Independent Natural Extension

Gert de Cooman[1], Enrique Miranda[2], and Marco Zaffalon[3]

[1] SYSTeMS, Ghent University, Belgium
gert.decooman@ugent.be
[2] University of Oviedo, Spain
mirandaenrique@uniovi.es
[3] IDSIA, Manno (Lugano), Switzerland
zaffalon@idsia.ch

Abstract We introduce a general definition for the independence of a number of finite-valued variables, based on coherent lower previsions. Our definition has an epistemic flavour: it arises from personal judgements that a number of variables are irrelevant to one another. We show that a number of already existing notions, such as strong independence, satisfy our definition. Moreover, there always is a least-committal independent model, for which we provide an explicit formula: the *independent natural extension*. Our central result is that the independent natural extension satisfies so-called marginalisation, associativity and *strong factorisation* properties. These allow us to relate our research to more traditional ways of defining independence based on factorisation.

Key words: Epistemic irrelevance, epistemic independence, independent natural extension, strong product, factorisation.

1 Motivation

In the literature on probability we can recognise two major approaches to defining independence. In the Kolmogorovian tradition, independence is defined by requiring a probability model to satisfy a factorisation property. We call this the *formalist approach*. It views independence as a mathematical property of the model under consideration. On the other hand, the tradition of subjective probability follows an alternative route by regarding independence as an assessment: it is a subject who for instance regards two events as independent, because he judges that learning about the occurrence of one of them will not affect his beliefs about the other. We call this the *epistemic approach*.

We investigate the relationships between the formalist and epistemic approaches to independence in a generalised setting that allows probabilities to be imprecisely specified. We consider a finite number of logically independent variables X_n assuming values in respective finite sets $\mathscr{X}_n$, $n \in N$. We want to express that these variables are independent, in the sense that learning the values of some of them will not affect the beliefs about the remaining ones. We base our analysis on *coherent lower previsions*, which are lower expectation functionals equivalent to closed convex sets of probability mass functions. In the case of precise probability, we refer to an expectation functional as a *linear prevision*.

After discussing the basic notational set-up in Sec. 2, we introduce the formalist approach in Sec. 3. We define three factorisation properties with increasing strength:

productivity, *factorisation*, and *strong factorisation*. For the product of linear previsions—the classical independence notion—all these properties coincide. For lower previsions, the *strong product* is a straightforward generalisation obtained by taking a lower envelope of products of linear previsions. We show that the strong product is strongly factorising.

In Sec. 4 we move on to the epistemic approach. We introduce two notions: *many-to-many independence*, where a subjects judges that knowing the value of any subset of the variables $\{X_n\colon n \in N\}$ is irrelevant to any other subset; and the weaker notion of *many-to-one independence*, where any subset of the variables of $\{X_n\colon n \in N\}$ is judged to be irrelevant to any other single variable. We show that the strong product is a many-to-many (and hence a many-to-one) independent product of its marginals, and that it is uniquely so in the case of linear previsions.

There is no such uniqueness for lower previsions: the strong product is only one of the generally infinitely many possible independent products. In Sec. 5, we focus on the point-wise smallest ones: the least-committal many-to-many, and the least-committal many-to-one, independent products of given marginals. It is an important result of our analysis that these two independent products turn out to be the same object. We call it the *independent natural extension*. The independent natural extension generalises to any finite number of variables a definition given by Walley for two variables [7, Sec. 9.3]. Observe that in the case of two variables, there is no need to distinguish between many-to-one and many-to-many independence.

The relation with the formalist approach comes to the fore in our next result: the independent natural extension is strongly factorising. We go somewhat further in Sec. 6, where we show that a factorising lower prevision must be a many-to-one independent product. Under some conditions, we also show that a strongly factorising lower prevision must be a many-to-many independent product. And since we already know that the smallest many-to-one independent product is the independent natural extension, we deduce that when looking for least-committal models, it is equivalent whether we focus on factorisation or on being an independent product. This allows us to establish a solid bridge between the formalist and epistemic approaches.

In a number of other results we provide useful properties of assorted independent products. Most notably, we show that each independent product (strong, many-to-many, many-to-one, and the independent natural extension) is in some sense associative, and that the operation of marginalisation preserves the type of independent product. We also give an explicit formula for the independent natural extension, as well as simplified expressions in a number of interesting particular cases.

In order to keep this paper reasonably short, we have to assume that the reader has a good working knowledge of the basics of Walley's theory of coherent lower previsions [7]. For a fairly detailed discussion of the coherence notions and results needed in the context of this paper, we refer to [4,5]. An interesting study of some of the notions considered here was done by Vicig [6] for coherent lower *probabilities*.

2 Set-up and Basic Notation

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

For a coherent lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X})$, there is a corresponding closed convex set of dominating linear previsions, or credal set, $\mathscr{M}(\underline{P})$, and a corresponding set of extreme points $\operatorname{ext}(\mathscr{M}(\underline{P}))$. Then $\underline{P}$ is the lower envelope of both $\mathscr{M}(\underline{P})$ and $\operatorname{ext}(\mathscr{M}(\underline{P}))$:

$$\underline{P}(f) = \min\{P(f)\colon P \in \mathscr{M}(\underline{P})\} = \min\{P(f)\colon P \in \operatorname{ext}(\mathscr{M}(\underline{P}))\} \text{ for all } f \in \mathscr{L}(\mathscr{X}) .$$

For any linear prevision P on $\mathscr{L}(\mathscr{X})$, the corresponding mass function p is defined by $p(x) := P(\mathbb{I}_{\{x\}})$, $x \in \mathscr{X}$, and then of course $P(f) = \sum_{x\in\mathscr{X}} f(x)p(x)$.

If I is any subset of N, then denote by X_I the tuple of variables whose components are the X_i, $i \in I$. We denote by $\mathscr{X}_I := \times_{i\in I}\mathscr{X}_i$ the Cartesian product of the sets $\mathscr{X}_i$, which is the set of all maps x_I from I to $\bigcup_{i\in I}\mathscr{X}_i$ such that $x_I(i) \in \mathscr{X}_i$ for all $i \in I$. The elements of $\mathscr{X}_I$ are generically denoted by x_I or z_I, with corresponding components $x_i := x_I(i)$ or $z_i := z_I(i)$, $i \in I$. We will assume that the variables X_i are logically independent, which means that for each non-empty subset I of N, X_I may assume all values in $\mathscr{X}_I$. We can then consider X_I to be a variable on $\mathscr{X}_I$.

We will frequently use the simplifying device of identifying a gamble on $\mathscr{X}_I$ with a gamble on $\mathscr{X}_N$, namely its cylindrical extension. To give an example, if $\mathscr{K} \subseteq \mathscr{L}(\mathscr{X}_N)$, this trick allows us to consider $\mathscr{K} \cap \mathscr{L}(\mathscr{X}_I)$ as the set of those gambles in $\mathscr{K}$ that depend only on the variable X_I. As another example, this device allows us to identify the indicator gambles $\mathbb{I}_{\{x_R\}}$ and $\mathbb{I}_{\{x_R\}\times\mathscr{X}_{N\setminus R}}$, and therefore also the events $\{x_R\}$ and $\{x_R\}\times\mathscr{X}_{N\setminus R}$. More generally, for any event $A \subseteq \mathscr{X}_R$, we can identify the gambles $\mathbb{I}_A$ and $\mathbb{I}_{A\times\mathscr{X}_{N\setminus R}}$, and therefore also the events A and $A\times\mathscr{X}_{N\setminus R}$. In the same spirit, a lower prevision on all gambles in $\mathscr{L}(\mathscr{X}_I)$ can be identified with a lower prevision defined on the set of corresponding gambles on $\mathscr{X}_N$ (those that depend on X_I only), a subset of $\mathscr{L}(\mathscr{X}_N)$.

If $\underline{P}_N$ is a coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$, then for any non-empty subset I of N we can consider its *$\mathscr{X}_I$-marginal* $\underline{P}_I$ as the coherent lower prevision on $\mathscr{L}(\mathscr{X}_I)$ defined by $\underline{P}_I(f) := \underline{P}_N(f)$ for all gambles f on $\mathscr{X}_I$.

3 The Formal Approach

3.1 Basic Definitions

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

The first is a stronger, symmetrised version of the notion of 'forward factorisation' introduced elsewhere [3].

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

In a paper [3] on laws of large numbers for coherent lower previsions, which generalises and subsumes most known versions in the literature, we prove that the condition of forward factorisation (which is implied by the present productivity condition) is sufficient for a law of large numbers to hold.

Theorem 1 (Weak law of large numbers). *Let the coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ be productive. Let $\varepsilon > 0$ and consider arbitrary gambles h_n on $\mathscr{X}_n$, $n \in N$. Let B be a common bound for the ranges of these gambles and let $\min h_n \leq m_n \leq \underline{P}_N(h_n) \leq \overline{P}_N(h_n) \leq M_n \leq \max h_n$ for all $n \in N$. Then*

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

Next comes a generalisation of the linear independence condition that was inspired by, and found to be quite useful in the context of, our research on credal networks [2].

Definition 2 (Factorisation). *Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. We call this lower prevision* (i) factorising *if for all $o \in N$ and all non-empty $I \subseteq N \setminus \{o\}$, all $g \in \mathscr{L}(\mathscr{X}_o)$ and all non-negative $f_i \in \mathscr{L}(\mathscr{X}_i)$, $i \in I$, $\underline{P}_N(f_I g) = \underline{P}_N(f_I \underline{P}_N(g))$, where $f_I := \prod_{i\in I} f_i$; and* (ii) strongly factorising *if $\underline{P}_N(fg) = \underline{P}_N(f\underline{P}_N(g))$ for all $g \in \mathscr{L}(\mathscr{X}_O)$ and non-negative $f \in \mathscr{L}(\mathscr{X}_I)$, where I and O are any disjoint proper subsets of N.*

Consider a real interval $\underline{\overline{a}} := [\underline{a}, \overline{a}]$ and a real number b, then we define $\underline{\overline{a}} \odot b$ to be equal to $\underline{a}b$ if $b \geq 0$, and equal to $\overline{a}b$ if $b \leq 0$. It then follows from the coherence of $\underline{P}_N$ that we also get $\underline{P}_N(f_I \underline{P}_N(g)) = \underline{\overline{P}}_N(f_I) \odot \underline{P}_N(g)$ and $\underline{P}_N(f\underline{P}_N(g)) = \underline{\overline{P}}_N(f) \odot \underline{P}_N(g)$ in Definition 2.

In general, the following relationships hold between these properties. It can be shown by means of counterexamples that the implications are strict.

Proposition 1. *Consider a coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$. If $\underline{P}_N$ is strongly factorising, then it is factorising, and if $\underline{P}_N$ is factorising, then it is productive.*

We now look at a number of special cases.

3.2 The Product of Linear Previsions

If we have linear previsions P_n on $\mathscr{L}(\mathscr{X}_n)$ with corresponding mass functions p_n, then their *product* $S_N := \times_{n\in N} P_n$ is defined as the linear prevision on $\mathscr{L}(\mathscr{X}_N)$ with mass function p_N defined by $p_N(x_N) := \prod_{n\in N} p_n(x_n)$ for all $x_N \in \mathscr{X}_N$, so

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

One of the very useful properties of the product of linear previsions, is that it is associative in the following sense.

Proposition 2. *Consider arbitrary linear previsions P_n on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$.*

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

Importantly, for linear previsions, all the properties introduced in Sec. 3.1 coincide.

Proposition 3. *Consider any linear prevision P_N on $\mathscr{L}(\mathscr{X}_N)$. Then the following statements are equivalent:* (i) $P_N = \times_{n\in N} P_n$ *is the product of its marginals* P_n, $n \in N$; (ii) $P_N(\prod_{n\in N} f_n) = \prod_{n\in N} P_N(f_n)$ *for all* f_n *in* $\mathscr{L}(\mathscr{X}_n)$, $n \in N$; (iii) P_N *is strongly factorising;* (iv) P_N *is factorising; and* (v) P_N *is productive.*

3.3 The Strong Product of Coherent Lower Previsions

In a similar vein, if we have coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, then [1,7] their *strong product* $\underline{S}_N := \times_{n\in N} \underline{P}_n$ is defined as the coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$ that is the lower envelope of the set of independent products:

$$\{\times_{n\in N} P_n \colon (\forall n \in N) P_n \in \mathrm{ext}(\mathscr{M}(\underline{P}_n))\} \ .$$

So for every $f \in \mathscr{L}(\mathscr{X}_N)$:

$$\underline{S}_N(f) = \inf\{\times_{n\in N} P_n(f) \colon (\forall n \in N) P_n \in \mathrm{ext}(\mathscr{M}(\underline{P}_n))\} \ .$$

The set $\mathrm{ext}(\mathscr{M}(\underline{S}_N))$ has the following nice characterisation, which guarantees that the infimum in the equation above is actually a minimum:

$$\mathrm{ext}(\mathscr{M}(\underline{S}_N)) = \{\times_{n\in N} P_n \colon (\forall n \in N) P_n \in \mathrm{ext}(\mathscr{M}(\underline{P}_n))\} \ .$$

Like the product of linear previsions, the strong product of lower previsions satisfies the following very interesting marginalisation and associativity properties:

Proposition 4. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$.*

(i) *For any non-empty subset R of N, $\underline{S}_R$ is the $\mathscr{X}_R$-marginal of $\underline{S}_N$: $\underline{S}_N(g) = \underline{S}_R(g)$ for all gambles g on $\mathscr{X}_R$;*

(ii) *For any partition N_1 and N_2 of N, $\times_{n\in N_1\cup N_2} \underline{P}_n = (\times_{n\in N_1} \underline{P}_n) \times (\times_{n\in N_2} \underline{P}_n)$, or in other words, $\underline{S}_N = \underline{S}_{N_1} \times \underline{S}_{N_2}$.*

This readily leads to the conclusion that the strong product of lower previsions shares many of the interesting properties of the product of linear previsions:

Proposition 5. *The strong product $\underline{S}_N$ is strongly factorising, and therefore factorising and productive. As a consequence, it satisfies the weak law of large numbers of Thm. 1.*

This ends our discussion of the formalist approach to independence for coherent lower previsions. We next turn to the treatment of independence following an epistemic approach, where independence is considered to be an assessment a subject makes.

4 Epistemic Irrelevance and Independence

Consider two disjoint proper subsets I and O of N. We say that a subject judges that X_I *is epistemically irrelevant to* X_O when he assumes that learning which value X_I assumes in $\mathscr{X}_I$ will not affect his beliefs about X_O.

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

$$\underline{P}_{O\cup I}(h|x_I) := \underline{P}_N(h(\cdot, x_I)) \text{ for all gambles } h \text{ on } \mathscr{X}_{O\cup I} \ . \tag{1}$$

4.1 Epistemic Many-to-Many Independence

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

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

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

where $\underline{P}_{O\cup I}(\cdot|X_I)$ is the conditional lower prevision on $\mathscr{L}(\mathscr{X}_{O\cup I})$ given by Eq. (1).

Definition 3. *A coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ is called* many-to-many independent *if it is coherent with the family $\mathscr{I}(\underline{P}_N)$. For a collection of coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, any many-to-many independent coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ that coincides with the $\underline{P}_n$ on their domains $\mathscr{L}(\mathscr{X}_n), n \in N$ is called a* many-to-many independent product *of these marginals.*

4.2 Epistemic Many-to-One Independence

There is weaker notion of independence that we will consider here. We say that a subject judges the variables X_n, $n \in N$ to be *epistemically many-to-one independent* when he assumes that learning the value of any number of these variables will not affect his beliefs about any *single* other. In other words, if he judges for any $o \in N$ and any non-empty subset I of $N \setminus \{o\}$ that X_I is epistemically irrelevant to X_o.

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

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

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

$$\underline{P}_{\{o\}\cup I}(h|x_I) := \underline{P}_N(h(\cdot,x_I)) = \underline{P}_o(h(\cdot,x_I)) \text{ for all gambles } h \text{ on } \mathscr{X}_{\{o\}\cup I} \ ,$$

where of course $\underline{P}_o$ is the $\mathscr{X}_o$-marginal lower prevision of $\underline{P}_N$. So we see that the family of conditional lower previsions $\mathscr{N}(\underline{P}_N)$ only depends on the joint $\underline{P}_N$ through its $\mathscr{X}_n$-marginals $\underline{P}_n$, $n \in N$. This, of course, explains our notation for it.

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

Obviously, if a joint lower prevision $\underline{P}_N$ is many-to-many independent, then it is also many-to-one independent. Another immediate property is that any independent product of a number of lower previsions must have these lower previsions as its marginals:

Proposition 6. *If the coherent lower prevision $\underline{P}_N$ on $\mathscr{L}(\mathscr{X}_N)$ is a many-to-one independent product of coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, then $\underline{P}_N(g) = \underline{P}_n(g)$ for all $g \in \mathscr{L}(\mathscr{X}_n)$ and for all $n \in N$.*

Moreover, independent products satisfy a number of basic marginalisation and associativity properties.

Proposition 7. *Consider arbitrary coherent lower previsions $\underline{P}_n$, $n \in N$. Let $\underline{P}_N$ be any many-to-one independent product and $\underline{Q}_N$ any many-to-many independent product of the marginals $\underline{P}_n$, $n \in N$. Let R and S be any proper subsets of N.*

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

A basic coherence result [7, Thm. 7.1.6] states that taking lower envelopes of a family of coherent conditional lower previsions again produces coherent conditional lower previsions. This implies that many-to-many independence and many-to-one independence are preserved by taking lower envelopes. There is also another interesting way of concluding that a given coherent lower prevision is a many-to-one independent product.

Proposition 8. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$, and let $\underline{Q}_1$ and $\underline{Q}_2$ be coherent lower previsions on $\mathscr{L}(\mathscr{X}_N)$ with these marginals $\underline{P}_n$. Let $\underline{Q}_3$ be any coherent lower prevision on $\mathscr{L}(\mathscr{X}_N)$ such that $\underline{Q}_1 \leq \underline{Q}_3 \leq \underline{Q}_2$. Then* (i) *if $\underline{Q}_1$ and $\underline{Q}_2$ are many-to-one independent products, then so is $\underline{Q}_3$; and* (ii) *if $\underline{Q}_1$ and $\underline{Q}_2$ are factorising, then so is $\underline{Q}_3$.*

We deduce that a convex combination of many-to-one independent products of the same given marginals is again a many-to-one independent product of these marginals.

5 Independent Natural Extension

The (strong) product turns out to be the central notion when we want to take independent products of linear previsions, as the following proposition makes clear.

Proposition 9. *Any collection of linear previsions P_n, $n \in N$ has a unique many-to-many independent product and a unique many-to-one independent product, and both of these are equal to their strong product $S_N = \times_{n \in N} P_n$.*

However, when the marginals we want to combine are lower rather than linear previsions, the situation is decidedly more complex, as we intend to show in the rest of this section. We begin by showing that there always is at least one many-to-many (and therefore also many-to-one) independent product:

Proposition 10. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Then their strong product $\times_{n\in N}\underline{P}_n$ is a many-to-many and many-to-one independent product of its marginals $\underline{P}_n$. As a consequence, the collection $\mathscr{N}(\underline{P}_N)$ of conditional lower previsions $\underline{P}_{\{o\}\cup I}(\cdot|X_I)$ is coherent.*

Because all the sets $\mathscr{X}_n$ are finite, we can invoke Walley's Finite Extension Theorem [7, Thm. 8.1.9] to conclude that there always is a *point-wise smallest* joint lower prevision $\underline{E}_N$ that is coherent with the coherent family $\mathscr{N}(\underline{P}_N)$. So there always is a smallest many-to-one independent product. Interestingly, this coherent lower prevision $\underline{E}_N$ can be proved to be also a many-to-many independent product. Summarising:

Theorem 2 (Independent natural extension). *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. They always have a point-wise smallest many-to-one independent product, and a point-wise smallest many-to-many independent product, and these products coincide. We call this smallest independent product the* independent natural extension *of the marginals $\underline{P}_n$, and denote it by $\underline{E}_N := \otimes_{n\in N}\underline{P}_n$.*

Since the strong product $\times_{n\in N}\underline{P}_n$ is a many-to-one independent product of the marginals $\underline{P}_n$, $n \in N$ by Prop. 10, it has to dominate the independent natural extension $\otimes_{n\in N}\underline{P}_n$: i.e., we have $\times_{n\in N}\underline{P}_n \geq \otimes_{n\in N}\underline{P}_n$. But these products do not coincide in general: Walley [7, Sect. 9.3.4] discusses an example where the many-to-one independent natural extension is not a lower envelope of independent linear products, and as a consequence does not coincide with the strong product.

The independent natural extension can be derived from the marginals constructively. The following theorem establishes a workable expression for it.

Theorem 3. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. Then for all gambles f on $\mathscr{X}_N$:*

$$\underline{E}_N(f) = \sup_{\substack{h_n \in \mathscr{L}(\mathscr{X}_N)\\ n\in N}} \inf_{z_N \in \mathscr{X}_N} \left[f(z_N) - \sum_{n\in N} [h_n(z_N) - \underline{P}_n(h_n(\cdot, z_{N\setminus\{n\}}))] \right]. \tag{2}$$

The special independent products introduced so far satisfy a monotonicity property:

Proposition 11. *Let $\underline{P}_n$ and $\underline{Q}_n$ be coherent lower previsions on $\mathscr{L}(\mathscr{X}_n)$ such that $\underline{P}_n \leq \underline{Q}_n$, $n \in N$. Then $\otimes_{n\in N}\underline{P}_n \leq \otimes_{n\in N}\underline{Q}_n$ and $\times_{n\in N}\underline{P}_n \leq \times_{n\in N}\underline{Q}_n$.*

Like the strong product, the independent natural extension satisfies very useful marginalisation and associativity properties.

Theorem 4. *Consider arbitrary coherent lower previsions $\underline{P}_n$ on $\mathscr{L}(\mathscr{X}_n)$, $n \in N$.*

(i) *For any non-empty subset R of N, $\underline{E}_R$ is the $\mathscr{X}_R$-marginal of $\underline{E}_N$: $\underline{E}_N(f) = \underline{E}_R(f)$ for all gambles f on $\mathscr{X}_R$;*

(ii) *For any partition* N_1 *and* N_2 *of* N, $\otimes_{n\in N_1\cup N_2}\underline{P}_n = (\otimes_{n\in N_1}\underline{P}_n)\otimes(\otimes_{n\in N_2}\underline{P}_n)$.

Using the associativity of the independent natural extension, and the factorising character of the strong product, we are led to the following practically important conclusion:

Theorem 5. *Consider arbitrary coherent marginal lower previsions* $\underline{P}_n$ *on* $\mathscr{L}(\mathscr{X}_n)$, $n \in N$. *Then their independent natural extension* $\otimes_{n\in N}\underline{P}_n$ *is strongly factorising, and therefore factorising and productive.*

Amongst other things, this implies that, like the strong product, the independent natural extension satisfies the weak law of large numbers of Thm. 1, and as a consequence also any many-to-one independent product.

When some of the marginals are linear or vacuous previsions, the expression for the independent natural extension in Eq. (2) simplifies. Because of the associativity result in Thm. 4, it suffices to consider the case of two variables X_1 and X_2, so $N = \{1,2\}$.

Proposition 12. *Let* P_1 *be any linear prevision on* $\mathscr{L}(\mathscr{X}_1)$, *and let* $\underline{P}_2$ *be any coherent lower prevision on* $\mathscr{L}(\mathscr{X}_2)$. *Let* $\underline{P}_{\{1,2\}}$ *be any (many-to-many) independent product of* P_1 *and* $\underline{P}_2$. *Then for all gambles* f *on* $\mathscr{X}_1 \times \mathscr{X}_2$*:*

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

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

Proposition 13. *Let* $\underline{P}_1^{A_1}$ *be the vacuous lower prevision on* $\mathscr{L}(\mathscr{X}_1)$ *relative to the non-empty set* $A_1 \subseteq \mathscr{X}_1$, *and let* $\underline{P}_2$ *be any coherent lower prevision on* $\mathscr{L}(\mathscr{X}_2)$. *Then for all gambles* f *on* $\mathscr{X}_1 \times \mathscr{X}_2$*:*

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

6 Factorisation and Independence

Since we know from Prop. 5 and Thm. 5 that both the strong product and the independent natural extension are strongly factorising, we wonder if we can use factorising lower previsions as many-to-one or many-to many independent products.

Theorem 6. *Consider an arbitrary coherent lower prevision* $\underline{P}_N$ *on* $\mathscr{L}(\mathscr{X}_N)$. *If it is factorising, then it is a many-to-one independent product.*

In the same vein, one could expect strong factorisation to be sufficient for being a many-to-many independent product. So far, we are only able to prove this for coherent lower previsions that satisfy an extra positivity property. It is still unclear whether we can extend the theorem below to general strongly factorising joints.

Theorem 7. *Let* $\underline{P}_N$ *be a strongly factorising coherent lower prevision on* $\mathscr{L}(\mathscr{X}_N)$. *If* $\overline{P}_N(\{x_n\}) > 0$ *for all* $x_n \in \mathscr{X}_n$ *and* $n \in N$, *then* $\underline{P}_N$ *is many-to-many independent.*

7 Conclusions

This paper is an attempt to lay down the foundations for a general notion of independence of finite-valued variables for imprecise probability models. We have taken the epistemic stance in that we regard a number of variables as independent when they are irrelevant to one another; and we have distinguished many-to-many from many-to-one independence based on whether the independence judgements affect sets of variables or single variables, respectively. This distinction turns out to vanish when we focus on least-committal models, as there is a unique smallest model for any (and therefore both) type of requirement: we have called it the independent natural extension. This generalises a proposal by Walley originally made for the case of two variables [7, Sec. 9.3]. Moreover, the independent natural extension satisfies a practically important strong factorisation property. This brings the independent natural extension closer to more traditional definitions of independence that are indeed based on factorisation, and it should make it easier to work with epistemic independence, since we can impose it through factorisation. Finally, other interesting contributions in this paper concern the relationship between epistemic independence and strong independence, and the proof of a number of their basic properties.

Acknowledgements. This work is supported by Swiss NSF grants n. 200020-116674/1 and 200020-121785/1, projects TIN2008-06796-C04-01 and MTM2007-61193, and SBO project 060043 of the IWT-Vlaanderen.

References

1. Fabio G. Cozman. Constructing sets of probability measures through Kuznetsov's independence condition. In G. de Cooman, T. L. Fine, and T. Seidenfeld, editors, *ISIPTA '01 – Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, pages 104–111. Shaker Publishing, Maastricht, 2000.
2. Gert de Cooman, Filip Hermans, Alessandro Antonucci, and Marco Zaffalon. Epistemic irrelevance in credal nets: the case of imprecise Markov trees. *International Journal of Approximate Reasoning*, 2009. Accepted for publication.
3. Gert de Cooman and Enrique Miranda. Weak and strong laws of large numbers for coherent lower previsions. *Journal of Statistical Planning and Inference*, 138(8):2409–2432, 2008.
4. Enrique Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.
5. Enrique Miranda. Updating coherent lower previsions on finite spaces. *Fuzzy Sets and Systems*, 160(9):1286–1307, 2009.
6. Paolo Vicig. Epistemic independence for imprecise probabilities. *International Journal of Approximate Reasoning*, 24(3):235–250, 2000.
7. Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.