underline{P}_1^{A_1} \otimes \underline{P}_2 \geq (\underline{P}_1^{A_1} \otimes \underline{P}_2)(\underline{P}_{\{1,2\}}(\cdot|X_1)) = \underline{P}_1^{A_1}(\underline{P}_{\{1,2\}}(\cdot|X_1))$, where the equality follows from Corollary 15. So we find that

$$(\underline{P}_1^{A_1} \times \underline{P}_2)(f) \geq (\underline{P}_1^{A_1} \otimes \underline{P}_2)(f) \geq \min_{x_1 \in A_1} \underline{P}_2(f(x_1,\cdot))$$

for every gamble f on $\mathscr{X}_1 \times \mathscr{X}_2$.

To prove that the equalities hold, consider any gamble f on $\mathscr{X}_{\{1,2\}}$. Since A_1 is finite there is some $x_1^* \in A_1$ such that $\underline{P}_2(f(x_1^*,\cdot)) = \min_{x_1 \in A_1} \underline{P}_2(f(x_1,\cdot))$. Moreover, it follows from the coherence of the lower prevision $\underline{P}_2$ that there is some linear prevision $P_2 \in \operatorname{ext}(\mathscr{M}(\underline{P}_2))$ such that $P_2(f(x_1^*,\cdot)) = \underline{P}_2(f(x_1^*,\cdot))$. Let $P_1 := \underline{P}_1^{\{x_1^*\}}$ denote the (degenerate) linear prevision on $\mathscr{L}(\mathscr{X}_1)$, all of whose probability mass lies in x_1^*. Observe that $P_1 \in \operatorname{ext}(\mathscr{M}(\underline{P}_1^{A_1}))$. Then the definition of the strong product implies that

$$\min_{x_1 \in A_1} \underline{P}_2(f(x_1,\cdot)) = \underline{P}_2(f(x_1^*,\cdot)) = P_2(f(x_1^*,\cdot)) = (P_1 \times P_2)(f) \geq (\underline{P}_1^{A_1} \times \underline{P}_2)(f).$$

We turn to the second statement. Let $\underline{P}$ be any factorising product of $\underline{P}_1^{A_1}$ and $\underline{P}_2$. Consider any gamble $f \in \mathscr{L}(\mathscr{X}_{\{1,2\}})$, and let us show that the equalities hold. It suffices to give a proof for non-negative f, since for arbitrary gambles we only need to add a non-negative constant [coherence guarantees that we can take additive real numbers out of the lower prevision operator]. Let a_1 and b_1 be elements of A_1 such that $\underline{P}_2(f(a_1,\cdot)) = \min_{x_1 \in A_1} \underline{P}_2(f(x_1,\cdot))$ and $\overline{P}_2(f(b_1,\cdot)) = \max_{x_1 \in A_1} \overline{P}_2(f(x_1,\cdot))$. Then

$$\begin{aligned}
&\underline{P}(f - \underline{P}_2(f(a_1,\cdot))) \\
&\quad = \underline{P}\left(\sum_{x_1 \in \mathscr{X}_1} \mathbb{I}_{\{x_1\}}[f(x_1,\cdot) - \underline{P}_2(f(a_1,\cdot))]\right) \\
&\quad \geq \sum_{x_1 \in \mathscr{X}_1} \underline{P}(\mathbb{I}_{\{x_1\}}[f(x_1,\cdot) - \underline{P}_2(f(a_1,\cdot))]) \\
&\quad = \sum_{x_1 \in A_1} \underline{P}(\mathbb{I}_{\{x_1\}}[f(x_1,\cdot) - \underline{P}_2(f(a_1,\cdot))]) + \sum_{x_1 \in A_1^c} \underline{P}(\mathbb{I}_{\{x_1\}}[f(x_1,\cdot) - \underline{P}_2(f(a_1,\cdot))]) \\
&\quad = \sum_{x_1 \in A_1} \underline{P}_1(\mathbb{I}_{\{x_1\}})[\underline{P}_2(f(x_1,\cdot)) - \underline{P}_2(f(a_1,\cdot))] + 0 \geq 0,
\end{aligned}$$

where the first inequality follows from the coherence of $\underline{P}$ and the last equality holds because $\underline{P}$ is assumed to be factorising and because $\underline{P}_1(\{x_1\}) = \overline{P}_1(\{x_1\}) = 0$ for every $x_1 \notin A_1$. Hence $\underline{P}(f) \geq \min_{x_1 \in A_1} \underline{P}_2(f(x_1,\cdot))$. Using the conjugacy between $\underline{P}$ and $\overline{P}$, we get $\overline{P}(f) \leq \max_{x_1 \in A_1} \overline{P}_2(f(x_1,\cdot))$.

On the other hand,

$$\overline{P}_2(f(b_1,\cdot)) = \overline{P}_1(\{b_1\})\overline{P}_2(f(b_1,\cdot)) = \overline{P}(\mathbb{I}_{\{b_1\}} f(b_1,\cdot)) \leq \overline{P}(f) \leq \max_{x_1 \in A_1} \overline{P}_2(f(x_1,\cdot)),$$

where the first equality holds because $\overline{P}_1(\{b_1\}) = 1$, and the second because $\underline{P}$ is assumed to be factorising. The first inequality holds because f is non-negative. We conclude that for any gamble f on $\mathscr{X}_{\{1,2\}}$, $\overline{P}(f) = \max_{x_1 \in A_1} \overline{P}_2(f(x_1,\cdot))$, and therefore also $\underline{P}(f) = \min_{x_1 \in A_1} \underline{P}_2(f(x_1,\cdot))$, by conjugacy.

We conclude with a proof for the third statement. Assume that $\underline{P}_2$ is the vacuous lower prevision $\underline{P}_2^{A_2}$ with respect to A_2, and let $\underline{P}$ be any coherent lower prevision on $\mathscr{L}(\mathscr{X}_{\{1,2\}})$ with marginals $\underline{P}_1^{A_1}$ and $\underline{P}_2^{A_2}$. Because it is coherent, $\underline{P}$ satisfies

$$\underline{P}(A_1 \times A_2) = 1 - \overline{P}(A_1^c \times \mathscr{X}_2 \cup \mathscr{X}_1 \times A_2^c) \geq 1 - \overline{P}_1^{A_1}(A_1^c) - \overline{P}_2^{A_2}(A_2^c) = 1,$$

and therefore it dominates the vacuous lower prevision $\underline{P}_{\{1,2\}}^{A_1\times A_2}$ relative to $A_1\times A_2$, which also is the independent natural extension $\underline{P}_1^{A_1}\otimes\underline{P}_2^{A_2}$ of $\underline{P}_1^{A_1}$ and $\underline{P}_2^{A_2}$, by the second statement.

There are now two possibilities. Either A_1 has more than one element. To prove that $\underline{P}$ is an independent product of $\underline{P}_1^{A_1}$ and $\underline{P}_2^{A_2}$, we use Corollary 14(ii). Consider any $z_1\in\mathscr{X}_1$ and any gamble f on $\mathscr{X}_2$. If $z_1\notin A_1$, then we find by inspection that $\underline{P}_{\{1,2\}}^{A_1\times A_2}(\mathbb{I}_{\{z_1\}}[f-\underline{P}_2(f)])=\overline{P}_{\{1,2\}}^{A_1\times A_2}(\mathbb{I}_{\{z_1\}}[f-\underline{P}_2(f)])=0$, and therefore $\underline{P}(\mathbb{I}_{\{z_1\}}[f-\underline{P}_2(f)])=0$. If $z_1\in A_1$, then

$$\begin{aligned}0=\underline{P}_{\{1,2\}}^{A_1\times A_2}(\mathbb{I}_{\{z_1\}}[f-\underline{P}_2(f)])&\leq\underline{P}(\mathbb{I}_{\{x_1\}}[f-\underline{P}_2(f)])\\&\leq\underline{P}(\mathbb{I}_{\{x_1\}}[\max f-\min f])=[\max f-\min f]\underline{P}_1^{A_1}(\{x_1\})=0.\end{aligned}$$

Here, the first and the last equalities hold because A_1 has more than one element.

Or A_1 has only one element x_1, and then Lemma 41 implies that $\underline{P}(f)=\underline{P}_2^{A_2}(f(x_1,\cdot))$ for all gambles f on $\mathscr{X}_1\times\mathscr{X}_2$, so there is only one coherent lower prevision that has these marginals. But in this case the marginal $\underline{P}_1^{A_1}$ is a linear prevision, and Proposition 25 tells us that the marginals $\underline{P}_1^{A_1}$ and $\underline{P}_2^{A_2}$ have only one independent product, and it is equal to $\underline{P}$. $\square$

Lemma 41. *Let $\underline{P}_1^{\{x_1\}}$ be the vacuous lower prevision on $\mathscr{L}(\mathscr{X}_1)$ relative to the singleton $\{x_1\}\subseteq\mathscr{X}_1$, and let $\underline{P}_2$ be any coherent lower prevision on $\mathscr{L}(\mathscr{X}_2)$. Let $\underline{P}$ be any coherent lower prevision with these marginals. Then $\underline{P}$ is unique and given by $\underline{P}(f)=\underline{P}_2(f(x_1,\cdot))$ for all gambles f on $\mathscr{X}_1\times\mathscr{X}_2$.*

Proof. First of all, consider any gamble g on $\mathscr{X}_1\times\mathscr{X}_2$ and any $z_1\in\mathscr{X}_1\setminus\{x_1\}$. We show that $\underline{P}(\mathbb{I}_{\{z_1\}}g)=\overline{P}(\mathbb{I}_{\{z_1\}}g)=0$. Indeed, by coherence of $\underline{P}$:

$$\begin{aligned}0&=\underline{P}_1^{\{x_1\}}(\mathbb{I}_{\{z_1\}})\min g(z_1,\cdot)=\overline{P}_1^{\{x_1\}}(\mathbb{I}_{\{z_1\}})\min g(z_1,\cdot)\\&\leq\underline{P}(\mathbb{I}_{\{z_1\}}g(z_1,\cdot))=\underline{P}(\mathbb{I}_{\{z_1\}}g)\\&\leq\overline{P}(\mathbb{I}_{\{z_1\}}g)=\overline{P}(\mathbb{I}_{\{z_1\}}g(z_1,\cdot))\\&\leq\overline{P}_1^{\{x_1\}}(\mathbb{I}_{\{z_1\}})\max g(z_1,\cdot)=\underline{P}_1^{\{x_1\}}(\mathbb{I}_{\{z_1\}})\max g(z_1,\cdot)=0.\end{aligned}$$

For any gamble f on $\mathscr{X}_1\times\mathscr{X}_2$ we then infer from the coherence of $\underline{P}$ that

$$\begin{aligned}\underline{P}(\mathbb{I}_{\{x_1\}}f)&=\underline{P}(\mathbb{I}_{\{x_1\}}f)+\sum_{z_1\in\mathscr{X}_1\setminus\{x_1\}}\underline{P}(\mathbb{I}_{\{z_1\}}f)\\&\leq\underline{P}\left(\sum_{z_1\in\mathscr{X}_1}\mathbb{I}_{\{z_1\}}f\right)=\underline{P}(f)\\&\leq\underline{P}(\mathbb{I}_{\{x_1\}}f)+\sum_{z_1\in\mathscr{X}_1\setminus\{x_1\}}\overline{P}(\mathbb{I}_{\{z_1\}}f)=\underline{P}(\mathbb{I}_{\{x_1\}}f).\end{aligned}$$

But this tells us that indeed

$$\underline{P}(f)=\underline{P}(\mathbb{I}_{\{x_1\}}f)=\underline{P}(\mathbb{I}_{\{x_1\}}f(x_1,\cdot))=\underline{P}(f(x_1,\cdot))=\underline{P}_2(f(x_1,\cdot)).\qquad\square$$

Proof of Proposition 27. We begin with the strong product. Because of its marginalisation and associativity properties [Proposition 8], it clearly suffices to consider the case $N=\{1,2\}$, and to show that $\underline{S}_{\{1,2\}}$ is externally additive. So consider arbitrary f_1 in $\mathscr{L}(\mathscr{X}_1)$ and f_2 in $\mathscr{L}(\mathscr{X}_2)$, then indeed:

$$\begin{aligned}\underline{S}_N(f_1+f_2)&=\inf\{(P_1\times P_2)(f_1+f_2)\colon P_1\in\operatorname{ext}(\mathscr{M}(\underline{P}_1))\text{ and }P_2\in\operatorname{ext}(\mathscr{M}(\underline{P}_2))\}\\&=\inf\{P_1(f_1)+P_2(f_2)\colon P_1\in\operatorname{ext}(\mathscr{M}(\underline{P}_1))\text{ and }P_2\in\operatorname{ext}(\mathscr{M}(\underline{P}_2))\}\\&=\inf\{P_1(f_1)\colon P_1\in\operatorname{ext}(\mathscr{M}(\underline{P}_1))\}+\inf\{P_2(f_2)\colon P_2\in\operatorname{ext}(\mathscr{M}(\underline{P}_2))\}\\&=\underline{P}_1(f_1)+\underline{P}_2(f_2).\end{aligned}$$

We finish by considering the independent natural extension. Here too, because of its marginalisation and associativity properties [Theorems 18 and 23], it clearly suffices to consider the case $N=\{1,2\}$, and to show that $\underline{E}_{\{1,2\}}$ is externally additive. So consider arbitrary f_1 in $\mathscr{L}(\mathscr{X}_1)$ and f_2 in $\mathscr{L}(\mathscr{X}_2)$. Since we know that $\underline{E}_{\{1,2\}}$ is dominated by $\underline{S}_{\{1,2\}}$, we see that $\underline{E}_{\{1,2\}}(f_1+f_2) \leq \underline{S}_{\{1,2\}}(f_1+f_2) = \underline{P}_1(f_1)+\underline{P}_2(f_2)$. To prove the converse inequality, it suffices to use the coherence of $\underline{E}_{\{1,2\}}$ to deduce that

$$\underline{E}_{\{1,2\}}(f_1+f_2) \geq \underline{E}_{\{1,2\}}(f_1)+\underline{E}_{\{1,2\}}(f_2) = \underline{P}_1(f_1)+\underline{P}_2(f_2),$$

where the equality follows from Corollary 15. □

Proof of Theorem 28. We use Corollary 14(ii). Consider arbitrary $o \in N$, $I \subseteq N \setminus \{o\}$, $x_I \in \mathscr{X}_I$ and $g \in \mathscr{L}(\mathscr{X}_o)$. Then, since $\underline{P}_N$ is factorising and has $\mathscr{X}_o$-marginal $\underline{P}_o$:

$$\underline{P}_N(\mathbb{I}_{\{x_I\}}[g-\underline{P}_o(g)]) = \underline{P}_N(\mathbb{I}_{\{x_I\}}\underline{P}_N(g-\underline{P}_o(g))) = \underline{P}_N(0) = 0. \qquad \square$$

Proof of Theorem 29. By Theorem 1, $\underline{P}_N$ is weakly coherent with the family $\mathscr{I}(\underline{P}_N)$ if and only if it is pairwise coherent with each of its members. Let us therefore establish the coherence of $\underline{P}_N$ with each conditional lower prevision $\underline{P}_O(\cdot|X_I)$ (taken separately, for each pair of disjoint subsets I and O of N). Again using Theorem 1, we see we have to show that

$$\underline{P}_N(\mathbb{I}_{\{x_I\}}[f-\underline{P}_O(f|x_I)]) = 0 \text{ for all } f \in \mathscr{L}(\mathscr{X}_O) \text{ and all } x_I \in \mathscr{X}_I.$$

We see that indeed:

$$\underline{P}_N(\mathbb{I}_{\{x_I\}}[f-\underline{P}_O(f|x_I)]) = \underline{P}_N(\mathbb{I}_{\{x_I\}}[f-\underline{P}_N(f)]) = \underline{P}_N(\mathbb{I}_{\{x_I\}}\underline{P}_N(f-\underline{P}_N(f))) = \underline{P}_N(0) = 0,$$

where the first equality follows from the definition of the conditional lower prevision $\underline{P}_O(\cdot|X_I)$, and the second one from the strongly factorising character of $\underline{P}_N$. □

Proof of Proposition 31. To prove (i), we use Corollary 14(ii). Observe that $\underline{Q}_1$, $\underline{Q}_2$ and $\underline{Q}_3$ all have the same marginals $\underline{P}_n$. Consider arbitrary $o \in N$, $I \subseteq N \setminus \{o\}$, $x_I \in \mathscr{X}_I$ and $g \in \mathscr{L}(\mathscr{X}_o)$. Then it follows from the inequalities $\underline{Q}_1 \leq \underline{Q}_3 \leq \underline{Q}_2$ that

$$0 = \underline{Q}_1(\mathbb{I}_{\{x_I\}}[g-\underline{P}_o(g)]) \leq \underline{Q}_3(\mathbb{I}_{\{x_I\}}[g-\underline{P}_o(g)]) \leq \underline{Q}_2(\mathbb{I}_{\{x_I\}}[g-\underline{P}_o(g)]) = 0,$$

where the equalities hold because $\underline{Q}_1$ and $\underline{Q}_2$ are many-to-one independent products. We deduce that $\underline{Q}_3$ is a many-to-one independent product too.

To prove (ii), consider arbitrary $o \in N$, $I \subseteq N \setminus \{o\}$, $f_o \in \mathscr{L}(\mathscr{X}_o)$ and non-negative $f_i \in \mathscr{L}(\mathscr{X}_i)$, $i \in N \setminus \{o\}$. Let, for ease of notation $f_I := \prod_{i \in I} f_i$. Since $\underline{Q}_1$ and $\underline{Q}_2$ are factorising, we deduce from Proposition 5 that $\underline{Q}_1(f_I) = \prod_{i \in I} \underline{P}_i(f_i) = \underline{Q}_2(f_I)$ and $\overline{Q}_1(f_I) = \prod_{i \in I} \overline{P}_i(f_i) = \overline{Q}_2(f_I)$, whence $\underline{Q}_3(f_I) = \prod_{i \in I} \underline{P}_i(f_i)$ and $\overline{Q}_3(f_I) = \prod_{i \in I} \overline{P}_i(f_i)$. Moreover,

$$\overline{\underline{Q}}_1(f_I) \odot \underline{P}_o(f_o) = \underline{Q}_1(f_o f_I) \leq \underline{Q}_3(f_o f_I) \leq \underline{Q}_2(f_o f_I) = \overline{\underline{Q}}_2(f_I) \odot \underline{P}_o(f_o).$$

Hence $\underline{Q}_3(f_o f_I) = \overline{\underline{Q}}_3(f_I) \odot \underline{P}_o(f_o) = \underline{Q}_3(f_I \underline{P}_o(f_o))$. Applying Proposition 5 again, we deduce that $\underline{Q}_3$ is also factorising.

The proof of (iii) is similar to that of (ii). Finally, to prove (iv) use that the inequality $\underline{Q}_3(\sum_{r \in R} f_r) \geq \sum_{r \in R} \underline{P}_r(f_r)$ follows from the super-additivity of the coherent lower prevision $\underline{Q}_3$, and that the converse inequality follows from $\underline{Q}_3 \leq \underline{Q}_2$. □

Proof of Proposition 32. (i)⇒(ii). Assume that $\underline{P}_N$ is weakly coherent with the family $\mathscr{I}(\underline{P}_N)$. Consider any disjoint proper subsets I and O of N, $g \in \mathscr{L}(\mathscr{X}_O)$ and non-negative $f \in \mathscr{L}(\mathscr{X}_I)$. By Theorem 1, it follows from the assumption that $\underline{P}_N$ is coherent with $\underline{P}_{O\cup I}(\cdot|X_I)$, and this implies the equality $\underline{P}_N(\mathbb{I}_{\{x_I\}}[g-\underline{P}_N(g)]) = 0$ for every $x_I \in \mathscr{X}_I$. The coherence [super-additivity] of $\underline{P}_N$ and the non-negativity of f then imply that:

$$\underline{P}_N(f[g-\underline{P}_N(g)]) \geq \sum_{x_I \in \mathscr{X}_I} f(x_I)\underline{P}_N(\mathbb{I}_{\{x_I\}}[g-\underline{P}_N(g)]) = 0.$$

Hence $\underline{P}_N$ is productive.

(ii)⇒(i). Assume $\underline{P}_N$ is productive. Consider any disjoint proper subsets I and O of N, $g \in \mathscr{L}(\mathscr{X}_O)$ and non-negative $f \in \mathscr{L}(\mathscr{X}_I)$. The assumption implies in particular that

$\underline{P}_N(\mathbb{I}_{\{x_I\}}[g-\underline{P}_N(g)]) \geq 0$ for all $x_I \in \mathscr{X}_I$ and all $g \in \mathscr{L}(\mathscr{X}_O)$. If there were some $x_I \in \mathscr{X}_I$ such that $\underline{P}_N(\mathbb{I}_{\{x_I\}}[g-\underline{P}_N(g)]) > 0$, then the coherence of $\underline{P}_N$ would imply that

$$0 = \underline{P}_N(g-\underline{P}_N(g)) = \underline{P}_N\left(\sum_{x_I\in\mathscr{X}_I} \mathbb{I}_{\{x_I\}}[g-\underline{P}_N(g)]\right) \geq \sum_{x_I\in\mathscr{X}_I} \underline{P}_N(\mathbb{I}_{\{x_I\}}[g-\underline{P}_N(g)]) > 0,$$

a contradiction. So we infer from Theorem 1 that $\underline{P}_N$ is weakly coherent with the family $\mathscr{I}(\underline{P}_N)$.

To complete the proof, if $\underline{P}_N$ is many-to-many independent then it is in strongly, and therefore also weakly, coherent with the family $\mathscr{I}(\underline{P}_N)$, and therefore productive as well. On the other hand, the first part of this proposition, together with Corollary 14, shows that if $\underline{P}_N$ is productive it is in particular many-to-one independent. Finally, the last statement is a consequence of Ref. [22, Theorem 11], which shows that all the conditioning events have positive lower probability then weak and strong coherence are equivalent. □

APPENDIX B. COUNTEREXAMPLES

In this Appendix, we have gathered a few examples with additional information on the notions introduced in this paper.

Example 2 (Factorisation properties are not preserved by taking lower envelopes)**.** Let $N := \{1,2\}$ and $\mathscr{X}_1 := \mathscr{X}_2 := \{0,1\}$. Consider the linear marginals P_1 and Q_1 for X_1 defined by $P_1(\{0\}) := P_1(\{1\}) := 1/2$, $Q_1(\{1\}) := 1$ and $Q_1(\{0\}) := 0$. Similarly, consider the linear marginals P_2 and Q_2 for X_2 defined by $P_2(\{0\}) := P_2(\{1\}) := 1/2$, $Q_2(\{1\}) := 1$ and $Q_2(\{0\}) := 0$.

Let $P_{\{1,2\}}$ be the product of P_1 and P_2, and let $Q_{\{1,2\}}$ be the product of Q_1 and Q_2. It follows from Proposition 7 that $P_{\{1,2\}}$ and $Q_{\{1,2\}}$ are factorising and therefore strongly factorising [for $N = \{1,2\}$, these notions are equivalent].

Now let $\underline{P}_{\{1,2\}}$ be the lower envelope of $P_{\{1,2\}}$ and $Q_{\{1,2\}}$. Consider the gamble $f := \mathbb{I}_{\{0\}}$ on $\mathscr{X}_1$ and the gamble $g := \mathbb{I}_{\{0\}} - \mathbb{I}_{\{1\}}$ on $\mathscr{X}_2$. Then $P_1(f) = 1/2$, $Q_1(f) = 0$, $P_2(g) = 0$ and $Q_2(g) = -1$, and therefore

$$\underline{P}_{\{1,2\}}(fg) = \min\{P_{\{1,2\}}(fg), Q_{\{1,2\}}(fg)\} = \min\left\{\frac{1}{2}\cdot 0, 0\cdot -1\right\} = \min\{0,0\} = 0,$$

while $\underline{P}_{\{1,2\}}(g) = \min\{P_2(g), Q_2(g)\} = \min\{0,-1\} = -1$, and

$$\underline{P}_{\{1,2\}}(f\underline{P}_{\{1,2\}}(g)) = \underline{P}_{\{1,2\}}(-f) = \min\{P_1(-f), Q_1(-f)\} = \min\left\{-\frac{1}{2}, 0\right\} = -\frac{1}{2}.$$

Hence $\underline{P}_{\{1,2\}}$ is not (strongly) factorising. ♦

Example 3 (The strong product is not the greatest independent product; many-to-many independent $\not\Rightarrow$ factorising; many-to-many independent $\not\Rightarrow$ externally additive; productive $\not\Rightarrow$ strongly factorising)**.** Consider the possibility spaces $\mathscr{X}_1 = \mathscr{X}_2 = \{0,1\}$, and let $\underline{P}_1$ and $\underline{P}_2$ be the vacuous lower previsions on $\mathscr{L}(\mathscr{X}_1)$ and $\mathscr{L}(\mathscr{X}_2)$, respectively. From the second statement in Proposition 26, the strong product $\underline{S}_{\{1,2\}} := \underline{P}_1 \times \underline{P}_2$ is the vacuous lower prevision on $\mathscr{L}(\mathscr{X}_{\{1,2\}})$, and it coincides with the independent natural extension $\underline{E}_{\{1,2\}} := \underline{P}_1 \otimes \underline{P}_2$.

Let $\underline{Q}_{\{1,2\}}$ be the vacuous lower prevision relative to $\{(0,0),(1,1)\}$. This lower prevision strictly dominates the strong product $\underline{S}_{\{1,2\}}$: we have for instance that

$$\underline{Q}_{\{1,2\}}(\{(0,0),(1,1)\}) = 1 > 0 = \underline{S}_{\{1,2\}}(\{(0,0),(1,1)\}).$$

Yet $\underline{Q}_{\{1,2\}}$ is a many-to-one independent product of the marginals $\underline{P}_1$ and $\underline{P}_2$. [Since $N = \{1,2\}$ it is then also many-to-many independent.] To prove this, use the third statement in Proposition 26. Applying Proposition 32, it is productive too.

The lower prevision $\underline{Q}_{\{1,2\}}$ is not factorising. To see this, consider the non-negative gambles $f := \mathbb{I}_{\{1\}} + \mathbb{I}_{\{0,1\}}$ on $\mathscr{X}_1$ and $g := \mathbb{I}_{\{0\}} + \mathbb{I}_{\{0,1\}}$ on $\mathscr{X}_2$. Then $\underline{Q}_{\{1,2\}}(fg) = \min\{2 \cdot 1, 1 \cdot 2\} = 2$, whereas $\underline{Q}_{\{1,2\}}(f)\underline{Q}_{\{1,2\}}(g) = 1 \cdot 1 = 1$. As a consequence, it is not strongly factorising either.

On the other hand, if we consider the gambles $f_1 := \mathbb{I}_{\{0\}}$ on $\mathscr{X}_1$ and $f_2 := \mathbb{I}_{\{1\}}$ on $\mathscr{X}_2$ we see that $f_1 + f_2 \geq \mathbb{I}_{\{(0,0),(1,1)\}}$ and therefore $\underline{Q}_{\{1,2\}}(f_1 + f_2) = 1 > 0 = \underline{P}_1(f_1) + \underline{P}_2(f_2)$. This shows that not every many-to-many independent product is externally additive. ⧫

Example 4 (Externally additive $\not\Rightarrow$ strongly externally additive; factorising $\not\Rightarrow$ strongly factorising)**.** Let $N := \{1,2,3\}$ and consider the binary variables X_1, X_2 and X_3 assuming values in $\mathscr{X}_1 = \mathscr{X}_2 = \mathscr{X}_3 = \{0,1\}$. Consider the corresponding marginal lower previsions $\underline{P}_1$, $\underline{P}_2$ and $\underline{P}_3$ given by

$$\underline{P}_j(f_j) := \frac{1}{2} f_j(0) + \frac{2}{5} f_j(1) + \frac{1}{10} \min\{f_j(0), f_j(1)\} \text{ for } f_j \in \mathscr{L}(\mathscr{X}_j) \text{ and } j = 1,2,3.$$

Walley [30, Example 9.3.4] has shown that the independent natural extension $\underline{E}_{\{1,2\}}$ of $\underline{P}_1$ and $\underline{P}_2$ is the lower envelope of the set of linear previsions P_k with mass functions p_k, $k = 1, \dots, 6$ given by

k	$p_k(1,1)$	$p_k(1,0)$	$p_k(0,1)$	$p_k(0,0)$
1	$1/4$	$1/4$	$1/4$	$1/4$
2	$1/5$	$1/5$	$3/10$	$3/10$
3	$1/5$	$3/10$	$1/5$	$3/10$
4	$4/25$	$6/25$	$6/25$	$9/25$
5	$2/11$	$3/11$	$3/11$	$3/11$
6	$2/9$	$2/9$	$2/9$	$1/3$.

On the other hand, the strong product $\underline{S}_{\{1,2\}}$ of $\underline{P}_1$ and $\underline{P}_2$ is the lower envelope of the set of linear previsions $\{P_1, P_2, P_3, P_4\}$.

Let P_7 and P_8 be the linear previsions on $\mathscr{X}_3$ whose respective mass functions p_7 and p_8 are determined by $p_7(1) = 2/5$ and $p_8(1) = 1/2$, so $\underline{P}_3$ is the lower envelope of P_7 and P_8. Let $\underline{Q}_{\{1,2,3\}}$ be the lower envelope of the following set of linear previsions on $\mathscr{L}(\mathscr{X}_{\{1,2,3\}})$:

$$\{P_1 \times P_7, P_2 \times P_7, P_3 \times P_7, P_4 \times P_7, P_1 \times P_8, P_2 \times P_8, P_3 \times P_8, P_4 \times P_8, P_5 \times P_8\}.$$

Then $\underline{Q}_{\{1,2,3\}} = \min\{\underline{S}_{\{1,2,3\}}, P_5 \times P_8\}$, where $\underline{S}_{\{1,2,3\}}$ is the strong product of $\underline{P}_1$, $\underline{P}_2$ and $\underline{P}_3$: because of the associativity of the strong product, established in Proposition 8, the strong product $\underline{S}_{\{1,2,3\}}$ is the lower envelope of the set of linear previsions:

$$\{P_1 \times P_7, P_2 \times P_7, P_3 \times P_7, P_4 \times P_7, P_1 \times P_8, P_2 \times P_8, P_3 \times P_8, P_4 \times P_8\}.$$

To see that $\underline{Q}_{\{1,2,3\}}$ is not strongly externally additive, let f be the indicator of the set $\{(0,0),(1,1)\} \times \mathscr{X}_3$ and let g be the indicator of the set $\mathscr{X}_{\{1,2\}} \times \{1\}$. Then

$$\underline{Q}_{\{1,2,3\}}(f) + \underline{Q}_{\{1,2,3\}}(g) = P_5(\{(0,0),(1,1)\}) + P_7(\{1\}) = \frac{5}{11} + \frac{2}{5} = \frac{47}{55},$$

whereas

$$\begin{aligned} \underline{Q}_{\{1,2,3\}}(f+g) &= \min\{\underline{S}_{\{1,2,3\}}(f+g), P_5 \times P_8(f+g)\} \\ &= \min\{\underline{S}_{\{1,2,3\}}(f+g), P_5(\{(0,0),(1,1)\}) + P_8(\{1\})\} \\ &= \min\left\{\frac{1}{2} + \frac{2}{5}, \frac{5}{11} + \frac{1}{2}\right\} = \frac{9}{10} > \frac{47}{55}. \end{aligned}$$

Let us show that $\underline{Q}_{\{1,2,3\}}$ is not productive, from which it follows, taking into account Proposition 6, that it is not strongly factorising either. Let f be minus the indicator of the set $\{1\} \times \mathscr{X}_{\{2,3\}}$ and g minus the indicator of the set $\mathscr{X}_1 \times \{1\} \times \mathscr{X}_3$. Then it follows that

$$\begin{aligned}\underline{Q}_{\{1,2,3\}}(g) &= \min\{\underline{S}_{\{1,2,3\}}(g), P_5 \times P_8(g)\}\\ &= \min\{-\overline{S}_{\{1,2\}}(\mathscr{X}_1 \times \{1\}), -P_5(\mathscr{X}_1 \times \{1\})\}\\ &= \min\left\{-\frac{1}{2}, -\frac{5}{11}\right\} = -\frac{1}{2},\end{aligned}$$

and as a consequence

$$\begin{aligned}\underline{Q}_{\{1,2,3\}}(f[g-\underline{Q}_{\{1,2,3\}}(g)]) &\leq P_5 \times P_8(f[g-\underline{Q}_{\{1,2,3\}}(g)])\\ &= P_5(\{(1,1)\}) - \frac{1}{2}P_5(\{1\} \times \mathscr{X}_2)\\ &= \frac{2}{11} - \frac{1}{2}\frac{5}{11} = -\frac{1}{22} < 0.\end{aligned}$$

We now show that, nevertheless, $\underline{Q}_{\{1,2,3\}}$ is externally additive. For this, use that $\underline{Q}_{\{1,2,3\}}$ is dominated by $\underline{S}_{\{1,2,3\}}$ and that both these coherent lower previsions have $\underline{P}_1$, $\underline{P}_2$ and $\underline{P}_3$ as their marginals, and apply Proposition 31(iv).

Now, the associativity of the independent natural extension [Theorem 23] and its being dominated by the strong product imply that

$$\underline{E}_{\{1,2,3\}} \leq \underline{E}_{\{1,2\}} \otimes \underline{P}_3 \leq \underline{E}_{\{1,2\}} \times \underline{P}_3 = \min\left\{P_i \times P_j \colon i = 1,\ldots,6, j = 7,8\right\} \leq \underline{Q}_{\{1,2,3\}}.$$

Applying Proposition 31(i), we deduce that $\underline{Q}_{\{1,2,3\}}$ is a many-to-one independent product, and from Proposition 31(ii) we infer that it is factorising. ⧫

Example 5 (Factorising $\not\Rightarrow$ Kuznetsov; strongly factorising $\not\Rightarrow$ strongly Kuznetsov)**.** Consider two binary variables X_1, X_2 assuming values in the set $\mathscr{X}_1 = \mathscr{X}_2 = \{0,1\}$. Let $\underline{P}_1, \underline{P}_2$ be the marginal lower previsions from Example 4. Consider the gambles $f := \mathbb{I}_{\{0\}} - \mathbb{I}_{\{1\}}$ on $\mathscr{X}_1$ and $g := \mathbb{I}_{\{0\}} - \mathbb{I}_{\{1\}}$ on $\mathscr{X}_2$. Then $\underline{P}_1(f) = \underline{P}_2(g) = 0$ and $\overline{P}_1(f) = \overline{P}_2(g) = 1/5$. As a consequence, $\underline{\overline{P}}_1(f) \boxtimes \underline{\overline{P}}_2(g) = [0, 1/25]$, whereas, considering the linear previsions $P_1, \ldots, P_6$ in that example, we see that their independent natural extension provides the following value for the lower bound:

$$\underline{E}_{\{1,2\}}(fg) = \min\left\{0,0,0,\frac{1}{25},-\frac{1}{11},\frac{1}{9}\right\} = -\frac{1}{11}.$$

This shows that the independent natural extension $\underline{E}_{\{1,2\}}$, which is factorising by Theorem 22, is not Kuznetsov. Moreover, in this example where $N = \{1,2\}$, factorisation is equivalent to strong factorisation, and being Kuznetsov is equivalent to being strongly Kuznetsov. ⧫

Example 6 (Many-to-one independent $\not\Rightarrow$ many-to-many independent; factorising $\not\Rightarrow$ strongly factorising)**.** Let $N := \{1,2,3\}$ and consider the binary variables X_1, X_2 and X_3 assuming values in $\mathscr{X}_1 = \mathscr{X}_2 = \mathscr{X}_3 = \{0,1\}$, and consider the corresponding marginal lower previsions $\underline{P}_1$, $\underline{P}_2$ and $\underline{P}_3$ given by

$$\underline{P}_j(f_j) := \frac{1}{2}f_j(0) + \frac{2}{5}f_j(1) + \frac{1}{10}\min\{f_j(0), f_j(1)\} \text{ for all } f_j \in \mathscr{L}(\mathscr{X}_j)$$

for $j = 1,2,3$. Let $\underline{E}_{\{1,2,3\}}$ denote their independent natural extension and $\underline{S}_{\{1,2,3\}}$ their strong product.

Define the coherent lower prevision $\underline{Q}_{\{1,2,3\}}$ on $\mathscr{L}(\mathscr{X}_{\{1,2,3\}})$ as the convex mixture $\underline{Q}_{\{1,2,3\}} := 1/2(\underline{E}_{\{1,2,3\}} + \underline{S}_{\{1,2,3\}})$. It follows from Proposition 31 that $\underline{Q}_{\{1,2,3\}}$ is factorising, and a many-to-one independent product. We are going to prove that $\underline{Q}_{\{1,2,3\}}$ is not a many-to-many independent product. It will then follow from Theorem 29 that it is not strongly factorising either.

Consider the conditional lower prevision $\underline{Q}_{\{1,2,3\}}(\cdot|X_3)$ derived from the joint lower prevision $\underline{Q}_{\{1,2,3\}}$ using the epistemic irrelevance of X_3 to $X_{\{1,2\}}$:

$$\underline{Q}_{\{1,2,3\}}(f|x_3) := \underline{Q}_{\{1,2,3\}}(f(\cdot,x_3)) \text{ for all } x_3 \in \mathscr{X}_3 \text{ and all } f \in \mathscr{L}(\mathscr{X}_{\{1,2,3\}}).$$

In order to show that it $\underline{Q}_{\{1,2,3\}}$ is not a many-to-many independent product, it suffices to show that it is not weakly coherent with this conditional lower prevision $\underline{Q}_{\{1,2,3\}}(\cdot|X_3)$. Consider the event $A := \{(0,0),(1,1)\}$ that $X_1 = X_2$, and the corresponding indicator $g := \mathbb{I}_A$ on $\mathscr{X}_{\{1,2\}}$. It follows from Ref. [30, Example 9.3.4] that $\underline{E}_{\{1,2,3\}}(A) = 5/11$ and $\underline{S}_{\{1,2,3\}}(A) = 1/2$, so

$$\underline{Q}_{\{1,2,3\}}(A) = \frac{1}{2}(\underline{E}_{\{1,2,3\}}(A) + \underline{S}_{\{1,2,3\}}(A)) = \frac{1}{2}\left(\frac{1}{2} + \frac{5}{11}\right) = \frac{21}{44}.$$

Let $x_3 = 0$. Since both $\underline{E}_{\{1,2,3\}}$ and $\underline{S}_{\{1,2,3\}}$ are strongly factorising [by Theorem 24 and Proposition 8(iv), respectively], we see that

$$\begin{aligned}\underline{E}_{\{1,2,3\}}(\mathbb{I}_{\{x_3\}}[g - \underline{Q}_{\{1,2,3\}}(g)]) &= \overline{P}_3(\{0\})\underline{E}_{\{1,2,3\}}(g - \underline{Q}_{\{1,2,3\}}(g)) \\ &= \frac{3}{5}\left(\frac{5}{11} - \frac{21}{44}\right) = -\frac{3}{220},\end{aligned}$$

whereas

$$\underline{S}_{\{1,2,3\}}(\mathbb{I}_{\{x_3\}}[g - \underline{Q}_{\{1,2,3\}}(g)]) = \underline{P}_3(\{0\})\underline{S}_{\{1,2,3\}}(g - \underline{Q}_{\{1,2,3\}}(g)) = \frac{1}{2}\left(\frac{1}{2} - \frac{21}{44}\right) = \frac{1}{88}.$$

As a consequence, we deduce that

$$\underline{Q}_{\{1,2,3\}}(\mathbb{I}_{\{x_3\}}[g - \underline{Q}_{\{1,2,3\}}(g)]) = \frac{1}{2}\left(-\frac{3}{220} + \frac{1}{88}\right) = -\frac{1}{880} < 0,$$

so $\underline{Q}_{\{1,2,3\}}$ is not coherent with $\underline{Q}_{\{1,2,3\}}(\cdot|X_3)$. This also shows that $\underline{Q}_{\{1,2,3\}}$ is not productive, applying Proposition 32.

Finally, note that, from Example 5 we can deduce that the independent natural extension $\underline{E}_{\{1,2,3\}}$ is not Kuznetsov, and since $\underline{Q}_{\{1,2,3\}}$ has the same marginals we deduce that it is not Kuznetsov either: it suffices to take the same gambles f and g from that example. ♦

This example shows that if we consider a many-to-one independent product $\underline{Q}_N$ of some given marginals $\underline{P}_n$, $n \in N$, and a partition of N given by sets R and S, then $\underline{Q}_N$ need not be coherent with the conditional lower previsions $\underline{Q}_{R\cup S}(\cdot|X_S)$ and $\underline{Q}_{R\cup S}(\cdot|X_R)$. In this sense the associativity properties satisfied by the strong product and the independent natural extension do not extend towards their convex combinations.

GHENT UNIVERSITY, SYSTEMS RESEARCH GROUP, TECHNOLOGIEPARK - ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM

E-mail address: gert.decooman@ugent.be

UNIVERSITY OF OVIEDO, DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH, C-CALVO SOTELO, S/N, 33007 OVIEDO, SPAIN

E-mail address: mirandaenrique@uniovi.es

IDSIA, GALLERIA 2, CH-6928 MANNO (LUGANO), SWITZERLAND

E-mail address: zaffalon@idsia.ch