# Independent products in infinite spaces

Enrique Miranda[a,*], Marco Zaffalon[b]

[a] *University of Oviedo, Department of Statistics and Operations Research, C-Calvo Sotelo, s/n, 33007 Oviedo, Spain*
[b] *Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA)*
*Galleria 2, 6928 Manno (Lugano), Switzerland*

**Abstract**

Probabilistic independence, intended as the mutual irrelevance of given variables, can be solidly founded on a notion of self-consistency of an uncertainty model, in particular when probabilities go imprecise. There is nothing in this approach that prevents it from being adopted in very general setups, and yet it has mostly been detailed for variables taking finitely many values. In this mathematical study, we complement previous research by exploring the extent to which such an approach can be generalised. We focus in particular on the independent products of two variables. We characterise the main notions, including some of factorisation and productivity, in the general case where both spaces can be infinite and show that, however, there are situations—even in the case of precise probability—where no independent product exists. This is not the case as soon as at least one space is finite. We study in depth this case at the frontiers of well-behaviour detailing the relations among the most important notions; we show for instance that being an independent product is equivalent to a certain productivity condition. Then we step back to the general case: we give conditions for the existence of independent products and study ways to get around its inherent limitations.

*Keywords:* Independence, irrelevance, imprecise probability, coherence, infinite spaces, factorisation.

## 1. Introduction

Independence is a founding notion for very many probabilistic models and applications. Despite its widespread use, there are substantial aspects of this notion that are still troublesome, in particular when we stick to the traditional approach to defining independence—the one based on requiring that a joint probability factorises. To make an example, if we say that events $A$ and $B$ are independent when $P(A \cap B) = P(A)P(B)$, what we get is that $A$ is independent of any other event $B$ if $P(A) = 0$, *including the event* $B = A^c$! This happens because the traditional definition neglects the issues originated by zero probabilities. This affects the case of finite spaces of possibility and even more the models based on infinite spaces, where it is very common that each element in the space has zero probability. The situation becomes more complex when we consider models that allow for imprecisely specified probabilities. In this case an event $A$ has both a lower probability $\underline{P}(A)$ and an upper probability $\overline{P}(A)$, so it can happen that $\underline{P}(A) = 0 < \overline{P}(A)$. It is clear that in this case we should have a uniform way to deal with independence that works irrespective of the positivity of probabilities. The imprecise case poses a series of other challenges as well: it has been shown that there are many possible definitions of independence in such a generalised setup [2, 7]; very often imprecise probability models are made of sets of finitely additive probabilities, which create additional complications w.r.t. the more regular countably additive probabilities.

It has been Walley to illustrate [24, Chapter 9] that all these issues can be nicely and uniformly addressed by a shift of paradigm in the way we define independence. Walley's approach is based on joining two pre-existing ideas. The first, which has its roots in the subjective approach to probability, as well as in the artificial intelligence community, is regarding independence as the mutual irrelevance of two events (or variables). The second, especially due to de Finetti, is that probabilistic models can be founded on a notion of self-consistency—most often called *coherence*. In this paper

*Corresponding author
*Email addresses:* `mirandaenrique@uniovi.es` (Enrique Miranda), `zaffalon@idsia.ch` (Marco Zaffalon)

we shall, for the most part, refer to Walley's coherence notion [24, Section 7.1.4(b)]. We can think of coherence as a stronger way to define probabilistic models than through a joint distribution. In fact, the existence of a joint distribution compatible with some marginal and conditional distributions, can equivalently be formulated as a self-consistency requirement, which is however weaker than coherence (see, e.g., [17, Section 4]). This weakness shows up dramatically just when there are events with zero probability, when we work in infinite spaces, when we deal with finitely additive probabilities, and of course also when we deal with imprecise probability. This is the reason why the stronger notion of coherence is at the basis of a more powerful approach to defining indepedence.

So how do we formulate independence through coherence? Consider variables $X_1$ and $X_2$, taking values from $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. We just say that a joint probabilistic model for these variables is an *independent product* of the marginal information we have about $X_1$ and $X_2$ if it is coherent with our assessment that knowledge of one of the two variables does not affect knowledge about the other variable. It is an independent product, in other words, if it is not inconsistent with two assessments of irrelevance: the irrelevance of $X_1$ to $X_2$ and of $X_2$ to $X_1$. Given that coherence is a notion defined under very general assumptions, we have automatically a well-posed way to discuss the notion of independence across all the situations mentioned above. As we said, approaching independence from the point of view of coherence has its roots in Walley's seminal book, which laid down the main ideas. Vicig [23] then pursued similar ideas using a weaker notion of coherence by Williams [25] (we call it *W-coherence*); his work addressed the case of infinite spaces of possibility, while restricting the attention to the special case of lower probabilities rather than expectations (or, as we call them, *previsions*). Finally, de Cooman and ourselves [10, 12] have studied the case of finite spaces of possibility, while allowing for more than two variables; the work focused in particular on the least-committal (or least precise, or weakest) independent product, which is called the *independent natural extension*, and on ways to relate it to different forms of factorisation. The independent natural extension is an important concept as it is the only independent product that is solely based on the mutual irrelevance of the variables under consideration.

In this paper, we analyse the independent products of two variables aiming at the greatest possible generality—thus covering the case of infinite spaces of possibility—as well as at establishing firm relations with known notions of factorisation—thus investigating the extent to which the traditional notion of independence and the notion based on coherence are compatible. After some preliminary concepts are given in Section 2, we start work on independent products in Section 3. First, we remark by an example that two given marginals may not admit *any* independent product when both spaces $\mathcal{X}_1, \mathcal{X}_2$ are infinite, not even in the case of precise probability: this shows that there are limits to the possibility to define independence in the general case. As a consequence, the independent natural extension may not exist either, unlike the case of finite spaces [12]. Yet, when it exists, we show that it can be characterised as the intersection of the two sets of probabilities that express irrelevance of $X_1$ to $X_2$ and of $X_2$ to $X_1$: this is a remarkably simple form of the independent natural extension that we expect to lead to a substantial increase of its use in practice. We also recall two main factorisation properties of an uncertainty model in the case of imprecise probabilities: one indeed called *factorisation* and a weaker one called *productivity*; we give new and simpler formulations of this second property, and relate it to the generalisation of Bayes' rule for the imprecise case (which is referred to as GBR for short).

Then we try to understand what is the most general setup where independent products are always well defined. We show in Section 4 that for this to be the case it is enough that one of the spaces $\mathcal{X}_1, \mathcal{X}_2$ is finite. It turns out that this case also delimits the boundary of well-behaviour: we show, for instance, that in this case the independent products are characterised as the (closed and convex) sets of probabilities that are included in the one that defines the independent natural extension; moreover, that productivity and being an independent product are equivalent notions. We show also that among them there is the *strong product*, which is a well-known extension of probabilistic independence to imprecise probability obtained through sets of precise-probability (stochastically) independent models. This nicely relates to the factorisation condition, as we show that any model not weaker than the independent natural extension and not stronger than the strong product is factorising. While doing this, we give simple sufficient conditions for the equality of the independent natural extension and the strong product, which are also equivalent conditions for the former to coincide with one of the models that express irrelevance of $X_1$ to $X_2$ or vice versa. Finally, we show that when both $\mathcal{X}_1$ and $\mathcal{X}_2$ are finite, the strong product is the most precise independent product that factorises. This can easily be extended to any finite number of variables, thus closing a few open problems from [12].

In Section 5 we step back to the general case where both $\mathcal{X}_1, \mathcal{X}_2$ can be infinite and propose two alternative paths to deal with it. In the first, we give sufficient conditions for the existence of an independent product. In the second, we consider weakening Walley's coherence notion so as to allow for a (weaker) notion of independence to exist in general. In particular, motivated by Vicig's work [23], we turn to Williams' coherence notion, which we reformulate in Walley's framework: we call it W-coherence. We show that W-independent products always exist and are those that upon conditioning by GBR yield models that express the irrelevance of $X_1$ to $X_2$ and of $X_2$ to $X_1$. On the other hand, we give a number of examples showing that W-independent products appear to be unreasonable because of the weakness of W-coherence: for instance, any precise model that assigns zero probability to the singletons is a W-independent product of every pair of marginals. We thus stop focusing on W-coherence and rather consider *irrelevant products*, that is, models that are coherent with just one irrelevance assessment (of $X_1$ to $X_2$ or vice versa, but not both), which are actually enough for quite some useful applications (e.g., [1, 8]). These products always exist and are given a simple characterisation. Therefore they can as well be considered an alternative avenue, and one that is close to the best we can do, in case independent products do not exist.

In order to improve readibility, all the proofs and a few more technical results have been relegated to Appendix A and Appendix B. The latter, in particular, considers what happens to independent products in case we use a consistency notion stronger than Walley's: that of *conglomerable coherence*, which we have proposed in [18]. The appendix shows that, under relatively weak assumptions, both this and Walley's coherence notion originate the same independent products, so that the study carried out in this paper under Walley's notion can be largely regarded as addressing the more stringent notion too.

## 2. Coherent lower previsions

Given a possibility space $\Omega$, a *gamble* is a bounded real-valued function $f : \Omega \to \mathbb{R}$. We shall denote by $\mathcal{L}(\Omega)$ the set of all gambles on $\Omega$, and by $\mathcal{L}^+(\Omega) := \{f \in \mathcal{L}(\Omega) : f \geq 0, f \neq 0\}$ the set of all positive gambles.

*Definition* 1 (**Coherent lower previsions**). Consider a possibility space $\Omega$. A lower prevision is a functional $\underline{P} : \mathcal{K} \to \mathbb{R}$ on some subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$. If $\mathcal{K}$ is linear, $\underline{P}$ is called *coherent* when:

- $\underline{P}(f) \geq \inf_{\omega\in\Omega} f(\omega)$;
- $\underline{P}(\lambda f) = \lambda \underline{P}(f)$;
- $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$,

for any $f, g \in \mathcal{K}$ and any $\lambda > 0$. When $\mathcal{K} = \mathcal{L}(\Omega)$ and the third condition holds with equality for every $f, g \in \mathcal{L}(\Omega)$, $\underline{P}$ is called a *linear prevision* and is denoted by $P$.

The behavioural interpretation of $\underline{P}(f)$ is that of the supremum acceptable buying price for the gamble $f$. Given a lower prevision $\underline{P}$ with domain $\mathcal{K}$, its associated *credal set* is given by $\mathcal{M}(\underline{P}) := \{P : P(f) \geq \underline{P}(f)\ \forall f \in \mathcal{K}\}$. Then it holds that $\underline{P}$ is coherent if and only if it is the lower envelope of its associated credal set, that is, iff $\underline{P}(f) = \min\{P(f) : P \in \mathcal{M}(\underline{P})\}\ \forall f \in \mathcal{K}$. Moreover, $\mathcal{M}(\underline{P})$ is compact in the weak-* topology and convex, and a one-to-one correspondence can be established between coherent lower previsions and weak-* compact and convex sets of linear previsions. The conjugate of a coherent lower prevision, given by $\overline{P}(f) := -\underline{P}(-f)$ for any gamble $f$, is called a coherent *upper* prevision. In the case of events, it holds that $\overline{P}(A) = 1 - \underline{P}(A^c)\ \forall A \subseteq \Omega$.

One interesting instance of coherent lower previsions are the *vacuous* ones, which are those such that $\underline{P}(f) = \inf f$ for every gamble $f$, and that model complete ignorance. At the other side of the spectrum we find linear previsions, which correspond to the absence of imprecision.

Conditional lower previsions are defined in a way analogous to the unconditional case:

*Definition* 2 (**Conditional lower previsions**). Consider a possibility space $\Omega$, and let $\mathcal{B}$ be a partition of $\Omega$. A *separately coherent conditional lower prevision* $\underline{P}(\cdot|\mathcal{B})$ is a functional from $\mathcal{L}(\Omega) \times \mathcal{B}$ to $\mathbb{R}$ satisfying the following conditions for every $f, g \in \mathcal{L}(\Omega), B \in \mathcal{B}$ and every $\lambda > 0$:

- $\underline{P}(f|B) \geq \inf_{\omega\in B} f(\omega)$ (avoiding sure loss).
- $\underline{P}(\lambda f|B) = \lambda \underline{P}(f|B)$ (positive homogeneity).
- $\underline{P}(f + g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$ (super-additivity).

Given a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$, the *support* of a gamble $f$ is given by $S(f) := \{B \in \mathcal{B} : Bf \neq 0\}$.

Note that an unconditional lower prevision $\underline{P}$ corresponds to the particular case where the partition is $\mathcal{B} = \{\Omega\}$. Given a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ and a gamble $f$, we shall use the notations $G(f|B) := I_B(f - \underline{P}(f|B))$ and $G(f|\mathcal{B}) := \sum_{B \in \mathcal{B}} G(f|B) = \sum_{B \in \mathcal{B}} I_B(f - \underline{P}(f|B))$, where $I_B$ denotes the indicator function of set $B$. Given a partition $\mathcal{B}$ of $\Omega$, a gamble is called *$\mathcal{B}$-measurable* when it is constant on the elements of $\mathcal{B}$. One instance is the gamble $\underline{P}(f|\mathcal{B})$, which takes the value $\underline{P}(f|B)$ on the elements of $B$; then we can also express $G(f|\mathcal{B}) = f - \underline{P}(f|\mathcal{B})$.

We shall use the following consequence of separate coherence in the proofs [24, Theorem 6.2.6($\ell$)]:

$$\underline{P}(fg|\mathcal{B}) = f\underline{P}(g|\mathcal{B}) \text{ for every } \mathcal{B}\text{-measurable } f \in \mathcal{L}^+(\Omega), \text{ and every } g \in \mathcal{L}(\Omega). \tag{1}$$

In this paper, we shall focus on the particular case where $\Omega = \mathcal{X}_1 \times \mathcal{X}_2$ and we condition on the partitions $\{\{x_1\} \times \mathcal{X}_2 : x_1 \in \mathcal{X}_1\}, \{\mathcal{X}_1 \times \{x_2\} : x_2 \in \mathcal{X}_2\}$ of $\Omega$. To simplify the notation, we shall use $\underline{P}(f|x_1), \underline{P}(f|x_2)$ to denote $\underline{P}(f|\{x_1\} \times \mathcal{X}_2), \underline{P}(f|\mathcal{X}_1 \times \{x_2\})$, respectively. Similarly, $\underline{P}(\cdot|\mathcal{X}_1), \underline{P}(\cdot|\mathcal{X}_2)$ will refer to the lower previsions conditional on the partitions $\{\{x_1\} \times \mathcal{X}_2 : x_1 \in \mathcal{X}_1\}, \{\mathcal{X}_1 \times \{x_2\} : x_2 \in \mathcal{X}_2\}$ of $\mathcal{X}_1 \times \mathcal{X}_2$. We shall say that a gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$ is *$\mathcal{X}_1$-measurable* when it is constant on the elements of the partition $\{\{x_1\} \times \mathcal{X}_2 : x_1 \in \mathcal{X}_1\}$, that is, when $f(x_1, x_2) = f(x_1, x_2') \; \forall x_1 \in \mathcal{X}_1, x_2, x_2' \in \mathcal{X}_2$; the notion of $\mathcal{X}_2$-measurable gamble is defined similarly. Note that the set of $\mathcal{X}_1$-measurable gambles is in a one-to-one correspondence with $\mathcal{L}(\mathcal{X}_1)$, and similarly the set of $\mathcal{X}_2$-measurable gambles is in a one-to-one correspondence with $\mathcal{L}(\mathcal{X}_2)$. Then if we have a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, its *marginals* $\underline{P}_1, \underline{P}_2$ are its restrictions to the sets of $\mathcal{X}_1, \mathcal{X}_2$-measurable gambles, respectively. Taking the previous comment into account, we shall sometimes say that they are defined on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$.

We shall consider a number of coherence notions for conditional lower previsions. They are defined in the following way.

*Definition* 3 (**Coherence**). Consider separately coherent conditional lower previsions $\underline{P}_1(\cdot|\mathcal{B}_1), \dots, \underline{P}_m(\cdot|\mathcal{B}_m)$ on $\mathcal{L}(\Omega)$. They are called:

- *coherent* when for every $f_0, f_1, \dots, f_m \in \mathcal{L}(\Omega), j \in \{1, \dots, m\}$ and $B_0 \in \mathcal{B}_j$, there is some $B \in \cup_{i=1}^m S_i(f_i) \cup \{B_0\}$ such that
$$\sup_{\omega \in B} \left[ \sum_{i=1}^m G_i(f_i|\mathcal{B}_i) - G(f_0|B_0) \right](\omega) \geq 0; \tag{2}$$

- *weakly coherent* when for every $f_0, f_1, \dots, f_m \in \mathcal{L}(\Omega), j \in \{1, \dots, m\}$ and $B_0 \in \mathcal{B}_j$, it holds that
$$\sup_{\omega \in \Omega} \left[ \sum_{i=1}^m G_i(f_i|\mathcal{B}_i) - G(f_0|B_0) \right](\omega) \geq 0; \tag{3}$$

- *W-coherent*[1] when Eq. (2) holds for every $f_0, f_1, \dots, f_m \in \mathcal{L}(\Omega), j \in \{1, \dots, m\}$ and $B_0 \in \mathcal{B}_j$ such that $S_i(f_i)$ is finite for every $i = 1, \dots, m$;

- *weakly W-coherent* when Eq. (3) holds for every $f_0, f_1, \dots, f_m \in \mathcal{L}(\Omega), j \in \{1, \dots, m\}$ and $B_0 \in \mathcal{B}_j$ such that $S_i(f_i)$ is finite for every $i = 1, \dots, m$.

One particular case of interest in this paper is that where we study the coherence of one conditional and one unconditional lower prevision:

**Proposition 1.** *[24, Theorem 6.5.3] Let $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ be a coherent lower prevision and a separately coherent conditional lower prevision on $\mathcal{L}(\Omega)$. They are coherent if and only if*

$$\underline{P}(G(f|B)) = 0 \; \forall f \in \mathcal{L}(\Omega), B \in \mathcal{B}; \tag{GBR}$$
$$\underline{P}(G(f|\mathcal{B})) \geq 0 \; \forall f \in \mathcal{L}(\Omega). \tag{CNG}$$

*In the particular case where $\mathcal{B}$ is finite, $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are coherent if and only they satisfy generalised Bayes rule (GBR): in fact, (CNG), which is a condition of so-called conglomerability, follows from (GBR) in that case.*

[1] This is a restricted version of Williams' coherence obtained when a lower prevision is only allowed to be conditional on a partition of $\Omega$. This is not required in Williams' original formulation; we use such a restricted version in this paper in order to compare it with Walley's theory, which is instead based on conditioning partitions. See Remark 1 in Section 5 for a discussion about this point with regard to independence.

One particular instance we shall use repeatedly in the proofs is the following:

**Proposition 2.** *[24, Sections 6.7.2 and 6.7.3] Let $\underline{P}(\cdot|\mathcal{B})$ be a separately coherent lower prevision on $\mathcal{L}(\Omega)$, and let $\underline{P}$ be a coherent lower prevision on the set of $\mathcal{B}$-measurable gambles.*

*(a) The smallest coherent lower prevision on $\mathcal{L}(\Omega)$ that is coherent with $\underline{P}(\cdot|\mathcal{B})$ and coincides with $\underline{P}$ on $\mathcal{B}$-measurable gambles is given by $\underline{P}(\underline{P}(\cdot|\mathcal{B}))$.*

*(b) If $\underline{P}(\cdot|\mathcal{B}) = P(\cdot|\mathcal{B})$ is linear, then $\underline{P}(P(\cdot|\mathcal{B}))$ is the only coherent lower prevision that is coherent with $P(\cdot|\mathcal{B})$ and coincides with $\underline{P}$ on $\mathcal{B}$-measurable gambles.*

## 3. Independent products

Consider two variables $X_1, X_2$ taking values in respective spaces $\mathcal{X}_1, \mathcal{X}_2$. We shall assume throughout that $X_1, X_2$ are *logically independent* [14], meaning that their joint $(X_1, X_2)$ can assume any value in the product space $\mathcal{X}_1 \times \mathcal{X}_2$. Let $\underline{P}_1, \underline{P}_2$ be the coherent lower previsions on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$ that model the *marginal* information about the variables $X_1, X_2$. Our goal in this paper is to establish how to combine these two marginal lower previsions into a joint coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, under an assumption of independence between the two variables.

There are several ways in which independence can be modelled when we consider sets of probabilities, or coherent lower previsions [2, 7]. In this paper, we shall found the notion of independence on the assessment of mutual irrelevance between two variables, which in our opinion is the most sound from the point of view of a behavioural theory of probability. Consider two possibility spaces $\mathcal{X}_1, \mathcal{X}_2$, and let $\underline{P}_1, \underline{P}_2$ be two coherent lower previsions with respective domains $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$. Define the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ by expressing their mutual irrelevance, which is to say that[2]

$$\underline{P}_1(f|x_2) := \underline{P}_1(f(\cdot, x_2)) \text{ and } \underline{P}_2(f|x_1) := \underline{P}_2(f(x_1, \cdot)) \; \forall f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2), x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2. \tag{4}$$

The behavioural interpretation of these conditional lower previsions is that the supremum buying price for a gamble $f$, conditional on the observation of $x_2 \in \mathcal{X}_2$, coincides with the unconditional supremum buying price for the gamble $f(\cdot, x_2)$ (and similarly if we condition on the observation of the value in $\mathcal{X}_1$). In other words, we are expressing that a subject's beliefs do not change with observations (see [5, 6, 8] for applications of this way to define independence). Note also that the supports of a gamble $f$ with respect to the conditional lower previsions above are given by $S_1(f) := \{x_2 \in \mathcal{X}_2 : f(\cdot, x_2) \neq 0\}$ and $S_2(f) := \{x_1 \in \mathcal{X}_1 : f(x_1, \cdot) \neq 0\}$.

*Definition* 4 (**Independent product**). A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ is said to be an *independent product* (of its marginals $\underline{P}_1, \underline{P}_2$) if $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent.

Since coherence is preserved by lower envelopes [24, Theorem 7.1.6], we can give the following notion:

*Definition* 5 (**Independent natural extension**). Given coherent lower previsions $\underline{P}_1, \underline{P}_2$ on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$, we call their *independent natural extension* $\underline{P}_1 \otimes \underline{P}_2$ the smallest independent product of these marginals.

The first thing we should remark is that the given marginals $\underline{P}_1, \underline{P}_2$ may not possess any independent product, and as a consequence the independent natural extension may not exist:[3]

*Example* 1. Consider $\mathcal{X}_1 := \mathbb{N}$ and $\mathcal{X}_2 := -\mathbb{N}$. Let $P_1, P_2$ be two linear previsions with respective domains $\mathcal{L}(\mathbb{N}), \mathcal{L}(-\mathbb{N})$ such that their restrictions to events satisfy $P_1(\{n\}) = P_2(-\{n\}) = 0$ for every natural number $n$. It follows from Proposition 2(b) that the only coherent lower prevision with marginal $P_2$ that is coherent with $P_1(\cdot|-\mathbb{N})$ is $P_2(P_1(\cdot|-\mathbb{N}))$. The same proposition implies that the only coherent lower prevision with marginal $P_1$ that is coherent with $P_2(\cdot|\mathbb{N})$ is given by $P_1(P_2(\cdot|\mathbb{N}))$. Thus, in order to show that $P_1, P_2$ do not have any independent product it suffices to show that $P_1(P_2(\cdot|\mathbb{N}))$ and $P_2(P_1(\cdot|-\mathbb{N}))$ do not coincide.

[2]Recall that Walley's theory of coherence, which is the one we are considering in this paper, requires the conditional and unconditional lower previsions to be defined on gambles over the same space; in this case, the domain of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ and $\underline{P}_2(\cdot|\mathcal{X}_1)$ is $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, even if $\underline{P}_1(\cdot|\mathcal{X}_2)$ and $\underline{P}_2(\cdot|\mathcal{X}_1)$ are derived from the coherent lower previsions $\underline{P}_1, \underline{P}_2$ that have respective domains $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$, in the manner described below. See again the comments after Eq. (1).

[3]This example is based on the Cantelli-Lévy paradox discussed by Walley in [24, Section 7.3.4].

Consider the set $A := \{(n,-m) \in \mathbb{N} \times -\mathbb{N} : n - m \leq 0\}$. Then $P_1(P_2(A|\mathbb{N})) = P_1(1) = 1$, since for every natural number $n$,

$$P_2(A|n) = P_2(\{-m : (n,-m) \in A\}) = P_2(\{-m : m \geq n\}) = 1.$$

On the other hand, $P_2(P_1(A|-\mathbb{N})) = P_2(0) = 0$, since for every natural number $m$,

$$P_1(A|-m) = P_1(\{n : (n,-m) \in A\}) = P_1(\{1,\dots,m\}) = 0.$$

We conclude therefore that $P_1, P_2$ do not have an independent product. ♦

Next, we take advantage of the reduction theorem in Walley [24, Theorem 7.1.5], which shows that the coherence of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ can be decomposed into two subproblems: the coherence of $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ and the weak coherence of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$. The latter is characterised in the following result, which also gives a new characterisation of the independent natural extension. Here, and in what follows, we shall use the simplified notations

$$\underline{P}_1(\underline{P}_2) := \underline{P}_1(\underline{P}_2(\cdot|\mathcal{X}_1)) \text{ and } \underline{P}_2(\underline{P}_1) := \underline{P}_2(\underline{P}_1(\cdot|\mathcal{X}_2)).$$

This also helps to stress that the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are derived from $\underline{P}_1, \underline{P}_2$ by means of an assessment of irrelevance.

**Theorem 3.** *Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with marginals $\underline{P}_1, \underline{P}_2$.*

*(a) $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ are coherent if and only if $\underline{P} \geq \underline{P}_2(\underline{P}_1)$.*

*(b) $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly coherent if and only if $\underline{P} \geq \max\{\underline{P}_2(\underline{P}_1), \underline{P}_1(\underline{P}_2)\}$.*

*(c) If $\underline{P}_1, \underline{P}_2$ have an independent product, the smallest one is the smallest coherent lower prevision that dominates both $\underline{P}_1(\underline{P}_2)$ and $\underline{P}_2(\underline{P}_1)$. In other words, $\mathcal{M}(\underline{P}_1 \otimes \underline{P}_2) = \mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1))$.*

A formal study of independence was made in [12], in terms of the factorisation properties of independent products. We recall the following definitions:

*Definition* 6 (**Factorisation and productivity**). A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ is *factorising* when

$$\underline{P}(f_1 f_2) = \underline{P}(f_1 \underline{P}(f_2)) \text{ and } \underline{P}(g_1 g_2) = \underline{P}(g_2 \underline{P}(g_1)) \ \forall f_1 \in \mathcal{L}^+(\mathcal{X}_1), g_1 \in \mathcal{L}(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2), g_2 \in \mathcal{L}^+(\mathcal{X}_2),$$

and it is *productive* when

$$\underline{P}(f_1 \cdot (f_2 - \underline{P}(f_2))) \geq 0 \text{ and } \underline{P}(g_2 \cdot (g_1 - \underline{P}(g_1))) \geq 0 \ \forall f_1 \in \mathcal{L}^+(\mathcal{X}_1), g_1 \in \mathcal{L}(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2), g_2 \in \mathcal{L}^+(\mathcal{X}_2).$$

Factorising-like properties of independent products have been established for events in a number of works [2, 23]. Since the behavioural theory of imprecise probabilities is established in terms of gambles, we think that the most sound definition of factorisation is the one considered above, which implies in particular the *strong factorisation property* from [23, Section 3.5]. With respect to productivity, when formulated for sequences of variables and not only two of them it has been linked to weak and strong laws of large numbers in [9].

Note also that a factorising lower prevision $\underline{P}$ satisfies in particular that $\underline{P}(fg) = \underline{P}(f)\underline{P}(g)$ for every $f \in \mathcal{L}^+(\mathcal{X}_1), g \in \mathcal{L}^+(\mathcal{X}_2)$. As we shall see in Example 3 later on, that condition is not sufficient for factorisation [4].

Next we are going to provide a characterisation of these conditions that will help us later in establishing their relationship with independence.

**Proposition 4.** *Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, with marginals $\underline{P}_1, \underline{P}_2$:*

[4]Note also that a similar condition, requiring $\underline{P}$(fg)=$\underline{P}$(f)$\underline{P}$(g) for any pair of gambles f,g will not be related to independence: it can be checked for instance that a vacuous coherent lower prevision will be an independent product of its (vacuous) marginals (this follows for instance from Proposition 5 in this paper), and it is factorizing in the sense we consider in the paper. However, it does not factorise in the sense considered above: if we take for instance the vacuous coherent lower prevision on $\{0,1\} \times \{2,3\}$, $f = I_0 - I_1$ and $g = I_2 - I_3$, then $\underline{P}(fg) = -1$ while $\underline{P}(f) = -1 = \underline{P}(g)$, and therefore $\underline{P}(f) \cdot \underline{P}(g) = 1 \neq \underline{P}(fg)$.

(a) $\underline{P}$ *is productive* $\Leftrightarrow \begin{cases} \underline{P}(A_1(f_2 - \underline{P}(f_2))) = 0 & \forall A_1 \subseteq \mathcal{X}_1, f_2 \in \mathcal{L}(\mathcal{X}_2) \\ \underline{P}(A_2(f_1 - \underline{P}(f_1))) = 0 & \forall A_2 \subseteq \mathcal{X}_2, f_1 \in \mathcal{L}(\mathcal{X}_1) \end{cases}$

$\Leftrightarrow \begin{cases} \underline{P}(f_1 f_2) \geq \underline{P}(f_1 \underline{P}(f_2)) & \forall f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2) \\ \underline{P}(g_1 g_2) \geq \underline{P}(g_2 \underline{P}(g_1)) & \forall g_2 \in \mathcal{L}^+(\mathcal{X}_2), g_1 \in \mathcal{L}(\mathcal{X}_1). \end{cases}$

(b) *If* $\underline{P}$ *is factorising, then it is productive.*

(c) *If* $\underline{P}$ *dominates both* $\underline{P}_1(\underline{P}_2)$ *and* $\underline{P}_2(\underline{P}_1)$*, then it is productive.*

(d) *If* $\underline{P}$ *is productive, then it satisfies (GBR) with* $\underline{P}_1(\cdot|\mathcal{X}_2)$ *and* $\underline{P}_2(\cdot|\mathcal{X}_1)$.

To see that the converse of the second statement does not hold in general, we refer to [12, Example 3]. Let us show that the converse of the fourth statement does not hold in general either:

*Example* 2. Consider $\mathcal{X}_1 := \mathcal{X}_2 := \mathbb{N}$, and let $P$ be a linear prevision on $\mathcal{X}_1 \times \mathcal{X}_2$ whose restriction to events satisfies $P(\{n\} \times \mathbb{N}) = P(\mathbb{N} \times \{n\}) = 0$ for every $n$, and $P(\{2n-1 : n \in \mathbb{N}\} \times \{2n-1 : n \in \mathbb{N}\}) = \frac{1}{2} = P(\{2n : n \in \mathbb{N}\} \times \{2n : n \in \mathbb{N}\})$. Let $P_1, P_2$ be the marginals of $P$. Since $P_1(\{n\}) = 0$ for every $n$, then given a gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$, it holds that $P(G_1(f|n)) = P(I_{\{n\}}(f(\cdot, n) - P_1(f(\cdot, n)))) = 0$; in fact, denoting $g := f(\cdot, n) - P_1(f(\cdot, n))$, we get $0 = P(\{n\}) \inf g = P(I_{\{n\}} \inf g) \leq P(I_{\{n\}} g) \leq P(I_{\{n\}} \sup g) = P(\{n\}) \sup g = 0$. This shows that $P, P_1(\cdot|\mathcal{X}_2)$ satisfy (GBR). Similarly, $P$ also satisfies (GBR) with $P_2(\cdot|\mathcal{X}_1)$. To prove that it is not productive, take $f_2 \in \mathcal{L}(\mathcal{X}_2)$ given by $f_2 := I_B - I_{B^c}$, where $B := \{2n : n \in \mathbb{N}\}$, and take $A \subseteq \mathcal{X}_1$ be given by $A := \{2n-1 : n \in \mathbb{N}\}$. Then $P_2(f_2) = P_2(B) - P_2(B^c) = 0$, so that $P(A(f_2 - P_2(f_2))) = 0 - \frac{1}{2} < 0$, given that $\{(\text{even}, \text{even}), (\text{even}, \text{odd}), (\text{odd}, \text{even}), (\text{odd}, \text{odd})\}$ is a partition of $\mathbb{N} \times \mathbb{N}$. We conclude from this that $P$ is not productive. ⧫

## 4. Independent products when one of the spaces is finite

We proceed to show that, when one of the possibility spaces is finite, there always exists an independent product of the marginals $\underline{P}_1, \underline{P}_2$.

**Proposition 5.** *Consider a coherent lower prevision* $\underline{P}$ *with marginals* $\underline{P}_1, \underline{P}_2$*, and define* $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ *by Eq.* (4)*. Assume* $\mathcal{X}_1$ *is finite.*

(a) *The coherent lower prevision* $\underline{P} := \inf\{P_1(P_2) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$ *is an independent product of* $\underline{P}_1, \underline{P}_2$*. As a consequence,* $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ *are coherent.*

(b) $\underline{P}$ *is an independent product if and only if it dominates both* $\underline{P}_1(\underline{P}_2)$ *and* $\underline{P}_2(\underline{P}_1)$.

As a consequence, the independent natural extension $\underline{P}_1 \otimes \underline{P}_2$ is the smallest coherent lower prevision that dominates both $\underline{P}_1(\underline{P}_2)$ and $\underline{P}_2(\underline{P}_1)$, and we deduce from Theorem 3 that $\mathcal{M}(\underline{P}_1 \otimes \underline{P}_2) = \mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1))$. In this case, the proposition above guarantees that the intersection of the credal sets $\mathcal{M}(\underline{P}_1(\underline{P}_2)), \mathcal{M}(\underline{P}_2(\underline{P}_1))$ is always non-empty, or, in other words, that two coherent lower previsions $\underline{P}_1, \underline{P}_2$ always have an independent product when one of the possibility spaces is finite. This was mentioned without any proof in [24, Footnote 5 in Section 9.3].

Next, we study the connection between productivity and independence when one of the possibility spaces is finite. We establish the following:

**Proposition 6.** *Consider a coherent lower prevision* $\underline{P}$ *with marginals* $\underline{P}_1, \underline{P}_2$*, and assume* $\mathcal{X}_1$ *is finite. The following are equivalent.*

(a) $\underline{P}$ *is productive.*

(b) $\underline{P}$ *is an independent product.*

(c) $\underline{P}(f) \geq \max\{\min \underline{P}_2(f|\mathcal{X}_1), \inf \underline{P}_1(f|\mathcal{X}_2)\}\ \forall f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$.

This result, together with Proposition 4, allows us to deduce that when one of the possibility spaces is finite any factorising coherent lower prevision $\underline{P}$ is an independent product of its marginals. To see that the converse is not true, we refer to [12, Example 3].

From Theorem 3, any independent product (and in particular the independent natural extension) must dominate the two concatenations $\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$. Our next result characterises in which cases the independent natural extension coincides with one of them:

**Proposition 7.** *Consider $\underline{P}_1, \underline{P}_2$ coherent lower previsions on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$, where $\mathcal{X}_1$ is finite, and let $\underline{P}_1 \otimes \underline{P}_2$ denote their independent natural extension.*

- *(a)* $\underline{P}_1 \otimes \underline{P}_2 = \underline{P}_1(\underline{P}_2) \Leftrightarrow \underline{P}_1(\underline{P}_2) \geq \underline{P}_2(\underline{P}_1) \Leftrightarrow \underline{P}_1(\underline{P}_2)$ *productive* $\Leftrightarrow \underline{P}_2$ *linear or* $\underline{P}_1$ *vacuous.*
- *(b)* $\underline{P}_1 \otimes \underline{P}_2 = \underline{P}_2(\underline{P}_1) \Leftrightarrow \underline{P}_2(\underline{P}_1) \geq \underline{P}_1(\underline{P}_2) \Leftrightarrow \underline{P}_2(\underline{P}_1)$ *productive* $\Leftrightarrow \underline{P}_1$ *linear or* $\underline{P}_2$ *vacuous.*

In fact, when one of the marginals (say, $\underline{P}_1$) is linear, there is only one independent product: the concatenation $\underline{P}_2(P_1)$, because of Proposition 2(b). In particular, if the two marginals $P_1, P_2$ are linear, then it follows from the result above that the two concatenations coincide $P_1(P_2) = P_2(P_1)$ (remember that we are assuming in this section that $\mathcal{X}_1$ is finite, so this poses no contradiction with Example 1). However, in general $\underline{P}_1 \otimes \underline{P}_2$ need not be the only independent product. Indeed, taking into account that coherence is preserved by taking lower envelopes, one natural way to obtain independent products is to consider a family of independent linear products, and then take its lower envelope. This produces the following definition:

*Definition* 7. Given two marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ with respective domains $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$, their *strong product* is $\underline{P}_1 \boxtimes \underline{P}_2(f) := \inf\{P_1(P_2(f)) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$ for every gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$.

Taking into account our comments above, we have that $\underline{P}_1 \boxtimes \underline{P}_2(f) := \inf\{P_2(P_1(f)) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$ for every gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$. The strong product was called *type-1 product* in [24, Section 9.3.5] and [3]. It has been studied in a number of works in the literature [7, 12, 20] and it is a founding concept for *credal networks* [4]. It models a notion of independence called *strong independence* that is more restrictive than the one considered so far, as we can also see by the following result:

**Proposition 8.** *Consider a coherent lower prevision $\underline{P}$ with marginals $\underline{P}_1, \underline{P}_2$.*

- *(a) If $\underline{P}$ is the strong product of $\underline{P}_1, \underline{P}_2$, then it is an independent product.*
- *(b) The strong product $\underline{P}_1 \boxtimes \underline{P}_2$ is factorising, and so is any coherent lower prevision $\underline{P}$ bounded between $\underline{P}_1 \otimes \underline{P}_2$ and $\underline{P}_1 \boxtimes \underline{P}_2$.*

Now we are able to show that the factorisation on positive gambles does not guarantee that a coherent lower prevision is an independent product, and as a consequence nor does it imply that it is factorising.

*Example* 3. Consider $\mathcal{X}_1 = \mathcal{X}_2 := \{0, 1\}$ and the following probability mass functions on $\{(0,0), (0,1), (1,0), (1,1)\}$:

$$p_1 := (0,0,0,1), p_2 := \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right), p_3 := \left(0, 0, \frac{1}{2}, \frac{1}{2}\right), p_4 := \left(0, \frac{1}{2}, 0, \frac{1}{2}\right), p_5 := \left(\frac{3}{8}, \frac{1}{8}, \frac{1}{8}, \frac{3}{8}\right).$$

Call $P_i$ the linear prevision induced by $p_i$; we define the lower envelope of the linear previsions created in that way as

$$\underline{P} := \min\{P_1, P_2, P_3, P_4, P_5\}.$$

Note that the first four of these linear previsions satisfy stochastic independence, while the last one does not. Moreover, if we consider the marginals $\underline{P}_1, \underline{P}_2$ of $\underline{P}$, it can be checked that $\underline{P}_1 \boxtimes \underline{P}_2 = \min\{P_1, P_2, P_3, P_4\} \geq \underline{P}$.

Now, for every $f \in \mathcal{L}^+(\mathcal{X}_1), g \in \mathcal{L}^+(\mathcal{X}_2)$, we have that $\underline{P}_1 \boxtimes \underline{P}_2(fg) \leq P_5(fg)$. To prove this, note that

$$\begin{aligned}
P_2(fg) > P_5(fg) &\Leftrightarrow (f(1) - f(0))(g(0) - g(1)) > 0, \\
P_3(fg) > P_5(fg) &\Leftrightarrow (f(1) - f(0))(3g(0) + g(1)) > 0, \\
P_4(fg) > P_5(fg) &\Leftrightarrow (3f(0) + f(1))(g(1) - g(0)) > 0,
\end{aligned}$$

and that these three equations are incompatible, taking into account that $3f(0) + f(1)$ and $3g(0) + g(1)$ are non-negative. As a consequence, $\underline{P}(fg) = \underline{P}_1 \boxtimes \underline{P}_2(fg)$, and since the latter is a factorising coherent lower prevision by Proposition 8, we conclude that $\underline{P}(fg) = \underline{P}(f)\underline{P}(g)$ for every $f \in \mathcal{L}^+(\mathcal{X}_1), g \in \mathcal{L}^+(\mathcal{X}_2)$.

However, $\underline{P}$ is not factorising itself, because it is not an independent product. To see this, consider the gamble $f := -I_{(0,0)} + I_{\{(0,1),(1,0),(1,1)\}}$ and note that

$$\underline{P}(f) = \frac{1}{4} < \underline{P}_1(\underline{P}_2(f)) = \underline{P}_1(I_1) = \frac{1}{2}.$$

Applying Proposition 5(b), we deduce that $\underline{P}$ is not an independent product. Then Proposition 6 implies that $\underline{P}$ is not productive and from Proposition 4(b) we conclude that it is not factorising. ⧫

The independent natural extension does not coincide with the strong product in general [24, Example 9.3.4]. Our next result gives a sufficient condition for the equality between them:

**Proposition 9.** *Consider marginal coherent lower previsions $\underline{P}_1$ on $\mathcal{L}(\mathcal{X}_1)$ and $\underline{P}_2$ on $\mathcal{L}(\mathcal{X}_2)$, where $\mathcal{X}_1$ is finite. If either $\underline{P}_1$ or $\underline{P}_2$ is vacuous or linear, then $\underline{P}_1 \otimes \underline{P}_2 = \underline{P}_1 \boxtimes \underline{P}_2$.*

Recall that, from Proposition 7, we deduce that in this case the strong product (and the independent natural extension) coincide with one of the concatenations of the marginals.

The following diagram summarises the implications when one of the possibility spaces is finite:

$$\begin{array}{ccccc} & & \underline{P} \text{ factorising} & & \\ & & \Downarrow & & \\ \underline{P} \text{ productive} & \Longleftrightarrow & \underline{P} \text{ independent product} & \Longleftrightarrow & \underline{P} \geq \underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1) \\ & & \Downarrow & & \\ & & \underline{P} \text{ satisfies (GBR) with } \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1). & & \end{array}$$

Next we show that when both $\mathcal{X}_1, \mathcal{X}_2$ are finite, the strong product is the greatest factorising independent product.

**Proposition 10.** *Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with marginals $\underline{P}_1, \underline{P}_2$. If $\mathcal{X}_1, \mathcal{X}_2$ are finite and $\underline{P}$ is factorising, then it is dominated by $\underline{P}_1 \boxtimes \underline{P}_2$.*

This shows the incompatibility of the factorisation condition with an envelope-like theorem: not only the lower envelope of a family of factorising linear previsions will not be factorising in general (it will only be so when it produces the strong product), but also a factorising lower prevision (which by Proposition 8 will be any coherent lower prevision bounded between the independent natural extension and the strong product) will not be the lower envelope of the dominating factorising linear previsions, unless it coincides with the strong product.

## 5. Extension to the general case

In the previous section, we gave a simple characterisation of independent products of marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ when one of the underlying possibility spaces is finite. Now we focus on the general case, where both $\mathcal{X}_1, \mathcal{X}_2$ can be infinite. As we know from Example 1, in this case there may not be any independent product of two marginals, not even in the linear case. This motivates the following definition:

*Definition* 8 (**Compatibility**). Two marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$ are called *compatible* when they have an independent product.

Taking Example 1 into account, there are two main avenues we can pursue: one is establishing sufficient conditions for the compatibility of the marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$; the other is relaxing the definition of independent product. With respect to the first path, we can establish the following:

**Proposition 11.** *Consider two marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$. They are compatible when any of the following conditions holds:*

*(a) Either $\underline{P}_1$ or $\underline{P}_2$ is vacuous.*

(b) *Either $\mathcal{X}_1$ or $\mathcal{X}_2$ is finite.*

(c) $\mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1)) \neq \emptyset$ *and* $\underline{P}_1(x_1) > 0, \underline{P}_2(x_2) > 0$ *for every* $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$.

A sufficient condition for compatibility established by Walley in [24, Section 9.3.2] for linear marginals $P_1, P_2$ is that their restrictions to events are countably additive. Indeed, if $\underline{P}_1, \underline{P}_2$ are the lower envelopes of countably additive linear marginals they are always compatible: we can build the products of these linear marginals and then take their lower envelope in order to obtain an independent product. However, note that coherent lower previsions are the lower envelopes of sets of finitely additive previsions that need not be countably additive. We refer to [24, Section 6.9] for a more detailed discussion of countable additivity and conditional lower previsions.

On the other hand, the existence of a factorising joint does not imply that the two marginals are compatible, because if both referential spaces are infinite a factorising coherent lower prevision $\underline{P}$ need not be an independent product of its marginals (it need not even dominate both concatenations $\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$). To prove this, it suffices to reconsider Example 1. In that case, any of the concatenations $P_1(P_2), P_2(P_1)$ is factorising; consider the former concatenation (the remaining case is analogous):

$$P_1(P_2(f_1 f_2|\mathcal{X}_1)) = P_1(f_1 P_2(f_2|\mathcal{X}_1)) = P_1(f_1 P_2(f_2)) = P_1(f_1)P_2(f_2),$$
$$P_1(P_2(g_1 g_2|\mathcal{X}_1)) = P_1(g_1 P_2(g_2|\mathcal{X}_1)) = P_1(g_1 P_2(g_2)) = P_1(g_1)P_2(g_2),$$

for every $f_1 \in \mathcal{L}^+(\mathcal{X}_1), g_1 \in \mathcal{L}(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2), g_2 \in \mathcal{L}^+(\mathcal{X}_2)$. Note that the first equality in both lines holds because of Eq. (1), since $f_1$ and $g_1$ are $\mathcal{X}_1$-measurable (their positivity is actually not necessary for the result to hold given that $P_2$ is a linear prevision). This shows that $P_1(P_2)$, and also $P_2(P_1)$, are factorising. But we have seen that there are no independent products of the marginal linear previsions $P_1, P_2$.

Next we discuss ways in which the notion of coherence in the definition of independent products can be relaxed. One way is to consider the notion of *weak coherence* instead of that of coherence. This notion has been related to the problem of the compatibility of marginal and conditional assessments and to satisfiability in [16, Section 8]. Taking into account Theorem 3, we see that a coherent lower prevision $\underline{P}$ is weakly coherent with the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ induced by its marginals and irrelevance if and only if it dominates the two concatenations $\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$. In spite of this nice result, we think that weak coherence may be indeed too weak from the behavioural point of view, because it does not rule out some inconsistent assessments. See [24, Section 7.3.5] for an example.

Another way in which we can weaken the notion of coherence is to consider Williams' notion of coherence, as we detail in the following section. This approach is closely related to Vicig's study of independence in terms of conditional events and gambles [23].

*Remark* 1. As we have mentioned in Footnote 1 of Section 2, our treatment of Williams' coherence is more restrictive than Williams' original formulation: for, in order to make a comparison with Walley's coherence, we are embedding Williams' coherence into Walley's setup. As a consequence, our study of independence is, in this respect, more restrictive than Vicig's in [23], given that the latter employs Williams' definition of coherence.

In particular, Williams' coherence does not require the conditional lower previsions to be conditional on partitions, but rather on *conditional gambles* of the type $f|B$ that are assumed to be called off unless $B$ is observed.

Using this idea, given two marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ with respective domains $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$, we can use an assessment of irrelevance to define

$$\underline{P}_1(f_1|B_2) := \underline{P}_1(f_1) \ \forall f_1 \ \mathcal{X}_1\text{-measurable}, \ \forall B_2 \subseteq \mathcal{X}_2$$

and

$$\underline{P}_2(f_2|B_1) := \underline{P}_2(f_2) \ \forall f_2 \ \mathcal{X}_2\text{-measurable}, \ \forall B_1 \subseteq \mathcal{X}_1.$$

Then the conditional lower previsions we obtain in this manner would satisfy Williams' definition of coherence and this formulation would prevent the counterintuitive results we shall describe in Example 4 below. ⧫

*5.1. W-independent products*

*Definition* 9 (**Williams independent product and natural extension**). A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ is said to be a *W-independent product* of its marginals $\underline{P}_1, \underline{P}_2$ when $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are W-coherent. The smallest such product is called the *W-independent natural extension*.

Next, we establish that two marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ are always compatible if we consider Williams' notion of coherence:

**Lemma 12.** *Consider two coherent lower previsions $\underline{P}_1, \underline{P}_2$ on $\mathcal{L}(\mathcal{X}_1), \mathcal{L}(\mathcal{X}_2)$ and define $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ by means of Eq.* (4). *Then $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are W-coherent.*

This allows us to deduce the following:

**Proposition 13.** *Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with marginals $\underline{P}_1, \underline{P}_2$.*

*(a) $\underline{P}$ is a W-independent product of its marginals $\Leftrightarrow$ it satisfies (GBR) with both $\underline{P}_1(\cdot|\mathcal{X}_2)$ and $\underline{P}_2(\cdot|\mathcal{X}_1)$.*

*(b) The W-independent natural extension of $\underline{P}_1, \underline{P}_2$ is given by*

$$\underline{P}(f) := \sup\{\mu : f-\mu \geq G_1(g|\mathcal{X}_2)+G_2(h|\mathcal{X}_1) \text{ for some } g, h \in \mathcal{L}(\mathcal{X}_1\times\mathcal{X}_2) \text{ such that } S_1(g), S_2(h) \text{ are finite}\} \tag{5}$$

*for any gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$.*

*(c) If $\underline{P}$ is productive, then it is a W-independent product of its marginals.*

Despite this result, we think that Williams' notion of coherence is too weak to properly deal with independence in infinite spaces:[5] on the one hand, and unlike Walley's notion of coherence, $\underline{P}$ may be W-coherent with $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ and not have $\underline{P}_1, \underline{P}_2$ as marginals. To see this, it suffices to take into account that if $\overline{P}(x_1) = \overline{P}(x_2) = 0$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$, then $\underline{P}$ satisfies (GBR) with *any* conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2)$, $\underline{P}_2(\cdot|\mathcal{X}_1)$: in fact, $\overline{P}(x_i) = 0$ implies that $\underline{P}(I_{x_i} f) = \overline{P}(I_{x_i} f) = 0$ for all gambles $f$. This means that if we consider any marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ and define $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ by irrelevance, $\underline{P}$ will be a W-independent product. In particular, a linear prevision $P$ on $\mathcal{X}_1 \times \mathcal{X}_2$ that satisfies $P(x_1) = P(x_2) = 0$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$ is a W-independent product of *any* marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$. This shows that in the definition of W-independent products we need to explicitly require that $\underline{P}_1, \underline{P}_2$ are the marginals of $\underline{P}$.

However, even if we assume that $\underline{P}$ has marginals $\underline{P}_1, \underline{P}_2$, the W-coherence with $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ is still too weak: for instance, if $\mathcal{X}_1, \mathcal{X}_2$ are infinite and we consider two linear previsions $P_1, P_2$ such that $P_1(x_1) = P_2(x_2) = 0$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$, then any linear prevision $P$ with marginals $P_1, P_2$ will be a W-independent product, which does not seem very reasonable. From this observation and Example 1 we can also deduce that W-independent products are not necessarily independent products. Indeed, since W-coherence is less restrictive than Walley's, we see that the W-independent natural extension of $\underline{P}_1, \underline{P}_2$ is dominated by their independent natural extension, when the latter exists. However, the W-independent natural extension may be too uninformative, as we show next:

*Example* 4. Consider two infinite spaces $\mathcal{X}_1, \mathcal{X}_2$, and let $\underline{P}_1, \underline{P}_2$ be two marginal coherent lower previsions. It follows from Eq. (5) that $\underline{P}(f) = \sup\{\inf[f - G_1(g|\mathcal{X}_2) - G_2(h|\mathcal{X}_1)] : g, h \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2) \text{ such that } S_1(g), S_2(h) \text{ are finite}\}$. For any gambles $g, h \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with $S_1(g) = A_2$ finite and $S_2(h) = A_1$ finite, it holds that $G_1(g|\mathcal{X}_2) = 0 = G_2(h|\mathcal{X}_1)$ in $A_1^c \times A_2^c$. Thus, $\inf[f - G_1(g|\mathcal{X}_2) - G_2(h|\mathcal{X}_1)] \leq \inf_{A_1^c \times A_2^c}[f - G_1(g|\mathcal{X}_2) - G_2(h|\mathcal{X}_1)] = \inf_{A_1^c \times A_2^c} f$. As a consequence, $\underline{P}(f) \leq \sup\{\inf_{A_1^c \times A_2^c} f : A_1 \subseteq \mathcal{X}_1, A_2 \subseteq \mathcal{X}_2 \text{ finite}\}$. In particular, if we take $\mathcal{X}_1 := \mathcal{X}_2 := \mathbb{N}$ and the gamble $f$ given by $f(n,m) := \frac{1}{nm}$, it follows that $\underline{P}(f) = 0$ *irrespective of the marginal coherent lower previsions $\underline{P}_1, \underline{P}_2$ we start with*! ♦

Let us show that W-independent products need not be productive either:

[5] Again, assuming we consider conditional lower previsions on partitions, as in Walley's framework; see Remark 1.

*Example* 5. Consider two linear previsions $P, Q$ on $\mathcal{L}(\mathbb{N})$ such that $P(\{n\}) = 0 = Q(\{n\})$ for every $n$, $P(\{2n-1 : n \in \mathbb{N}\}) = 1 = Q(\{2n : n \in \mathbb{N}\})$. Define $\mathcal{X}_1 := \mathcal{X}_2 := \mathbb{N}$, $P_1 := \frac{P+Q}{2}$, and $P_2(\cdot|\mathcal{X}_1)$ as

$$P_2(f|n) := \begin{cases} P(f(n,\cdot)) & \text{if } n \text{ odd,} \\ Q(f(n,\cdot)) & \text{if } n \text{ even.} \end{cases}$$

Consider the coherent lower prevision $Q' := P_1(P_2(\cdot|\mathcal{X}_1))$. Its $\mathcal{X}_1$-marginal is $P_1$. On the other hand, given an $\mathcal{X}_2$-measurable gamble $f$,

$$P'(f) = P_1(\{2n-1 : n \in \mathbb{N}\})P(f) + P_1(\{2n : n \in \mathbb{N}\})Q(f) = (P(f)+Q(f))/2 = P_1(f).$$

From this we deduce that for every natural number $n$ it is $P'(\{n\}\times\mathbb{N}) = P'(\mathbb{N}\times\{n\}) = 0$, and as a consequence $P'$ is a W-independent product of its marginals. To prove that it is not productive, take $f := I_{\text{odd}} \in \mathcal{L}(\mathcal{X}_1), g := I_{\text{even}} \in \mathcal{L}(\mathcal{X}_2)$. Then

$$P'(f(g - P'(g))) = P'(fg) - P'(f)P'(g) = P'((\text{odd}, \text{even})) - P_1(\text{odd})P_1(\text{even}) = 0 - \frac{1}{4},$$

whence $P'$ is not productive. ♦

The following diagram summarises the relationships between the different conditions for arbitrary $\mathcal{X}_1, \mathcal{X}_2$:

$$\begin{array}{ccc} \underline{P} \text{ factorising} & \not\Leftarrow & \underline{P} \text{ independent product} \\ \Downarrow\not\Uparrow & \not\swarrow & \Downarrow \\ \underline{P} \text{ productive} & \Leftarrow & \underline{P} \geq \underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1) \\ \Downarrow\not\Uparrow & & \end{array}$$

$$\underline{P} \text{ satisfies (GBR) with } \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1) \iff \underline{P} \text{ W-indep. product.}$$

It is not difficult to find examples showing that the missing implications do not hold:

- [12, Example 3]: $\underline{P}$ productive, independent product $\not\Rightarrow$ $\underline{P}$ factorising.
- Example 1: $\underline{P}$ factorising $\not\Rightarrow$ $\underline{P} \geq \underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$. From this we deduce that $\underline{P}$ factorising $\not\Rightarrow$ $\underline{P}$ independent product and that $\underline{P}$ productive $\not\Rightarrow$ $\underline{P} \geq \underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$.
- Example 5: $\underline{P}$ W-independent product $\not\Rightarrow$ $\underline{P}$ productive.

There is only one open problem at this stage: we still need to establish whether any coherent lower prevision $\underline{P}$ that dominates both concatenations $\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$ of its marginals is an independent product. From Theorem 3, this will be the case as soon as the two conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent, that is, as soon as the marginals $\underline{P}_1, \underline{P}_2$ are compatible. But it is still to be established whether the compatibility of $\underline{P}_1, \underline{P}_2$ is equivalent to the credal sets $\mathcal{M}(\underline{P}_1(\underline{P}_2)), \mathcal{M}(\underline{P}_2(\underline{P}_1))$ having non-empty intersection. Our current conjecture is that this is not the case: by Theorem 3, $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly coherent if and only if there is some $\underline{P} \geq \underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$, so if we had the implication it would follow that $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly coherent if and only if they are coherent. It follows from the discussion in [24, Section 7.3] that this is not the case in general, and at this moment we see no reason why the equivalence should hold when the conditional lower previsions are defined by irrelevance.

### 5.2. *Irrelevant products*

Since W-independent products seem to be too weak to give a proper account of independence, we shall next consider the case where we stick to Walley's coherence but we only make an assumption of irrelevance of one of the variables towards the other. This gives rise to the following notion:

*Definition* 10 (**2-1 irrelevant product**). Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with marginals $\underline{P}_1, \underline{P}_2$, and define $\underline{P}_1(\cdot|\mathcal{X}_2)$ by means of Eq. (4). We say that $\underline{P}$ is a *2-1 irrelevant product* when it is coherent with $\underline{P}_1(\cdot|\mathcal{X}_2)$.

It follows immediately from Theorem 3 that $\underline{P}$ is a 2-1 product of its marginals $\underline{P}_1, \underline{P}_2$ if and only if it dominates $\underline{P}_2(\underline{P}_1)$. This implies that the smallest 2-1 irrelevant product is given by the *marginal extension* $\underline{P}_2(\underline{P}_1)$, which is the generalisation of the law of total probability to the imprecise case. To see that this is not the only 2-1 irrelevant product, consider the *2-1 strong product*, given by $\underline{S}_{2\text{-}1}(f) := \min\{P_2(P_1(f)) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$ for every $f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$. Note that these two products do not coincide in general (consider for instance $\underline{P}_1, \underline{P}_2$ as in [24, Example 9.3.4]).

## 6. Conclusions

The study of independence is often regarded as relatively settled in the case of precise probability. The word 'relatively' accounts for the subtle issues originated by conditioning on events of probability zero and that are relevant for finite spaces and even more for infinite ones. In the case of imprecise probability, the situation becomes more complex, as there is more than one extension of the notion of independence (e.g., see the surveys [2, 7]). Interestingly, one can uniformly address all the issues above, both in the precise and in the imprecise case, by a shift of paradigm in the definition of independence. The work in [12] did in fact investigate this avenue by regarding an independent product as a probabilistic model that is coherent with the assessments of mutual irrelevance between given variables. That work also related this approach to the more traditional one based on some factorisation condition. However, the study in [12] was restricted to the case of finite possibility spaces, an assumption that is not always met in practice—for instance in statistical applications. In this paper, we have analysed this still largely unexplored side of independence by considering the general case.

In particular, we have focused on the types of independent products that arise from two marginals using Walley's notion of coherence. We have showed that, contrary to the case of finite spaces, this type of independent product does not always exist. This applies even to the case of precise probability and, in the imprecise case, also to well-established notions such as strong independence. We have proved that we can nonetheless always build independent products when one of the possibility spaces is finite, and that in that case one can characterise independent products by means of a number of properties: productivity, which was employed in the derivation of the laws of large numbers in [9]; weak coherence, which is a weaker version of Walley's consistency axiom that has been related to satisfiability in [16]; and more generally we have proven that a joint model is an independent product if and only if it dominates the two concatenations of its marginals.

Since independent products do not exist in general, we have investigated two alternative models that guarantee their existence. On the one hand, we have considered the notion of coherence proposed by Williams, which in our opinion proves too weak to give a proper account of independence; on the other, we have considered a single assumption of irrelevance of one variable to the other (instead of both as in the case of independence), and showed that in that case we can use the extension to imprecision of the law of total probability to establish the joint model. Nevertheless, our study of Williams' coherence has assumed that lower previsions are conditional on partitions of the possibility space, which is only required in Walley's formulation. A more general approach, briefly discussed in Remark 1, might make it possible to define also informative products in the infinite case. A thorough study of this avenue of research is left as an open problem.

Note also that we have characterised the strong product as the greatest factorising product of given marginals, when all the possibility spaces are finite. It is a small step to extend our proof to the case of $n$ variables, and this allows to close a few open problems from [12]: we deduce that the strong product is the greatest joint model satisfying other conditions of interest, such as *strong factorisation* and *strong Kuznetsov independence*; that any factorising model satisfies the condition of *external additivity*; and that we can characterise factorising models as those lying between the independent natural extension and the strong product.

With respect to the open problems remaining from the analysis we have carried out in this paper, we point out three: one of them is the characterisation of independent products as models that dominate the two concatenations; if this held we would have an easy way to verify if a joint model is an independent product, and we would deduce the equivalence between weak coherence and coherence in this context. Another open problem would be the quest for other sufficient conditions for the existence of an independent product of two marginals. Although we have already provided a few in this paper, it would be interesting to determine other conditions, for instance in terms of envelope theorems. Finally, in some other works [18, 27] we have argued that there are cases where Walley's coherence notion should be strengthened into what we have called *conglomerable coherence*. We prove in Appendix B that when one possibility space is finite Walley's notion and conglomerable coherence coincide for the two conditionals defined by irrelevance, so that the analysis in this paper covers already such an extension. For the case where both spaces are infinite, we give a number of quite stringent sufficient conditions for the equality and yet it is an open problem to prove necessity at this point.

Last, since the applications of independence are widespread, we would like to study if our results can be employed to simplify the combination of marginal models in some problems of interest. We believe, for instance, that our new formulation of the independent natural extension in Theorem 3 has quite some potential to impact on algorithms and applications, with special regard to graphical models.

**Acknowledgements**

This work has been partly supported by project MTM2010-17844 and by Swiss NSF grants nos. 200020_137680 / 1 and 200021_146606 / 1. We would also like to thank the anonymous reviewer for a careful reading of the paper that led to many useful suggestions.

**Appendix A. Proofs and technical results**

*Proof of Theorem 3.* (a) We begin with the direct implication. If $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ are coherent, then $\underline{P}$ must dominate the smallest coherent lower prevision with marginal $\underline{P}_2$ that is coherent with $\underline{P}_1(\cdot|\mathcal{X}_2)$. By Proposition 2(a), this is given by $\underline{P}_2(\underline{P}_1)$.

Conversely, if $\underline{P}$ dominates $\underline{P}_2(\underline{P}_1)$, then given a gamble $f$ in $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ it holds that $\underline{P}(G_1(f|\mathcal{X}_2)) = \underline{P}(f - \underline{P}_1(f|\mathcal{X}_2)) \geq \underline{P}_2(\underline{P}_1(f - \underline{P}_1(f|\mathcal{X}_2)|\mathcal{X}_2)) = \underline{P}_2(\underline{P}_1(f|\mathcal{X}_2) - \underline{P}_1(f|\mathcal{X}_2)) = 0$, where the penultimate passage holds because $\underline{P}_1(f|\mathcal{X}_2)$ is $\mathcal{X}_2$-measurable. In particular, it follows that $\underline{P}(G_1(f|x_2)) \geq 0$ for every $x_2 \in \mathcal{X}_2$, and $\underline{P}(G_1(fI_A|\mathcal{X}_2)) \geq 0$ for all $A \subseteq \mathcal{X}_2$. Assume ex-absurdo that there is some gamble $h$ on $\mathcal{X}_1 \times \mathcal{X}_2$ and some $x_2 \in \mathcal{X}_2$ such that $\underline{P}(G_1(h|x_2)) > 0$. Then if we define the $\mathcal{X}_1$-measurable gamble $g$ by $g(x_1', x_2') := h(x_1', x_2)$ for every $(x_1', x_2') \in \mathcal{X}_1 \times \mathcal{X}_2$, we obtain that $\underline{P}(G_1(g|x_2)) > 0$ and $\underline{P}(G_1(gI_{x_2^c}|\mathcal{X}_2)) \geq 0$. It follows that $\underline{P}(G_1(g|\mathcal{X}_2)) \geq \underline{P}(G_1(gI_{x_2}|\mathcal{X}_2)) + \underline{P}(G_1(gI_{x_2^c}|\mathcal{X}_2)) = \underline{P}(G_1(g|x_2)) + \underline{P}(G_1(gI_{x_2^c}|\mathcal{X}_2)) > 0$, while on the other hand $G_1(g|\mathcal{X}_2)(x_1', x_2') = h(x_1', x_2) - \underline{P}_1(h(\cdot, x_2))$ for every $(x_1', x_2') \in \mathcal{X}_1 \times \mathcal{X}_2$, and therefore $\underline{P}(G_1(g|\mathcal{X}_2)) = \underline{P}_1(h(\cdot, x_2) - \underline{P}_1(h(\cdot, x_2))) = 0$, taking into account that $\underline{P}_1$ is the marginal of $\underline{P}$. This is a contradiction. We deduce thus that $\underline{P}(G_1(f|x_2)) = 0$ for every gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$ and every $x_2 \in \mathcal{X}_2$. Since $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ satisfy (CNG) and (GBR), we deduce from Proposition 1 that they are coherent.

(b) This follows from the first statement and the equivalence between weak coherence and pairwise coherence established in [16, Theorem 1].

(c) By the reduction theorem in [24, Theorem 7.1.5], $\underline{P}$ is an independent product of $\underline{P}_1, \underline{P}_2$ if and only if $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly coherent and $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent. From the second statement, the weak coherence of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ holds if and only if $\underline{P} \geq \max\{\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)\}$. On the other hand, if there is some coherent lower prevision $\underline{Q}$ that is coherent with $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$, we deduce in particular that these two conditional lower previsions are coherent. Thus, the smallest independent product $\underline{P}_1 \otimes \underline{P}_2$ is the smallest coherent lower prevision that dominates both $\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$. As a consequence, its credal set $\mathcal{M}(\underline{P}_1 \otimes \underline{P}_2)$ is the greatest credal set included in $\mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1))$. Since the latter is a weak *-compact and convex set of linear previsions, we deduce that $\mathcal{M}(\underline{P}_1 \otimes \underline{P}_2) = \mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1))$. □

**Lemma 14.** *Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, with marginals $\underline{P}_1, \underline{P}_2$. The following are equivalent:*

*(a1)* $\underline{P}(f_1(f_2 - \underline{P}(f_2))) \geq 0 \; \forall f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$.

*(a2)* $\underline{P}(f_1 f_2) \geq \underline{P}(f_1 \underline{P}(f_2)) \; \forall f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$.

*(a3)* $\underline{P}(A_1(f_2 - \underline{P}_2(f_2))) = 0 \; \forall A_1 \subseteq \mathcal{X}_1, f_2 \in \mathcal{X}_2$.

*Similarly, the following are equivalent:*

*(b1)* $\underline{P}(f_2(f_1 - \underline{P}(f_1))) \geq 0 \; \forall f_1 \in \mathcal{L}(\mathcal{X}_1), f_2 \in \mathcal{L}^+(\mathcal{X}_2)$.

*(b2)* $\underline{P}(f_2 f_1) \geq \underline{P}(f_2 \underline{P}(f_1)) \; \forall f_1 \in \mathcal{L}(\mathcal{X}_1), f_2 \in \mathcal{L}^+(\mathcal{X}_2)$.

*(b3)* $\underline{P}(A_2(f_1 - \underline{P}_1(f_1))) = 0 \; \forall A_2 \subseteq \mathcal{X}_2, f_1 \in \mathcal{X}_1$.

*Proof.* Let us make a circular proof of the equivalence (a1) ⇔ (a2) ⇔ (a3). The proof of the equivalence (b1) ⇔ (b2) ⇔ (b3) is analogous.

(a1)⇒(a2) Consider $f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$. Then $\underline{P}(f_1 f_2) \geq \underline{P}(f_1(f_2 - \underline{P}(f_2))) + \underline{P}(f_1 \underline{P}(f_2)) \geq \underline{P}(f_1 \underline{P}(f_2))$, where the first inequality follows from the super-additivity of $\underline{P}$ and the second from (a1).

(a2)⇒(a3) Given $A_1 \subseteq \mathcal{X}_1$ and $f_2 \in \mathcal{L}(\mathcal{X}_2)$, condition (a2) implies that $\underline{P}(A_1(f_2 - \underline{P}(f_2))) \geq \underline{P}(A_1(\underline{P}(f_2 - \underline{P}(f_2)))) = 0$, where the equality follows from [24, Proposition 2.6.1(b,c)]. In particular, we also have that $\underline{P}(A_1^c(f_2 - \underline{P}(f_2))) \geq 0$, and since $0 = \underline{P}(f_2 - \underline{P}(f_2)) \geq \underline{P}(A_1(f_2 - \underline{P}(f_2))) + \underline{P}(A_1^c(f_2 - \underline{P}(f_2))) \geq 0$, we deduce that $\underline{P}(A_1(f_2 - \underline{P}(f_2))) = 0$.

(a3)⇒(a1) Consider $f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$. If $f_1$ is simple, $f_1 = \sum_{i=1}^n x_i I_{A_i}$, it follows from super-additivity that $\underline{P}(f_1(f_2 - \underline{P}(f_2))) = \underline{P}((\sum_{i=1}^n x_i I_{A_i})(f_2 - \underline{P}(f_2))) = \underline{P}(\sum_{i=1}^n x_i I_{A_i}(f_2 - \underline{P}(f_2))) \geq \sum_{i=1}^n x_i \underline{P}(A_i(f_2 - \underline{P}(f_2))) = 0$. On the other hand, given $f_1 \in \mathcal{L}^+(\mathcal{X}_1)$, there is a sequence $(h_n)_{n\in\mathbb{N}}$ of simple gambles in $\mathcal{L}^+(\mathcal{X}_1)$ that converges uniformly towards $f_1$. As a consequence, the sequence $(h_n(f_2 - \underline{P}(f_2)))_{n\in\mathbb{N}}$ converges uniformly towards $f_1(f_2 - \underline{P}(f_2))$, and since coherent lower previsions are continuous under uniform convergence by [24, Theorem 2.6.1($\ell$)], $\underline{P}(f_1(f_2 - \underline{P}(f_2))) = \lim_{n\to\infty} \underline{P}(h_n(f_2 - \underline{P}(f_2))) \geq 0$. □

**Lemma 15.** *Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, with marginals $\underline{P}_1, \underline{P}_2$. Then*

$$\underline{P} \geq \underline{P}_1(\underline{P}_2) \Rightarrow \underline{P}(f_1(f_2 - \underline{P}(f_2))) \geq 0 \;\forall f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2),$$

*and similarly*

$$\underline{P} \geq \underline{P}_2(\underline{P}_1) \Rightarrow \underline{P}(f_2(f_1 - \underline{P}(f_1))) \geq 0 \;\forall f_1 \in \mathcal{L}(\mathcal{X}_1), f_2 \in \mathcal{L}^+(\mathcal{X}_2).$$

*Proof.* Let us prove the first equation; the proof of the second is similar. Consider $f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$. Then $\underline{P}(f_1(f_2 - \underline{P}(f_2))) \geq \underline{P}_1(\underline{P}_2(f_1(f_2 - \underline{P}(f_2)))) = \underline{P}_1(f_1 \underline{P}_2(f_2 - \underline{P}_2(f_2))) = \underline{P}_1(f_1 \cdot 0) = 0$, where the first equality follows from Eq. (1). □

*Proof of Proposition 4.* (a) This follows from the equivalences (a1)⇔(a2)⇔(a3) and (b1)⇔(b2)⇔(b3) established in Lemma 14.

(b) Use the equivalence between the first and the third statements in (a).

(c) Use Lemma 15.

(d) This follows immediately using the equivalences (a1)⇔(a3) and (b1)⇔(b3) from Lemma 14, considering the particular cases of $A_1 \coloneqq \{x_1\}, A_2 \coloneqq \{x_2\}$. □

**Lemma 16.** *Consider marginal linear previsions $P_1, P_2$, and define $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$ by irrelevance. Assume $\mathcal{X}_1$ is finite. Then:*

*(a)* $P_1(P_2) = P_2(P_1)$.

*(b)* $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$ *are coherent.*

*(c)* $P_1(P_2) = P_2(P_1)$ *is the only independent product of* $P_1, P_2$.

*Proof.* (a) By Proposition 2(b), the smallest (and in this case, the only) coherent lower prevision with $\mathcal{X}_2$-marginal $P_2$ that is in addition coherent with $P_1(\cdot|\mathcal{X}_2)$ is given by the concatenation $P_2(P_1)$. Let us show that $P_2(P_1)$ is also coherent with $P_2(\cdot|\mathcal{X}_1)$. Since $\mathcal{X}_1$ is finite, this holds if and only if $P_2(P_1), P_2(\cdot|\mathcal{X}_1)$ satisfy (GBR). Consider thus a gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$, $x_1 \in \mathcal{X}_1$. It holds that $P_2(P_1(I_{x_1}(f - P_2(f|x_1)))) = P_2(P_1(I_{x_1}(g - P_2(g)))) = P_2((g - P_2(g))P_1(x_1)) = P_1(x_1)(P_2(g - P_2(g))) = 0$, where the gamble $g$ in $\mathcal{L}(\mathcal{X}_2)$ is given by $g(x_2) \coloneqq f(x_1, x_2)$. Thus, $P_2(P_1), P_2(\cdot|\mathcal{X}_1)$ are coherent, and applying Proposition 2(b), $P_2(P_1)$ must coincide with $P_1(P_2)$, which is the only linear prevision that is coherent with $P_2(\cdot|\mathcal{X}_1)$.

(b) Since $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$ are linear, they are coherent if and only if for every pair of gambles $f, g \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$, there exists some $B \in S_1(f) \cup S_2(g)$ such that [6]

$$\sup_{\omega \in B} [G_1(f|\mathcal{X}_2) + G_2(g|\mathcal{X}_1)](\omega) \geq 0. \tag{A.1}$$

Fix gambles $f, g$; in order to simplify the notation, let $h$ denote the gamble $G_1(f|\mathcal{X}_2) + G_2(g|\mathcal{X}_1)$. If $\sup h > 0$, then there must be some $B \in S_1(f) \cup S_2(g)$ such that $\sup_{\omega \in B} h(\omega) > 0$, so Eq. (A.1) holds. On the other hand, if $S_2(g) = \emptyset$, then Eq. (A.1) follows from the separate coherence of $P_1(\cdot|\mathcal{X}_2)$.

Assume next that $S_2(g) \neq \emptyset$ and that $\sup_B h < 0$ for every $B \in S_1(f) \cup S_2(g)$, whence $\sup h \leq 0$. Consider $x_1 \in S_2(g)$, and denote $B_{x_1} := \{x_1\} \times \mathcal{X}_2$. Consider also the linear prevision $P := P_1(P_2) = P_2(P_1)$. It follows from the first statement and Theorem 3(a) that this linear prevision is coherent with each of $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$.

If $\sup_{\omega \in B_{x_1}} h(\omega) = -\delta < 0$, then $P(h) = P(B_{x_1} h) + P(B_{x_1}^c h) \leq -\delta P(B_{x_1})$, using that $h \leq -\delta$ on $B_{x_1}$ and $h \leq 0$ on $B_{x_1}^c$. On the other hand, the coherence of $P, P_1(\cdot|\mathcal{X}_2)$ implies, via Proposition 1, that $P(G_1(f|\mathcal{X}_2)) \geq 0$, and similarly the coherence of $P, P_2(\cdot|\mathcal{X}_1)$ implies that $P(G_2(g|\mathcal{X}_1)) \geq 0$. Applying the linearity of $P$, we deduce that $P(G_1(f|\mathcal{X}_2) + G_2(g|\mathcal{X}_1)) = P(G_1(f|\mathcal{X}_2)) + P(G_2(g|\mathcal{X}_1)) \geq 0$. We conclude that

$$P(B_{x_1}) = P_1(x_1) = 0 \text{ for every } x_1 \in S_2(g). \tag{A.2}$$

Consider now $x_2 \in S_1(f)$, and let $\sup_{\mathcal{X}_1 \times \{x_2\}} h = -\varepsilon < 0$. The equality $P(G_1(f|x_2)) = 0$ implies that there must be some $x_1 \in \mathcal{X}_1$ such that $G_1(f|x_2)(x_1, x_2) \geq -\frac{\varepsilon}{2}$, and we can assume without loss of generality that $P_1(x_1) > 0$, because $0 = P(G_1(f|x_2)) = \sum_{x_1 \in \mathcal{X}_1, P_1(x_1) > 0, G_1(f|x_2)(x_1,x_2) \geq -\frac{\varepsilon}{2}} P_1(x_1)(G_1(f|x_2)(x_1, x_2)) + \sum_{x_1 \in \mathcal{X}_1, P_1(x_1) > 0, G_1(f|x_2)(x_1,x_2) < -\frac{\varepsilon}{2}} P_1(x_1)(G_1(f|x_2)(x_1, x_2))$.

Fix this $x_1$. Since $h(x_1, x_2) = [G_1(f|\mathcal{X}_2) + G_2(g|\mathcal{X}_1)](x_1, x_2) \leq -\varepsilon < -\frac{\varepsilon}{2} \leq G_1(f|x_2)(x_1, x_2)$, it must be $G_2(g|\mathcal{X}_1)(x_1, x_2) < 0$, whence $x_1 \in S_2(g)$. Thus, we have found an element in the support $S_2(g)$ with positive probability. This contradicts Eq. (A.2). As a consequence, $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$ are coherent.

(c) By Proposition 2(b), $P_1(P_2)$ is the only coherent lower prevision that is coherent with $P_2(\cdot|\mathcal{X}_1)$ and has marginal $P_1$, and $P_2(P_1)$ is the only coherent lower prevision that is coherent with $P_1(\cdot|\mathcal{X}_2)$ and has marginal $P_2$. Since by the first statement $P_1(P_2) = P_2(P_1) =: P$, we deduce that this linear prevision is the only one to be weakly coherent with $P_2(\cdot|\mathcal{X}_1), P_1(\cdot|\mathcal{X}_2)$. Applying the second statement, we conclude that $P, P_2(\cdot|\mathcal{X}_1), P_1(\cdot|\mathcal{X}_2)$ are coherent and that $P$ is the only independent product. □

*Proof of Proposition 5.* (a) Consider $P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2$. By Lemma 16, $P_1(P_2), P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$ are coherent. Since $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are the lower envelopes of the family $\{P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$, the envelope theorem in [24, Theorem 7.1.6] implies that $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent with $\underline{P} := \inf\{P_1(P_2) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$.

(b) Applying the reduction theorem [24, Theorem 7.1.5], $\underline{P}$ is an independent product of its marginals if and only if $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent and $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly coherent. The first of these two conditions follows from the first statement. The equivalence between the statement in (b) and weak coherence follows from Theorem 3. □

*Proof of Proposition 6.* Let us make a circular proof.

(a)⇒(b) Since by Proposition 5(a) the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2)$ and $\underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent, it follows from the reduction theorem in [24, Theorem 7.1.5] that $\underline{P}$ is an independent product of its marginals $\underline{P}_1, \underline{P}_2$ if and only if $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_1(\cdot|\mathcal{X}_1)$ are weakly coherent. By [16, Theorem 1], this weak coherence is equivalent to the pairwise coherence of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ on the one hand, and of $\underline{P}, \underline{P}_2(\cdot|\mathcal{X}_1)$ on the other. By Proposition 4, if $\underline{P}$

[6]The equivalence between this condition and Eq. (2) follows from the fact that for any conditional linear prevision $P(\cdot|\mathcal{B})$ it holds that $G(-f|\mathcal{B}) = -G(f|\mathcal{B})$, and also $G(f|\mathcal{B}) + G(g|\mathcal{B}) = G(f + g|\mathcal{B})$. Recall also that we are using in this paper Walley's approach, whence a linear conditional prevision is conglomerable, unlike those considered by de Finetti in [14].

is productive, it satisfies (GBR) with both these conditional lower previsions. Since $\mathcal{X}_1$ is finite, we deduce in particular that $\underline{P}, \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent. Thus, the only thing we need to show is that $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ satisfy (CNG), that is, that for every gamble $f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ it holds that $\underline{P}(G_1(f|\mathcal{X}_2)) \geq 0$. Fix then $f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$.

Assume first that the set of gambles $\{f(\cdot, x_2) : x_2 \in \mathcal{X}_2\} \subseteq \mathcal{L}(\mathcal{X}_1)$ is finite and of size $n$. Let us denote these gambles by $g_1, \dots, g_n \in \mathcal{L}(\mathcal{X}_1)$. Then we can define a partition of $\mathcal{X}_2$ by the sets $B_1, \dots, B_n$, where $B_i := \{x_2 \in \mathcal{X}_2 : f(\cdot, x_2) = g_i\}$. As a consequence, $G_1(f|\mathcal{X}_2)(x_1, x_2) = \sum_{i=1}^n I_{B_i}(x_2)(g_i(x_1) - \underline{P}_1(g_i)) \ \forall x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$. Applying the super-additivity and productivity of $\underline{P}$, we deduce that

$$\underline{P}(G_1(f|\mathcal{X}_2)) = \underline{P}\left(\sum_{i=1}^n I_{B_i}(g_i - \underline{P}_1(g_i))\right) \geq \sum_{i=1}^n \underline{P}(B_i(g_i - \underline{P}_1(g_i))) \geq 0.$$

Next, given an arbitrary gamble $f$, there exists some natural number $k$ such that $-k \leq \inf f \leq \sup f \leq k$. For every $n$, there exists a finite partition $\mathcal{B}_n := \{I_1^n, \dots, I_{k_n}^n\}$ of $[-k, k]$ with intervals $I_j^n$ of length smaller than $\frac{1}{n}$. Let us define the gamble $f_n$ on $\mathcal{X}_1 \times \mathcal{X}_2$ by $f_n(x_1, x_2) := \inf I_j^n$ if $f(x_1, x_2) \in I_j^n$, for every $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. Then since the partition $\mathcal{B}_n$ is finite, the set $\{f_n(\cdot, x_2) : x_2 \in \mathcal{X}_2\} \subseteq \mathcal{L}(\mathcal{X}_1)$ is finite. Moreover, by construction $|f_n(x_1, x_2) - f(x_1, x_2)| \leq \frac{1}{n}$ for every $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$, from which we deduce that $|G_1(f_n|\mathcal{X}_2) - G_1(f|\mathcal{X}_2)| \leq \frac{2}{n}$ for all $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. We see then that the sequence $(G_1(f_n|\mathcal{X}_2))_{n \in \mathbb{N}}$ converges uniformly towards $G_1(f|\mathcal{X}_2)$. By [24, Theorem 2.6.1($\ell$)], $\underline{P}(G_1(f|\mathcal{X}_2)) = \lim_{n \to \infty} \underline{P}(G_1(f_n|\mathcal{X}_2)) \geq 0$. This implies that $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ are coherent and as a consequence that $\underline{P}$ is an independent product.

(b)⇒(c) If $\underline{P}$ is an independent product, it follows from the characterisation of coherence in [24, Theorem 6.5.2] that for every gamble $f$ it must be $\underline{P}(f) \geq \min \underline{P}_2(f|\mathcal{X}_1)$ and $\underline{P}(f) \geq \inf \underline{P}_1(f|\mathcal{X}_2)$.

(c)⇒(a) Consider $A_1 \subseteq \mathcal{X}_1$, a gamble $f_2 \in \mathcal{X}_2$ and let $I_{A_1}(f_2 - \underline{P}_2(f_2)) \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$. Then $\underline{P}(I_{A_1}(f_2 - \underline{P}_2(f_2))) \geq \min \underline{P}_2(I_{A_1}(f_2 - \underline{P}_2(f_2))|\mathcal{X}_1) = \min[I_{A_1}\underline{P}_2(f_2 - \underline{P}_2(f_2))] = 0$, where the penultimate passage follows from Eq. (1). Similarly, $\underline{P}(I_{A_1^c}(f_2 - \underline{P}_2(f_2))) \geq 0$ and since $0 = \underline{P}(f_2 - \underline{P}_2(f_2)) \geq \underline{P}(I_{A_1}(f_2 - \underline{P}_2(f_2))) + \underline{P}(I_{A_1^c}(f_2 - \underline{P}_2(f_2)))$, it must be $\underline{P}(I_{A_1}(f_2 - \underline{P}_2(f_2))) = 0$. An analogous reasoning allows us to deduce that $\underline{P}(I_{A_2}(f_1 - \underline{P}_1(f_1))) = 0$ for every $A_2 \subseteq \mathcal{X}_2$ and every gamble $f_1$ on $\mathcal{X}_1$. Applying Proposition 4(a) we deduce that $\underline{P}$ is productive. □

*Proof of Proposition 7.* First of all, by Proposition 5(b) the independent natural extension $\underline{P}_1 \otimes \underline{P}_2$ is the smallest coherent lower prevision that dominates the two concatenations $\underline{P}_1(\underline{P}_2), \underline{P}_2(\underline{P}_1)$. Hence, it coincides with one of them if and only if this concatenation dominates the other one.

On the other hand, Proposition 6 implies that a concatenation is an independent product if and only if it is productive, and by Proposition 5 it is an independent product if and only if it coincides with the independent natural extension.

Let us establish now the remaining equivalences.

(a) If $\underline{P}_2$ is linear, then $\underline{P}_1(P_2)$ is the only coherent lower prevision with marginal $\underline{P}_1$ that is coherent with $P_2(\cdot|\mathcal{X}_1)$, and as a consequence it coincides with the independent natural extension.

If $\underline{P}_1$ is vacuous, then it is easy to prove that $\underline{P}_2(f|x_1) \geq \underline{P}_2(\underline{P}_1(f))$ for every gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$ and every $x_1 \in \mathcal{X}_1$. This implies that $\inf_{x_1 \in \mathcal{X}_1} \underline{P}_2(f|x_1) \geq \underline{P}_2(\underline{P}_1(f|\mathcal{X}_2))$, or, in other words, that $\underline{P}_1(\underline{P}_2(f)) \geq \underline{P}_2(\underline{P}_1(f))$. Applying Proposition 5(b) (which is where the finiteness of $\mathcal{X}_1$ comes into play), we obtain that $\underline{P}_1(\underline{P}_2)$ is an independent product. That it is the smallest one follows trivially because if there were a smaller one, then it could not dominate $\underline{P}_1(\underline{P}_2)$. We deduce that $\underline{P}_1(\underline{P}_2)$ coincides with the independent natural extension.

We turn now to the other implication. Assume ex-absurdo that $\underline{P}_1$ is not vacuous and $\underline{P}_2$ is non-linear. Then there are subsets $A_1 \subseteq \mathcal{X}_1, A_2 \subseteq \mathcal{X}_2$ such that $\underline{P}_1(A_1) \in (0, 1)$ [11, Theorem 8] (because $\underline{P}_1$ is non-vacuous) and $\underline{P}_2(A_2) < \overline{P}_2(A_2)$ [24, Corollary 3.2.3] (because $\underline{P}_2$ is non-linear). Then Eq. (1) implies that $G_1(I_{A_1 \times A_2}|\mathcal{X}_2) = I_{A_2}(I_{A_1} - \underline{P}_1(A_1)) = I_{A_2}G_1(A_1)$, and it is easy to prove that $\underline{P}_2(I_{A_2}G_1(A_1)|\mathcal{X}_1) = G_1(A_1)(\underline{P}_2(A_2)I_{A_1} + \overline{P}_2(A_2)I_{A_1^c})$. Taking this into account, we deduce that

$$\underline{P}_1(\underline{P}_2(G_1(I_{A_1 \times A_2}|\mathcal{X}_2))|\mathcal{X}_1) \leq P_1(G_1(A_1)(\underline{P}_2(A_2)I_{A_1} + \overline{P}_2(A_2)I_{A_1^c})),$$

where $P_1 \geq \underline{P}_1$ is a linear prevision satisfying $P_1(A_1) = \underline{P}_1(A_1) \in (0,1)$, so that $G_1(A_1) = A_1 - \underline{P}_1(A_1) = A_1 - P_1(A_1)$. Since $P_1(G_1(A_1)(\underline{P}_2(A_2)I_{A_1} + \overline{P}_2(A_2)I_{A_1^c})) = P_1(A_1)(1-P_1(A_1))(\underline{P}_2(A_2) - \overline{P}_2(A_2)) < 0$, we deduce that $\underline{P}_1(\underline{P}_2(G_1(I_{A_1\times A_2}|\mathcal{X}_2))) < 0$ and therefore $\underline{P}_1(\underline{P}_2)$ is not coherent with $\underline{P}_1(\cdot|\mathcal{X}_2)$. Thus, it is not an independent product and in particular it does not coincide with the independent natural extension.

(b) The proof is analogous to that of the first statement. □

*Proof of Proposition 8.* (a) This is actually point (a) of Proposition 5.

(b) Let us show that the strong product $\underline{P}_1 \boxtimes \underline{P}_2$ is factorising. Consider gambles $f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$, and let us show that $(\underline{P}_1 \boxtimes \underline{P}_2)(f_1f_2) = \underline{P}_1(f_1\underline{P}_2(f_2))$ (the other case is similar). Since from the first statement the strong product is an independent product, it dominates $\underline{P}_1(\underline{P}_2)$, and consequently $(\underline{P}_1 \boxtimes \underline{P}_2)(f_1f_2) \geq \underline{P}_1(\underline{P}_2(f_1f_2)) = \underline{P}_1(f_1\underline{P}_2(f_2))$, where the equality follows from Eq. (1). Consider on the other hand a linear prevision $P_2 \geq \underline{P}_2$ such that $P_2(f_2) = \underline{P}_2(f_2)$, and another linear prevision $P_1 \geq \underline{P}_1$ such that $P_1(f_1\underline{P}_2(f_2)) = \underline{P}_1(f_1\underline{P}_2(f_2))$. Then by construction $(\underline{P}_1 \boxtimes \underline{P}_2)(f_1f_2) \leq P_1(P_2(f_1f_2)) = P_1(f_1P_2(f_2)) = P_1(f_1\underline{P}_2(f_2)) = \underline{P}_1(f_1\underline{P}_2(f_2))$, whence $(\underline{P}_1 \boxtimes \underline{P}_2)(f_1f_2) = \underline{P}_1(f_1\underline{P}_2(f_2))$. Thus, $\underline{P}_1 \boxtimes \underline{P}_2$ is factorising.

Next, assume that $\underline{P}_1 \otimes \underline{P}_2 \leq \underline{P} \leq \underline{P}_1 \boxtimes \underline{P}_2$. Consider gambles $f_1 \in \mathcal{L}^+(\mathcal{X}_1), f_2 \in \mathcal{L}(\mathcal{X}_2)$, and let us show that $\underline{P}(f_1f_2) = \underline{P}_1(f_1\underline{P}_2(f_2))$ (the other case is similar). Then $\underline{P}(f_1f_2) \geq (\underline{P}_1 \otimes \underline{P}_2)(f_1f_2) \geq \underline{P}_1(\underline{P}_2(f_1f_2)) = \underline{P}_1(f_1\underline{P}_2(f_2))$, using again Eq. (1), and on the other hand $\underline{P}(f_1f_2) \leq (\underline{P}_1 \boxtimes \underline{P}_2)(f_1f_2) = \underline{P}_1(f_1\underline{P}_2(f_2))$, taking into account that $\underline{P}_1 \boxtimes \underline{P}_2$ is factorising. Thus, $\underline{P}(f_1f_2) = \underline{P}_1(f_1\underline{P}_2(f_2))$, and from this we deduce that $\underline{P}$ is factorising. □

*Proof of Proposition 9.* Consider a gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$. Assume for instance that $\underline{P}_2$ is vacuous, so that $\underline{P}_1 \otimes \underline{P}_2 = \underline{P}_2(\underline{P}_1)$ by Proposition 7. Then for every $\varepsilon > 0$ there exists some $x_2 \in \mathcal{X}_2$ such that $\underline{P}_1(f|x_2) \leq \inf_{x_2\in\mathcal{X}_2} \underline{P}_1(f|x_2) + \varepsilon = \underline{P}_2(\underline{P}_1(f)) + \varepsilon$.

Let $P_2$ be the linear prevision associated with the degenerate distribution on $x_2$. Then $P_2 \in \mathcal{M}(\underline{P}_2)$ because $\underline{P}_2$ is vacuous. Take $P_1 \in \mathcal{M}(\underline{P}_1)$ such that $P_1(f(\cdot, x_2)) = \underline{P}_1(f(\cdot, x_2))$. Then by definition of the strong product $(\underline{P}_1 \boxtimes \underline{P}_2)(f) \leq P_2(P_1(f|\mathcal{X}_2)) = P_1(f(\cdot, x_2)) = \underline{P}_1(f|x_2) \leq \underline{P}_2(\underline{P}_1(f)) + \varepsilon = (\underline{P}_1 \otimes \underline{P}_2)(f) + \varepsilon$. Since this holds for every $\varepsilon > 0$ we deduce that $\underline{P}_1 \boxtimes \underline{P}_2 \leq \underline{P}_1 \otimes \underline{P}_2$ and as a consequence they are equal. The proof of the equality when $\underline{P}_1$ is vacuous is analogous.

Finally, the equality when one of the marginals (say, $\underline{P}_1$) is linear follows from Proposition 2(b). □

*Proof of Proposition 10.* Let $\mathcal{M}_i$ denote the set of exposed points of $\mathcal{M}(\underline{P}_i), i = 1, 2$. It follows from Straszewicz's theorem [21] that $\mathcal{M}(\underline{P}_i)$ equals the closed convex hull of $\mathcal{M}_i$ for $i = 1, 2$, whence $\underline{P}_i = \inf \mathcal{M}_i$ for $i = 1, 2$ and as a consequence

$$\underline{P}_1 \boxtimes \underline{P}_2 = \inf\{P_1(P_2) : P_i \in \mathcal{M}(\underline{P}_i)\} = \inf\{P_1(P_2) : P_i \in \mathcal{M}_i\}. \tag{A.3}$$

What we want to show is that $\underline{P} \leq P_1(P_2)$ for every $P_1 \in \mathcal{M}_1, P_2 \in \mathcal{M}_2$. Consider then $P_i \in \mathcal{M}_i, i = 1, 2$. By definition of exposed point there is some gamble $f_i \in \mathcal{L}(\mathcal{X}_i)$ such that $\underline{P}_i(f_i) = P_i(f_i) < P_i'(f_i)$ for every $P_i' \neq P_i$ in $\mathcal{M}(\underline{P}_i)$. Moreover, using the constant additivity of coherent lower previsions established in [24, Thm. 2.6.1(c)], we can assume without loss of generality that $\min f_i > 0$.

If $\underline{P}$ is a factorising lower prevision, then it must satisfy $\underline{P}(f_1f_2) = \underline{P}(f_1\underline{P}(f_2)) = \underline{P}_1(f_1)\underline{P}_2(f_2)$ for any pair of non-negative gambles $f_1, f_2$, and as a consequence there must be some linear prevision $P \geq \underline{P}$ such that $P(f_1f_2) = \underline{P}_1(f_1)\underline{P}_2(f_2) = P_1(f_1)P_2(f_2)$. Let us show that it must be $P = P_1(P_2)$.

Since $\mathcal{X}_1, \mathcal{X}_2$ are finite, we can write $P = P(P(\cdot|\mathcal{X}_1))$, where $P(\cdot|x_1)$ is derived from $P$ by Bayes' rule when $P(x_1) > 0$ and can be defined arbitrarily otherwise. Moreover, since $P \geq \underline{P}$ and the latter is an independent product because it is factorising, it follows that $P(f|x_1) \geq \underline{P}(f|x_1) = \underline{P}_2(f|x_1) = \underline{P}_2(f(x_1, \cdot))$ for every $f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ and every $x_1 \in \mathcal{X}_1$, where the first equality holds because $\underline{P}$ is coherent with $\underline{P}_2(\cdot|\mathcal{X}_1)$. As a consequence, $P(f_1f_2) = P(P(f_1f_2|\mathcal{X}_1)) = P(f_1P(f_2|\mathcal{X}_1)) \geq P(f_1\underline{P}(f_2)) = P(f_1)\underline{P}(f_2) = \underline{P}(f_1)\underline{P}(f_2)$, and we can only have the equality provided that $P(f_2|x_1) = P_2(f_2)$ whenever $P(x_1) > 0$, because $\min f_1 > 0$. But since $P(\cdot|x_1) \in \mathcal{M}(\underline{P}_2)$, the equality above can only hold provided that $P(\cdot|x_1) = P_2$ for every $x_1 \in \mathcal{X}_1$ with $P_1(x_1) > 0$. Since we can define $P(\cdot|x_1)$ arbitrarily when $P(x_1) = 0$, we deduce that $P(\cdot|\mathcal{X}_1)$ must be constant on $P_2$.

A similar reasoning shows that $P(\cdot|\mathcal{X}_2)$ must be constant on $P_1$, whence $P_1$ must be the $\mathcal{X}_1$-marginal of $P$.

From this we deduce that $P = P(P(\cdot|\mathcal{X}_1)) = P_1(P_2)$. Therefore, $\underline{P} \leq P_1(P_2)$ for every $P_1 \in \mathcal{M}_1, P_2 \in \mathcal{M}_2$, and applying Eq. (A.3) we deduce that $\underline{P} \leq \underline{P}_1 \boxtimes \underline{P}_2$. □

*Proof of Proposition 11.* (a) By [24, Section 7.3.7], $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent whenever any of them is vacuous. If for instance $\underline{P}_1$ is vacuous, then we can reason as in the proof of Proposition 7(a) and establish that $\underline{P}_1(\underline{P}_2) \geq \underline{P}_2(\underline{P}_1)$. Then Theorem 3 and the reduction theorem in [24, Theorem 7.1.5] imply that $\underline{P}_1(\underline{P}_2)$ is an independent product of $\underline{P}_1, \underline{P}_2$.

(b) This is Proposition 5(a).

(c) If $\mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1)) \neq \emptyset$, then it follows from Theorem 3 that the lower envelope $\underline{P}$ of $\mathcal{M}(\underline{P}_1(\underline{P}_2)) \cap \mathcal{M}(\underline{P}_2(\underline{P}_1)) \neq \emptyset$ is a coherent lower prevision that is weakly coherent with $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$. Moreover, by construction the marginals of $\underline{P}$ dominate $\underline{P}_1, \underline{P}_2$, whence $\underline{P}(x_1) > 0, \underline{P}(x_2) > 0$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$. In that case, we can apply [15, Theorem 11] to deduce that $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent, and in particular that $\underline{P}_1, \underline{P}_2$ are compatible. □

Next we state an extension of a couple of results from Walley's theory ([24, Theorem 7.1.5] and [16, Theorem 1]) to Williams' notion of coherence. The proof is analogous to those of these results.

**Proposition 17.** *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ and consider conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$.*

*(a) $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are W-coherent $\Leftrightarrow$ $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are W-coherent and $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly W-coherent.*

*(b) $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are weakly W-coherent if and only if $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ and $\underline{P}, \underline{P}_2(\cdot|\mathcal{X}_1)$ are both W-coherent.*

*Proof of Lemma 12.* Consider first of all the case of linear $P_1, P_2$. They are W-coherent if and only if for every $g, h \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with finite support it holds that

$$\sup_{S_1(g) \cup S_2(h)} G_1(g|\mathcal{X}_2) + G_2(h|\mathcal{X}_1) \geq 0. \tag{A.4}$$

It is quite immediate to show that the linear prevision $P_1(P_2)$ satisfies (GBR) with both $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$. As a consequence,

$$P_1(P_2(G_1(g|\mathcal{X}_2) + G_2(h|\mathcal{X}_1))) = \sum_{x_2 \in S_1(g)} P_1(P_2(G_1(g|x_2))) + \sum_{x_1 \in S_2(h)} P_1(P_2(G_2(h|x_1))) \geq 0. \tag{A.5}$$

Assume that the supremum in (A.4) is equal to $-\delta < 0$. This means that $G_1(g|\mathcal{X}_2) + G_2(h|\mathcal{X}_1) \leq -\delta$ over $S_1(g) \cup S_2(h)$ and it equals zero elsewhere; whence the only possibility for (A.5) to hold is that $P_1(P_2(S_1(g) \cup S_2(h))) = 0$, whence $P_2(S_1(g)) = 0 = P_1(S_2(h))$. For a given (fixed) $x_1 \in S_2(h)$, the separate coherence of $P_2(\cdot|\mathcal{X}_1)$ implies that the set $A_{x_1} := \{x_2 \in \mathcal{X}_2 : G_2(h|\mathcal{X}_1)(x_1, x_2) \geq -\frac{\delta}{2}\}$ is non-empty. Let us show that any $x_2 \in A_{x_1}$ must belong to $S_1(g)$ too. Assume by contradiction that there is $x_2 \in A_{x_1} \setminus S_1(g)$. Then $G_1(g|\mathcal{X}_2)(x_1, x_2) = 0$ while $G_2(h|\mathcal{X}_1)(x_1, x_2) \geq -\frac{\delta}{2}$, since $x_1 \in S_2(h)$ and $x_2 \in A_{x_1}$. From this we obtain that the supremum in (A.4) is greater than or equal to $-\frac{\delta}{2}$, which contradicts our assumption that it is equal to $-\delta$. We deduce that any $x_2 \in A_{x_1}$ belongs to $S_1(g)$, whence $P_2(A_{x_1}) = 0$. Define now the gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$ by $f(x_1', x_2') := h(x_1, x_2') \; \forall (x_1', x_2') \in \mathcal{X}_1 \times \mathcal{X}_2$. This gamble is $\mathcal{X}_2$-measurable, so $P_2(f) = P(h(x_1, \cdot)) = P(h|x_1)$. Moreover, $0 = P_2(f) - P_2(f) = P_2(f - P_2(h|x_1)) = P(f - P_2(h|x_1)) = P(I_{A_{x_1}^c}(G_2(h|x_1))) \leq -\frac{\delta}{2} P(A_{x_1}^c) = -\frac{\delta}{2} < 0$, which is a contradiction. As a consequence, the supremum in (A.4) is non-negative and $P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1)$ are W-coherent.

Now, given coherent lower previsions $\underline{P}_1, \underline{P}_2$, the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ they induce by irrelevance are the lower envelopes of the family of conditional linear previsions $\{P_1(\cdot|\mathcal{X}_2), P_2(\cdot|\mathcal{X}_1) : P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2\}$. Since W-coherence is preserved by taking lower envelopes [25, Theorem 2], we deduce that the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent. □

*Proof of Proposition 13.* (a) From Proposition 17 and Lemma 12 $\underline{P}$ is a W-independent product of its marginals $\underline{P}_1, \underline{P}_2$ if and only if $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ are W-coherent and $\underline{P}, \underline{P}_2(\cdot|\mathcal{X}_1)$ are W-coherent. By [22, p. 363], this is equivalent to $\underline{P}$ satisfying (GBR) with both $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$.

(b) Let us show that $\underline{P}$ is the smallest coherent lower prevision that satisfies (GBR) with both $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$.

The coherence of $\underline{P}$ follows from its definition and from the separate coherence of $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$. Let us prove next that $\underline{P}$ satisfies (GBR) with $\underline{P}_1(\cdot|\mathcal{X}_2)$ (the proof for $\underline{P}_2(\cdot|\mathcal{X}_1)$ is similar). Consider a gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$ and $x_2 \in \mathcal{X}_2$. Then by choosing $g := I_{x_2} f$, $h := 0$ in Eq. (5) we obtain that $\underline{P}(G_1(f|x_2)) \geq 0$. If it were $\underline{P}(G_1(f|x_2)) > 0$, then there would be some $\mu > 0$ and gambles $g, h$ with finite support such that $G_1(f|x_2) - \mu \geq G_1(g|\mathcal{X}_2) + G_2(h|\mathcal{X}_1)$, a contradiction with the W-coherence of $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ that has been established in Lemma 12.

Consider now another coherent lower prevision $\underline{Q}$ that satisfies (GBR) with both $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$. If there is some gamble $f$ on $\mathcal{X}_1 \times \mathcal{X}_2$ and some $\varepsilon > 0$ such that $\underline{Q}(f) + \varepsilon < \underline{P}(f)$, then by Eq. (5) there are gambles $g, h$ with finite supports such that $f - \underline{Q}(f) - \varepsilon \geq G_1(g|\mathcal{X}_2) + G_2(h|\mathcal{X}_1)$, whence $-\varepsilon \geq G_1(g|\mathcal{X}_2) + G_2(h|\mathcal{X}_1) - (f - \underline{Q}(f))$, and thus $\underline{Q}$ is not W-coherent with $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$, a contradiction. Thus, $\underline{P}$ is the W-independent natural extension of $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$.

(c) This is a consequence of the first statement and point (d) in Proposition 4. □

**Appendix B. Conglomerably coherent independent products**

In [18], we argued that Walley's definition of coherence may be too weak to fully capture the behavioural implications of given assessments, and that if we accept Walley's conglomerative principle we should use the notion of conglomerable coherence instead. In this appendix, we study what kind of independent products are originated by such a more stringent notion of coherence. In order to do this, we must first give a number of preliminary notions. We start with a notion of coherence, originally put forward by Williams [25], for sets of so-called *desirable gambles*:

*Definition* 11 (**Coherence for gambles**). Let $\mathcal{R} \subseteq \mathcal{L}(\Omega)$ be a set of gambles. We consider the following rationality axioms for desirability:

D1. $\mathcal{L}^+(\Omega) \subseteq \mathcal{R}$.

D2. $0 \notin \mathcal{R}$.

D3. $f \in \mathcal{R}, \lambda > 0 \Rightarrow \lambda f \in \mathcal{R}$.

D4. $f, g \in \mathcal{R} \Rightarrow f + g \in \mathcal{R}$.

A set of gambles $\mathcal{R}$ satisfying these four axioms is called *coherent*, and its elements *desirable*.

It follows from these axioms that if $f$ belongs to a coherent set of gambles $\mathcal{R}$ and $g \geq f$, then it must be $g \in \mathcal{R}$. One particular coherent set of desirable gambles is the set of *strictly desirable gambles* induced by a coherent lower prevision $\underline{P}$, which is given by

$$\mathcal{R} := \mathcal{L}^+ \cup \{f \in \mathcal{L} : \underline{P}(f) > 0\}.$$

Walley developed his theory on top of Williams', in particular by introducing a further axiom, besides D1–D4, which has to do with a property of conglomerability of an uncertainty model:

*Definition* 12 (**Conglomerability for gambles**). Let $\mathcal{R}$ be a coherent set of desirable gambles and $\mathcal{B}$ a partition of $\Omega$. $\mathcal{R}$ is called *$\mathcal{B}$-conglomerable* when it satisfies the following axiom:

D5. $f \in \mathcal{L}, Bf \in \mathcal{R} \cup \{0\}$ for all $B \in \mathcal{B} \Rightarrow f \in \mathcal{R} \cup \{0\}$.

Observe that D5 is a consequence of D4 when $\mathcal{B}$ is finite.

The notion of conglomerability is at the basis of the following definition [19]:

*Definition* 13 (**Conglomerable natural extension for gambles**). Given a set of desirable gambles $\mathcal{R}$ and a partition $\mathcal{B}$ of $\Omega$, the *$\mathcal{B}$-conglomerable natural extension* of $\mathcal{R}$, if it exists, is the smallest set $\mathcal{F}$ that contains $\mathcal{R}$ and satisfies D1–D5.

Now we focus on a collection of conditional lower previsions $\underline{P}(\cdot|\mathcal{B}_1), \dots, \underline{P}(\cdot|\mathcal{B}_m)$, where $\mathcal{B}_1, \dots, \mathcal{B}_m$ are partitions of $\Omega$. A set of desirable gambles $\mathcal{R}$ induces a conditional lower prevision $\underline{P}(\cdot|\mathcal{B}_i)$ on $\mathcal{L}(\Omega)$ by means of the formula

$$\underline{P}(f|B_i) \coloneqq \sup\{\mu : B_i(f - \mu) \in \mathcal{R}\}, \tag{B.1}$$

whenever $f \in \mathcal{L}(\Omega)$ and $B_i \in \mathcal{B}_i$.

*Definition* 14 (**Conglomerable coherence for lower previsions**). Let $\underline{P}(\cdot|\mathcal{B}_1), \dots, \underline{P}(\cdot|\mathcal{B}_m)$ be conditional lower previsions. They are called *conglomerably coherent* if they can be induced through (B.1) from a coherent set of desirable gambles that is conglomerable w.r.t. all the partitions $\mathcal{B}_1, \dots, \mathcal{B}_m$.

It is argued in [18] that the notion of conglomerable coherence, which is more restrictive than Walley's notion of coherence from Definition 3, is the one we should use when investigating the consistency of a number of conditional lower previsions in case we accept Walley's conglomerative principle. In this sense, we should talk of *conglomerably coherent independent products*: these are the coherent lower previsions $\underline{P}$ on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ that satisfy conglomerable coherence with the conditional lower previsions $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ that we can derive from the marginals of $\underline{P}$ using the mutual irrelevance of $X_1$ and $X_2$.

The next result shows that, when one of the spaces is finite, conglomerable coherence reduces to coherence to the extent of definining independent products: this is to say that restricting the attention to Walley's coherence, as we have done in the main body of this paper, is enough to cover also the extension to conglomerable coherence:

**Theorem 18.** *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with marginals $\underline{P}_1, \underline{P}_2$. Assume that $\mathcal{X}_1$ is finite, and define $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ by epistemic irrelevance. Then:*

$$\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1) \text{ conglomerably coherent} \Leftrightarrow \underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1) \text{ coherent.}$$

*Proof.* The direct implication follows taking into account that conglomerable coherence is stronger than coherence, by [18, Theorem 9]. Let us prove the converse. Let $\mathcal{R}, \mathcal{R}_1, \mathcal{R}_2$ denote the sets of strictly desirable gambles associated with $\underline{P}, \underline{P}_1, \underline{P}_2$, respectively. Then it follows from the assumption of epistemic irrelevance that the conglomerable natural extensions (with respect to $\mathcal{X}_2$ and $\mathcal{X}_1$, respectively) of the sets of strictly desirable gambles associated with $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are given by

$$\mathcal{F}_1 := \{0 \neq f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2) : f(\cdot, x_2) \in \mathcal{R}_1 \cup \{0\} \; \forall x_2 \in \mathcal{X}_2\}, \tag{B.2}$$

$$\mathcal{F}_2 := \{0 \neq f \in \mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2) : f(x_1, \cdot) \in \mathcal{R}_2 \cup \{0\} \; \forall x_1 \in \mathcal{X}_1\}. \tag{B.3}$$

It has been established in [18, Theorem 10] that the set of gambles $\mathcal{R} \oplus \mathcal{F}_1 \oplus \mathcal{F}_2$ induces $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$, and that it is a coherent set. Since $\mathcal{X}_1$ is finite, it only remains to prove that this set is $\mathcal{X}_2$-conglomerable to deduce that $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are conglomerably coherent.

Let $f \neq 0$ be a gamble on $\mathcal{X}_1 \times \mathcal{X}_2$ such that $I_{x_2} f$ belongs to $\mathcal{R} \oplus \mathcal{F}_1 \oplus \mathcal{F}_2 \cup \{0\}$ for every $x_2 \in \mathcal{X}_2$. Consider $x_2$ such that $I_{x_2} f \neq 0$. Then, there are gambles $h \in \mathcal{R} \cup \{0\}, h_1 \in \mathcal{F}_1 \cup \{0\}, h_2 \in \mathcal{F}_2 \cup \{0\}$ such that $I_{x_2} f = h + h_1 + h_2$. Note that by construction $h_1 \geq G_1(h_1|\mathcal{X}_2), h_2 \geq G_2(h_2|\mathcal{X}_1)$, and then the coherence of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ implies that $\underline{P}(h_1) \geq 0, \underline{P}(h_2) \geq 0$.

Let us consider a number of possibilities:

(a) If $\underline{P}(h) > 0$, then $\underline{P}(I_{x_2} f) \geq \underline{P}(h) + \underline{P}(h_1) + \underline{P}(h_2) > 0$, whence $I_{x_2} f \in \mathcal{R}$.

(b) If (a) does not hold, then it must be $h \geq 0$. In that case, we can redefine $h_1 := h_1 + h \in \mathcal{F}_1 \cup \{0\}$ and we obtain that $I_{x_2} f = h_1 + h_2 \in \mathcal{F}_1 \oplus \mathcal{F}_2$.

   Let us define $A_1 \coloneqq \{x_1 \in \mathcal{X}_1 : h_2(x_1, \cdot) \geq 0\}$. We can assume without loss of generality that $h_2 = h_2 I_{A_1^c}$; otherwise, we can redefine $h_1 := h_1 + h_2 I_{A_1} \geq h_1 \in \mathcal{F}_1$ and $h_2 := h_2 I_{A_1^c} \in \mathcal{F}_2 \cup \{0\}$. To see the latter, consider that $h_2(x_1, \cdot) \in \mathcal{R}_2 \cup \{0\}$ for all $x_1 \in \mathcal{X}_1$; when we turn some of these $h_2(x_1, \cdot)$-gambles into zero, they still belong to $\mathcal{R}_2 \cup \{0\}$, so $h_2 I_{A_1^c}$ is still in $\mathcal{F}_2$ unless it is identically equal to zero.

   Now, if $A_1^c = \emptyset$, then we obtain with the transformation above that $I_{x_2} f = h_1 \in \mathcal{F}_1$. Assume next that $A_1^c \neq \emptyset$ and therefore that $h_2 \neq 0$. Then it follows that for every $x_1 \in A_1^c$ we have that $\underline{P}_2(h_2(x_1, \cdot)) > 0$.

Since $A_1^c \subseteq \mathcal{X}_1$ is finite, we can define $\varepsilon := \min_{x_1 \in A_1^c} \underline{P}_2(h_2(x_1, \cdot))$, so that $\varepsilon > 0$. It follows that

$$h_2 I_{A_1^c} \geq \sum_{x_1 \in A_1^c} I_{x_1}(h_2(x_1, \cdot) - (\underline{P}_2(h_2(x_1, \cdot) - \varepsilon)) = G_2(h_2 I_{A_1^c} | \mathcal{X}_1) + \varepsilon I_{A_1^c}.$$

At this point, there are two possibilities:

(b1) If $\underline{P}(A_1^c) > 0$, then

$$\underline{P}(h_2) = \underline{P}(h_2 I_{A_1^c}) \geq \underline{P}(G_2(h_2 I_{A_1^c} | \mathcal{X}_1)) + \varepsilon \underline{P}(A_1^c) > 0,$$

taking into account that $\underline{P}(G_2(h_2 I_{A_1^c} | \mathcal{X}_1)) \geq 0$ by the coherence of $\underline{P}, \underline{P}_2(\cdot|\mathcal{X}_1)$, and also that $\varepsilon \underline{P}(A_1^c) > 0$. Thus, in that case also $\underline{P}(I_{x_2} f) \geq \underline{P}(h_1) + \underline{P}(h_2) > 0$, whence $I_{x_2} f \in \mathcal{R}$.

(b2) If on the other hand $\underline{P}(A_1^c) = 0$, we can also assume without loss of generality that $h_1 = h_1 I_{x_2}$ and $h_2 = h_2 I_{x_2}$.

- To prove this, assume that there is no $x_2' \neq x_2$ such that $h_1(\cdot, x_2') \neq 0$. Then it would already be the case that $h_1 = h_1 I_{x_2}$ and as a consequence $h_2 = I_{x_2} f - h_1 = I_{x_2}(f - h_1)$, and therefore $h_2 = h_2 I_{x_2}$.
- Assume now that there is $x_2' \neq x_2$ such that $h_1(\cdot, x_2') \neq 0$. Since

$$h_1(x_1, x_2') = (I_{x_2} f - h_2)(x_1, x_2') = 0 \; \forall x_1 \in A_1,$$

we deduce that $h_1(\cdot, x_2') = h_1(\cdot, x_2') I_{A_1^c}$, whence $\underline{P}_1(h_1(\cdot, x_2')) = \underline{P}_1(h_1(\cdot, x_2') I_{A_1^c}) \leq 0$, using that $\underline{P}_1(A_1^c) = \underline{P}(A_1^c) = 0$ together with [19, Lemma 1]. Since $h_1(\cdot, x_2') \in \mathcal{R}_1 \cup \{0\}$ and it is non-zero by assumption, this can only be if $h_1(\cdot, x_2') \geq 0$. Taking into account that $h_1(\cdot, x_2') = (I_{x_2} f - h_2)(\cdot, x_2') = -h_2(\cdot, x_2')$, this means that $h_2(\cdot, x_2') \leq 0$ for every $x_2' \neq x_2$.
Let $A_2 := \{x_2' \in \mathcal{X}_2 : x_2' \neq x_2, h_1(\cdot, x_2') \neq 0\}$, and redefine $h_1 := h_1 - h_1 I_{A_2} \in \mathcal{F}_1 \cup \{0\}, h_2 := h_2 + h_1 I_{A_2} \geq h_2 \in \mathcal{F}_2 \cup \{0\}$. Then $I_{x_2} f = h_1 + h_2$ and $h_1 = h_1 I_{x_2}$ by construction. Reasoning as at the end of previous case we deduce that also $h_2 = h_2 I_{x_2}$.

Now, if $h_2 I_{x_2} \in \mathcal{F}_2 \cup \{0\}$, it means that for every $x_1 \in \mathcal{X}_1$ the gamble $h_2(x_1, \cdot)$ belongs to $\mathcal{R}_2 \cup \{0\}$. But this gamble can only be non-zero on $x_2$, and as a consequence it must be $h_2(x_1, x_2) \geq 0$. From this we deduce that $h_2 \geq 0$, and since we assumed before that $h_2 = h_2 I_{A_1^c}$, we conclude that $h_2 = 0$. Thus, $I_{x_2} f = h_1 \in \mathcal{F}_1 \cup \{0\}$, and since this gamble is different from zero we deduce that $I_{x_2} f \in \mathcal{F}_1$.

We see then that whenever $I_{x_2} f$ is non-zero then it must belong to either $\mathcal{R}$ or $\mathcal{F}_1$. Both these sets are $\mathcal{X}_2$-conglomerable; to see that $\mathcal{R}$ is conglomerable, use that the coherence of $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2)$ implies that they are conglomerably coherent by [19, Theorem 25(i)], and then apply [18, Theorem 9(2)]. As a consequence,

$$f = \sum_{x_2 \in \mathcal{X}_2 : I_{x_2} f \in \mathcal{R}} I_{x_2} f + \sum_{x_2 \in \mathcal{X}_2 : I_{x_2} f \notin \mathcal{R}} I_{x_2} f \in \mathcal{R} \oplus \mathcal{F}_1 \subseteq \mathcal{R} \oplus \mathcal{F}_1 \oplus \mathcal{F}_2.$$

This shows that $\mathcal{R} \oplus \mathcal{F}_1 \oplus \mathcal{F}_2$ is $\mathcal{X}_2$-conglomerable and as a consequence $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are conglomerably coherent. □

It is instead an open problem whether coherence and conglomerable coherence are equivalent under independence in general: we do not know yet if it is possible for two marginals $\underline{P}_1, \underline{P}_2$ to have an independent product $\underline{P}$ while $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are not conglomerably coherent. Nevertheless, we have established a number of sufficient conditions for this equivalence, which show that the two notions are very close also in this case. We must first establish the following lemma:

**Lemma 19.** *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\Omega)$ and consider $A \subseteq \Omega$ such that $\overline{P}(A) = 0$. Given two gambles $f, g$ such that $f I_{A^c} = g I_{A^c}$, it holds that $\underline{P}(f) = \underline{P}(g)$.*

*Proof.* We begin by showing that for any real number $k$ it holds that $\underline{P}(f) = \underline{P}(f + kI_A)$. To see this, note that

$$\underline{P}(kI_A) = \begin{cases} k\underline{P}(A) = 0 \text{ if } k \geq 0 \\ -k\overline{P}(A) = 0 \text{ if } k < 0 \end{cases} \quad \text{and, similarly,} \quad \overline{P}(kI_A) = \begin{cases} k\overline{P}(A) = 0 \text{ if } k \geq 0 \\ -k\underline{P}(A) = 0 \text{ if } k < 0. \end{cases}$$

Since $\underline{P}(f) + \underline{P}(kI_A) \leq \underline{P}(f + kI_A) \leq \underline{P}(f) + \overline{P}(kI_A)$, we deduce that $\underline{P}(f) = \underline{P}(f + kI_A)$. Now, if $f = g$ on $A^c$, by letting $\underline{k} := \min\{\inf f, \inf g\}, \bar{k} := \max\{\sup f, \sup g\}$, we deduce that

$$fI_{A^c} + \underline{k}I_A \leq f, g \leq fI_{A^c} + \bar{k}I_A,$$

and as a consequence $\underline{P}(f) = \underline{P}(g) = \underline{P}(fI_{A^c})$. □

**Theorem 20.** *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$ with marginals $\underline{P}_1, \underline{P}_2$. Define $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ by epistemic irrelevance. Then $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are conglomerably coherent if and only if they are coherent when any of the following conditions holds:*

1. *$\underline{P}_1(x_1) > 0, \underline{P}_2(x_2) > 0$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$.*
2. *$\overline{P}_1(x_1) = 0 = \overline{P}_2(x_2)$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$.*
3. *Either $\underline{P}_1$ or $\underline{P}_2$ is vacuous.*

*Proof.* Since conglomerable coherence is stronger than coherence by [18, Theorem 9], we only need to show that the converse holds under any of the hypotheses of the theorem.

1. If $\underline{P}, \underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are coherent, then they are in particular weakly coherent. Since $\underline{P}(x_1) > 0, \underline{P}(x_2) > 0$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$, we can apply [18, Theorem 13] to deduce that $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$ are also conglomerably coherent.

2. Consider next the case where $\overline{P}_1(x_1) = 0 = \overline{P}_2(x_2)$ for every $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$. Let $\mathcal{F}_1, \mathcal{F}_2$ be the sets of gambles given by Eqs. (B.2) and (B.3). It has been showed in [18, Theorem 10] that $\mathcal{F}_1 \oplus \mathcal{F}_2$ is a coherent set that induces $\underline{P}_1(\cdot|\mathcal{X}_2), \underline{P}_2(\cdot|\mathcal{X}_1)$. Let us prove that this set is $\mathcal{X}_2$-conglomerable (the proof of $\mathcal{X}_1$-conglomerability is analogous).

   Consider a gamble $f \neq 0$ such that $I_{x_2}f \in \mathcal{F}_1 \oplus \mathcal{F}_2 \cup \{0\}$ for every $x_2 \in \mathcal{X}_2$. Take $x_2$ such that $I_{x_2}f \neq 0$. Then there are $h_1 \in \mathcal{F}_1 \cup \{0\}, h_2 \in \mathcal{F}_2 \cup \{0\}$ such that $I_{x_2}f = h_1 + h_2$. Let us prove that $I_{x_2}f$ belongs to $\mathcal{F}_1$.

   There are a number of possibilities:

   - If $h_1 \geq 0$, then $I_{x_2}f \geq h_2$, and as a consequence it belongs to $\mathcal{F}_2$. This means that for every $x_1 \in \mathcal{X}_1$ the gamble $I_{x_2}f(x_1, \cdot)$ belongs to $\mathcal{R}_2 \cup \{0\}$. Since this gamble is only non-zero on $(x_1, x_2)$, we deduce that $f(x_1, x_2) \geq 0$ for every $x_1 \in \mathcal{X}_1$. Thus, $I_{x_2}f \geq 0$ and as a consequence it belongs to $\mathcal{F}_1$.
   - If $h_2 \geq 0$ then $I_{x_2}f \geq h_1$ and as a consequence it belongs to $\mathcal{F}_1$.
   - Consider next the case where $h_1 \not\geq 0$ and $h_2 \not\geq 0$, and define

     $$A_1 := \{x_1 \in \mathcal{X}_1 : h_2(x_1, \cdot) \geq 0\}, A_2 := \{x_2' \in \mathcal{X}_2 : h_1(\cdot, x_2') \geq 0\}.$$

   Let $g_1 := h_1 I_{A_2^c}, g_2 := h_2 I_{A_1^c}$. Since by construction $g_1, g_2 \neq 0$, we deduce that $g_1 \in \mathcal{F}_1, g_2 \in \mathcal{F}_2$. Moreover, $g_1 + g_2 \leq h_1 + h_2 = I_{x_2}f$, and $\underline{P}_2(g_2|x_1) = \underline{P}_2(g_2(x_1, \cdot)) > 0$ for every $x_1 \in S_1(g_2)$.

     – Let us show that we can assume without loss of generality that $S_2(g_2) \cap \{x_2\}^c = S_2(g_1) \cap \{x_2\}^c$:

       ∗ If $x_2' \notin S_2(g_1), x_2' \neq x_2$, then $g_1(\cdot, x_2') = 0$, whence $g_2(\cdot, x_2') \leq I_{x_2}(x_2')f(\cdot, x_2') - g_1(\cdot, x_2') = 0$. If we redefine then $g_2$ by $g_2(x_1, x_2') := 0 \; \forall x_2' \notin S_2(g_1), x_2' \neq x_2$, we still have that $g_2 \in \mathcal{F}_2$ (because the new gamble dominates the one we had before) as well as that $g_1 + g_2 \leq I_{x_2}f$ (because $(g_1 + g_2)(\cdot, x_2') = 0 = I_{x_2}(x_2')f(\cdot, x_2')$) , and also $x_2' \notin S_2(g_2)$.

* On the other hand, given $x_2' \in S_2(g_1) \cap \{x_2\}^c$, then $\underline{P}_1(g_1(\cdot, x_2')) > 0$, and as a consequence, $\overline{P}_1(g_2(\cdot, x_2')) \leq \overline{P}_1(-g_1(\cdot, x_2')) = -\underline{P}_1(g_1(\cdot, x_2')) < 0$, whence it must be $g_2(\cdot, x_2') \neq 0$, i.e., $x_2' \in S_2(g_2)$.

Therefore, for every $x_2' \neq x_2$ such that $x_2' \in S_2(g_2)$, we have that $\overline{P}_1(g_2(\cdot, x_2')) < 0$. If $x_2 \notin S_2(g_2)$, then we would have that $\overline{P}_1(g_2|x_2') < 0$ for every $x_2' \in S_2(g_2)$ and $\underline{P}_1(g_2|x_1) > 0$ for every $x_1 \in S_1(g_2)$. This is a contradiction with point (ii) of the characterisation of coherence established in [24, Theorem 7.3.6]. We conclude from that theorem that

$$x_2 \in S_2(g_2) \text{ and } \underline{P}_1(g_2(\cdot, x_2)) \geq 0. \tag{B.4}$$

– Let us prove now that also $x_2 \in S_2(g_1)$. Assume ex-absurdo that $g_1(\cdot, x_2) = 0$. Let us see that we can modify the gamble $g_2$ in order to contradict the coherence of the set of desirable gambles $\mathcal{F}_1 \oplus \mathcal{F}_2$. Define $g_2' := I_{x_2^c} g_2$ on $\mathcal{X}_1 \times \mathcal{X}_2$. This gamble belongs to $\mathcal{F}_2$:
  * Given $x_1 \in S_1(g_2)$, it holds that $\underline{P}_2(g_2'(x_1, \cdot)) = \underline{P}_2(g_2(x_1, \cdot)) > 0$, using Lemma 19. As a consequence also $x_1 \in S_1(g_2')$ and $g_2'(x_1, \cdot) \in \mathcal{R}_2 \cup \{0\}$.
  * If $x_1 \notin S_1(g_2)$, then $g_2(x_1, \cdot) = 0$, whence also $g_2'(x_1, \cdot) = 0 \in \mathcal{R}_2 \cup \{0\}$.

This means that $g_2' \in \mathcal{F}_2 \cup \{0\}$ and as a consequence $g_1 + g_2' \in \mathcal{F}_1 \oplus \mathcal{F}_2$. But $g_1 + g_2' \leq 0$: on the one hand,

$$x_2 \notin S_2(g_1) \Rightarrow g_1(\cdot, x_2) = 0 \text{ and } g_2'(\cdot, x_2) = g_2(\cdot, x_2) I_{x_2^c}(x_2) = 0, \text{ so } g_1(\cdot, x_2) + g_2'(\cdot, x_2) = 0,$$

and for every $x_2' \neq x_2$, it holds that

$$g_1(\cdot, x_2') + g_2'(\cdot, x_2') = g_1(\cdot, x_2') + g_2(\cdot, x_2') \leq I_{x_2}(x_2') f(\cdot, x_2') = 0.$$

Since we are assuming that $g_1 \neq 0$, we obtain a contradiction with the coherence of $\mathcal{F}_1 \oplus \mathcal{F}_2$. We conclude that

$$x_2 \in S_2(g_1) \text{ and therefore } \underline{P}_1(g_1(\cdot, x_2)) > 0. \tag{B.5}$$

Thus,

$$\underline{P}_1(f(\cdot, x_2)) \geq \underline{P}_1(g_1(\cdot, x_2) + g_2(\cdot, x_2)) \geq \underline{P}_1(g_1(\cdot, x_2)) + \underline{P}_1(g_2(\cdot, x_2)) > 0,$$

where the last inequality is a consequence of Eqs. (B.4), (B.5). From this we conclude that $I_{x_2} f \in \mathcal{F}_1$.

We see then that $I_{x_2} f$ belongs to $\mathcal{F}_1$ for every $x_2 \in \mathcal{X}_2$ with $I_{x_2} f \neq 0$. Since $\mathcal{F}_1$ is $\mathcal{X}_2$-conglomerable, we deduce that $f$ also belongs to $\mathcal{F}_1 \subseteq \mathcal{F}_1 \oplus \mathcal{F}_2$. Thus, $\mathcal{F}_1 \oplus \mathcal{F}_2$ is $\mathcal{X}_2$-conglomerable. The proof of the $\mathcal{X}_1$-conglomerability of $\mathcal{F}_1 \oplus \mathcal{F}_2$ is similar.

3. Finally, if $\underline{P}_1$ is vacuous, then the set $\mathcal{F}_1$ it defines by means of Eq. (B.2) is given by $\mathcal{L}^+$, whence $\mathcal{F}_1 \oplus \mathcal{F}_2 = \mathcal{F}_2$ which is both $\mathcal{X}_1$- and $\mathcal{X}_2$-conglomerable. The proof if $\underline{P}_2$ is vacuous is analogous. □

## References


[1] A. Benavoli, M. Zaffalon, and E. Miranda. Robust filtering through coherent lower previsions. *IEEE Transactions on Automatic Control*, 56(7):1567–1581, 2011.

[2] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5:165–181, 2000.

[3] F. Cozman. Irrelevance and independence relations in Quasi-Bayesian networks. In G.F. Coopen and S. Moral, editors, *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*, pages 89–96, 1998.

[4] F. Cozman. Credal networks. *Artificial Intelligence*, 120:199–233, 2000.

[5] F. Cozman and P. Walley. Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence*, 45(1–2):173–195, 2005.

[6] C. de Campos and F. Cozman. Computing lower and upper expectations under epistemic independence. *International Journal of Approximate Reasoning*, 44(3):244–260, 2007.

[7] L. M. de Campos and S. Moral. Independence concepts for convex sets of probabilities. In *Proceedings of 11th. Conference on Uncertainty in Artificial Intelligence*, pages 108–115, San Mateo (EEUU), 1995.

[8] G. de Cooman, F. Hermans, A. Antonucci, and M. Zaffalon. Epistemic irrelevance in credal nets: the case of imprecise Markov trees. *International Journal of Approximate Reasoning*, 51(9):1029–1052, 2010.

[9] G. de Cooman and E. Miranda. Weak and strong laws of large numbers for coherent lower previsions. *Journal of Statistical Planning and Inference*, 138(8):2409–2432, 2008.

[10] G. de Cooman and E. Miranda. Irrelevant and independent natural extension for sets of desirable gambles. *Journal of Artificial Intelligence Research*, 45:601–640, 2012.

[11] G. de Cooman and E. Miranda. Lower previsions induced by filter maps. *Journal of Mathematical Analysis and Applications*, 410:101–116, 2014.

[12] G. de Cooman, E. Miranda, and M. Zaffalon. Independent natural extension. *Artificial Intelligence*, 175(12–13):1911–1950, 2011.

[13] B. de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.

[14] B. de Finetti. *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester, 1974–1975. English translation of [13], two volumes.

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

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

[17] E. Miranda and M. Zaffalon. Conditional models: coherence and inference through sequences of joint mass functions. *Journal of Statistical Planning and Inference*, 140(7):1805–1833, 2010.

[18] E. Miranda and M. Zaffalon. Conglomerable coherence. *International Journal of Approximate Reasoning*, 54(9):1322–1350, 2013.

[19] E. Miranda, M. Zaffalon, and G. de Cooman. Conglomerable natural extension. *International Journal of Approximate Reasoning*, 53(8):1200–1227, 2012.

[20] S. Moral and A. Cano. Strong conditional independence for credal sets. *Annals of Mathematics and Artificial Intelligence*, 35(1–4):295–321, 2002.

[21] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[22] P. Vicig, M. Zaffalon, and F. Cozman. Notes on 'Notes on conditional previsions'. *International Journal of Approximate Reasoning*, 44(3):358–365, 2007.

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

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

[25] P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Reprinted in [26].

[26] P. M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of [25].

[27] M. Zaffalon and E. Miranda. Probability and time. *Artificial Intelligence*, 198(1):1–51, 2013.